# Mass Spectrometry Method Development and Application to Investigate Plant-Microbe and Microbiome-Host Symbiosis

By

Caitlin Keller

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2019

Date of final oral examination:

The dissertation is approved by the following members of the Final Oral Committee
Lingjun Li, Professor, Chemistry and Pharmacy
Joshua Coon, Professor, Chemistry and Biomolecular Chemistry
Michal Sussman, Professor, Biochemistry
Cameron Currie, Professor, Bacteriology

**Acknowledgements**

First, I would like to thank my graduate school advisor Dr. Lingjun Li. Her support has been critical to my success in graduate school and this research would not have been possible without her. She has let me pursue a variety of research topics during by graduate school career based upon my interests and provided me with the necessary guidance to become an independent scientist. I appreciate all her time and effort to support my graduate school career and develop my professional interests. I truly appreciate the opportunity to have her as my advisor as she genuinely cares about every student. I will always be thankful for this opportunity.

I would also like to thank my committee members: Professor Joshua Coon, Professor Michael R. Sussman, and Professor Cameron Currie. In addition to being my committee members for my thesis background oral exam, original research preproposal, and thesis defense, they have also offered insightful advice on my research and been a part of my professional development. Professor Cameron Currie is also a collaborator for my microbiome project, along with the Frederico Rey lab. For the *Medicago truncatula* studies, I collaborate with Profoessor Jean-Michel Ané. I would like to thank all of my collaborators in the various labs that I have worked with for their willingness to work with me and all of their help over the years. All the collaborative work has increased by passion for bioanalytical research.

I would also like to thank the Morgridge Wisconsin Distinguished Graduate Fellowship in Biotechnology and NSF funding for the Medicago studies that have supported my graduate school career. Additionally, I have twice received the conference presentation funds award through the graduate school to support my travel to the annual American Society for Mass Spectrometry conference. I have also received the Chemistry Department Travel Award, which supported my travels to the Chicago Mass Spec day conference and the Travel Stipend Award

for the 33rd Asilomar Conference on Mass Spectrometry. These conferences helped me to grow as a scientist and researcher and helped me to keep up with new technologies in the mass spectrometry field. I would also like to thank funding from the University of Wisconsin Graduate School.

I would also like to thank all the opportunities that I had to perform research during my undergraduate career. Those experiences sparked my interest in performing research, especially analytical chemistry research, and influenced my decision to go to graduate school. Initially, I had the opportunity to explore analytical chemistry research with Dr. William Steinecker at the Institute for the Development of Commercialized Advanced Sensor Technology at the University of Dayton Research Institute. Here, I was introduced to analytical chemistry research. I also had two summer research experience through the Research Experience for Undergrads (REU) program. The first REU with Professor Eric Ross at Colorado State University gave me the opportunity to learn about research in biology. The second REU with Professor Gary Lorigan at Miami University gave me experience in biophysical research. I would like to thank all my undergraduate mentors for developing my research skills and interest in bioanalytical chemistry.

I would also like to acknowledge the Li lab members that I have meet and worked with during my time in graduate school. I especially would like to thank Dr. Erin Gemperline, who was my graduate student mentor when I joined the Li lab. She introduced me the mass spectrometry techniques I have used in my research and initiated the plant and microbiome projects. I would also like to thank Li lab members that have been very helpful in collaborative projects. Pingli Wei has worked with me on the microbiome projects. Additionally, Nhu Vu and Jericha Mill have worked with me on the plant and microbiome projects respectively. I would also like to thank Dr.

Amanda Buchberger, Jill Johnson, and Kellen Delany, co-workers in the lab with me, who have offered thoughtful advice for my research.

The School of Pharmacy's Analytical Instrumentation Center (AIC) has also provided insightful guidance and suggestions on my research. I would like to thank Molly and Cameron in the AIC for their help and for maintaining and troubleshooting the instruments in the AIC. They have been very supportive whenever I had questions about instrumentation maintenance and troubleshooting or about my research.

Finally, I would like to thank my family for their support during my graduate school career. Graduate school was not always easy, and my family's support helped to overcome the challenges. My parents have always been supportive of my goals and their support has been critical to me achieving my goals. Whether it was a midnight call asking how to get a bat out of my apartment or a midday call about my most recent research failure, my parents were always there to calm me down and tell me it will be alright.

## Table of Contents

# Mass Spectrometry Method Development and Application to Investigate Plant-Microbe and Microbiome-Host Symbiosis

Caitlin Keller
Under the supervision of Professor Lingjun Li
University of Wisconsin-Madison

## Abstract

Mass Spectrometry is a powerful tool to analyze different biomolecules in a tissue section. Both liquid chromatography-mass spectrometry (LC-MS) and mass spectrometry imaging (MSI) techniques are useful for large-scale detection and identification of small molecules, peptides, and proteins in samples, with MSI additionally being able to determine the spatial distribution of molecules within a tissue section. In this thesis, LC-MS and MSI approaches were applied to study symbiotic relationships. In the plant *Medicago truncatula*, a symbiotic relationship with rhizobia bacteria in specialized organs, root nodules, of the plant results in the reduction of atmospheric nitrogen into ammonia (termed biological nitrogen fixation). MSI methods were developed to investigate the small molecule and peptide content in the root nodules. Furthermore, a developed MSI method was applied to study salt stress, which negatively affects development of the symbiotic relationship in the root nodules. The other relationship explored was the complex interactions between the gut microbiome and its host. To investigate this complicated relationship, multiomics studies, which combine two or more single omics studies, such as metabolomics and proteomics, were utilized to achieve a more comprehensive analysis of the gut microbiome. Developed multiomic approaches were applied to study the response of the gut microbiome to pathogenic infection. Overall, this work develops a multifaceted approach to investigate symbiotic relationships, resulting in new method development, improved MSI sample preparation protocols, and increased knowledge about molecular players involved in the symbiosis.

# Chapter 1

# Introduction and Research Summary

**Introduction and Research Summary**

Mass spectrometry (MS) is a robust and sensitive technique for qualitative and quantitative measurements of the various types of biomolecules present in a tissue. Liquid chromatography (LC)- MS can detect and, with tandem MS (MS/MS) capabilities, identify hundreds to thousands of molecules in a single run, which is advantageous for large-scale profiling and relative quantitation of molecules. However, the spatial resolution of the molecules in the tissue is lost during the extraction process. Mass spectrometry imaging (MSI) profiles molecules directly in a tissue section, allowing for determination of the spatial localization of molecules in a tissue. Although there are many ionization methods for MSI, matrix-assisted laser desorption/ionization (MALDI) is very commonly used for MSI. By combining LC-MS/MS and MALDI-MSI, large-scale detection, quantitation, and localization of biomolecules are possible. This work studies plant-bacteria symbiosis and the microbiome-host relationship via LC-MS/MS and MALDI-MSI to investigate the various biomolecules involved in these important biological systems. In Chapter 1, a general overview of the work described in this thesis is provided. Chapter 2 describes the use of MS and MSI techniques to study metabolites, peptides, and proteins in plant science.[1] A sample preparation procedure for MALDI-MSI of peptides in the root nodules of the plant *Medicago truncatula* can be found in Appendix I, and a discussion of the challenges in metabolite identification and the various tools available to aid in metabolite identification can be found in Appendix II.

MALDI-MSI has exhibited significant utility in the localization of metabolites in plants, including *Medicago truncatula*.[2, 3] A study of the metabolites involved in biological nitrogen fixation in *Medicago truncatula* root nodules through MALDI-MSI and LC-MS/MS can be found in Appendix III. MALDI-MSI can also be used to study the molecular changes due to salt stress,

which negatively affects root nodule formation. As root nodule formation is critical for the symbiosis with rhizobia bacteria, a better understanding of the metabolites involved in salt stress would be beneficial. Metabolites were imaged in root nodules grown in low and high salt concentrations to determine the distribution changes in the root nodules in response to salt stress, the details of which are in Appendix IV. The study of salt stress in the root nodule was continued with a novel MALDI source capable of higher spatial resolution. Efforts in improving the spatial resolution in MALDI-MSI look to decrease pixel size in order to resolve smaller molecular features. In Chapter 3, an atmospheric pressure (AP)-MALDI source was investigated as it allowed for higher spatial resolution imaging. Initially AP-MALDI imaging of metabolites in root nodules of *Medicago truncatula* was compared to vacuum MALDI-MSI. Then, AP-MALDI-MSI was applied to study salt stress in *Medicago truncatula* root nodules.[4]

Although MALDI-MSI is commonly applied to study metabolites, imaging of endogenous peptides with MALDI-MSI is less commonly employed.[5] As endogenous peptides have critical roles in the plant, including root nodule development,[6] a MALDI-MSI method to image peptides in plants was developed. Chapter 4 describes the initial study developing a MALDI-MSI method to study peptides in *Medicago truncatula* root nodules.[7] Different mutant lines with altered peptide content were compared to wild-type plants, and peptides were identified with LC-MS/MS. While numerous peptides were imaged in this study, nodule-specific cysteine-rich (NCR) peptides, which are critical for proper formation of the symbiotic relationship in the root nodules, were not identified. As a result, Chapter 5 describes the optimization of MALDI-MSI sample preparation through inclusion of sample washing procedures to image NCR peptides. Multiple NCR peptides were identified with the optimized strategy and LC-MS/MS was used to confirm identifications. Appendix V describes the optimization of LC-MS/MS sample preparation procedures on

*Medicago truncatula* seedling plants to identify endogenous peptides from plant extractions. Multiple extractions were attempted with various sample clean-up preparation steps.

The other biological system studied in this work is the gut microbiome, which has critical roles in human health. Due to the complexity of the microbiome, multiomics studies are beneficial as they allow for characterization of different molecule types in the microbiome. For example, metagenomic analysis of changes in microbiome composition can be combined with metabolomics or metaproteomics studies.[8, 9] Chapter 6 describes the cecum extraction optimization for combining metabolomics, peptidomics, and proteomics studies into a single multiomics analysis. Three different extraction methods commonly employed for metabolomics, peptidomics, and/or proteomics studies were compared. Peptidomics is less commonly combined with other omics to study the microbiome. By including peptidomics with metabolomics and proteomcis, more detailed knowledge of molecular processes occurring in the gut can be obtained, including information about communication in the gut via signaling peptides.

The optimal extraction found in Chapter 6 was then applied to two multiomcs studies investigating the microbiome's response to infection by either *Salmonella enterica* Typhimurium or *Candida albicans*. Chapter 7 describes the metagenomics and metabolomics study utilizing LC-MS to look at compounds produced by the microbiome after pathogenic infection.[10] A model of the human gut microbiome was introduced to germ-free mice as a model system. Humanized, infected mice were then compared to humanized mice with no infection and mice monocolonized with the pathogen. In Chapter 8, the study was further expanded to investigate the metabolomic, peptidomic, and proteomic changes in the microbiome due to infection. Both *Salmonella enterica* Typhimurium and *Candida albicans* infections were again utilized with the humanized mice model

system. The inclusion of peptidomics to the workflow allowed profiling of neuropeptides involved the response to infection in the gut.

The multiomics studies on the microbiome focused on identifying molecular compounds in the cecum. In order to localize target compounds to the mucosal layer, where members of the microbiome reside, MALDI-MSI of mouse intestine samples was explored in Appendix VI. Different instrumentation platforms, including time of flight instruments capable of high spatial resolution, as well as different matrices were compared. Additionally, a PVDF membrane set-up was explored to investigate bacteria interactions via MALDI-MSI. Bacteria are grown in wells on top of the PVDF membrane, and molecular communication occurs within the membrane. The ability of MALDI-MSI to detect signaling molecules produced by the bacteria in the PVDF membrane is described in Appendix VII.

Although the majority of the work presented in this thesis focuses on microbial relationships with their host, other projects utilizing MS techniques to study various biological systems were performed. Appendix VIII describes the optimization of MALDI-MSI workflow to detect endogenous peptides in the pancreas islets of Langerhans. These islets produce many endogenous peptides, including insulin, and therefore pancreas tissue is a good target tissue for MALDI-MSI.[11, 12] Appendix IX describes the use of MALDI-MSI to image a Gd chelate alkylphosphocholine analog molecule in tumor, liver, and kidneys. The Gd chelate molecule is based upon previous alkylphosphocholine analog technology that allows for tumor specific uptake of the molecule for magnetic resonance imaging.[13] MALDI-MSI was used to visualize the drug distribution in various tissues, and to investigate other molecular changes occurring due addition of the molecule. Finally, Appendix X describes the use of MALDI-MSI to investigate the lipid molecules providing insulation in the electric eel.

Chapter 9 provides overall conclusions derived from the studies presented in this thesis and describes potential future directions. Chapter 10 describes the MALDI-MSI studies on *Medicago truncatula* root nodules, including both the small molecule changes due to salt stress and the optimization of MALDI-MSI to study NCR peptides, for a broader audience as part of the Wisconsin Initiative for Scientific Literacy project.

**References**

1. Gemperline, E.; Keller, C.; Li, L. J., Mass Spectrometry in Plant-omics. *Analytical Chemistry* **2016,** *88* (7), 3422-3434.

2. Lee, Y. J.; Perdian, D. C.; Song, Z. H.; Yeung, E. S.; Nikolau, B. J., Use of mass spectrometry for imaging metabolites in plants. *Plant Journal* **2012,** *70* (1), 81–95.

3. Gemperline, E.; Jayaraman, D.; Maeda, J.; Ane, J. M.; Li, L., Multifaceted investigation of metabolites during nitrogen fixation in Medicago via high resolution MALDI-MS imaging and ESI-MS. *J Am Soc Mass Spectrom* **2015,** *26* (1), 149-58.

4. Keller, C.; Maeda, J.; Jayaraman, D.; Chakraborty, S.; Sussman, M. R.; Harris, J. M.; Ané, J.-M.; Li, L., Comparison of Vacuum MALDI and AP-MALDI Platforms for the Mass Spectrometry Imaging of Metabolites Involved in Salt Stress in Medicago truncatula. *Frontiers in Plant Science* **2018,** *9*, 1238.

5. Kaspar, S.; Peukert, M.; Svatos, A.; Matros, A.; Mock, H. P., MALDI-imaging mass spectrometry - An emerging technique in plant biology. *Proteomics* **2011,** *11* (9), 1840-50.

6. Tavormina, P.; De Coninck, B.; Nikonorova, N.; De Smet, I.; Cammue, B. P., The Plant Peptidome: An Expanding Repertoire of Structural Features and Biological Functions. *Plant Cell* **2015,** *27* (8), 2095-118.

7. Gemperline, E.; Keller, C.; Jayaraman, D.; Maeda, J.; Sussman, M. R.; Ane, J. M.; Li, L., Examination of Endogenous Peptides in Medicago truncatula Using Mass Spectrometry Imaging. *J Proteome Res* **2016,** *15* (12), 4403-4411.

8. Verberkmoes, N. C.; Russell, A. L.; Shah, M.; Godzik, A.; Rosenquist, M.; Halfvarson, J.; Lefsrud, M. G.; Apajalahti, J.; Tysk, C.; Hettich, R. L.; Jansson, J. K., Shotgun metaproteomics of the human distal gut microbiota. *Isme Journal* **2009,** *3* (2), 179-189.

9. Erickson, A. R.; Cantarel, B. L.; Lamendella, R.; Darzi, Y.; Mongodin, E. F.; Pan, C.; Shah, M.; Halfvarson, J.; Tysk, C.; Henrissat, B.; Raes, J.; Verberkmoes, N. C.; Fraser, C. M.;

Hettich, R. L.; Jansson, J. K., Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. *PLoS One* **2012,** *7* (11), e49138.

10. Bratburd, J. R.; Keller, C.; Vivas, E.; Gemperline, E.; Li, L.; Rey, F. E.; Currie, C. R., Gut Microbial and Metabolic Responses to Salmonella enterica Serovar Typhimurium and Candida albicans. *MBio* **2018,** *9* (6).

11. Stewart, K. W.; Phillips, A. R.; Whiting, L.; Jullig, M.; Middleditch, M. J.; Cooper, G. J., A simple and rapid method for identifying and semi-quantifying peptide hormones in isolated pancreatic islets by direct-tissue matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* **2011,** *25* (22), 3387-95.

12. Minerva, L.; Clerens, S.; Baggerman, G.; Arckens, L., Direct profiling and identification of peptide expression differences in the pancreas of control and ob/ob mice by imaging mass spectrometry. *Proteomics* **2008,** *8* (18), 3763-3774.

13. Weichert, J. P.; Clark, P. A.; Kandela, I. K.; Vaccaro, A. M.; Clarke, W.; Longino, M. A.; Pinchuk, A. N.; Farhoud, M.; Swanson, K. I.; Floberg, J. M.; Grudzinski, J.; Titz, B.; Traynor, A. M.; Chen, H.-E.; Hall, L. T.; Pazoles, C. J.; Pickhardt, P. J.; Kuo, J. S., Alkylphosphocholine Analogs for Broad-Spectrum Cancer Imaging and Therapy. *Science Translational Medicine* **2014,** *6* (240), 240ra75.

# Chapter 2

# Mass Spectrometry in Plant-omics

Adapted from: Gemperline E, Keller C, Li L. Mass Spectrometry in Plant-omics. *Analytical Chemistry* **2016,** *88* (7), 3422-3434.

Keywords: Mass Spectrometry; Mass spectrometry Imaging; LC-MS; metabolomics; proteomics; peptidomics

**Abstract**

Plant-omics is rapidly becoming an important field of study in the scientific community due to the urgent need to address many of the most important questions facing humanity today with regard to agriculture, medicine, biofuels, environmental decontamination, ecological sustainability, etc. High performance mass spectrometry is a dominant tool for interrogating the metabolomes, peptidomes, and proteomes of a diversity of plant species under various conditions, revealing key insights into the functions and mechanisms of plant biochemistry.

**Plant-omics**

Plant science is one of the oldest[1-3] areas of scientific research but is still an extremely important and rapidly developing area of study. Plant life plays an essential role in these challenging issues. For example, growing robust crops with increased yield not only affects the agriculture industry, but also plays a role in environmental sustainability and contamination. Non-food crop plants are being studied and developed as biofuels for sustainable energy.[4-7] Additionally, the ability of plants to withstand attacks from pests and pathogens in their environment is being studied in the medical community in light of the rise of drug-resistant strains of infections.[8-11] Plant-derived natural products provide a potentially rich source of drug candidates and target compounds for drug discovery and development.[12-17] Research on plants is essential to finding solutions to these major issues, and the key to finding viable solutions lies in integrated, holistic studies and collaboration between scientists.

In molecular biology, the term "–ome" refers to the study of the global or total changes in an organism's DNA or genes (genome), proteins (proteome), metabolites (metabolome), etc. -Omics (genomics, proteomics, metabolomics, etc.) has come to refer generally to the study of large,

comprehensive biological datasets focused on investigating changes in these -omes. Therefore, plant-omics can be described as a holistic study of any or all –omes, specifically in plants. **Figure 1** depicts the different classes of -omics and the interactions between them. Multi-omics (*i.e.* a combination of genomics, transcriptomics, proteomics, peptidomics, and/or metabolomics) strategies could provide large-scale insights into complex plant systems that could otherwise be misinterpreted if only one –ome was interrogated.

Recent years have seen an explosion of new -omics technologies that have the potential to enable ground-breaking discoveries in plant sciences.[18] In the field of genomics, next generation sequencing and advances in genome analysis have helped to lay a foundation for other, downstream -omics studies. Transcriptomics, proteomics, peptidomics, and metabolomics can provide further insights into the inner workings of plant cells, cell-cell communication, and even plant-environment interactions.

Traditional biological tools for studying plant-omics include SNP genotyping for genomics, RNA-Seq and gene expression microarrays/ gene chips for transcriptomics, gel electrophoresis, ELISAs (enzyme-linked immunosorbent assay), protein microarrays and chromatography for proteomics, and chromatography and NMR (nuclear magnetic resonance) for metabolomics. These assays are still widely used in plant science research, often in combination or as a supplement to mass spectrometry analysis, especially when it comes to proteomics and metabolomics.

This Feature will focus on three –omic classes, proteomics, peptidomics, and metabolomics, which have made great advances in the plant-omics field due to the innovation of mass spectrometry. It will summarize the current state of mass spectrometry in plant-omic research, highlighting the different mass spectrometry techniques, especially mass spectrometry

imaging, discussing several examples, and providing an outlook for where the field is headed. This is not intended to serve as a comprehensive literature review but rather to provide readers with an overview of how mass spectrometry can be used to benefit plant science research.

**Chromatography- Mass Spectrometry in Plant-omics**

Coupling chromatography to mass spectrometry, via either liquid chromatography-mass spectrometry (LC-MS) or gas chromatography-mass spectrometry (GC-MS), is a powerful technique for analyzing proteins and peptides (LC-MS), and small molecules (both LC-MS and GC-MS) as they provide a great depth of information and offer higher sensitivity, selectivity, and greater structural determination capabilities compared to methods such as IR and Raman spectroscopy. GC-MS detects small volatile or derivatized nonvolatile molecules at lower concentrations compared to other techniques for structural characterization, such as NMR.[19] Similarly, protein gels have a limited separation capacity and detection ability often suffers from bias to a specific class or size of proteins, which decreases the number of proteins that can be identified compared to MS.[20]

Proteomics

Proteomics is used to describe the comprehensive study of the proteins present in an organism, or a particular part of the organism. Thus, proteomics provides a more comprehensive understanding of the molecular components produced and utilized by organisms to sustain various processes required for life. Traditional plant proteomic studies utilize gel electrophoresis in the workflow. However, gel electrophoresis suffers from several limitations including difficulty in analyzing highly basic or acidic proteins, bias towards more abundant proteins and limited

dynamic range, making low abundance proteins hard to detect, and an inability to separate insoluble membrane proteins.[20, 21] The higher sensitivity, selectivity, and structural determination capabilities of LC-MS make it an advantageous technique for a wide range of applications in plant proteomics, including descriptive proteomics of a sample, comparative proteomics using quantification techniques, post-translation modifications, and protein-protein interactions.[22]

Proteomics studies can be classified as either top-down or bottom-up. A study by Lagrain *et al.* used a top down approach to identify high molecular weight glutenin subunits (HMW-GS), which are responsible for variability in wheat quality. HMW-GS have long repetitive sequences with few cleavage sites for trypsin, making the bottom-up approach difficult due to low protein coverage, meaning isoforms and post translation modifications are difficult to distinguish. Relative average mass matching was used to identify HMW-GS along with ω5-gliadins via LC-MS analysis. Overall, they demonstrated a useful method for the characterization of HMW-GS via top-down MS.[23] Top-down studies provide high accuracy protein structure determination, maps of post translational modifications (PTMs), and information on single amino acid polymorphisms.[24] Bottom-up analysis is more prevalent than top-down analysis for a variety of reasons, despite the usefulness of the top-down approach. Intact proteins are harder to separate than peptides as they are generally less soluble, meaning that separation of proteins prior to MS analysis is a challenge. Advances in the separation of intact proteins can improve the number of proteins identified in top down studies. For example, Tran *et al*. separated proteins with solution isoelectric focusing, gel-eluted liquid fraction entrapment electrophoresis, and nanocapillary liquid chromatography prior to MS analysis to reach a 20 fold increase in proteome coverage,[25] and Valeja *et al*. developed a 3D liquid chromatography separation involving hydrophobic interaction chromatography, ion exchange chromatography, and reversed phase chromatography to improve

protein identifications with top down analysis.[26] However, these methods have yet to be applied to plant science. Also, top down approach requires more sample than bottom up method due to low ion counts for fragmentation, and a limited number of bioinformatics tools are available for data analysis.[24]

Currently, the bottom-up approach is used much more frequently than the top-down approach, and one benefit of the bottom-up approach is that it enables the identification of proteins in complex mixtures.[21] A study by Li *et al.* investigated the changes in the proteome of *Arabidopsis thaliana* (Arabidopsis) seedlings in response to strigolactones, which are hormones that control shoot branching, as well as other aspects of growth and development. The study used iTRAQ, a common chemical label for $MS^2$ quantification, for relative proteome comparisons between a wild-type and a mutant version of Arabidopsis in the absence and presence of a strigolactone synthetic analog. Of the 2095 proteins identified across all samples in the wild-type, 19 were found to be reproducibly up-regulated (fold change greater than 1.25) and 18 were reproducibly down-regulated (25% or greater decrease). The up-regulated and down-regulated proteins for the mutant were nine and two respectively. This study identified proteins not previously known to be in the strigolactone pathway.[27]

The identification of post-translational modifications (PTMs) is a key area of proteomics as PTMs play a large role in the function of a protein. For example, phosphorylation sites regulate cell differentiation and signaling networks, and play a role in substrate specificity. In bottom-up proteomics, a specific PTM, such as phosphorylation, is enriched for after digestion, by affinity purification for example, and then LC-MS/MS analysis is performed to identify the peptides with that PTM.[28] Rose *et al.* used the model legume, *Medicago truncatula* (Medicago), to investigate phosphorylation changes due to Nod factors, which are necessary for rhizobia infection and nodule

development in the symbiotic relationship. The study identified 13,506 phosphosites in 7739 proteins and determined that Nod factor (NF) initiation resulted in changes in phosphorylation, but not overall protein levels. **Figure 2** shows the proteome and phosphoproteome changes in the wild-type, nfp mutant, and dmi3 mutant plants. Changes in phosphorylation sites in wild-type plants but not in the nfp mutant were analyzed in order to determine changes in NF signal transduction. The combination of transcriptional, translational, and post-translational experiments performed resulted in a wealth of new information regarding the Nod factor signaling cascade.[29]

The depth of information provided by bottom-up proteomics studies comes at the cost of complexity in the experimental workflow and data processing. After protein extraction, the sample can undergo multiple steps, including digestion, chemical labeling, PTM enrichment, fractionation, and desalting prior to LC-MS/MS analysis depending on the scope of the study. Similarly, much effort is required after the MS acquisition in data processing. Well-annotated gene sequences are necessary in order to create predicted $MS^2$ theoretical fragmentation spectra for matching to experimental data. Obtaining well-annotated gene sequences is one of the more complex parts of plant proteomics data analysis as the size of plant genomes can vary tremendously. For example, sequencing studies estimate genome sizes from approximately 420 Mb for rice,[30-32] to 2300 Mb for maize,[33] to 17,000 Mb for wheat.[34] In comparison to the human genome (3000 Mb), most plant genomes are just as complex, and many of the larger plant genomes are not yet sequenced.[35] Plant genomes that are not sequenced or not well-annotated pose a challenge to a data analysis process that requires well-annotated gene sequences. Also, software to match peptides to the original protein must consider the false discovery rate and how to assemble proteins from a list of peptide spectral matches when peptides may match more than one protein.[36]

Peptidomics

Peptidomics, the study of endogenous peptides produced by an organism, is a relatively new and underdeveloped field branching off of proteomics; therefore there is a growing need to investigate the role endogenous peptides play in an organism on a molecular level.[37-41] Known plant peptides include phytohormones (with signaling roles in both defense and non-defense processes), a wide range of defense peptides, and protease inhibitors. As signaling molecules, plant peptides play roles in cell division, development, nodulation, reproduction, and defense.[41, 42] Thus, peptides affect nearly all parts of the plant and a better understanding of them could lead to improvements in crop yields and pharmaceutical products.[43]

As it stands, relatively few examples illustrate the usefulness of peptidomics analysis. Ohyama *et al.* studied the peptides excreted by Arabidopsis plants into the liquid culture growth media by LC-MS analysis. To validate their method, CLE44 overexpressor plants, which encode for a known secreted peptide, were grown. Analysis by LC-MS/MS revealed the expected secreted peptide. They then investigated an uncharacterized gene family encoding secreted peptides. Analysis of an overexpressed model gene for this family revealed a 15 amino acid peptide with PTMs.[44] Chen *et al.* investigated defense signaling peptides in tomato plants. Endogenous peptides before and after stress induction were compared using a hypothetical peptide database created from predicted cleavage sites. Analysis revealed 14 new peptides and one known peptide that were up-regulated after wounding by more than 2 fold. One novel peptide, CAPE1, was found to have an expression pattern similar to a known peptide (systemin) and was investigated further.[45] Haruta *et al*. investigated the interaction of the peptide RALF in *Arabidopsis thaliana* using a metabolic $^{15}N$ labeling technique and MS analysis to determine the signaling pathway for the regulation of cell expansion. The study found that RALF interacts with the receptor FERONIA to

suppress cell elongation. Furthermore, phosphoproteome analysis showed that the RALF-FERONIA interaction phosphorylated $H^+$-adenosine triphosphate 2, which inhibits proton transport, giving a mechanism for extracellular alkanization caused by RALF.[46] As peptidomics becomes more wide spread, studies will be able to go beyond discovery experiments to further explore the functional roles of peptides in plants.

Peptidomics employs similar workflow, and data processing steps as proteomics. Although a trypsin digest is not necessary, fractionation and other separation steps are necessary for complex samples. Also, data processing requires a tandem mass spectra database for identification of the peptides. Database searching is complicated by the fact that the cleavage sites that create the peptides are not always known, and there is no trypsin digest to create predictable cleavages and -COOH groups at the C-terminus, which complicates spectra interpretation.[43] If all partial sequences from the protein database are compiled, the database size increases tremendously, thus increasing the database searching time, the number of false positives in the decoy database, and the score needed to confidently identify positive hits.[45] Additionally, plant peptidomics is further complicated by the low abundance of many plant peptides, which is further challenged by endogenous proteolytic degradation.[43] Careful sample preparation and data processing will help to provide high quality results that improve current knowledge on plant peptidomics.

Metabolomics

As the end products of various biochemical processes catalyzed by enzymes, metabolites provide useful molecular insight into an organism's biochemistry at a given time. With molecular weights typically under 1000 Da, metabolites are often classified as small molecule analytes.

Metabolites are either primary metabolites, meaning that they are essential to a plant's growth and development, or secondary metabolites, which are diverse chemicals produced by plants that may have specific impacts on other organism. Metabolomics in plants is of particular importance as plants produce many nutrients and vitamins, and genetic engineering of plants to produce more nutrients and vitamins in the edible portions of the plant is an important area of plant metabolomics research.[47] The fact that metabolites cannot be deduced simply from genetic data, unlike transcripts and proteins, creates a challenge for the field of metabolomics. Metabolomics uses a range of analytical techniques, including MS, NMR spectroscopy, and laser-induced fluorescence detection, of which MS is a selective and sensitive method commonly used, to characterize the metabolites present in an organism at a given time.[48]

The information learned from metabolomics studies in plants has the potential for great impact in improving plant response to stress, agricultural efficiency, and food quality.[47] For example, Oms-Oliu *et al.* studied the metabolites present in tomato fruit before, during, and after ripening. GC-MS analysis revealed major changes during the ripening process, such as an increases glucose, fructose, cell wall components, and amino acids. Additionally, metabolites like mannose and citramalic were found to strongly indicate a post-harvest state.[49] Metabolomics can also provide insights into a plant's response to stress. Zhang *et al.* used a combination of transcriptomics and metabolomics to study the effect of drought conditions on *Medicago truncatula*. Many thousands of transcripts changed in response to drought and GC-MS analysis identified changes to hundreds of metabolites. The combination of the two datasets helped to characterize the regulation of the metabolomics pathways under drought conditions.[50] An alternative ionization approach worth mentioning is leaf-spray mass spectrometry, a fairly new technique, based on paper-spray technology, where plant tissues cut in the shape of a triangle

produce a spray into the inlet of the mass spectrometry from either an electrolyte present in the tissue or from an added solvent upon application of an electric field.[51] Leaf spray has been used in a number of studies for detection of a variety of metabolites.[52]

Sample preparation starts with quenching the plant tissue, which stops the metabolic processes in a cell through the use of low temperatures, such as flash-freezing with liquid nitrogen.[53] An extraction of the desired metabolites follows quenching, and separation of metabolites into subclasses prior to LC or GC-MS helps with the analysis of complex samples. A wide concentration range of metabolites in conjunction with a large number of metabolites with differing chemical properties make metabolite samples complex, and challenging to analyze.[54] Data processing of the complex datasets in metabolomics experiments can influence the quality of the results.[55] Also, the indirect relationship between metabolomics and genomics creates a challenge in confidently identifying metabolites. A combination of high-resolution, accurate mass matching, MS/MS fragmentation data, and retention time comparisons between pure standards or database information and experimental data are needed for identification of metabolites.[56]

**Mass Spectrometry Imaging**

One of the exciting new developments in mass spectrometry that is becoming more widely used for plant-omics is mass spectrometry imaging (MSI). MSI shows great promise for biological studies because it allows for molecular analysis of tissue while retaining information about the spatial distribution of the different analytes in the tissue sample.[57] The conventional tissue extraction methods described above for proteomic, peptidomic, and metabolomic studies do not provide the spatial information that MSI can provide; however, liquid extraction is still the method of choice for large-scale studies as MSI is lower throughput and less reproducible when acquiring

a large amount of samples. Several other biological techniques allow for the visualization of macromolecules in various cellular structures, such as immunolocalization or other techniques that rely on the interactions between the analyte and an external probe, and therefore require clear knowledge of the analytes of interest and specifically target larger molecules such as proteins. Mass spectrometry lends itself to discovery experiments in which a multitude of analytes, from metabolites to proteins, can be detected and studied without prior knowledge of sample contents. Herein, we discuss four different MSI techniques (MALDI, SIMS, DESI, and LAESI) and how they are being applied to plant-omics. **Table 1** summarizes the optimal analytes, mass range, and spatial resolutions of these four different MSI techniques.

Before discussing the different ionization sources used for MSI, there are a variety of different types of mass analyzers that offer distinct advantages and disadvantages when it comes to MSI experiments. Arguably the most common type of mass analyzers used for MSI are time-of-flight (TOF) and TOF/TOF analyzers in which the *m/z* is determined by the time the ions take to travel from the ionization source through the TOF tube.[58] A wide mass range of molecules (theoretically unlimited) can be detected with these types of instruments with fast analysis times and MS/MS capabilities for molecular identification, however the mass accuracy and mass resolving power are low compared to other instrument types.

Orbitrap Fourier transform mass spectrometers are gaining popularity in MSI applications. In an Orbitrap mass spectrometer, *m/z* values are determined by measuring the axial oscillation frequency of ions moving back and forth along a spindle-like electrode, where the spinning ions generate a small current that is detected at either end of the electrode, and the mass of the ion is related to the oscillation frequency.[59] While the Orbitrap provides superior mass resolution and mass accuracy, disadvantages of the commercially available MALDI-Orbitrap compared to TOF

instruments are the larger laser beam size, which limits the spatial resolution of the images to approximately 75 μm (although custom instrument modifications have produced laser beams as small as 3 μm spatial resolution), slower acquisition speed thus longer analysis time, and limited mass range (maximum of 4000 Da as higher *m/z* compounds are difficult to retain during the orbital rotation). Additionally, the commercial MALDI-LTQ Orbitrap has capabilities for CID (collision-induced dissociation) and HCD (high-energy collisional dissociation) MS/MS fragmentation, which both can be used to obtain more complete fragmentation coverage of metabolites and peptides and improved molecular identifications. Hybrid methods have been developed by the Young Jin Lee group that utilize the hybrid linear ion trap and Orbitrap portions of the MALDI-LTQ Orbitrap instrument to reduce the data acquisition time while collecting even more molecular information by interspersing ion trap scans and MS/MS scans during the longer Orbitrap scan using a spiral step plate motion.[60]

Fourier transform ion cyclotron resonance (FT-ICR) mass analyzers are also used for some MSI applications. In this type of instrument, ions are trapped in a magnetic field and excited by an electric field, thus leaving the ions rotating at their cyclotron frequency that generates a current which is detected by the mass spectrometer. FT-ICR mass analyzers can provide extremely accurate mass measurements, has MS/MS capabilities, and have a higher upper mass limit of detection compared to Orbitraps; however, like the Orbitrap, suffer from long analysis times.

Matrix-Assisted Laser Desorption/Ionization

Matrix-assisted laser desorption/ionization (MALDI)-MSI is the most popular type of MSI. A workflow comparing the sample preparation, data acquisition, and examples of raw data for LC-MS and MALDI-MSI workflows is shown in **Figure 3**. For this technique, tissues of interest are

trimmed or harvested from the plant and immediately frozen before being sectioned on a cryostat into slices that are typically 10-20 μm thick.[61] Next, the tissue sections are thaw-mounted onto a stainless steel conductive plate, indium-tin-oxide (ITO)-coated conductive glass slides, or regular glass slides, depending on instrumentation.[62] The next step, MALDI matrix selection and application, is critical to the experimental workflow as the type of matrix chosen will determine the types of analytes that can be ionized and the application method will control the matrix crystal size which in turn defines the maximum image resolution.[63, 64] Conventional matrices include 2,5-dihydroxybenzoic acid (DHB) and α-cyano-4-hydroxycinnamic acid (CHCA)[65] for positive mode and 9-aminoacridine (9-AA)[66] for negative imaging mode. Less traditional matrices such as $TiO_2$, gold, or silver nanoparticles[67-69], 1,5-diaminonapthalene (DAN)[70], 2,3,4,5-tetrakis(3′,4′-dihydroxylphenyl)thiophene (DHPT)[71], and 1,8-bis(dimethyl-amino) naphthalene (DMAN)[72, 73] are being used more frequently and are reported to improve spectral quality, crystallization and vacuum stability.[74] Spatial resolution and reproducibility of results are limited by the matrix crystal size and application consistency, among other instrumental parameters such as raster step size and laser beam diameter.[75] Matrix can be applied manually with an artist airbrush, with automated systems[76], via sublimation[77], or filtered onto tissue through a small (~20 μm) sieve or mesh.[78, 79] The matrix allows for ions to be generated, which is essential for mass spectrometric analyses as the mass spectrometer can only detect charged particles.

After the matrix is applied to the sample and the analytes are co-crystallized with matrix, the sample plate is loaded into the mass spectrometer on a moveable X-Y stage. A laser is fired at the sample plate, causing the matrix to be ablated and form a gas plume in which matrix and analyte ions are formed. The moveable X-Y stage allows for the laser to raster across the sample and mass spectra to be collected at every pre-defined raster point. After completing the 2D raster, ion images

for each mass in the spectrum can be shown and the software will display the relative abundance of each ion as a colored map of signal intensity across the raster area.

One advantage MALDI-MSI for biological analysis is that it can generate larger ions, such as peptides and proteins, which is one of the reasons why MALDI is the most widely-used MSI technique.[57, 80-82] The matrix absorbs much of the incident laser energy, providing soft ionization for analytes , which allows for the ionization of larger molecules (*m/z* over 100 kDa).[83] Unfortunately, while there are advantages for the analysis in the high mass range, there are disadvantages in the low mass range. One disadvantage to using MALDI is that the matrix itself produces ions that can interfere or mask analyte ions in the small molecule mass range. This can be somewhat avoided using high resolution instrumentation or novel matrices.[67, 70-72, 84-88] There are several excellent reviews that highlight the application of MALDI-MSI (and other MSI techniques) to plant metabolomics, peptidomics, and proteomics.[89-94]

As previously stated, MALDI-MSI is the most highly used type of MS imaging, possibly due to its versatility and availability. MALDI-MSI can be used to detect a wide range of molecules from metabolites to peptides to proteins; although there are currently far more studies using MALDI-MSI for plant metabolomics. Our lab recently used MALDI-MSI to study the symbiotic relationship between the model legume, *Medicago truncatula,* and the nitrogen fixing bacteria, *Sinorhizobium meliloti*. We compared combinations of wild-type (wt) and mutant strains of plant (*dnf1*) and bacteria (*fixJ*) and identified metabolites, such as heme, various amino acids, and other organic acids that are present in functional, nitrogen-fixing nodules (wt-wt) and absent from non-functional nodules (wt-*fixJ*, *dnf1*-wt, or *dnf1-fixJ*).[95, 96] **Figure 4** shows representative images of some of the metabolites detected in this study. The information gained from using this technique could provide valuable insight into the biological nitrogen fixation process in legume plants by not

only identifying metabolites that may be involved in this process, but by also localizing these molecules within the root nodule structures. Collaborative work between the Young Jin Lee and Basil Nikolau groups highlights MALDI-MSI for subcellular-level imaging of metabolites in maize leaves, providing insights into the known asymmetric metabolism associated with different tissue types within these leaves.[97] This proof-of-principle study of understanding molecular details and localizations at the sub-cellular level shows exciting promise for deeper interrogation of the biological mechanisms and implications. Plant peptidomics is a relatively under-explored area in mass spectrometry, and especially so in mass spectrometry imaging. One of the very few reports on plant peptide imaging focuses on imaging cyclotides in petunias. Using MALDI-MSI, Poth *et al.* were able to detect novel cyclotides (defense-related cyclic plant peptides) from petunia leaves.[98] The goal with this study would be to genetically enrich major crop species to be able to express cyclic peptides of their own, thus enhancing crop protection. While mass spectrometry analysis is the gold standard for proteomic analysis in general, there are very few reports using MALDI-MSI for plant proteomics. A previous review highlights MALDI-MSI of proteins in soybean cotyledons in a proof-of –principle figure,[92] and a known allergenic protein in peaches was shown to be localized to the skin of the peach using MSI,[99] but no further applications of MSI to plant proteomics have been reported so far.

Secondary Ion Mass Spectrometry

Secondary ion mass spectrometry (SIMS) is a long established technique and, unlike MALDI, it does not require special preparation or matrix application, although some optional methods, such as the addition of a matrix or a thin coat of metal nanoparticles, can improve the

imaging results.[100, 101] Remarkably, SIMS is able to reliably achieve less than 1 μm spatial resolution, making it the method of choice for subcellular investigations.

Like MALDI-MSI, in SIMS-MSI the sample is put under high vacuum. Instead of ionizing with a laser like in MALDI, the sample on the surface is bombarded with high energy primary ions which facilitate the ionization of the analytes.[102, 103] The ionized analytes, or secondary ions, are sputtered from the sample surface into the mass analyzer for analysis. The highly focused ion beam in SIMS provides excellent spatial resolution (less than 1 μm); however, the primary ion source is fairly limited to molecules under 1000 Da because it is a high energy beam that easily fragments larger molecules, and for this reason, SIMS is especially popular for elemental and atomic analyses.[96, 104] SIMS-MSI of primary and secondary metabolites in pea plants and *Arabidopsis thaliana* with sub-2 μm spatial resolution has been reported by Seyer *et al*., as shown in **Figure 5**.[105] A new sample preparation method was developed for SIMS-MSI of Arabidopsis and the method was applied to confirm variations in flavonoid content in seeds from different Arabidopsis mutants. While SIMS is one of the oldest MS techniques, its application to plant-omics is lacking and could be significantly expanded in the future.

Desorption Electrospray Ionization

Although the spatial information of a sample is preserved with MSI, MSI has still been mainly viewed as an invasive process until the development of ambient ionization techniques, such as desorption electrospray ionization (DESI). DESI is a simple, ambient ionization technique that channels charged solvent droplets and ions from an electrospray source onto the surface of the sample, and the impact yields gaseous analyte ions.[106] Images can either be collected directly from the sample surface or indirectly via imprints. Like SIMS, DESI also does not require matrix to

create ions and actually requires little to no sample preparation at all other than sectioning and mounting the sample.[107] One disadvantage DESI has compared to vacuum MS methods, SIMS and MALDI, is that it has lower spatial resolution of approximately 180–200 μm.[107]

Like the other MSI techniques mentioned above, DESI has been primarily utilized for metabolomics studies. Recently published work by Tata *et al*. used DESI-MSI to study changes in the potato plant metabolome in response to pathogen invasion.[108] Their method of imprinting the sample on tape and performing DESI-MS and DESI-MSI on the imprint allowed for simple and rapid metabolomics profiling with minimal sample preparation, and resulted in the first qualitative study of plant defense against phytopathogen invasion via ambient mass spectrometry. Hemalatha and Pradeep present a thorough analysis of molecular signatures (primarily metabolites) from a variety of plant species and tissues using DESI-MSI.[109] This work identified metabolites with implications in varietal differences, toxic metabolite production, metabolites during plant growth cycles, plant defense, etc.

Laser Ablation Electrospray Ionization

Laser ablation electrospray ionization (LAESI) is another type of ambient ionization and utilizes the natural water content of cells and tissues as the matrix.[110] With LAESI, a laser is used to ablate molecules in the tissue which are then captured by the electrospray and generate analyte ions.[111] The Vertes group invented and developed LAESI-MSI for plant sciences. They have demonstrated the application of LAESI-MSI for live, untreated tissue analysis of metabolites on plant leaves[112] and have even demonstrated the use of LAESI for 3D MSI of metabolites in *S. lynise* and *A. squarrosa* leaf tissue as shown in **Figure 6**.[113] 3D imaging of molecular distributions can elucidate the correlation between biochemical processes and the spatial organization of a

biological tissue. More recently this group has coupled LAESI-MSI with ion mobility mass spectrometry (IM-MS) of metabolites on *P. peltatum* leaves.[114] They were able to use this technique to separate structural isomers which is very important as isomers can often have extremely different biological activities. This technique can theoretically be applied to plant peptidomics and proteomics as well, but has so far only been demonstrated for the direct detection of these larger molecules on animal tissue.[115]

**Data Analysis**

Mass spectrometry experiments have the capacity to generate huge amounts of data. Sorting through and analyzing data is often the greatest bottleneck of MS and MSI experiments. There have been many different software tools developed for the analysis of metabolomics, peptidomics, or proteomics datasets.

Metabolomics data processing typically begins with matching obtained accurate masses to online metabolite databases such as METLIN,[116] HMDB (Human Metabolome Database),[117, 118] KEGG (Kyoto Encyclopedia of Genes and Genomes),[119] PubChem,[120] MassBank,[121] MMCD (Madison Metabolomics Consortium Database),[122] LIPID MAPS,[123, 124] and more. Database searching will often result in multiple potential identifications for a given mass; therefore, to further narrow down identification assignments, tandem mass spectra collected for each metabolite of interest can be compared to metabolite standards, tandem mass databases, or *in silico* fragmentation tools such as MetFrag.[125] Instrument vendors have released software packages for metabolomics analyses such as ProgenesisQI (Waters Corporation), Sieve (Thermo Scientific), ProfileAnalysis (Bruker Daltonics), to name a few, as well as open-source software tools like XCMS[126, 127] and MetaboAnalyst.[128-130] These tools allow for data alignment, normalization,

automatic feature selection, some statistical analysis, and even metabolite identification prediction based on MS/MS data, although confident identifications is still lacking. Software tool development for metabolomics is growing area that could still use further development as many metabolites share common fragment masses with no consistent fragmentation rules for different functional groups (unlike peptide fragmentation) that would make accurate fragmentation prediction achievable at this time, leaving many ambiguous identifications.

Accurate mass matching can also be used to identify peptides and proteins, although the more widely-accepted approach is to match MS/MS data to sequenced genomes, using software tools like Mascot, or by *de novo* sequencing, with software packages like PEAKS[131] which is also useful for endogenous peptidomics studies.[131] Several programs, such as BLAST (Basic Local Alignment Search Tool)[132] and MEME,[133] have been developed for homology searches by comparing putative peptide sequences against a database of closely related species which can provide insight into key evolutionary and functional roles of peptides.[134] In comparison to metabolomics, the more mature field of proteomics has many software tools, such as Mascot (by Matrix Science), COMPASS (Coon OMSSA Proteomic Analysis Software Suite),[135] Proteome Discoverer (by Thermo Scientific), SkyLine,[136] Morpheus,[137] etc., available that perform protein identifications based on sequencing of peptide fragment ions and protein database searching.

MSI data processing is often less straightforward and less automated than traditional extraction-based analyses. MSI data processing typically involves manually extracting ion images for each specific *m/z* in the mass spectrum and physically checking the quality of the image by making sure the analyte's spatial distribution is localized to the tissue with no contaminating ions interfering with the quality of the image. A list of *m/z* values with high quality MS images can then be compiled for further interpretation. Vendor specific MSI analysis software (e.g., FlexImaging

from Bruker Daltonics, ImageQuest from Thermo Scientific, HDImaging from Waters Corporation, and TissueView from Applied Biosystems/MDS) or open source software such as BioMap are commonly used to extract ion images for selected analytes from the collected data. Fairly recently, MSiReader, a readily available MSI data processing software, has been developed that has drastically decreased the MSI data processing time in comparison to manual data processing, allowing for higher throughput and faster overall analysis.[138] Large-scale MSI datasets can easily be tens of GB in size for a single experiment and management of these enormous data files causes significant challenges in data processing and analysis. Several key techniques have been developed to specifically address the information and analysis of large MSI datasets, such as OpenMSI[139] and memory efficient algorithms for principal component analysis.[140]

Many bioinformatics tools exist for analyzing a specific type of –omics datasets. Recently, researchers have been working toward developing plant-specific databases and software tools that can integrate multiple –omics datasets. The Sumner group reported a tandem mass spectral library of plant natural products, created using authentic standards and purified compounds.[141] The database focuses primarily on plant secondary metabolites, and contains retention time data as well as tandem mass spectra collected at six different collision energies. Other plant specific databases include the Medicago PhosphoProtein Database[142], The Plant Proteome Database for *Arabidopsis thaliana* and maize[143], the Rice Proteome Database[144], and Promex which contains tryptic peptide mass spectra information for a dozen plant species[145]. In addition to plant specific databases for metabolomics or proteomics data, other software tools and databases have been developed to integrate multiple –omics datasets for a more comprehensive systems-biology view of the results. The Sumner group has developed MedicCyc specifically for *Medicago truncatula* metabolic pathway reconstruction with over 250 pathways of related metabolites, enzymes, and genes.[146]

Similarly, AraCyc[147], MetaCyc[148], the Plant Metabolic Network (Plant Metabolic Network (PMN) on www.plantcyc.org) , and KEGG[149] are pathway tools for various other plant species.

**Outlook**

With regard to an integrated –omics strategy, there are obvious concerns about the complexity and dynamic range of the different classes of analytes being studied. These challenges have motivated the development of analytical workflows incorporating multi-dimensional front-end separation strategies and more advanced instrumentation offering superior separation capabilities and enhanced MS detection with faster scanning rates, better sensitivity, and higher performance. MALDI-MSI instrumentation with spatial resolution down to approximately 20 μm is readily available. Custom built instrumentation can achieve a spatial resolution of < 3 μm and it is only a matter of time before these custom instruments become commercially available.[150-153] Improvements to ambient ionization techniques can also provide more biologically relevant information about the plant systems. Adopting these new technologies with important new capabilities, along with careful sample preparation, can provide dramatic improvements when integrated into -omic workflows.

In addition to technological advances in instrumentation, the simultaneous development, advancement, and integration of closely related –omics approaches can guide researchers to a better understanding and characterization of various biological processes and signaling mechanisms in plants. Integrated genomics and metabolomics or metabolomics and transcriptomics studies have recently been reported, and we expect more exciting studies of this nature in the future.[154-156] By integrating multiple -omics technologies it may be possible to obtain extensive molecular-level information and elucidate biological functions and underlying

mechanisms. Plant-specific bioinformatics tools for integrating multi-omics strategies are beginning to emerge, but continuous development of wide-reaching, user-friendly software tools would drastically improve data analysis strategies and holistic knowledge.

A more widespread adoption of MSI as a technique for –omics studies will provide spatial information of where analytes are located within the plant. Knowing, for example, where metabolites and the proteins/enzymes that interact with them are located at different developmental stages could reveal key insights into biological processes and mechanisms within the plant. In addition to studying plant development, MSI can also provide a way to localize differences in analytes between tissue types, disease states, genetic differences, or following genetic manipulation. It could be possible to quickly ascertain the functional potential of plants for various biotechnological applications such as bioenergy production, environmental decontamination, natural product production, and many other important applications. Collaborations between experts in various plant-omic methodologies, plant scientists, and bioinformaticians, along with ever-advancing technological innovations, impart an exciting outlook for the future of plant science.

We believe that the application of mass spectrometry technologies to plant systems-biology will greatly accelerate the progress of this field. It allows for high-throughput, large-scale analyses on the metabolomic, peptidomic, and proteomic level that can correlate with genomic information. Mass spectrometry can be used to compare the metabolic and proteomic profiles of wild-type plants to a variety of genetic mutants where knocking-out certain genes will have a known or unknown physiological response. A multi-omics profile comparison between mutants with known and unknown responses will allow researchers to infer the previously unknown function if the –omics profiles correlate well with those of known function. New software and bioinformatics tools

are needed to interrogate enormous datasets generated with these high-throughput, information-rich mass spectrometry approaches. The new insights gleaned from these large-scale systems biology experiments can lead to new hypotheses and novel findings that will ultimately advance our knowledge of fundamental plant biology and plant biotechnology.

**Acknowledgements**

Support for this research was provided by the University of Wisconsin–Madison (UW-Madison), Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation (WARF). E.G. acknowledges a National Science Foundation (NSF) Graduate Research Fellowship (DGE-1256259). LL acknowledges an H. I. Romnes Faculty Research Fellowship and a Vilas Distinguished Achievement Professorship with funding provided by the WARF and UW-Madison School of Pharmacy.

**References**

1. Taylor, T., Bacteria and Their Relations to Plant Culture. *Science* **1881,** *2* (65), 443-4.

2. Wiegand, K. M., Plant Embryo-Sacs. *Science* **1900,** *12* (296), 347.

3. Smith, E. F.; Townsend, C. O., A Plant-Tumor of Bacterial Origin. *Science* **1907,** *25* (643), 671-3.

4. Quinn, L. D.; Gordon, D. R.; Glaser, A.; Lieurance, D.; Flory, S. L., Bioenergy Feedstocks at Low Risk for Invasion in the USA: a "White List" Approach. *Bioenergy Research* **2015,** *8*, 471-481.

5. Tadege, M.; Chen, F.; Murray, J.; Wen, J.; Ratet, P.; Udvardi, M. K.; Dixon, R. A.; Mysore, K. S., Control of Vegetative to Reproductive Phase Transition Improves Biomass Yield and Simultaneously Reduces Lignin Content in Medicago truncatula. *BioEnergy Research* **2015,** *8* (2), 857-867.

6. Kumar, S.; Kumaria, S.; Tandon, P., Efficient In vitro Plant Regeneration Protocol from Leaf Explant of Jatropha curcas L - A Promising Biofuel Plant. *J Plant Biochem Biot* **2010,** *19* (2), 275-277.

7. Hung, K. H.; Chiang, T. Y.; Chiu, C. T.; Hsu, T. W.; Ho, C. W., Isolation and characterization of microsatellite loci from a potential biofuel plant Miscanthus sinensis (Poaceae). *Conserv Genet* **2009,** *10* (5), 1377-1380.

8. Dewick, P. M., *Medicinal Natural Products: A Biosynthetic Approach*. 3 ed.; Wiley: Chichester, 2009.

9. Russo, P.; Frustaci, A.; Del Bufalo, A.; Fini, M.; Cesario, A., Multitarget drugs of plants origin acting on Alzheimer's disease. *Current medicinal chemistry* **2013,** *20* (13), 1686-93.

10. Kano, S., Artemisinin-based combination therapies and their introduction in Japan. *Kansenshogaku Zasshi* **2014,** *88* (3 Suppl 9-10), 18-25.

11. Kittakoop, P.; Mahidol, C.; Ruchirawat, S., Alkaloids as Important Scaffolds in Therapeutic Drugs for the Treatments of Cancer, Tuberculosis, and Smoking Cessation. *Curr Top Med Chem* **2014,** *14* (2), 239-252.

12. Dias, D. A.; Urban, S.; Roessner, U., A historical overview of natural products in drug discovery. *Metabolites* **2012,** *2* (2), 303-36.

13. Zheng, Y.; Jiang, X.; Gao, F.; Song, J.; Sun, J.; Wang, L.; Sun, X.; Lu, Z.; Zhang, H., Identification of plant-derived natural products as potential inhibitors of the Mycobacterium tuberculosis proteasome. *BMC complementary and alternative medicine* **2014,** *14*, 400.

14. Lucas, D. M.; Still, P. C.; Perez, L. B.; Grever, M. R.; Kinghorn, A. D., Potential of plant-derived natural products in the treatment of leukemia and lymphoma. *Current drug targets* **2010,** *11* (7), 812-22.

15. Lee, K. H., Anticancer drug design based on plant-derived natural products. *Journal of biomedical science* **1999,** *6* (4), 236-50.

16. Hung, H. Y.; Qian, K.; Morris-Natschke, S. L.; Hsu, C. S.; Lee, K. H., Recent discovery of plant-derived anti-diabetic natural products. *Natural product reports* **2012,** *29* (5), 580-606.

17. Dholwani, K. K.; Saluja, A. K.; Gupta, A. R.; Shah, D. R., A review on plant-derived natural products and their analogs with anti-tumor activity. *Indian journal of pharmacology* **2008,** *40* (2), 49-58.

18. Ehrhardt, D. W.; Frommer, W. B., New technologies for 21st century plant science. *The Plant cell* **2012,** *24* (2), 374-94.

19. Dunn, W. B.; Bailey, N. J. C.; Johnson, H. E., Measuring the metabolome: current analytical technologies. *Analyst* **2005,** *130* (5), 606-625.

20. Roe, M. R.; Griffin, T. J., Gel-free mass spectrometry-based high throughput proteomics: tools for studying biological response of proteins and proteomes. *Proteomics* **2006,** *6* (17), 4678-87.

21. Zhang, Y. Y.; Fonslow, B. R.; Shan, B.; Baek, M. C.; Yates, J. R., Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical Reviews* **2013,** *113* (4), 2343-2394.

22. Jorrin-Novo, J. V.; Pascual, J.; Sanchez-Lucas, R.; Romero-Rodriguez, M. C.; Rodriguez-Ortega, M. J.; Lenz, C.; Valledor, L., Fourteen years of plant proteomics reflected in Proteomics: Moving from model species and 2DE-based approaches to orphan species and gel-free platforms. *Proteomics* **2015,** *15* (5-6), 1089-1112.

23. Lagrain, B.; Brunnbauer, M.; Rombouts, I.; Koehler, P., Identification of intact high molecular weight glutenin subunits from the wheat proteome using combined liquid chromatography-electrospray ionization mass spectrometry. *PLoS One* **2013,** *8* (3), e58682.

24. Armirotti, A.; Damonte, G., Achievements and perspectives of top-down proteomics. *Proteomics* **2010,** *10* (20), 3566-3576.

25. Tran, J. C.; Zamdborg, L.; Ahlf, D. R.; Lee, J. E.; Catherman, A. D.; Durbin, K. R.; Tipton, J. D.; Vellaichamy, A.; Kellie, J. F.; Li, M.; Wu, C.; Sweet, S. M.; Early, B. P.; Siuti, N.; LeDuc, R. D.; Compton, P. D.; Thomas, P. M.; Kelleher, N. L., Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* **2011,** *480* (7376), 254-8.

26. Valeja, S. G.; Xiu, L.; Gregorich, Z. R.; Guner, H.; Jin, S.; Ge, Y., Three dimensional liquid chromatography coupling ion exchange chromatography/hydrophobic interaction chromatography/reverse phase chromatography for effective protein separation in top-down proteomics. *Anal Chem* **2015,** *87* (10), 5363-71.

27. Li, Z.; Czarnecki, O.; Chourey, K.; Yang, J.; Tuskan, G. A.; Hurst, G. B.; Pan, C. L.; Chen, J. G., Strigolactone-Regulated Proteins Revealed by iTRAQ-Based Quantitative Proteomics in Arabidopsis. *Journal of Proteome Research* **2014,** *13* (3), 1359-1372.

28. Grimsrud, P. A.; Swaney, D. L.; Wenger, C. D.; Beauchene, N. A.; Coon, J. J., Phosphoproteomics for the Masses. *Acs Chemical Biology* **2010,** *5* (1), 105-119.

29. Rose, C. M.; Venkateshwaran, M.; Volkening, J. D.; Grimsrud, P. A.; Maeda, J.; Bailey, D. J.; Park, K.; Howes-Podoll, M.; den Os, D.; Yeun, L. H.; Westphall, M. S.; Sussman, M. R.; Ane, J. M.; Coon, J. J., Rapid Phosphoproteomic and Transcriptomic Changes in the Rhizobia-legume Symbiosis. *Molecular & Cellular Proteomics* **2012,** *11* (9), 724-744.

30. Yu, J.; Hu, S. N.; Wang, J.; Wong, G. K. S.; Li, S. G.; Liu, B.; Deng, Y. J.; Dai, L.; Zhou, Y.; Zhang, X. Q.; Cao, M. L.; Liu, J.; Sun, J. D.; Tang, J. B.; Chen, Y. J.; Huang, X. B.; Lin, W.; Ye, C.; Tong, W.; Cong, L. J.; Geng, J. N.; Han, Y. J.; Li, L.; Li, W.; Hu, G. Q.; Huang, X. G.; Li, W. J.; Li, J.; Liu, Z. W.; Li, L.; Liu, J. P.; Qi, Q. H.; Liu, J. S.; Li, L.; Li, T.; Wang, X. G.; Lu, H.; Wu, T. T.; Zhu, M.; Ni, P. X.; Han, H.; Dong, W.; Ren, X. Y.; Feng,

X. L.; Cui, P.; Li, X. R.; Wang, H.; Xu, X.; Zhai, W. X.; Xu, Z.; Zhang, J. S.; He, S. J.; Zhang, J. G.; Xu, J. C.; Zhang, K. L.; Zheng, X. W.; Dong, J. H.; Zeng, W. Y.; Tao, L.; Ye, J.; Tan, J.; Ren, X. D.; Chen, X. W.; He, J.; Liu, D. F.; Tian, W.; Tian, C. G.; Xia, H. G.; Bao, Q. Y.; Li, G.; Gao, H.; Cao, T.; Wang, J.; Zhao, W. M.; Li, P.; Chen, W.; Wang, X. D.; Zhang, Y.; Hu, J. F.; Wang, J.; Liu, S.; Yang, J.; Zhang, G. Y.; Xiong, Y. Q.; Li, Z. J.; Mao, L.; Zhou, C. S.; Zhu, Z.; Chen, R. S.; Hao, B. L.; Zheng, W. M.; Chen, S. Y.; Guo, W.; Li, G. J.; Liu, S. Q.; Tao, M.; Wang, J.; Zhu, L. H.; Yuan, L. P.; Yang, H. M., A draft sequence of the rice genome (Oryza sativa L. ssp indica). *Science* **2002,** *296* (5565), 79-92.

31. Matsumoto, T.; Wu, J. Z.; Kanamori, H.; Katayose, Y.; Fujisawa, M.; Namiki, N.; Mizuno, H.; Yamamoto, K.; Antonio, B. A.; Baba, T.; Sakata, K.; Nagamura, Y.; Aoki, H.; Arikawa, K.; Arita, K.; Bito, T.; Chiden, Y.; Fujitsuka, N.; Fukunaka, R.; Hamada, M.; Harada, C.; Hayashi, A.; Hijishita, S.; Honda, M.; Hosokawa, S.; Ichikawa, Y.; Idonuma, A.; Iijima, M.; Ikeda, M.; Ikeno, M.; Ito, K.; Ito, S.; Ito, T.; Ito, Y.; Ito, Y.; Iwabuchi, A.; Kamiya, K.; Karasawa, W.; Kurita, K.; Katagiri, S.; Kikuta, A.; Kobayashi, H.; Kobayashi, N.; Machita, K.; Maehara, T.; Masukawa, M.; Mizubayashi, T.; Mukai, Y.; Nagasaki, H.; Nagata, Y.; Naito, S.; Nakashima, M.; Nakama, Y.; Nakamichi, Y.; Nakamura, M.; Meguro, A.; Negishi, M.; Ohta, I.; Ohta, T.; Okamoto, M.; Ono, N.; Saji, S.; Sakaguchi, M.; Sakai, K.; Shibata, M.; Shimokawa, T.; Song, J. Y.; Takazaki, Y.; Terasawa, K.; Tsugane, M.; Tsuji, K.; Ueda, S.; Waki, K.; Yamagata, H.; Yamamoto, M.; Yamamoto, S.; Yamane, H.; Yoshiki, S.; Yoshihara, R.; Yukawa, K.; Zhong, H. S.; Yano, M.; Sasaki, T.; Yuan, Q. P.; Shu, O. T.; Liu, J.; Jones, K. M.; Gansberger, K.; Moffat, K.; Hill, J.; Bera, J.; Fadrosh, D.; Jin, S. H.; Johri, S.; Kim, M.; Overton, L.; Reardon, M.; Tsitrin, T.; Vuong, H.; Weaver, B.; Ciecko, A.; Tallon, L.; Jackson, J.; Pai, G.; Van Aken, S.; Utterback, T.; Reidmuller, S.; Feldblyum, T.; Hsiao, J.; Zismann, V.; Iobst, S.; de Vazeille, A. R.; Buell, C. R.; Ying, K.; Li, Y.; Lu, T. T.; Huang, Y. C.; Zhao, Q.; Feng, Q.; Zhang, L.; Zhu, J. J.; Weng, Q. J.; Mu, J.; Lu, Y. Q.; Fan, D. L.; Liu, Y. L.; Guan, J. P.; Zhang, Y. J.; Yu, S. L.; Liu, X. H.; Zhang, Y.; Hong, G. F.; Han, B.; Choisne, N.; Demange, N.; Orjeda, G.; Samain, S.; Cattolico, L.; Pelletier, E.; Couloux, A.; Segurens, B.; Wincker, P.; D'Hont, A.; Scarpelli, C.; Weissenbach, J.; Salanoubat, M.; Quetier, F.; Yu, Y.; Kim, H. R.; Rambo, T.; Currie, J.; Collura, K.; Luo, M. Z.; Yang, T. J.; Ammiraju, J. S. S.; Engler, F.; Soderlund, C.; Wing, R. A.; Palmer, L. E.; de la Bastide, M.; Spiegel, L.; Nascimento, L.; Zutavern, T.; O'Shaughnessy, A.; Dike, S.; Dedhia, N.; Preston, R.; Balija, V.; McCombie, W. R.; Chow, T. Y.; Chen, H. H.; Chung, M. C.; Chen, C. S.; Shaw, J. F.; Wu, H. P.; Hsiao, K. J.; Chao, Y. T.; Chu, M. K.; Cheng, C. H.; Hour, A. L.; Lee, P. F.; Lin, S. J.; Lin, Y. C.; Liou, J. Y.; Liu, S. M.; Hsing, Y. I.; Raghuvanshi, S.; Mohanty, A.; Bharti, A. K.; Gaur, A.; Gupta, V.; Kumar, D.; Ravi, V.; Vij, S.; Kapur, A.; Khurana, P.; Khurana, P.; Khurana, J. P.; Tyagi, A. K.; Gaikwad, K.; Singh, A.; Dalal, V.; Srivastava, S.; Dixit, A.; Pal, A. K.; Ghazi, I. A.; Yadav, M.; Pandit, A.; Bhargava, A.; Sureshbabu, K.; Batra, K.; Sharma, T. R.; Mohapatra, T.; Singh, N. K.; Messing, J.; Nelson, A. B.; Fuks, G.; Kavchok, S.; Keizer, G.; Llaca, E. L. V.; Song, R. T.; Tanyolac, B.; Young, S.; Il, K. H.; Hahn, J. H.; Sangsakoo, G.; Vanavichit, A.; de Mattos, L. A. T.; Zimmer, P. D.; Malone, G.; Dellagostin, O.; de Oliveira, A. C.; Bevan, M.; Bancroft, I.; Minx, P.; Cordum, H.; Wilson, R.; Cheng, Z. K.; Jin, W. W.; Jiang, J. M.; Leong, S. A.; Iwama, H.; Gojobori, T.; Itoh, T.; Niimura, Y.; Fujii, Y.; Habara, T.; Sakai, H.; Sato, Y.; Wilson, G.; Kumar, K.; McCouch, S.; Juretic, N.; Hoen, D.; Wright, S.; Bruskiewich, R.; Bureau, T.; Miyao, A.; Hirochika, H.; Nishikawa, T.; Kadowaki, K.;

Sugiura, M.; Project, I. R. G. S., The map-based sequence of the rice genome. *Nature* **2005,** *436* (7052), 793-800.

32. Goff, S. A.; Ricke, D.; Lan, T. H.; Presting, G.; Wang, R.; Dunn, M.; Glazebrook, J.; Sessions, A.; Oeller, P.; Varma, H.; Hadley, D.; Hutchison, D.; Martin, C.; Katagiri, F.; Lange, B. M.; Moughamer, T.; Xia, Y.; Budworth, P.; Zhong, J.; Miguel, T.; Paszkowski, U.; Zhang, S.; Colbert, M.; Sun, W. L.; Chen, L.; Cooper, B.; Park, S.; Wood, T. C.; Mao, L.; Quail, P.; Wing, R.; Dean, R.; Yu, Y.; Zharkikh, A.; Shen, R.; Sahasrabudhe, S.; Thomas, A.; Cannings, R.; Gutin, A.; Pruss, D.; Reid, J.; Tavtigian, S.; Mitchell, J.; Eldredge, G.; Scholl, T.; Miller, R. M.; Bhatnagar, S.; Adey, N.; Rubano, T.; Tusneem, N.; Robinson, R.; Feldhaus, J.; Macalma, T.; Oliphant, A.; Briggs, S., A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). *Science* **2002,** *296* (5565), 92-100.

33. Schnable, P. S.; Ware, D.; Fulton, R. S.; Stein, J. C.; Wei, F.; Pasternak, S.; Liang, C.; Zhang, J.; Fulton, L.; Graves, T. A.; Minx, P.; Reily, A. D.; Courtney, L.; Kruchowski, S. S.; Tomlinson, C.; Strong, C.; Delehaunty, K.; Fronick, C.; Courtney, B.; Rock, S. M.; Belter, E.; Du, F.; Kim, K.; Abbott, R. M.; Cotton, M.; Levy, A.; Marchetto, P.; Ochoa, K.; Jackson, S. M.; Gillam, B.; Chen, W.; Yan, L.; Higginbotham, J.; Cardenas, M.; Waligorski, J.; Applebaum, E.; Phelps, L.; Falcone, J.; Kanchi, K.; Thane, T.; Scimone, A.; Thane, N.; Henke, J.; Wang, T.; Ruppert, J.; Shah, N.; Rotter, K.; Hodges, J.; Ingenthron, E.; Cordes, M.; Kohlberg, S.; Sgro, J.; Delgado, B.; Mead, K.; Chinwalla, A.; Leonard, S.; Crouse, K.; Collura, K.; Kudrna, D.; Currie, J.; He, R.; Angelova, A.; Rajasekar, S.; Mueller, T.; Lomeli, R.; Scara, G.; Ko, A.; Delaney, K.; Wissotski, M.; Lopez, G.; Campos, D.; Braidotti, M.; Ashley, E.; Golser, W.; Kim, H.; Lee, S.; Lin, J.; Dujmic, Z.; Kim, W.; Talag, J.; Zuccolo, A.; Fan, C.; Sebastian, A.; Kramer, M.; Spiegel, L.; Nascimento, L.; Zutavern, T.; Miller, B.; Ambroise, C.; Muller, S.; Spooner, W.; Narechania, A.; Ren, L.; Wei, S.; Kumari, S.; Faga, B.; Levy, M. J.; McMahan, L.; Van Buren, P.; Vaughn, M. W.; Ying, K.; Yeh, C. T.; Emrich, S. J.; Jia, Y.; Kalyanaraman, A.; Hsia, A. P.; Barbazuk, W. B.; Baucom, R. S.; Brutnell, T. P.; Carpita, N. C.; Chaparro, C.; Chia, J. M.; Deragon, J. M.; Estill, J. C.; Fu, Y.; Jeddeloh, J. A.; Han, Y.; Lee, H.; Li, P.; Lisch, D. R.; Liu, S.; Liu, Z.; Nagel, D. H.; McCann, M. C.; SanMiguel, P.; Myers, A. M.; Nettleton, D.; Nguyen, J.; Penning, B. W.; Ponnala, L.; Schneider, K. L.; Schwartz, D. C.; Sharma, A.; Soderlund, C.; Springer, N. M.; Sun, Q.; Wang, H.; Waterman, M.; Westerman, R.; Wolfgruber, T. K.; Yang, L.; Yu, Y.; Zhang, L.; Zhou, S.; Zhu, Q.; Bennetzen, J. L.; Dawe, R. K.; Jiang, J.; Jiang, N.; Presting, G. G.; Wessler, S. R.; Aluru, S.; Martienssen, R. A.; Clifton, S. W.; McCombie, W. R.; Wing, R. A.; Wilson, R. K., The B73 maize genome: complexity, diversity, and dynamics. *Science* **2009,** *326* (5956), 1112-5.

34. A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science* **2014,** *345* (6194), 1251788.

35. Hamilton, J. P.; Buell, C. R., Advances in plant genome sequencing. *Plant Journal* **2012,** *70* (1), 177-190.

36. Nesvizhskii, A. I., A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. *Journal of Proteomics* **2010,** *73* (11), 2092-2123.

37. Clynen, E.; Baggerman, G.; Veelaert, D.; Cerstiaens, A.; Van der Horst, D.; Harthoorn, L.; Derua, R.; Waelkens, E.; De Loof, A.; Schoofs, L., Peptidomics of the pars intercerebralis-corpus cardiacum complex of the migratory locust, Locusta migratoria. *Eur J Biochem* **2001,** *268* (7), 1929-39.

38. Verhaert, P.; Uttenweiler-Joseph, S.; de Vries, M.; Loboda, A.; Ens, W.; Standing, K. G., Matrix-assisted laser desorption/ionization quadrupole time-of-flight mass spectrometry: an elegant tool for peptidomics. *Proteomics* **2001,** *1* (1), 118-31.

39. Buchberger, A.; Yu, Q.; Li, L., Advances in Mass Spectrometric Tools for Probing Neuropeptides. *Annu Rev Anal Chem (Palo Alto Calif)* **2015**.

40. Dallas, D. C.; Guerrero, A.; Parker, E. A.; Robinson, R. C.; Gan, J. N.; German, J. B.; Barile, D.; Lebrilla, C. B., Current peptidomics: Applications, purification, identification, quantification, and functional analysis. *Proteomics* **2015,** *15* (5-6), 1026-1038.

41. Matsubayashi, Y.; Sakagami, Y., Peptide hormones in plants. In *Annual Review of Plant Biology*, Annual Reviews: Palo Alto, 2006; Vol. 57, pp 649-674.

42. Yamaguchi, Y.; Huffaker, A., Endogenous peptide elicitors in higher plants. *Current Opinion in Plant Biology* **2011,** *14* (4), 351-357.

43. Farrokhi, N.; Whitelegge, J. P.; Brusslan, J. A., Plant peptides and peptidomics. *Plant Biotechnology Journal* **2008,** *6* (2), 105-134.

44. Ohyama, K.; Ogawa, M.; Matsubayashi, Y., Identification of a biologically active, small, secreted peptide in Arabidopsis by in silico gene screening, followed by LC-MS-based structure analysis. *Plant Journal* **2008,** *55* (1), 152-160.

45. Chen, Y. L.; Lee, C. Y.; Cheng, K. T.; Chang, W. H.; Huang, R. N.; Nam, H. G.; Chen, Y. R., Quantitative Peptidomics Study Reveals That a Wound-Induced Peptide from PR-1 Regulates Immune Signaling in Tomato. *Plant Cell* **2014,** *26* (10), 4135-4148.

46. Haruta, M.; Sabat, G.; Stecker, K.; Minkoff, B. B.; Sussman, M. R., A peptide hormone and its receptor protein kinase regulate plant cell expansion. *Science* **2014,** *343* (6169), 408-11.

47. Okazaki, Y.; Saito, K., Recent advances of metabolomics in plant biotechnology. *Plant Biotechnology Reports* **2012,** *6* (1), 1-15.

48. Lei, Z. T.; Huhman, D. V.; Sumner, L. W., Mass Spectrometry Strategies in Metabolomics. *Journal of Biological Chemistry* **2011,** *286* (29), 25435-25442.

49. Oms-Oliu, G.; Hertog, M.; Van de Poel, B.; Ampofo-Asiama, J.; Geeraerd, A. H.; Nicolai, B. M., Metabolic characterization of tomato fruit during preharvest development, ripening, and postharvest shelf-life. *Postharvest Biology and Technology* **2011,** *62* (1), 7-16.

50. Zhang, J. Y.; de Carvalho, M. H. C.; Torres-Jerez, I.; Kang, Y.; Allen, S. N.; Huhman, D. V.; Tang, Y. H.; Murray, J.; Sumner, L. W.; Udvardi, M. K., Global reprogramming of transcription and metabolism in Medicago truncatula during progressive drought and after rewatering. *Plant Cell and Environment* **2014,** *37* (11), 2553-2576.

51. Liu, J. J.; Wang, H.; Cooks, R. G.; Ouyang, Z., Leaf Spray: Direct Chemical Analysis of Plant Material and Living Plants by Mass Spectrometry. *Analytical Chemistry* **2011,** *83* (20), 7608-7613.

52. Klampfl, C. W.; Himmelsbach, M., Direct ionization methods in mass spectrometry: An overview. *Analytica Chimica Acta* **2015,** *890*, 44-59.

53. Vuckovic, D., Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry. *Analytical and Bioanalytical Chemistry* **2012,** *403* (6), 1523-1548.

54. Kueger, S.; Steinhauser, D.; Willmitzer, L.; Giavalisco, P., High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. *Plant Journal* **2012,** *70* (1), 39-50.

55. Katajamaa, M.; Oresic, M., Data processing for mass spectrometry-based metabolomics. *Journal of Chromatography A* **2007,** *1158* (1-2), 318-328.

56. Creek, D. J.; Dunn, W. B.; Fiehn, O.; Griffin, J. L.; Hall, R. D.; Lei, Z. T.; Mistrik, R.; Neumann, S.; Schymanski, E. L.; Sumner, L. W.; Trengove, R.; Wolfender, J. L., Metabolite identification: are you sure? And how do your peers gauge your confidence? *Metabolomics* **2014,** *10* (3), 350-353.

57. Balluff, B.; Schone, C.; Hofler, H.; Walch, A., MALDI imaging mass spectrometry for direct tissue analysis: technological advancements and recent applications. *Histochemistry and cell biology* **2011,** *136* (3), 227-44.

58. Cohen, M. Z., A historical overview of the phenomenologic movement. *Image J Nurs Sch* **1987,** *19* (1), 31-4.

59. Makarov, A., Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. *Anal Chem* **2000,** *72* (6), 1156-1162.

60. Perdian, D. C.; Lee, Y. J., Imaging MS methodology for more chemical information in less data acquisition time utilizing a hybrid linear ion trap-orbitrap mass spectrometer. *Anal Chem* **2010,** *82* (22), 9393-400.

61. Crossman, L.; McHugh, N. A.; Hsieh, Y. S.; Korfmacher, W. A.; Chen, J. W., Investigation of the profiling depth in matrix-assisted laser desorption/ionization imaging mass spectrometry. *Rapid Commun Mass Sp* **2006,** *20* (2), 284-290.

62. Schwartz, S. A.; Reyzer, M. L.; Caprioli, R. M., Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation. *J Mass Spectrom* **2003,** *38* (7), 699-708.

63. Kaletas, B. K.; van der Wiel, I. M.; Stauber, J.; Dekker, L. J.; Guzel, C.; Kros, J. M.; Luider, T. M.; Heeren, R. M. A., Sample preparation issues for tissue imaging by imaging MS. *Proteomics* **2009,** *9* (10), 2622-2633.

64. Gemperline, E.; Rawson, S.; Li, L., Optimization and comparison of multiple MALDI matrix application methods for small molecule mass spectrometric imaging. *Anal Chem* **2014,** *86* (20), 10030-5.

65. Chen, R.; Cape, S. S.; Sturm, R. M.; Li, L., Mass spectrometric imaging of neuropeptides in decapod crustacean neuronal tissues. *Methods Mol Biol* **2010,** *656*, 451-63.

66. Burrell, M. M.; Earnshaw, C. J.; Clench, M. R., Imaging Matrix Assisted Laser Desorption Ionization Mass Spectrometry: a technique to map plant metabolites within tissues at high spatial resolution. *J Exp Bot* **2007,** *58* (4), 757-763.

67. Shrivas, K.; Hayasaka, T.; Sugiura, Y.; Setou, M., Method for simultaneous imaging of endogenous low molecular weight metabolites in mouse brain using TiO2 nanoparticles in nanoparticle-assisted laser desorption/ionization-imaging mass spectrometry. *Analytical Chemistry* **2011,** *83* (19), 7283-9.

68. Altelaar, A. F.; Klinkert, I.; Jalink, K.; de Lange, R. P.; Adan, R. A.; Heeren, R. M.; Piersma, S. R., Gold-enhanced biomolecular surface imaging of cells and tissue by SIMS and MALDI mass spectrometry. *Anal Chem* **2006,** *78* (3), 734-42.

69. Hayasaka, T.; Goto-Inoue, N.; Zaima, N.; Shrivas, K.; Kashiwagi, Y.; Yamamoto, M.; Nakamoto, M.; Setou, M., Imaging mass spectrometry with silver nanoparticles reveals the distribution of fatty acids in mouse retinal sections. *J Am Soc Mass Spectrom* **2010,** *21* (8), 1446-54.

70. Thomas, A.; Charbonneau, J. L.; Fournaise, E.; Chaurand, P., Sublimation of new matrix candidates for high spatial resolution imaging mass spectrometry of lipids: enhanced information in both positive and negative polarities after 1,5-diaminonapthalene deposition. *Analytical Chemistry* **2012,** *84* (4), 2048-54.

71. Chen, S.; Chen, L.; Wang, J.; Hou, J.; He, Q.; Liu, J.; Xiong, S.; Yang, G.; Nie, Z., 2,3,4,5-Tetrakis(3',4'-dihydroxylphenyl)thiophene: a new matrix for the selective analysis of low molecular weight amines and direct determination of creatinine in urine by MALDI-TOF MS. *Analytical Chemistry* **2012,** *84* (23), 10291-7.

72. Shroff, R.; Rulisek, L.; Doubsky, J.; Svatos, A., Acid-base-driven matrix-assisted mass spectrometry for targeted metabolomics. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (25), 10092-6.

73. Shroff, R.; Svatos, A., Proton sponge: a novel and versatile MALDI matrix for the analysis of metabolites using mass spectrometry. *Analytical chemistry* **2009,** *81* (19), 7954-9.

74. Ye, H.; Gemperline, E.; Li, L., A vision for better health: Mass spectrometry imaging for clinical diagnostics. *Clinica chimica acta; international journal of clinical chemistry* **2012**.

75. Goodwin, R. J. A., Sample preparation for mass spectrometry imaging: Small mistakes can lead to big consequences. *J Proteomics* **2012,** *75* (16), 4893-4911.

76. Baluya, D. L.; Garrett, T. J.; Yost, R. A., Automated MALDI matrix deposition method with inkjet printing for imaging mass spectrometry. *Analytical Chemistry* **2007,** *79* (17), 6862-7.

77. Hankin, J. A.; Barkley, R. M.; Murphy, R. C., Sublimation as a method of matrix application for mass spectrometric imaging. *J Am Soc Mass Spectrom* **2007,** *18* (9), 1646-52.

78. Puolitaival, S. M.; Burnum, K. E.; Cornett, D. S.; Caprioli, R. M., Solvent-free matrix dry-coating for MALDI Imaging of phospholipids. *J Am Soc Mass Spectrom* **2008,** *19* (6), 882-886.

79. Trimpin, S.; Herath, T. N.; Inutan, E. D.; Wager-Miller, J.; Kowalski, P.; Claude, E.; Walker, J. M.; Mackie, K., Automated Solvent-Free Matrix Deposition for Tissue Imaging by Mass Spectrometry. *Analytical Chemistry* **2010,** *82* (1), 359-367.

80. Chang, W. C.; Huang, L. C. L.; Wang, Y. S.; Peng, W. P.; Chang, H. C.; Hsu, N. Y.; Yang, W. B.; Chen, C. H., Matrix-assisted laser desorption/ionization (MALDI) mechanism revisited. *Anal Chim Acta* **2007,** *582* (1), 1-9.

81. Karas, M.; Kruger, R., Ion formation in MALDI: The cluster ionization mechanism. *Chem Rev* **2003,** *103* (2), 427-439.

82. Knochenmuss, R., A quantitative model of ultraviolet matrix-assisted laser desorption/ionization including analyte ion generation. *Analytical Chemistry* **2003,** *75* (10), 2199-2207.

83. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T., Protein and polymer analyses up to m/z 100000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun Mass Sp* **1988,** *2* (8), 151-153.

84. Ye, H.; Gemperline, E.; Li, L., A vision for better health: mass spectrometry imaging for clinical diagnostics. *Clin Chim Acta* **2013,** *420*, 11-22.

85. Fitzgerald, J. J. D.; Kunnath, P.; Walker, A. V., Matrix-enhanced secondary ion mass spectrometry (ME-SIMS) using room temperature ionic liquid matrices. *Analytical Chemistry* **2010,** *82* (11), 4413-4419.

86. Lemaire, R.; Tabet, J. C.; Ducoroy, P.; Hendra, J. B.; Salzet, M.; Fournier, I., Solid ionic matrixes for direct tissue analysis and MALDI Imaging. *Analytical Chemistry* **2006,** *78* (3), 809-819.

87. Shroff, R.; Svatos, A., Proton sponge: a novel and versatile MALDI matrix for the analysis of metabolites using mass spectrometry. *Analytical Chemistry* **2009,** *81* (19), 7954-9.

88. Mainini, V.; Bovo, G.; Chinello, C.; Gianazza, E.; Grasso, M.; Cattoretti, G.; Magni, F., Detection of high molecular weight proteins by MALDI imaging mass spectrometry. *Mol Biosyst* **2013,** *9* (6), 1101-1107.

89. Lee, Y. J.; Perdian, D. C.; Song, Z. H.; Yeung, E. S.; Nikolau, B. J., Use of mass spectrometry for imaging metabolites in plants. *Plant Journal* **2012,** *70* (1), 81-95.

90. Bjarnholt, N.; Li, B.; D'Alvise, J.; Janfelt, C., Mass spectrometry imaging of plant metabolites - principles and possibilities. *Natural product reports* **2014**.

91. Chatterji, B.; Pich, A., MALDI imaging mass spectrometry and analysis of endogenous peptides. *Expert review of proteomics* **2013,** *10* (4), 381-8.

92. Grassl, J.; Taylor, N. L.; Millar, A. H., Matrix-assisted laser desorption/ionisation mass spectrometry imaging and its development for plant protein imaging. *Plant methods* **2011,** *7* (1), 21.

93. Kaspar, S.; Peukert, M.; Svatos, A.; Matros, A.; Mock, H. P., MALDI-imaging mass spectrometry - An emerging technique in plant biology. *Proteomics* **2011,** *11* (9), 1840-50.

94. Matros, A.; Mock, H. P., Mass spectrometry based imaging techniques for spatially resolved analysis of molecules. *Front Plant Sci* **2013,** *4*, 89.

95. Gemperline, E.; Jayaraman, D.; Maeda, J.; Ane, J. M.; Li, L., Multifaceted investigation of metabolites during nitrogen fixation in Medicago via high resolution MALDI-MS imaging and ESI-MS. *J Am Soc Mass Spectrom* **2015,** *26* (1), 149-58.

96. Ye, H.; Gemperline, E.; Venkateshwaran, M.; Chen, R.; Delaux, P. M.; Howes-Podoll, M.; Ane, J. M.; Li, L., MALDI mass spectrometry-assisted molecular imaging of metabolites during nitrogen fixation in the Medicago truncatula-Sinorhizobium meliloti symbiosis. *The Plant journal : for cell and molecular biology* **2013,** *75* (1), 130-45.

97. Korte, A. R.; Yandeau-Nelson, M. D.; Nikolau, B. J.; Lee, Y. J., Subcellular-level resolution MALDI-MS imaging of maize leaf metabolites by MALDI-linear ion trap-Orbitrap mass spectrometer. *Anal Bioanal Chem* **2015,** *407* (8), 2301-9.

98. Poth, A. G.; Mylne, J. S.; Grassl, J.; Lyons, R. E.; Millar, A. H.; Colgrave, M. L.; Craik, D. J., Cyclotides Associate with Leaf Vasculature and Are the Products of a Novel Precursor in Petunia (Solanaceae). *J Biol Chem* **2012,** *287* (32), 27033-27046.

99. Cavatorta, V.; Sforza, S.; Mastrobuoni, G.; Pieraccini, G.; Francese, S.; Moneti, G.; Dossena, A.; Pastorello, E. A.; Marchelli, R., Unambiguous characterization and tissue localization of Pru P 3 peach allergen by electrospray mass spectrometry and MALDI imaging. *J Mass Spectrom* **2009,** *44* (6), 891-7.

100. Wu, K. J.; Odom, R. W., Matrix-enhanced secondary ion mass spectrometry: a method for molecular analysis of solid surfaces. *Anal Chem* **1996,** *68* (5), 873-82.

101. Kim, Y. P.; Oh, E.; Hong, M. Y.; Lee, D.; Han, M. K.; Shon, H. K.; Moon, D. W.; Kim, H. S.; Lee, T. G., Gold nanoparticle-enhanced secondary ion mass spectrometry imaging of peptides on self-assembled monolayers. *Anal Chem* **2006,** *78* (6), 1913-20.

102. Fletcher, J. S.; Lockyer, N. P.; Vickerman, J. C., Developments in molecular SIMS depth profiling and 3D imaging of biological systems using polyatomic primary ions. *Mass spectrometry reviews* **2011,** *30* (1), 142-74.

103. Liebl, H., Sims Instrumentation and Imaging Techniques. *Scanning* **1980,** *3* (2), 79-89.

104. Sumner, L. W.; Lei, Z.; Nikolau, B. J.; Saito, K., Modern plant metabolomics: advanced natural product gene discoveries, improved technologies, and future prospects. *Natural product reports* **2015,** *32* (2), 212-29.

105. Seyer, A.; Einhorn, J.; Brunelle, A.; Laprevote, O., Localization of flavonoids in seeds by cluster time-of-flight secondary ion mass spectrometry imaging. *Anal Chem* **2010,** *82* (6), 2326-33.

106. Takats, Z.; Wiseman, J. M.; Gologan, B.; Cooks, R. G., Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science* **2004,** *306* (5695), 471-473.

107. Ifa, D. R.; Wu, C. P.; Ouyang, Z.; Cooks, R. G., Desorption electrospray ionization and other ambient ionization methods: current progress and preview. *Analyst* **2010,** *135* (4), 669-681.

108. Tata, A.; Perez, C. J.; Hamid, T. S.; Bayfield, M. A.; Ifa, D. R., Analysis of metabolic changes in plant pathosystems by imprint imaging DESI-MS. *J Am Soc Mass Spectrom* **2015,** *26* (4), 641-8.

109. Hemalatha, R. G.; Pradeep, T., Understanding the molecular signatures in leaves and flowers by desorption electrospray ionization mass spectrometry (DESI MS) imaging. *Journal of agricultural and food chemistry* **2013,** *61* (31), 7477-87.

110. Nemes, P.; Vertes, A., Laser ablation electrospray ionization for atmospheric pressure, in vivo, and imaging mass spectrometry. *Analytical Chemistry* **2007,** *79* (21), 8098-8106.

111. Chen, Z. Y.; Bogaerts, A.; Vertes, A., Phase explosion in atmospheric pressure infrared laser ablation from water-rich targets. *Appl Phys Lett* **2006,** *89* (4).

112. Nemes, P.; Barton, A. A.; Li, Y.; Vertes, A., Ambient molecular imaging and depth profiling of live tissue by infrared laser ablation electrospray ionization mass spectrometry. *Analytical Chemistry* **2008,** *80* (12), 4575-4582.

113. Nemes, P.; Barton, A. A.; Vertes, A., Three-Dimensional Imaging of Metabolites in Tissues under Ambient Conditions by Laser Ablation Electrospray Ionization Mass Spectrometry. *Analytical Chemistry* **2009,** *81* (16), 6668-6675.

114. Li, H.; Smith, B. K.; Mark, L.; Nemes, P.; Nazarian, J.; Vertes, A., Ambient molecular imaging by laser ablation electrospray ionization mass spectrometry with ion mobility separation. *Int J Mass Spectrom* **2015,** *377*, 681-689.

115. Kiss, A.; Smith, D. F.; Reschke, B. R.; Powell, M. J.; Heeren, R. M., Top-down mass spectrometry imaging of intact proteins by laser ablation ESI FT-ICR MS. *Proteomics* **2014,** *14* (10), 1283-9.

116. Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G., METLIN: a metabolite mass spectral database. *Therapeutic drug monitoring* **2005,** *27* (6), 747-51.

117. Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M. A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, D. D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; Macinnis, G. D.; Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L., HMDB: the Human Metabolome Database. *Nucleic Acids Res* **2007,** *35* (Database issue), D521-6.

118. Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorndahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A., HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res* **2013,** *41* (Database issue), D801-7.

119. Kanehisa, M., A database for post-genome analysis. *Trends in genetics : TIG* **1997,** *13* (9), 375-6.

120. Bolton, E. E.; Wang, Y. L.; Thiessen, P. A.; Bryant, S. H., PubChem: Integrated Platform of Small Molecules and Biological Activities. *Ann Rep Comp Chem* **2010,** *4*, 217-241.

121. Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Tanaka, K.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T., MassBank: a public repository for sharing mass spectral data for life sciences. *J Mass Spectrom* **2010,** *45* (7), 703-714.

122. Cui, Q.; Lewis, I. A.; Hegeman, A. D.; Anderson, M. E.; Li, J.; Schulte, C. F.; Westler, W. M.; Eghbalnia, H. R.; Sussman, M. R.; Markley, J. L., Metabolite identification via the Madison Metabolomics Consortium Database. *Nat Biotechnol* **2008,** *26* (2), 162-164.

123. Fahy, E.; Subramaniam, S.; Brown, H. A.; Glass, C. K.; Merrill, A. H.; Murphy, R. C.; Raetz, C. R. H.; Russell, D. W.; Seyama, Y.; Shaw, W.; Shimizu, T.; Spener, F.; van Meer, G.; VanNieuwenhze, M. S.; White, S. H.; Witztum, J. L.; Dennis, E. A., A comprehensive classification system for lipids. *J Lipid Res* **2005,** *46* (5), 839-861.

124. Fahy, E.; Subramaniam, S.; Murphy, R. C.; Nishijima, M.; Raetz, C. R. H.; Shimizu, T.; Spener, F.; van Meer, G.; Wakelam, M. J. O.; Dennis, E. A., Update of the LIPID MAPS comprehensive classification system for lipids. *J Lipid Res* **2009,** *50*, S9-S14.

125. Wolf, S.; Schmidt, S.; Muller-Hannemann, M.; Neumann, S., In silico fragmentation for computer assisted identification of metabolite mass spectra. *Bmc Bioinformatics* **2010,** *11*.

126. Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G., XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* **2006,** *78* (3), 779-87.

127. Gowda, H.; Ivanisevic, J.; Johnson, C. H.; Kurczy, M. E.; Benton, H. P.; Rinehart, D.; Nguyen, T.; Ray, J.; Kuehl, J.; Arevalo, B.; Westenskow, P. D.; Wang, J.; Arkin, A. P.; Deutschbauer, A. M.; Patti, G. J.; Siuzdak, G., Interactive XCMS Online: simplifying advanced metabolomic data processing and subsequent statistical analyses. *Anal Chem* **2014,** *86* (14), 6931-9.

128. Xia, J. G.; Psychogios, N.; Young, N.; Wishart, D. S., MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res* **2009,** *37*, W652-W660.

129. Xia, J. G.; Mandal, R.; Sinelnikov, I. V.; Broadhurst, D.; Wishart, D. S., MetaboAnalyst 2.0-a comprehensive server for metabolomic data analysis. *Nucleic Acids Res* **2012,** *40* (W1), W127-W133.

130. Xia, J. G.; Sinelnikov, I. V.; Han, B.; Wishart, D. S., MetaboAnalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Res* **2015,** *43* (W1), W251-W257.

131. Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G., PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid communications in mass spectrometry : RCM* **2003,** *17* (20), 2337-42.

132. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic Local Alignment Search Tool. *J Mol Biol* **1990,** *215* (3), 403-410.

133. Bailey, T. L.; Boden, M.; Buske, F. A.; Frith, M.; Grant, C. E.; Clementi, L.; Ren, J. Y.; Li, W. W.; Noble, W. S., MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **2009,** *37*, W202-W208.

134. Baggerman, G.; Liu, F.; Wets, G.; Schoofs, L., Bioinformatic analysis of peptide precursor proteins. *Annals of the New York Academy of Sciences* **2005,** *1040*, 59-65.

135. Wenger, C. D.; Phanstiel, D. H.; Lee, M. V.; Bailey, D. J.; Coon, J. J., COMPASS: A suite of pre- and post-search proteomics software tools for OMSSA. *Proteomics* **2011,** *11* (6), 1064-1074.

136. MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J., Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **2010,** *26* (7), 966-968.

137. Wenger, C. D.; Coon, J. J., A Proteomics Search Algorithm Specifically Designed for High-Resolution Tandem Mass Spectra. *Journal of Proteome Research* **2013,** *12* (3), 1377-1386.

138. Robichaud, G.; Garrard, K. P.; Barry, J. A.; Muddiman, D. C., MSiReader: an open-source interface to view and analyze high resolving power MS imaging files on Matlab platform. *Journal of the American Society for Mass Spectrometry* **2013,** *24* (5), 718-21.

139. Rubel, O.; Greiner, A.; Cholia, S.; Louie, K.; Bethel, E. W.; Northen, T. R.; Bowen, B. P., OpenMSI: a high-performance web-based platform for mass spectrometry imaging. *Anal Chem* **2013,** *85* (21), 10354-61.

140. Race, A. M.; Steven, R. T.; Palmer, A. D.; Styles, I. B.; Bunch, J., Memory efficient principal component analysis for the dimensionality reduction of large mass spectrometry imaging data sets. *Anal Chem* **2013,** *85* (6), 3071-8.

141. Lei, Z.; Jing, L.; Qiu, F.; Zhang, H.; Huhman, D.; Zhou, Z.; Sumner, L. W., Construction of an Ultrahigh Pressure Liquid Chromatography-Tandem Mass Spectral Library of Plant Natural Products and Comparative Spectral Analyses. *Anal Chem* **2015**.

142. Grimsrud, P. A.; den Os, D.; Wenger, C. D.; Swaney, D. L.; Schwartz, D.; Sussman, M. R.; Ane, J. M.; Coon, J. J., Large-Scale Phosphoprotein Analysis in Medicago truncatula Roots Provides Insight into in Vivo Kinase Activity in Legumes. *Plant Physiol* **2010,** *152* (1), 19-28.

143. Sun, Q.; Zybailov, B.; Majeran, W.; Friso, G.; Olinares, P. D. B.; van Wijk, K. J., PPDB, the Plant Proteomics Database at Cornell. *Nucleic Acids Res* **2009,** *37*, D969-D974.

144. Komatsu, S., Rice proteome database: A step toward functional analysis of the rice genome. *Plant Mol Biol* **2005,** *59* (1), 179-190.

145. Wienkoop, S.; Staudinger, C.; Hoehenwarter, W.; Weckwerth, W.; Egelhofer, V., ProMEX - a mass spectral reference database for plant proteomics. *Front Plant Sci* **2012,** *3*.

146. Urbanczyk-Wochniak, E.; Sumner, L. W., MedicCyc: a biochemical pathway database for Medicago truncatula. *Bioinformatics* **2007,** *23* (11), 1418-1423.

147. Mueller, L. A.; Zhang, P. F.; Rhee, S. Y., AraCyc: A biochemical pathway database for Arabidopsis. *Plant Physiol* **2003,** *132* (2), 453-460.

148. Zhang, P. F.; Foerster, H.; Tissier, C. P.; Mueller, L.; Paley, S.; Karp, P. D.; Rhee, S. Y., MetaCyc and AraCyc. Metabolic pathway databases for plant research. *Plant Physiol* **2005,** *138* (1), 27-37.

149. Kanehisa, M., Toward pathway engineering: a new database of genetic and molecular pathways. *Science & Technology Japan* **1996,** *59*, 34-38.

150. Spengler, B.; Hubert, M., Scanning microprobe matrix-assisted laser desorption ionization (SMALDI) mass spectrometry: instrumentation for sub-micrometer resolved LDI and MALDI surface analysis. *J Am Soc Mass Spectrom* **2002,** *13* (6), 735-48.

151. Rompp, A.; Spengler, B., Mass spectrometry imaging with high resolution in mass and space. *Histochemistry and cell biology* **2013,** *139* (6), 759-783.

152. Koestler, M.; Kirsch, D.; Hester, A.; Leisner, A.; Guenther, S.; Spengler, B., A high-resolution scanning microprobe matrix-assisted laser desorption/ionization ion source for imaging analysis on an ion trap/Fourier transform ion cyclotron resonance mass spectrometer. *Rapid Commun Mass Sp* **2008,** *22* (20), 3275-3285.

153. Boggio, K. J.; Obasuyi, E.; Sugino, K.; Nelson, S. B.; Agar, N. Y. R.; Agar, J. N., Recent advances in single-cell MALDI mass spectrometry imaging and potential clinical impact. *Expert review of proteomics* **2011,** *8* (5), 591-604.

154. Li, L.; Hur, M.; Lee, J. Y.; Zhou, W.; Song, Z.; Ransom, N.; Demirkale, C. Y.; Nettleton, D.; Westgate, M.; Arendsee, Z.; Iyer, V.; Shanks, J.; Nikolau, B.; Wurtele, E. S., A systems biology approach toward understanding seed composition in soybean. *BMC genomics* **2015,** *16 Suppl 3*, S9.

155. Kim, T.; Dreher, K.; Nilo-Poyanco, R.; Lee, I.; Fiehn, O.; Lange, B. M.; Nikolau, B. J.; Sumner, L.; Welti, R.; Wurtele, E. S.; Rhee, S. Y., Patterns of metabolite changes identified from large-scale gene perturbations in Arabidopsis using a genome-scale metabolic network. *Plant Physiol* **2015,** *167* (4), 1685-98.

156. Quanbeck, S. M.; Brachova, L.; Campbell, A. A.; Guan, X.; Perera, A.; He, K.; Rhee, S. Y.; Bais, P.; Dickerson, J. A.; Dixon, P.; Wohlgemuth, G.; Fiehn, O.; Barkan, L.; Lange, I.; Lange, B. M.; Lee, I.; Cortes, D.; Salazar, C.; Shuman, J.; Shulaev, V.; Huhman, D. V.; Sumner, L. W.; Roth, M. R.; Welti, R.; Ilarslan, H.; Wurtele, E. S.; Nikolau, B. J., Metabolomics as a hypothesis-generating functional genomics tool for the annotation of Arabidopsis thaliana genes of "unknown function". *Front Plant Sci* **2012,** *3*.

## Tables

**Table 1.** Comparison of MSI approaches detailing ionization methods, mass range, and spatial resolution, and optimal analytes.

| | **MALDI** | **SIMS** | **DESI** | **LAESI** |
|---|---|---|---|---|
| **Ionization Source** | UV/IR laser (Soft ionization) | Ion gun (Hard ionization) | Solvent spray (Soft ionization) | Solvent spray (Soft ionization) |
| **Mass Range (Da)** | 0-70,000 | 0-2,000 (static mode)<br>0-300 (dynamic mode) | 0-5,000 | 0-66,000 |
| **Lateral Resolution (μm)** | 20-500 | 0.5-50 (static)<br>0.04-0.5 (dynamic) | 100-1000 | 200-400 |
| **Molecular Classes** | metabolites, lipids, peptides, proteins | elements, metabolites, lipids, small peptides | metabolites, lipids, small peptides | metabolites, lipids, peptides, proteins |

**Figures**

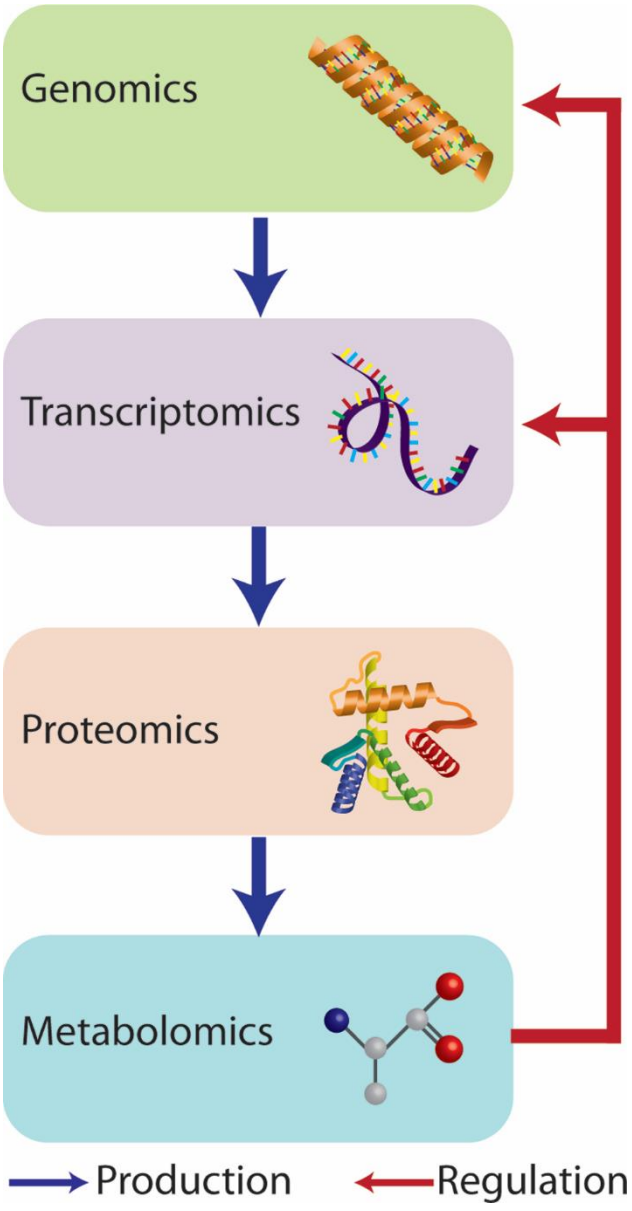



**Figure 1.** Interaction between genomics, transcriptomics, proteomics, and metabolomics.

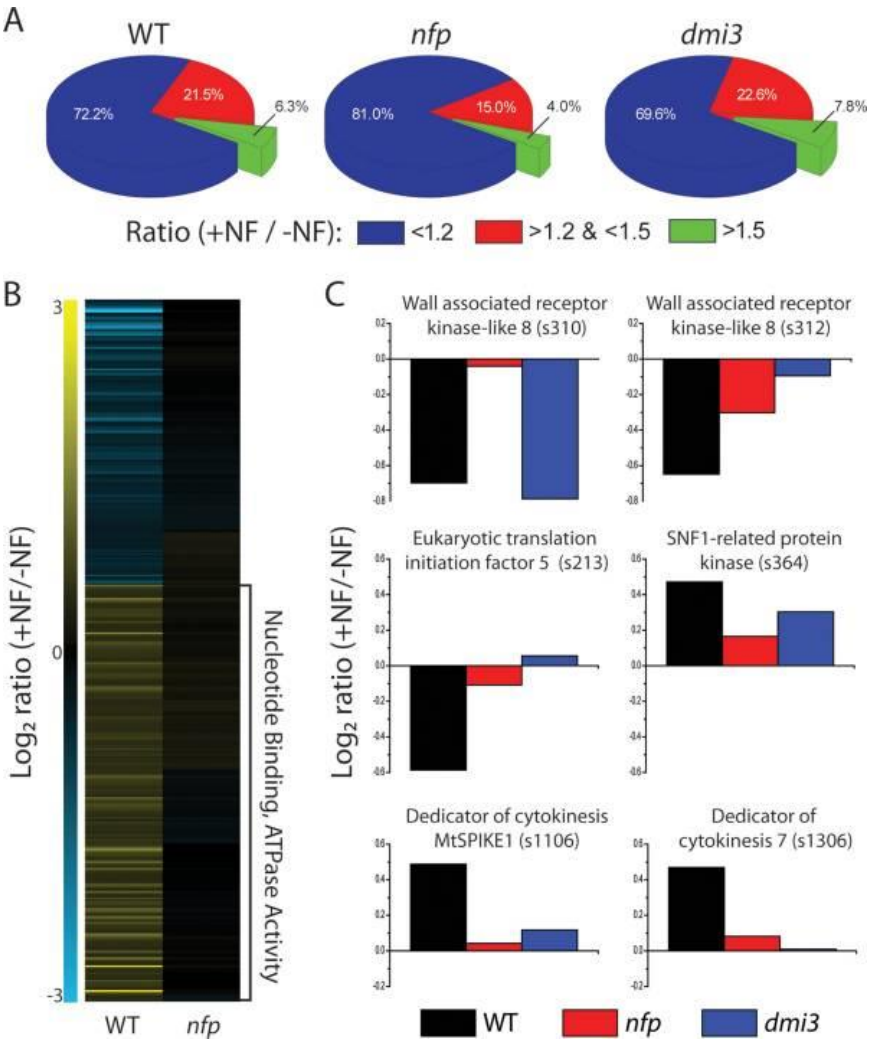



**Figure 2.** Global view of changes in the proteome and phosphoproteome from all the experiments combined. **(A)** Pie charts displaying the distribution of phosphorylation changes for wild-type (WT) and *nfp* and *dmi3 mutants*. Wild-type plants readily respond to NF as 6.3% of phosphoisoforms are altered more than 1.5-fold. *Nfp* displayed a lower response, 4.0%, but the presence of these changes provides evidence of a separate NF receptor sensing NF. *Dmi3 showed* more NF response than wild-type plants, as the phosphorylation state of 7.8% of the phosphoisoforms was altered more than 1.5-fold. ***(*B)** Heatmap of phosphoisoforms altered in wild-type and not in *nfp*. Phophoisofroms exhibiting a fold change more than 1.35 in wild-type plants and less than 1.25 in *nfp* were grouped *via* hierarchical clustering. Functional analysis revealed phosphoisoforms, which demonstrate an up-regulation in phosphorylation upon NF treatment, are enriched for the terms "nucleotide binding" and "ATPase activity". **(C)** Representative proteins significantly altered in wild-type and not in *nfp*. Six phosphoisoforms that were significantly ($p <$ 0.05, Student's *t* test, assuming equal variances) altered more than 1.35-fold in wild-type plants and less than 1.25-fold in *nfp* demonstrate the use of comparing wild-type and mutant measurements. These examples contain phosphoisoforms, which display both *nfp* dependent regulation and *dmi3* dependent regulation. Reprinted with permission from Ref. [29]. Copyright 2012 The American Society for Biochemistry and Molecular Biology, Inc.

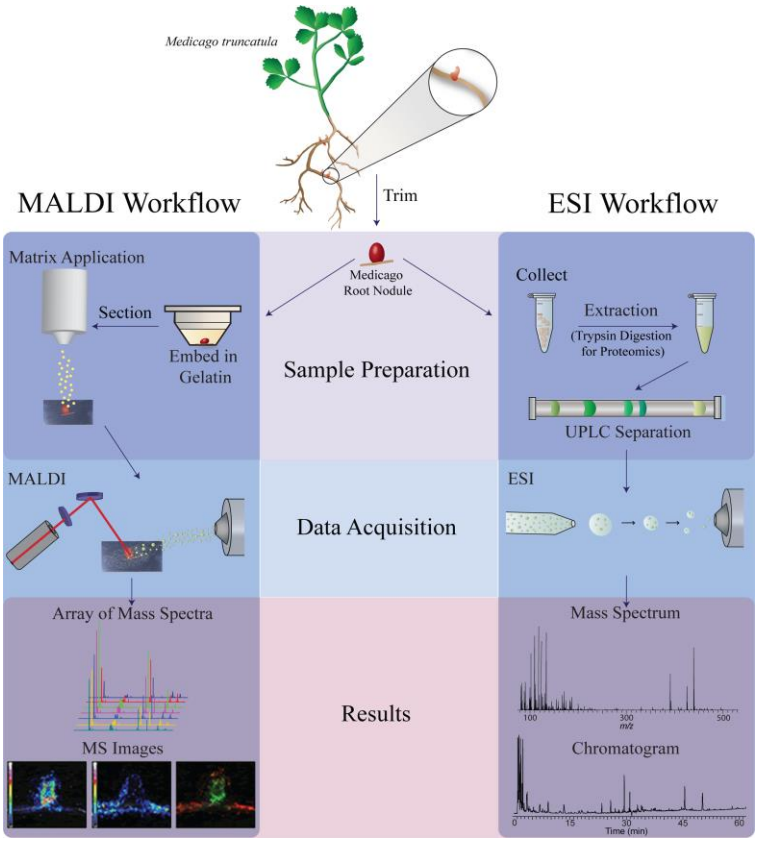



**Figure 3.** Comparison of MALDI-MSI (left) and LC-MS (right) Workflows. In both types of analysis, the tissue of interest, the root nodule in this example, is trimmed from live plants and flash frozen to preserve the analytes of interest. For LC-MS, many root nodules are pooled, followed by analyte extraction and additional steps of trypsin digestion and desalting for proteomics experiments. For MALDI-MSI, a single root nodule is embedded in a stabilizing media, such as gelatin, cryo-sectioned onto a glass slide, and matrix is applied. For LC-MS, the plant extract is injected onto the UPLC column for LC separation and analyzed by the mass spectrometer yielding a mass spectrum and a chromatogram. For MALDI-MSI, the glass slide is inserted into the mass spectrometer where a mass spectrum is acquired in a raster pattern across the entire surface of the tissue resulting in a large array of mass spectra that are coordinated with spatial information. Software can then compile the array of mass spectra into heat-map-like images based on the intensity of each *m/z* at each corresponding location on the tissue.

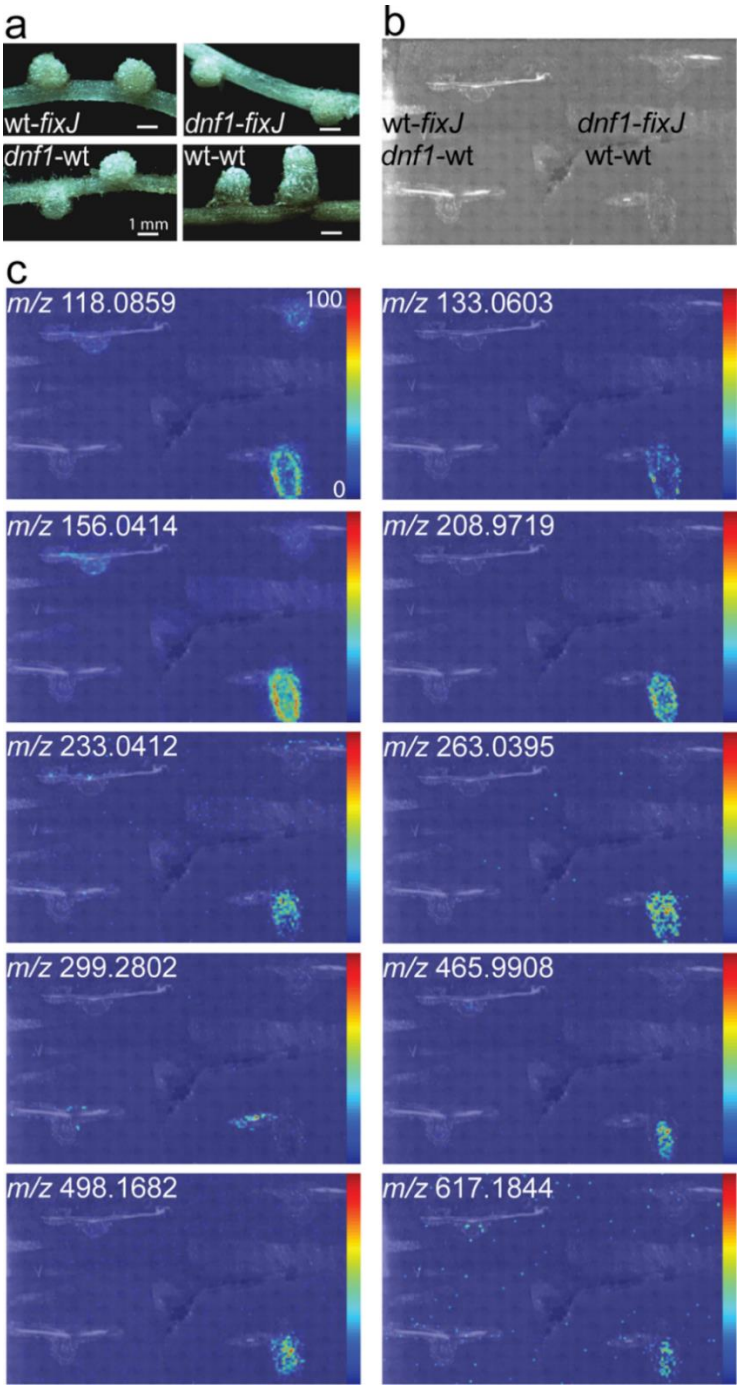



**Figure 4.** MALDI-MSI of metabolites in Medicago. **(A)** A photograph of all four of the sample types that were used in this study: wild-type Medicago/ *fixJ* rhizobia mutant (wt-*fixJ*), *dnf1* Medicago mutant/ *fixJ* rhizobia mutant (*dnf1-fixJ*), *dnf1* Medicago mutant/ wild-type rhizobia (*dnf1*-wt), and wild-type Medicago/ wild-type rhizobia (wt-wt). Where wt-wt samples are capable of nitrogen fixation and samples containing either the plant or bacterial mutant (wt-*fixJ*, *dnf1*-wt, or *dnf1-fixJ*) are incapable of nitrogen fixation. **(B)** Thaw-mounted Medicago sections on a glass slide that was covered with DHB matrix prior to MALDI-MSI. **(C)** Representative ion images of metabolites with distinct spatial distributions that were found only in the wt-wt samples that are capable of fixing nitrogen, indicating that these metabolites may play a role in the biological nitrogen fixation process. Scale bar = 1 mm**.** Reprinted with permission from Ref. [95]. Copyright 2015 Springer.

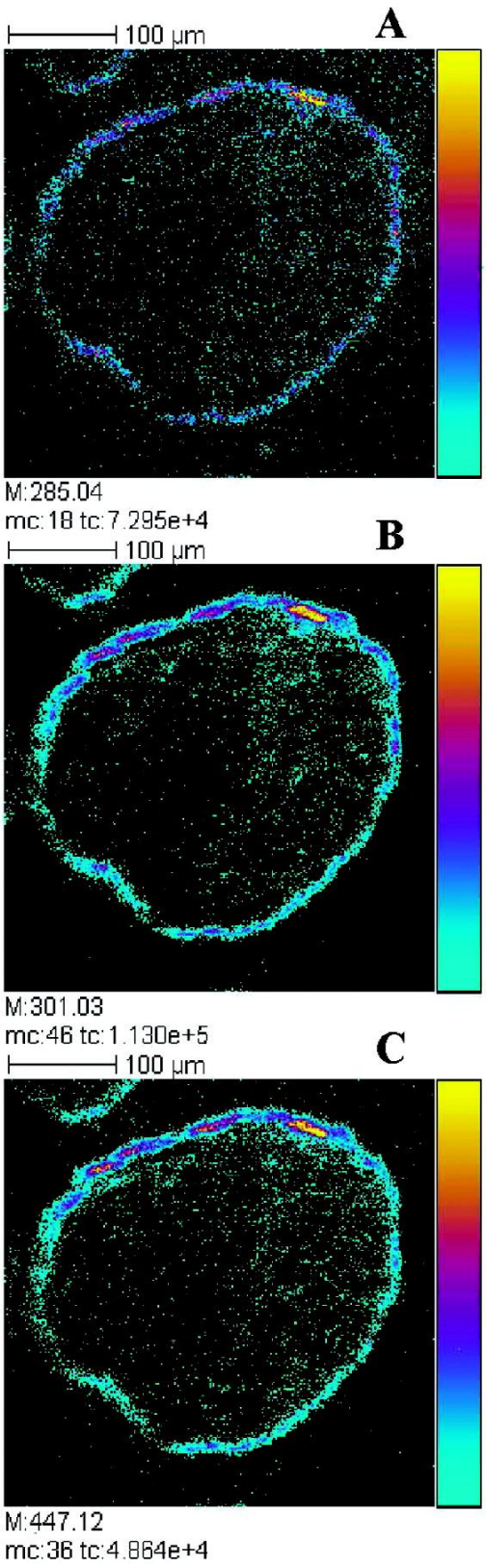



**Figure 5.** TOF-SIMS negative ion images of an *A. thaliana* seed section, embedded in a polyester resin: **(A)** *m/z* 285.04, **(B)** *m/z* 301.03, **(C)** *m/z* 447.1 (field of view 400 µm × 400 µm, 256 × 256 pixels, pixel size 1.56 µm, fluence 1.5 × 1012 ions · cm-2). Color scale bars, with amplitude in number of counts, are indicated to the right of each ion image. The amplitude of the color scale corresponds to the maximum number of counts (mc) and could be read as [0, mc]. TC is the total number of counts recorded for the specified *m/z* (sum of counts in all the pixels). The ion signals from the coat of another seed section appear in the top left of the images. Reprinted with permission from Ref. [105]. Copyright 2010 American Chemical Society.

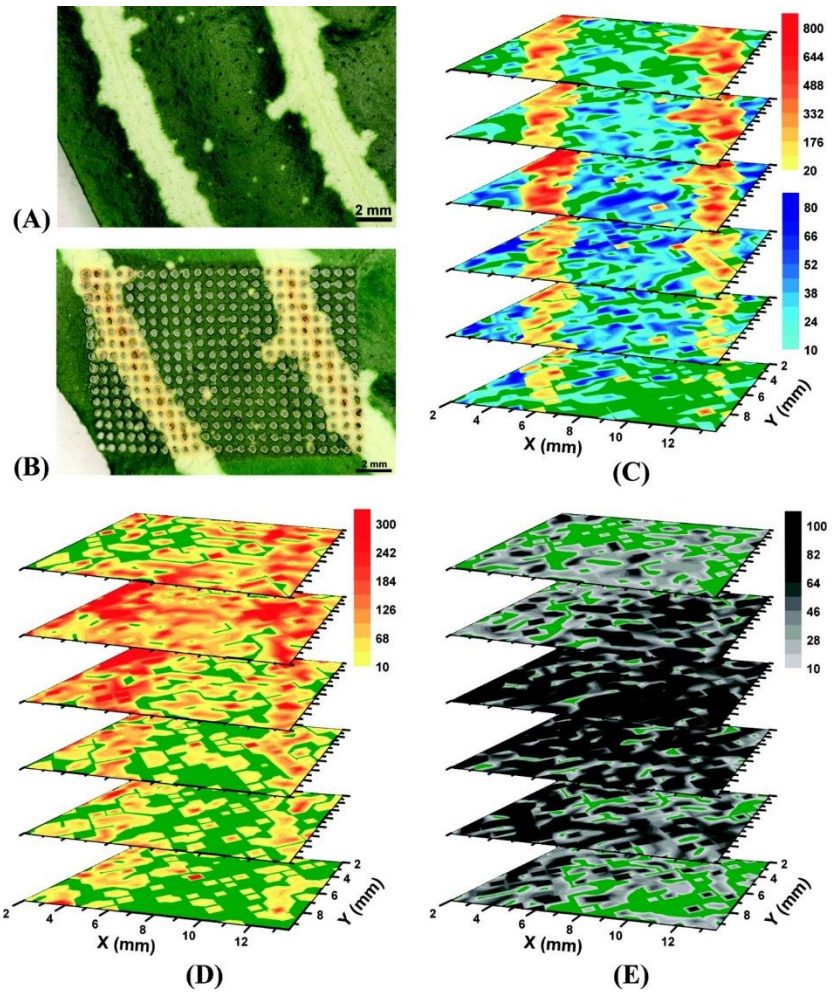



**Figure 6.** Metabolites in relation to tissue architecture captured by LAESI 3D imaging MS. Optical image of *A. squarrosa* leaves **(A)** before and **(B)** after analysis. **(C)** LAESI 3D imaging MS distribution of kaempferol/luteolin with *m/z* 287.0 (yellow/orange scale) followed the variegation pattern. Chlorophyll a with *m/z* 893.5 (blue scale) accumulated in the mesophyll layers. **(D)** Acacetin with *m/z* 285.0 showed higher abundance in the yellow sectors of the second and third layers with a homogeneous distribution in the others. **(E)** Kaempferol-(diacetyl coumarylrhamnoside) with *m/z* 663.2 accumulated in the mesophyll layers (third and fourth) with uniform lateral distributions. Reprinted with permission from Ref. [113]. Copyright 2009 American Chemical Society.

# Chapter 3

# Comparison of Vacuum MALDI and AP-MALDI Platforms for the Mass Spectrometry Imaging of Metabolites Involved in Salt Stress in *Medicago truncatula*

Adapted from: Keller C, Maeda J, Jayaraman D, Chakraborty S, Sussman MR, Harris J, Ané JM, Li L. Comparison of Vacuum MALDI and AP-MALDI Platforms for the Mass Spectrometry Imaging of Metabolites Involved in Salt Stress in *Medicago truncatula*. *Frontiers in Plant Science*. **2018**, 9, 1238

Keyword: AP-MALDI; mass spectrometric imaging; MSI; *Medicago truncatula*; salt stress; metabolomics

**Abstract**

Matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) is routinely used to determine the spatial distributions of various biomolecules in tissues. Recently, there has been an increased interest in creating higher resolution images using sources with more focused beams. One such source, an atmospheric pressure (AP) MALDI source from MassTech, has a laser capable of reaching spatial resolutions of 10 μm. Here, the AP-MALDI source coupled with a Q Exactive HF Orbitrap platform is compared to the commercial MALDI LTQ Orbitrap XL system using *Medicago truncatula* root nodules. AP-MALDI parameters, such as the S lens value, capillary temperature, and spray voltage, were optimized on the Q Exactive-HF platform for optimal detection of plant metabolites. The performance of the two systems was evaluated for sensitivity, spatial resolution, and overall ability to detect plant metabolites. The commercial MALDI LTQ Orbitrap XL was superior regarding the number of compounds detected, as at least two times more *m/z* were detected compared to the AP-MALDI system. However, although the AP-MALDI source requires a spatial resolution higher than 10 μm to get the best signal, the spatial resolution at 30 μm is still superior compared to the 75 μm spatial resolution achieved on the MALDI platform. The AP-MALDI system was also used to investigate the metabolites present in *Medicago truncatula* roots and root nodules under high salt and low salt conditions. A discriminative analysis with SCiLS software revealed *m/z* ions specific to the control and salt conditions. This analysis revealed 44 *m/z* ions present at relatively higher abundances in the control samples, and 77 *m/z* enriched in the salt samples. Liquid chromatography-tandem MS was performed to determine the putative molecular identities of some of the mass ions enriched in each sample, including, asparagine, adenosine, and nicotianamine in the control samples, and arginine and soyasaponin I in the salt treated samples.

## Introduction

Matrix-assisted laser desorption/ionization (MALDI) is commonly used as a soft ionization technique to study a wide range of biomolecules. A powerful application of MALDI-mass spectrometry (MS) is the ability to determine the spatial distribution of molecules in a tissue slice by mass spectrometry imaging (MSI).[1] MALDI-MSI has been used to study a wide range of biomolecules, from small molecule metabolites,[2, 3] to neuropeptides,[4] and intact proteins.[5]

More recently, atmospheric pressure MALDI (AP-MALDI) was introduced, increasing the ease of sample preparation and allowing for analysis of volatile molecules, as the sample no longer needs to be placed under vacuum prior to analysis.[6, 7] Since then, AP-MALDI has been used to detect tryptic peptides,[8-10] pesticides,[11] oligosaccharides,[12] and proteolytic fragments.[13] Tandem MS has also been coupled with AP-MALDI ionization, which provides the ability to fragment molecules and use the fragmentation patterns to identify biomolecules.[14, 15] AP-MALDI is also capable of performing imaging experiments. The handling of samples at atmospheric pressure is an advantage of the technique, as shown by the imaging of lipids with a matrix that sublimes under higher vacuum.[16] As lipids ionize readily, multiple studies have reported imaging lipids with AP-MALDI.[17-19] Other applications of AP-MALDI-MSI include imaging of secondary metabolites in licorice rhizome[20] and neuropeptides in crustaceans.[18]

Recent developments in MALDI-MSI have been directed at lowering the minimum spatial resolution. Lowering spatial resolution allows for increased resolution of fine molecular features and for single-cell MALDI-MSI.[21] Although historically MALDI-MSI imaging experiments have been carried out above the low μm spatial resolution requirement for single cell imaging, recently instrument advances have lowered the minimum raster step size to allow imaging at higher spatial resolution.[22-24] AP-MALDI sources have also been introduced with optimized geometry to allow

for high spatial resolution, including a scanning microprobe AP-MALDI (TransMIT GMbH, Griessen Germany). The AP-SMALDI source has allowed for imaging of metabolites and lipids in a variety of samples at 5-20 μm spatial resolution.[17, 20, 25-27]

The AP-MALDI (ng) UHR system (MassTech Inc., Columbia, MD) is another AP-MALDI source capable of imaging at high spatial resolutions. The source is compact, allowing for easy and fast switching between ESI and AP-MALDI. High spatial resolution is achieved through an Nd:YAG 355 nm laser with a laser spot size of 10 μm and a maximum output frequency of 10 kHz. In the source, the sample plate is approximately 2 mm away from the heated MS inlet capillary.[28-30] The laser operates as a continuous raster along the rows of the sample. TARGET software (MassTech Inc., Columbia, MD) controls the source settings. ImageQuest software (Thermo Scientific, Waltham, MA) is used to correlate the XY coordinates to the MS spectra in the raw file to create a molecular map of target analytes distributed on the tissue section.

MALDI-MSI is now commonly used to investigate metabolite distribution in various plant tissues.[2, 31] In *Medicago truncatula* (Medicago), which forms a symbiotic relationship with rhizobia for biological nitrogen fixation, MALDI-MSI has been applied to root nodules to provide insight into the metabolites involved in biological nitrogen fixation.[3, 32] In addition to studying metabolites involved in biological nitrogen fixation, metabolite distribution changes in the root nodules due to stress can also be investigated with MALDI-MSI. Salt stress results in an energy cost to plants, as they reallocate more of their energy to physiological changes that allow continued function under stress.[33] This cost of energy for plants under salt stress translates into economic costs to farmers due to reduced yields.[33] The ability of legumes to form symbiotic root nodules is highly sensitive even to mild concentrations of salt that do not affect other aspects of plant growth.[34, 35] Consequently, identifying metabolite changes within symbiotic nodules under salt

stress or non-stress conditions may help us to understand why this symbiosis is so strikingly affected by moderate levels of salt stress. Salinity tolerance in Medicago has been studied by adding NaCl and monitoring metabolite changes with activity assays.[36] Gas Chromatography-MS has also been used to investigate the metabolic profile of severe salt stress.[37]

Here, the ability of an AP-MALDI (ng) UHR source coupled to a high resolution accurate mass platform to study metabolites in Medicago root nodules will be investigated. As a stand-alone source that can attach to multiple instruments, the AP-MALDI (ng) UHR source is a promising alternative to a traditional dedicated MALDI source for labs that might not have the ability to obtain a dedicated MALDI platform. Thus, a study to compare the performance of the AP-MALDI (ng) UHR source to a traditional MALDI platform and to demonstrate the application of the source to investigate metabolite changes due to stress is a valuable evaluation of the performance of the source. Initially, optimized AP-MALDI MSI of root nodules was compared to MALDI-MSI of root nodules on a commercial MALDI LTQ Orbitrap XL system. The AP-MALDI system was then used to study the metabolic response to salt stress through imaging at high spatial resolution. This study analyzed the localization changes of metabolites in Medicago root nodules during salt stress.

## Experimental Methods

### Materials

2,5-dihydroxybenzoic acid (DHB) was purchased through Acros Organics (Fisher Scientific), and α-Cyano-4-hydroxycinnamic acid (CHCA) through Sigma Aldrich. Methanol, acetonitrile, chloroform, and formic acid were purchased through Fisher Chemical (Fisher

Scientific). A Millipore system was used for double distilled water. Plain microscope slides were obtained from Fisher Scientific, and indium tin oxide coated glass slides from Delta Technologies.

Plant Growth

Seeds of *Medicago truncatula* cv. Jemalong A17 were acid scarified, surface sterilized, and vernalized for two overnights at 4°C. Seedlings were germinated at room temperature and transferred to sterilized growth pouches which contained 10 ml of Modified Nodulation Medium (MNM) which was modified from Buffered Nodulation Medium (BNM)[38] with addition of 1 mM KCl containing 100 mM of sodium chloride. Plants grown only in MNM medium were used as controls. The pouches were placed in a transparent box in a growth chamber with 16 hours light for four days. The roots were inoculated with 1 ml of *Sinorhizobium meliloti* (Rm1021) ($OD_{600}$ of 0.1), grown for another three weeks and nodules were harvested for subsequent analysis.

MSI Sample Preparation

Root nodules from control and high salt plants were trimmed from the plants with 2-4 mm of the surrounding root. Nodules were embedded in 100 mg/mL gelatin and frozen on dry ice. Nodules were sectioned at 16 um thickness on a Microm HM 525 cryostat (Thermo Scientific) at -20$^{o}$C. Sections were thaw-mounted onto plain glass microscope slides for analysis on the MALDI LTQ Orbitrap XL or indium tin oxide coated glass slides for analysis on the AP-MALDI QE-HF system. A TM Sprayer (HTX Technologies, LLC, Carrboro, NC, USA) was used to apply DHB and CHCA matrix. DHB matrix (40 mg/mL in 50% methanol, 0.1% formic acid) was applied with a 24 pass TM Sprayer method (30 s dry time in between passes, 90$^{o}$ rotation between passes and the spacing offset in between every two passes, 3 mm spacing, 1250 velocity, 80$^{o}$C temperature, and 0.05 mL/min flow rate). CHCA matrix (10 mg/mL in 70% acetonitrile, 0.1% formic acid) was applied with a 4 pass TM Sprayer method (30 s dry time in between passes, 90$^{o}$ rotation between

passes and the spacing offset in between every two passes, 1.5 mm spacing, 1200 velocity, 75°C temperature, and 0.24 mL/min flow rate). Matrix covered samples were stored in a dry box at -20°C until analysis.

Vacuum MALDI MSI

MALDI-MSI was performed on a MALDI LTQ Orbitrap XL (referred to as MALDI) mass spectrometer (Thermo Scientific, Waltham, MA, USA) equipped with a nitrogen laser in positive ion mode. LTQ Tune software (Thermo Scientific, Waltham, MA, USA) and Xcalibur (Thermo Scientific, Waltham, MA, USA) were used to select the imaging region and step size and the instrument parameters respectively. The laser energy for DHB was set at either 15 or 20 μJ (later replicates needed higher laser energy to get the same signal level as earlier replicates) and the laser energy for CHCA was set at 10 μJ. Imaging was performed on three biological replicates with technical replicates at 75 μm raster step size. The mass range was set to 100-1000 *m/z* and the resolution to 60,000. Two microscans were averaged at each pixel.

AP-MALDI MSI

AP-MALDI experiments were performed on an AP-MALDI (ng) UHR ion source (MassTech Inc., Columbia, MD) coupled to a Q Exactive-HF (Thermo Scientific, Waltham, MA, USA). Initially, the S-lens RF value, capillary temperature, and spray voltage parameters were optimized on-tissue. Imaging experiments were conducted in positive ion mode for 100-1,000 *m/z* with 60,000 resolution, 2 microscans, 1E6 AGC target, 100 ms maximum injection time, 3.25 kV spray voltage, 350°C capillary temperature, and 70% for the S-lens RF value. For DHB covered sections 40% laser energy was used, and for CHCA sections 25% laser energy was used on the AP-MALDI control software. Experiments were conducted at 30 μm raster size. TARGET ng

software (MassTech Inc., Columbia, MD) was used to set the imaging area, raster size, and laser energy. Tune software (Thermo Scientific, Waltham, MA, USA) was used to acquire data.

MSI Data Analysis

MSiReader software[39] was used to create peak lists and generate images from the data. Briefly, the interrogated zone was drawn around the nodule and root and compared to the reference zone of a matrix only area. Each technical replicate was analyzed individually in MSiReader, and all data was normalized to the total ion current. For each nodule, *m/z* in more than 15% of the total area of the interrogated zone (the root and root nodule) and less than 5% of the total area of the reference zone (matrix area) were pulled out for the MALDI data sets. *M/z* in 10% of the interrogated zone and less than 5% of the reference zone were pulled out for AP-MALDI data. The analysis was performed using a +/- 5 ppm window. Low numbers for the interrogated zone percentages were selected to ensure that peaks localized to a small region of the sample (and not just peaks localized to the entire sample) were pulled out. A low percentage for the reference zone was used to have a strict cut-off for removing matrix peaks. Different interrogated region percentages were used for the two platforms due to the difference in signal intensity and number of peaks pulled out between the two. For the MALDI system, 15% was used over 10% as using 10% pulled out many more noise peaks compared to 15%. Also, as the AP-MALDI system detected hundreds of peaks (compared to the over 2,000 mass spectral peaks detected by the MALDI system), the extra *m/z* were easier to manually verify for the AP-MALDI platform. As the AP-MALDI system produced fewer images, a lower cutoff threshold was used to generate as many good images as possible. Peak lists for biological replicates were generated by combining the technical replicate peak lists and combining duplicates (*m/z* within 5 ppm). Peak lists from the three biological replicates were combined and duplicates combined (5 ppm error) to create peak

lists for the two platforms with each matrix. All peak lists were manually validated by visual inspection of the resulting ion images. Peak lists were imported into SCiLs software (Bruker, Bremen, Germany) along with the data for statistical analysis. Centroid data was imported with linear interpolation at a mass accuracy of 0.0005 Da for the mass axis settings. Data was normalized to the total ion current after importing, and all analysis were performed with normalization to the total ion current. The discriminative analysis was performed using receiver operating characteristic (ROC) on both the nodules and root together and the nodule and root separately. For the analysis, individual spectra from three biological replicate nodules were used for each class. The two classes were control and salt. As the control and salt nodules are not necessarily the same size, a random subset of 500 spectra for each class was used for the analysis. A 5 ppm interval width and the validated peak lists were used for the analysis. Hypothesis tests were performed on all individual spectra of the same three biological replicates as the ROC test. The entire root nodule and root area for the control and salt samples were used for the test. Specifically, the t-test was used with a 5 ppm window. The peak lists generated in MSiReader was again used for the test. Principal component analysis (PCA) was performed using the mean spectra of each region (each region was drawn around the root and root nodule from a technical replicate) with a 5.0 ppm interval width, 5 components, and unit variance scaling.

Sample Extraction

Approximately 100 control nodules and 100 salt treated nodules (with 2-4 mm of surrounding root) were trimmed from the plants and flash frozen. Nodules were ground with a mortar and pestle under liquid nitrogen. A methanol/chloroform/water (Milli-Q) extraction was performed by adding in order 3 parts methanol (600 μL), 1 part chloroform (200 μL), and 4 parts water (800 μL). Samples were vortexed briefly and centrifuged for 10 min at 5000 x g and 4$^{o}$C.

The upper aqueous layer was removed, 4 parts methanol were added, and the extraction was vortexed briefly. The extraction was centrifuged again at 1500 x g for 5 min and 4°C. The supernatant (organic layer) was removed. The aqueous and organic fractions were dried down in a speedvac and saved in a -80 °C freezer prior to analysis.

LC-MS/MS for Identifications

Aqueous samples were resuspended in optima grade water with 0.1 % FA at 10 mg/mL. LC-MS/MS was performed with a Dionex Ultimate 3000 UHPLC system (Thermo Scientific, Waltham, MA, USA) equipped with a Kintex C18 column (2.1 mm internal diameter x 150 mm length, 1.7 μm particle size; Phenomenex, Torrance, CA, USA) with a corresponding guard column, and a Q Exactive MS (Thermo Scientific, Waltham, MA, USA). For separation, the column temperature was 35°C, and the mobile phases were optima grade water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). A 35-minute gradient at a flow rate of 0.3 mL/min with the following conditions was used: 0-5 min, held at 1% B; 5-10 min, linear gradient from 1-3% B; 10-18 min, linear gradient from 3-40% B; 18-22 min, linear gradient from 40-80% B; 22-27 min, column cleaning at 95% B; and 27-35 min, re-equilibration at 1% B. The injection volume was 4 μL, and the samples were kept at 10°C during analysis. The MS was operated in the positive ion mode with a scan range of *m/z* 100-1500 using a top 5 method for MS/MS. A target list, which included *m/z* more prevalent in either the control nodules or salt nodules, was used to acquire MS/MS on target *m/z*. If less than 5 *m/z* on the target list were found, then the most abundant *m/z* were chosen. The MS parameters were as follows: 70,000 resolution, 1 E6 AGC, and 100 ms maximum injection time. The settings for HCD MS/MS were as follows: 35,000 resolution, 1 E5 AGC, 100 ms max inject time, 15 s dynamic exclusion, and collision energies of 30, 35, and 40 for injections 1,2 and 3 respectively. MetFrag[40] was used to analyze the MS/MS

results by searching the $[M+H]^+$, $[M+Na]^+$, and $[M+K]^+$ adducts against the KEGG database with 5 ppm error tolerance. The *in silico* fragmentation was matched up to the top 20-40 experimental fragments of the MS/MS spectra at a 5 ppm and 0.01 Da tolerances. MS/MS spectra from the mzCloud high-resolution MS/MS database was used where possible to validate the MetFrag identification. For mzCloud analysis, LC-MS/MS results were loaded into Compound Discoverer software (Thermo Scientific). Briefly, raw files were aligned with adaptive curve setting with 5 ppm mass and 1.0 min retention time tolerances. Unknown compounds were detected with a 5 ppm mass tolerance, 3 signal to noise ratio, and 1,000,000 minimum peak intensity, and then grouped with 5 ppm mass and 0.1 min retention time tolerances. A search against the mzCloud database was then performed against all activation types with a 25 activation energy tolerance, and the intensity threshold set to true. Identifications were made if the top result in MetFrag explained almost all the major fragments and there were no other strong results in the lower scoring MetFrag results (score less than 0.8 for all other hits) or if the top result in MetFrag explained almost all the major fragments, and the compound discoverer MS/MS for this compound matched almost exactly. Arginine, soyasaponin I, asparagine, and adenosine standards were obtained to verify identifications. MS/MS parameters were the same as described for the extractions.

## Results

### AP-MALDI parameter optimization

Initially, the AP-MALDI source and QE-HF MS parameters were optimized by metabolite profiling on tissue sections. The laser energy was optimized on matrix areas by increasing the laser energy until increasing the energy no longer increased yield of matrix ions. A wide range of S-lens, capillary temperature, and spray voltage values were tested by profiling on sections of control

nodules. Ions initially enter the instrument through a heated capillary. The spray voltage is applied to the source, in this case to the plate, to assist ions into the MS. The S-lens is an ion guide behind the heated capillary consisting of a series of stacked rings that operates as a radio frequency (RF) device to capture and focus ions into a beam. Typically, larger molecules need a higher S-lens value to be efficiently transferred into the mass analyzer.[41] The temperature on the heated capillary, the voltage on the plate, and the RF value on the S-lens, all play a role in optimal detection of ions. **Figure 1(A-C)** shows the profiling results from adjusting the S-lens, spray voltage, and capillary temperature with CHCA as the matrix. Adjusting the S-lens and spray voltage values did not reveal any clear trend during the profiling experiments, although most *m/z* had an increase in signal at 80% for the S-lens RF value. The capillary temperature showed a stronger trend, as increasing the temperature resulted in an increased signal. Supplemental Figure 1 (A-C) shows similar optimization graphs for optimizing instrument parameters with DHB as the matrix.

A smaller subset of instrument parameters was further tested by performing imaging experiments on control nodules. The base parameters were 70% S-lens, 3.0 kV spray voltage, and 300$^{o}$C capillary temperature. Parameters were then individually adjusted above and below these values. **Figure 1(D-F)** shows the imaging results from adjusting the S-lens, spray voltage, and capillary temperature with CHCA as the matrix. Supplementary Figure 1 (D-F) shows the imaging optimization results with DHB. The S-lens value still showed inconsistent results as higher *m/z* tended to increase slightly with higher S-lens while lower *m/z* decreased slightly with higher S-lens. Thus, a middle value of 70% was chosen for future experiments. In the imaging experiments, increasing the spray voltage did tend to increase the signal, especially for higher *m/z* (above 600), which were not very noticeable with the lower spray voltages. Furthermore, increasing the capillary temperature increased signal, which was consistent with the profiling results. Parameters

of 3.25 kV spray voltage and 350ºC capillary temperature were chosen to give the best signal, especially for higher *m/z*. The same instrument parameters were chosen for both DHB and CHCA matrices as there was not a noticeable difference between the two matrices. Spray voltages above 3.25 kV were not attempted as higher voltages could cause discharge on the AP-MALDI electronics.

CHCA matrix showed higher overall signal and better coverage of the lipid region (*m/z* above 600) compared to imaging experiments with DHB as the matrix. In **Supplemental Figure S1**, only *m/z* 104.1073 and *m/z* 133.0607 had a signal above 10,000, and this occurred at higher spray voltages and capillary temperatures. In the CHCA imaging experiments, all but the two highest *m/z* (682.0023 and 719.9575) had a signal above 10,000 at high capillary temperatures and spray voltages. In the DHB graphs in **Supplemental Figure S1**, *m/z* above 600 were not included as there was minimal signal in this region with DHB matrix.

Comparison of AP-MALDI and Vacuum MALDI sources

Imaging experiments using control nodules on both the AP-MALDI and MALDI platforms were conducted to compare the performance of the two instruments for MSI. **Figure 2(A-H)** compares the MALDI and AP-MALDI platforms. **Figure 2(A)** shows the number of *m/z* detected that resulted in good images (i.e., not matrix peaks and signal predominantly in plant sample) for control nodules in both the MALDI and AP-MALDI systems with both DHB and CHCA matrices. It should be noted that if a *m/z* is not detected in one platform but is detected in the other, it is likely the case that the *m/z* is present in both samples, but the signal intensity in one platform was below the threshold needed to pull out the *m/z* with MSiReader. Interestingly, DHB provided better results on the MALDI, and CHCA gave better results on the AP-MALDI system. While instrumentation differences could play a role in why different matrices were best for the two

platforms, the fact that the CHCA matrix application method was a wetter method than the DHB method could also play a role. The wetter CHCA method could potentially extract metabolites form the tissue at a higher concentration. As the AP-MALDI had a lower signal than the MALDI platform, improved extraction could benefit the AP-MALDI platform more than the MALDI platform. The MALDI system detected significantly more *m/z* than the AP-MALDI system as it detected almost twice the number of *m/z* using CHCA and over six times the number of *m/z* using DHB as matrix, respectively. **Figure 2(B,C)** shows the overlap between the *m/z* detected in MALDI and AP-MALDI systems for DHB matrix and CHCA matrix. Both matrices had similar numbers of *m/z* shared between the two instruments. However, there were more unique *m/z* than shared *m/z*, especially for CHCA. **Figure 2(D,E)** compares the DHB and CHCA matrices for both the MALDI and AP-MALDI platforms. While both platforms have many *m/z* that are shared in both matrices, there is a high number of *m/z* only detected in one of the matrices, especially for DHB on the MALDI and CHCA on the AP-MALDI due to their higher number of detected *m/z* compared to the other matrix. Thus, the DHB and CHCA matrices were complementary to each other. The PCA plot in **Figure 2(F)** shows that the different matrix and platform conditions (MALDI DHB, MALDI CHCA, AP-MALDI DHB, AP-MALDI CHCA) all separate out into groups. The technical replicates group close together in most cases. While there is more variation in the biological replicates, the biological replicates from each matrix/platform experimental group are close enough together to separate them from the other matrix/platform experimental groups. **Figure 2(G,H)** shows example spectra for control nodules with DHB matrix for the MALDI **(G)** and AP-MALDI **(H)** platforms. The spectra show clear differences in *m/z* and intensity, which supports the separation of the different experimental conditions in the PCA plot. Example spectra averaged over control nodules with CHCA as the matrix for each platform are shown in

**Supplemental Figure S2**. **Supplemental Tables S1** and **S2** list the *m/z* unique to control nodules imaged with the AP-MALDI system for DHB and CHCA respectively. These *m/z* were compared to the matches to the mzCloud database from the LC-MS/MS data. The putative identifications from this analysis are shown in **Supplemental Table S3**. From these putative identifications, the AP-MALDI data is potentially detecting more acids as nicotinic acid/ picolinic acid, pyroglutamic acid, aspartic acid, DL-α-aminosuberic acid, and pantothenic acid were all putatively identified from the *m/z* unique to AP-MALDI control nodules. Further MS/MS data collection and analysis would be necessary to verify the identity of these acids or identify additional compounds unique the AP-MALDI control nodules. The high number of unique *m/z* between the two sources is potentially due to the instrumental differences. One hypothesis for the differences is that the AP-MALDI had fewer in-source fragmentation products as other studies observed that AP-MALDI is a soft ionization method with decreased and more consistent fragmentation.[6, 42] As the MALDI detected many more *m/z* and had a higher overall signal compared to the AP-MALDI, the *m/z* solely detected in the MALDI experiments could be due to an increased sensitivity. Also, the MALDI has a nitrogen laser, which operates at 337 nm, whereas the AP-MALDI uses an Nd/YAG laser (355 nm). The differing beam profiles of these lasers[43] could be causing some of the differences in *m/z* detected between the sources. The different efficiencies of the two instruments could also affect the detected ions. As the two platforms have two different Orbitrap instruments, the differences in ion transfer, detection, and fragmentation efficiencies can potentially result in some of the observed differences.

Despite the lower number of shared *m/z* between the MALDI and AP-MALDI, the distributions of the shared *m/z* were similar. **Figure 3** compares the spatial distribution of *m/z* detected in both sources. Two representative images were selected from the shared *m/z* (see the

Venn diagrams in **Figure 2(B,C)**) for each matrix. The *m/z* were chosen for their good normalized signal intensity in both platforms to make the images easy to compare, and an attempt was made to get an *m/z* spread evenly across the 100-1,000 range. For each *m/z*, the images acquired with the AP-MALDI have similar distributions as the images obtained on the MALDI system. However, the box and whisker plots showing the unnormalized intensity reveal a wide gap in the intensity of the signals between the AP-MALDI and MALDI. The optical images for the samples shown in **Figure 3** are shown in **Supplemental Figure S3**, and the box and whisker plots for three biological replicates of the AP-MALDI control nodules, and three biological replicates of the MALDI control nodules are shown in **Supplemental Figure S4**. After normalization to the total ion current, signals between the AP-MALDI and MALDI are much more comparable in the box and whisker plots. Although the overall signal on the AP-MALDI was lower than the MALDI, the AP-MALDI QE-HF instrument was still capable of determining the spatial distribution of small molecules in Medicago root nodules.

Metabolites changing due to salt stress

Overall, the quality of MS spectra obtained from salt nodules was consistent with the quality of MS spectra obtained from control nodules despite the abundance of sodium in the salt nodules. The total ion current was very similar between the control and salt nodules. For example, in one biological replicate the total ion current was 2.6E5 for control nodules versus 2.3E5 for salt nodules with CHCA and 9.6 E4 for control nodules versus 4.0E4 for salt nodules with DHB. Example spectra for salt nodules with both matrices are shown in **Supplemental Figure S5** (control nodule spectra are located in **Figure 2** for DHB and **Supplemental Figure S2** for CHCA). The largest difference between the control and salt samples was the abundance of sodium adducts in the salt samples. The higher tolerance of MALDI systems to salt could potentially account for

the fact that ion suppression due to the high salt concentrations in this study did not severely decrease the signal in the high salt samples. SCiLS software was used for statistical analysis of MSI data obtained from control and salt root nodules. ROC analysis was performed to generate area under the curve (AUC) values for specific *m/z*. ROC curves are generated by plotting the sensitivity (true positive rate) versus 100-specificity (false positive rate) for the ability of a single *m/z* value to discriminate between two conditions. AUC values, which range from 0 to 1, are calculated from the ROC curve for a specific *m/z*. By importing a peak list, an AUC value was generated from its respective ROC curve for each *m/z* in the list. An AUC cut-off of 0.75 was utilized as this resulted in a list of *m/z* that showed distinct differences between the control and salt nodules. As AUC values closest to 0.5 are less discriminative, an AUC value halfway in between 0.5 and 1 was chosen to give numerous *m/z* that were different between the control and salt root nodules. The ROC test was run on the entire nodule and root sample, just the nodule, and only the root to find *m/z* values that are discriminative to either the salt or control condition in specific regions of the sample (compared to *m/z* values present in the entire root and nodule sample). The discriminative analysis was compared to the t-test results, which were only performed on the entire root and nodule region as the t-test was less sensitive to the area selected. **Figure 4** compares the number of *m/z* selected from either control or salt samples using the three analysis methods: manual analysis, ROC analysis, or the t-test. The number of *m/z* by manual analysis was determined by looking through images for each biological replicate and selecting *m/z* that only showed signal in either the control or salt condition. The final number for the manual analysis only shows *m/z* selected in all three biological replicates.

In **Figure 4**, the results of the SCiLS analysis to find *m/z* solely in the control nodules **(A-C)** and *m/z* strictly in the salt nodules **(D-F)** are shown. The t-test found the highest number *m/z*

specific to either the control or salt root nodules, with well over half of the input *m/z* having p values less than 0.001. A number of these significant *m/z* did not appear to be changed in the images by naked eye, making it very hard to sort some *m/z* into either the control or salt group. The discriminative analysis test and manual analysis provided a more practical number of *m/z* to focus on. In most cases, the *m/z* selected in the manual and ROC analysis were found to be significant by the t-test. Differences between the manual and ROC analysis can likely be attributed to low signal in one or more biological replicates and inconsistencies in the manual sorting. Consequently, the ROC test was selected to look at the differences between the control and salt roots and root nodules. **Supplemental Table S4** lists the *m/z* and AUC values for *m/z* with AUC>0.75 for the ROC analysis on the control roots and root nodules. **Supplemental Table S5** provides the *m/z* and AUC values for *m/z* with AUC>0.75 in the ROC analysis on the salt roots and root nodules.

After combining the DHB and CHCA results, removing isotope peaks, and removing images with high background signal, 44 targets from control samples, and 77 targets from salt samples were selected. Overall, a minority of images with an AUC above 0.75 were removed due to high signal in the background. **Figure 5** shows representative images from control targets, and **Figure 6** shows images for selected targets from the salt treated root nodules. Most *m/z* with an AUC higher than 0.75 show signal uniformly distributed throughout the nodule or throughout the nodule and root. Only a couple of *m/z* values, which had an AUC higher than 0.75 just in the roots, did not show any distribution in the nodule. The images show distinct differences between the control and salt nodules with AUC's above 0.75, demonstrating the power of the ROC analysis.

Identification of metabolites from the LC-MS/MS data was performed by searching *m/z* against the KEGG database and using a combination of *in silico* fragmentation (MetFrag) and

matching to the mzCloud high-resolution MS/MS database. For the MetFrag analysis, compounds that yielded theoretical fragments matching the highest number of fragments in the experimental MS/MS spectra were considered putative identifications. If more than one compound matched the major fragments, then an attempt was made to narrow down to one candidate with MS/MS spectra in the mzCloud database. **Table 1** shows the identifications from the control list of *m/z* with AUC > 0.75. Adenosine was the best option in MetFrag results, and nicotianamine was the only KEGG hit within 5 ppm for its *m/z* (the *in silico* fragmentation results did match well), but asparagine was the highest scored MetFrag result with two good options behind it. Glycylglycine was second but was ruled out with its MS/MS spectra in mzCloud. The 3$^{rd}$ MetFrag result, N-carbamoylsarcosine, was not in mzCloud. The MS/MS spectra for asparagine in mzCloud was nearly identical to the experimental MS/MS, so it was putatively identified. Both asparagine and nicotianamine had AUC values higher than 0.75 in the nodules, while adenosine only had an AUC value higher than 0.75 in the roots, although it was also detected in the nodules. **Table 2** shows the identifications in salt roots and root nodules. Arginine was detected with an AUC higher than 0.75 in the salt nodules with MS/MS that closely matched the database spectra in mzCloud. For *m/z* 365.1045, the AUC was very high in the root nodules, roots, and roots and nodules combined, but the MS/MS was only able to distinguish the *m/z* as the sodium adduct of a disaccharide as multiple sugars ranked very high in the MetFrag analysis. Soyasaponin I was also detected as a sodium adduct and interestingly was only located to the outer portion of the root nodules and in the roots in salt nodules. **Figure 7** shows the AP-MALDI images for the *m/z* identified in **Tables 1-2.** The experimental MS/MS spectra for the identifications are shown in **Supplemental Figure S6**. Arginine, soyasaponin I, asparagine, and adenosine experimental MS/MS spectra were compared to that of obtained standards for verification of the identification. The MS/MS spectra for the

standards are in **Supplemental Figure S7**. Retention times matched closely between the experimental data and the obtained standards (values are provided in **Tables 1-2** and **Supplemental Figure S7**).

**Discussion**

Here, an AP-MALDI (ng) UHR source was utilized for imaging of Medicago root nodules at 30 μm spatial resolution. The spatial resolution provided by the AP-MALDI source is much higher than the conventional MALDI, which is 75 μm spatial resolution without oversampling. The AP-MALDI source is also compatible with multiple mass spectrometers. Here, a high-resolution accurate mass QE-HF orbitrap instrument is utilized, offering even higher mass accuracy and resolution compared to the commercial MALDI system. Furthermore, the coupling of the AP-MALDI system to a high-resolution accurate mass Orbitrap system offers distinct advantage over commercial MALDI-TOF instruments, in terms of its high mass accuracy and resolution for confident identification of small molecule metabolites.

To maximize the *m/z* detected with the AP-MALDI source, parameters were carefully optimized. The parameters selected for imaging (high capillary temperature and spray voltage) maximized the detection of most *m/z* ions. However, even with the optimized parameters, the signal in the current AP-MALDI setup was at least one order of magnitude lower than the signal with the MALDI. A previous study comparing AP and vacuum MALDI on peptides and protein digests spots revealed that although signal increased two-fold in the vacuum system, the noise level increased at a similar rate, resulting in a similar signal to noise ratios between the two.[8] While a full limit of detection and signal to noise analysis was not conducted here, the MALDI detected significantly more *m/z* than the AP-MALDI, indicating a higher sensitivity for the MALDI system.

The MALDI's superior performance regarding signal intensity and detection of *m/z* provides a powerful instrument for comprehensive analysis of tissue sections. However, the lower signal did not prevent imaging of many ions with the AP-MALDI system and its higher spatial resolution provides the ability to analyze samples with fine molecular features that may be difficult to resolve with the lower spatial resolution of the MALDI instrument. In addition, the 10 kHz laser on the AP-MALDI significantly increases the speed of image acquisition compared to the 60 Hz laser on the MALDI. A 50 x 70 pixel grid on the AP-MALDI took 26.60 minutes to image, resulting in 2.192 pixels/s acquisition speed. However, on the MALDI, a 29 x 34 pixel grid took 65.02 minutes, giving an acquisition speed of 0.2527 pixels/s. Thus, the AP-MALDI is more than 8 times faster than the MALDI. To acquire the 50 x 70 grid of the AP-MALDI, the MALDI would take 230.6 minutes compared to the 26.60 minutes of the AP-MALDI. Therefore, the AP-MALDI has an advantage over the MALDI system regarding the speed of acquisition. Furthermore, as at best half of the *m/z* detected with the AP-MALDI were also detected with the MALDI, the AP-MALDI-MSI results are complementary to the MALDI imaging results. The AP-MALDI source allows for the detection of additional small molecules and potentially labile small molecules that are not compatible with vacuum MALDI sources. By performing MALDI-MSI studies with both sources, one could increase the coverage of the metabolome in MALDI-MSI studies.

The AP-MALDI QE-HF system was used to study the metabolite changes due to salt stress with high spatial resolution and high mass accuracy. SCiLS software was used to perform statistical analysis on the MSI data to confidently assign *m/z* discriminative to the control and salt conditions. Although the t-test (p-value <0.001) gave the largest number of *m/z* as its output, the percentage of input *m/z* that were selected as significant was very high, and for some *m/z*, it was not apparent to the naked eye which group (either control or salt) was higher. Here, discriminative

analysis using an ROC test was chosen as this test gave *m/z* with a signal that was consistently distinctive to either the salt or control group and mostly avoided *m/z* with only slight changes or changes in only one biological replicate. The discriminative analysis is also beneficial over manual analysis as it avoids potential inconsistencies in sorting. A random subset of spectra was used for the analysis as the salt nodules were typically smaller than the control nodules, meaning that using all spectra would result in a different number of spectra in each class. Although using multiple spectra per sample creates a large subset to generate ROC curves, it should be noted that individual spectra from the same sample are not independent. Furthermore, the ROC curve analysis and the t-test have two different meanings. The ROC curve is looking for *m/z* that discriminate between conditions (often healthy versus diseased tissue) whereas the t-test looking for *m/z* that have significant changes between the two conditions. While a p-value<0.001 and an AUC>0.75 are not the same and provide different explanations about the data, the objective here was to compare their ability to select whichever *m/z* are changing between the conditions. As manual analysis is laborious, a statistical test to select changing *m/z* to focus identification efforts on is beneficial.

Previous studies have found changes in amino acids, organic acids, and sugars due to salt stress.[36, 37] Although sugars and amino acids were identified here as differing in the salt and control nodules, a potential pitfall of this study is that some of the metabolic differences observed could be due to the increased sodium levels in the salt samples. This creates a paradox observation where the same compound is higher in the control nodules for the $[M+H]^+$ and $[M+K]^+$ adducts, but higher in the salt treated samples for the $[M+Na]^+$ adduct. The identifications of asparagine and nicotianamine in control nodule samples show this fluctuation as they had an AUC>0.75 in the control nodules for the $[M+H]^+$ adduct but the *m/z* that accurate mass matched to the sodium adduct was higher in the salt target list (MS/MS data was not able to confirm presence in salt nodules).

Similarly, the disaccharide sodium adduct was identified in the salt treated samples, but based on accurate mass matching, the potassium adduct was shown upregulated in the control target list. In addition, the inability of traditional MALDI-MSI to separate isobaric compounds prevented identification of different sugars. As a result, the changes in sugar content was difficult to determine as the changes in the availability of sodium for adduction and the isobaric nature of the sugars complicated assignment significantly. However, in most cases, one can still identify metabolites changing due to salt stress (and not due to the differences in sodium adduct formation). For example, both the $[M+H]^+$ and $[M+Na]^+$ adducts of arginine were on the salt target list, indicating that this change is due to the effects of the stress and not due to changes in sodium availability. Although some of the compounds with AUC>0.75 are likely due to changes in sodium levels and not due to the salt stress, there are still many targets discovered that do not show the relative intensity change between the control and salt nodules with different adducts (i.e. $[M+H]^+$ adduct higher in control and $[M+Na]^+$ higher in salt). These changes in the relative intensity between different adduct species can be determined by looking at the images for the control and salt nodules on the same intensity scale for each adduct species. Thus, AP-MALDI-MSI provides a viable technique to study metabolite changes in salt stress in Medicago nodules.

We observed increased accumulation of arginine in the salt-stressed nodules. Accumulation of arginine is often seen in plants subjected to various environmental stresses, and exogenous arginine helps to tolerate the harmful effects of salt stress.[44, 45] Arginine metabolism plays a crucial role in salt tolerance in plants as discussed below. Arginine is synthesized from the non-proteinogenic amino acid ornithine. *N*-acetylglutamate synthase (NAGS) is an enzyme that catalyzes the first reaction during ornithine biosynthesis, and overexpression of the gene encoding NAGS improves salt tolerance in tomato plants.[46] Arginase catalyzes the initial reaction of arginine

degradation, and a loss of activity of this enzyme is associated with increased salt tolerance, presumably via accumulation of beneficial molecules, such as, nitric oxide (NO) and polyamines.[46-48] Ornithine δ-aminotransferase, another enzyme involved in arginine catabolism shows increased activity under salt stress.[46] The arginine decarboxylase (ADC) enzyme converts arginine to agmatine, which is a precursor of polyamines. Spermine is a polyamine often involved in salt tolerance, and its deficiency leads to salt hypersensitivity.[49] Spermine accumulation is low in salt-treated roots in a genetic background where arginine decarboxylase activity is reduced compared to the wild-type, implicating this enzyme in salt-acclimation.[50] In salt-tolerant rice, expression of the *ADC* gene is induced in the presence of salinity.[51] Single nucleotide polymorphisms (SNPs) associated with *ADC* showed a strong correlation with multiple environmental factors, such as, salinity, drought, and soil nitrogen, placing this enzyme as an essential regulator of plant-environment interactions.[52] Arginine is also involved in the production of NO with the latter implicated in salt tolerance.[53, 54] Exogenous NO, in the form of its donor S-nitroso-N-acetylpenicillamine (SNAP), alleviates the adverse effects of salt stress, presumably by upregulating Reactive Oxygen Species (ROS)- scavenging enzymes and enhancing the accumulation of osmolytes.[55] It is suggested that the accumulation of NO and other Reactive Nitrogen Species (RNS) cause nitrosative stress, which is essential for salt "priming".[47] Altogether, these results suggest an essential position of arginine metabolism in salt stress responses.

We also found an enhanced accumulation of soyasaponin I. Saponins are amphipathic glycosides found in many plant species.[56] A salt-tolerant genotype of soybean accumulates high amounts of group B saponin, alluding to its possible role in salt tolerance.[57] These findings validate our technique and demonstrate that it can be used to address significant biological questions.

Here, the AP-MALDI-MSI analysis of metabolites in salt stress demonstrated the ability of the AP-MALDI (ng) UHR source to image metabolites with high resolution in both mass and space. Despite the lower number of detected compounds due to a reduced sensitivity compared to the vacuum MALDI MS platform, a respectable number of *m/z* values were found to change in root nodules between the control and salt conditions. The spatial resolution used here was not quite at the level of single-cell imaging, but with further optimization higher spatial resolutions could be achieved as the source has the potential for 5-10 μm imaging. Overall, the AP-MALDI QE-HF platform is a robust system for analyzing small molecules, and when combined with the ease of changing between AP-MALDI and ESI on a single mass spectrometer, the source makes for a useful alternative to a traditional dedicated MALDI instrument. The custom-designed source is a cost-effective substitute for a traditional MALDI platform, allowing labs to perform imaging experiments on mass spectrometers currently used with ESI. Furthermore, the complementary detection of *m/z* between the AP-MALDI and MALDI platforms allows for wider coverage of metabolites. On-going development for a new generation of a sub-atmospheric pressure MALDI source from MassTech will offer improved sensitivity, and with continued ease of switching between ESI and MALDI operation, would allow for more comprehensive metabolome characterization of these important model systems.

**Acknowledgments**

The authors would like to acknowledge MassTech Inc. for providing the AP-MALDI (ng) source. This work was supported in part by funding from the National Science Foundation (NSF) Division of Integrative Organismal Systems (IOS) RESEARCH PGR award #1546742, University of Wisconsin-Madison Graduate School and the Wisconsin Alumni Research Foundation

(WARF), a Vilas Distinguished Achievement Professorship to L.L., and a NSF grant to JMA (NSF#0701846). LL acknowledges funding support from NIH through grants R56MH110215 and R01 DK071801. The MALDI-Orbitrap LTQ XL and Q Exactive instruments were purchased through an NIH shared instrument grant (NCRR S10RR029531 to LL).

**References**

1. Caprioli, R. M.; Farmer, T. B.; Gile, J., Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* **1997,** *69* (23), 4751-60.

2. Lee, Y. J.; Perdian, D. C.; Song, Z.; Yeung, E. S.; Nikolau, B. J., Use of mass spectrometry for imaging metabolites in plants. *Plant J* **2012,** *70* (1), 81-95.

3. Gemperline, E.; Jayaraman, D.; Maeda, J.; Ane, J. M.; Li, L., Multifaceted investigation of metabolites during nitrogen fixation in Medicago via high resolution MALDI-MS imaging and ESI-MS. *J Am Soc Mass Spectrom* **2015,** *26* (1), 149-58.

4. Chen, R.; Li, L., Mass spectral imaging and profiling of neuropeptides at the organ and cellular domains. *Anal Bioanal Chem* **2010,** *397* (8), 3185-93.

5. Chaurand, P.; Norris, J. L.; Cornett, D. S.; Mobley, J. A.; Caprioli, R. M., New developments in profiling and imaging of proteins from tissue sections by MALDI mass spectrometry. *Journal of Proteome Research* **2006,** *5* (11), 2889-2900.

6. Laiko, V. V.; Baldwin, M. A.; Burlingame, A. L., Atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry. *Anal Chem* **2000,** *72* (4), 652-7.

7. Laiko, V. V.; Moyer, S. C.; Cotter, R. J., Atmospheric pressure MALDI/ion trap mass spectrometry. *Anal Chem* **2000,** *72* (21), 5239-43.

8. Schneider, B. B.; Lock, C.; Covey, T. R., AP and vacuum MALDI on a QqLIT instrument. *J Am Soc Mass Spectrom* **2005,** *16* (2), 176-82.

9. Gudlavalleti, S. K.; Sundaram, A. K.; Razumovski, J.; Doroshenko, V., Application of atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry for rapid identification of Neisseria species. *J Biomol Tech* **2008,** *19* (3), 200-4.

10. Nguyen, J.; Russell, S. C., Targeted proteomics approach to species-level identification of Bacillus thuringiensis spores by AP-MALDI-MS. *J Am Soc Mass Spectrom* **2010,** *21* (6), 993-1001.

11. Mahale, V.; Singh, A.; Phadke, G. S.; Ghanate, A. D.; Oulkar, D. P.; Banerjee, K.; Panchagnula, V., Determination of Triazines and Triazoles in Grapes Using Atmospheric Pressure Matrix-Assisted Laser Desorption/Ionization High-Resolution Mass Spectrometry. *J AOAC Int* **2017,** *100* (3), 640-646.

12. Chong, S. L.; Nissila, T.; Ketola, R. A.; Koutaniemi, S.; Derba-Maceluch, M.; Mellerowicz, E. J.; Tenkanen, M.; Tuomainen, P., Feasibility of using atmospheric pressure matrix-assisted laser desorption/ionization with ion trap mass spectrometry in the analysis of acetylated xylooligosaccharides derived from hardwoods and Arabidopsis thaliana. *Anal Bioanal Chem* **2011,** *401* (9), 2995-3009.

13. Grasso, G.; Rizzarelli, E.; Spoto, G., AP/MALDI-MS complete characterization of the proteolytic fragments produced by the interaction of insulin degrading enzyme with bovine insulin. *J Mass Spectrom* **2007,** *42* (12), 1590-8.

14. Mayrhofer, C.; Krieger, S.; Raptakis, E.; Allmaier, G., Comparison of vacuum matrix-assisted laser desorption/ionization (MALDI) and atmospheric pressure MALDI (AP-MALDI) tandem mass spectrometry of 2-dimensional separated and trypsin-digested glomerular proteins for database search derived identification. *J Proteome Res* **2006,** *5* (8), 1967-78.

15. Ito, E.; Tominaga, A.; Waki, H.; Miseki, K.; Tomioka, A.; Nakajima, K.; Kakehi, K.; Suzuki, M.; Taniguchi, N.; Suzuki, A., Structural characterization of monosialo-, disialo- and trisialo-gangliosides by negative ion AP-MALDI-QIT-TOF mass spectrometry with MS(n) switching. *Neurochem Res* **2012,** *37* (6), 1315-24.

16. Jackson, S. N.; Muller, L.; Roux, A.; Oktem, B.; Moskovets, E.; Doroshenko, V. M.; Woods, A. S., AP-MALDI Mass Spectrometry Imaging of Gangliosides Using 2,6-Dihydroxyacetophenone. *J Am Soc Mass Spectrom* **2018**.

17. Schober, Y.; Guenther, S.; Spengler, B.; Rompp, A., Single cell matrix-assisted laser desorption/ionization mass spectrometry imaging. *Anal Chem* **2012,** *84* (15), 6293-7.

18. Chen, B.; OuYang, C.; Tian, Z.; Xu, M.; Li, L., A high resolution atmospheric pressure matrix-assisted laser desorption/ionization-quadrupole-orbitrap MS platform enables in situ analysis of biomolecules by multi-mode ionization and acquisition. *Anal Chim Acta* **2018,** *1007*, 16-25.

19. Desbenoit, N.; Walch, A.; Spengler, B.; Brunelle, A.; Rompp, A., Correlative mass spectrometry imaging, applying time-of-flight secondary ion mass spectrometry and atmospheric pressure matrix-assisted laser desorption/ionization to a single tissue section. *Rapid Commun Mass Spectrom* **2018,** *32* (2), 159-166.

20. Li, B.; Bhandari, D. R.; Janfelt, C.; Rompp, A.; Spengler, B., Natural products in Glycyrrhiza glabra (licorice) rhizome imaged at the cellular level by atmospheric pressure matrix-assisted laser desorption/ionization tandem mass spectrometry imaging. *Plant J* **2014,** *80* (1), 161-71.

21. Baker, T. C.; Han, J.; Borchers, C. H., Recent advancements in matrix-assisted laser desorption/ionization mass spectrometry imaging. *Curr Opin Biotechnol* **2017,** *43*, 62-69.

22. Zavalin, A.; Yang, J.; Caprioli, R., Laser beam filtration for high spatial resolution MALDI imaging mass spectrometry. *J Am Soc Mass Spectrom* **2013,** *24* (7), 1153-6.

23. Zavalin, A.; Yang, J.; Haase, A.; Holle, A.; Caprioli, R., Implementation of a Gaussian beam laser and aspheric optics for high spatial resolution MALDI imaging MS. *J Am Soc Mass Spectrom* **2014,** *25* (6), 1079-82.

24. Korte, A. R.; Yandeau-Nelson, M. D.; Nikolau, B. J.; Lee, Y. J., Subcellular-level resolution MALDI-MS imaging of maize leaf metabolites by MALDI-linear ion trap-Orbitrap mass spectrometer. *Anal Bioanal Chem* **2015,** *407* (8), 2301-9.

25. Bhandari, D. R.; Shen, T.; Rompp, A.; Zorn, H.; Spengler, B., Analysis of cyathane-type diterpenoids from Cyathus striatus and Hericium erinaceus by high-resolution MALDI MS imaging. *Anal Bioanal Chem* **2014,** *406* (3), 695-704.

26. Bhandari, D. R.; Schott, M.; Rompp, A.; Vilcinskas, A.; Spengler, B., Metabolite localization by atmospheric pressure high-resolution scanning microprobe matrix-assisted laser desorption/ionization mass spectrometry imaging in whole-body sections and individual organs of the rove beetle Paederus riparius. *Anal Bioanal Chem* **2015,** *407* (8), 2189-201.

27. Khalil, S. M.; Rompp, A.; Pretzel, J.; Becker, K.; Spengler, B., Phospholipid Topography of Whole-Body Sections of the Anopheles stephensi Mosquito, Characterized by High-Resolution Atmospheric-Pressure Scanning Microprobe Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging. *Anal Chem* **2015,** *87* (22), 11309-16.

28. Kellersberger, K. A.; Tan, P. V.; Laiko, V. V.; Doroshenko, V. M.; Fabris, D., Atmospheric pressure MALDI-fourier transform mass spectrometry. *Anal Chem* **2004,** *76* (14), 3930-4.

29. Tan, P. V.; Laiko, V. V.; Doroshenko, V. M., Atmospheric pressure MALDI with pulsed dynamic focusing for high-efficiency transmission of ions into a mass spectrometer. *Anal Chem* **2004,** *76* (9), 2462-9.

30. Moskovets, E.; Misharin, A.; Laiko, V.; Doroshenko, V., A comparative study on the analytical utility of atmospheric and low-pressure MALDI sources for the mass spectrometric characterization of peptides. *Methods* **2016,** *104*, 21-32.

31. Kaspar, S.; Peukert, M.; Svatos, A.; Matros, A.; Mock, H. P., MALDI-imaging mass spectrometry - An emerging technique in plant biology. *Proteomics* **2011,** *11* (9), 1840-50.

32. Ye, H.; Gemperline, E.; Venkateshwaran, M.; Chen, R.; Delaux, P. M.; Howes-Podoll, M.; Ane, J. M.; Li, L., MALDI mass spectrometry-assisted molecular imaging of metabolites

during nitrogen fixation in the Medicago truncatula-Sinorhizobium meliloti symbiosis. *Plant J* **2013,** *75* (1), 130-45.

33. Munns, R.; Gilliham, M., Salinity tolerance of crops - what is the cost? *New Phytol* **2015,** *208* (3), 668-73.

34. Soussi, M.; Ocaña, A.; Lluch, C., Effects of salt stress on growth, photosynthesis and nitrogen fixation in chick-pea (Cicer arietinum L.). *Journal of Experimental Botany* **1998,** *49* (325), 1329-1337.

35. Zahran, H. H., Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol Mol Biol Rev* **1999,** *63* (4), 968-89, table of contents.

36. Lopez, M.; Herrera-Cervera, J. A.; Iribarne, C.; Tejera, N. A.; Lluch, C., Growth and nitrogen fixation in Lotus japonicus and Medicago truncatula under NaCl stress: nodule carbon metabolism. *J Plant Physiol* **2008,** *165* (6), 641-50.

37. Del-Saz, N. F.; Florez-Sarasa, I.; Clemente-Moreno, M. J.; Mhadhbi, H.; Flexas, J.; Fernie, A. R.; Ribas-Carbo, M., Salinity tolerance is related to cyanide-resistant alternative respiration in Medicago truncatula under sudden severe stress. *Plant Cell Environ* **2016,** *39* (11), 2361-2369.

38. Ehrhardt, D. W.; Atkinson, E. M.; Long, S. R., Depolarization of alfalfa root hair membrane potential by Rhizobium meliloti Nod factors. *Science* **1992,** *256* (5059), 998-1000.

39. Robichaud, G.; Garrard, K. P.; Barry, J. A.; Muddiman, D. C., MSiReader: An Open-Source Interface to View and Analyze High Resolving Power MS Imaging Files on Matlab Platform. *Journal of the American Society for Mass Spectrometry* **2013,** *24* (5), 718-721.

40. Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S., MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform* **2016,** *8*, 3.

41. Eliuk, S.; Makarov, A., Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annu Rev Anal Chem (Palo Alto Calif)* **2015,** *8*, 61-80.

42. Schulz, E.; Karas, M.; Rosu, F.; Gabelica, V., Influence of the matrix on analyte fragmentation in atmospheric pressure MALDI. *Journal of the American Society for Mass Spectrometry* **2006,** *17* (7), 1005-1013.

43. Holle, A.; Haase, A.; Kayser, M.; Hohndorf, J., Optimizing UV laser focus profiles for improved MALDI performance. *J Mass Spectrom* **2006,** *41* (6), 705-16.

44. Rare, E., Stress physiology: The functional significance of the accumulation of nitrogen-containing compounds. *Journal of Horticultural Science* **1990,** *65* (3), 231-243.

45. Kakkar, R. K.; Bhaduri, S.; Rai, V. K.; Kumar, S., Amelioration of NaCl stress by arginine in rice seedlings: changes in endogenous polyamines. *Biologia Plantarum* **2000,** *43* (3), 419-422.

46. Winter, G.; Todd, C. D.; Trovato, M.; Forlani, G.; Funck, D., Physiological implications of arginine metabolism in plants. *Front Plant Sci* **2015,** *6*, 534.

47. Molassiotis, A.; Tanou, G.; Diamantidis, G., NO says more than 'YES' to salt tolerance: Salt priming and systemic nitric oxide signaling in plants. *Plant Signal Behav* **2010,** *5* (3), 209-12.

48. Masson, P. H.; Takahashi, T.; Angelini, R., Editorial: Molecular Mechanisms Underlying Polyamine Functions in Plants. *Front Plant Sci* **2017,** *8*, 14.

49. Yamaguchi, K.; Takahashi, Y.; Berberich, T.; Imai, A.; Miyazaki, A.; Takahashi, T.; Michael, A.; Kusano, T., The polyamine spermine protects against high salt stress in Arabidopsis thaliana. *FEBS Lett* **2006,** *580* (30), 6783-8.

50. Kasinathan, V.; Wingler, A., Effect of reduced arginine decarboxylase activity on salt tolerance and on polyamine formation during salt stress in Arabidopsis thaliana. *Physiol Plant* **2004,** *121* (1), 101-107.

51. Chattopadhyay, M. K.; Gupta, S.; Sengupta, D. N.; Ghosh, B., Expression of arginine decarboxylase in seedlings of indica rice (Oryza sativa L.) cultivars as affected by salinity stress. *Plant Mol Biol* **1997,** *34* (3), 477-83.

52. Guerrero, J.; Andrello, M.; Burgarella, C.; Manel, S., Soil environment is a key driver of adaptation in Medicago truncatula: new insights from landscape genomics. *New Phytol* **2018,** *219* (1), 378-390.

53. Farooq, M.; Siddique, K. H. M.; Schubert, S., Role of Nitric Oxide in Improving Plant Resistance Against Salt Stress. In *Ecophysiology and Responses of Plants under Salt Stress*, Ahmad, P.; Azooz, M. M.; Prasad, M. N. V., Eds. Springer New York: New York, NY, 2013; pp 413-424.

54. Domingos, P.; Prado, A. M.; Wong, A.; Gehring, C.; Feijo, J. A., Nitric oxide: a multitasked signaling gas in plants. *Mol Plant* **2015,** *8* (4), 506-20.

55. Ahmad, P.; Abdel Latef, A. A.; Hashem, A.; Abd Allah, E. F.; Gucel, S.; Tran, L. S., Nitric Oxide Mitigates Salt Stress by Regulating Levels of Osmolytes and Antioxidant Enzymes in Chickpea. *Front Plant Sci* **2016,** *7*, 347.

56. Osbourn, A., Saponins and plant defence — a soap story. *Trends in Plant Science* **1996,** *1* (1), 4-9.

57. Wu, W.; Zhang, Q.; Zhu, Y.; Lam, H. M.; Cai, Z.; Guo, D., Comparative metabolic profiling reveals secondary metabolites correlated with soybean salt tolerance. *J Agric Food Chem* **2008,** *56* (23), 11132-8.

## Tables

**Table 1.** Identifications from Control Roots and Root Nodules with AUC>0.75

| **m/z; Retention time (min)** | **Distribution** | **AUC >0.75 Location** | **Identification; Adduct Identified** | **Literature Molecular Weight** | **Delta ppm** |
|---|---|---|---|---|---|
| 133.0606; 1.05 | Nodule | Nodule and Root | Asparagine $[M+H]^+$ | 132.0535 | -1.48 |
| 268.1034; 3.55 | Nodule and Root | Root | Adenosine $[M+H]^+$ | 267.0968 | -2.53 |
| 304.1493; 1.05 | Nodule and Root | Nodule | Nicotianamine $[M+H]^+$ | 303.1430 | -3.31 |

**Table 2.** Identifications from Salt Treated Roots and Root Nodules with AUC>0.75

| ***m/z*; Retention Time (min)** | **Distribution** | **AUC >0.75 Location** | **Identification; Adduct Identified** | **Literature Molecular Weight** | **Delta ppm** |
|---|---|---|---|---|---|
| 175.1186; 1.04 | Nodule | Nodule | Arginine $[M+H]^+$ | 174.1117 | -1.81 |
| 365.1045; 1.18 | Root and Nodule | Root and Nodule | Disaccharide $[M+Na]^+$ | 342.1162 | 2.74 |
| 965.5076; 20.46 | Root and Outer Nodule | Root | Soyasaponin I $[M+Na]^+$ | 942.5188 | -0.44 |

## Figures

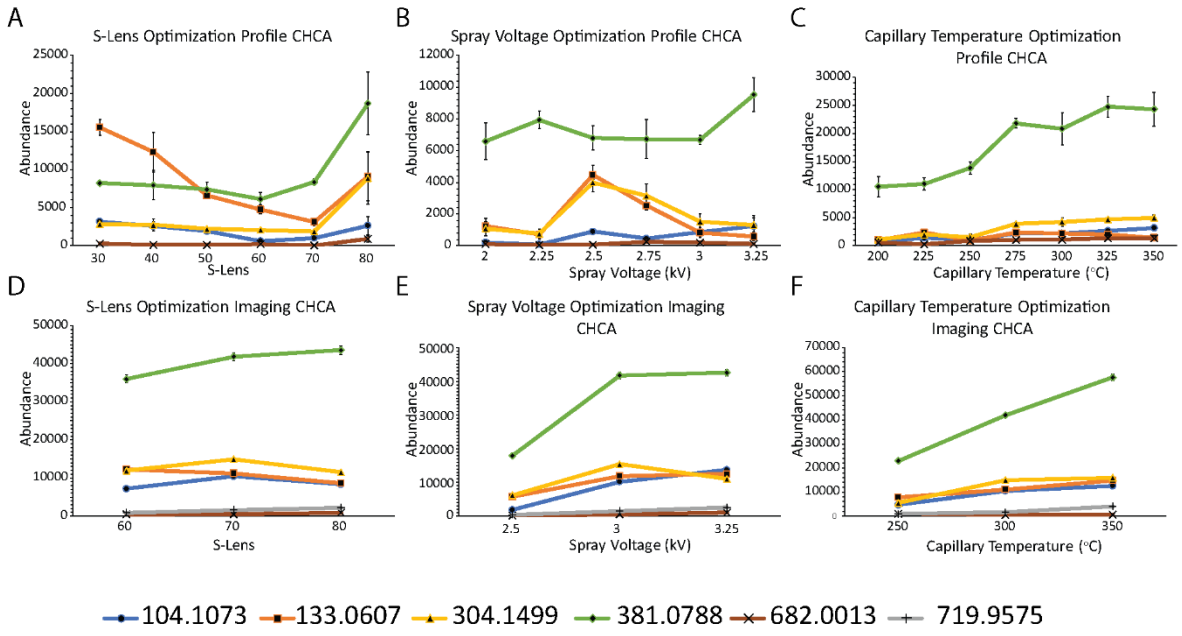



**Figure 1**. QE-HF parameter optimization graphs with CHCA as the matrix. **(A-C)** show optimization of S-lens **(A)**, spray voltage **(B)**, and capillary temperature **(C)** by profiling on tissue. **(D-F)** show optimization of S-lens **(D)**, spray voltage **(E)**, and capillary temperature **(F)** by imaging individual control nodules. Different *m/z* values are indicated by line color and data point shape. Error bars indicate standard error of the mean.

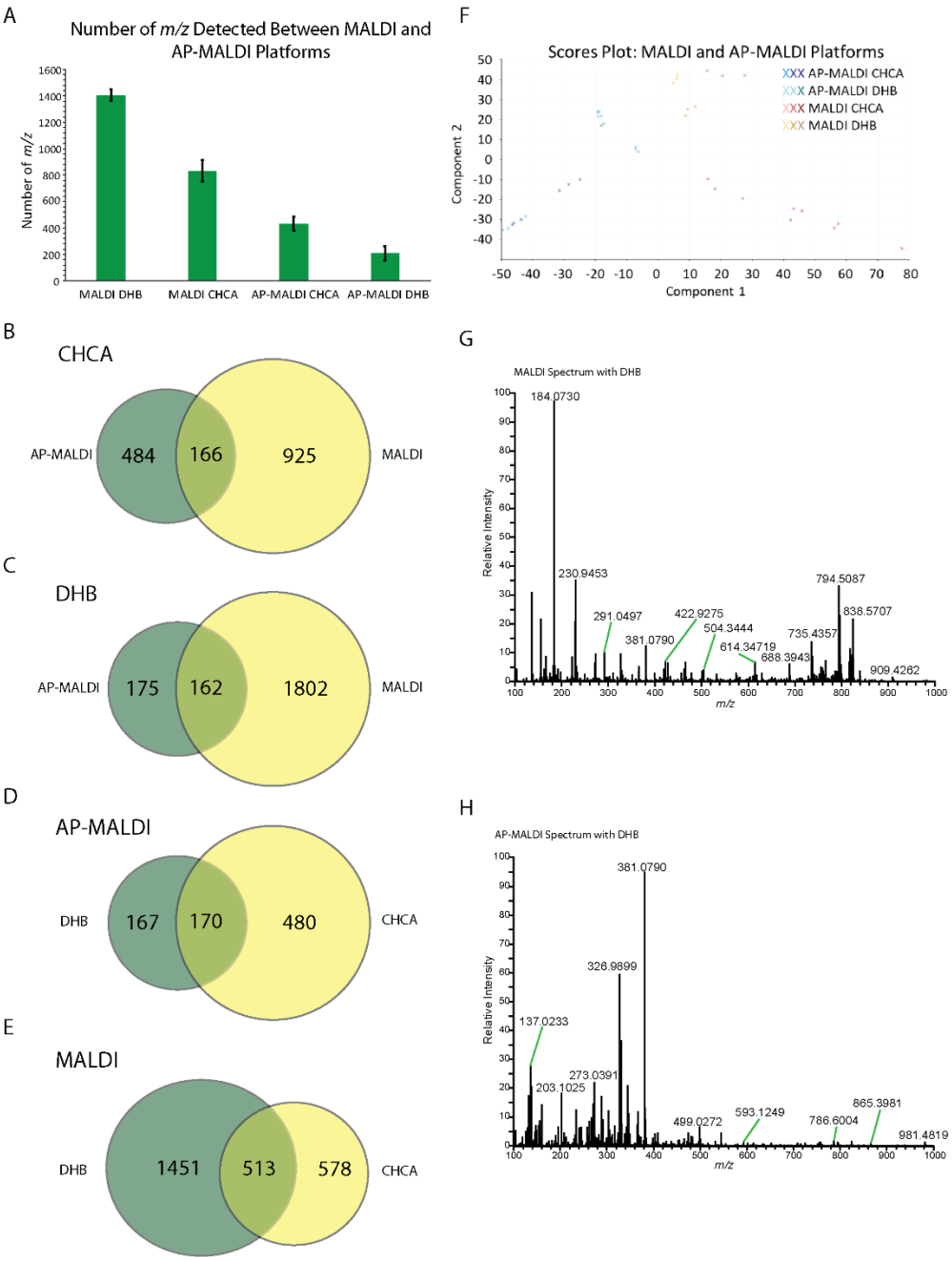



**Figure 2**. Comparison of the vacuum MALDI and AP-MALDI QE-HF systems for imaging of control nodules. In **(A)**, the number of *m/z* detected for both systems with CHCA and DHB as matrices are shown. Error bars show the standard error of the mean. **(B)** and **(C)** show Venn diagrams for the overlap in detected *m/z* values between the two systems for CHCA **(B)** and DHB **(C)**. Venn diagrams comparing the overlap between *m/z* observed with DHB and CHCA matrices are shown for the AP-MALDI **(D)** and MALDI **(E)**. The PCA plot for all the biological and technical replicates of control nodules imaged with either the AP-MALDI or MALDI platform with either DHB or CHCA is shown in **(F)**. For each condition technical replicates are all the same color and biological replicates are differing shades of a color. Example spectra averaged over the nodule with the DHB matrix are shown for the MALDI **(G)** and AP-MALDI **(H)** platforms.

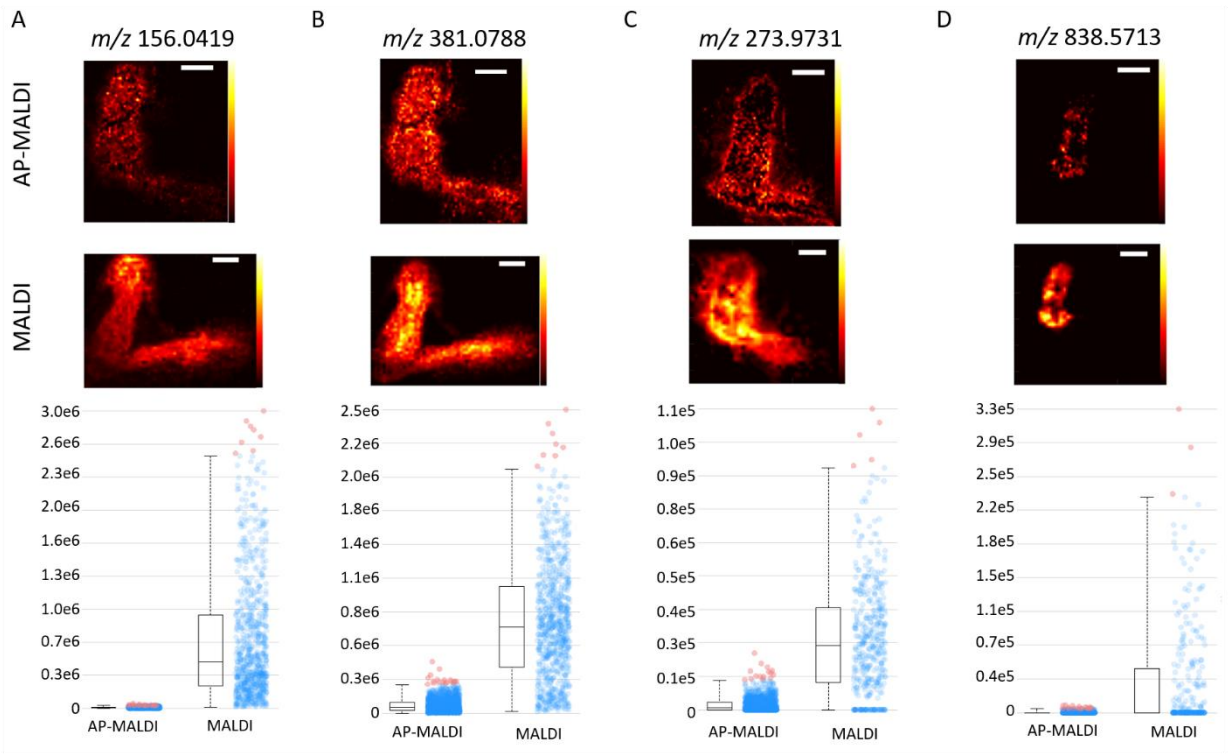



**Figure 3**. Comparison of images detected in both the AP-MALDI and MALDI platforms. Each part in **(A-D)** depicts a different *m/z* with a +/- 5 ppm window. For each part, the AP-MALDI image is shown on the top, the MALDI image in the middle, and the box and whisker plot on the bottom. **(A,B)** are from the DHB matrix data and **(C,D)** are from the CHCA matrix data. The white scale bar corresponds to 1 mm.

A Control *m/z*

| | Manual | ROC Test | t-test |
|---|---|---|---|
| DHB | 14 | 10 | 322 |
| CHCA | 38 | 50 | 531 |

D Salt *m/z*

| | Manual | ROC Test | t-test |
|---|---|---|---|
| DHB | 14 | 35 | 295 |
| CHCA | 78 | 121 | 471 |

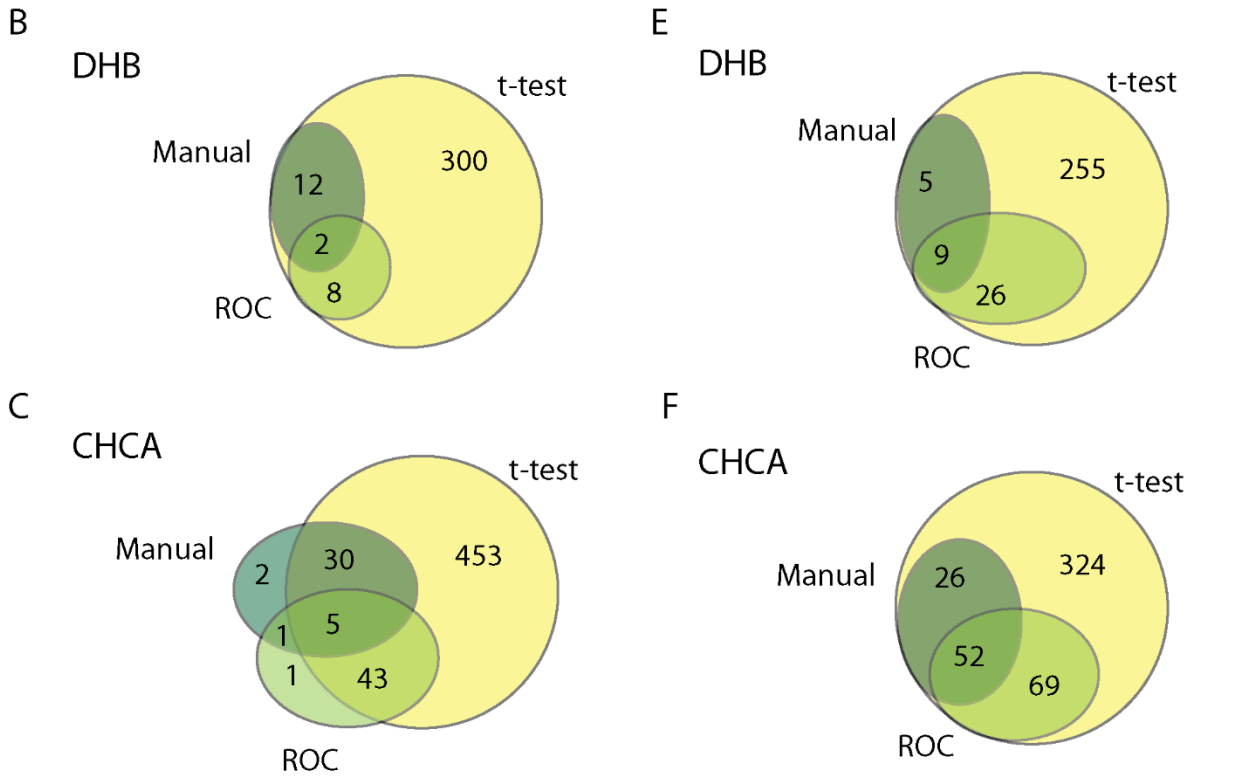



**Figure 4**. Overview of the SCiLS statistical analysis on the control *vs.* salt nodules and roots. **(A)** gives the number of significant *m/z* values determined for three analysis types in control nodules: manual analysis, discriminative analysis in SCiLS (ROC), and hypothesis test (t-test) in SCiLS software. **(B)** and **(C)** compare the three types of analysis for DHB and CHCA matrix respectively using Venn diagrams. **(D-F)** gives the same data as **(A-C)** only for significant *m/z* in the salt samples.

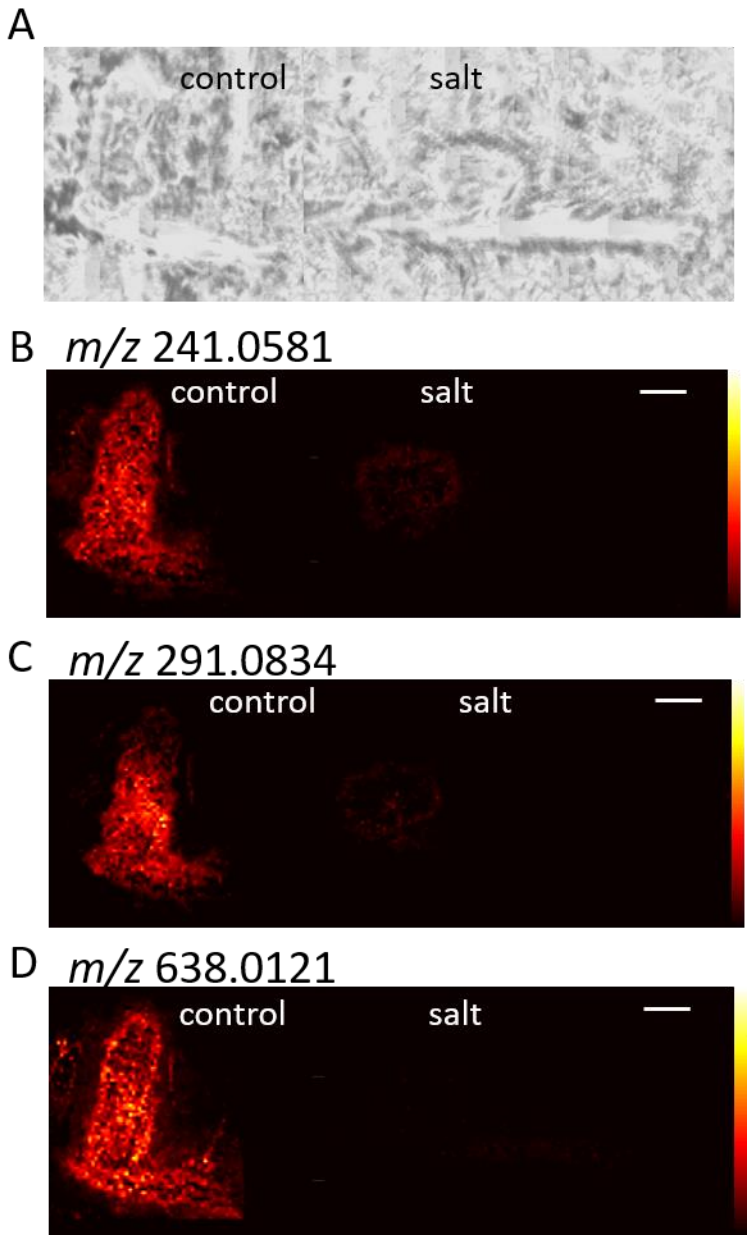



**Figure 5.** Example images for control *m/z* with AUC values above 0.75. The optical image is shown in **(A)** and **(B-D)** show three different *m/z*. CHCA was the matrix for all images shown. The white scale bar indicates 1 mm.

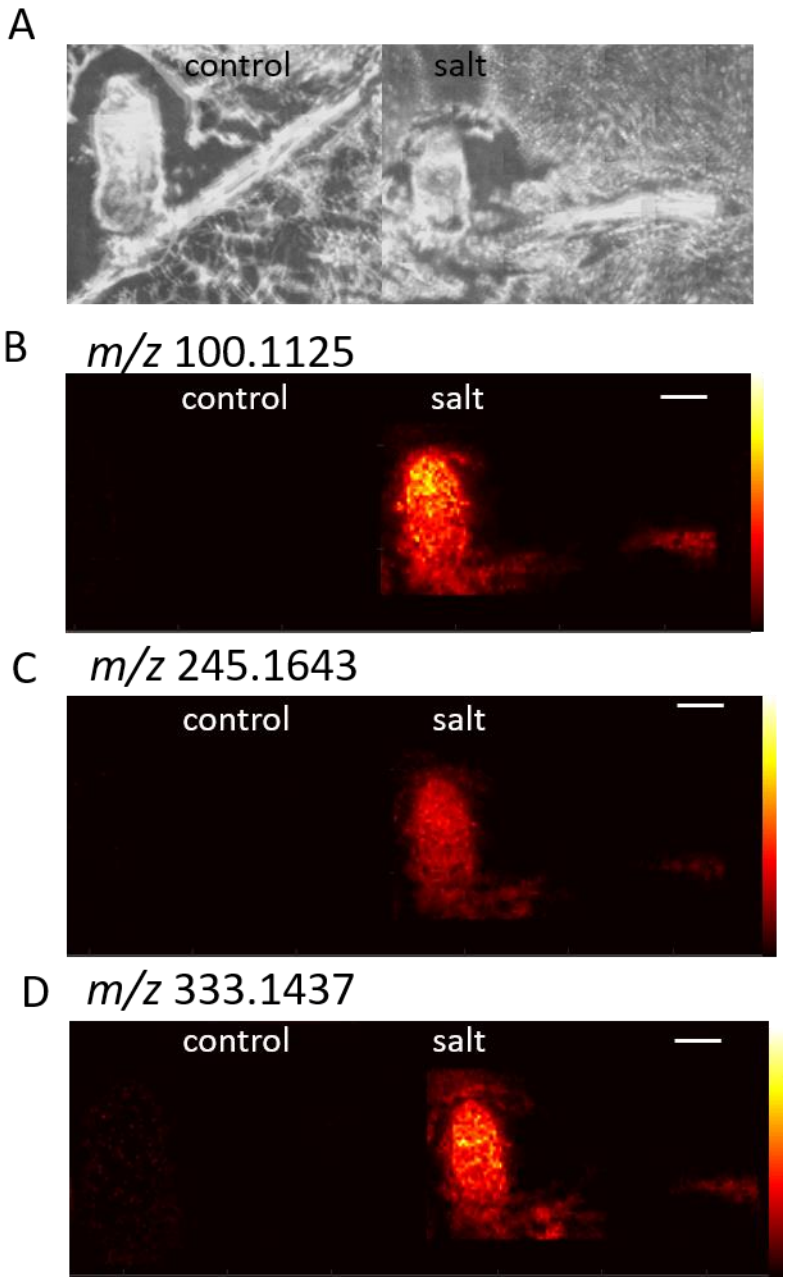



**Figure 6**. Example images for selected *m/z* ions in salt treated nodules with AUC values above 0.75. The optical image is shown in **(A)** and **(B-D)** show three different *m/z*. CHCA was the matrix for all images shown. The white scale bar indicates 1 mm.

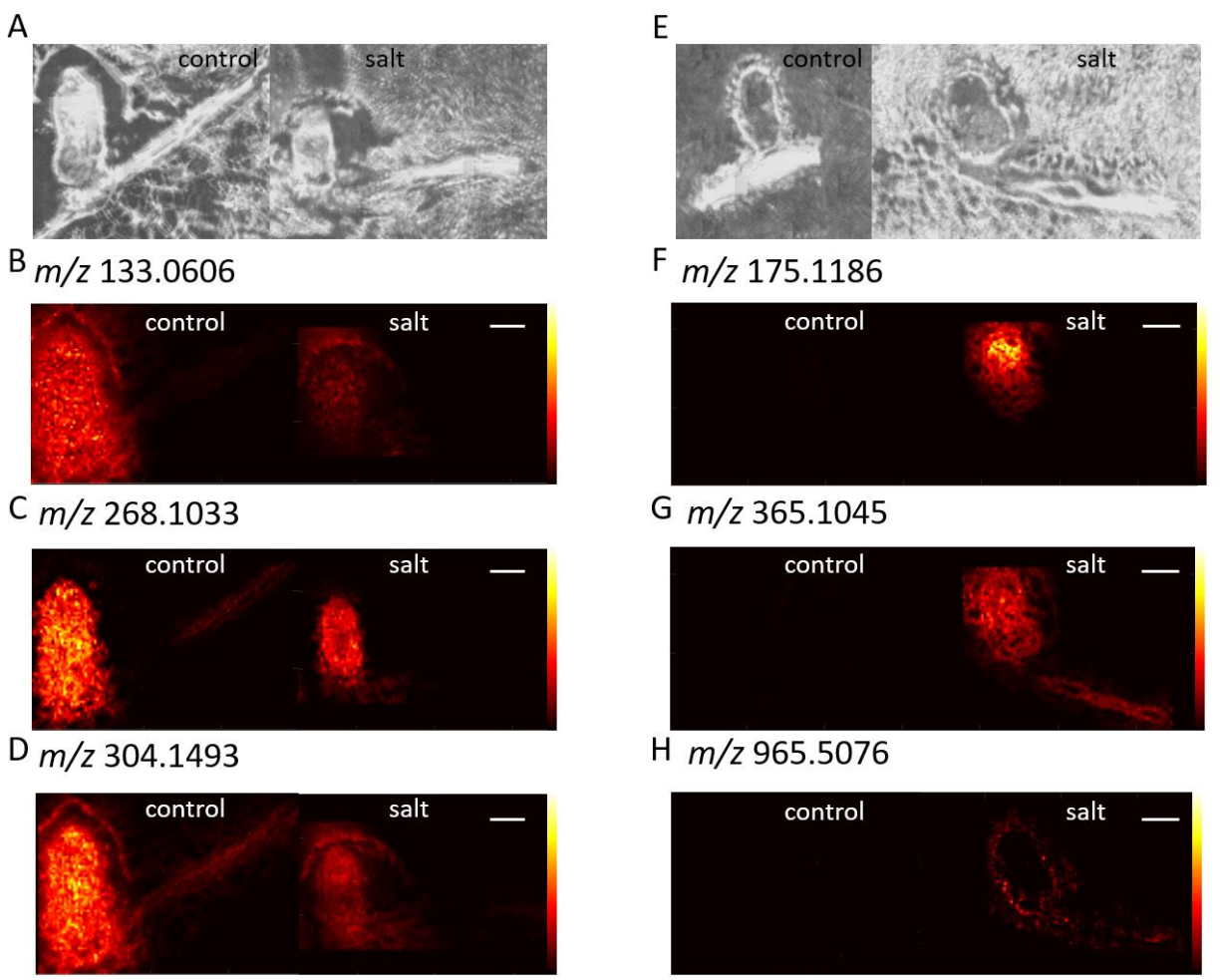



**Figure 7.** AP-MALDI MSI images for the identifications in Table 1 **(B-D)** and Table 2 **(F-H)**. **(A)** shows the optical image for the control identification **(B-D)** and **(E)** shows the optical images for salt identifications **(F-H)**. Images in **(B-D,F-G)** were with CHCA as the matrix and **(H)** was with DHB as the matrix. The white scale bar corresponds to 1 mm.

## Supplemental Material

## Supplementary Tables

**Supplemental Table S1:** *m/z* unique to AP-MALDI Control Nodules with DHB

| | | | | |
|---|---|---|---|---|
| 104.0708 | 229.9373 | 340.9816 | 398.1584 | 518.9724 |
| 112.0874 | 230.9454 | 342.1171 | 398.9512 | 528.1338 |
| 112.1123 | 230.9903 | 342.9976 | 401.0478 | 545.2182 |
| 116.0344 | 247.1650 | 343.9917 | 404.1760 | 546.2217 |
| 120.0658 | 260.1606 | 344.1226 | 405.0925 | 547.1189 |
| 123.0552 | 268.9454 | 346.0036 | 405.1976 | 549.1352 |
| 124.0391 | 269.0762 | 349.0305 | 407.0582 | 558.9439 |
| 126.1278 | 275.1022 | 350.9663 | 410.1071 | 564.1346 |
| 127.0391 | 286.1394 | 354.9612 | 421.0875 | 577.1299 |
| 129.1387 | 288.1553 | 355.0419 | 422.0924 | 591.1094 |
| 130.0499 | 289.1397 | 357.0366 | 432.1708 | 594.1280 |
| 132.0656 | 291.1541 | 360.1496 | 436.0178 | 595.1402 |
| 134.0448 | 294.1544 | 363.9376 | 438.0316 | 601.0376 |
| 134.0640 | 294.9765 | 364.9457 | 439.0977 | 611.0624 |
| 139.0586 | 294.9998 | 365.0269 | 440.1015 | 614.0917 |
| 144.1020 | 296.1125 | 365.9530 | 440.9983 | 619.0474 |
| 146.1175 | 298.1281 | 366.1080 | 441.1135 | 621.1307 |
| 159.1127 | 299.1267 | 366.9444 | 446.1866 | 633.0265 |
| 161.0922 | 300.0022 | 366.9609 | 454.0269 | 637.0579 |
| 161.1842 | 305.1532 | 367.9519 | 455.1156 | 663.1665 |
| 162.0762 | 306.1655 | 367.9646 | 455.3515 | 707.0617 |
| 175.1077 | 307.0322 | 368.0970 | 462.9225 | 709.0772 |
| 187.1078 | 307.0785 | 368.1608 | 464.0125 | 715.5100 |
| 188.0706 | 308.0164 | 368.9594 | 465.0215 | 716.5216 |
| 188.0916 | 310.9940 | 369.1649 | 465.3343 | 725.0723 |
| 190.1072 | 311.9983 | 376.9900 | 466.0256 | 763.0278 |
| 191.1027 | 317.0052 | 378.0581 | 474.9822 | 922.6170 |
| 195.0864 | 318.1169 | 380.0090 | 475.1764 | |
| 204.1060 | 319.0209 | 381.9484 | 476.1606 | |
| 205.1182 | 319.1133 | 383.9597 | 476.9981 | |
| 212.1127 | 325.1128 | 384.9547 | 490.1762 | |
| 212.8419 | 327.9970 | 384.9717 | 490.9772 | |
| 214.8403 | 328.1165 | 385.9751 | 491.1295 | |
| 217.1545 | 331.0207 | 386.9874 | 492.9935 | |
| 220.1176 | 338.9662 | 388.1808 | 498.0162 | |
| 222.0398 | 339.9702 | 392.0970 | 509.1400 | |
| 222.9883 | 340.1024 | 395.1073 | 517.0115 | |

**Supplemental Table S2.** *m/z* unique to AP-MALDI Control Nodules with CHCA

| | | | | | | |
|---|---|---|---|---|---|---|
| 100.1123 | 204.1057 | 284.0985 | 384.0423 | 452.9873 | 624.3857 | 787.4876 |
| 104.0709 | 204.2261 | 285.0128 | 384.0804 | 453.3355 | 626.0345 | 788.0855 |
| 112.0504 | 205.0815 | 285.9432 | 385.0368 | 453.9911 | 628.8844 | 788.6091 |
| 112.0867 | 205.1181 | 286.1388 | 386.0261 | 455.3511 | 629.4537 | 789.0971 |
| 112.1123 | 205.1332 | 288.0259 | 387.0726 | 457.0188 | 635.9955 | 789.3439 |
| 114.0916 | 206.0844 | 288.1549 | 388.0401 | 458.0219 | 639.0148 | 790.1003 |
| 115.0948 | 207.1485 | 289.1392 | 390.0569 | 465.1389 | 640.0102 | 790.3482 |
| 116.0343 | 208.0965 | 289.1538 | 392.0517 | 467.1552 | 642.0284 | 795.1314 |
| 116.1070 | 210.1230 | 291.0834 | 392.0721 | 473.0886 | 644.0246 | 797.1187 |
| 120.0654 | 213.0738 | 291.0965 | 392.1592 | 474.1033 | 644.0461 | 798.1219 |
| 123.0553 | 214.1550 | 291.1698 | 394.1752 | 475.1757 | 646.3471 | 805.4128 |
| 124.0392 | 215.1023 | 293.1117 | 395.1784 | 476.1573 | 648.6280 | 813.5006 |
| 126.1278 | 216.1129 | 294.1539 | 396.1546 | 476.1852 | 649.6311 | 814.5034 |
| 127.0388 | 218.1490 | 298.1157 | 398.0137 | 482.1534 | 650.0572 | 825.0373 |
| 127.1309 | 219.0970 | 299.1265 | 398.1576 | 483.1563 | 654.1804 | 826.0421 |
| 129.0656 | 220.1175 | 301.0937 | 400.1369 | 489.0780 | 662.9688 | 827.0527 |
| 129.1386 | 221.9714 | 303.0472 | 401.0349 | 490.1361 | 664.6224 | 828.0559 |
| 130.0498 | 222.1114 | 304.1493 | 403.0650 | 490.1744 | 679.5972 | 829.0554 |
| 132.0655 | 222.9750 | 304.1650 | 404.1743 | 494.9752 | 679.9850 | 831.0730 |
| 133.0606 | 223.9694 | 305.1529 | 404.9798 | 495.9785 | 680.6010 | 833.0887 |
| 134.0444 | 223.9870 | 305.9154 | 405.0791 | 496.9737 | 682.0008 | 834.0916 |
| 134.0639 | 225.9855 | 305.9328 | 405.1149 | 496.9911 | 683.0036 | 835.0805 |
| 136.0616 | 226.0109 | 306.1548 | 407.8975 | 497.9767 | 683.9989 | 849.4024 |
| 137.0648 | 227.0628 | 306.1651 | 408.9971 | 497.9949 | 688.0153 | 851.4179 |
| 139.0579 | 230.0211 | 307.9711 | 409.0545 | 506.0121 | 693.6128 | 852.4216 |
| 141.1382 | 230.1500 | 310.8090 | 409.0784 | 509.1292 | 694.6163 | 856.0602 |
| 144.0802 | 234.0186 | 311.1127 | 409.3455 | 510.1312 | 699.1038 | 862.9948 |
| 146.0809 | 234.1233 | 313.1414 | 412.0161 | 510.9700 | 701.1188 | 865.3966 |
| 146.1174 | 235.0224 | 316.1491 | 413.0114 | 511.1447 | 702.1230 | 866.3997 |
| 146.1649 | 235.1268 | 319.0422 | 413.0283 | 512.1470 | 703.3453 | 867.0116 |
| 147.0762 | 236.0167 | 320.1805 | 414.0147 | 517.0025 | 703.9840 | 868.0132 |
| 147.1682 | 237.1016 | 321.0576 | 414.0327 | 517.1328 | 714.5057 | 869.0265 |
| 148.0598 | 238.1053 | 322.1055 | 420.1536 | 520.1471 | 715.5090 | 871.0416 |
| 149.0632 | 239.1177 | 324.0955 | 421.0778 | 526.1426 | 716.5209 | 871.4719 |
| 150.0579 | 242.0619 | 328.1166 | 421.1490 | 528.0924 | 717.5244 | 872.0449 |
| 150.0772 | 242.1495 | 329.9327 | 422.0834 | 537.1949 | 718.5353 | 873.0441 |
| 156.0763 | 243.0371 | 331.0838 | 423.1647 | 543.1319 | 719.5402 | 874.0484 |
| 157.0969 | 245.0053 | 331.1041 | 423.3609 | 545.2181 | 722.9734 | 888.1434 |
| 159.0759 | 246.0162 | 331.2122 | 424.1679 | 546.9379 | 723.9738 | 889.9673 |
| 159.1127 | 247.0571 | 332.0875 | 424.3651 | 552.9716 | 728.5221 | 893.0158 |
| 160.1805 | 247.1067 | 333.0993 | 424.9926 | 554.0502 | 730.5369 | 894.0149 |
| 161.0920 | 249.1226 | 336.2267 | 425.9968 | 555.0546 | 732.5521 | 908.9959 |
| 161.1781 | 249.1586 | 343.1082 | 429.0238 | 555.0899 | 733.5565 | 910.0017 |
| 161.1839 | 250.1909 | 343.1159 | 430.0086 | 555.2791 | 738.5058 | 910.9977 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 162.0759 | 252.1437 | 344.0521 | 431.0268 | 556.0672 | 740.5210 | 932.1352 |
| 172.1807 | 255.1367 | 344.1025 | 432.1694 | 556.9310 | 741.5253 | 933.1386 |
| 173.0916 | 258.0737 | 346.0464 | 433.0785 | 557.0705 | 742.5363 | 937.4175 |
| 174.0755 | 258.1095 | 346.0676 | 433.1324 | 558.0826 | 743.0919 | 947.9587 |
| 174.1967 | 258.1441 | 347.9257 | 434.1336 | 566.2024 | 743.5404 | 948.9536 |
| 175.0709 | 259.1283 | 348.0698 | 435.0577 | 568.2189 | 744.5524 | 948.9634 |
| 175.1076 | 260.1599 | 348.0830 | 435.0904 | 572.0820 | 745.1072 | 949.9581 |
| 175.1188 | 261.1225 | 348.1548 | 435.1099 | 574.1463 | 745.5557 | 949.9673 |
| 175.1438 | 263.1595 | 348.1699 | 435.9701 | 576.2209 | 746.5662 | 950.9765 |
| 177.0978 | 263.9671 | 349.0528 | 436.1486 | 576.9213 | 753.1290 | 951.9804 |
| 178.1336 | 264.2064 | 349.0939 | 436.9736 | 581.1860 | 753.4657 | 963.4711 |
| 185.0914 | 265.9610 | 349.2229 | 436.9925 | 583.1648 | 754.1328 | 979.4644 |
| 186.0759 | 266.9644 | 349.9226 | 437.1519 | 590.9266 | 755.0891 | 981.4787 |
| 186.1963 | 267.9593 | 350.1852 | 437.3407 | 592.1016 | 755.4808 | 982.4848 |
| 187.1076 | 267.9667 | 353.0842 | 437.9686 | 592.9255 | 756.4919 | |
| 188.0546 | 267.9764 | 360.0471 | 437.9861 | 598.0398 | 759.3706 | |
| 188.0706 | 268.1034 | 360.0607 | 438.1641 | 600.0558 | 761.5862 | |
| 188.0909 | 268.9623 | 360.1494 | 439.1682 | 601.0597 | 772.5834 | |
| 188.1756 | 269.1068 | 361.1518 | 439.3560 | 607.0943 | 773.5860 | |
| 189.1343 | 269.9570 | 363.0686 | 440.1442 | 608.0769 | 775.3656 | |
| 191.1023 | 272.0791 | 365.0083 | 440.3602 | 608.3909 | 778.4789 | |
| 193.0970 | 278.1126 | 365.1047 | 441.8558 | 609.0801 | 781.0488 | |
| 193.1330 | 280.1751 | 367.0235 | 443.0032 | 610.0374 | 781.1424 | |
| 197.0664 | 280.9924 | 369.0397 | 446.1460 | 610.0771 | 783.0635 | |
| 203.1174 | 281.9560 | 373.0300 | 446.1850 | 610.1919 | 784.0678 | |
| 203.2226 | 282.1193 | 377.1453 | 450.1276 | 612.2075 | 785.0669 | |
| 204.0862 | 283.9713 | 378.0934 | 450.9689 | 615.9323 | 787.0815 | |

**Supplemental Table S3:** mzCloud Putative Identifications from *m/z* unique to AP-MALDI Control Nodules

| ***m/z*** | **Matrix** | **Compound** | **mzCloud Score** | **Theoretical *m/z*** | **Delta ppm** |
|---|---|---|---|---|---|
| 123.0552 | DHB | Nicotinamide | 89.3 | 123.0553 | -0.64 |
| 124.0391 | DHB | Nicotinic acid | 94.8 | 124.0393 | -1.33 |
| | | Picolinic acid | 85.6 | | |
| 127.0391 | DHB | Phloroglucinol | 87.1 | 127.0390 | 0.78 |
| | | 4-Hydroxy-6-methyl-2-pyrone | 85.5 | | |
| | | Maltol | 85 | | |
| | | 5-Hydroxymethyl-2-furaldehyde | 84.8 | | |
| | | Pyrogallol | 84.4 | | |
| 130.0499 | DHB | L-Pyroglutamic acid | 84.1 | 130.0499 | 0.15 |
| | | D-(+)-Pyroglutamic Acid | 83.3 | | |
| 132.0656 | DHB | trans-4-Hydroxy-L-proline | 87.5 | 132.0655 | 0.83 |
| | | cis-4-Hydroxy-D-proline | 85.9 | | |
| | | 3-Hydroxy-L-proline | 83.1 | | |
| 134.0448 | DHB | L-Aspartic acid | 90 | 134.0448 | 0.34 |
| | | D-(-)-Aspartic acid | 81.6 | | |
| 144.1020 | DHB | DL-Stachydrine | 91.6 | 144.1019 | 0.73 |
| | | 1-Aminocyclohexanecarboxylic acid | 82.5 | | |
| 175.1077 | DHB | N-Acetylornithine | 83.6 | 175.1077 | -0.28 |
| 190.1072 | DHB | DL-α-Aminosuberic acid | 92.4 | 190.1074 | -1.18 |
| 220.1176 | DHB | Pantothenic acid | 80.9 | 220.1179 | -1.49 |

**Supplemental Table S4**. Control *m/z* with AUC values from ROC test (AUC>0.75 is shown in bold)

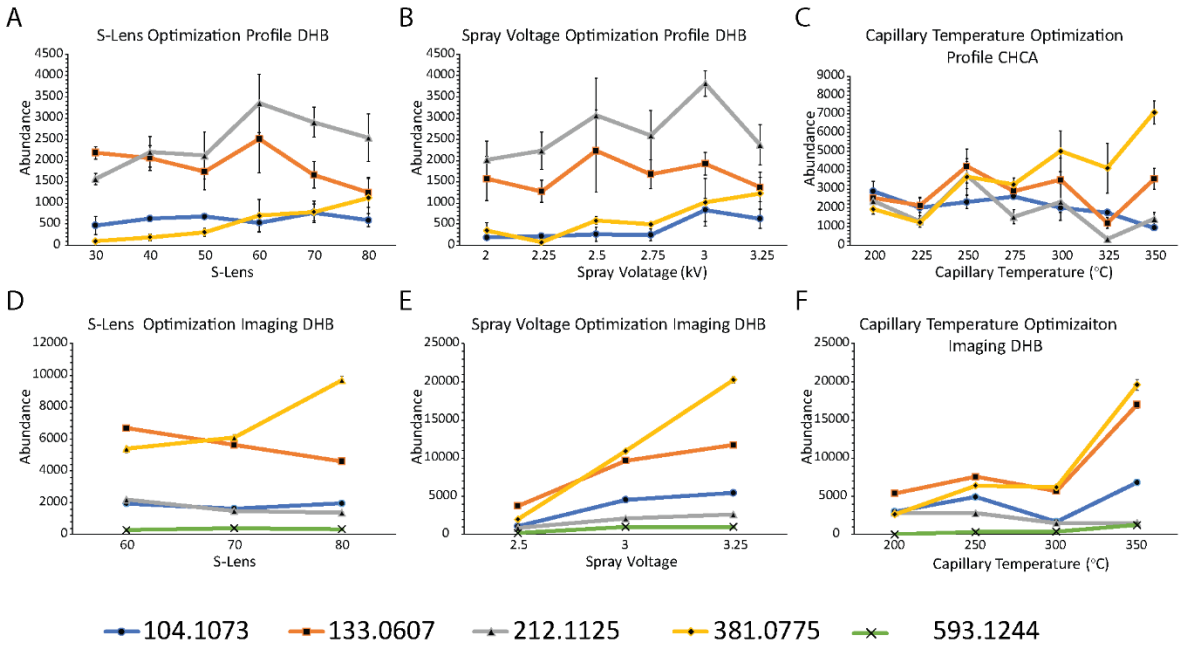
| ***m/z*** | **Matrix** | **Root and Nodule AUC** | **Nod AUC** | **Root AUC** |
|---|---|---|---|---|
| 130.0498 | CHCA | 0.6620 | **0.7589** | 0.5633 |
| | DHB | 0.7004 | **0.8089** | 0.5480 |
| 133.0606 | CHCA | **0.7658** | **0.8240** | **0.9364** |
| | DHB | **0.8136** | **0.8767** | **0.8125** |
| 147.0762 | CHCA | 0.7195 | **0.7878** | 0.6612 |
| | DHB | 0.6710 | **0.7730** | 0.5501 |
| 156.0417 | CHCA | 0.6602 | **0.7834** | 0.6662 |
| 203.1021 | CHCA | 0.6999 | **0.8201** | **0.8381** |
| | DHB | **0.7925** | **0.8774** | **0.7630** |
| 208.9722 | DHB | 0.7267 | **0.7541** | 0.6220 |
| | CHCA | **0.7618** | **0.9130** | 0.7079 |
| 210.1230 | CHCA | 0.5777 | 0.4565 | **0.7644** |
| 212.8419 | CHCA | **0.8253** | **0.8056** | **0.9193** |
| | DHB | 0.7372 | 0.6924 | **0.7599** |
| 217.1543 | CHCA | 0.6235 | 0.4708 | **0.8688** |
| 234.0186 | CHCA | **0.8857** | **0.8510** | **0.9278** |
| | DHB | 0.7279 | 0.6941 | **0.8285** |
| 234.1233 | CHCA | 0.7015 | **0.7614** | 0.6093 |
| 236.0167 | CHCA | 0.6551 | 0.5439 | **0.7878** |
| 241.0581 | CHCA | 0.7087 | **0.8401** | 0.6886 |
| 247.1645 | CHCA | 0.6302 | 0.4728 | **0.8907** |
| 265.9610 | CHCA | 0.7280 | **0.8497** | 0.7238 |
| 267.9593 | CHCA | 0.7157 | **0.8170** | 0.7139 |
| 268.1034 | CHCA | 0.6700 | 0.7295 | **0.7961** |
| 271.9749 | CHCA | **0.8285** | **0.8368** | **0.8982** |
| 273.9726 | CHCA | 0.6839 | 0.6240 | **0.7800** |
| 291.0834 | CHCA | 0.7100 | **0.7826** | 0.6110 |
| 304.1493 | CHCA | 0.6541 | **0.7793** | 0.7057 |
| 342.1054 | CHCA | 0.7248 | **0.8780** | 0.6874 |
| 367.0235 | CHCA | **0.7569** | **0.7888** | 0.6173 |
| 381.0785 | CHCA | **0.7961** | **0.8339** | **0.8269** |
| | DHB | **0.7688** | **0.7827** | 0.7149 |
| 383.0761 | CHCA | 0.6893 | 0.6784 | **0.7626** |
| 408.9971 | CHCA | **0.7999** | **0.8510** | 0.6438 |
| 413.0114 | CHCA | **0.7527** | **0.7939** | 0.6543 |
| 435.9701 | CHCA | 0.7161 | **0.8251** | 0.6023 |
| 452.9873 | CHCA | **0.7568** | **0.7930** | 0.6987 |
| 467.1552 | CHCA | 0.6906 | **0.7624** | 0.6511 |
| 475.1757 | CHCA | 0.7071 | **0.7537** | 0.6311 |
| | DHB | 0.7342 | **0.7924** | 0.7417 |
| 492.9589 | CHCA | **0.8393** | **0.9240** | **0.8058** |
| 494.9567 | CHCA | **0.7750** | **0.8359** | **0.7868** |
| 498.9730 | CHCA | 0.6822 | 0.6566 | **0.7741** |
| 511.1447 | CHCA | 0.7052 | 0.7286 | **0.7916** |
| 608.0769 | CHCA | 0.7005 | 0.7157 | **0.7983** |
| 638.0121 | CHCA | **0.7812** | **0.8324** | 0.7416 |
| 719.9575 | CHCA | **0.8165** | **0.8750** | **0.8125** |
| 797.1187 | CHCA | 0.7286 | 0.6792 | **0.7958** |
| 871.0416 | CHCA | 0.7025 | 0.6683 | **0.8045** |
| 946.9554 | CHCA | 0.7364 | **0.7511** | 0.7287 |

**Supplemental Table S5**. Salt Sample *m/z* and AUC values from ROC test (AUC>0.75 is shown in bold)

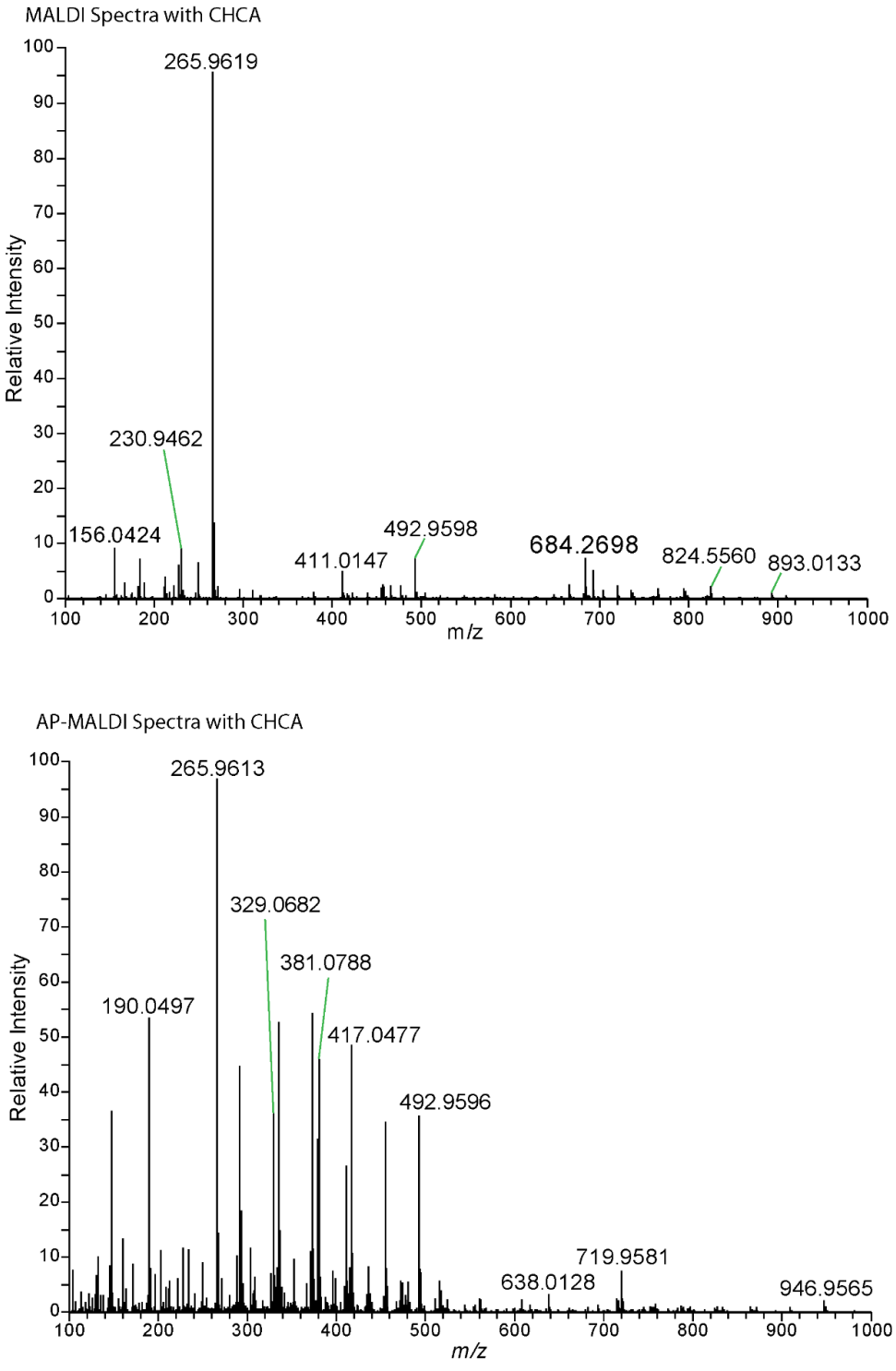
| ***m/z*** | **Matrix** | **Root and Nodule AUC** | **Nod AUC** | **Root AUC** |
|---|---|---|---|---|
| 100.1125 | CHCA | **0.9760** | **0.9971** | **0.9731** |
| 114.1279 | CHCA | **0.7705** | **0.8820** | 0.5946 |
| 116.0711 | DHB | 0.6468 | 0.5967 | **0.7509** |
| 120.0656 | CHCA | 0.6211 | 0.5876 | **0.7999** |
| | DHB | 0.7137 | 0.6559 | **0.7697** |
| 129.0658 | CHCA | 0.7272 | **0.8138** | 0.6004 |
| 130.0861 | CHCA | 0.7248 | **0.8752** | 0.5149 |
| 138.1024 | CHCA | 0.6926 | 0.7222 | **0.7550** |
| 140.0681 | CHCA | **0.9255** | **0.9674** | **0.8472** |
| | DHB | **0.7636** | 0.7303 | **0.7906** |
| 144.1016 | CHCA | **0.9615** | **0.9926** | **0.9826** |
| | DHB | **0.8495** | **0.7944** | **0.9200** |
| 146.1173 | CHCA | **0.9317** | **0.9808** | **0.9624** |
| | DHB | **0.9546** | **0.9807** | **0.9541** |
| 147.1128 | CHCA | 0.6716 | **0.8265** | 0.3990 |
| | DHB | **0.8220** | **0.7762** | **0.8069** |
| 155.0424 | CHCA | **0.8203** | **0.9546** | 0.5694 |
| 158.1174 | CHCA | **0.7639** | **0.8938** | 0.5917 |
| 162.1122 | CHCA | **0.8265** | **0.8836** | 0.7294 |
| 166.0834 | CHCA | **0.9936** | **0.9892** | **0.9911** |
| | DHB | **0.8504** | **0.8074** | **0.9236** |
| 175.1186 | CHCA | 0.7310 | **0.9085** | 0.4436 |
| 175.1439 | CHCA | 0.6729 | **0.7801** | 0.4987 |
| 177.0245 | CHCA | 0.7393 | **0.8484** | 0.5319 |
| 178.1335 | CHCA | **0.7837** | **0.8549** | **0.8081** |
| 182.0575 | CHCA | **0.9805** | **0.9928** | **0.9892** |
| | DHB | **0.8188** | **0.7591** | **0.8969** |
| 184.0558 | CHCA | **0.8960** | **0.9014** | **0.9062** |
| 184.0731 | CHCA | 0.7368 | 0.7470 | **0.7502** |
| 188.0551 | CHCA | 0.6059 | 0.5651 | **0.8055** |
| | DHB | 0.6374 | 0.5313 | **0.7636** |
| 189.1343 | CHCA | 0.7387 | **0.8670** | 0.5051 |
| 192.9984 | CHCA | **0.7815** | **0.9323** | 0.5397 |
| 197.1005 | CHCA | 0.6549 | **0.7785** | 0.5038 |
| 203.0523 | CHCA | **0.8095** | **0.7911** | **0.8591** |
| | DHB | **0.7615** | 0.6737 | **0.8432** |
| 203.1499 | CHCA | 0.6804 | **0.7982** | 0.5281 |
| 203.2227 | CHCA | 0.7199 | **0.7513** | 0.6783 |
| 205.0814 | CHCA | 0.7231 | **0.7811** | **0.8106** |
| | DHB | 0.6378 | 0.5202 | **0.7721** |
| 213.9639 | DHB | 0.7118 | 0.6421 | **0.7851** |
| 216.1128 | CHCA | 0.7138 | 0.6310 | **0.9471** |
| 217.0678 | CHCA | **0.8911** | **0.9738** | **0.7819** |
| | DHB | **0.8319** | **0.7829** | **0.8977** |
| 217.1907 | CHCA | 0.6594 | **0.7632** | 0.5115 |
| 225.0842 | CHCA | **0.8210** | **0.9612** | 0.5783 |
| 227.0634 | CHCA | **0.8271** | **0.8022** | **0.9853** |
| | DHB | **0.7622** | 0.6581 | **0.8910** |
| 229.0316 | CHCA | 0.6656 | **0.7677** | 0.5026 |
| 237.1018 | CHCA | 0.7110 | **0.7605** | 0.6427 |

| | | | | |
|---|---|---|---|---|
| 240.0270 | CHCA | **0.7703** | **0.7946** | **0.7580** |
| 243.0373 | CHCA | 0.6299 | 0.6202 | **0.8381** |
| | DHB | 0.6507 | 0.5341 | **0.7675** |
| 243.1486 | CHCA | **0.8065** | **0.8850** | 0.7113 |
| 245.1643 | CHCA | **0.9482** | **0.9888** | **0.9322** |
| 247.2010 | CHCA | **0.7651** | **0.8256** | 0.7098 |
| 249.0453 | DHB | 0.6308 | 0.5200 | **0.7528** |
| | CHCA | 0.6303 | 0.5538 | **0.8355** |
| 252.0025 | CHCA | **0.7774** | **0.7928** | **0.8145** |
| 252.1226 | DHB | **0.7569** | 0.7182 | **0.7974** |
| 254.1382 | DHB | **0.8045** | **0.7842** | **0.8538** |
| 258.1093 | CHCA | 0.6887 | **0.7868** | 0.5977 |
| 265.0191 | CHCA | 0.7377 | 0.6720 | **0.9718** |
| | DHB | 0.6810 | 0.5457 | **0.8356** |
| 269.0390 | CHCA | 0.7041 | **0.7926** | 0.6164 |
| 275.1097 | CHCA | 0.6852 | **0.7775** | 0.5268 |
| 280.1175 | DHB | **0.7767** | 0.7233 | **0.8407** |
| 280.9934 | CHCA | 0.5772 | 0.5141 | **0.7998** |
| 281.1490 | CHCA | **0.7611** | **0.7739** | **0.7526** |
| 283.1621 | CHCA | **0.7874** | **0.9054** | 0.6008 |
| 287.1383 | CHCA | **0.8025** | **0.8765** | 0.7014 |
| 289.1539 | CHCA | **0.9697** | **0.9959** | **0.9817** |
| 291.1695 | CHCA | **0.9047** | **0.9600** | **0.8386** |
| 298.1280 | DHB | **0.8764** | **0.8663** | **0.9165** |
| 309.1777 | CHCA | **0.8643** | **0.9366** | **0.7755** |
| 326.1314 | CHCA | **0.9047** | **0.9834** | **0.7846** |
| | DHB | 0.7463 | 0.7332 | **0.7741** |
| 333.1437 | CHCA | **0.9297** | **0.9683** | **0.9362** |
| 335.1592 | CHCA | **0.8789** | **0.9431** | **0.7921** |
| 341.0977 | DHB | 0.6174 | 0.5056 | **0.7580** |
| 350.9867 | DHB | **0.8010** | 0.7287 | **0.8894** |
| 363.0894 | DHB | 0.6808 | 0.6157 | **0.7591** |
| 365.1045 | CHCA | **0.9677** | **0.9996** | **0.9009** |
| | DHB | **0.8485** | **0.7890** | **0.9167** |
| 404.0224 | CHCA | 0.6581 | **0.7548** | 0.5255 |
| 408.0923 | DHB | 0.6259 | 0.5134 | **0.7692** |
| 434.2067 | CHCA | **0.7664** | **0.8212** | 0.6312 |
| 455.1151 | CHCA | **0.8399** | **0.7996** | **0.9717** |
| | DHB | **0.7546** | 0.6837 | **0.8792** |
| 463.0089 | CHCA | **0.7660** | 0.7430 | **0.8099** |
| 576.1295 | CHCA | **0.7570** | **0.8387** | 0.7213 |
| 737.1545 | CHCA | 0.7176 | **0.7817** | 0.6549 |
| 781.1445 | CHCA | **0.7959** | **0.8534** | **0.7611** |
| 849.4227 | CHCA | 0.6876 | 0.6065 | **0.8483** |
| | DHB | 0.6726 | 0.5646 | **0.8060** |
| 965.5076 | DHB | 0.6811 | 0.6102 | **0.7778** |

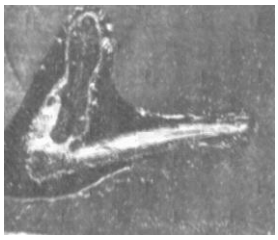
## Supplemental Figures



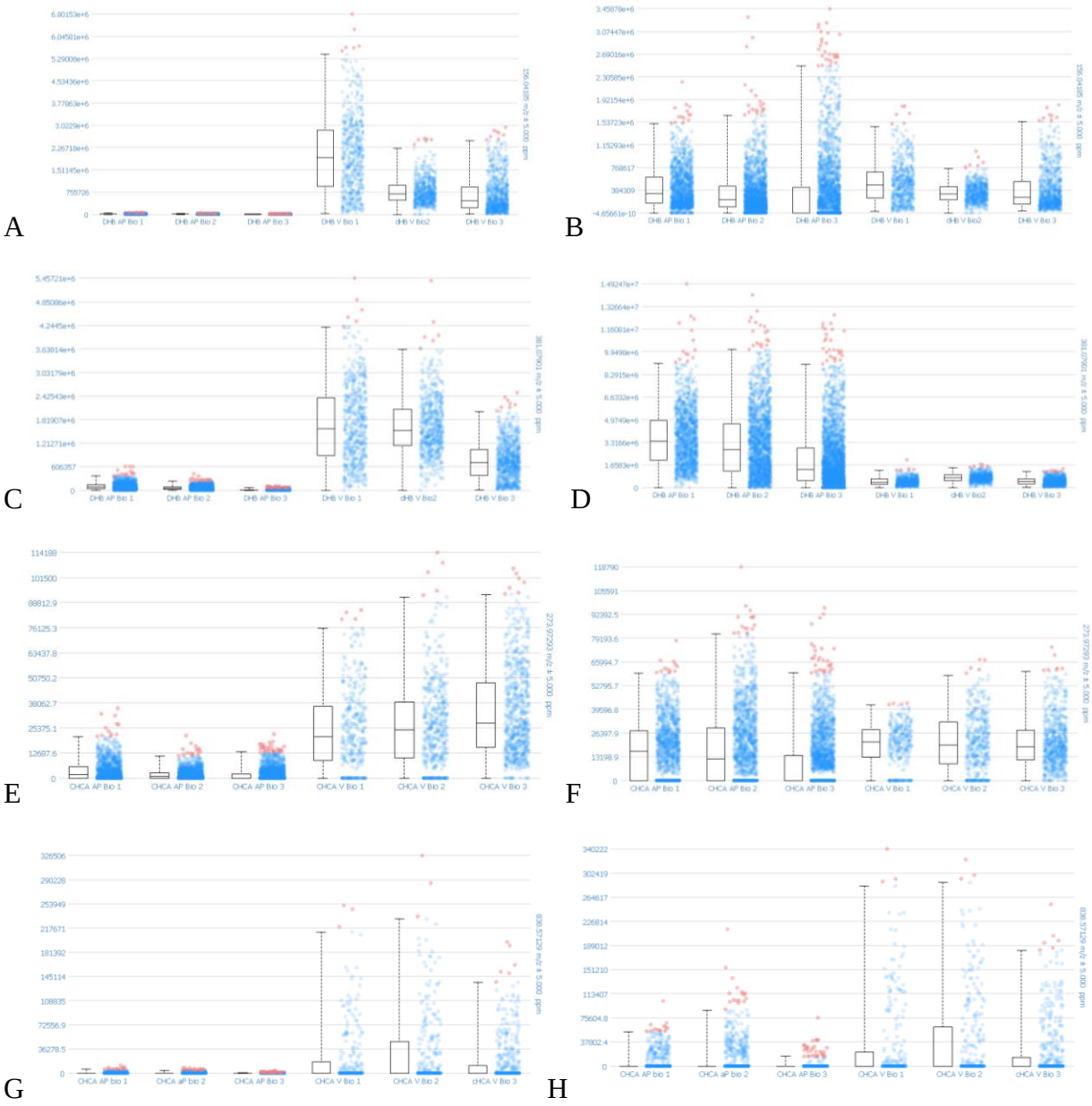

**Supplemental Figure S1**. AP-MALDI and QE-HF Parameter Optimization graphs for DHB as the matrix. **(A-C)** show optimization of S-lens **(A)**, spray voltage **(B)**, and capillary temperature **(C)** by profiling on tissue. **(D-F)** show optimization of S-lens **(D)**, spray voltage **(E)**, and capillary temperature **(F)** by imaging individual control root nodules.



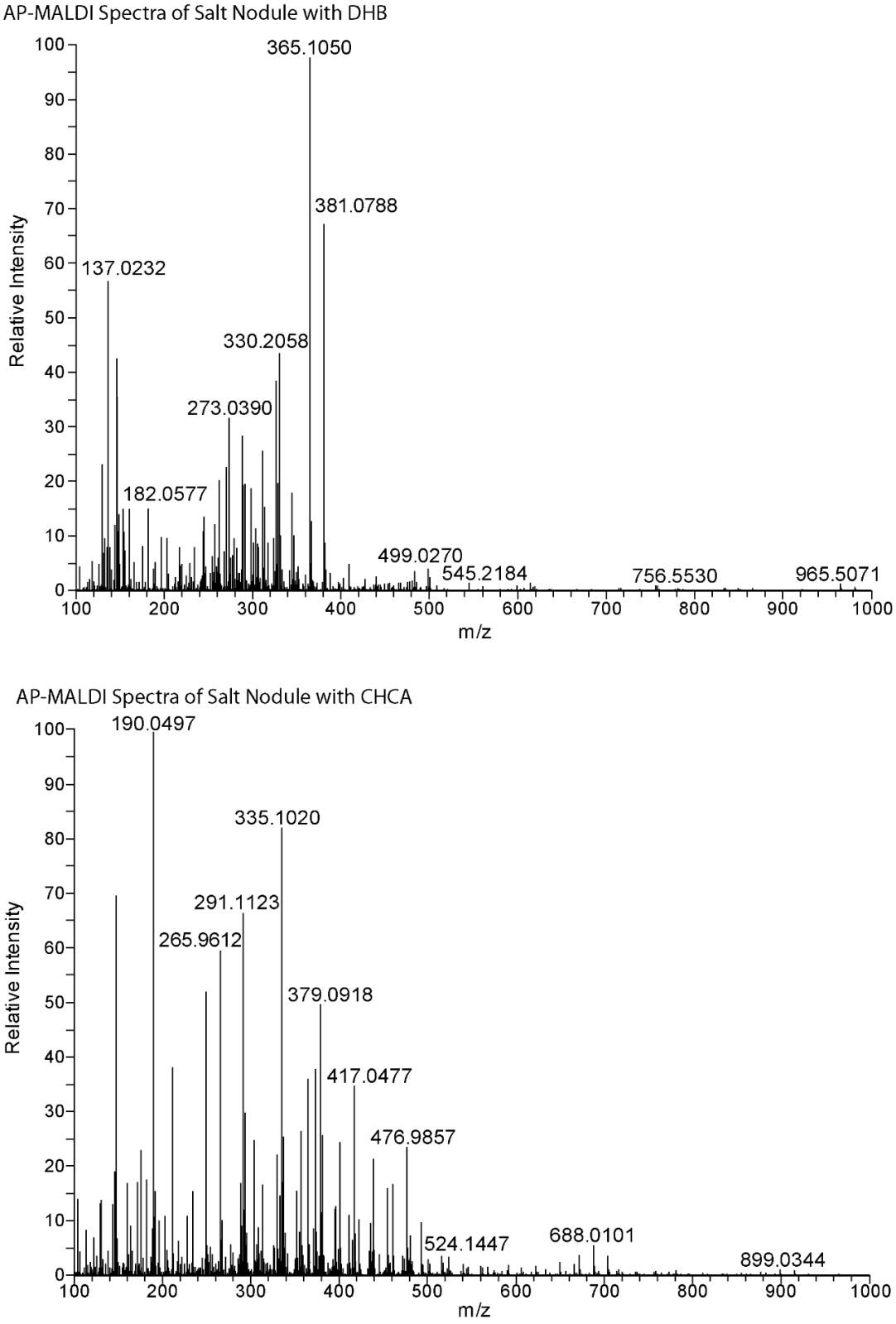

**Supplemental Figure S2.** Example spectra with CHCA as the matrix for control root nodules imaged with the MALDI platform and AP-MALDI platform. Spectra were averaged over the entire root nodule.

A  B 

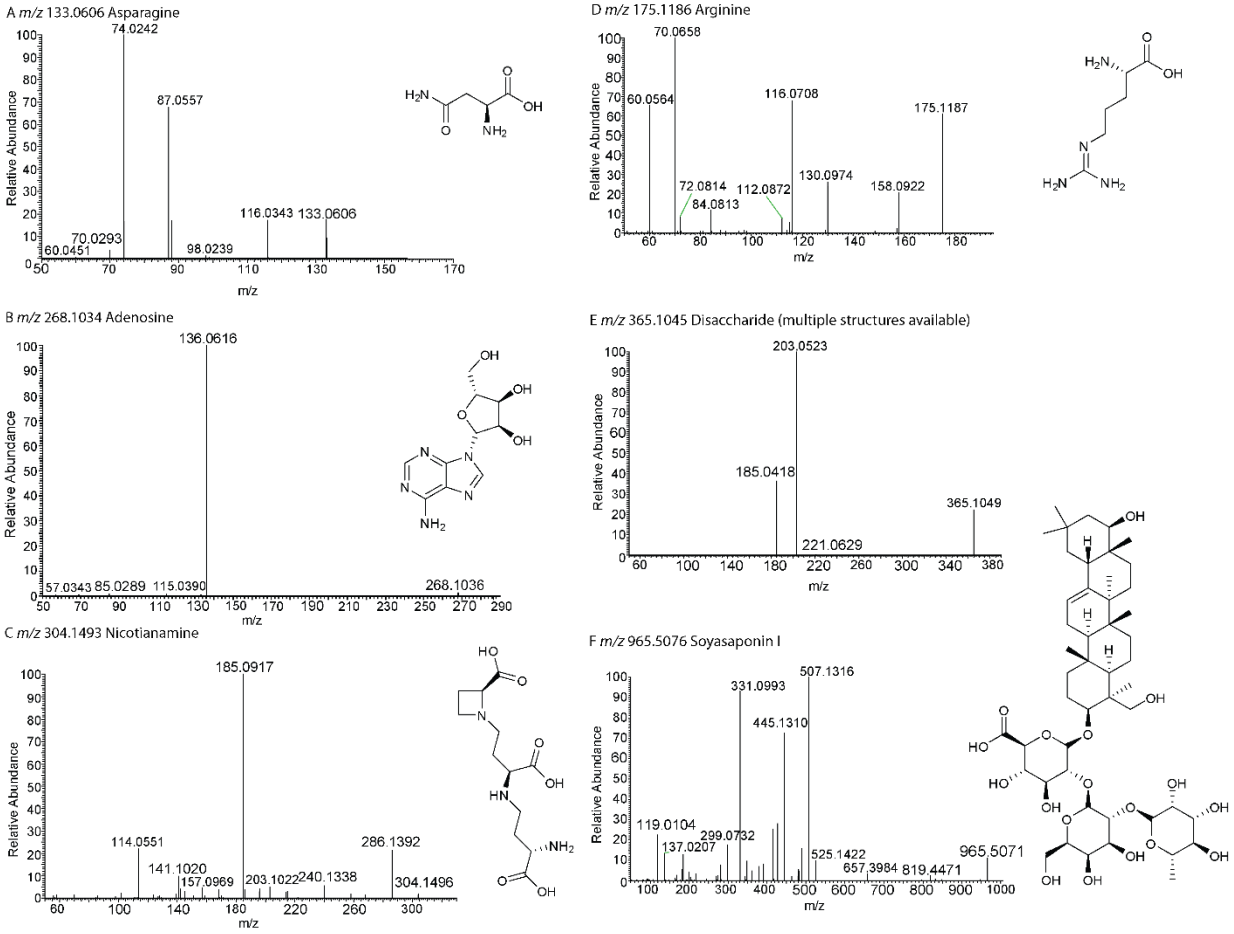
**Supplemental Figure S3.** Optical Images of Nodules from Figure 2. In **(A)**, the optical image corresponds to MALDI DHB data. In **(B)**, the optical image corresponds to AP-MALDI data, and MALDI CHCA data (serial sections of one control nodules were taken for these runs).



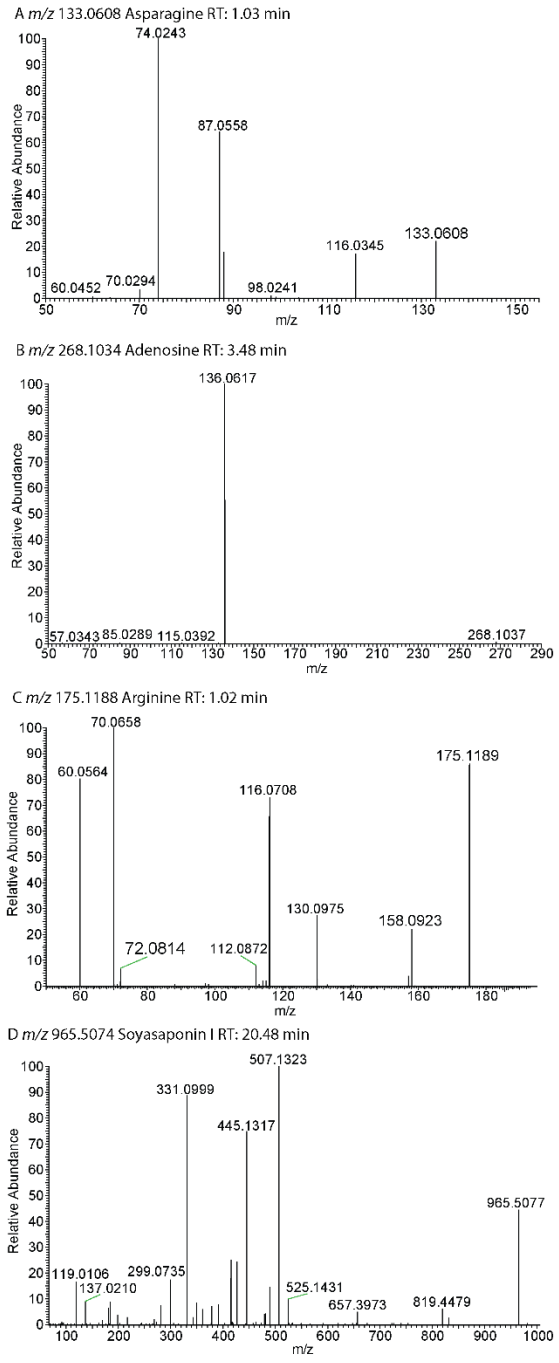

**Supplemental Figure S4.** Box and Whisker Plots comparing signal between the AP-MALDI and MALDI data. The plots show all three biological replicates for both instruments. The unnormalized data is shown in **(A,C,E,G)** and the normalized data is shown in **(B,D,F,H)**. The same *m/z* from **Figure 3** are shown. V indicates MALDI data and AP indicates AP-MALDI data.




**Supplemental Figure S5.** Example spectra for salt root nodules with DHB and CHCA as the matrix. Spectra are averaged over the root nodule.




**Supplemental Figure S6**. MS/MS spectra used for identifications. **(A-C)** show identifications from the control root nodules in **Table 1** and **(D-F)** show identifications from the salt root nodules in **Table 2.**




**Supplemental Figure S7.** MS/MS of standards for selected identifications. **(A)** and **(B)** are for identifications from control root nodules (asparagine and adenosine respectively) and **(C)** and **(D)** are for identifications from salt root nodules (arginine and soyasaponin I respectively).

# Chapter 4

# Examination of Endogenous Peptides in *Medicago truncatula* using Mass Spectrometry Imaging

Adapted from: Gemperline E, Keller C, Jayaraman D, Maeda J, Sussman MR, Ane JM, Li L. Examination of Endogenous Peptides in Medicago truncatula Using Mass Spectrometry Imaging. *J Proteome Res* **2016,** *15* (12), 4403-4411.

Keywords: MALDI-MSI; peptidomics; *Medicago truncatula*; plant science; Orbitrap; proteins

**Abstract**

Plant science is an important, rapidly developing area of study. Within plant science, one area of study that has grown tremendously with recent technological advances, such as mass spectrometry, is the field of plant-omics; however, plant peptidomics is relatively underdeveloped in comparison with proteomics and metabolomics. Endogenous plant peptides can act as signaling molecules and have been shown to affect cell division, development, nodulation, reproduction, symbiotic associations, and defense reactions. There is a growing need to uncover the role of endogenous peptides on a molecular level. Mass spectrometric imaging (MSI) is a valuable tool for biological analyses as it allows for the detection of thousands of analytes in a single experiment and also displays spatial information for the detected analytes. Despite the prediction of a large number of plant peptides, their detection and imaging with spatial localization and chemical specificity is currently lacking. Here we analyzed the endogenous peptides and proteins in *Medicago truncatula* using matrix-assisted laser desorption/ionization (MALDI)−MSI. Hundreds of endogenous peptides and protein fragments were imaged, with interesting peptide spatial distribution changes observed between plants in different developmental stages.

**Introduction**

Plant sciences play a significant role in mitigating three main challenges facing humanity in the 21st century, viz. food, energy, and environment (including pollution and climate change).[1] These challenges are intricately linked to each other. For instance, climate change affects crop yield, which consequently affects food and energy supply. Besides food, plants are also used for the production of therapeutic and antimicrobial products that are used to treat various human diseases such as cancer, Alzheimer's disease, and high cholesterol.[2-6] Plants are also used for

biofuel production and thereby act as "$CO_2$ mitigators" due to the lower carbon footprint of plant-derived energy compared with energy produced from petroleum or natural gas.[2] While significant progress has been made in the field of plant sciences, there remains a huge potential for improvement, which needs to be achieved to feed the burgeoning human population in the face of dwindling availability of arable land and water. For instance, despite the availability of close to 400,000 species of flowering plants, only a small fraction, about 200 species, have been domesticated for food and feed purposes. Among these, only 12 species contribute >75% of the food consumed across the world.[2, 7] Furthermore, with the vagaries of climate, plants are subjected to new or increased incidence of biotic and abiotic stresses. While it is true that the traditional method of plant genetics has led to most of the crop varieties we use today, relying on these traditional techniques alone will not satisfy our future needs for food, energy, and a stable environment. To produce better crops for the future, it is imperative that we understand the molecular processes that govern plant growth and development and their responses to the environment. The advent of "omics" technologies has facilitated our understanding of the molecular underpinnings of these complex traits like never before.

Among the omics technologies, the field of plant proteomics is rapidly growing and focuses on the study of proteins and enzymes expressed within various plant tissues. Limitations of more traditional gel electrophoresis-based proteomic methods include difficulty in analyzing highly basic or acidic proteins, bias toward more abundant proteins, and limited dynamic range. These limitations make low abundance proteins difficult to detect.[8, 9] Mass spectrometry is an advantageous technique for plant proteomic analyses due the higher sensitivity, selectivity, and structural determination capabilities of this technique.[10]

Branching off of proteomics, plant peptidomics, or the study of the endogenous peptides produced by a plant, is a relatively new and underdeveloped field.[11-14] As signaling molecules, plant peptides have been shown to affect cell division, development, nodulation, reproduction, symbiotic associations, and defense reactions.[14-18] Secreted peptides can act at low nanomolar concentrations, and the mature plant peptides are usually processed from larger polypeptides that undergo extensive proteolysis and posttranslational modifications (PTMs);[17] therefore, the discovery and identification of bioactive signaling peptides represent a significant analytical challenge.

Mass spectrometric imaging (MSI) is a valuable tool for biological analyses because it allows for molecular analysis of tissue while retaining information about the spatial distribution of different analytes found within the tissue.[19] In an MSI experiment, a laser is fired at the tissue sample in a predefined raster pattern, resulting in an array of mass spectra that can be compiled into a 2D distribution map for each mass measured. An advantage of MSI is that it lends itself to discovery experiments as it allows for the mass analysis of thousands of analytes in a single experiment and provides spatial information for the detected ions.

In the past decade, MSI has been increasingly utilized for the analysis of plant metabolomics.[19-25] However, as previously mentioned, plant peptidomics is a relatively under-explored area in mass spectrometry and especially mass spectrometry imaging. A literature search resulted in one report of matrix-assisted laser desorption/ionization (MALDI)-MSI being used for the analysis of cyclotides in petunias.[26] Additionally, while MS analysis is the gold standard for proteomic analysis, very few reports use MALDI−MSI for plant proteomics. Studies using MALDI −MSI for plant proteomics include a known allergenic protein in peaches shown to be exclusively found in the outer skin of the peach,[27] a lipid-transfer protein imaged in tomato seeds,[28] and proof-

of-principle images of proteins in soybean cotyledons or barley grains.[22, 29] So far, no further applications of MALDI−MSI to plant proteomics have been reported.

Here we present a study using MALDI−MSI to investigate endogenous peptides and proteins in the model legume, *Medicago truncatula* (Medicago). This legume forms a symbiotic association with rhizobial bacteria that are housed in plant-derived organs, the root nodules. Inside these nodules, rhizobia are enclosed in unique structures, symbiosomes. These symbiosomes provide a conducive environment for the oxygen-sensitive rhizobial nitrogenase enzyme, thereby enabling the fixation of atmospheric nitrogen into a plant-available form.[30, 31] Nodule formation requires substantial energy expenditure by the plants and is therefore under tight regulation by local and systemic endogenous signals.[32, 33] Nodule development is also affected by environmental factors and, in particular, when nitrogen is available in the soil. Several signaling peptides that play a critical role in nodulation have been identified. For instance, a number of CLAVATA/ESR-related (CLE) peptides such as MtCLE12 and 13 of Medicago, LjCLE-RS1 and 2 of *Lotus japonicas*, and GmRIC1 and 2 of soybean act systemically as negative regulators of nodulation.[33-37] These CLE peptides are induced upon rhizobial inoculation and are predicted to be the root-derived signals that, upon perception in the shoots, lead to the production of the shoot-derived inhibitor (SDI) that is transported to the roots, ultimately affecting nodulation. Recent evidence suggests that cytokinins are one of the shoot-derived inhibitors.[38] Besides CLE, peptides such as nodule-specific cysteine-rich (NCR) peptides, C-terminally encoded peptide (CEP), rapid alkalinization factor (RALF), and devil/rotundifolia (ROT)-four-like (DVL/RTFL) play roles in a wide range of functions including bacteroid differentiation, nodule development, and infection thread formation and progression.[39-42] Because of the importance of these peptides in legume symbiosis, we used MALDI−MSI to analyze 1 week old Medicago seedling roots and mature

Medicago roots and root nodules. MSI was performed using both of the most common MALDI matrices, 2,5-dihydroxybenzoic acid (DHB) and α-cyano-4-hydroxycinnamic acid (CHCA), in the mass range from *m/z* 900−4000.[43] In addition to studying the Medicago peptidome and proteome at different ages, wild-type plants were also compared with well-characterized Medicago mutants that are known to lack certain classes of peptides (*dnf*1−1) or overexpress certain classes of peptides (*35S*:*MtCLE13*). In addition to the MALDI−MSI experiments, parallel ESI−MS experiments were conducted to obtain accurate mass and high-quality tandem mass information for the *m/z* values detected via MSI. *De novo* sequencing was performed on the peptides detected with ESI using PEAKS software.[44] Hundreds of endogenous peptides and protein fragments were imaged and showed interesting spatial distribution differences between plants at different growth stages and between wild-type and the various mutants.

**Experimental Methods**

Plant Growth and Inoculation with Rhizobia

*Medicago truncatula* seeds of wild-type (cv. Jemalong A17), *dnf*1−1, and *35S*:*CLE13* were scarified with pure sulfuric acid for 8 min and sterilized with 8% (w/v) calcium hypochlorite solution for 2 min and germinated on agar supplemented with gibberellic acid. Freshly germinated, 1 day old seedlings were transferred to square plates containing nitrogen-free modified Fåhraeus medium,[45] after which the seedlings were grown in a growth chamber on modified Fåhraeus medium that was overlaid with germination paper. For nodule sampling, 10 seedlings were placed per plate, and the root part of the plate was covered with aluminum foil. The plates were placed vertically on a shelf at room temperature under fluorescent light. After 5 days of growth, the roots were inoculated with *Sinorhizobium meliloti* Rm1021 (wild-type)[46] at an $OD_{600}$ of 0.1 and then

returned to the light shelf. At 3 weeks postinoculation, nodules were separated from the roots and ground to a fine powder in liquid nitrogen. For seedling sampling, 100 1 day old seedlings were transferred to a plate containing nitrogen-free modified Fåhraeus medium. After 7 days, entire seedlings were collected, immediately frozen, and ground with liquid nitrogen.

Sample Preparation for MALDI

For the mature plants, root nodules with approximately 2 to 3 mm of surrounding root were excised from the plant. The excised tissue was placed in a plastic cup, covered with gelatin (100 mg/mL in double-distilled water), and frozen gently with dry ice. For the seedlings, 2 to 3 cm long portions of the root were cut from the seedling, embedded in gelatin, and frozen as detailed above. The frozen tissue was then sectioned into16 μm slices using a cryostat at −20 °C and thaw-mounted onto standard glass microscope slides. A TM Sprayer (HTX Technologies, Carrboro, NC) was used to apply MALDI-matrix to the samples. The TM Sprayer method for DHB (40 mg/mL DHB in 50:50 water: methanol) was as follows: 80 °C, 0.1 mL/min flow rate, 24 passes- rotate and offset, 3 mm spacing, and a velocity of 1250 mm/min. The TM Sprayer method for CHCA (10 mg/mL CHCA in 50:50 water/acetonitrile) was as follows: 90 °C, 0.2 mL/min flow rate, eight passes- rotate and offset, 3 mm spacing, and a velocity of 1100 mm/min. DHB and CHCA were purchased from Sigma-Aldrich (St. Louis, MO).

MALDI-Orbitrap MSI

A high-resolution, accurate mass (≤5 ppm error) MALDI-LTQ Orbitrap hybrid mass spectrometer (Thermo Scientific, Waltham, MA) was used for MSI of both mature nodules and 1 week old seedlings. Multiple technical replicates of three or more biological replicates were imaged in the positive ion mode using a mass range of *m/z* 900−4000 and a mass resolution of 60,000. Mass spectra were collected across the surface of the sample with a raster step size of 75

μm. Peptide images were extracted automatically using MSiReader.[47] In brief, the plant tissue was selected as the "region of interest" and the matrix was selected as the reference region. Masses from the matrix region were removed, and a list of *m/z* values detected in the plant samples was automatically generated. Ion images for each of the masses in the list were automatically extracted using the "generate an image for each peak in a list" function in MSiReader. Each extracted image was then manually confirmed as a true ion image, excluding isotope and matrix ion images.

Tissue Extraction

Approximately 50−100 root nodules with 2 to 3 mm of surrounding root were detached from mature Medicago plants or ∼50 1 week old Medicago seedlings were removed from the growth media and placed into a prechilled mortar. The tissue was flash-frozen with liquid nitrogen and ground to powder with the mortar and pestle. The powder was transferred to a prechilled 13 mL PTFE-coated centrifuge tube. The endogenous peptides were extracted with 3:1:4 methanol/chloroform: water (v/v), followed by brief vortexing and centrifugation for 10 min at 4 °C and 4700 rpm. The resulting aqueous supernatant was collected and dried in a SpeedVac. An additional four parts methanol were added to the remaining solution, followed by brief vortexing and centrifuging for 5 min at 4 °C and 4700 rpm. The organic layer was removed from the protein pellet, and both fractions were dried in a SpeedVac. The samples were stored at −80 °C until analysis.

Q-Exactive for ESI-MS

To acquire LC−ESI−MS/MS data, Medicago root nodule or seedling extracts were initially subjected to SCX fractionation on a Waters Alliance HPLC using a PolySulfethyl A column (2.1 mm internal diameter × 200 mm length, 5 μm particle size with 300 Å pore size; PolyLC, Columbia, MD). The mobile phases were (A) 10 mM ammonium formate in 75% water/25%

acetonitrile at pH ~6.8 and (B) 500 mM ammonium formate in 75% water/25% acetonitrile at pH 3. The samples were separated over 80 min under the following conditions: 0−15 min, 0% B; 15−45 min, 0−50% B; 45−55 min, 50−100% B; 55−65 min, 100% B; 65−65.5 min, 100−0% B, and finally re-equilibration at 0% B for 14.5 min. The column temperature was 30 °C, the flow rate was 0.2 mL/min, and the injection volume was 100 µL. Fractions were collected every 6 min between 10−70 min of the gradient. The fractions were combined and dried down three times with pure water to remove excess salts. Following SCX fractionation, the samples were resuspended in either water (aqueous fractions) or acetonitrile (organic and protein fractions) to a final concentration of 0.34 µg/µL (a total of 1.2 µg loaded onto the column). The samples were then separated on a NanoAcquity UPLC apparatus (Waters, Milford, MA) using a self-packed column (75 µm internal diameter ×160 mm length, 1.7 µm particle size with 130 Å pore size). The mobile phases were (A) water with 0.1% formic acid and (B) acetonitrile with 0.1% formic acid. The samples were separated over 108 min under the following conditions: 0−2 min, 0−4% B; 2−70 min, 4−30% B; 70−71 min, 30−75% B; 71−81 min, 75% B; 81−82 min, 75−95% B; 82−92 min, 75−95% B; 92−93 min, 95−0% B. The system was re-equilibrated at 0% B for 15 min. The flow rate was 0.3 mL/min, and the injection volume was 3.5 µL. The samples were kept at 4 °C during the analysis. MS/MS data were acquired in positive ion mode on an ESI-Q-Exactive Orbitrap mass spectrometer (Thermo Scientific). A top-15 data-dependent analysis (DDA) method was used with the full MS scan range from *m/z* 300−2000; an isolation window of 2.0; an intensity threshold of $5.0 \times 10^2$ for triggering MS2; exclusion of 1, 8, >8 charged species; a normalized collision energy of 30; a dynamic exclusion of 30 s; an MS1 resolution of 35,000; and an MS2 resolution of 17,500.

MS/MS spectra were *de novo* sequenced and matched to proteins using PEAKS software (Bioinformatics Solutions) with a parent mass tolerance of 20.0 ppm, fragment tolerance of 0.01

Da, no enzyme, and five variable post-translational modifications (amidation, oxidation (M), hydroxylation, arabinosylation, and acetylation (protein N-term)). Three variable PTMs were allowed per peptide, and spectra were matched against the NCBI Medicago database. The ALC cutoff score was set at 50% for *de novo* sequenced peptides, and an FDR of 0.1% was used for protein matches.

**Results and Discussion**

MALDI-Orbitrap MS Imaging

This study utilized wild-type Medicago and well-characterized Medicago mutants *dnf*1−1 and *35S*:*MtCLE13*.[48, 49] A complete summary of the putative peptide *m/z* values detected can be found in the Supporting Information (**Tables S1**, **S2**, **S3**, and **S4**); *m/z* values detected in multiple samples were averaged in the reported lists. These tables detail a list of *m/z* values detected in each replicate, a comparison of the peptides detected using CHCA compared with DHB, and a comparison of the peptides detected in the seedling roots compared with the mature roots and nodules.

CHCA and DHB were chosen as complementary matrices for MALDI−MSI. Photographs showing the difference between application of DHB and CHCA on root nodules can be found in **Figure S1**. There was a surprisingly small percentage of peptide masses detected using both DHB and CHCA, as shown in the Venn diagrams in **Figure 1**. A greater number of peptides was detected using CHCA as the matrix in comparison with DHB. Sinapic acid (SA), a matrix more often used to analyze higher molecular weight species such as peptides and proteins, was also used but did not show improved signal or detectability compared with CHCA. Representative images of putative peptides detected with CHCA or DHB matrices in the mature Medicago roots and root

nodules are shown in **Figure 2**. These representative images display ions with different spatial distributions in the plant root and nodules, which could provide further insight into the function of these putative peptides/proteins within the plant.

Putative peptides in Medicago roots were also compared in different stages of plant development. **Figure 3** presents representative putative peptide images showing distinct distribution patterns in the Medicago seedlings and the mature roots and root nodules. Some of the detected species show similar distribution patterns when we compare the seedlings to the mature plants (**Figure 3a**); however, other putative peptides seem to shift their localization from the root to the root nodules as they develop into older plants (**Figure 3b,c**). This implies that these putative peptides/proteins may play a role in nodule development or other nodule-related processes. **Figure 4** shows Venn diagrams comparing the numbers of putative peptides detected in the seedlings compared to the mature plants. Interestingly, a greater number of peptides was detected in the young seedling plants compared with the mature plants, regardless of the MALDI matrix used for ionization, and there was little overlap between the peptides detected in the plants in either stage of development.

In addition to comparing wild-type Medicago roots/root nodules in two different stages of development, wild-type plants were also compared to two different mutants. In the first mutant line, *35S*:*MtCLE13*, the CLAVATA3/endosperm-surrounding region (CLE) family of peptides is overexpressed; therefore, this mutant was thought to be a good model for MSI method development to diminish the challenge of trying to detect low-concentration peptides. **Figure 5** displays putative peptides that were detected in the *35S*:*MtCLE13* plants but not in the wild-type plants. It is thought that these putative peptides could belong to the CLE peptide family but are too low in concentration to be detected via MALDI−MSI in the wild-type plants. We also compared

the wild-type plants to *dnf*1−1 mutants, which develop stunted, nonfunctional nodules. **Figure 6** shows putative peptides that were detected in the wild-type plants but were absent from the *dnf*1−1 plants. This suggests that these putative peptides may play a role in nodule development or function.

Peptide/Protein Identification

Peptides and proteins can sometimes be identified by accurate-mass-matching; however, the more widely accepted approach for identification is to match MS/MS data to sequenced genomes or by *de novo* sequencing with software packages like PEAKS.[44] Furthermore, MALDI generally produces poor MS/MS fragmentation due to limited sample amounts and the fact that typically only singly-charged ions are generated. Therefore, MALDI−MSI results were matched to LC−MS/MS results for identifications as described below. **Table 1** details a list of *m/z* values detected via MALDI−MSI and LC−MS/MS that were able to be *de novo* sequenced using PEAKS. Since MALDI−MSI typically generates +1 charged ions and LC−MS typically generates +2 or +3 charged peptide ions, the molecular weight of each peptide detected was calculated from the acquired *m/z* to compare data across ionization methods. Using the molecular weights, the PEAKS *de novo* sequencing data generated from LC−MS/MS experiments was searched for peptide masses detected via MALDI−MSI. Since these calculations were made and multiple ionization sources were used, a mass error of <10 ppm was allowed for confident peptide identification via *de novo* sequencing. The annotated MS/MS spectra used for *de novo* sequencing can be found in the Supporting Information (**Figure S2**).

In addition to de novo sequencing, PEAKS also allows for MS/MS matching to sequenced genomes. PEAKS was used to generate a mass list from the LC−MS/MS data of unique peptide masses that matched proteins from the genome data within a 0.1% false discovery rate. A .fasta

file of the complete *Medicago truncatula* proteome was generated using the NCBI NR Database and used for peptide identification. Using this method, 10 imaged peptides were identified. Example images and corresponding LC−MS/MS spectra for these unique peptides are shown in **Figure 7** (all additional unique peptides with corresponding LC−MS/MS spectra are shown in the Supporting Information, **Figure S3**). Peptides corresponding to the ferritin, an iron storage protein, were identified. Iron is an essential component of nitrogenase, the bacterial enzyme that converts atmospheric dinitrogen to ammonia.[50] Depending upon the stage of nodule development the concentration and distribution of iron vary. In addition to accumulating in infected cells, ferritin also accumulates in uninfected cells within nodules. This increased concentration of ferritin may facilitate iron incorporation into nitrogenase.[51, 52] Besides ferritin, we also detected aquaporins, which are the predominant members of major intrinsic proteins (MIPs) commonly implicated in the transport of water, glycerol, and ammonia.[53] These plant aquaporins are divided into five subfamilies, including the tonoplast intrinsic proteins (TIPs). At least seven different TIPs have been identified in Medicago and, except for *TIP1g*, TIPs are expressed at low levels in 14 day old nodules.[54] The exception, *TIP1g*, is localized to the tonoplast in the infection zone (the zone where rhizobia are released into nodule cells) of nodules, whereas in the nitrogen-fixation zone *TIP1g* is redirected toward the symbiosome membrane. Knocking down the expression of *TIP1g* affected symbiosome maturation to the nitrogen-fixing stage, and it is hypothesized that this may be the result of altered water availability.[54] Further analysis will be necessary to identify the TIP we detected here, but based on the expression pattern of Medicago TIPs, it is highly likely that it is *TIP1g*. In addition to these peptides, a wound-induced basic family protein was identified. Although the function of this protein is still unknown, it should be noted that this family of proteins

is known to be upregulated in soybean nodules.[55] A complete list of the imaged unique peptides is shown in **Table 2**.

## Conclusions and Future Directions

We have demonstrated the benefits of using MALDI and ESI for the complementary detection and identification of endogenous peptides and protein fragments in plants. Using both CHCA and DHB as MALDI matrices for MSI greatly increased the coverage of peptides/protein fragments that were detected. We noticed interesting differences in the overall numbers of peptides detected between seedlings and mature plants. In addition to the difference in overall peptides, we also observed changes in the spatial distributions of peptides detected in both the seedlings and mature plants. Because of the low concentrations of endogenous peptides in wild-type plant tissues, peptide enrichment strategies might be a valuable next step for targeting and detecting specific classes of peptides in a biologically relevant manner. Other sample preparation approaches are in development for reducing the number of protein fragments that are detected and improving the detection of endogenous peptides. Additional biological studies examining the peptides/protein fragments detected in this study could reveal more insights into the functions of these peptides within the plant in different stages of development.

## Acknowledgments

This work was supported in part by funding from the National Science Foundation (NSF) Division of Integrative Organismal Systems (IOS) RESEARCH PGR award #1546742, University of Wisconsin- Madison Graduate School and the Wisconsin Alumni Research Foundation (WARF), a Romnes Faculty Research Fellowship, and a Vilas Distinguished Achievement

Professorship to L.L., and an NSF grant to JMA (NSF#0701846). E.G. acknowledges an NSF Graduate Research Fellowship (DGE-1256259). The MALDI-Orbitrap and Q-Exactive instruments were purchased through an NIH shared instrument grant

(NCRR S10RR029531).

**References**

1. Ehrhardt, D. W.; Frommer, W. B., New technologies for 21st century plant science. *Plant Cell* **2012,** *24* (2), 374-94.

2. Moshelion, M.; Altman, A., Current challenges and future perspectives of plant and agricultural biotechnology. *Trends Biotechnol* **2015,** *33* (6), 337-42.

3. Dey, A.; Mukherjee, A., Therapeutic potential of bryophytes and derived compounds against cancer. *Journal of Acute Disease* **2015,** *4* (3), 236-248.

4. Perumal Samy, R.; Gopalakrishnakone, P., Therapeutic Potential of Plants as Anti-microbials for Drug Discovery. *Evid Based Complement Alternat Med* **2010,** *7* (3), 283-94.

5. Glickman-Simon, R.; Schneider, C., Homeopathy for Depression, Music for Postoperative Recovery, Red Yeast Rice for High Cholesterol, Acupuncture for Seasonal Allergic Rhinitis, and Ginger for Osteoarthritis. *Explore (NY)* **2016,** *12* (4), 287-91.

6. Mikulass, K. R.; Nagy, K.; Bogos, B.; Szegletes, Z.; Kovacs, E.; Farkas, A.; Varo, G.; Kondorosi, E.; Kereszt, A., Antimicrobial nodule-specific cysteine-rich peptides disturb the integrity of bacterial outer and inner membranes and cause loss of membrane potential. *Ann Clin Microbiol Antimicrob* **2016,** *15* (1), 43.

7. Food and Agriculture Organization of the United Nations Statistics Division.

8. Roe, M. R.; Griffin, T. J., Gel-free mass spectrometry-based high throughput proteomics: tools for studying biological response of proteins and proteomes. *Proteomics* **2006,** *6* (17), 4678-87.

9. Zhang, Y. Y.; Fonslow, B. R.; Shan, B.; Baek, M. C.; Yates, J. R., Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical Reviews* **2013,** *113* (4), 2343-2394.

10. Jorrin-Novo, J. V.; Pascual, J.; Sanchez-Lucas, R.; Romero-Rodriguez, M. C.; Rodriguez-Ortega, M. J.; Lenz, C.; Valledor, L., Fourteen years of plant proteomics reflected in Proteomics: Moving from model species and 2DE-based approaches to orphan species and gel-free platforms. *Proteomics* **2015,** *15* (5-6), 1089-1112.

11. Verhaert, P.; Uttenweiler-Joseph, S.; de Vries, M.; Loboda, A.; Ens, W.; Standing, K. G., Matrix-assisted laser desorption/ionization quadrupole time-of-flight mass spectrometry: an elegant tool for peptidomics. *Proteomics* **2001,** *1* (1), 118-31.

12. Buchberger, A.; Yu, Q.; Li, L., Advances in Mass Spectrometric Tools for Probing Neuropeptides. *Annu Rev Anal Chem (Palo Alto Calif)* **2015**.

13. Dallas, D. C.; Guerrero, A.; Parker, E. A.; Robinson, R. C.; Gan, J. N.; German, J. B.; Barile, D.; Lebrilla, C. B., Current peptidomics: Applications, purification, identification, quantification, and functional analysis. *Proteomics* **2015,** *15* (5-6), 1026-1038.

14. Matsubayashi, Y.; Sakagami, Y., Peptide hormones in plants. In *Annual Review of Plant Biology*, Annual Reviews: Palo Alto, 2006; Vol. 57, pp 649-674.

15. Lindsey, K., Plant peptide hormones: The long and the short of it. *Current Biology* **2001,** *11* (18), R741-R743.

16. Kondo, T.; Sawa, S.; Kinoshita, A.; Mizuno, S.; Kakimoto, T.; Fukuda, H.; Sakagami, Y., A plant peptide encoded by CLV3 identified by in situ MALDI-TOF MS analysis. In *Science*, United States, 2006; Vol. 313, pp 845-8.

17. Farrokhi, N.; Whitelegge, J. P.; Brusslan, J. A., Plant peptides and peptidomics. *Plant Biotechnology Journal* **2008,** *6* (2), 105-134.

18. Yamaguchi, Y.; Huffaker, A., Endogenous peptide elicitors in higher plants. *Current Opinion in Plant Biology* **2011,** *14* (4), 351-357.

19. Balluff, B.; Schone, C.; Hofler, H.; Walch, A., MALDI imaging mass spectrometry for direct tissue analysis: technological advancements and recent applications. *Histochemistry and Cell Biology* **2011,** *136* (3), 227-44.

20. Gemperline, E.; Jayaraman, D.; Maeda, J.; Ane, J. M.; Li, L., Multifaceted investigation of metabolites during nitrogen fixation in Medicago via high resolution MALDI-MS imaging and ESI-MS. *J Am Soc Mass Spectrom* **2015,** *26* (1), 149-58.

21. Lee, Y. J.; Perdian, D. C.; Song, Z. H.; Yeung, E. S.; Nikolau, B. J., Use of mass spectrometry for imaging metabolites in plants. *Plant Journal* **2012,** *70* (1), 81–95.

22. Kaspar, S.; Peukert, M.; Svatos, A.; Matros, A.; Mock, H. P., MALDI-imaging mass spectrometry - An emerging technique in plant biology. *Proteomics* **2011,** *11* (9), 1840-50.

23. Gemperline, E.; Li, L., MALDI-mass spectrometric imaging for the investigation of metabolites in Medicago truncatula root nodules. *J Vis Exp* **2014,** (85).

24. Ye, H.; Gemperline, E.; Venkateshwaran, M.; Chen, R.; Delaux, P. M.; Howes-Podoll, M.; Ane, J. M.; Li, L., MALDI mass spectrometry-assisted molecular imaging of metabolites during nitrogen fixation in the *Medicago truncatula-Sinorhizobium meliloti* symbiosis. *Plant J.* **2013,** *75* (1), 130–45.

25. Bjarnholt, N.; Li, B.; D'Alvise, J.; Janfelt, C., Mass spectrometry imaging of plant metabolites - principles and possibilities. *Nat Prod Rep* **2014,** *31* (6), 818–837.

26. Poth, A. G.; Mylne, J. S.; Grassl, J.; Lyons, R. E.; Millar, A. H.; Colgrave, M. L.; Craik, D. J., Cyclotides associate with leaf vasculature and are the products of a novel precursor in petunia (Solanaceae). *J Biol Chem* **2012,** *287* (32), 27033-46.

27. Cavatorta, V.; Sforza, S.; Mastrobuoni, G.; Pieraccini, G.; Francese, S.; Moneti, G.; Dossena, A.; Pastorello, E. A.; Marchelli, R., Unambiguous characterization and tissue localization of Pru P 3 peach allergen by electrospray mass spectrometry and MALDI imaging. *J Mass Spectrom* **2009,** *44* (6), 891-7.

28. Bencivenni, M.; Faccini, A.; Zecchi, R.; Boscaro, F.; Moneti, G.; Dossena, A.; Sforza, S., Electrospray MS and MALDI imaging show that non-specific lipid-transfer proteins (LTPs) in tomato are present as several isoforms and are concentrated in seeds. *J Mass Spectrom* **2014,** *49* (12), 1264-71.

29. Grassl, J.; Taylor, N. L.; Millar, A. H., Matrix-assisted laser desorption/ionisation mass spectrometry imaging and its development for plant protein imaging. *Plant Methods* **2011,** *7*, 11.

30. Mus, F.; Crook, M. B.; Garcia, K.; Costas, A. G.; Geddes, B. A.; Kouri, E. D.; Paramasivan, P.; Ryu, M. H.; Oldroyd, G. E. D.; Poole, P. S.; Udvardi, M. K.; Voigt, C. A.; Ane, J. M.; Peters, J. W., Symbiotic Nitrogen Fixation and the Challenges to Its Extension to Nonlegumes. *Applied and Environmental Microbiology* **2016,** *82* (13), 3698-3710.

31. Venkateshwaran, M.; Volkening, J. D.; Sussman, M. R.; Ane, J. M., Symbiosis and the social network of higher plants. *Curr Opin Plant Biol* **2013,** *16* (1), 118-27.

32. McKay, I. A.; Djordjevic, M. A., Production and Excretion of Nod Metabolites by Rhizobium leguminosarum bv. trifolii Are Disrupted by the Same Environmental Factors That Reduce Nodulation in the Field. *Appl Environ Microbiol* **1993,** *59* (10), 3385-92.

33. Reid, D. E.; Ferguson, B. J.; Gresshoff, P. M., Inoculation- and nitrate-induced CLE peptides of soybean control NARK-dependent nodule formation. *Mol Plant Microbe Interact* **2011,** *24* (5), 606-18.

34. Mortier, V.; Den Herder, G.; Whitford, R.; Van de Velde, W.; Rombauts, S.; D'Haeseleer, K.; Holsters, M.; Goormachtig, S., CLE Peptides Control Medicago truncatula Nodulation Locally and Systemically. *Plant Physiology* **2010,** *153* (1), 222-237.

35. Saur, I. M.; Oakes, M.; Djordjevic, M. A.; Imin, N., Crosstalk between the nodulation signaling pathway and the autoregulation of nodulation in Medicago truncatula. *New Phytol* **2011,** *190* (4), 865-74.

36. Okamoto, S.; Ohnishi, E.; Sato, S.; Takahashi, H.; Nakazono, M.; Tabata, S.; Kawaguchi, M., Nod factor/nitrate-induced CLE genes that drive HAR1-mediated systemic regulation of nodulation. *Plant Cell Physiol* **2009,** *50* (1), 67-77.

37. Lim, C. W.; Lee, Y. W.; Hwang, C. H., Soybean nodule-enhanced CLE peptides in roots act as signals in GmNARK-mediated nodulation suppression. *Plant Cell Physiol* **2011,** *52* (9), 1613-27.

38. Sasaki, T.; Suzaki, T.; Soyano, T.; Kojima, M.; Sakakibara, H.; Kawaguchi, M., Shoot-derived cytokinins systemically regulate root nodulation. *Nat Commun* **2014,** *5*, 4983.

39. Kondorosi, E.; Mergaert, P.; Kereszt, A., A paradigm for endosymbiotic life: cell differentiation of Rhizobium bacteria provoked by host plant factors. *Annu Rev Microbiol* **2013,** *67*, 611-28.

40. Bedinger, P. A.; Pearce, G.; Covey, P. A., RALFs: peptide regulators of plant growth. *Plant Signal Behav* **2010,** *5* (11), 1342-6.

41. Imin, N.; Mohd-Radzman, N. A.; Ogilvie, H. A.; Djordjevic, M. A., The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in Medicago truncatula. *Journal of Experimental Botany* **2013,** *64* (17), 5395-5409.

42. Valdivia, E. R.; Hertweck, K. L.; Cho, S. K.; Walker, J. C., DVL/RTFL. In *Handbook of Biologically Active Peptides*, 2 ed.; Elsevier Inc: 2013; pp 15-19.

43. Chen, R.; Cape, S. S.; Sturm, R. M.; Li, L., Mass spectrometric imaging of neuropeptides in decapod crustacean neuronal tissues. *Methods Mol Biol* **2010,** *656*, 451-63.

44. Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G., PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun Mass Spectrom* **2003,** *17* (20), 2337-42.

45. Catoira, R.; Galera, C.; de Billy, F.; Penmetsa, R.; Journet, E.; Maillet, F.; Rosenberg, C.; Cook, D.; Gough, C.; Denarie, J., Four genes of *Medicago truncatula* controlling components of a nod factor transduction pathway. *The Plant Cell* **2000,** *12* (9), 1647–1665.

46. Oke, V.; Long, S. R., Bacteroid formation in the rhizobium-legume symbiosis. *Curr Opin Microbiol* **1999,** *2* (6), 641–6.

47. Robichaud, G.; Garrard, K. P.; Barry, J. A.; Muddiman, D. C., MSiReader: An Open-Source Interface to View and Analyze High Resolving Power MS Imaging Files on Matlab Platform. *Journal of the American Society for Mass Spectrometry* **2013,** *24* (5), 718-721.

48. Mitra, R. M.; Long, S. R., Plant and bacterial symbiotic mutants define three transcriptionally distinct stages in the development of the *Medicago truncatula*/*Sinorhizobium meliloti* symbiosis. *Plant Physiology* **2004,** *134* (2), 595–604.

49. Wang, D.; Griffitts, J.; Starker, C.; Fedorova, E.; Limpens, E.; Ivanov, S.; Bisseling, T.; Long, S. R., A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* **2010,** *327* (5969), 1126–1129.

50. Brear, E. M.; Day, D. A.; Smith, P. M., Iron: an essential micronutrient for the legume-rhizobium symbiosis. *Front Plant Sci* **2013,** *4*, 359.

51. Lucas, M. M.; Sype, G. V. D.; Herouart, D.; Hernandez, M. J.; Puppo, A.; Felipe, M. R. D., Immunolocalization of ferritin in determinate and indeterminate legume root nodules. *Protoplasma* **1998,** *204* (1), 61-70.

52. Ragland, M.; Theil, E. C., Ferritin (mRNA, protein) and iron concentrations during soybean nodule development. *Plant Mol Biol* **1993,** *21* (3), 555-60.

53. Chaumont, F.; Tyerman, S. D., Aquaporins: highly regulated channels controlling plant water relations. *Plant Physiol* **2014,** *164* (4), 1600-18.

54. Gavrin, A.; Kaiser, B. N.; Geiger, D.; Tyerman, S. D.; Wen, Z.; Bisseling, T.; Fedorova, E. E., Adjustment of host cells for accommodation of symbiotic bacteria: vacuole defunctionalization, HOPS suppression, and TIP1g retargeting in Medicago. *Plant Cell* **2014,** *26* (9), 3809-22.

55. Li, X.; Xu, J.; Yu, G.; Luo, L., A wound-induced small polypeptide gene family is upregulated in soybean nodules. *Chinese Sci. Bull* **2013,** *58* (9), 1003-1009.

## Tables

**Table 1.** List of *m*/*z* values that were *de novo* sequenced using PEAKS

| **De Novo Sequenced Peptides (molecular weight)** | | | |
|---|---|---|---|
| **MALDI Measured** | **LC-MS Measured** | **Δ ppm** | **Peptide Sequence** |
| 908.4234 | 908.4240 | 0.66 | FGGSTVEVN |
| 962.4813 | 962.4743 | 7.31 | TVNEEKLM |
| 990.5163 | 990.5134 | 2.89 | PSPPLRGEP |
| | | | EPPPHVTSK |
| 1006.5114 | 1006.5195 | 8.09 | TVGKGAHAGPP |
| 1008.4694 | 1008.4624 | 6.98 | DDARPPQGGP |
| 1021.5575 | 1021.5556 | 1.85 | AVGKDYRLT |
| 1068.5733 | 1068.5791 | 5.42 | CKVWPLPGK |
| | | | CKVWPPLGK |
| 1094.5280 | 1094.5244 | 3.28 | PQTEAPAVGAP |
| 1106.5081 | 1106.4993 | 7.92 | YNDQDTPVR |
| 1106.5255 | 1106.5356 | 9.10 | TDSSAPGGFLR |
| 1112.5974 | 1112.5938 | 3.21 | STGGVAAPRAVQ |
| | | | TSGVQAPRGPK |
| 1159.6191 | 1159.6084 | 9.20 | TEAATATPAVTK |
| 1170.6070 | 1170.6133 | 5.38 | VLDPGDSDLLK |
| | | | LVDPGDSDLLK |
| | | | GPDDSDLVLLK |
| | | | DPGDSDLVLLK |
| | | | DPGSDDLVLLK |
| 1175.5925 | 1175.5935 | 0.88 | TGAEGKVHSYK |
| 1179.6039 | 1179.6135 | 8.10 | KAPPPVADDTK |
| 1186.5763 | 1186.5830 | 5.68 | TVGNGPVEASGLS |
| 1192.5108 | 1192.5107 | 0.09 | HGGTEDPVTSGH |
| 1300.6790 | 1300.6775 | 1.12 | QSSHSPVLVKGF |
| 1300.6790 | 1300.6697 | 7.12 | TVGAVDAVTLMPQ |
| 1308.6282 | 1308.6211 | 5.45 | SYFANAQPQQR |
| 1342.6713 | 1342.6663 | 3.72 | QSVKMTNAHSLQ |
| 1376.7577 | 1376.7598 | 1.55 | VSLALVCSPVPHR |
| 1728.9683 | 1728.9597 | 4.95 | KPLNVELGFKAVAAGLC |

**Table 2.** List of *m/z* values of unique peptides for which MS images were acquired

| [M+H] | Protein Group | Protein Accession | Peptide | Protein Description |
|---|---|---|---|---|
| 965.51694 | 47 | gi\|657381377, gi\|657381378 | C.TVIDAPGHR.D | GTP-binding elongation factor Tu family protein |
| | 47 | gi\|357496973 | C.TVIDAPGHR.D | Elongation factor 1-alpha |
| 1132.53084 | 215 | gi\|657377737 | K.ANENKPVMTE | Wound-inducible basic family protein |
| 1160.61634 | 43 | gi\|657399288 | A.ETAATATPAVTK.S | NADPH-protochlorophyllide oxidoreductase |
| 1162.59554 | 31 | gi\|357474991 | M.A(+42.01)ASGEEKKIST.S | Low-temperature inducible |
| | 31 | gi\|355508836 | M.A(+42.01)ASGEEKKIST.S | Nodulin-like protein |
| | 31 | gi\|388522163 | M.A(+42.01)ASGEEKKIST.S | Unknown |
| 1171.62114 | 71 | gi\|388504426, gi\|217073834, gi\|388496542 | S.IVDPGDSDIIK.T | Unknown |
| | 71 | gi\|657371310, gi\|355496648 | S.IVDPGDSDIIK.T | Ribosomal protein L7Ae/L30e?S12e?Gadd45 family protein |
| | 71 | gi\|357502689 | S.IVDPGDSDIIK.T | 60S ribosomal protein L30 |
| 1180.62144 | 148 | gi\|388522541, gi\|388492578 | K.AVAPPPVADDTK.A | Unknown |
| | 148 | gi\|657373135 | K.AVAPPPVADDTK.A | carboxy-terminal region remorin |
| 1293.67314 | 149 | gi\|388507838, gi\|217073043 | T.GVIFEPFEEVK.K | Unknown |
| | 149 | gi\|357468557 | T.GVIFEPFEEVK.K | Ferritin-3 |
| | 149 | gi\|355505618 | T.GVIFEPFEEVK.K | Ferritin |
| | 149 | gi\|657385619, gi\|355518020, gi\|355498374 | T.GVLFEPFEEVK.K | Ferritin |
| | 149 | gi\|388491178, gi\|217073544, gi\|388499902, gi\|217073522 | T.GVLFEPFEEVK.K | Unknown |
| | 149 | gi\|357492793 | T.GVLFEPFEEVK.K | Ferritin-2 |
| | 149 | gi\|357506141 | T.GVIFEPFEEVK.K | Ferritin-1 |
| 1337.67014 | 61 | gi\|657379753 | G.TPQEATHPDTLK.A | Tonoplast intrinsic protein |
| | 61 | gi\|32363409, gi\|9716259 | G.TPQEATHPDTLK.A | Probable aquaporin TIP-type |
| 1483.75454 | 13 | gi\|357471525, gi\|355507102, gi\|404332436, gi\|404332513, gi\|355501595, gi\|357512583, gi\|404332359, gi\|357502811, gi\|124360830, gi\|355496709 | A.AFRVSPQPGVPAEE.A | Ribulose biphosphate carboxylase large chain domain protein |
| | 13 | gi\|543174105, gi\|153012229 | A.AFRVSPQPGVPAEE.A | Ribulose- 1,5-biphosphate carboxylase/ oxygenase large subunit (chloroplast) |
| 1976.02434 | 11 | gi\|355501329, gi\|357512051 | N.NKNNPNLFNNLVYTPLT.I | Basic 7S globulin-like protein |
| | 11 | gi\|87240526 | N.NKNNPNLFNNLVYTPLT.I | Peptidase A1, pepsin |

**Figures**

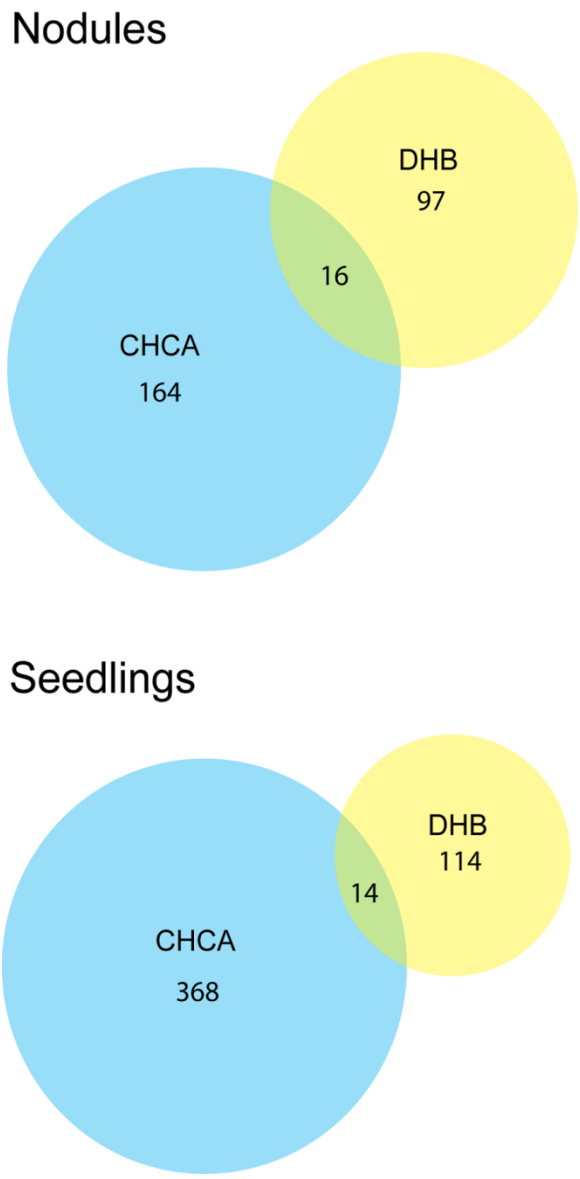



**Figure 1.** Comparison of the numbers of endogenous peptides detected using either CHCA or DHB as MALDI matrices in both mature Medicago roots/root nodules and seedling roots.

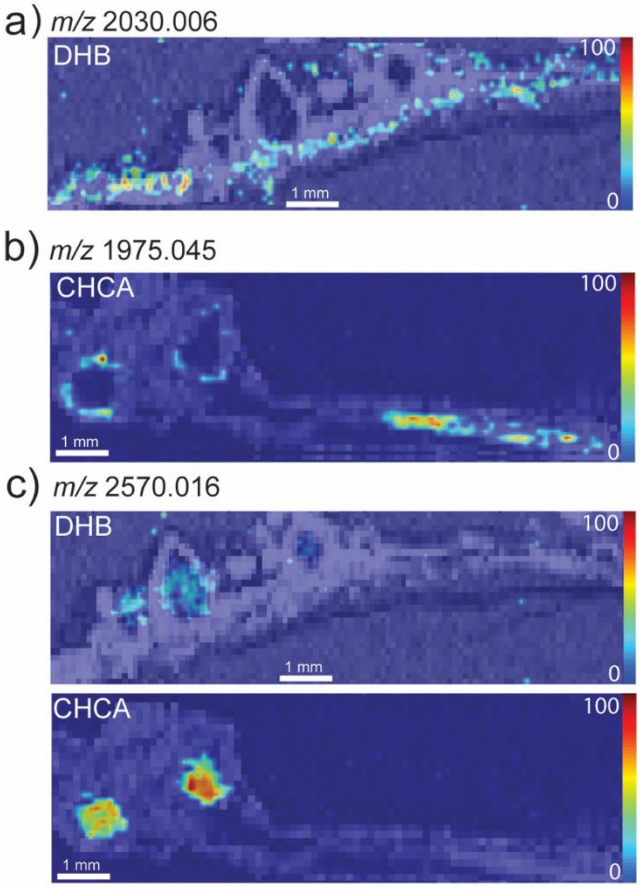



**Figure 2.** Representative images of putative peptides detected with CHCA and/or DHB matrices in mature Medicago roots and root nodules showing different spatial distributions throughout the root and nodule portions of the plant. a) *m*/*z* 2030.006 was only detected using DHB as the matrix and is localized to the plant root. b) *m*/*z* 1975.045 was only detected using CHCA as the matrix and is localized to the plant root and outer portion of the nodules. c) *m*/*z* 2570.016 was detected using both DHB and CHCA as the matrix and is localized to the nodules. Intensity scale = high (red) to low (blue).

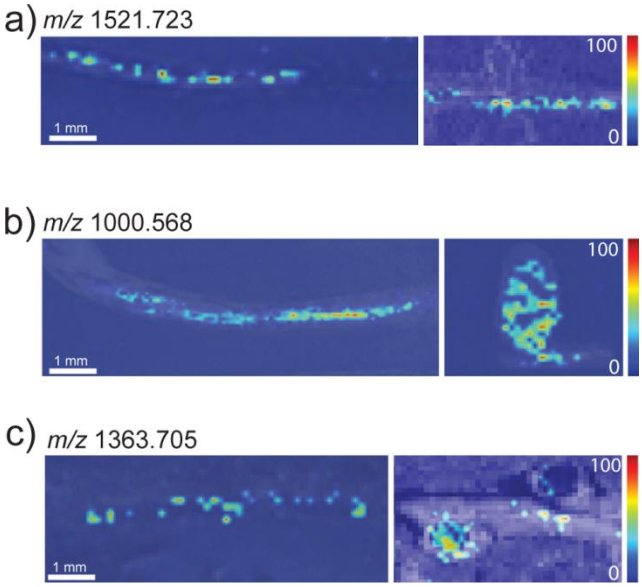



**Figure 3.** Representative putative peptide images showing distinct distribution patterns in the Medicago seedlings and the mature roots and root nodules. a) *m*/*z* 1521.723 shows a similar distribution pattern in both the seedlings and the mature plants. b) *m*/*z* 1000.568 and c) *m*/*z* 1363.705 represent two of the detected peptides that show a shift in their localization from the root to the root nodules as they develop into older plants.

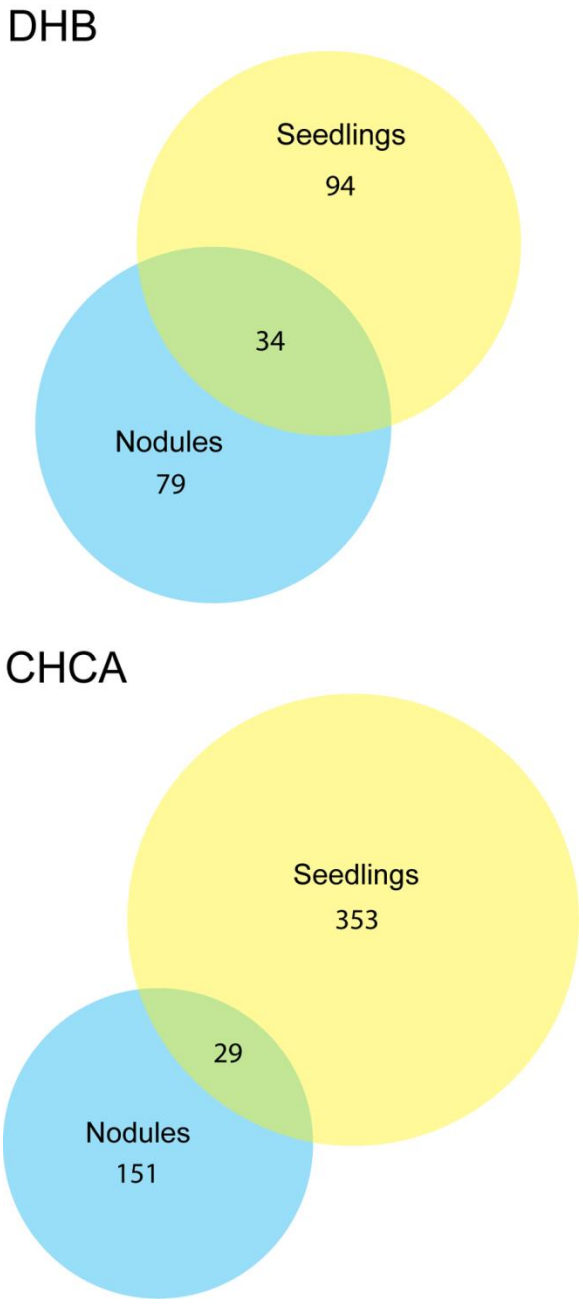



**Figure 4.** Comparison of the endogenous peptides detected in Medicago roots at different stages of development. The diagrams show the numbers of peptides present uniquely in the young seedlings compared to the mature plants, as well as the number of detected peptides that are present at both stages of development. The results are shown for the peptides detected using both the DHB and CHCA matrices.

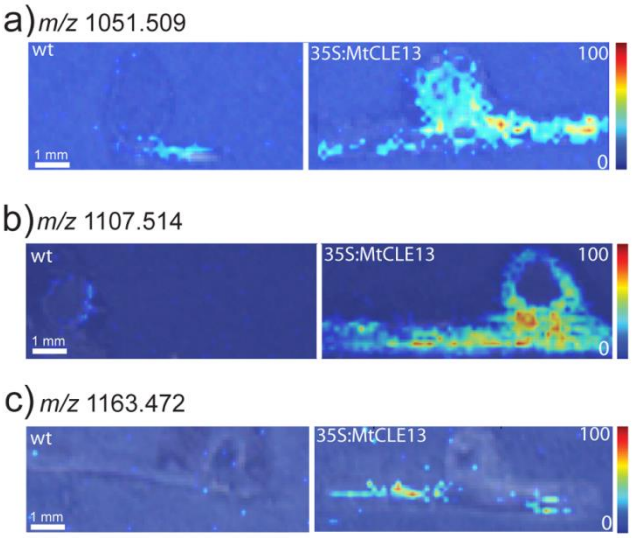



**Figure 5.** Putative peptides that were detected in the *35S:MtCLE13* plants (demonstrating over-expression of CLAVATA3/endosperm-surrounding region (CLE) peptides), but are present in much lower concentrations in the wild-type (wt) plants. a) *m/z* 1051.509 is located in both the plant root and nodule. b) *m/z* 1107.514 is distributed to the plant root and outer nodule. c) *m/z* 1163.472 is localized to the plant root.

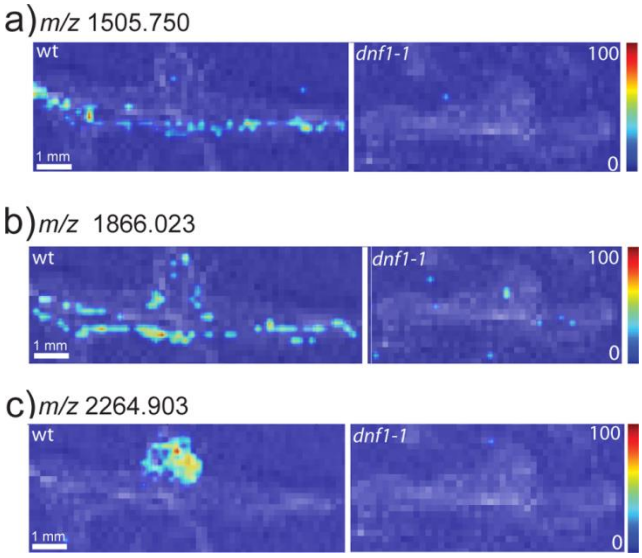



**Figure 6.** Putative peptides that were detected in the wild-type (wt) plants but were absent from the *dnf1-1* plants. *dnf1-1* mutants develop stunted, non-functional nodules, suggesting that these putative peptides may play a role in nodule development or function. a) *m*/*z* 1505.750 is localized to the root. b) *m*/*z* 1866.023 is distributed to the plant root and outer nodule. c) *m*/*z* 2264.903 is localized to the plant nodule.

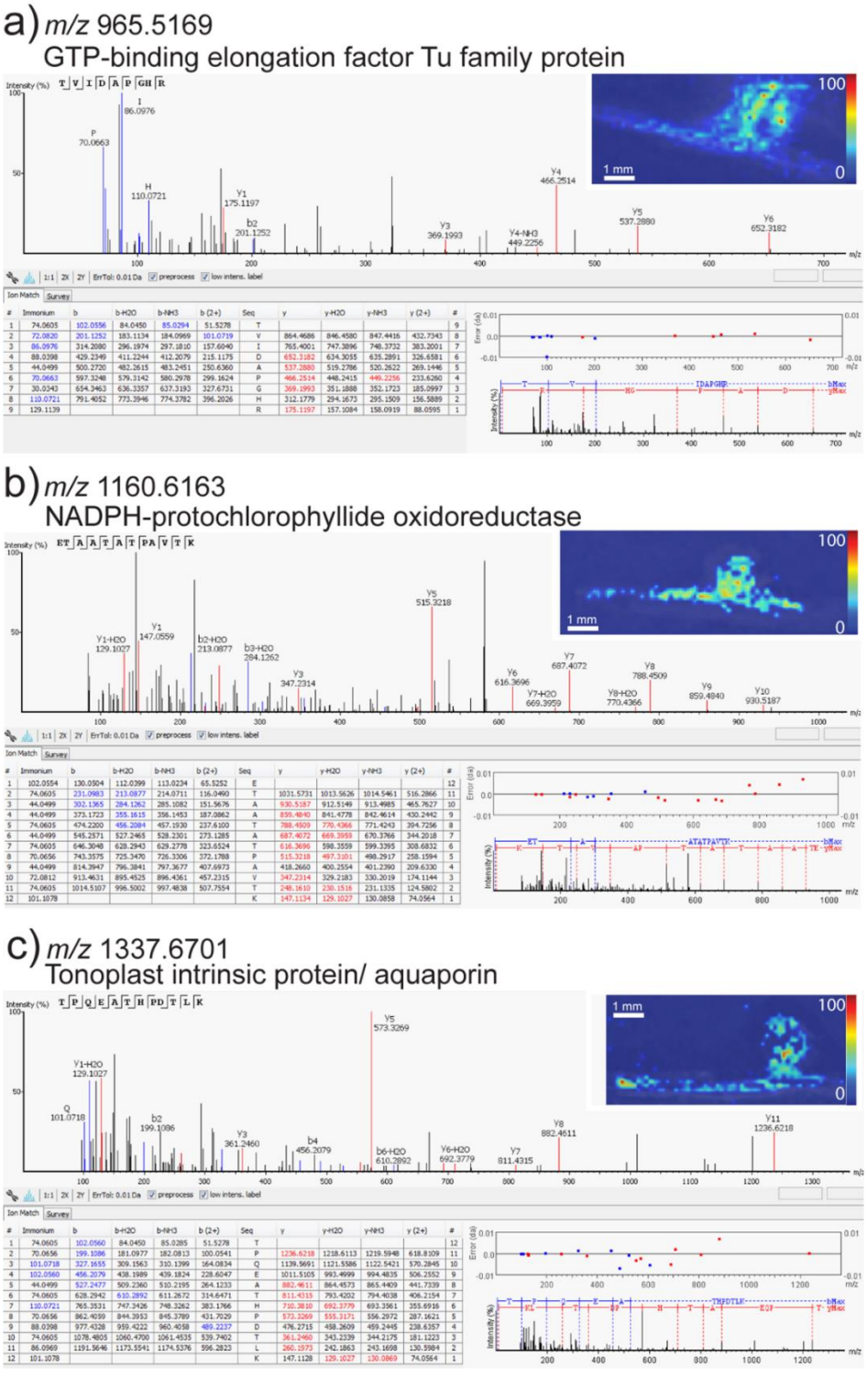



**Figure 7.** Example images and corresponding LC-MS/MS spectra unique peptides of known Medicago proteins. The MS/MS with annotated *de novo* sequencing, the predicted protein, and the corresponding MS image are shown.

**Supplemental Information**

**Supplemental Table 1:** *m/z* Values Detected in Medicago Mature Nodules (DHB matrix)

| | | |
|---|---|---|
| 905.4537 | 1023.281 | 1566.143 |
| 909.4324 | 1025.552 | 1572.195 |
| 911.4474 | 1027.453 | 1586.21 |
| 911.4487 | 1033.416 | 1678.653 |
| 915.2552 | 1051.509 | 1866.025 |
| 919.4321 | 1067.483 | 1953.057 |
| 921.4461 | 1069.581 | 1976.043 |
| 931.2298 | 1091.54 | 1991.013 |
| 935.4247 | 1107.514 | 1992.016 |
| 937.421 | 1113.605 | 2012.997 |
| 944.3733 | 1127.566 | 2030.006 |
| 947.4988 | 1143.539 | 2068.085 |
| 949.5163 | 1145.474 | 2107.087 |
| 951.4013 | 1146.603 | 2233.785 |
| 951.4952 | 1157.631 | 2233.815 |
| 959.5727 | 1160.63 | 2249.759 |
| 960.3487 | 1163.472 | 2249.769 |
| 961.2422 | 1176.603 | 2249.78 |
| 963.4731 | 1177.543 | 2264.903 |
| 963.4891 | 1193.517 | 2532.049 |
| 965.5098 | 1227.394 | 2532.064 |
| 967.4706 | 1261.757 | 2554.031 |
| 967.5171 | 1285.702 | 2555.034 |
| 975.5466 | 1287.696 | 2570.004 |
| 978.3575 | 1299.713 | 2571.023 |
| 979.4901 | 1321.696 | 2592.988 |
| 981.4835 | 1331.538 | 2646.091 |
| 981.5228 | 1337.67 | 2646.107 |
| 991.526 | 1363.705 | 2668.073 |
| 993.5034 | 1379.684 | 2669.077 |
| 995.4849 | 1407.586 | 2684.047 |
| 997.4996 | 1505.75 | 2964.376 |
| 1000.571 | 1521.721 | 3292.523 |
| 1007.5 | 1537.104 | 3468.472 |
| 1009.477 | 1538.119 | 3589.13 |
| 1011.477 | 1540.129 | |
| 1012.476 | 1542.148 | |
| 1017.32 | 1545.165 | |
| 1019.438 | 1564.13 | |

**Supplemental Table 2**: *m/z* Values Detected in Medicago Seedlings (DHB matrix)

| | | | |
|---|---|---|---|
| 900.716 | 984.5932 | 1129.487 | 1323.657 |
| 902.9289 | 991.5242 | 1137.595 | 1331.53 |
| 905.4476 | 992.4176 | 1138.643 | 1337.667 |
| 908.653 | 995.4606 | 1143.534 | 1339.651 |
| 909.4295 | 995.4786 | 1149.63 | 1341.72 |
| 909.5313 | 995.5015 | 1151.646 | 1359.651 |
| 911.4458 | 997.4768 | 1160.625 | 1363.706 |
| 911.6553 | 1000.566 | 1171.615 | 1379.679 |
| 912.4511 | 1002.228 | 1176.599 | 1483.771 |
| 913.6711 | 1007.519 | 1177.544 | 1521.726 |
| 917.6023 | 1012.463 | 1187.586 | 2275.595 |
| 917.7035 | 1013.506 | 1193.519 | |
| 919.4289 | 1014.4 | 1199.489 | |
| 921.4426 | 1019.435 | 1209.509 | |
| 935.4044 | 1019.441 | 1209.57 | |
| 937.3417 | 1022.567 | 1213.471 | |
| 937.42 | 1030.371 | 1225.543 | |
| 937.5868 | 1049.491 | 1233.724 | |
| 944.9339 | 1051.46 | 1247.74 | |
| 947.4997 | 1051.508 | 1261.756 | |
| 949.3856 | 1061.451 | 1263.771 | |
| 949.5151 | 1065.467 | 1264.775 | |
| 951.4385 | 1067.482 | 1269.723 | |
| 951.9594 | 1068.611 | 1271.68 | |
| 953.5624 | 1069.489 | 1275.544 | |
| 955.5785 | 1081.501 | 1275.773 | |
| 958.5155 | 1089.466 | 1283.738 | |
| 959.5718 | 1091.543 | 1283.74 | |
| 960.9113 | 1095.536 | 1285.695 | |
| 963.4907 | 1098.95 | 1291.539 | |
| 965.5076 | 1106.567 | 1293.664 | |
| 967.5114 | 1107.516 | 1299.712 | |
| 972.5361 | 1107.533 | 1301.687 | |
| 975.5453 | 1108.567 | 1301.695 | |
| 979.4683 | 1111.496 | 1305.721 | |
| 981.4825 | 1113.489 | 1307.682 | |
| 982.4828 | 1114.495 | 1309.637 | |
| 982.9018 | 1119.584 | 1313.727 | |
| 984.4889 | 1127.47 | 1321.693 | |

**Supplemental Table 3**: *m/z* Values Detected in Medicago Mature Nodules (CHCA matrix)

| | | | | |
|---|---|---|---|---|
| 905.2795 | 1072.924 | 1658.07 | 2249.779 | 3137.98 |
| 907.9993 | 1085.008 | 1667.522 | 2249.787 | 3180.993 |
| 909.4316 | 1088.895 | 1674.041 | 2250.781 | 3221.998 |
| 911.4473 | 1091.542 | 1685.534 | 2267.696 | 3223.003 |
| 911.9465 | 1097.344 | 1696.026 | 2327.734 | 3253.032 |
| 915.0165 | 1107.002 | 1697.512 | 2345.733 | 3264.049 |
| 921.447 | 1120.026 | 1722.613 | 2363.736 | 3280.027 |
| 922.2633 | 1122.972 | 1753.166 | 2405.747 | 3300.032 |
| 930.9908 | 1135.999 | 1757.556 | 2429.747 | 3342.04 |
| 931.229 | 1142.017 | 1799.566 | 2447.758 | 3384.057 |
| 935.2898 | 1146.479 | 1817.555 | 2459.779 | 3426.098 |
| 939.6062 | 1157.989 | 1829.578 | 2465.335 | 3442.078 |
| 941.279 | 1160.628 | 1859.564 | 2471.756 | 3469.074 |
| 946.9636 | 1173.962 | 1866.655 | 2481.323 | 3588.153 |
| 952.9745 | 1176.604 | 1871.587 | 2487.318 | 3605.134 |
| 959.5741 | 1180.612 | 1889.598 | 2513.765 | 3631.127 |
| 965.3006 | 1196.584 | 1925.694 | 2533.076 | 3668.153 |
| 965.5102 | 1239.636 | 1931.607 | 2554.068 | 3673.139 |
| 968.9476 | 1321.696 | 1965.194 | 2570.028 | 3708.155 |
| 970.6381 | 1337.669 | 1967.631 | 2591.811 | 3750.18 |
| 977.967 | 1343.68 | 1975.045 | 2592.012 | 3792.174 |
| 982.1727 | 1346.703 | 1979.607 | 2614.994 | 3835.191 |
| 983.5484 | 1500.976 | 1991.017 | 2630.966 | 3876.199 |
| 984.5936 | 1504.106 | 1997.027 | 2646.117 | 3913.249 |
| 985.5633 | 1505.469 | 1997.619 | 2669.11 | 3919.21 |
| 986.5531 | 1510.123 | 2003.626 | 2676.822 | |
| 987.5796 | 1516.14 | 2021.624 | 2684.072 | |
| 993.9414 | 1535.481 | 2039.63 | 2706.056 | |
| 998.1492 | 1538.12 | 2063.637 | 2723.86 | |
| 1000.57 | 1540.506 | 2069.734 | 2772.87 | |
| 1007.31 | 1547.481 | 2087.747 | 2795.87 | |
| 1009.917 | 1565.489 | 2103.72 | 2813.878 | |
| 1013.595 | 1571.477 | 2105.639 | 2855.892 | |
| 1014.126 | 1589.489 | 2153.681 | 2934.919 | |
| 1014.586 | 1607.501 | 2183.671 | 2939.948 | |
| 1037.322 | 1613.488 | 2195.692 | 3018.944 | |
| 1063.028 | 1625.513 | 2225.686 | 3102.001 | |
| 1064.561 | 1643.524 | 2233.803 | 3118.979 | |
| 1072.646 | 1655.5 | 2233.805 | 3119.983 | |

**Supplemental Table 4**: *m/z* Values Detected in Medicago Seedlings (CHCA matrix)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 900.0429 | 946.0156 | 995.3091 | 1031.885 | 1084.059 | 1164.602 | 1254.997 | 1312.063 | 1431.946 | 1620.041 |
| 900.7189 | 946.0592 | 998.1507 | 1032.379 | 1085.005 | 1166.051 | 1255.93 | 1313.015 | 1434.578 | 1625.951 |
| 901.2039 | 946.1873 | 999.9355 | 1034.016 | 1087.949 | 1168.115 | 1255.938 | 1315.075 | 1437.965 | 1626.05 |
| 903.2207 | 946.945 | 1000.57 | 1034.953 | 1088.892 | 1176.603 | 1256.005 | 1317.763 | 1444.031 | 1641.923 |
| 906.2865 | 946.9562 | 1002.381 | 1035.018 | 1089.076 | 1179.953 | 1256.939 | 1318.026 | 1447.915 | 1642.033 |
| 906.9794 | 952.9702 | 1003.059 | 1036.117 | 1091.023 | 1182.03 | 1256.998 | 1319.256 | 1449.843 | 1655.055 |
| 908.0881 | 955.0269 | 1004.173 | 1037.9 | 1096.972 | 1182.96 | 1260.071 | 1319.395 | 1450.05 | 1657.997 |
| 908.6536 | 956.0947 | 1006.027 | 1038.075 | 1097.123 | 1182.97 | 1260.955 | 1321.088 | 1456.066 | 1658.009 |
| 908.9494 | 956.563 | 1007.309 | 1040.03 | 1098.031 | 1182.975 | 1261.007 | 1321.7 | 1463.081 | 1659.004 |
| 908.9954 | 957.0875 | 1008.082 | 1040.969 | 1100.981 | 1184.035 | 1261.015 | 1325.031 | 1465.817 | 1669.974 |
| 909.089 | 958.9227 | 1008.33 | 1041.044 | 1104.55 | 1184.09 | 1261.074 | 1327.705 | 1466.024 | 1671.029 |
| 911.9435 | 958.977 | 1009.895 | 1041.07 | 1106.997 | 1190.139 | 1261.952 | 1328.034 | 1467.531 | 1671.986 |
| 912.9873 | 958.9775 | 1013.027 | 1041.975 | 1107.005 | 1200.066 | 1262.022 | 1328.035 | 1472.038 | 1685.951 |
| 913.1099 | 960.2165 | 1014.1 | 1044.91 | 1109.375 | 1203.134 | 1262.03 | 1331.051 | 1479.056 | 1690.08 |
| 913.9911 | 961.9917 | 1014.125 | 1046.045 | 1110.065 | 1204.007 | 1262.954 | 1333.74 | 1481.994 | 1706.046 |
| 915.0145 | 963.1372 | 1014.405 | 1047.052 | 1110.639 | 1204.946 | 1267.033 | 1335.369 | 1482.004 | 1714.004 |
| 916.9893 | 963.4962 | 1015.132 | 1049.048 | 1112.943 | 1205.949 | 1267.972 | 1337.062 | 1485.07 | 1729.976 |
| 917.0057 | 965.5124 | 1015.404 | 1049.989 | 1113.097 | 1206.111 | 1274.039 | 1337.37 | 1488.004 | 1759.031 |
| 919.1544 | 966.2718 | 1015.911 | 1050.928 | 1118.702 | 1209.076 | 1276.988 | 1337.676 | 1488.014 | 1774.003 |
| 920.2496 | 968.0701 | 1018.04 | 1050.993 | 1120.023 | 1209.179 | 1277.927 | 1338.304 | 1493.868 | 1802.054 |
| 921.0227 | 968.6219 | 1019.032 | 1051.928 | 1122.969 | 1217.048 | 1280.063 | 1338.684 | 1501.045 | 1818.027 |
| 922.2627 | 970.5785 | 1020.144 | 1051.986 | 1123.022 | 1219.105 | 1283.005 | 1343.014 | 1503.993 | 1831.064 |
| 924.0615 | 972.5372 | 1020.205 | 1052.358 | 1126.04 | 1219.974 | 1283.944 | 1343.681 | 1504.104 | 1846.079 |
| 926.9777 | 973.0639 | 1021.029 | 1053.87 | 1129.069 | 1219.988 | 1286.072 | 1347.016 | 1504.114 | 1847.048 |
| 929.083 | 974.9517 | 1021.927 | 1055.054 | 1131.979 | 1220.92 | 1290.018 | 1348.423 | 1505.758 | 1862.055 |
| 930.9727 | 976.1923 | 1022.101 | 1056.008 | 1132.597 | 1225.059 | 1293.08 | 1353.039 | 1507.061 | 1890.105 |
| 930.9889 | 977.9646 | 1022.205 | 1056.946 | 1135.995 | 1225.156 | 1296.035 | 1354.277 | 1509.843 | 1906.08 |
| 932.9616 | 979.1085 | 1024.304 | 1059.072 | 1138.002 | 1226.937 | 1296.087 | 1358.987 | 1517.009 | 1934.131 |
| 932.9841 | 979.9669 | 1025.059 | 1063.019 | 1142.014 | 1235.083 | 1299.922 | 1359.654 | 1520.085 | 1950.105 |
| 933.0531 | 980.1142 | 1025.096 | 1065.963 | 1143.015 | 1236.888 | 1300.107 | 1361.793 | 1522.733 | 1978.157 |
| 933.967 | 984.3707 | 1026.063 | 1066.076 | 1144.654 | 1238.089 | 1300.346 | 1374.959 | 1536.061 | 1994.137 |
| 936.2214 | 984.5948 | 1027.999 | 1066.904 | 1146.133 | 1239.027 | 1302.048 | 1375.45 | 1538.899 | |
| 936.3012 | 986.5516 | 1028.146 | 1068.332 | 1146.611 | 1239.964 | 1303.283 | 1377.766 | 1539.057 | |
| 936.9573 | 988.0891 | 1028.936 | 1069.033 | 1147.954 | 1241.128 | 1304.429 | 1405.817 | 1543.755 | |
| 936.9966 | 991.0467 | 1029.006 | 1071.98 | 1148.025 | 1242.908 | 1308.661 | 1415.971 | 1552.025 | |
| 938.238 | 991.527 | 1030.073 | 1072.647 | 1148.119 | 1245.046 | 1309.055 | 1415.973 | 1552.037 | |
| 940.1204 | 992.9824 | 1030.374 | 1072.92 | 1157.986 | 1245.984 | 1309.061 | 1421.791 | 1553.869 | |
| 941.1574 | 993.923 | 1031.078 | 1075.047 | 1160.629 | 1246.051 | 1311.333 | 1428.057 | 1583.427 | |
| 942.9493 | 993.9378 | 1031.38 | 1082.052 | 1161.416 | 1252.063 | 1312.002 | 1429.067 | 1597.898 | |

**Supplemental Figures**

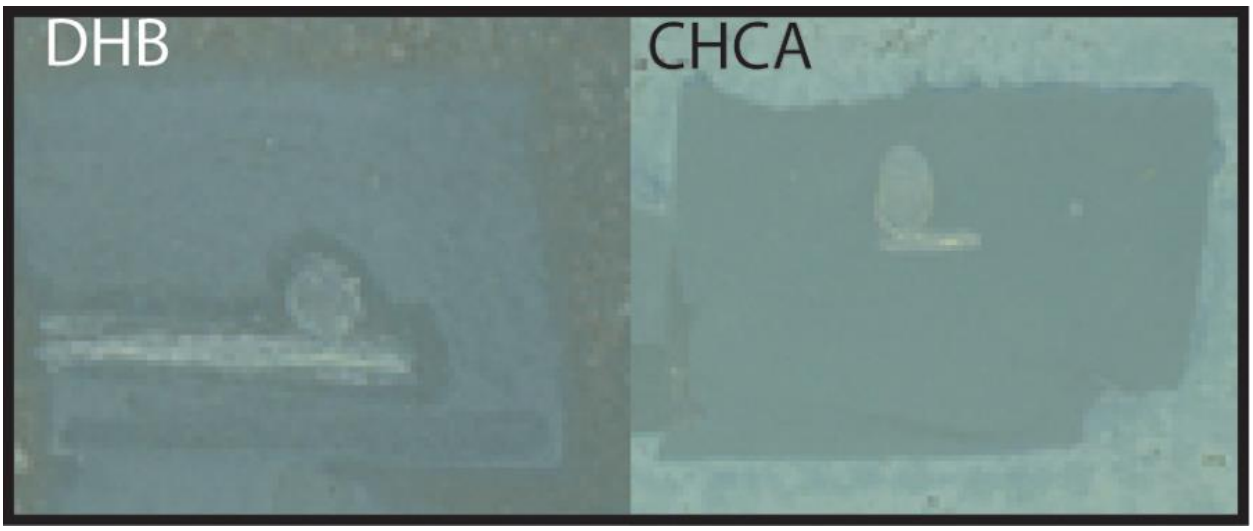
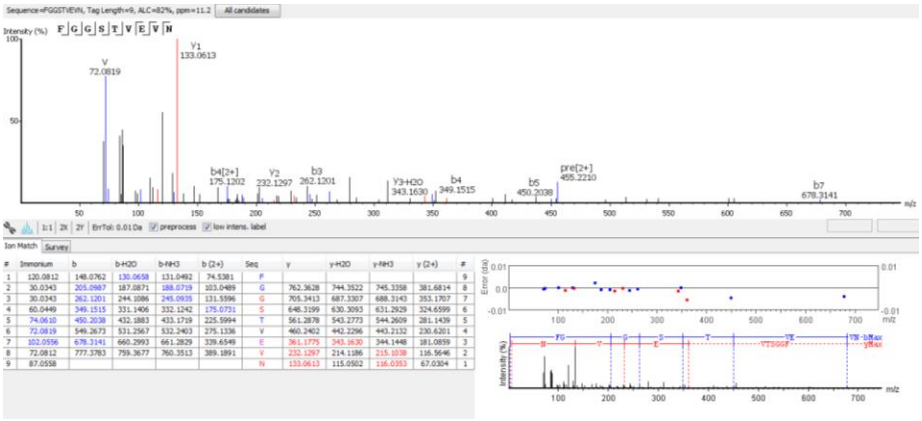


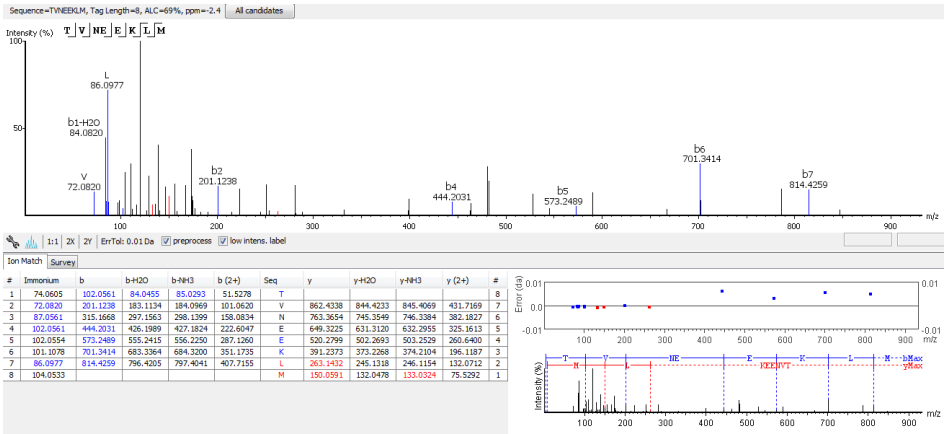

**Figure S1**: Photographs comparing DHB and CHCA matrix application on root nodule slices

**Figure S2:** Annotated MS/MS spectra from *de novo* sequencing.

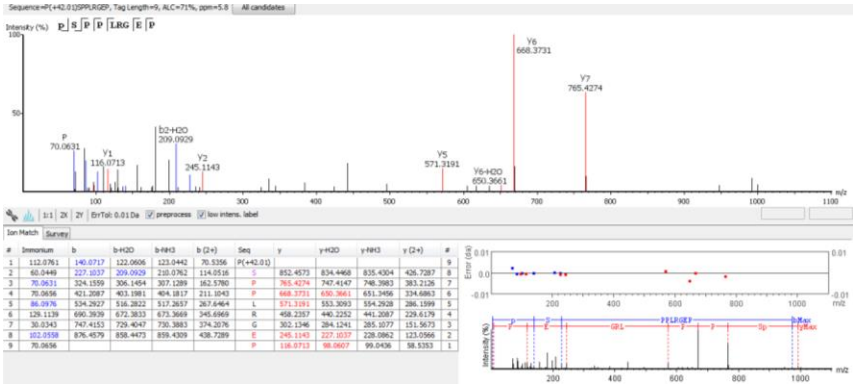
MW 908.4238 LC-MS/MS ALC=82%



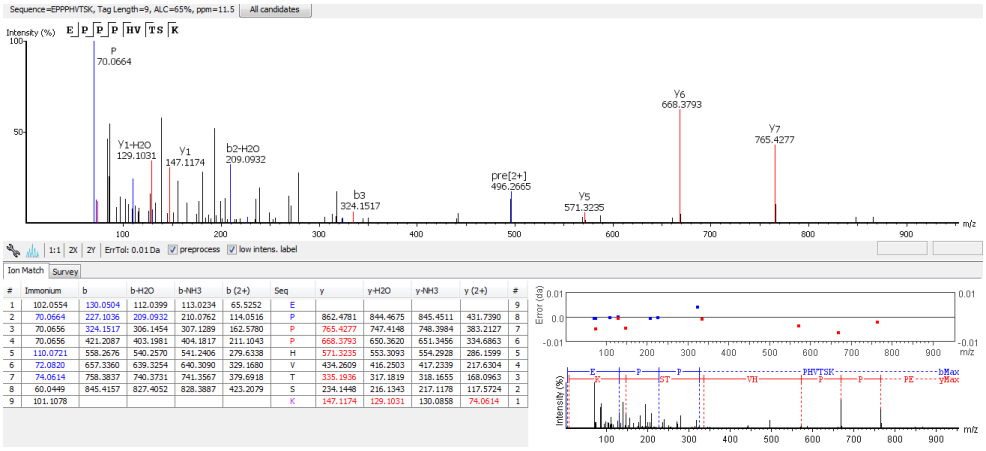
MW 962.4813 LC-MS/MS ALC=69%



MW 990.5163 LC-MS/MS ALC=71%

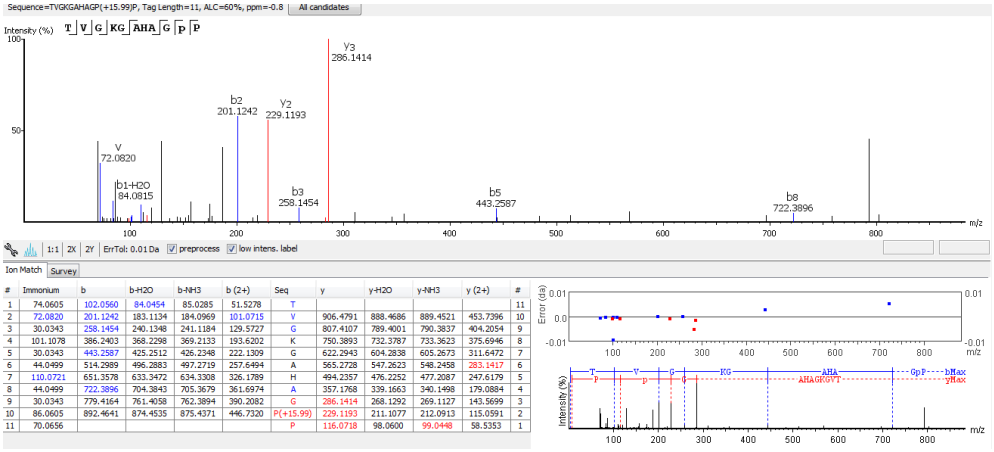


MW 990.5163 LC-MS/MS ALC=65%



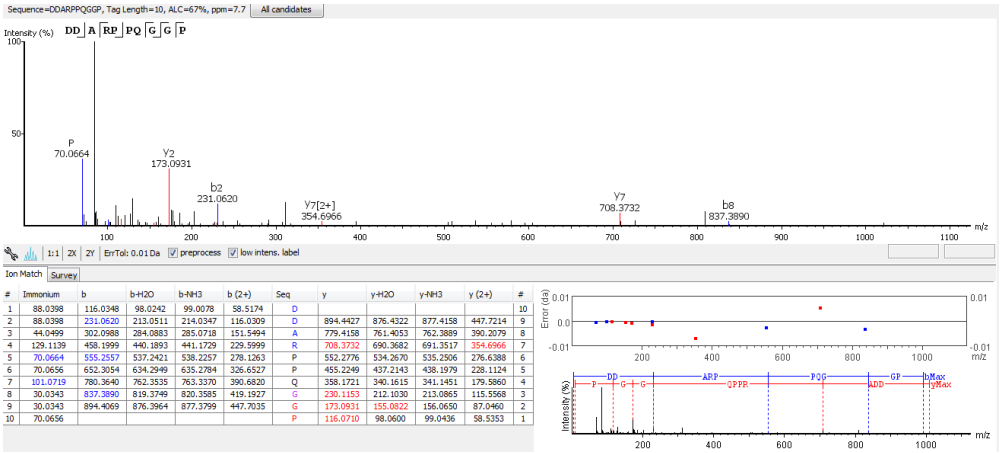

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 102.0554 | 130.0504 | 112.0399 | 113.0234 | 65.5252 | E | | | | | 9 |
| 2 | 70.0664 | 227.1036 | 209.0932 | 210.0762 | 114.0516 | P | 862.4781 | 844.4675 | 845.4511 | 431.7390 | 8 |
| 3 | 70.0656 | 324.1517 | 306.1454 | 307.1289 | 162.5780 | P | 765.4277 | 747.4148 | 748.3984 | 383.2127 | 7 |
| 4 | 70.0656 | 421.2087 | 403.1981 | 404.1817 | 211.1043 | P | 668.3793 | 650.3620 | 651.3456 | 334.6863 | 6 |
| 5 | 110.0721 | 558.2676 | 540.2570 | 541.2406 | 279.6338 | H | 571.3235 | 553.3093 | 554.2928 | 286.1599 | 5 |
| 6 | 72.0820 | 657.3360 | 639.3254 | 640.3090 | 329.1680 | V | 434.2609 | 416.2503 | 417.2339 | 217.6304 | 4 |
| 7 | 74.0614 | 758.3837 | 740.3731 | 741.3567 | 379.6918 | T | 335.1936 | 317.1819 | 318.1655 | 168.0963 | 3 |
| 8 | 60.0449 | 845.4157 | 827.4052 | 828.3887 | 423.2079 | S | 234.1448 | 216.1343 | 217.1178 | 117.5724 | 2 |
| 9 | 101.1078 | | | | | K | 147.1174 | 129.1031 | 130.0858 | 74.0614 | 1 |

MW 1006.5114 LC-MS/MS ALC=60%



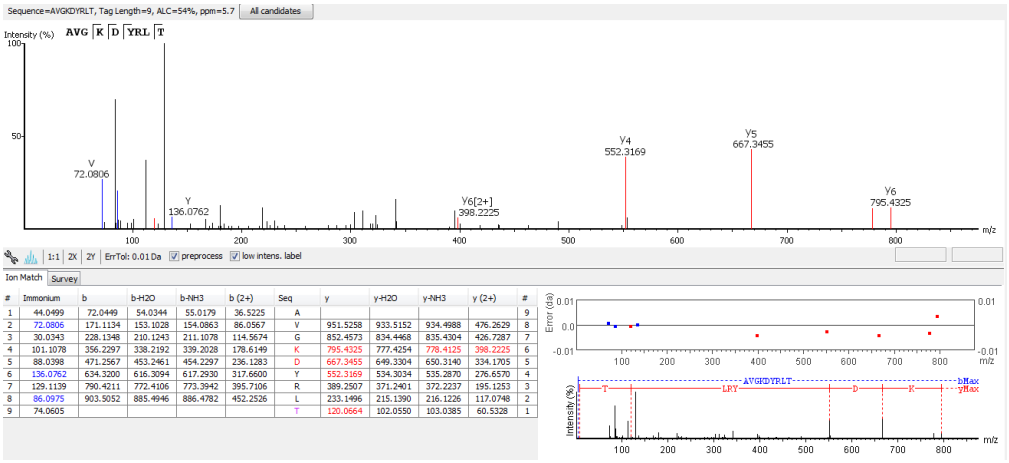

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 74.0605 | 102.0560 | 84.0454 | 85.0285 | 51.5278 | T | | | | | 11 |
| 2 | 72.0820 | 201.1242 | 183.1134 | 184.0969 | 101.0715 | V | 906.4791 | 888.4686 | 889.4521 | 453.7396 | 10 |
| 3 | 30.0343 | 258.1454 | 240.1348 | 241.1184 | 129.5727 | G | 807.4107 | 789.4001 | 790.3837 | 404.2054 | 9 |
| 4 | 101.1078 | 386.2403 | 368.2298 | 369.2133 | 193.6202 | K | 750.3893 | 732.3787 | 733.3623 | 375.6946 | 8 |
| 5 | 30.0343 | 443.2587 | 425.2512 | 426.2348 | 222.1309 | G | 622.2943 | 604.2838 | 605.2673 | 311.6472 | 7 |
| 6 | 44.0499 | 514.2989 | 496.2883 | 497.2719 | 257.6494 | A | 565.2728 | 547.2623 | 548.2458 | 283.1417 | 6 |
| 7 | 110.0721 | 651.3578 | 633.3472 | 634.3308 | 326.1789 | H | 494.2357 | 476.2252 | 477.2087 | 247.6179 | 5 |
| 8 | 44.0499 | 722.3896 | 704.3843 | 705.3679 | 361.6974 | A | 357.1768 | 339.1663 | 340.1498 | 179.0884 | 4 |
| 9 | 30.0343 | 779.4164 | 761.4058 | 762.3894 | 390.2082 | G | 286.1414 | 268.1292 | 269.1127 | 143.5699 | 3 |
| 10 | 86.0605 | 892.4641 | 874.4535 | 875.4371 | 446.7320 | P(+15.99) | 229.1193 | 211.1077 | 212.0913 | 115.0591 | 2 |
| 11 | 70.0656 | | | | | P | 116.0718 | 98.0600 | 99.0448 | 58.5353 | 1 |

MW 1008.4694 LC-MS/MS ALC=67%



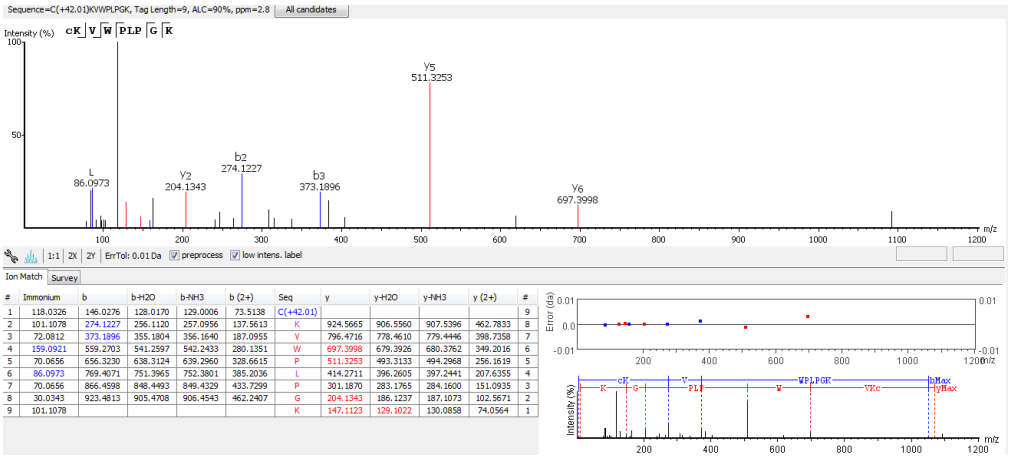

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 88.0398 | 116.0348 | 98.0242 | 99.0078 | 58.5174 | D | | | | | 10 |
| 2 | 88.0398 | 231.0620 | 213.0511 | 214.0347 | 116.0309 | D | 894.4427 | 876.4322 | 877.4158 | 447.7214 | 9 |
| 3 | 44.0499 | 302.0988 | 284.0883 | 285.0718 | 151.5494 | A | 779.4158 | 761.4053 | 762.3889 | 390.2079 | 8 |
| 4 | 129.1139 | 458.1999 | 440.1893 | 441.1729 | 229.5999 | R | 708.3732 | 690.3682 | 691.3517 | 354.6966 | 7 |
| 5 | 70.0664 | 555.2557 | 537.2421 | 538.2257 | 278.1263 | P | 552.2776 | 534.2670 | 535.2506 | 276.6388 | 6 |
| 6 | 70.0656 | 652.3054 | 634.2949 | 635.2784 | 326.6527 | P | 455.2249 | 437.2143 | 438.1979 | 228.1124 | 5 |
| 7 | 101.0719 | 780.3640 | 762.3535 | 763.3370 | 390.6820 | Q | 358.1721 | 340.1615 | 341.1451 | 179.5860 | 4 |
| 8 | 30.0343 | 837.3890 | 819.3749 | 820.3585 | 419.1927 | G | 230.1153 | 212.1030 | 213.0865 | 115.5568 | 3 |
| 9 | 30.0343 | 894.4069 | 876.3964 | 877.3799 | 447.7035 | G | 173.0931 | 155.0822 | 156.0650 | 87.0460 | 2 |
| 10 | 70.0656 | | | | | P | 116.0710 | 98.0600 | 99.0436 | 58.5353 | 1 |

MW 1021.5575 LC-MS/MS ALC=54%



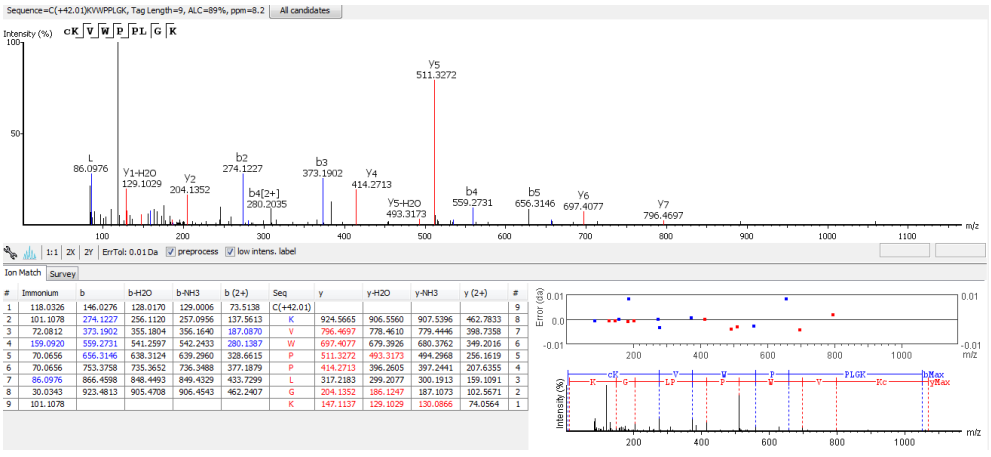

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 44.0499 | 72.0449 | 54.0344 | 55.0179 | 36.5225 | A | | | | | 9 |
| 2 | 72.0806 | 171.1134 | 153.1028 | 154.0863 | 86.0567 | V | 951.5258 | 933.5152 | 934.4988 | 476.2629 | 8 |
| 3 | 30.0343 | 228.1348 | 210.1243 | 211.1078 | 114.5674 | G | 852.4573 | 834.4468 | 835.4304 | 426.7287 | 7 |
| 4 | 101.1078 | 356.2297 | 338.2192 | 339.2028 | 178.6149 | K | 795.4325 | 777.4254 | 778.4125 | 398.2225 | 6 |
| 5 | 88.0398 | 471.2567 | 453.2461 | 454.2297 | 236.1283 | D | 667.3455 | 649.3304 | 650.3140 | 334.1705 | 5 |
| 6 | 136.0762 | 634.3200 | 616.3094 | 617.2930 | 317.6600 | Y | 552.3169 | 534.3034 | 535.2870 | 276.6570 | 4 |
| 7 | 129.1139 | 790.4211 | 772.4106 | 773.3942 | 395.7106 | R | 389.2507 | 371.2401 | 372.2237 | 195.1253 | 3 |
| 8 | 86.0975 | 903.5052 | 885.4946 | 886.4782 | 452.2526 | L | 233.1496 | 215.1390 | 216.1226 | 117.0748 | 2 |
| 9 | 74.0605 | | | | | T | 120.0664 | 102.0550 | 103.0385 | 60.5328 | 1 |

MW 1068.5733 LC-MS/MS ALC=90%



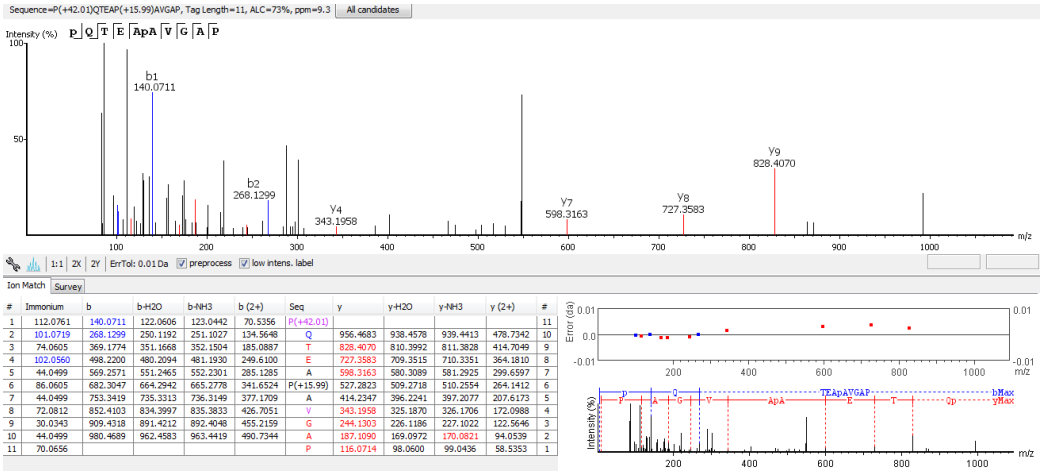

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 118.0326 | 146.0276 | 128.0170 | 129.0006 | 73.5138 | C(+42.01) | | | | | 9 |
| 2 | 101.1078 | 274.1227 | 256.1120 | 257.0956 | 137.5613 | K | 924.5665 | 906.5560 | 907.5396 | 462.7833 | 8 |
| 3 | 72.0812 | 373.1896 | 355.1804 | 356.1640 | 187.0955 | V | 796.4716 | 778.4610 | 779.4446 | 398.7358 | 7 |
| 4 | 159.0921 | 559.2703 | 541.2597 | 542.2433 | 280.1351 | W | 697.3998 | 679.3926 | 680.3762 | 349.2016 | 6 |
| 5 | 70.0656 | 656.3230 | 638.3124 | 639.2960 | 328.6615 | P | 511.3253 | 493.3133 | 494.2968 | 256.1619 | 5 |
| 6 | 86.0973 | 769.4071 | 751.3965 | 752.3801 | 385.2036 | L | 414.2711 | 396.2605 | 397.2441 | 207.6355 | 4 |
| 7 | 70.0656 | 866.4598 | 848.4493 | 849.4329 | 433.7299 | P | 301.1870 | 283.1765 | 284.1600 | 151.0935 | 3 |
| 8 | 30.0343 | 923.4813 | 905.4708 | 906.4543 | 462.2407 | G | 204.1343 | 186.1237 | 187.1073 | 102.5671 | 2 |
| 9 | 101.1078 | | | | | K | 147.1123 | 129.1022 | 130.0858 | 74.0564 | 1 |

MW 1068.5733 LC-MS/MS ALC=89%



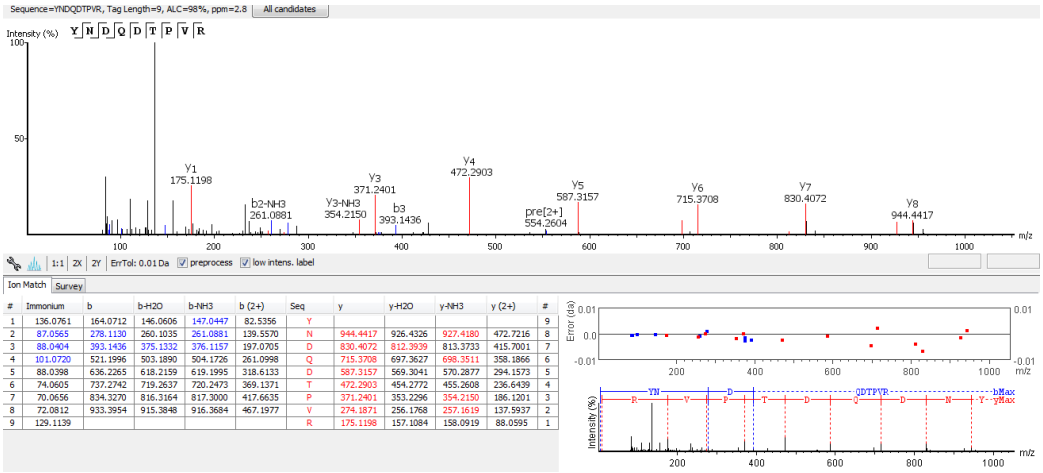

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 118.0326 | 146.0276 | 128.0170 | 129.0006 | 73.5138 | C(+42.01) | | | | | 9 |
| 2 | 101.1078 | 274.1227 | 256.1120 | 257.0956 | 137.5613 | K | 924.5665 | 906.5560 | 907.5396 | 462.7833 | 8 |
| 3 | 72.0812 | 373.1902 | 355.1804 | 356.1640 | 187.0870 | V | 796.4697 | 778.4610 | 779.4446 | 398.7358 | 7 |
| 4 | 159.0920 | 559.2731 | 541.2597 | 542.2433 | 280.1387 | W | 697.4077 | 679.3926 | 680.3762 | 349.2016 | 6 |
| 5 | 70.0656 | 656.3146 | 638.3124 | 639.2960 | 328.6615 | P | 511.3272 | 493.3173 | 494.2968 | 256.1619 | 5 |
| 6 | 70.0656 | 753.3758 | 735.3652 | 736.3488 | 377.1879 | P | 414.2713 | 396.2605 | 397.2441 | 207.6355 | 4 |
| 7 | 86.0976 | 866.4598 | 848.4493 | 849.4329 | 433.7299 | L | 317.2183 | 299.2077 | 300.1913 | 159.1091 | 3 |
| 8 | 30.0343 | 923.4813 | 905.4708 | 906.4543 | 462.2407 | G | 204.1352 | 186.1247 | 187.1073 | 102.5671 | 2 |
| 9 | 101.1078 | | | | | K | 147.1137 | 129.1029 | 130.0866 | 74.0564 | 1 |

MW 1094.5280 LC-MS/MS ALC=73%



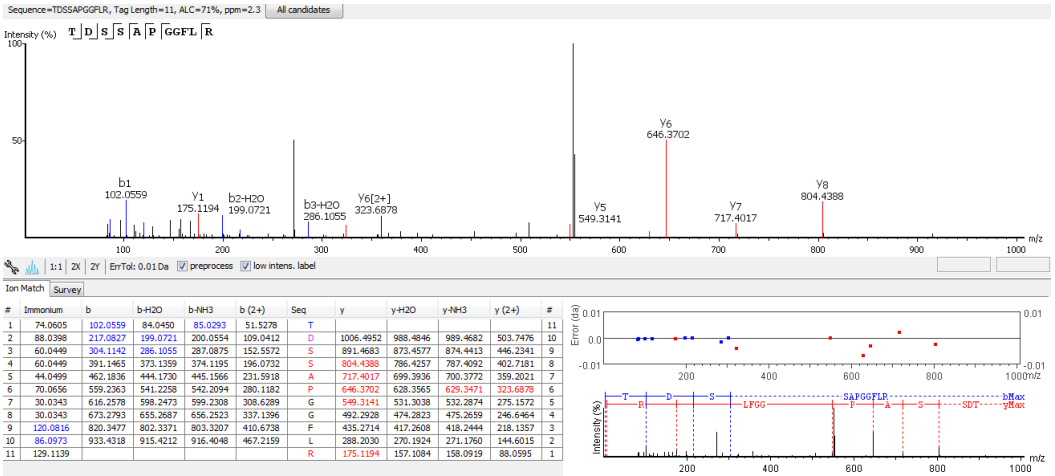

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 112.0761 | 140.0711 | 122.0606 | 123.0442 | 70.5356 | P(+42.01) | | | | | 11 |
| 2 | 101.0719 | 268.1299 | 250.1192 | 251.1027 | 134.5648 | Q | 956.4683 | 938.4578 | 939.4413 | 478.7342 | 10 |
| 3 | 74.0605 | 369.1774 | 351.1668 | 352.1504 | 185.0887 | T | 828.4070 | 810.3992 | 811.3828 | 414.7049 | 9 |
| 4 | 102.0560 | 498.2200 | 480.2094 | 481.1930 | 249.6100 | E | 727.3583 | 709.3515 | 710.3351 | 364.1810 | 8 |
| 5 | 44.0499 | 569.2571 | 551.2465 | 552.2301 | 285.1285 | A | 598.3163 | 580.3089 | 581.2925 | 299.6597 | 7 |
| 6 | 86.0605 | 682.3047 | 664.2942 | 665.2778 | 341.6524 | P(+15.99) | 527.2823 | 509.2718 | 510.2554 | 264.1412 | 6 |
| 7 | 44.0499 | 753.3419 | 735.3313 | 736.3149 | 377.1709 | A | 414.2347 | 396.2241 | 397.2077 | 207.6173 | 5 |
| 8 | 72.0812 | 852.4103 | 834.3997 | 835.3833 | 426.7051 | V | 343.1958 | 325.1870 | 326.1706 | 172.0988 | 4 |
| 9 | 30.0343 | 909.4318 | 891.4212 | 892.4048 | 455.2159 | G | 244.1303 | 226.1186 | 227.1022 | 122.5646 | 3 |
| 10 | 44.0499 | 980.4689 | 962.4583 | 963.4419 | 490.7344 | A | 187.1090 | 169.0972 | 170.0821 | 94.0539 | 2 |
| 11 | 70.0656 | | | | | P | 116.0714 | 98.0600 | 99.0436 | 58.5353 | 1 |

MW 1106.5081 LC-MS/MS ALC=98%



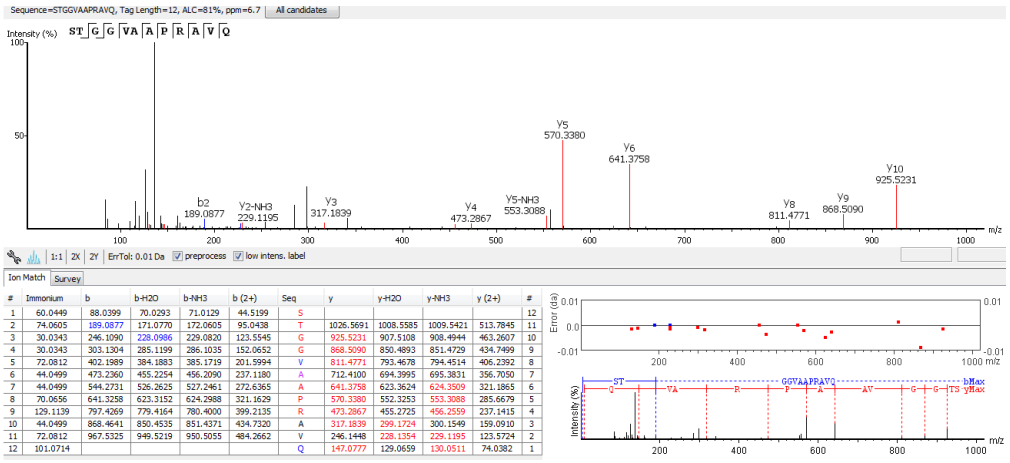

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 136.0761 | 164.0712 | 146.0606 | 147.0447 | 82.5356 | Y | | | | | 9 |
| 2 | 87.0565 | 278.1130 | 260.1035 | 261.0881 | 139.5570 | N | 944.4417 | 926.4326 | 927.4180 | 472.7216 | 8 |
| 3 | 88.0404 | 393.1436 | 375.1332 | 376.1157 | 197.0705 | D | 830.4072 | 812.3939 | 813.3733 | 415.7001 | 7 |
| 4 | 101.0720 | 521.1996 | 503.1890 | 504.1726 | 261.0998 | Q | 715.3708 | 697.3627 | 698.3511 | 358.1866 | 6 |
| 5 | 88.0398 | 636.2265 | 618.2159 | 619.1995 | 318.6133 | D | 587.3157 | 569.3041 | 570.2877 | 294.1573 | 5 |
| 6 | 74.0605 | 737.2742 | 719.2637 | 720.2473 | 369.1371 | T | 472.2903 | 454.2772 | 455.2608 | 236.6439 | 4 |
| 7 | 70.0656 | 834.3270 | 816.3164 | 817.3000 | 417.6635 | P | 371.2401 | 353.2296 | 354.2150 | 186.1201 | 3 |
| 8 | 72.0812 | 933.3954 | 915.3848 | 916.3684 | 467.1977 | V | 274.1871 | 256.1768 | 257.1619 | 137.5937 | 2 |
| 9 | 129.1139 | | | | | R | 175.1198 | 157.1084 | 158.0919 | 88.0595 | 1 |

MW 1106.5255 LC-MS/MS ALC=71%



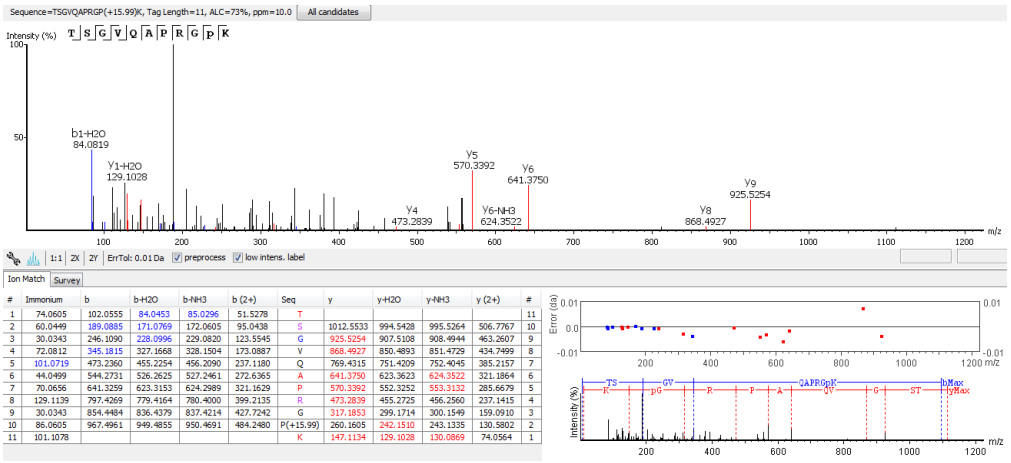

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 74.0605 | 102.0559 | 84.0450 | 85.0293 | 51.5278 | T | | | | | 11 |
| 2 | 88.0398 | 217.0827 | 199.0721 | 200.0554 | 109.0412 | D | 1006.4952 | 988.4846 | 989.4682 | 503.7476 | 10 |
| 3 | 60.0449 | 304.1142 | 286.1055 | 287.0875 | 152.5572 | S | 891.4683 | 873.4577 | 874.4413 | 446.2341 | 9 |
| 4 | 60.0449 | 391.1465 | 373.1359 | 374.1195 | 196.0732 | S | 804.4388 | 786.4257 | 787.4092 | 402.7181 | 8 |
| 5 | 44.0499 | 462.1836 | 444.1730 | 445.1566 | 231.5918 | A | 717.4017 | 699.3936 | 700.3772 | 359.2021 | 7 |
| 6 | 70.0656 | 559.2363 | 541.2258 | 542.2094 | 280.1182 | P | 646.3702 | 628.3565 | 629.3471 | 323.6878 | 6 |
| 7 | 30.0343 | 616.2578 | 598.2473 | 599.2308 | 308.6289 | G | 549.3141 | 531.3038 | 532.2874 | 275.1572 | 5 |
| 8 | 30.0343 | 673.2793 | 655.2687 | 656.2523 | 337.1396 | G | 492.2928 | 474.2823 | 475.2659 | 246.6464 | 4 |
| 9 | 120.0816 | 820.3477 | 802.3371 | 803.3207 | 410.6738 | F | 435.2714 | 417.2608 | 418.2444 | 218.1357 | 3 |
| 10 | 86.0973 | 933.4318 | 915.4212 | 916.4048 | 467.2159 | L | 288.2030 | 270.1924 | 271.1760 | 144.6015 | 2 |
| 11 | 129.1139 | | | | | R | 175.1194 | 157.1084 | 158.0919 | 88.0595 | 1 |

MW 1112.5974 LC-MS/MS ALC=81%



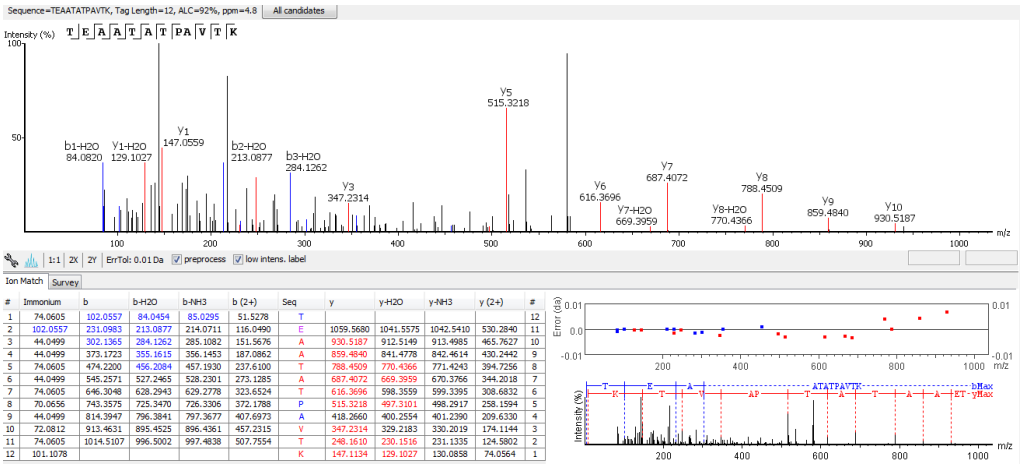

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 60.0449 | 88.0399 | 70.0293 | 71.0129 | 44.5199 | S | | | | | 12 |
| 2 | 74.0605 | 189.0877 | 171.0770 | 172.0605 | 95.0438 | T | 1026.5691 | 1008.5585 | 1009.5421 | 513.7845 | 11 |
| 3 | 30.0343 | 246.1090 | 228.0986 | 229.0820 | 123.5545 | G | 925.5231 | 907.5108 | 908.4944 | 463.2607 | 10 |
| 4 | 30.0343 | 303.1304 | 285.1199 | 286.1035 | 152.0652 | G | 868.5090 | 850.4893 | 851.4729 | 434.7499 | 9 |
| 5 | 72.0812 | 402.1989 | 384.1883 | 385.1719 | 201.5994 | V | 811.4771 | 793.4678 | 794.4514 | 406.2392 | 8 |
| 6 | 44.0499 | 473.2360 | 455.2254 | 456.2090 | 237.1180 | A | 712.4100 | 694.3995 | 695.3831 | 356.7050 | 7 |
| 7 | 44.0499 | 544.2731 | 526.2625 | 527.2461 | 272.6365 | A | 641.3758 | 623.3624 | 624.3509 | 321.1865 | 6 |
| 8 | 70.0656 | 641.3258 | 623.3152 | 624.2988 | 321.1629 | P | 570.3380 | 552.3253 | 553.3088 | 285.6679 | 5 |
| 9 | 129.1139 | 797.4269 | 779.4164 | 780.4000 | 399.2135 | R | 473.2867 | 455.2725 | 456.2559 | 237.1415 | 4 |
| 10 | 44.0499 | 868.4641 | 850.4535 | 851.4371 | 434.7320 | A | 317.1839 | 299.1724 | 300.1549 | 159.0910 | 3 |
| 11 | 72.0812 | 967.5325 | 949.5219 | 950.5055 | 484.2662 | V | 246.1448 | 228.1354 | 229.1195 | 123.5724 | 2 |
| 12 | 101.0714 | | | | | Q | 147.0777 | 129.0659 | 130.0511 | 74.0382 | 1 |

MW 1112.5974 LC-MS/MS ALC=73%



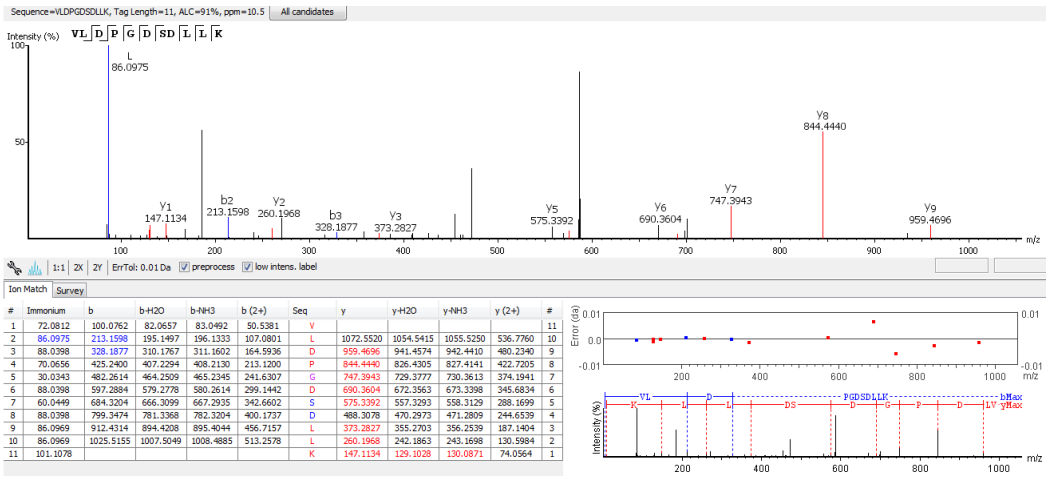

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 74.0605 | 102.0555 | 84.0453 | 85.0296 | 51.5278 | T | | | | | 11 |
| 2 | 60.0449 | 189.0885 | 171.0769 | 172.0605 | 95.0438 | S | 1012.5533 | 994.5428 | 995.5264 | 506.7767 | 10 |
| 3 | 30.0343 | 246.1090 | 228.0996 | 229.0820 | 123.5545 | G | 925.5254 | 907.5108 | 908.4944 | 463.2607 | 9 |
| 4 | 72.0812 | 345.1815 | 327.1668 | 328.1504 | 173.0887 | V | 868.4927 | 850.4893 | 851.4729 | 434.7499 | 8 |
| 5 | 101.0719 | 473.2360 | 455.2254 | 456.2090 | 237.1180 | Q | 769.4315 | 751.4209 | 752.4045 | 385.2157 | 7 |
| 6 | 44.0499 | 544.2731 | 526.2625 | 527.2461 | 272.6365 | A | 641.3750 | 623.3623 | 624.3522 | 321.1864 | 6 |
| 7 | 70.0656 | 641.3259 | 623.3153 | 624.2989 | 321.1629 | P | 570.3392 | 552.3252 | 553.3132 | 285.6679 | 5 |
| 8 | 129.1139 | 797.4269 | 779.4164 | 780.4000 | 399.2135 | R | 473.2839 | 455.2725 | 456.2560 | 237.1415 | 4 |
| 9 | 30.0343 | 854.4484 | 836.4379 | 837.4214 | 427.7242 | G | 317.1853 | 299.1714 | 300.1549 | 159.0910 | 3 |
| 10 | 86.0605 | 967.4961 | 949.4855 | 950.4691 | 484.2480 | P(+15.99) | 260.1605 | 242.1510 | 243.1335 | 130.5802 | 2 |
| 11 | 101.1078 | | | | | K | 147.1134 | 129.1028 | 130.0869 | 74.0564 | 1 |

MW 1159.6191 LC-MS/MS ALC=92%



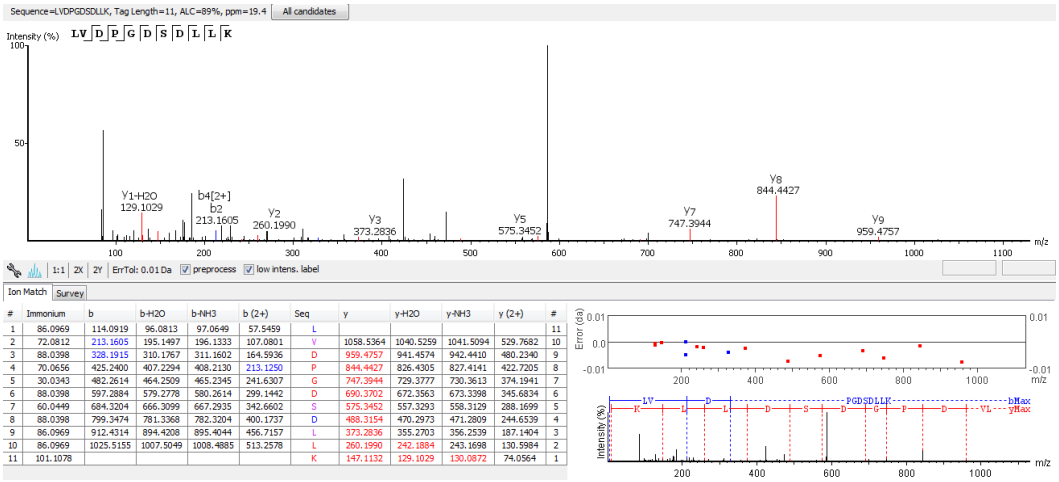

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 74.0605 | 102.0557 | 84.0454 | 85.0295 | 51.5278 | T | | | | | 12 |
| 2 | 102.0557 | 231.0983 | 213.0877 | 214.0711 | 116.0490 | E | 1059.5680 | 1041.5575 | 1042.5410 | 530.2840 | 11 |
| 3 | 44.0499 | 302.1365 | 284.1262 | 285.1082 | 151.5676 | A | 930.5187 | 912.5149 | 913.4985 | 465.7627 | 10 |
| 4 | 44.0499 | 373.1723 | 355.1615 | 356.1453 | 187.0862 | A | 859.4840 | 841.4778 | 842.4614 | 430.2442 | 9 |
| 5 | 74.0605 | 474.2200 | 456.2084 | 457.1930 | 237.6100 | T | 788.4509 | 770.4366 | 771.4243 | 394.7256 | 8 |
| 6 | 44.0499 | 545.2571 | 527.2465 | 528.2301 | 273.1285 | A | 687.4072 | 669.3959 | 670.3766 | 344.2018 | 7 |
| 7 | 74.0605 | 646.3048 | 628.2943 | 629.2778 | 323.6524 | T | 616.3696 | 598.3559 | 599.3395 | 308.6832 | 6 |
| 8 | 70.0656 | 743.3575 | 725.3470 | 726.3306 | 372.1788 | P | 515.3218 | 497.3101 | 498.2917 | 258.1594 | 5 |
| 9 | 44.0499 | 814.3947 | 796.3841 | 797.3677 | 407.6973 | A | 418.2660 | 400.2554 | 401.2390 | 209.6330 | 4 |
| 10 | 72.0812 | 913.4631 | 895.4525 | 896.4361 | 457.2315 | V | 347.2314 | 329.2183 | 330.2019 | 174.1144 | 3 |
| 11 | 74.0605 | 1014.5107 | 996.5002 | 997.4838 | 507.7554 | T | 248.1610 | 230.1516 | 231.1335 | 124.5802 | 2 |
| 12 | 101.1078 | | | | | K | 147.1134 | 129.1027 | 130.0858 | 74.0564 | 1 |

MW 1170.6070 LC-MS/MS ALC=91%



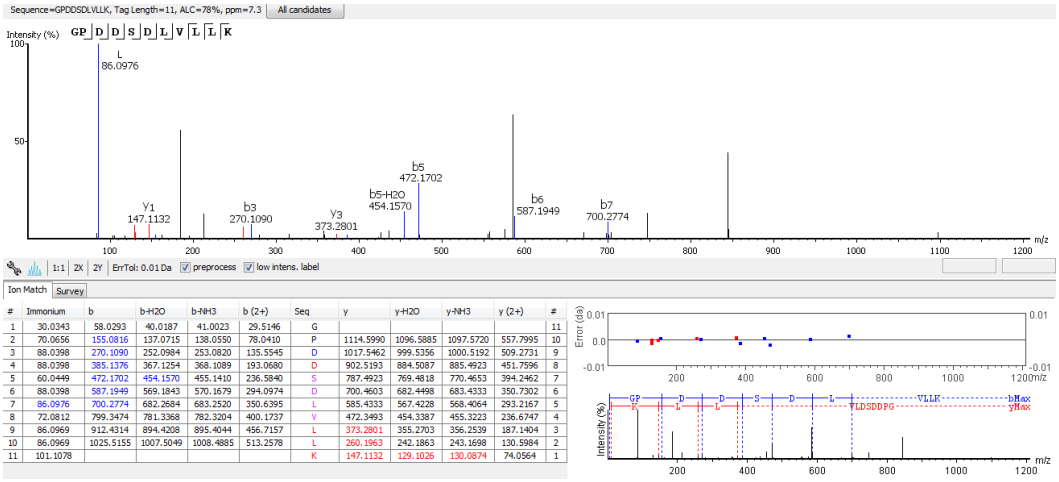

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 72.0812 | 100.0762 | 82.0657 | 83.0492 | 50.5381 | V | | | | | 11 |
| 2 | 86.0975 | 213.1598 | 195.1497 | 196.1333 | 107.0801 | L | 1072.5520 | 1054.5415 | 1055.5250 | 536.7760 | 10 |
| 3 | 88.0398 | 328.1877 | 310.1767 | 311.1602 | 164.5936 | D | 959.4696 | 941.4574 | 942.4410 | 480.2340 | 9 |
| 4 | 70.0656 | 425.2400 | 407.2294 | 408.2130 | 213.1200 | P | 844.4440 | 826.4305 | 827.4141 | 422.7205 | 8 |
| 5 | 30.0343 | 482.2614 | 464.2509 | 465.2345 | 241.6307 | G | 747.3943 | 729.3777 | 730.3613 | 374.1941 | 7 |
| 6 | 88.0398 | 597.2884 | 579.2778 | 580.2614 | 299.1442 | D | 690.3604 | 672.3563 | 673.3398 | 345.6834 | 6 |
| 7 | 60.0449 | 684.3204 | 666.3099 | 667.2935 | 342.6602 | S | 575.3392 | 557.3293 | 558.3129 | 288.1699 | 5 |
| 8 | 88.0398 | 799.3474 | 781.3368 | 782.3204 | 400.1737 | D | 488.3078 | 470.2973 | 471.2809 | 244.6539 | 4 |
| 9 | 86.0969 | 912.4314 | 894.4208 | 895.4044 | 456.7157 | L | 373.2827 | 355.2703 | 356.2539 | 187.1404 | 3 |
| 10 | 86.0969 | 1025.5155 | 1007.5049 | 1008.4885 | 513.2578 | L | 260.1968 | 242.1863 | 243.1698 | 130.5984 | 2 |
| 11 | 101.1078 | | | | | K | 147.1134 | 129.1028 | 130.0871 | 74.0564 | 1 |

MW 1170.6070 LC-MS/MS ALC=89%



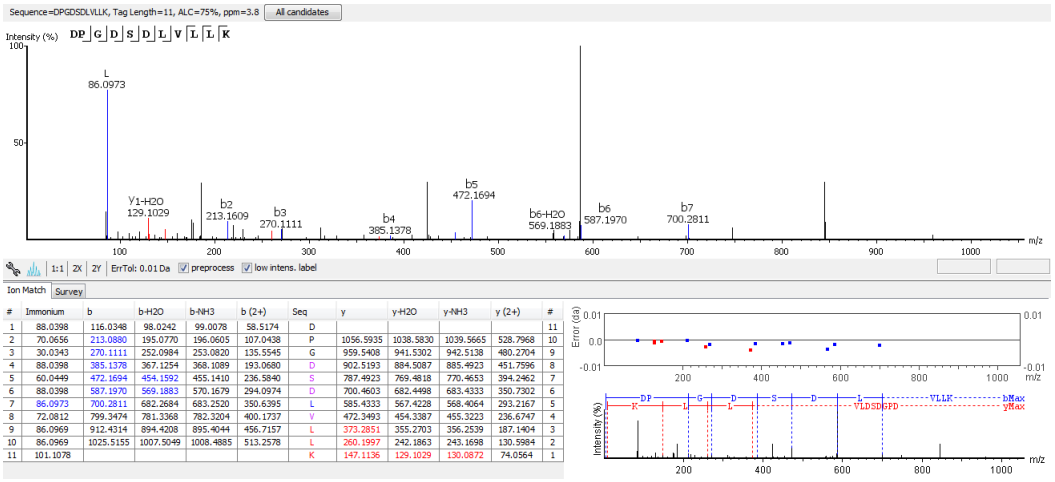

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 86.0969 | 114.0919 | 96.0813 | 97.0649 | 57.5459 | L | | | | | 11 |
| 2 | 72.0812 | 213.1605 | 195.1497 | 196.1333 | 107.0801 | V | 1058.5364 | 1040.5259 | 1041.5094 | 529.7682 | 10 |
| 3 | 88.0398 | 328.1915 | 310.1767 | 311.1602 | 164.5936 | D | 959.4757 | 941.4574 | 942.4410 | 480.2340 | 9 |
| 4 | 70.0656 | 425.2400 | 407.2294 | 408.2130 | 213.1250 | P | 844.4427 | 826.4305 | 827.4141 | 422.7205 | 8 |
| 5 | 30.0343 | 482.2614 | 464.2509 | 465.2345 | 241.6307 | G | 747.3944 | 729.3777 | 730.3613 | 374.1941 | 7 |
| 6 | 88.0398 | 597.2884 | 579.2778 | 580.2614 | 299.1442 | D | 690.3702 | 672.3563 | 673.3398 | 345.6834 | 6 |
| 7 | 60.0449 | 684.3204 | 666.3099 | 667.2935 | 342.6602 | S | 575.3452 | 557.3293 | 558.3129 | 288.1699 | 5 |
| 8 | 88.0398 | 799.3474 | 781.3368 | 782.3204 | 400.1737 | D | 488.3154 | 470.2973 | 471.2809 | 244.6539 | 4 |
| 9 | 86.0969 | 912.4314 | 894.4208 | 895.4044 | 456.7157 | L | 373.2836 | 355.2703 | 356.2539 | 187.1404 | 3 |
| 10 | 86.0969 | 1025.5155 | 1007.5049 | 1008.4885 | 513.2578 | L | 260.1990 | 242.1884 | 243.1698 | 130.5984 | 2 |
| 11 | 101.1078 | | | | | K | 147.1132 | 129.1029 | 130.0872 | 74.0564 | 1 |

MW 1170.6070 LC-MS/MS ALC=78%



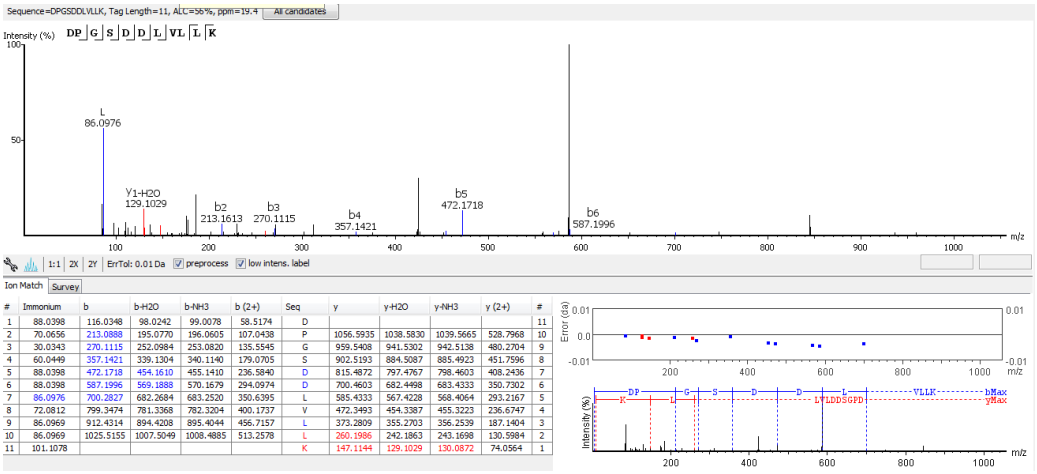

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 30.0343 | 58.0293 | 40.0187 | 41.0023 | 29.5146 | G | | | | | 11 |
| 2 | 70.0656 | 155.0816 | 137.0715 | 138.0550 | 78.0410 | P | 1114.5990 | 1096.5885 | 1097.5720 | 557.7995 | 10 |
| 3 | 88.0398 | 270.1090 | 252.0984 | 253.0820 | 135.5545 | D | 1017.5462 | 999.5356 | 1000.5192 | 509.2731 | 9 |
| 4 | 88.0398 | 385.1376 | 367.1254 | 368.1089 | 193.0680 | D | 902.5193 | 884.5087 | 885.4923 | 451.7596 | 8 |
| 5 | 60.0449 | 472.1702 | 454.1570 | 455.1410 | 236.5840 | S | 787.4923 | 769.4818 | 770.4653 | 394.2462 | 7 |
| 6 | 88.0398 | 587.1949 | 569.1843 | 570.1679 | 294.0974 | D | 700.4603 | 682.4498 | 683.4333 | 350.7302 | 6 |
| 7 | 86.0976 | 700.2774 | 682.2684 | 683.2520 | 350.6395 | L | 585.4333 | 567.4228 | 568.4064 | 293.2167 | 5 |
| 8 | 72.0812 | 799.3474 | 781.3368 | 782.3204 | 400.1737 | V | 472.3493 | 454.3387 | 455.3223 | 236.6747 | 4 |
| 9 | 86.0969 | 912.4314 | 894.4208 | 895.4044 | 456.7157 | L | 373.2801 | 355.2703 | 356.2539 | 187.1404 | 3 |
| 10 | 86.0969 | 1025.5155 | 1007.5049 | 1008.4885 | 513.2578 | L | 260.1963 | 242.1863 | 243.1698 | 130.5984 | 2 |
| 11 | 101.1078 | | | | | K | 147.1132 | 129.1026 | 130.0874 | 74.0564 | 1 |

MW 1170.6070 LC-MS/MS ALC=75%



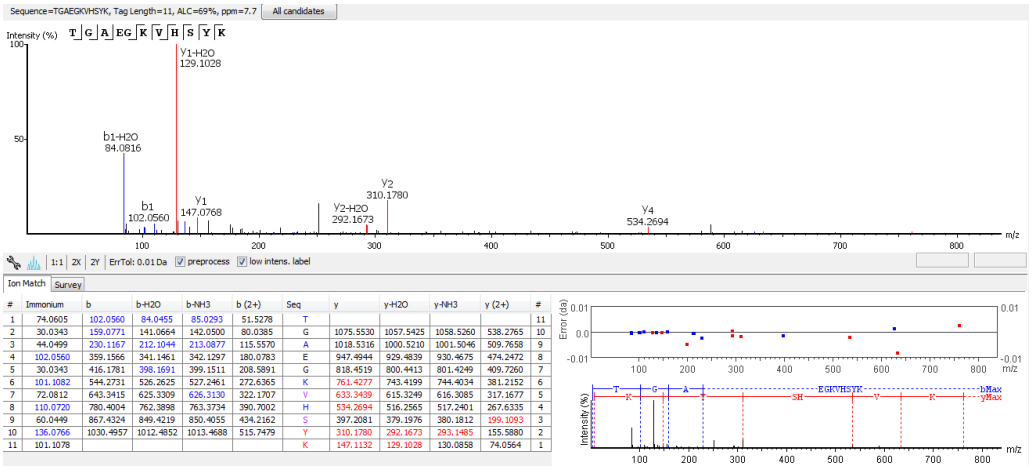

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 88.0398 | 116.0348 | 98.0242 | 99.0078 | 58.5174 | D | | | | | 11 |
| 2 | 70.0656 | 213.0880 | 195.0770 | 196.0605 | 107.0438 | P | 1056.5935 | 1038.5830 | 1039.5665 | 528.7968 | 10 |
| 3 | 30.0343 | 270.1111 | 252.0984 | 253.0820 | 135.5545 | G | 959.5408 | 941.5302 | 942.5138 | 480.2704 | 9 |
| 4 | 88.0398 | 385.1378 | 367.1254 | 368.1089 | 193.0680 | D | 902.5193 | 884.5087 | 885.4923 | 451.7596 | 8 |
| 5 | 60.0449 | 472.1694 | 454.1592 | 455.1410 | 236.5840 | S | 787.4923 | 769.4818 | 770.4653 | 394.2462 | 7 |
| 6 | 88.0398 | 587.1970 | 569.1883 | 570.1679 | 294.0974 | D | 700.4603 | 682.4498 | 683.4333 | 350.7302 | 6 |
| 7 | 86.0973 | 700.2811 | 682.2684 | 683.2520 | 350.6395 | L | 585.4333 | 567.4228 | 568.4064 | 293.2167 | 5 |
| 8 | 72.0812 | 799.3474 | 781.3368 | 782.3204 | 400.1737 | V | 472.3493 | 454.3387 | 455.3223 | 236.6747 | 4 |
| 9 | 86.0969 | 912.4314 | 894.4208 | 895.4044 | 456.7157 | L | 373.2851 | 355.2703 | 356.2539 | 187.1404 | 3 |
| 10 | 86.0969 | 1025.5155 | 1007.5049 | 1008.4885 | 513.2578 | L | 260.1997 | 242.1863 | 243.1698 | 130.5984 | 2 |
| 11 | 101.1078 | | | | | K | 147.1136 | 129.1029 | 130.0872 | 74.0564 | 1 |

MW 1170.6070 LC-MS/MS ALC=56%



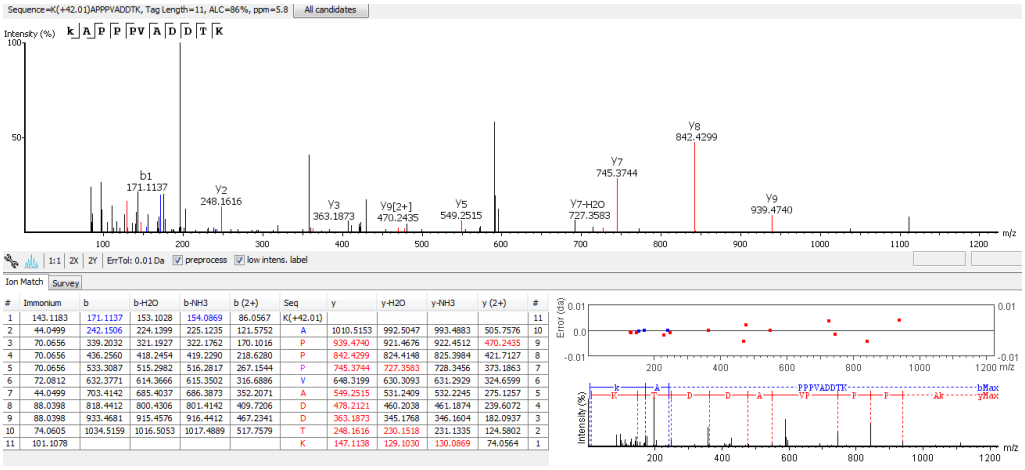

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 88.0398 | 116.0348 | 98.0242 | 99.0078 | 58.5174 | D | | | | | 11 |
| 2 | 70.0656 | 213.0888 | 195.0770 | 196.0605 | 107.0438 | P | 1056.5935 | 1038.5830 | 1039.5665 | 528.7968 | 10 |
| 3 | 30.0343 | 270.1115 | 252.0984 | 253.0820 | 135.5545 | G | 959.5408 | 941.5302 | 942.5138 | 480.2704 | 9 |
| 4 | 60.0449 | 357.1421 | 339.1304 | 340.1140 | 179.0705 | S | 902.5193 | 884.5087 | 885.4923 | 451.7596 | 8 |
| 5 | 88.0398 | 472.1718 | 454.1610 | 455.1410 | 236.5840 | D | 815.4872 | 797.4767 | 798.4603 | 408.2436 | 7 |
| 6 | 88.0398 | 587.1996 | 569.1888 | 570.1679 | 294.0974 | D | 700.4603 | 682.4498 | 683.4333 | 350.7302 | 6 |
| 7 | 86.0976 | 700.2827 | 682.2684 | 683.2520 | 350.6395 | L | 585.4333 | 567.4228 | 568.4064 | 293.2167 | 5 |
| 8 | 72.0812 | 799.3474 | 781.3368 | 782.3204 | 400.1737 | V | 472.3493 | 454.3387 | 455.3223 | 236.6747 | 4 |
| 9 | 86.0969 | 912.4314 | 894.4208 | 895.4044 | 456.7157 | L | 373.2809 | 355.2703 | 356.2539 | 187.1404 | 3 |
| 10 | 86.0969 | 1025.5155 | 1007.5049 | 1008.4885 | 513.2578 | L | 260.1986 | 242.1863 | 243.1698 | 130.5984 | 2 |
| 11 | 101.1078 | | | | | K | 147.1144 | 129.1029 | 130.0872 | 74.0564 | 1 |

MW 1175.5925 LC-MS/MS ALC=69%



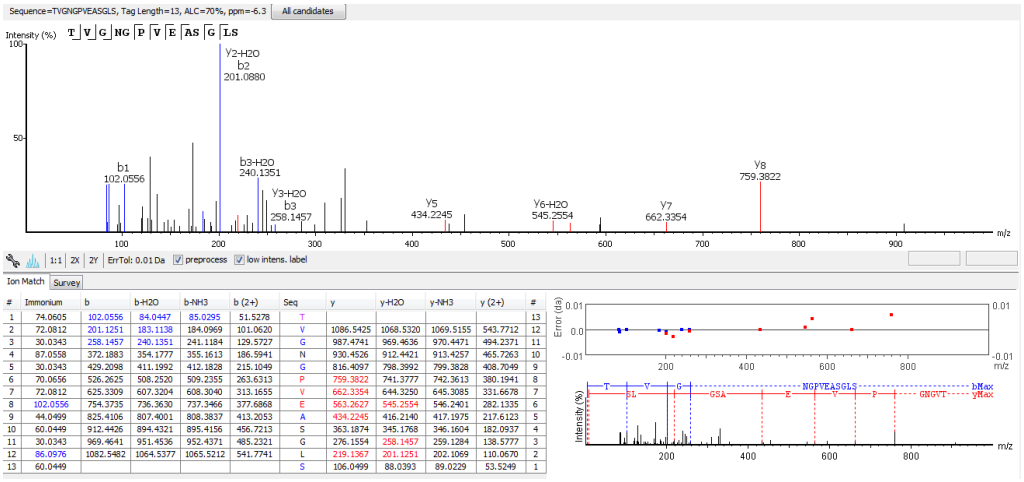

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 74.0605 | 102.0560 | 84.0455 | 85.0293 | 51.5278 | T | | | | | 11 |
| 2 | 30.0343 | 159.0771 | 141.0664 | 142.0500 | 80.0385 | G | 1075.5530 | 1057.5425 | 1058.5260 | 538.2765 | 10 |
| 3 | 44.0499 | 230.1167 | 212.1044 | 213.0877 | 115.5570 | A | 1018.5316 | 1000.5210 | 1001.5046 | 509.7658 | 9 |
| 4 | 102.0560 | 359.1566 | 341.1461 | 342.1297 | 180.0783 | E | 947.4944 | 929.4839 | 930.4675 | 474.2472 | 8 |
| 5 | 30.0343 | 416.1781 | 398.1691 | 399.1511 | 208.5891 | G | 818.4519 | 800.4413 | 801.4249 | 409.7260 | 7 |
| 6 | 101.1082 | 544.2731 | 526.2625 | 527.2461 | 272.6365 | K | 761.4277 | 743.4199 | 744.4034 | 381.2152 | 6 |
| 7 | 72.0812 | 643.3415 | 625.3309 | 626.3130 | 322.1707 | V | 633.3439 | 615.3249 | 616.3085 | 317.1677 | 5 |
| 8 | 110.0720 | 780.4004 | 762.3898 | 763.3734 | 390.7002 | H | 534.2694 | 516.2565 | 517.2401 | 267.6335 | 4 |
| 9 | 60.0449 | 867.4324 | 849.4219 | 850.4055 | 434.2162 | S | 397.2081 | 379.1976 | 380.1812 | 199.1093 | 3 |
| 10 | 136.0766 | 1030.4957 | 1012.4852 | 1013.4688 | 515.7479 | Y | 310.1780 | 292.1673 | 293.1485 | 155.5880 | 2 |
| 11 | 101.1078 | | | | | K | 147.1132 | 129.1028 | 130.0858 | 74.0564 | 1 |

MW 1179.6039 LC-MS/MS ALC=86%



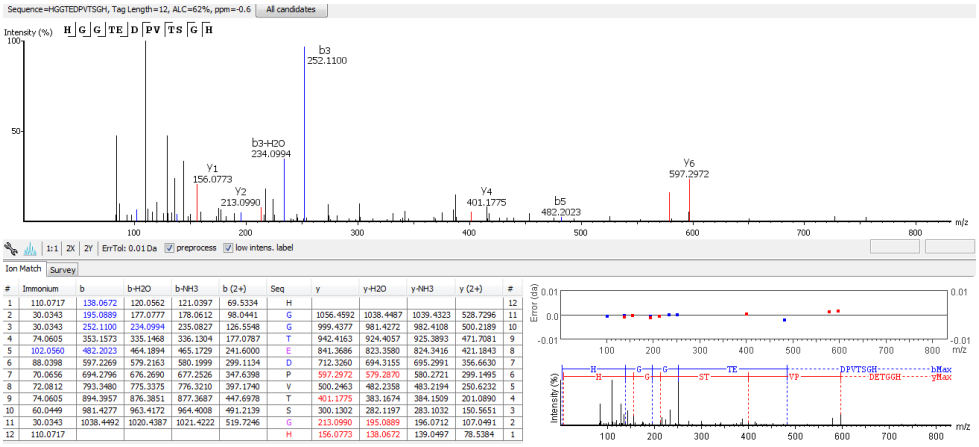

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 143.1183 | 171.1137 | 153.1028 | 154.0869 | 86.0567 | K(+42.01) | | | | | 11 |
| 2 | 44.0499 | 242.1506 | 224.1399 | 225.1235 | 121.5752 | A | 1010.5153 | 992.5047 | 993.4883 | 505.7576 | 10 |
| 3 | 70.0656 | 339.2032 | 321.1927 | 322.1762 | 170.1016 | P | 939.4740 | 921.4676 | 922.4512 | 470.2435 | 9 |
| 4 | 70.0656 | 436.2560 | 418.2454 | 419.2290 | 218.6280 | P | 842.4299 | 824.4148 | 825.3984 | 421.7127 | 8 |
| 5 | 70.0656 | 533.3087 | 515.2982 | 516.2817 | 267.1544 | P | 745.3744 | 727.3583 | 728.3456 | 373.1863 | 7 |
| 6 | 72.0812 | 632.3771 | 614.3666 | 615.3502 | 316.6886 | V | 648.3199 | 630.3093 | 631.2929 | 324.6599 | 6 |
| 7 | 44.0499 | 703.4142 | 685.4037 | 686.3873 | 352.2071 | A | 549.2515 | 531.2409 | 532.2245 | 275.1257 | 5 |
| 8 | 88.0398 | 818.4412 | 800.4306 | 801.4142 | 409.7206 | D | 478.2121 | 460.2038 | 461.1874 | 239.6072 | 4 |
| 9 | 88.0398 | 933.4681 | 915.4576 | 916.4412 | 467.2341 | D | 363.1873 | 345.1768 | 346.1604 | 182.0937 | 3 |
| 10 | 74.0605 | 1034.5159 | 1016.5053 | 1017.4889 | 517.7579 | T | 248.1616 | 230.1518 | 231.1335 | 124.5802 | 2 |
| 11 | 101.1078 | | | | | K | 147.1138 | 129.1030 | 130.0869 | 74.0564 | 1 |

MW 1186.5763 LC-MS/MS ALC=70%



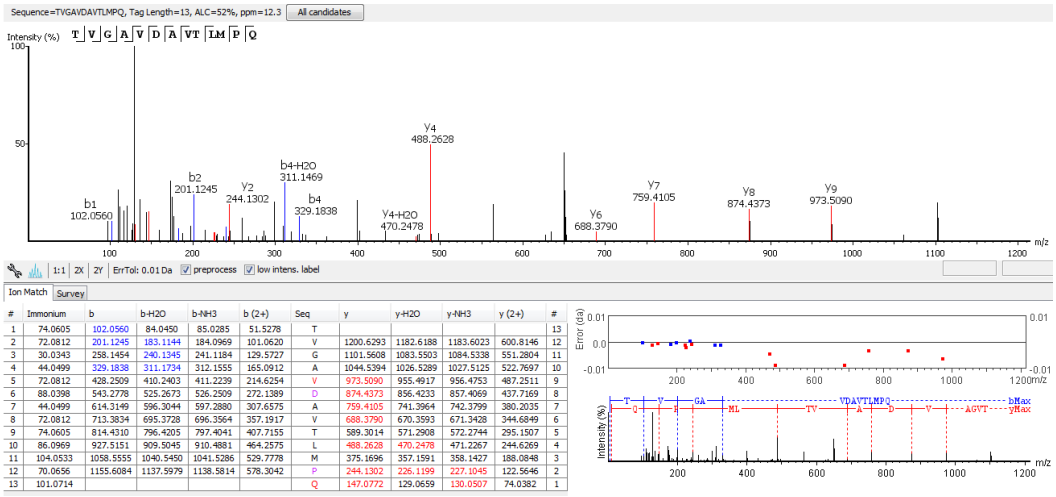

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 74.0605 | 102.0556 | 84.0447 | 85.0295 | 51.5278 | T | | | | | 13 |
| 2 | 72.0812 | 201.1251 | 183.1138 | 184.0969 | 101.0620 | V | 1086.5425 | 1068.5320 | 1069.5155 | 543.7712 | 12 |
| 3 | 30.0343 | 258.1457 | 240.1351 | 241.1184 | 129.5727 | G | 987.4741 | 969.4636 | 970.4471 | 494.2371 | 11 |
| 4 | 87.0558 | 372.1883 | 354.1777 | 355.1613 | 186.5941 | N | 930.4526 | 912.4421 | 913.4257 | 465.7263 | 10 |
| 5 | 30.0343 | 429.2098 | 411.1992 | 412.1828 | 215.1049 | G | 816.4097 | 798.3992 | 799.3828 | 408.7049 | 9 |
| 6 | 70.0656 | 526.2625 | 508.2520 | 509.2355 | 263.6313 | P | 759.3822 | 741.3777 | 742.3613 | 380.1941 | 8 |
| 7 | 72.0812 | 625.3309 | 607.3204 | 608.3040 | 313.1655 | V | 662.3354 | 644.3250 | 645.3085 | 331.6678 | 7 |
| 8 | 102.0556 | 754.3735 | 736.3630 | 737.3466 | 377.6868 | E | 563.2627 | 545.2554 | 546.2401 | 282.1335 | 6 |
| 9 | 44.0499 | 825.4106 | 807.4001 | 808.3837 | 413.2053 | A | 434.2245 | 416.2140 | 417.1975 | 217.6123 | 5 |
| 10 | 60.0449 | 912.4426 | 894.4321 | 895.4156 | 456.7213 | S | 363.1874 | 345.1768 | 346.1604 | 182.0937 | 4 |
| 11 | 30.0343 | 969.4641 | 951.4536 | 952.4371 | 485.2321 | G | 276.1554 | 258.1457 | 259.1284 | 138.5777 | 3 |
| 12 | 86.0976 | 1082.5482 | 1064.5377 | 1065.5212 | 541.7741 | L | 219.1367 | 201.1251 | 202.1069 | 110.0670 | 2 |
| 13 | 60.0449 | | | | | S | 106.0499 | 88.0393 | 89.0229 | 53.5249 | 1 |

MW 1192.5108 LC-MS/MS ALC=62%



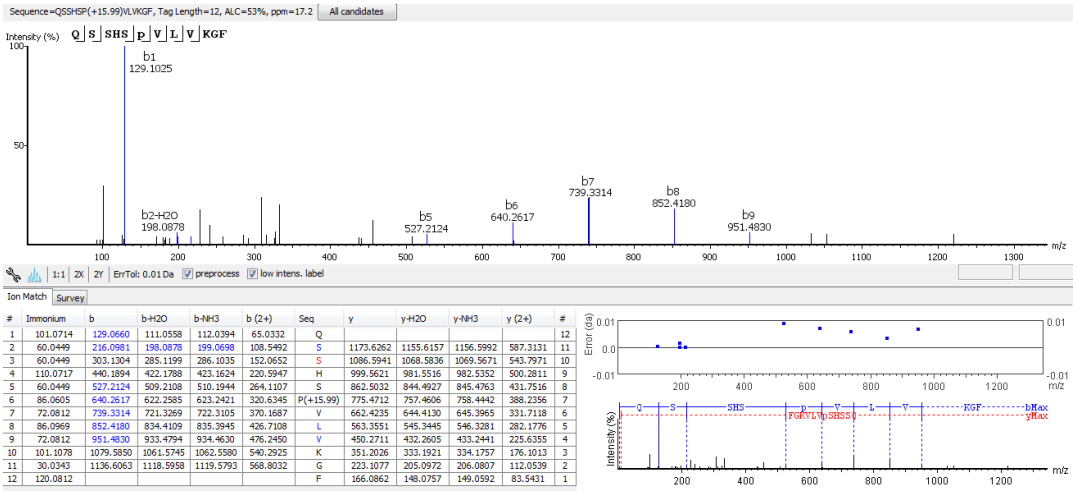

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 110.0717 | 138.0672 | 120.0562 | 121.0397 | 69.5334 | H | | | | | 12 |
| 2 | 30.0343 | 195.0889 | 177.0777 | 178.0612 | 98.0441 | G | 1056.4592 | 1038.4487 | 1039.4323 | 528.7296 | 11 |
| 3 | 30.0343 | 252.1100 | 234.0994 | 235.0827 | 126.5548 | G | 999.4377 | 981.4272 | 982.4108 | 500.2189 | 10 |
| 4 | 74.0605 | 353.1573 | 335.1468 | 336.1304 | 177.0787 | T | 942.4163 | 924.4057 | 925.3893 | 471.7081 | 9 |
| 5 | 102.0560 | 482.2023 | 464.1894 | 465.1729 | 241.6000 | E | 841.3686 | 823.3580 | 824.3416 | 421.1843 | 8 |
| 6 | 88.0398 | 597.2269 | 579.2163 | 580.1999 | 299.1134 | D | 712.3260 | 694.3155 | 695.2991 | 356.6630 | 7 |
| 7 | 70.0656 | 694.2796 | 676.2690 | 677.2526 | 347.6398 | P | 597.2972 | 579.2870 | 580.2721 | 299.1495 | 6 |
| 8 | 72.0812 | 793.3480 | 775.3375 | 776.3210 | 397.1740 | V | 500.2463 | 482.2358 | 483.2194 | 250.6232 | 5 |
| 9 | 74.0605 | 894.3957 | 876.3851 | 877.3687 | 447.6978 | T | 401.1775 | 383.1674 | 384.1509 | 201.0890 | 4 |
| 10 | 60.0449 | 981.4277 | 963.4172 | 964.4008 | 491.2139 | S | 300.1302 | 282.1197 | 283.1032 | 150.5651 | 3 |
| 11 | 30.0343 | 1038.4492 | 1020.4387 | 1021.4222 | 519.7246 | G | 213.0990 | 195.0889 | 196.0712 | 107.0491 | 2 |
| 12 | 110.0717 | | | | | H | 156.0773 | 138.0672 | 139.0497 | 78.5384 | 1 |

MW 1300.6790 LC-MS/MS ALC=52%



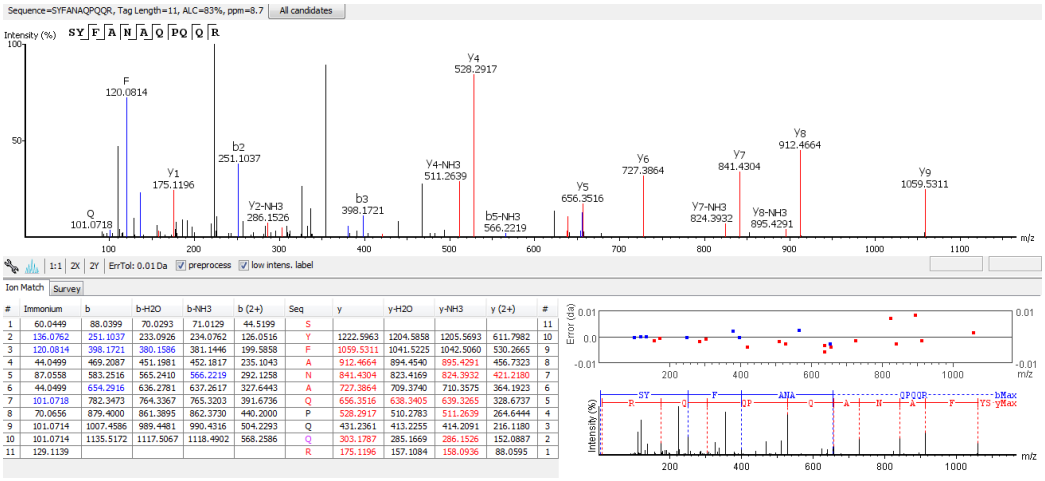

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 74.0605 | 102.0560 | 84.0450 | 85.0285 | 51.5278 | T | | | | | 13 |
| 2 | 72.0812 | 201.1245 | 183.1144 | 184.0969 | 101.0620 | V | 1200.6293 | 1182.6188 | 1183.6023 | 600.8146 | 12 |
| 3 | 30.0343 | 258.1454 | 240.1345 | 241.1184 | 129.5727 | G | 1101.5608 | 1083.5503 | 1084.5338 | 551.2804 | 11 |
| 4 | 44.0499 | 329.1838 | 311.1734 | 312.1555 | 165.0912 | A | 1044.5394 | 1026.5289 | 1027.5125 | 522.7697 | 10 |
| 5 | 72.0812 | 428.2509 | 410.2403 | 411.2239 | 214.6254 | V | 973.5090 | 955.4917 | 956.4753 | 487.2511 | 9 |
| 6 | 88.0398 | 543.2778 | 525.2673 | 526.2509 | 272.1389 | D | 874.4373 | 856.4233 | 857.4069 | 437.7169 | 8 |
| 7 | 44.0499 | 614.3149 | 596.3044 | 597.2880 | 307.6575 | A | 759.4105 | 741.3964 | 742.3799 | 380.2035 | 7 |
| 8 | 72.0812 | 713.3834 | 695.3728 | 696.3564 | 357.1917 | V | 688.3790 | 670.3593 | 671.3428 | 344.6849 | 6 |
| 9 | 74.0605 | 814.4310 | 796.4205 | 797.4041 | 407.7155 | T | 589.3014 | 571.2908 | 572.2744 | 295.1507 | 5 |
| 10 | 86.0969 | 927.5151 | 909.5045 | 910.4881 | 464.2575 | L | 488.2628 | 470.2478 | 471.2267 | 244.6269 | 4 |
| 11 | 104.0533 | 1058.5555 | 1040.5450 | 1041.5286 | 529.7778 | M | 375.1696 | 357.1591 | 358.1427 | 188.0848 | 3 |
| 12 | 70.0656 | 1155.6084 | 1137.5979 | 1138.5814 | 578.3042 | P | 244.1302 | 226.1199 | 227.1045 | 122.5646 | 2 |
| 13 | 101.0714 | | | | | Q | 147.0772 | 129.0659 | 130.0507 | 74.0382 | 1 |

MW 1300.6790 LC-MS/MS ALC=53%



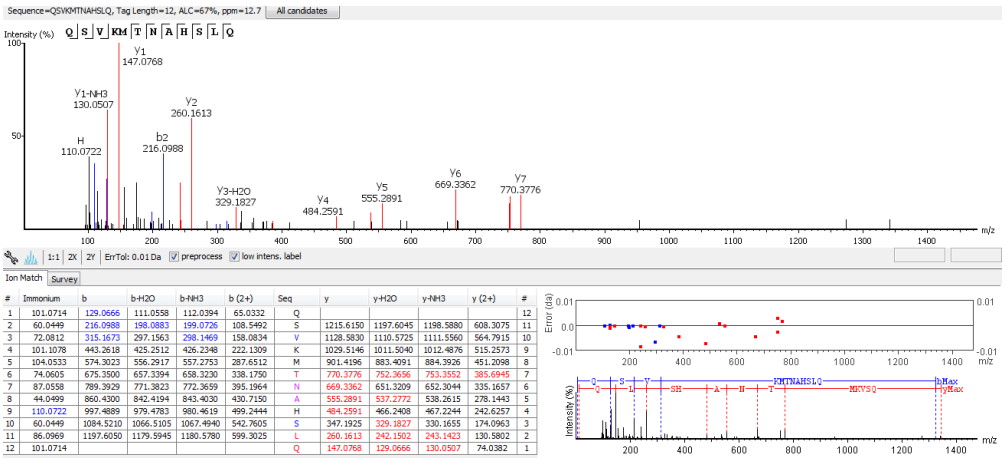

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 101.0714 | 129.0660 | 111.0558 | 112.0394 | 65.0332 | Q | | | | | 12 |
| 2 | 60.0449 | 216.0981 | 198.0878 | 199.0698 | 108.5492 | S | 1173.6262 | 1155.6157 | 1156.5992 | 587.3131 | 11 |
| 3 | 60.0449 | 303.1304 | 285.1199 | 286.1035 | 152.0652 | S | 1086.5941 | 1068.5836 | 1069.5671 | 543.7971 | 10 |
| 4 | 110.0717 | 440.1894 | 422.1788 | 423.1624 | 220.5947 | H | 999.5621 | 981.5516 | 982.5352 | 500.2811 | 9 |
| 5 | 60.0449 | 527.2124 | 509.2108 | 510.1944 | 264.1107 | S | 862.5032 | 844.4927 | 845.4763 | 431.7516 | 8 |
| 6 | 86.0605 | 640.2617 | 622.2585 | 623.2421 | 320.6345 | P(+15.99) | 775.4712 | 757.4606 | 758.4442 | 388.2356 | 7 |
| 7 | 72.0812 | 739.3314 | 721.3269 | 722.3105 | 370.1687 | V | 662.4235 | 644.4130 | 645.3965 | 331.7118 | 6 |
| 8 | 86.0969 | 852.4180 | 834.4109 | 835.3945 | 426.7108 | L | 563.3551 | 545.3445 | 546.3281 | 282.1776 | 5 |
| 9 | 72.0812 | 951.4830 | 933.4794 | 934.4630 | 476.2450 | V | 450.2711 | 432.2605 | 433.2441 | 225.6355 | 4 |
| 10 | 101.1078 | 1079.5850 | 1061.5745 | 1062.5580 | 540.2925 | K | 351.2026 | 333.1921 | 334.1757 | 176.1013 | 3 |
| 11 | 30.0343 | 1136.6063 | 1118.5958 | 1119.5793 | 568.8032 | G | 223.1077 | 205.0972 | 206.0807 | 112.0539 | 2 |
| 12 | 120.0812 | | | | | F | 166.0862 | 148.0757 | 149.0592 | 83.5431 | 1 |

MW 1308.6282 LC-MS/MS ALC=83%




| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 60.0449 | 88.0399 | 70.0293 | 71.0129 | 44.5199 | S | | | | | 11 |
| 2 | 136.0762 | 251.1037 | 233.0926 | 234.0762 | 126.0516 | Y | 1222.5963 | 1204.5858 | 1205.5693 | 611.7982 | 10 |
| 3 | 120.0814 | 398.1721 | 380.1586 | 381.1446 | 199.5858 | F | 1059.5311 | 1041.5225 | 1042.5060 | 530.2665 | 9 |
| 4 | 44.0499 | 469.2087 | 451.1981 | 452.1817 | 235.1043 | A | 912.4664 | 894.4540 | 895.4291 | 456.7323 | 8 |
| 5 | 87.0558 | 583.2516 | 565.2410 | 566.2219 | 292.1258 | N | 841.4304 | 823.4169 | 824.3932 | 421.2180 | 7 |
| 6 | 44.0499 | 654.2916 | 636.2781 | 637.2617 | 327.6443 | A | 727.3864 | 709.3740 | 710.3575 | 364.1923 | 6 |
| 7 | 101.0718 | 782.3473 | 764.3367 | 765.3203 | 391.6736 | Q | 656.3516 | 638.3405 | 639.3265 | 328.6737 | 5 |
| 8 | 70.0656 | 879.4000 | 861.3895 | 862.3730 | 440.2000 | P | 528.2917 | 510.2783 | 511.2639 | 264.6444 | 4 |
| 9 | 101.0714 | 1007.4586 | 989.4481 | 990.4316 | 504.2293 | Q | 431.2361 | 413.2255 | 414.2091 | 216.1180 | 3 |
| 10 | 101.0714 | 1135.5172 | 1117.5067 | 1118.4902 | 568.2586 | Q | 303.1787 | 285.1669 | 286.1526 | 152.0887 | 2 |
| 11 | 129.1139 | | | | | R | 175.1196 | 157.1084 | 158.0936 | 88.0595 | 1 |

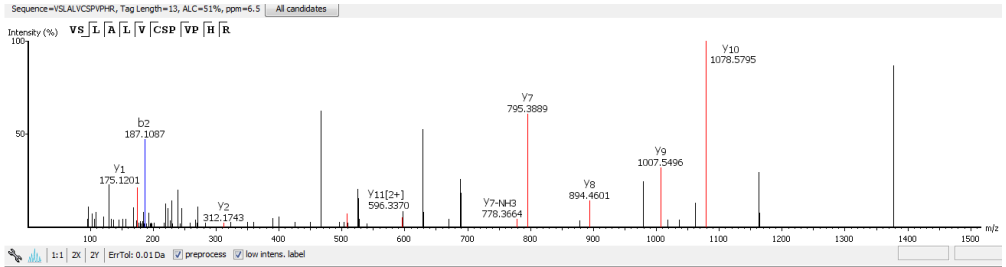
MW 1342.6713 LC-MS/MS ALC=67%



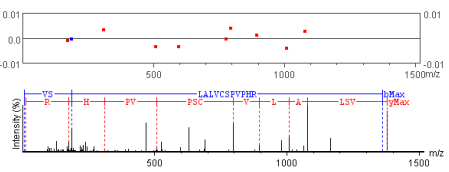

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 101.0714 | 129.0666 | 111.0558 | 112.0394 | 65.0332 | Q | | | | | 12 |
| 2 | 60.0449 | 216.0988 | 198.0883 | 199.0726 | 108.5492 | S | 1215.6150 | 1197.6045 | 1198.5880 | 608.3075 | 11 |
| 3 | 72.0812 | 315.1673 | 297.1563 | 298.1469 | 158.0834 | V | 1128.5830 | 1110.5725 | 1111.5560 | 564.7915 | 10 |
| 4 | 101.1078 | 443.2618 | 425.2512 | 426.2348 | 222.1309 | K | 1029.5146 | 1011.5040 | 1012.4876 | 515.2573 | 9 |
| 5 | 104.0533 | 574.3023 | 556.2917 | 557.2753 | 287.6512 | M | 901.4196 | 883.4091 | 884.3926 | 451.2098 | 8 |
| 6 | 74.0605 | 675.3500 | 657.3394 | 658.3230 | 338.1750 | T | 770.3776 | 752.3656 | 753.3552 | 385.6945 | 7 |
| 7 | 87.0558 | 789.3929 | 771.3823 | 772.3659 | 395.1964 | N | 669.3362 | 651.3209 | 652.3044 | 335.1657 | 6 |
| 8 | 44.0499 | 860.4300 | 842.4194 | 843.4030 | 430.7150 | A | 555.2891 | 537.2772 | 538.2615 | 278.1443 | 5 |
| 9 | 110.0722 | 997.4889 | 979.4783 | 980.4619 | 499.2444 | H | 484.2591 | 466.2408 | 467.2244 | 242.6257 | 4 |
| 10 | 60.0449 | 1084.5210 | 1066.5105 | 1067.4940 | 542.7605 | S | 347.1925 | 329.1827 | 330.1655 | 174.0963 | 3 |
| 11 | 86.0969 | 1197.6050 | 1179.5945 | 1180.5780 | 599.3025 | L | 260.1613 | 242.1502 | 243.1423 | 130.5802 | 2 |
| 12 | 101.0714 | | | | | Q | 147.0768 | 129.0666 | 130.0507 | 74.0382 | 1 |



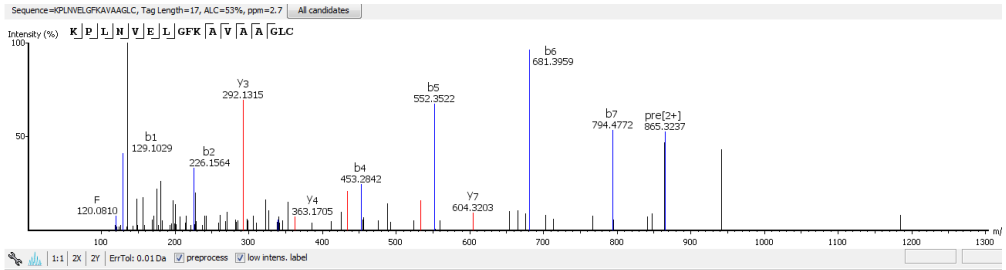

MW 1376.7577 LC-MS/MS ALC=51%




| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 72.0812 | 100.0762 | 82.0657 | 83.0492 | 50.5381 | V | | | | | 13 |
| 2 | 60.0449 | 187.1087 | 169.0977 | 170.0813 | 94.0541 | S | 1278.6986 | 1260.6881 | 1261.6716 | 639.8493 | 12 |
| 3 | 86.0969 | 300.1923 | 282.1818 | 283.1653 | 150.5962 | L | 1191.6666 | 1173.6561 | 1174.6396 | 596.3370 | 11 |
| 4 | 44.0499 | 371.2294 | 353.2189 | 354.2025 | 186.1147 | A | 1078.5795 | 1060.5720 | 1061.5555 | 539.7913 | 10 |
| 5 | 86.0969 | 484.3135 | 466.3029 | 467.2865 | 242.6567 | L | 1007.5496 | 989.5349 | 990.5185 | 504.2727 | 9 |
| 6 | 72.0812 | 583.3819 | 565.3713 | 566.3549 | 292.1909 | V | 894.4601 | 876.4509 | 877.4344 | 447.7307 | 8 |
| 7 | 76.0220 | 686.3911 | 668.3806 | 669.3641 | 343.6956 | C | 795.3889 | 777.3824 | 778.3664 | 398.1965 | 7 |
| 8 | 60.0449 | 773.4231 | 755.4125 | 756.3961 | 387.2115 | S | 692.3838 | 674.3732 | 675.3568 | 346.6919 | 6 |
| 9 | 70.0656 | 870.4759 | 852.4653 | 853.4489 | 435.7379 | P | 605.3517 | 587.3412 | 588.3248 | 303.1759 | 5 |
| 10 | 72.0812 | 969.5443 | 951.5338 | 952.5173 | 485.2722 | V | 508.3026 | 490.2885 | 491.2720 | 254.6495 | 4 |
| 11 | 70.0656 | 1066.5970 | 1048.5865 | 1049.5701 | 533.7985 | P | 409.2306 | 391.2200 | 392.2036 | 205.1153 | 3 |
| 12 | 110.0717 | 1203.6560 | 1185.6455 | 1186.6290 | 602.3280 | H | 312.1743 | 294.1673 | 295.1509 | 156.5889 | 2 |
| 13 | 129.1139 | | | | | R | 175.1201 | 157.1084 | 158.0919 | 88.0595 | 1 |




MW 1728.9683 LC-MS/MS ALC=53%




| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 101.1078 | 129.1029 | 111.0922 | 112.0758 | 65.0514 | K | | | | | 17 |
| 2 | 70.0656 | 226.1564 | 208.1450 | 209.1285 | 113.5778 | P | 1601.8719 | 1583.8615 | 1584.8450 | 801.4360 | 16 |
| 3 | 86.0969 | 339.2389 | 321.2290 | 322.2126 | 170.1198 | L | 1504.8191 | 1486.8086 | 1487.7921 | 752.9095 | 15 |
| 4 | 87.0558 | 453.2842 | 435.2720 | 436.2556 | 227.1413 | N | 1391.7351 | 1373.7246 | 1374.7081 | 696.3676 | 14 |
| 5 | 72.0812 | 552.3522 | 534.3404 | 535.3240 | 276.6755 | V | 1277.6921 | 1259.6816 | 1260.6652 | 639.3461 | 13 |
| 6 | 102.0554 | 681.3959 | 663.3829 | 664.3665 | 341.1967 | E | 1178.6238 | 1160.6133 | 1161.5968 | 589.8119 | 12 |
| 7 | 86.0969 | 794.4772 | 776.4670 | 777.4506 | 397.7388 | L | 1049.5812 | 1031.5707 | 1032.5542 | 525.2906 | 11 |
| 8 | 30.0343 | 851.4990 | 833.4885 | 834.4720 | 426.2495 | G | 936.4971 | 918.4865 | 919.4701 | 468.7485 | 10 |
| 9 | 120.0810 | 998.5674 | 980.5569 | 981.5405 | 499.7837 | F | 879.4756 | 861.4651 | 862.4487 | 440.2378 | 9 |
| 10 | 101.1078 | 1126.6624 | 1108.6519 | 1109.6354 | 563.8312 | K | 732.4072 | 714.3967 | 715.3802 | 366.7036 | 8 |
| 11 | 44.0499 | 1197.6995 | 1179.6890 | 1180.6725 | 599.3497 | A | 604.3203 | 586.3017 | 587.2853 | 302.6561 | 7 |
| 12 | 72.0812 | 1296.7679 | 1278.7574 | 1279.7410 | 648.8840 | V | 533.2792 | 515.2646 | 516.2482 | 267.1376 | 6 |
| 13 | 44.0499 | 1367.8051 | 1349.7946 | 1350.7781 | 684.4025 | A | 434.2051 | 416.1962 | 417.1798 | 217.6034 | 5 |
| 14 | 44.0499 | 1438.8422 | 1420.8317 | 1421.8152 | 719.9211 | A | 363.1705 | 345.1591 | 346.1427 | 182.0848 | 4 |
| 15 | 30.0343 | 1495.8636 | 1477.8531 | 1478.8367 | 748.4318 | G | 292.1315 | 274.1220 | 275.1056 | 146.5663 | 3 |
| 16 | 86.0969 | 1608.9476 | 1590.9371 | 1591.9207 | 804.9738 | L | 235.1111 | 217.1005 | 218.0841 | 118.0555 | 2 |

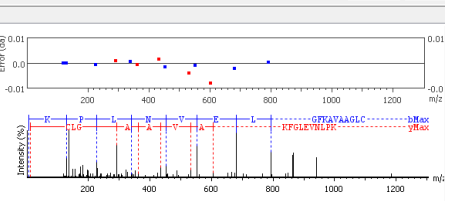
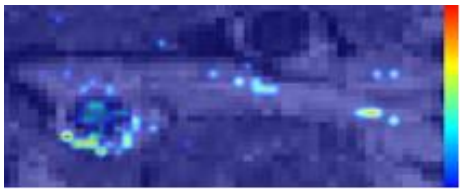
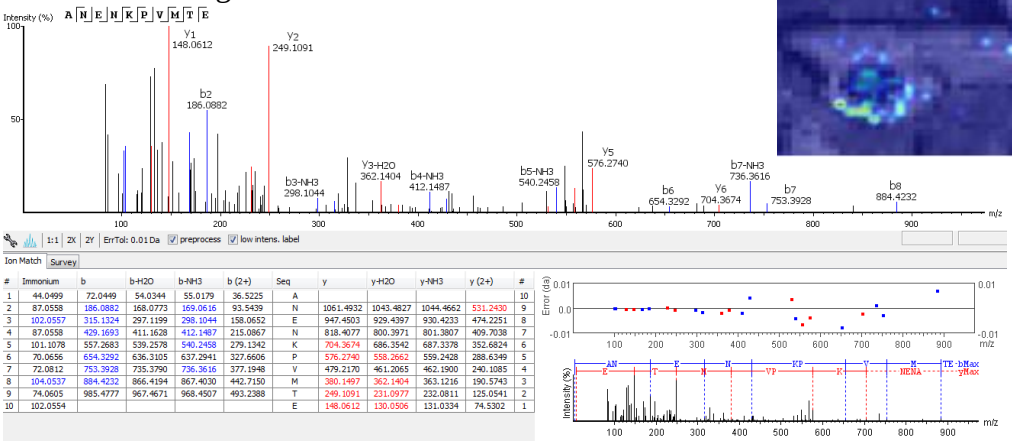


**Figure S3:** Annotated MS/MS spectra of unique peptides.

*m/z* 1132.53084, Wound-inducible Basic Family Protein MALDI-MSI

LC-MS/MS -10lgP=24.3

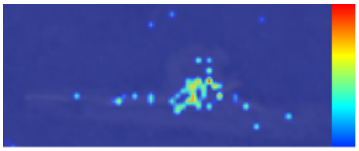


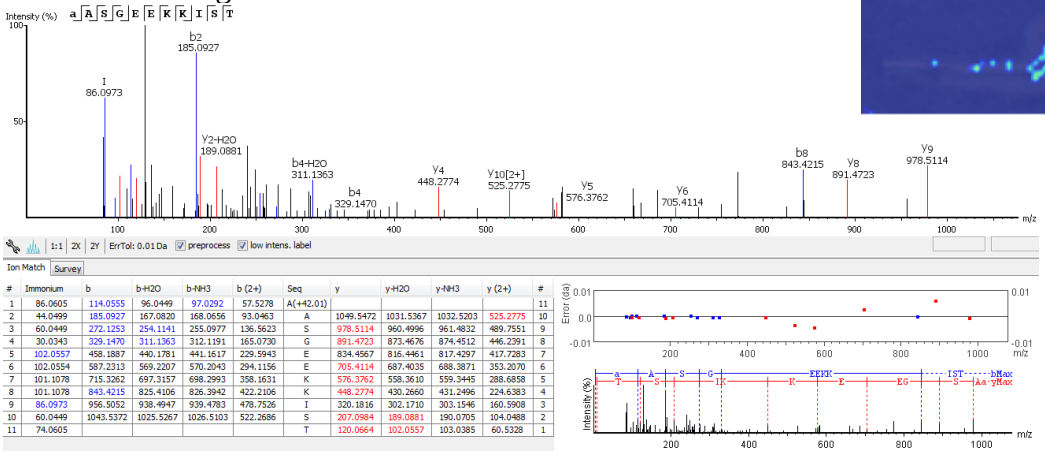


| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 44.0499 | 72.0449 | 54.0344 | 55.0179 | 36.5225 | A | | | | | 10 |
| 2 | 87.0558 | 186.0882 | 168.0773 | 169.0616 | 93.5439 | N | 1061.4932 | 1043.4827 | 1044.4662 | 531.2430 | 9 |
| 3 | 102.0557 | 315.1324 | 297.1199 | 298.1044 | 158.0652 | E | 947.4503 | 929.4397 | 930.4233 | 474.2251 | 8 |
| 4 | 87.0558 | 429.1693 | 411.1628 | 412.1487 | 215.0867 | N | 818.4077 | 800.3971 | 801.3807 | 409.7038 | 7 |
| 5 | 101.1078 | 557.2683 | 539.2578 | 540.2458 | 279.1342 | K | 704.3674 | 686.3542 | 687.3378 | 352.6824 | 6 |
| 6 | 70.0656 | 654.3292 | 636.3105 | 637.2941 | 327.6606 | P | 576.2740 | 558.2662 | 559.2428 | 288.6349 | 5 |
| 7 | 72.0812 | 753.3928 | 735.3790 | 736.3616 | 377.1948 | V | 479.2170 | 461.2065 | 462.1900 | 240.1085 | 4 |
| 8 | 104.0537 | 884.4232 | 866.4194 | 867.4030 | 442.7150 | M | 380.1497 | 362.1404 | 363.1216 | 190.5743 | 3 |
| 9 | 74.0605 | 985.4777 | 967.4671 | 968.4507 | 493.2388 | T | 249.1091 | 231.0977 | 232.0811 | 125.0541 | 2 |
| 10 | 102.0554 | | | | | E | 148.0612 | 130.0506 | 131.0334 | 74.5302 | 1 |

*m/z* 1162.59554, Low-temperature Inducible/ Nodulin-like Protein MALDI-MSI

LC-MS/MS -10lgP=29.2

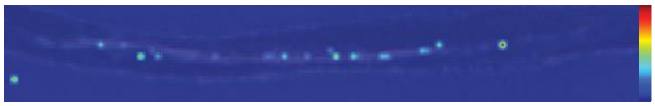


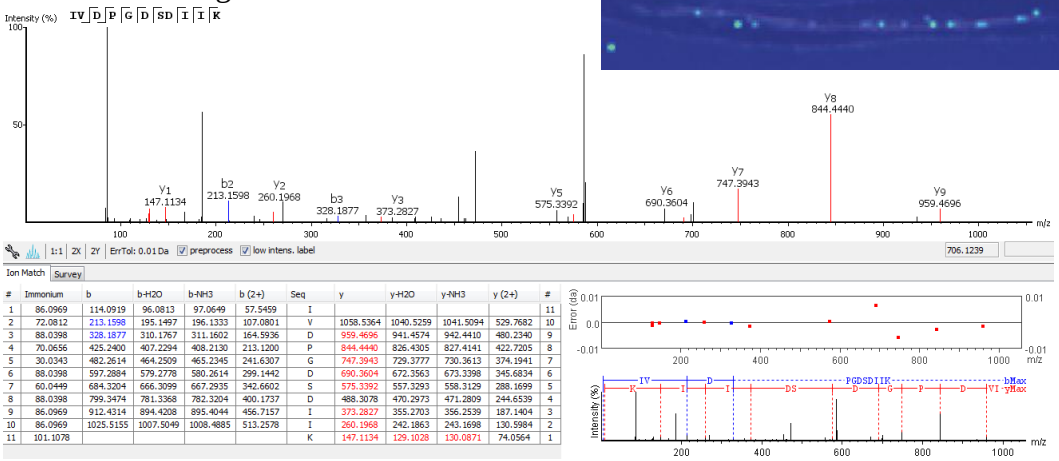


| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 86.0605 | 114.0555 | 96.0449 | 97.0292 | 57.5278 | A(+42.01) | | | | | 11 |
| 2 | 44.0499 | 185.0927 | 167.0820 | 168.0656 | 93.0463 | A | 1049.5472 | 1031.5367 | 1032.5203 | 525.2775 | 10 |
| 3 | 60.0449 | 272.1253 | 254.1141 | 255.0977 | 136.5623 | S | 978.5114 | 960.4996 | 961.4832 | 489.7551 | 9 |
| 4 | 30.0343 | 329.1470 | 311.1363 | 312.1191 | 165.0730 | G | 891.4723 | 873.4676 | 874.4512 | 446.2391 | 8 |
| 5 | 102.0557 | 458.1887 | 440.1781 | 441.1617 | 229.5943 | E | 834.4567 | 816.4461 | 817.4297 | 417.7283 | 7 |
| 6 | 102.0554 | 587.2313 | 569.2207 | 570.2043 | 294.1156 | E | 705.4114 | 687.4035 | 688.3871 | 353.2070 | 6 |
| 7 | 101.1078 | 715.3262 | 697.3157 | 698.2993 | 358.1631 | K | 576.3762 | 558.3610 | 559.3445 | 288.6858 | 5 |
| 8 | 101.1078 | 843.4215 | 825.4106 | 826.3942 | 422.2106 | K | 448.2774 | 430.2660 | 431.2496 | 224.6383 | 4 |
| 9 | 86.0973 | 956.5052 | 938.4947 | 939.4783 | 478.7526 | I | 320.1816 | 302.1710 | 303.1546 | 160.5908 | 3 |
| 10 | 60.0449 | 1043.5372 | 1025.5267 | 1026.5103 | 522.2686 | S | 207.0984 | 189.0881 | 190.0705 | 104.0488 | 2 |
| 11 | 74.0605 | | | | | T | 120.0664 | 102.0557 | 103.0385 | 60.5328 | 1 |

*m/z* 1171.62114, 60S Ribosomal Protein L30 MALDI-MSI

LC-MS/MS -10lgP=27.54

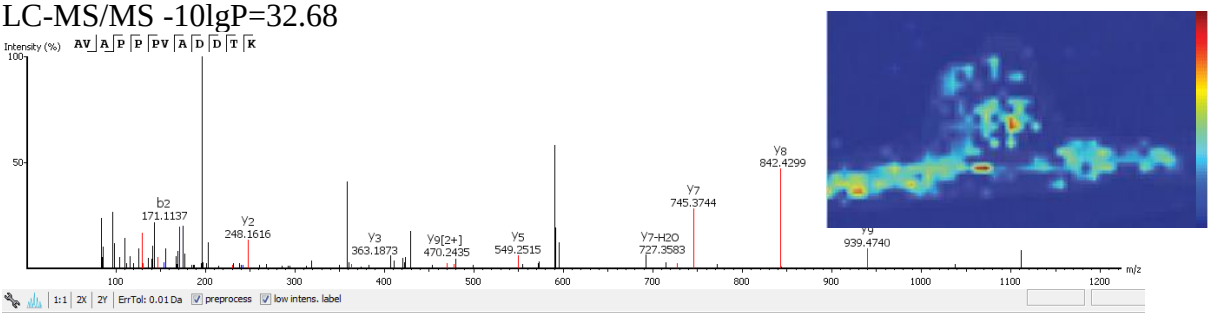




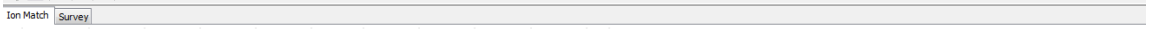
| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 86.0969 | 114.0919 | 96.0813 | 97.0649 | 57.5459 | I | | | | | 11 |
| 2 | 72.0812 | 213.1598 | 195.1497 | 196.1333 | 107.0801 | V | 1058.5364 | 1040.5259 | 1041.5094 | 529.7682 | 10 |
| 3 | 88.0398 | 328.1877 | 310.1767 | 311.1602 | 164.5936 | D | 959.4696 | 941.4574 | 942.4410 | 480.2340 | 9 |
| 4 | 70.0656 | 425.2400 | 407.2294 | 408.2130 | 213.1200 | P | 844.4440 | 826.4305 | 827.4141 | 422.7205 | 8 |
| 5 | 30.0343 | 482.2614 | 464.2509 | 465.2345 | 241.6307 | G | 747.3943 | 729.3777 | 730.3613 | 374.1941 | 7 |
| 6 | 88.0398 | 597.2884 | 579.2778 | 580.2614 | 299.1442 | D | 690.3604 | 672.3563 | 673.3398 | 345.6834 | 6 |
| 7 | 60.0449 | 684.3204 | 666.3099 | 667.2935 | 342.6602 | S | 575.3392 | 557.3293 | 558.3129 | 288.1699 | 5 |
| 8 | 88.0398 | 799.3474 | 781.3368 | 782.3204 | 400.1737 | D | 488.3078 | 470.2973 | 471.2809 | 244.6539 | 4 |
| 9 | 86.0969 | 912.4314 | 894.4208 | 895.4044 | 456.7157 | I | 373.2827 | 355.2703 | 356.2539 | 187.1404 | 3 |
| 10 | 86.0969 | 1025.5155 | 1007.5049 | 1008.4885 | 513.2578 | I | 260.1968 | 242.1863 | 243.1698 | 130.5984 | 2 |
| 11 | 101.1078 | | | | | K | 147.1134 | 129.1028 | 130.0871 | 74.0564 | 1 |

*m/z* 1180.62144, Carboxy-terminal Region Remorin MALDI-MSI

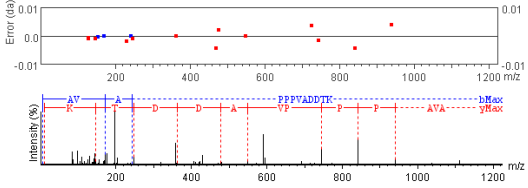
LC-MS/MS -10lgP=32.68





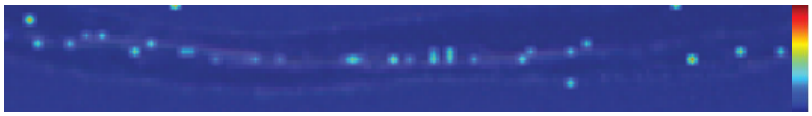



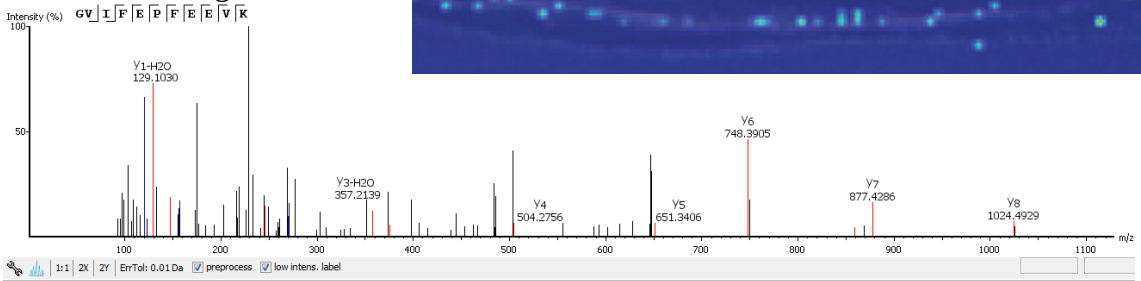

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 44.0499 | 72.0449 | 54.0344 | 55.0179 | 36.5225 | A | | | | | 12 |
| 2 | 72.0812 | 171.1137 | 153.1028 | 154.0869 | 86.0567 | V | 1109.5836 | 1091.5731 | 1092.5566 | 555.2918 | 11 |
| 3 | 44.0499 | 242.1506 | 224.1399 | 225.1235 | 121.5752 | A | 1010.5153 | 992.5047 | 993.4883 | 505.7576 | 10 |
| 4 | 70.0656 | 339.2032 | 321.1926 | 322.1762 | 170.1016 | P | 939.4740 | 921.4676 | 922.4512 | 470.2435 | 9 |
| 5 | 70.0656 | 436.2560 | 418.2454 | 419.2290 | 218.6280 | P | 842.4299 | 824.4148 | 825.3984 | 421.7127 | 8 |
| 6 | 70.0656 | 533.3087 | 515.2982 | 516.2817 | 267.1544 | P | 745.3744 | 727.3583 | 728.3456 | 373.1863 | 7 |
| 7 | 72.0812 | 632.3771 | 614.3666 | 615.3502 | 316.6886 | V | 648.3199 | 630.3093 | 631.2929 | 324.6599 | 6 |
| 8 | 44.0499 | 703.4142 | 685.4037 | 686.3873 | 352.2071 | A | 549.2515 | 531.2409 | 532.2245 | 275.1257 | 5 |
| 9 | 88.0398 | 818.4412 | 800.4306 | 801.4142 | 409.7206 | D | 478.2121 | 460.2038 | 461.1874 | 239.6072 | 4 |
| 10 | 88.0398 | 933.4681 | 915.4576 | 916.4412 | 467.2341 | D | 363.1873 | 345.1768 | 346.1604 | 182.0937 | 3 |
| 11 | 74.0605 | 1034.5157 | 1016.5052 | 1017.4888 | 517.7579 | T | 248.1616 | 230.1518 | 231.1335 | 124.5802 | 2 |
| 12 | 101.1078 | | | | | K | 147.1138 | 129.1030 | 130.0869 | 74.0564 | 1 |




*m/z* 1293.67314, Ferritin MALDI-MSI

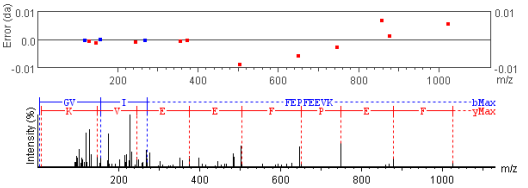
LC-MS/MS -10lgP=27.38



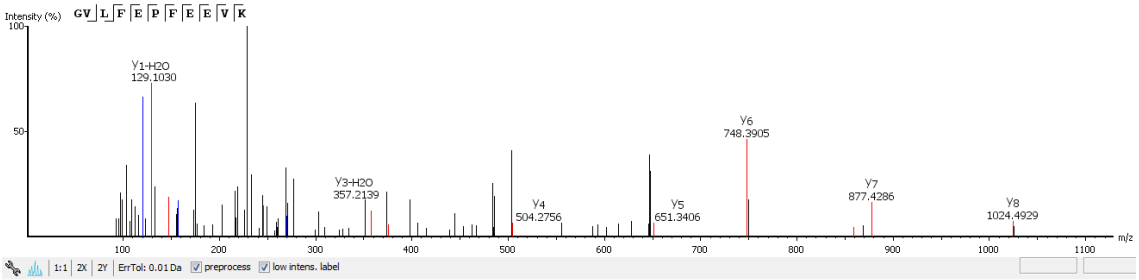


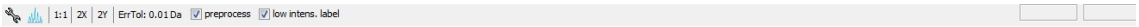


| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 30.0343 | 58.0293 | 40.0187 | 41.0023 | 29.5146 | G | | | | | 11 |
| 2 | 72.0812 | 157.0978 | 139.0871 | 140.0707 | 79.0489 | V | 1236.6510 | 1218.6405 | 1219.6240 | 618.8255 | 10 |
| 3 | 86.0969 | 270.1822 | 252.1712 | 253.1548 | 135.5909 | I | 1137.5825 | 1119.5720 | 1120.5555 | 569.2913 | 9 |
| 4 | 120.0817 | 417.2502 | 399.2396 | 400.2232 | 209.1251 | F | 1024.4929 | 1006.4880 | 1007.4716 | 512.7493 | 8 |
| 5 | 102.0554 | 546.2927 | 528.2822 | 529.2657 | 273.6464 | E | 877.4286 | 859.4128 | 860.4031 | 439.2151 | 7 |
| 6 | 70.0656 | 643.3455 | 625.3350 | 626.3185 | 322.1728 | P | 748.3905 | 730.3770 | 731.3605 | 374.6938 | 6 |
| 7 | 120.0812 | 790.4139 | 772.4033 | 773.3869 | 395.7069 | F | 651.3406 | 633.3242 | 634.3078 | 326.1674 | 5 |
| 8 | 102.0554 | 919.4565 | 901.4459 | 902.4295 | 460.2282 | E | 504.2756 | 486.2558 | 487.2394 | 252.6332 | 4 |
| 9 | 102.0554 | 1048.4990 | 1030.4885 | 1031.4720 | 524.7495 | E | 375.2242 | 357.2139 | 358.1968 | 188.1119 | 3 |
| 10 | 72.0812 | 1147.5675 | 1129.5570 | 1130.5405 | 574.2838 | V | 246.1823 | 228.1706 | 229.1542 | 123.5906 | 2 |
| 11 | 101.1078 | | | | | K | 147.1141 | 129.1030 | 130.0858 | 74.0564 | 1 |

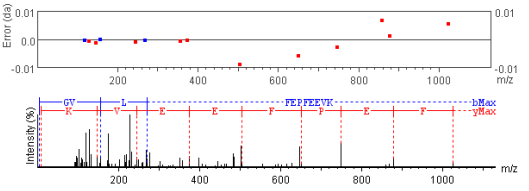



LC-MS/MS -10lgP=27.38

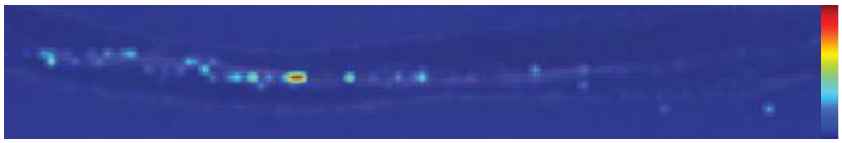


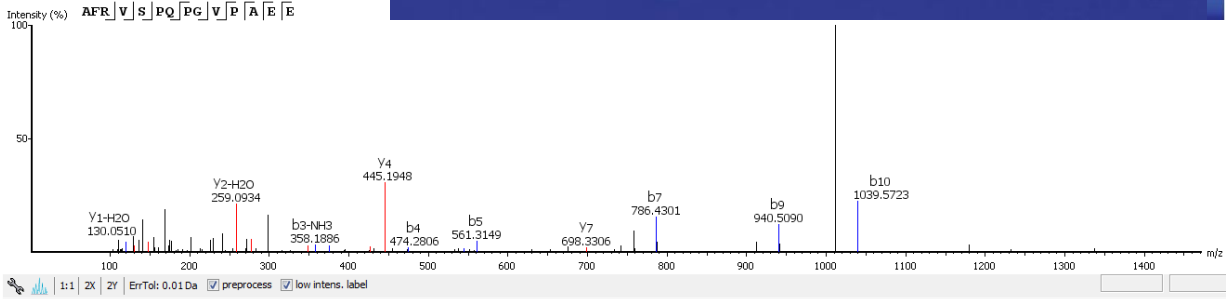

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 30.0343 | 58.0293 | 40.0187 | 41.0023 | 29.5146 | G | | | | | 11 |
| 2 | 72.0812 | 157.0978 | 139.0871 | 140.0707 | 79.0489 | V | 1236.6510 | 1218.6405 | 1219.6240 | 618.8255 | 10 |
| 3 | 86.0969 | 270.1822 | 252.1712 | 253.1548 | 135.5909 | L | 1137.5825 | 1119.5720 | 1120.5555 | 569.2913 | 9 |
| 4 | 120.0817 | 417.2502 | 399.2396 | 400.2232 | 209.1251 | F | 1024.4929 | 1006.4880 | 1007.4716 | 512.7493 | 8 |
| 5 | 102.0554 | 546.2927 | 528.2822 | 529.2657 | 273.6464 | E | 877.4286 | 859.4128 | 860.4031 | 439.2151 | 7 |
| 6 | 70.0656 | 643.3455 | 625.3350 | 626.3185 | 322.1728 | P | 748.3905 | 730.3770 | 731.3605 | 374.6938 | 6 |
| 7 | 120.0812 | 790.4139 | 772.4033 | 773.3869 | 395.7069 | F | 651.3406 | 633.3242 | 634.3078 | 326.1674 | 5 |
| 8 | 102.0554 | 919.4565 | 901.4459 | 902.4295 | 460.2282 | E | 504.2756 | 486.2558 | 487.2394 | 252.6332 | 4 |
| 9 | 102.0554 | 1048.4990 | 1030.4885 | 1031.4720 | 524.7495 | E | 375.2242 | 357.2139 | 358.1968 | 188.1119 | 3 |
| 10 | 72.0812 | 1147.5675 | 1129.5570 | 1130.5405 | 574.2838 | V | 246.1823 | 228.1706 | 229.1542 | 123.5906 | 2 |
| 11 | 101.1078 | | | | | K | 147.1141 | 129.1030 | 130.0858 | 74.0564 | 1 |



*m/z* 1483.75454, Ribulose Biphosphate Carboxylase/ Oxygenase Large Chain Domain Protein MALDI-MSI



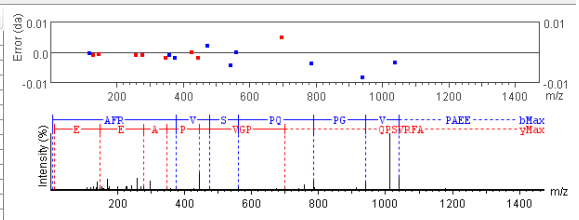
LC-MS/MS -10lgP=35.64




1:1 | 2X | 2Y | ErrTol: 0.01 Da | preprocess | low intens. label

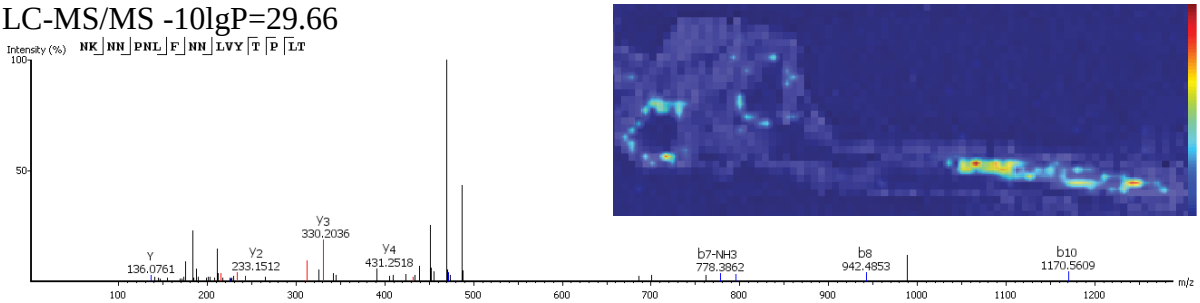
Ion Match | Survey

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 44.0499 | 72.0449 | 54.0344 | 55.0179 | 36.5225 | A | | | | | 14 |
| 2 | 120.0817 | 219.1133 | 201.1028 | 202.0863 | 110.0567 | F | 1412.7168 | 1394.7063 | 1395.6898 | 706.8584 | 13 |
| 3 | 129.1139 | 375.2163 | 357.2039 | 358.1886 | 188.1072 | R | 1265.6484 | 1247.6379 | 1248.6215 | 633.3242 | 12 |
| 4 | 72.0812 | 474.2806 | 456.2723 | 457.2559 | 237.6414 | V | 1109.5472 | 1091.5367 | 1092.5203 | 555.2736 | 11 |
| 5 | 60.0449 | 561.3149 | 543.3043 | 544.2924 | 281.1574 | S | 1010.4789 | 992.4683 | 993.4519 | 505.7394 | 10 |
| 6 | 70.0656 | 658.3676 | 640.3571 | 641.3406 | 329.6838 | P | 923.4468 | 905.4363 | 906.4199 | 462.2234 | 9 |
| 7 | 101.0714 | 786.4301 | 768.4156 | 769.3992 | 393.7131 | Q | 826.3940 | 808.3835 | 809.3671 | 413.6970 | 8 |
| 8 | 70.0656 | 883.4789 | 865.4684 | 866.4520 | 442.2395 | P | 698.3306 | 680.3250 | 681.3085 | 349.6678 | 7 |
| 9 | 30.0343 | 940.5090 | 922.4899 | 923.4734 | 470.7502 | G | 601.2827 | 583.2722 | 584.2557 | 301.1414 | 6 |
| 10 | 72.0812 | 1039.5723 | 1021.5583 | 1022.5419 | 520.2844 | V | 544.2613 | 526.2507 | 527.2343 | 272.6306 | 5 |
| 11 | 70.0656 | 1136.6216 | 1118.6111 | 1119.5946 | 568.8108 | P | 445.1948 | 427.1822 | 428.1659 | 223.0964 | 4 |
| 12 | 44.0499 | 1207.6587 | 1189.6482 | 1190.6317 | 604.3293 | A | 348.1422 | 330.1295 | 331.1131 | 174.5701 | 3 |
| 13 | 102.0554 | 1336.7013 | 1318.6908 | 1319.6743 | 668.8506 | E | 277.1042 | 259.0934 | 260.0760 | 139.0515 | 2 |
| 14 | 102.0554 | | | | | E | 148.0611 | 130.0510 | 131.0334 | 74.5302 | 1 |




*m/z* 1976.02434, Basic 7S Globulin/ Peptidase A1, Pepsin MALDI-MSI

LC-MS/MS -10lgP=29.66

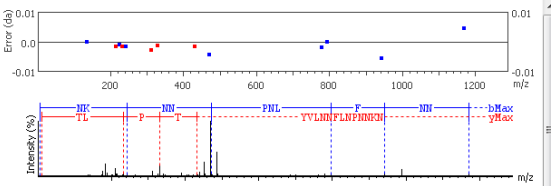





1:1 | 2X | 2Y | ErrTol: 0.01 Da | preprocess | low intens. label

Ion Match | Survey

| # | Immonium | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 87.0558 | 115.0508 | 97.0402 | 98.0238 | 58.0254 | N | | | | | 17 |
| 2 | 101.1078 | 243.1475 | 225.1351 | 226.1198 | 122.0729 | K | 1861.9806 | 1843.9701 | 1844.9536 | 931.4903 | 16 |
| 3 | 87.0558 | 357.1886 | 339.1781 | 340.1617 | 179.0943 | N | 1733.8857 | 1715.8752 | 1716.8588 | 867.4429 | 15 |
| 4 | 87.0558 | 471.2360 | 453.2210 | 454.2046 | 236.1158 | N | 1619.8428 | 1601.8323 | 1602.8158 | 810.4214 | 14 |
| 5 | 70.0656 | 568.2843 | 550.2737 | 551.2573 | 284.6422 | P | 1505.7998 | 1487.7893 | 1488.7728 | 753.3999 | 13 |
| 6 | 87.0558 | 682.3273 | 664.3167 | 665.3003 | 341.6636 | N | 1408.7471 | 1390.7366 | 1391.7201 | 704.8735 | 12 |
| 7 | 86.0969 | 795.4114 | 777.4008 | 778.3862 | 398.2057 | L | 1294.7041 | 1276.6936 | 1277.6771 | 647.8521 | 11 |
| 8 | 120.0812 | 942.4853 | 924.4692 | 925.4528 | 471.7399 | F | 1181.6201 | 1163.6096 | 1164.5931 | 591.3101 | 10 |
| 9 | 87.0558 | 1056.5227 | 1038.5122 | 1039.4957 | 528.7614 | N | 1034.5516 | 1016.5411 | 1017.5247 | 517.7758 | 9 |
| 10 | 87.0558 | 1170.5609 | 1152.5551 | 1153.5386 | 585.7828 | N | 920.5087 | 902.4982 | 903.4818 | 460.7544 | 8 |
| 11 | 86.0969 | 1283.6497 | 1265.6392 | 1266.6227 | 642.3248 | L | 806.4658 | 788.4553 | 789.4388 | 403.7329 | 7 |
| 12 | 72.0812 | 1382.7180 | 1364.7075 | 1365.6910 | 691.8590 | V | 693.3817 | 675.3712 | 676.3547 | 347.1909 | 6 |
| 13 | 136.0761 | 1545.7814 | 1527.7709 | 1528.7544 | 773.3907 | Y | 594.3134 | 576.3028 | 577.2864 | 297.6567 | 5 |
| 14 | 74.0605 | 1646.8291 | 1628.8186 | 1629.8021 | 823.9146 | T | 431.2518 | 413.2394 | 414.2230 | 216.1250 | 4 |
| 15 | 70.0656 | 1743.8818 | 1725.8713 | 1726.8549 | 872.4409 | P | 330.2036 | 312.1946 | 313.1754 | 165.6012 | 3 |
| 16 | 86.0969 | 1856.9658 | 1838.9553 | 1839.9388 | 928.9829 | L | 233.1512 | 215.1407 | 216.1226 | 117.0748 | 2 |
| 17 | 74.0605 | | | | | T | 120.0655 | 102.0550 | 103.0385 | 60.5328 | 1 |



# Chapter 5

# Enhancement of Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging of Nodule-Specific Cysteine-Rich Peptides with a Simple Washing Strategy

Adapted from: Keller C, Vu NQ, Kleven B, Chakraborty S, Jayaraman D, Maeda J, Mortier V, Sussman MR, Goormachtig S, Djordjevic MA, Ané JM, Li L. Enhancement of Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging of Nodule-Specific Cysteine-Rich Peptides with a Simple Washing Strategy. *The Plant Journal.* In preparation.

Keywords: MALDI-MSI; peptidomics; *Medicago truncatula*; Nodule-specific cysteine-rich peptides

**Summary**

Although MALDI-MSI is commonly used to determine the localization of metabolites in plant tissues, it is rarely used to analyze the distribution of endogenous peptides and even proteins in plants. Thus, a MALDI-MSI method was developed using *Medicago truncatula* root nodules to improve detection of endogenous peptides in plant tissue. A washing step of 70% ethanol for 30s, 100% ethanol for 30s, Carnoy solution for 2 min, and 100% ethanol for 30s was employed to fix proteins and remove lipid species, which ionize readily and can suppress protein signal. After washing, the ion intensities of numerous peptide *m/z* values were increased, revealing unique distributions of peptides in the nodules. Nodule-specific cysteine-rich (NCR) peptides were then identified via accurate mass matching to a predicted NCR peptide database and LC-MS/MS data obtained from a nodule peptide extraction. Sixteen *m/z* values observed in the MALDI-MSI results were matched to predicted NCR peptides, of which 3 were also identified by LC-MS/MS. Overall, this is a robust method that could easily be applied to the imaging of other signaling peptides across other plant species.

**Significance**

MALDI-MSI of plant peptides is a challenging and underdeveloped area, despite the essential roles of endogenous peptides in the plant. This work shows the optimization of a plant peptide MALDI-MSI procedure to detect Nodule-specific Cysteine-Rich peptides in *Medicago truncatula*, providing insights into the localization of these important peptides in the root nodules.

## Introduction

Endogenous peptides play many critical roles in plants, including growth and development, defense, nutrient deficiency, and symbiosis.[1, 2] The model legume, *Medicago truncatula* (Medicago), is highly studied due to its ability to form a symbiotic relationship with rhizobia bacteria. Rhizobia enter the roots and even plant cells via an endocytosis-like process. In parallel, cells of the root cortex divide and differentiate which leads to the development of a new organ, the root nodule. A mature nodule contains rhizobia differentiated into bacteroids within plant cells. These bacteroids are encapsulated within host-derived membranes, in compartments known as symbiosomes. In this differentiated form, the rhizobia can fix nitrogen *i.e.*, reduce atmospheric nitrogen into ammonia, a form of nitrogen usable to plants.[3]

Several endogenous peptides have been implicated in various stages of nodule development.[4] C-terminally encoded (CEP) peptides are involved in the development of nodules as well as lateral roots.[5, 6] CLAVATA3/embryo-surrounding region (CLE) peptides processed from pre-proteins encoded by *CLE12* and *CLE13* are involved in autoregulation of nodulation; both the genes are expressed in the dividing cells of a nodule primordium and later in the meristem of a mature nodule.[7, 8] Nodule-specific cysteine-rich (NCR) peptides are required for bacteroid differentiation.[9] Despite their critical roles in various biological processes, endogenous peptides can present a severe analytical challenge to analyze. They are often present at very low concentrations in the plant, are usually produced by peptidase action on larger polypeptide sequences and can have various post-translational modifications (PTMs).[8, 10]

Matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) is commonly used to visualize the spatial distribution of various biomolecules, including metabolites,[11] lipids,[12] neuropeptides,[13, 14] and proteins[15, 16] in various tissues. In MALDI-MSI, a

laser is fired at discrete points, or pixels, across tissue section, which is covered in a matrix to assist in ionization. Thus, a mass spectrum is created at each pixel in the tissue section, and the ion intensity for a specific *m/z* can be pulled out at each pixel to generate a heatmap showing the distribution of that *m/z* across the tissue section. Many reviews provide comprehensive details of sample preparation, instrumentation, and data analysis for MALDI-MSI.[17, 18]

In plants, small molecules have been investigated with MALDI-MSI in many different species.[19-23] For example, MALDI-MSI was used in *Medicago truncatula* root nodules to reveal metabolites involved in biological nitrogen fixation,[24, 25] in cotton embryos to reveal lipid distribution,[26] in maize to look at subcellular metabolomics,[27] and in *Arabidopsis thaliana* to localize glucosinolates, which have roles in plant defense, to the flowers and siliques.[28] Endogenous peptide localization with MALDI-MSI, however, is less frequently performed on plant tissues. Examples of plant peptide MSI include an initial study in Medicago of peptides present from 1000-4000 *m/z* in the roots and root nodules,[29] a study investigating the metabolite, peptide, and protein content of nematode induced gall tissues in tomato root,[30] proof of principle studies investigating peptides and small proteins in soybean cotyledons[31] and barley grains,[32] and cyclic endogenous peptides, the cyclotides, imaged in multiple species.[33, 34]

In this study, a MALDI-MSI experimental procedure was optimized for the detection and localization of NCR endogenous signaling peptides. Although peptides have previously been imaged with MALDI-MSI in Medicago,[29] no NCR peptides were identified, in part, because the mass range used (1000-4000 *m/z*) was not ideal for detection of NCR peptides. Here, a rapifleX MALDI TOF/TOF mass spectrometer was utilized to extend the mass range beyond 4000 *m/z* as the molecular weight of NCR peptides ranges from approximately 3-5 kDa.[35]

Furthermore, detection of NCR peptide signal was optimized by including a washing step in the sample preparation. The inclusion of additional sample preparation steps prior to matrix application to remove interfering species or modify compounds to increase their detection (*i.e.*, chemical derivatization or enzyme digestion) has been used to enhance the detection of various analyte(s) of interest.[36-38] For example, a simple strategy of washing tissues with organic solvents has been used to remove abundant lipid species to increase detection of proteins.[39] Here, a washing step was developed to increase NCR coverage as the removal of abundant lipid species alleviates the signal suppression, thus leading to increased signal intensity of higher molecular weight species. Furthermore, LC-MS/MS was employed to identify intact NCR endogenous peptides observed in the MALDI-MSI experiments. Overall, this workflow allows for the identification and localization of many NCR peptides in Medicago root nodules and could be applied to other plant species to study endogenous peptide distributions or changes in endogenous peptide distribution due to a mutation or a stressor.

**Experimental Procedures**

Plant Growth

For MALDI-MSI, *Medicago truncatula* wild-type (Jemalong A17) seeds were scarified with sulfuric acid for 15 minutes, surface-sterilized with 30% Clorox for 10 minutes, imbibed at room temperature for 4-5 hours and stratified at 4°C for 1-4 days. Overnight germinated seedlings were grown in growth pouches in modified Fahräeus medium [40] for five days and inoculated with *Sinorhizobium meliloti* (*Sm*1021) [41] at $OD_{600}$ 0.1, and grown for three weeks to obtain mature nodules. For LC-MS/MS, wild-type and *CLE13*-overexpressing seedlings were grown on modified

Fahräeus plates for five days and inoculated with *Sm*1021 at $OD_{600}$ 0.2. The seedlings were grown for two weeks and then processed for analysis.

To obtain the *35S:CLE13* line, *M. truncatula* Jemalong 5 young leaves from 4 week-old *in vitro* grown plants were used for *A. tumefaciens*-mediated transformation which was performed as described by Primard-Brisset et al. (2005). The binary vectors pB7WG2D-35S:MtCLE13 was electroporated in *A. tumefaciens* AGLO and a single colony was isolated after selection on YEB supplemented with 100 mg/L rifampicin and 100 mg/L spectinomycin. After plant transformation and cocultivation, transgenic tissue was selected on subsequently SHb3a (incubation), AgCIM (callogenesis), AgEIM (pro-embrygenesis), AgEMB (embryogenesis) and plant medium (plant development and rooting). All media, except the plant medium, contained between 200 to 800 mg/L augmentine during callogenesis and (pro) embryogenesis stages and 3 mg/L basta.[42, 43] *In vitro* regenerated transgenic shoots were further grown on plates for 1 month and then planted in earth and sand (3:1) in the growth chamber or shoot regenerated on PDM medium. Shoot regeneration was performed by cutting the shoot tip at the fourth internode from the shoot apex and transferring the sectioned shoot on PDM medium (http://www.isv.cnrsgif.fr/embo01/manuels/pdf/module2.pdf).

MALDI-MSI Sample Preparation

Root nodules were trimmed from the plant with 2-3 mm of the surrounding root, flash frozen on dry ice, and embedded in gelatin (100 mg/mL in Milli-Q ultrapure water, Millipure Sigma). Some nodules were processed with the fresh structure preserve setting on the Denator (Denator AB, Uppsala Science Park, Sweden) prior to embedding. Embedded nodules were saved in the -80$^{o}$C freezer until further processing. Nodules were sectioned at 16 µm thickness on a cryostat (Microm HM 525, Thermo Scientific) and thaw mounted on an indium tin oxide coated

glass slide (Bruker Daltonics, Billerica, MA). Solvents were from Fisher Scientific unless otherwise noted. Sections were either washed with a single 30 s wash of one of three solvents (ethanol (200 proof, Pharmco), chloroform, hexanes (mixture of isomers)) or with one of three gradient ethanol washes typically used in MALDI-MSI protein experiments (70% ethanol 30s, 100% ethanol 30 s; 70% ethanol 90s, 70% ethanol 90 s, 100% ethanol 90 s; 70% ethanol 30s, 100% ethanol 30 s; Carnoy's wash 2 min, 70% ethanol 30s). Ethanol was 200 proof (Pharmco), hexanes was ACS grade, chloroform was ACS grade, and Milli-Q ultrapure water was used. Sections were also left unwashed as controls. After washing, sections were dried in a vacuum desiccator for 15 minutes. 2,5-Dihydroxybenzoic acid (DHB, Acros Organics) matrix (40 mg/mL in 50% methanol, HPLC grade, 0.1% formic acid) was applied with a TM Sprayer (12 passes, 1250 velocity, 0.1 mL/min flow rate, CC pattern, 30 s dry, 10 psi, 3 mm spacing). Slides were saved at -20$^{o}$C until analysis.

MALDI-MSI Data Acquisition

Washed and Unwashed root nodule sections were analyzed on a rapifleX MALDI-TOF/TOF instrument (Bruker Daltonics, Billerica, MA) in both linear and reflectron mode. The reflectron method averaged 500 laser shots at 80% laser power and 5 kHz frequency from 1500-6000 *m/z* with 5.4x detector gain (2206 V replicate 1, and 2297 V replicate 2 and 3). Voltages were 20.000 kV for ion source 1, 2.720 kV for PIE, 11.800 kV for lens, 20.830 kV for reflector 1, 1.085 kV for reflector 2, and 8.650 kV for reflector 3. Linear mode analysis average 500 laser shots at 85% laser power and 10 kHz frequency from 1500-10000 *m/z* with 15.8 x detector gain (2720 V). Voltages were 20.000 kV for ion source 1, 1.550 kV for PIE, and 10.500 kV for Lens. All imaging experiments were conducted at 50 µm with the M5 small laser setting. Profiling experiments were conducted from 500-2000 *m/z* to determine lipid content after washing. The detector gain was set

at 5.3 x, 400 laser shots were averaged with laser frequency of 5 kHz, and voltages were the same as used for imaging experiments.

MALDI-MSI Data Analysis

Data was exported into SCiLS software (Bruker Daltonics, Billerica, MA) for data analysis. For reflectron data, raw data was imported with quadratic interpretation using 80,610 data points, and the auto resample method for the mass axis settings and a top hat baseline removal with width 200. For linear data, reduced data was imported with linear interpretation using 15,000 data points, and the auto resample method for the mass axis settings and a top hat baseline removal with width 200. Data were normalized to the total ion current, a weak denoising setting was used, and hotspot removal was enabled. Peaks were picked with the find peaks tool using the skyline spectra, which plots the the maximum intensity for each *m/z*, for each region (a unique region was drawn for each biological and technical replicate). The top 800 peaks were picked for reflectron samples at an interval width of 0.1 Da, and the top 500 peaks were selected for linear samples at an interval width of 2.0 Da. Peak lists were then manually validated by visual inspection of ion images. For the final reflectron peak list of *m/z* observed with the four-step wash, all peaks that were not the monoisotopic peak were removed and a signal threshold of 0.3 for the mean spectra or 1.0 for the skyline spectra was employed, meaning that the tallest isotopic peak for the peptide had to be above one of these thresholds. Segmentation was performed on roots and root nodules with weak denoising, normalization to the total ion current and 0.007% interval width.

Sample Extraction

For extraction, root nodules were trimmed from the plant with 2-3 mm of the surrounding root, flash frozen on dry ice, and saved at -80$^{o}$C until extraction. Nodules were ground to powder in a mortar and pestle under liquid nitrogen. Extraction solvent (40% ethanol) was added to the

wild-type and CLE nodules to reach a concentration of 10 mL/g sample, and the mixtures were probe sonicated with 15 s pulse, 30 s off cycle for 3 minutes' total pulse time. The extractions were centrifuged at 4,000 x g at 4$^{o}$C for 30 minutes, and the supernatant was dried down in a speed vac. Samples were resuspended in 50 mM tris buffer pH=8 at a peptide concentration of 3.5 mg/mL, reduced with dithiothreitol (final concentration 5 mM) for 60 min at room temperature, alkylated with iodoacetamide (final concentration 15 mM) for 30 minutes at room temperature in the dark. The volume of DTT was added a second time and 10% TFA added to a final concentration of 0.3% to quench. The samples were desalted with C18 SepPak and dried in the speed vac. SCX purification was performed on the samples with SCX spin tips (Protea Biosciences, Morgantown, WV) according to kit protocol 1, except SpinTip rinse was performed with 100 µL, the sample was reconstituted in 200 µL and loaded three times through the SpinTip, the sample was eluted with two 150 µL aliquots, and all centrifuges occurred for 1.5 min. Samples were dried down in a speed vac, and then three 1 mL aliquots of MilliQ water were added to the sample followed by drying down in the speed vac for desalting after SCX purification. The Pierce Quantitative Colorimetric Peptide Assay (Thermo Scientific) was used to assess peptide concentration. Samples were stored at -80$^{o}$C until analysis.

nanoLC-MS/MS Data Acquisition

Samples were resuspended at 1 µg/µL in Optima grade water with 0.1% formic acid. LC-MS/MS was performed on a nano-Aquity LC (Waters) coupled to Q Exactive mass spectrometer (Thermo Scientific). Separation occurred on a 15 cm long self-packed C18 nano-LC column at a flow rate of 0.35 µL/min. Mobile phase A was optima water with 0.1% formic acid and mobile phase B was optima acetonitrile with 0.1% formic acid. Injection amount was 2 µL, and samples were kept at 4$^{o}$C during analysis. The LC separation used 14 min trapping initially, followed by a

gradient from 0-4% B from 0 to 2 min, 4-35% B from 2-70 min, 35-75% B from 70-71 min, hold at 75% B from 71-81 min, 75-95% from 81-82 min, hold at 95% B from 82-92 min, 95-0% B from 92-93 min, and re-equilibration at 0% B for from 93-108 min. A top 15 DDA method was used for MS analysis. Full MS spectra were taken from 300-2000 *m/z* at 35,500 resolution, 1E6 AGC target, and 100 ms max injection time. The default charge state was selected as 2. MS/MS spectra were performed with HCD activation at a collision energy of 30 at 17,500 resolution, 1E5 AGC target, 200 ms max injection time and a 2.0 *m/z* isolation window. A dynamic exclusion of 30 s and a charge exclusion of unassigned, 1, and greater than or equal to 8 was used.

nanoLC-MS/MS Data Analysis

LC-MS/MS experiments were analyzed in PEAKS 7.0 software for *de novo* sequencing of MS/MS data and protein identification. Parameters include 10.0 ppm parent ion tolerance, 0.01 Da fragment ion tolerance, carbamidomethylation (57.02) as a fixed modification, and oxidation (M, 15.99 Da), amidation (-0.98 Da), acetylation (42.01 Da), and hydroxylation (15.99 Da) as variable post-translational modifications (PTM). Three variable PTMs were allowed per peptide, and sequenced peptides were searched against the UniProt *Medicago truncatula* database in no enzyme mode. A 1% FDR cut-off and one unique peptide per protein settings were used for data analysis. Peptides were manually verified to check that a majority of amino acids in the peptide sequence were fragmented.

## Results

Optimization of Washing Strategy for Detection of NCR Peptides

Initially, six different washes were attempted on the root nodule sections to test for signal improvement in the range typically observed for NCR peptides (3-5 kDa). The averaged spectra

for the roots and root nodules for each of the washes compared to the control (no wash) are shown in **Figure 1**. In **Figure 1(A)**, three gradient ethanol washes were compared to a control. The washes attempted were (i) two-step wash consisting of 70% ethanol for 30 s, and 100% ethanol for 30s, (ii) three-step wash composed of 70% ethanol for 90s, 70% ethanol for 90s, and 100% ethanol for 90s, (iii) four-step wash consisting of 70% ethanol for 30s, 100% ethanol for 30s, Carnoy solution for 2 min, and 100% ethanol for 30s. While all spectra showed similar peak distributions, especially from 2000-3500 *m/z*, the control nodules showed more peaks below 2000 *m/z*, and the washed nodules showed an increase in signal for peaks above 3500 *m/z*. As the range for NCR peptides falls mostly above 3500 *m/z* and are not expected to be below 2000 *m/z*, it was concluded that all three multi-step gradient washes attempted showed promise for increasing the signal of NCR peptides in MALDI-MSI. **Figure 1(B)**, shows the averaged spectra in reflectron mode over the root and root nodules for three single step washes of 30 s with either 100% ethanol, chloroform, or hexanes. These washes were tested as the longer multi-step washes could cause delocalization or wash away peptides, in which case, shorter washes could be advantageous. The hexanes wash and control spectra are very similar; as a result, washing with hexanes does not appear to alter the spectra. The chloroform and ethanol spectra, however, did show an increase in signal of *m/z* above 3500. Although the ethanol and chloroform washes were very similar, when overlaid, the ethanol wash showed slightly higher signal increases. The averaged spectra in linear mode for the roots and root nodules resulted in similar trends as the reflectron mode, with the ethanol, chloroform, two-step, three-step, and four-step washes increasing the signal of *m/z* above 3500 compared to the control (**Supporting Figure S1**). Washes are typically employed to remove lipid species, which readily ionize and are highly abundant and can, therefore, suppress signal of other species. To validate that lipid removal was causing the signal increase in peptides above 3500 *m/z*, root

nodules were profiled in the 500-2000 *m/z* region. This analysis showed that all the washes except the hexanes wash removed lipid species in the root nodules, with the three-step wash and the four-step wash removing the most lipids species. (**Supporting Figure S2**).

To further investigate the signal improvement after washing nodule sections with an organic solvent, additional experiments on biological replicates were performed for the single step ethanol wash, along with the two, three, and four-step gradient ethanol washes. As the ethanol and chloroform 30 s washes resulted in very similar results, the ethanol wash was chosen to be compared to the two-step, three-step, and four-step washes, which resulted in stronger signal increases compared to the one-step washes. Spectra in both linear and reflectron mode for the additional biological replicates are shown in **Supporting Figures S3-S6**. All washes showed improvement in the signal intensity of ions above 3500 *m/z*. To further compare the signal increases from the washes, box and whisker plots for specific *m/z*, along with the ion image for the *m/z*, were compared (**Figure 2**). The *m/z* shown in **Figure 2** were selected to provide multiple examples of *m/z* values across the typical range observed for NCR peptides. The box and whisker plots show an increase in signal for these four *m/z* values after washing with the two-step, three-step, and four-step washes. The images for the *m/z* also show a corresponding increase in signal after performing these washes.

Additionally, the ion images for the washes are very similar to the control ion images, indicating that delocalization of the peptides does not appear to be a problem. The signal for the ethanol washed nodules was more similar to the unwashed nodules. The longer three- and four-step washes generally showed a larger increase in signal for most of the *m/z* values above 3500. However, the washes did show a decrease in signal for some low-intensity *m/z* values between 2500 and 3000. The ethanol wash was the only wash that did not have a loss of specific *m/z* in the

2500-3000 range. This result indicated that the 70% ethanol step might result in a decrease of the signal for a handful of *m/z* values at the lower end of the NCR peptide range. As the number of *m/z* values that decreased with the 70% ethanol step was less than the number of *m/z* values above 3500 that increased with the step, washing with 70% ethanol initially is deemed advantageous. The four-step wash was, therefore, chosen as the best wash over the three-step wash, as the three-step wash had a longer 70% ethanol washing step, which was not necessary to increase signal, and negatively impacted the signal in the 2500-3000 *m/z* range more than the shorter 70% ethanol washing step (**Figure 2(A)**).

MALDI-MSI of NCR Endogenous Peptides

After conducting the four-step wash, 93 *m/z* values were observed in reflectron mode (all non-monoisotopic peaks were removed), and 106 *m/z* were observed in linear mode. In reflectron mode, 71 of the peaks were above 2500 *m/z* (start of potential NCR peaks), and in linear mode, 94 of the peaks were above 2500 *m/z*. Thus, linear mode resulted in slightly more peaks; however, reflectron mode allowed the isotopic pattern of peptides to be resolved. **Figure 3** shows example ion images detected in the washed nodules. The optical image for the root nodules in shown in **Figure 3(A)**. Segmentation analysis was performed on the root nodule to observe trends in the data, shown in **Figure 3(B)**. Initially, the root, which is in red, and the nodule, were separated. The root nodule was then further divided into the lower, shown in yellow, and upper, shown in blue, halves. The imaging data reflects this as many of the *m/z* are localized to either the entire nodule, as in **Figure 3(E)**, or the lower half of the nodule, as in **Figure 3(D,F)**, although some *m/z* do show increased signal in the mid to upper half of the nodule (**Figure 2(B)**). Furthermore, a couple *m/z* values below 2500 show localization to the roots and the outer edge of the nodule, as in **Figure**

**3(C)**. While these peptides in the outer rim of the nodules are below the mass range of NCR peptides, they could be a different class of endogenous peptide in the plant.

NCR peptide sequences, accessed from UniProt, were processed through SignalP [44] to generate predicted NCR sequences in order to perform accurate mass-matching to the MALDI-MSI detected peptides. **Supporting Table S3** shows the results of the SignalP analysis. Peptides from known or predicted NCR proteins were kept, and predicted mature chains over 100 residues were deleted, as these were very unlikely to be below 10,000 Da, which was the upper limit for acquiring imaging data in linear mode. The monoisotopic mass of each mature chain predicted peptide was calculated using Xcalibur Qual Browser software (Thermo Scientific, Waltham, MA) using the peptide/protein isotope calculator. The number of cysteine residues in the sequence was then counted, and the number of disulfide bonds predicted for the sequence based upon the number of cysteine residues. If there were 0-1 cysteine residues present, no disulfide bond was predicted, if there were 2-3 cysteine residues present, 1 disulfide bond was predicted, if there were 4-5 cysteine residues present, 2 disulfide bonds were predicted, and if 6 or more disulfide bonds were present, 3 disulfide bonds were predicted. Thus, the *m/z* values of mature peptide monoisotopic mass (M) and the singly protonated form $[M+H]^+$ were calculated based upon the number of disulfide bonds present. The $[M+H]^+$ *m/z* values were then used to match with the *m/z* values detected from imaging data acquired in reflectron mode (**Supporting Table S1**). **Table 1** provides the putative identifications based upon accurate mass matching to the predicted NCR mature peptides in **Supplemental Table 3**. Sixteen imaging *m/z* values were matched to NCR predicted peptides (or a group of NCR predicted peptides). **Table 3** shows the overlap between NCR peptides detected in this study and two earlier studies, including the comprehensive proteome analysis that we had previously performed [45].

LC-MS/MS for Identification of NCR Peptides

Given the roles of CLE peptides in maintaining the balance between cell proliferation and differentiation and NCR peptides in nodule differentiation, extractions were performed on root nodules from Medicago wild-type, and transgenic plants overexpressing *MtCLE13* for tandem MS (MS/MS) information to support predicted NCR identifications in imaging data. A 40% ethanol extraction was performed as this extraction previously had success with cysteine-rich cyclotide endogenous peptides, which have a similar mass range as NCR peptides.[46] Peptides were reduced and alkylated to improve MS/MS sequence coverage, and SCX spin tip purification was performed to remove potential contaminants. The wild-type root nodules resulted in 17 NCR identifications, while the CLE root nodules resulted in 22 identifications. Of these identifications, 15 were shared between the wild-type and CLE root nodules, leaving 2 unique identifications in the wild-type root nodules, and 7 unique identifications in the CLE root nodules. **Supporting Table S4** and **Supporting Table S5** provide the NCR peptides identified in wild-type and CLE root nodules with an FDR of 1%. Of the identified NCR peptides, 5 were potentially identified with the entire, intact mature peptide (A0A072U6V1, wt; A0A072U1G7, wt, and CLE; A0A072THJ9, CLE; G7JHR4, CLE; A7KH73, CLE). While the length of the detected intact NCR peptide, A0A072U6V1 matched the predicted Uniprot mature peptide, the NCR peptide A0A072U1G7, G7JHR4, and A7KH73 were predicted as transmembrane in Uniprot and therefore, did not have a predicted mature peptide. The NCR peptide A0A072THJ9 was identified starting at amino acid 24 in the sequence, but Uniprot predicts the start of the mature peptide at amino acid 23 in the sequence, which is potentially a one amino acid truncation.

The intact peptide mass with disulfide bonds was calculated for putative intact mature NCR peptides by subtracting the alkylation modification on cysteine residues (57.02146 Da), and then

adding in either two or three disulfide bonds by subtracting either 4 or 6 hydrogen atoms (1.007276) depending on whether there were 4 or 6 cysteine residues. Finally, the calculated $[M+H]^+$ values from the extraction were matched to the detected monoisotopic peaks in the four-step wash to identify NCR peptides in the MALDI-MSI data. From this, three NCR peptides were identified in the imaging data. These peptides are shown in **Table 2**, which contains sequence information and the difference between the calculated $[M+H]^+$ values from the extraction and the observed $[M+H]^+$ values from the MALDI-MSI. The images for these peptides are shown in **Figure 4**. The NCR peptide A0A072THJ9 is localized to the lower half of the nodules, while the other two identified NCR peptides are distributed throughout the entire nodules. LC-MS/MS analysis also detected fragments of mature NCR peptides. For these peptides, predicted mature peptide sequence in UniProt (if provided) was searched against the imaging data. The UniProt predicted mature peptide for A0A072THJ9, which was one amino acid longer than the observed mature peptide in the LC-MS/MS results, was matched to a detected peak in the imaging results (*m/z* 2646.17, shown in **Figure 2(A)**), indicating that there is a likelihood of some degradation occurring in both the MALDI-MSI and LC-MS/MS results. Another match to the imaging results from the UniProt predicted mature peptides was for NCR peptide A0A072U6V1, where the mature peptide and longest peptide revealed by LC-MS/MS analysis were precisely the same.

**Discussion**

Organic washes are typically used in MALDI-MSI to enhance the detection of proteins, whose signal can be suppressed by high concentration lipid species in the tissue.[39] Gradient ethanol washes are commonly used for this purpose of improving protein imaging and tissue fixation,[47-49] and has been utilized in one instance for neuropeptide MALDI-MSI imaging.[50] Therefore, three

variations of gradient ethanol washes were investigated for enhancement of signal in the 3-5 kDa range, including a (i) 70% ethanol, and 100% ethanol for 30s each, (ii) 70% ethanol, 70% ethanol, and 100% ethanol for 90s each, and (iii) 70% ethanol for 30s, 100% ethanol for 30s, Carnoy's solution for 2 min, and 100% ethanol for 30s. The water wash or water dipping step that is commonly used in gradient ethanol washes was omitted here due to concerns about high peptide solubility in water and subsequent signal loss and delocalization of peptides.

Additionally, single solvents (*i.e.*, chloroform, ethanol) tested previously for protein MALDI-MSI that can be used for lipid removal[39, 51] were tested along with a single step 100% ethanol wash as smaller endogenous peptides may not benefit from the longer gradient ethanol washes. The results showed, however, that the longer washes improved peptide signals in the 3-5 kDa range more than the shorter and single solvent washes. The larger size of the NCR peptides compared to other peptides, which can be less than 2 kDa, may account for why the longer washes are beneficial.

With the optimal wash of 70% ethanol for 30s, 100% ethanol for 30s, Carnoy's solution for 2 min, and 100% ethanol for 30s, multiple NCR peptides were putatively identified with either accurate mass matching to the SignalP predicted mature peptides or the root nodule extraction followed by LC-MS/MS. The accurate mass matching to SignalP predicted peptides identified more putative peptides than the LC-MS/MS results, 16 compared to 3, respectively, likely due to the higher number of NCR peptides in the SignalP predicted peptide database compared to the identified NCR peptides by MS/MS. However, a majority of the 71 potential NCR peptide peaks from the reflectron data of the four-step wash were not identified. Due to lack of sequence information in the MALDI-MSI results, we cannot confirm that these 71 peaks contain the two to three disulfide bonds indicative of NCR peptides. While accurate mass matching was carried out

with predicted NCR peptides, a previous LC-MS/MS study utilizing trypsin digestion was able to confirm approximately 50% of the predicted SignalP cleavage sites for the NCR peptides detected in their research, indicating that the predicted SignalP results should be reliable for most of the peptides.[52] Thus, the percentage of incorrectly predicted cleavage sites in the SignalP database should be low. Several identified peptides in **Table 1** have been previously detected in *M. truncatula* nodules,[45, 53] providing validation to our experimental technique. Degradation of the NCR peptides appears to be a more significant concern in this study, especially during sample preparation for LC-MS/MS analysis. Most of the NCR peptides identified via LC-MS/MS were fragments rather than full length, mature peptides.

Genetic studies conducted in *M. truncatula* have demonstrated the involvement of NCR211 and NCR169 in bacteroid differentiation. In the absence of NCR169, bacteroid differentiation is impaired, and the nodules show premature signs of senescence, while in the absence of NCR211, bacteroid differentiation appears to be normal, but the subsequent maintenance of intracellular bacteroids is impeded.[54, 55] We were unable to detect these peptides using our optimization technique. This lack of detection could be attributed to post-translational modification of these peptides resulting in a change in mass. Although atypical to cysteine-rich peptides, post-translational modifications such as arabinosylation, hydroxylation and sulfation occur in several secreted peptides in plants.[56] However, the possibility of degradation during processing of the nodule samples cannot be completely ruled out.

Further optimization of the extraction procedure, for example, by adding acid to the solvent or by adding protease inhibitors, may help reduce degradation. A majority of the degradation products observed in the LC-MS/MS results were not found in the MALDI-MSI results, indicating that most of the deterioration occurred in solution. One exception to this, however, was the most

extended peptide identified by LC-MS/MS for the NCR protein A0A072THJ9. This peptide was one residue shorter than the predicted mature NCR peptide, and both the LC-MS/MS identified peptide and the predicted full-length peptide were observed in the MALDI-MSI results. Based on this observation, although the MALDI-MSI results have less degradation than the LC-MS/MS results, there is still likely degradation occurring in the MALDI-MSI sample preparation.

Preliminary experiments utilizing a Denator on the nodules before embedding in gelatin for MALDI-MSI analysis were conducted to determine if this would be a potential way to limit degradation in the imaging of NCR peptides. **Supporting Figure S6** shows spectra averaged over the roots and root nodules after treatment with the Denator. The spectra still show that the *m/z* 2532.1 peak is taller than the *m/z* 2636.1 peak, indicating possible degradation of this peptide. This can potentially be attributed to the time between trimming the nodule from the plant to the start of the Denator process was not much shorter than flash freezing in gelatin due to the air evacuation step with the Denator system. Additionally, while the root nodules withstood the heat stabilization process, the roots were more fragile and decreased in volume after heat stabilization by the Denator. Thus, the Denator heat stabilization procedure needs to be further optimized to achieve the best results. The ratio of the two peaks (*m/z* 2532.1 and the *m/z* 2636.1 peak) could be useful, however, for assessing degradation in the sample.

Overall, the MALDI-MSI method developed here by including a four-step washing procedure on the root nodules before matrix application showed significant improvement in most peptide signals above *m/z* 2500, especially for *m/z* above 3500. Different localization patterns were observed for various peptide signals, which could be related to the different functions of these peptides. The developed method also enabled the MS imaging of 16 putative NCR peptides, of which three were also identified via LC-MS/MS. While further development is needed to identify

intact mature NCR peptides via LC-MS/MS, the MALDI-MSI results did not show nearly as many peptide degradation fragments, suggesting its advantage as an *in situ* direct tissue measurement approach. As a result, the MALDI-MSI method developed here could be extended and applied to other plant systems to increase our knowledge about spatial localization of these important signaling peptides and help to promote the prevalence of endogenous peptide MALDI-MSI in plant science.

## Acknowledgments

This work was supported in part by funding from the National Science Foundation (NSF) Division of Integrative Organismal Systems (IOS) RESEARCH PGR award #1546742, University of Wisconsin-Madison Graduate School and the Wisconsin Alumni Research Foundation (WARF), a Vilas Distinguished Achievement Professorship to L.L., and a NSF grant to JMA (NSF#0701846). The Q Exactive instrument was purchased through an NIH shared instrument grant (NCRR S10RR029531 to LL). The RapifleX MALDI TOF/TOF mass spectrometer was purchased through the support of an NIH shared instrument grant S10OD025084. NQV thanks National Institutes of Health-General Medical Sciences (T32GM008349) for funding. L.L. acknowledges a Vilas Distinguished Achievement Professorship and the Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

## References

1. Marmiroli, N.; Maestri, E., Plant peptides in defense and signaling. *Peptides* **2014,** *56*, 30-44.

2. de Bang, T. C.; Lundquist, P. K.; Dai, X.; Boschiero, C.; Zhuang, Z.; Pant, P.; Torres-Jerez, I.; Roy, S.; Nogales, J.; Veerappan, V.; Dickstein, R.; Udvardi, M. K.; Zhao, P. X.; Scheible, W. R., Genome-Wide Identification of Medicago Peptides Involved in Macronutrient Responses and Nodulation. *Plant Physiol* **2017,** *175* (4), 1669-1689.

3. Mus, F.; Crook, M. B.; Garcia, K.; Costas, A. G.; Geddes, B. A.; Kouri, E. D.; Paramasivan, P.; Ryu, M. H.; Oldroyd, G. E. D.; Poole, P. S.; Udvardi, M. K.; Voigt, C. A.; Ane, J. M.; Peters, J. W., Symbiotic Nitrogen Fixation and the Challenges to Its Extension to Nonlegumes. *Applied and Environmental Microbiology* **2016,** *82* (13), 3698-3710.

4. Djordjevic, M. A.; Mohd-Radzman, N. A.; Imin, N., Small-peptide signals that control root nodule number, development, and symbiosis. *Journal of Experimental Botany* **2015,** *66* (17), 5171-5181.

5. Imin, N.; Mohd-Radzman, N. A.; Ogilvie, H. A.; Djordjevic, M. A., The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in Medicago truncatula. *Journal of Experimental Botany* **2013,** *64* (17), 5395-5409.

6. Mohd-Radzman, N. A.; Binos, S.; Truong, T. T.; Imin, N.; Mariani, M.; Djordjevic, M. A., Novel MtCEP1 peptides produced in vivo differentially regulate root development in Medicago truncatula. *Journal of Experimental Botany* **2015,** *66* (17), 5289-5300.

7. Mortier, V.; Den Herder, G.; Whitford, R.; Van de Velde, W.; Rombauts, S.; D'Haeseleer, K.; Holsters, M.; Goormachtig, S., CLE Peptides Control Medicago truncatula Nodulation Locally and Systemically. *Plant Physiology* **2010,** *153* (1), 222-237.

8. Imin, N.; Patel, N.; Corcilius, L.; Payne, R. J.; Djordjevic, M. A., CLE peptide tri-arabinosylation and peptide domain sequence composition are essential for SUNN-dependent autoregulation of nodulation in Medicago truncatula. *New Phytol* **2018,** *218* (1), 73-80.

9. Pan, H.; Wang, D., Nodule cysteine-rich peptides maintain a working balance during nitrogen-fixing symbiosis. *Nat Plants* **2017,** *3* (5), 17048.

10. Farrokhi, N.; Whitelegge, J. P.; Brusslan, J. A., Plant peptides and peptidomics. *Plant Biotechnology Journal* **2008,** *6* (2), 105-134.

11. Khatib-Shahidi, S.; Andersson, M.; Herman, J. L.; Gillespie, T. A.; Caprioli, R. M., Direct molecular analysis of whole-body animal tissue sections by imaging MALDI mass spectrometry. *Analytical Chemistry* **2006,** *78* (18), 6448-6456.

12. Gode, D.; Volmer, D. A., Lipid imaging by mass spectrometry - a review. *Analyst* **2013,** *138* (5), 1289-315.

13. Zhang, Y.; DeLaney, K.; Hui, L.; Wang, J.; Sturm, R. M.; Li, L., A Multifaceted Mass Spectrometric Method to Probe Feeding Related Neuropeptide Changes in Callinectes sapidus and Carcinus maenas. *J Am Soc Mass Spectrom* **2018,** *29* (5), 948-960.

14. Monroe, E. B.; Annangudi, S. P.; Hatcher, N. G.; Gutstein, H. B.; Rubakhin, S. S.; Sweedler, J. V., SIMS and MALDI MS imaging of the spinal cord. *Proteomics* **2008,** *8* (18), 3746-54.

15. Caprioli, R. M.; Farmer, T. B.; Gile, J., Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* **1997,** *69* (23), 4751-60.

16. Casadonte, R.; Caprioli, R. M., Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. *Nature Protocols* **2011,** *6* (11), 1695-1709.

17. Buchberger, A. R.; DeLaney, K.; Johnson, J.; Li, L., Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal Chem* **2018,** *90* (1), 240-265.

18. Lee, Y. J.; Perdian, D. C.; Song, Z.; Yeung, E. S.; Nikolau, B. J., Use of mass spectrometry for imaging metabolites in plants. *Plant J* **2012,** *70* (1), 81-95.

19. Sturtevant, D.; Lee, Y. J.; Chapman, K. D., Matrix assisted laser desorption/ionization-mass spectrometry imaging (MALDI-MSI) for direct visualization of plant metabolites in situ. *Current Opinion in Biotechnology* **2016,** *37*, 53-60.

20. Feenstra, A. D.; Hansen, R. L.; Lee, Y. J., Multi-matrix, dual polarity, tandem mass spectrometry imaging strategy applied to a germinated maize seed: toward mass spectrometry imaging of an untargeted metabolome. *Analyst* **2015,** *140* (21), 7293-304.

21. Bjarnholt, N.; Li, B.; D'Alvise, J.; Janfelt, C., Mass spectrometry imaging of plant metabolites - principles and possibilities. *Nat Prod Rep* **2014,** *31* (6), 818–837.

22. Li, B.; Bhandari, D. R.; Janfelt, C.; Rompp, A.; Spengler, B., Natural products in Glycyrrhiza glabra (licorice) rhizome imaged at the cellular level by atmospheric pressure matrix-assisted laser desorption/ionization tandem mass spectrometry imaging. *Plant J* **2014,** *80* (1), 161-71.

23. Bhandari, D. R.; Wang, Q.; Li, B.; Friedt, W.; Rompp, A.; Spengler, B.; Gottwald, S., Histology-guided high-resolution AP-SMALDI mass spectrometry imaging of wheat-Fusarium graminearum interaction at the root-shoot junction. *Plant Methods* **2018,** *14*, 103.

24. Gemperline, E.; Jayaraman, D.; Maeda, J.; Ane, J. M.; Li, L., Multifaceted investigation of metabolites during nitrogen fixation in Medicago via high resolution MALDI-MS imaging and ESI-MS. *J Am Soc Mass Spectrom* **2015,** *26* (1), 149-58.

25. Ye, H.; Gemperline, E.; Venkateshwaran, M.; Chen, R.; Delaux, P. M.; Howes-Podoll, M.; Ane, J. M.; Li, L., MALDI mass spectrometry-assisted molecular imaging of metabolites during nitrogen fixation in the *Medicago truncatula-Sinorhizobium meliloti* symbiosis. *Plant J.* **2013,** *75* (1), 130–45.

26. Horn, P. J.; Korte, A. R.; Neogi, P. B.; Love, E.; Fuchs, J.; Strupat, K.; Borisjuk, L.; Shulaev, V.; Lee, Y. J.; Chapman, K. D., Spatial Mapping of Lipids at Cellular Resolution in Embryos of Cotton. *Plant Cell* **2012,** *24* (2), 622-636.

27. Korte, A. R.; Yandeau-Nelson, M. D.; Nikolau, B. J.; Lee, Y. J., Subcellular-level resolution MALDI-MS imaging of maize leaf metabolites by MALDI-linear ion trap-Orbitrap mass spectrometer. *Anal Bioanal Chem* **2015,** *407* (8), 2301-9.

28. Sarsby, J.; Towers, M. W.; Stain, C.; Cramer, R.; Koroleva, O. A., Mass spectrometry imaging of glucosinolates in Arabidopsis flowers and siliques. *Phytochemistry* **2012,** *77*, 110-118.

29. Gemperline, E.; Keller, C.; Jayaraman, D.; Maeda, J.; Sussman, M. R.; Ane, J. M.; Li, L., Examination of Endogenous Peptides in Medicago truncatula Using Mass Spectrometry Imaging. *J Proteome Res* **2016,** *15* (12), 4403-4411.

30. Barbosa, E. A.; Bonfim, M. F., Jr.; Bloch, C., Jr.; Engler, G.; Rocha, T.; de Almeida Engler, J., Imaging Mass Spectrometry of Endogenous Polypeptides and Secondary Metabolites from Galls Induced by Root-Knot Nematodes in Tomato Roots. *Mol Plant Microbe Interact* **2018,** *31* (10), 1048-1059.

31. Grassl, J.; Taylor, N. L.; Millar, A. H., Matrix-assisted laser desorption/ionisation mass spectrometry imaging and its development for plant protein imaging. *Plant Methods* **2011,** *7*, 11.

32. Kaspar, S.; Peukert, M.; Svatos, A.; Matros, A.; Mock, H. P., MALDI-imaging mass spectrometry - An emerging technique in plant biology. *Proteomics* **2011,** *11* (9), 1840-50.

33. Poth, A. G.; Mylne, J. S.; Grassl, J.; Lyons, R. E.; Millar, A. H.; Colgrave, M. L.; Craik, D. J., Cyclotides associate with leaf vasculature and are the products of a novel precursor in petunia (Solanaceae). *J Biol Chem* **2012,** *287* (32), 27033-46.

34. Slazak, B.; Kapusta, M.; Malik, S.; Bohdanowicz, J.; Kuta, E.; Malec, P.; Goransson, U., Immunolocalization of cyclotides in plant cells, tissues and organ supports their role in host defense. *Planta* **2016,** *244* (5), 1029-1040.

35. Farkas, A.; Maroti, G.; Kereszt, A.; Kondorosi, E., Comparative Analysis of the Bacterial Membrane Disruption Effect of Two Natural Plant Antimicrobial Peptides. *Front Microbiol* **2017,** *8*, 51.

36. Quiason, C. M.; Shahidi-Latham, S. K., Imaging MALDI MS of Dosed Brain Tissues Utilizing an Alternative Analyte Pre-extraction Approach. *Journal of the American Society for Mass Spectrometry* **2015,** *26* (6), 967-973.

37. Powers, T. W.; Neely, B. A.; Shao, Y.; Tang, H.; Troyer, D. A.; Mehta, A. S.; Haab, B. B.; Drake, R. R., MALDI imaging mass spectrometry profiling of N-glycans in formalin-fixed paraffin embedded clinical tissue blocks and tissue microarrays. *PLoS One* **2014,** *9* (9), e106255.

38. Wu, Q.; Comi, T. J.; Li, B.; Rubakhin, S. S.; Sweedler, J. V., On-Tissue Derivatization via Electrospray Deposition for Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging of Endogenous Fatty Acids in Rat Brain Tissues. *Anal Chem* **2016,** *88* (11), 5988-95.

39. Seeley, E. H.; Oppenheimer, S. R.; Mi, D.; Chaurand, P.; Caprioli, R. M., Enhancement of protein sensitivity for MALDI imaging mass spectrometry after chemical treatment of tissue sections. *Journal of the American Society for Mass Spectrometry* **2008,** *19* (8), 1069-1077.

40. Boisson-Dernier, A.; Chabaud, M.; Garcia, F.; Becard, G.; Rosenberg, C.; Barker, D. G., Agrobacterium rhizogenes-transformed roots of Medicago truncatula for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Mol Plant Microbe Interact* **2001,** *14* (6), 695-700.

41. Meade, H. M.; Long, S. R.; Ruvkun, G. B.; Brown, S. E.; Ausubel, F. M., Physical and genetic characterization of symbiotic and auxotrophic mutants of Rhizobium meliloti induced by transposon Tn5 mutagenesis. *J Bacteriol* **1982,** *149* (1), 114-22.

42. Trinh, T. H.; Ratet, P.; Kondorosi, E.; Durand, P.; Kamate, K.; Bauer, P.; Kondorosi, A., Rapid and efficient transformation of diploid Medicago truncatula and Medicago sativa ssp. falcata lines improved in somatic embryogenesis. *Plant Cell Rep* **1998,** *17* (5), 345-355.

43. Chabaud, M.; de Carvalho-Niebel, F.; Barker, D. G., Efficient transformation of Medicago truncatula cv. Jemalong using the hypervirulent Agrobacterium tumefaciens strain AGL1. *Plant Cell Rep* **2003,** *22* (1), 46-51.

44. Petersen, T. N.; Brunak, S.; von Heijne, G.; Nielsen, H., SignalP 4.0: discriminating signal peptides from transmembrane regions. In *Nat Methods*, United States, 2011; Vol. 8, pp 785-6.

45. Marx, H.; Minogue, C. E.; Jayaraman, D.; Richards, A. L.; Kwiecien, N. W.; Siahpirani, A. F.; Rajasekar, S.; Maeda, J.; Garcia, K.; Del Valle-Echevarria, A. R.; Volkening, J. D.; Westphall, M. S.; Roy, S.; Sussman, M. R.; Ane, J. M.; Coon, J. J., A proteomic atlas of the legume Medicago truncatula and its nitrogen-fixing endosymbiont Sinorhizobium meliloti. *Nat Biotechnol* **2016,** *34* (11), 1198-1205.

46. Serra, A.; Hemu, X.; Nguyen, G. K.; Nguyen, N. T.; Sze, S. K.; Tam, J. P., A high-throughput peptidomic strategy to decipher the molecular diversity of cyclic cysteine-rich peptides. *Sci Rep* **2016,** *6*, 23005.

47. Maier, S. K.; Hahne, H.; Gholami, A. M.; Balluff, B.; Meding, S.; Schoene, C.; Walch, A. K.; Kuster, B., Comprehensive identification of proteins from MALDI imaging. *Mol Cell Proteomics* **2013,** *12* (10), 2901-10.

48. Chaurand, P.; Norris, J. L.; Cornett, D. S.; Mobley, J. A.; Caprioli, R. M., New developments in profiling and imaging of proteins from tissue sections by MALDI mass spectrometry. *Journal of Proteome Research* **2006,** *5* (11), 2889-2900.

49. Spraggins, J. M.; Rizzo, D. G.; Moore, J. L.; Rose, K. L.; Hammer, N. D.; Skaar, E. P.; Caprioli, R. M., MALDI FTICR IMS of Intact Proteins: Using Mass Accuracy to Link Protein Images with Proteomics Data. *J Am Soc Mass Spectrom* **2015,** *26* (6), 974-85.

50. Hanrieder, J.; Ljungdahl, A.; Andersson, M., MALDI imaging mass spectrometry of neuropeptides in Parkinson's disease. *J Vis Exp* **2012,** (60).

51. Lemaire, R.; Wisztorski, M.; Desmons, A.; Tabet, J. C.; Day, R.; Salzet, M.; Fournier, I., MALDI-MS direct tissue analysis of proteins: Improving signal sensitivity using organic treatments. *Anal Chem* **2006,** *78* (20), 7145-53.

52. Durgo, H.; Klement, E.; Hunyadi-Gulyas, E.; Szucs, A.; Kereszt, A.; Medzihradszky, K. F.; Kondorosi, E., Identification of nodule-specific cysteine-rich plant peptides in endosymbiotic bacteria. *Proteomics* **2015,** *15* (13), 2291-5.

53. Liese, R.; Schulze, J.; Cabeza, R. A., Nitrate application or P deficiency induce a decline in Medicago truncatula N2-fixation by similar changes in the nodule transcriptome. *Scientific Reports* **2017,** *7*, 46264.

54. Horvath, B.; Domonkos, A.; Kereszt, A.; Szucs, A.; Abraham, E.; Ayaydin, F.; Boka, K.; Chen, Y.; Chen, R.; Murray, J. D.; Udvardi, M. K.; Kondorosi, E.; Kalo, P., Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the Medicago truncatula dnf7 mutant. *Proc Natl Acad Sci U S A* **2015,** *112* (49), 15232-7.

55. Kim, M.; Chen, Y.; Xi, J.; Waters, C.; Chen, R.; Wang, D., An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proc Natl Acad Sci U S A* **2015,** *112* (49), 15238-43.

56. Matsubayashi, Y., Post-translational modifications in secreted peptide hormones in plants. *Plant Cell Physiol* **2011,** *52* (1), 5-13.

## Tables

**Table 1.** Identification of NCR peptides in the MALDI-MSI data from the SignalP predicated mature NCR peptide chains. The ones highlighted in green were detected by Liese *et al.*, 2017 as "commonly regulated TUs". The yellow ones were not found in their study.

| Accession | Sequence | Predicted $[M+H]^+$ | Imaging $[M+H]^+$ | Δ Da |
|---|---|---|---|---|
| A0A072THJ9 | NIPGARCATDEDCPVGEKCIGGNCVE | 2646.11 | 2646.17 | 0.07 |
| A0A072UB12 | SIPCETTADCPVAVPPEYYKCMYKVCVLIR | 3387.60 | 3387.66 | 0.06 |
| A0A072V920 | SIPCKTRTQCPEKMCRLPKFVWCIDGSCVCA | 3496.60 | 3496.70 | 0.10 |
| G7JQH4 | VKIPCVKDTDCPTLPCPLYSKCVDGFCKMLSI | 3512.68 | 3512.76 | 0.08 |
| A0A072V961 | YVVMCEKDSDCVDSFCVPPNVPKCRVVCKCLPK | 3697.69 | 3697.78 | 0.09 |
| A0A072TZN8 | KCNSDAECRERWIMCPLETVVKCVEDECICVH | 3734.60 | 3734.72 | 0.12 |
| A0A072U6V1 | DSFRGCNKDTDCPEKFCSSPDVVRCIYIECYCI | 3802.58 | 3802.70 | 0.12 |
| G7JPA2 | EKECDTDADCRKKFAGANQHLLWCNNGYCECHTH | 3935.62 | 3935.73 | 0.11 |
| G7I5S5 | YVPCITVADCPPNTWFKIYRCEKGICRYHKLWIV | 4111.04 | 4111.04 | 0.00 |
| A0A072TT10 | DFPCKTKDDCAQQIDYIAECIIGFCRYFTPFEHPF | 4156.84 | 4156.84 | 0.00 |
| A0A072U1G7 | KCVSDAECSGQYMCPTLTVIKCIKDECVCIHYDHDKQ | 4199.82 | 4199.91 | 0.09 |
| A0A072UBM1 | FEECKEDADCHPVCSVPGCSNICTLPDVPTCIDNNCFCI | 4227.67 | 4227.76 | 0.09 |
| G7KV51 | FEECKEDADCHPVCSVPGCSNICTLPDVPTCIDNNCFCI | 4227.67 | | 0.09 |
| G7KV52 | FEECKEDADCHPVCSVPGCSNICTLPDVPTCIDNNCFCI | 4227.67 | | 0.09 |
| A0A072UWB0 | NSFNSKIVFTDCKTDKDCQNHRGFNFRCRKGNCVAKIR | 4446.14 | 4446.24 | 0.11 |
| A7KH78 | IKIFTEHRCRTDADCPARELPEYLKCQGGMCRLLIKKD | 4446.21 | | 0.03 |
| A0A072VTT4 | KDITCTVAGDCPNFFVCPPNNFVRCIRNLCKCRSLSYKQP | 4544.14 | 4543.95 | -0.19 |
| A7KH86 | FQPCVTTADCMKTLKTDENIWYECINDFCIPFPIPKGRK | 4561.15 | 4561.11 | -0.05 |
| A0A072UME6 | CDTDVDCQKKYPGPFEHLLKCIHGYCVCFPRNPGDSSGYPI | 4598.05 | 4598.09 | 0.04 |
| G7J0N2 | GIRCHDVSECPKGLYCNVGSHMECVKHQCKCIKNFEPIDLA | 4598.08 | | 0.01 |

**Table 2.** Identification of NCR peptides in the MALDI-MSI data using the LC-MS/MS results.

| Accession | Peptide | Expected $[M+H]^+$ from LC-MS/MS | Imaging $[M+H]^+$ | Δ Da |
|---|---|---|---|---|
| A0A072U6V1\| A0A072U6V1_MEDTR | DSFRGCNKDTDCPEKFCSSPDVV RCIYIECYCI | 3802.58 | 3802.70 | 0.12 |
| A0A072U1G7\| A0A072U1G7_MEDTR | KCVSDAECSGQYMCPTLTVIKCI KDECVCIHYDHDKQ | 4199.82 | 4199.91 | 0.09 |
| A0A072THJ9\|A0A072THJ9 _MEDTR | IPGARCATDEDCPVGEKCIGGNC VE | 2532.06 | 2532.13 | 0.07 |

**Table 3**. Overlap between the NCR peptides detected in two previous studies and the present study.

| **Protein ID** | **Gene ID (V 4.0)** | **Marx *et al.*, 2016** | **Liese *et al.*, 2017** |
|---|---|---|---|
| A0A072THJ9 | Medtr0115s0030 | | + |
| A0A072UB12 | Medtr7g064020 | | + |
| A0A072V920 | Medtr2g073170 | | + |
| G7JQH4 | 11413600 | | |
| A0A072V961 | Medtr2g450150 | + | + |
| A0A072TZN8 | Medtr7g063420 | + | |
| A0A072U6V1 | Medtr7g007940 | + | |
| G7JPA2 | Medtr4g033900 | + | |
| G7I5S5 | Medtr1g011245 | | |
| A0A072TT10 | Medtr8g064070 | + | + |
| A0A072U1G7 | Medtr7g063400 | + | |
| A0A072UBM1 | Medtr7g071720 | | |
| G7KV51 | | | |
| G7KV52 | | | |
| A0A072UWB0 | Medtr3g451370 | | |
| A7KH78 | Medtr6g055160 | + | + |
| A0A072VTT4 | Medtr1g052005 | + | |
| A7KH86 | Medtr5g055370 | + | + |
| A0A072UME6 | Medtr4g065840 | + | |
| G7J0N2 | Medtr3g031335 | + | + |

## Figures

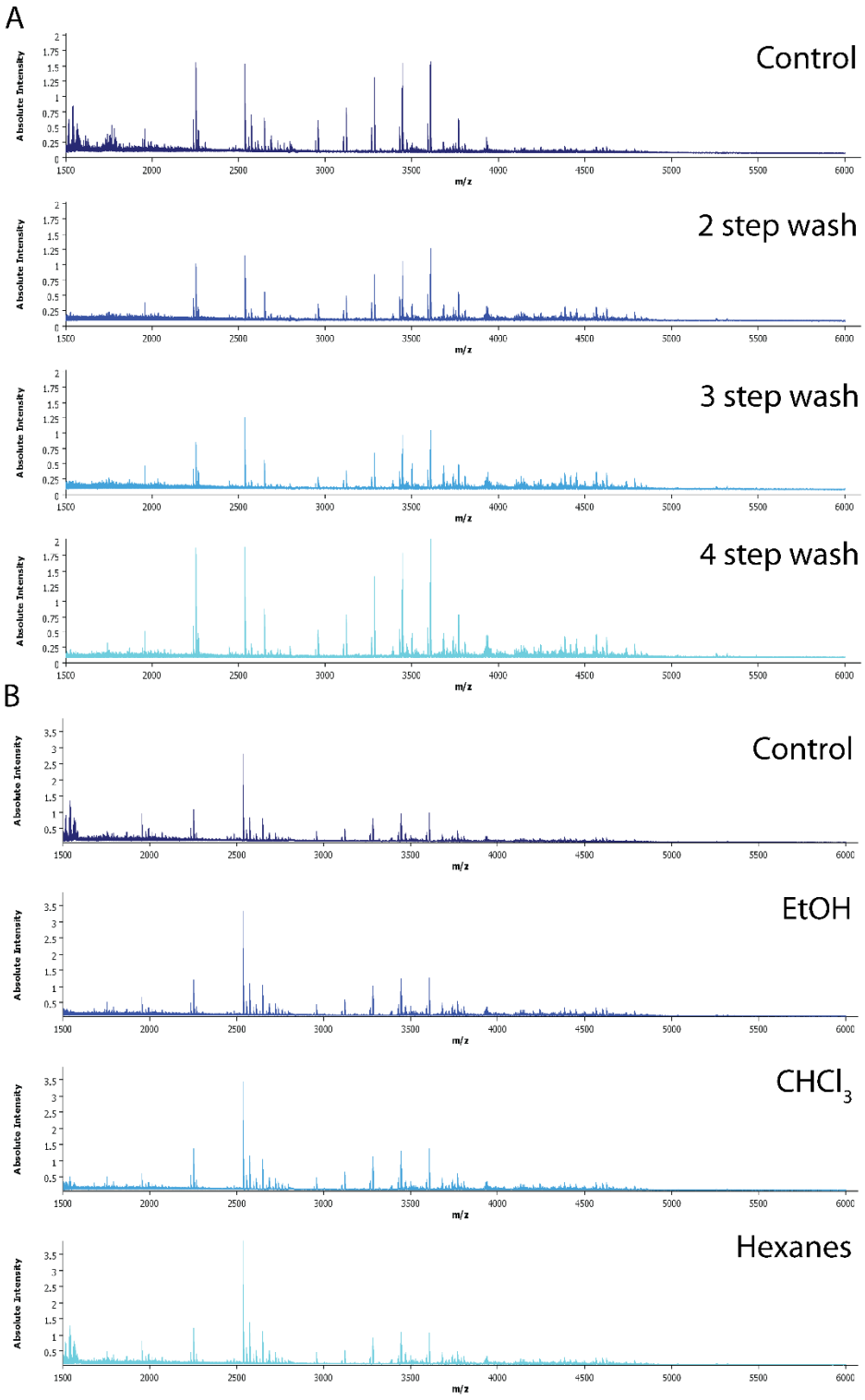



**Figure 1.** Reflectron mode mass spectra showing MALDI-MSI results for initial washes tested. In **(A)**, averaged spectra over the roots and root nodules for protein wash series, two-step wash, three-step wash, and four-step wash compared to the control. In **(B)**, averaged spectra over the roots and root nodules are shown for the single step wash series, ethanol, chloroform, and hexane are compared to a control.

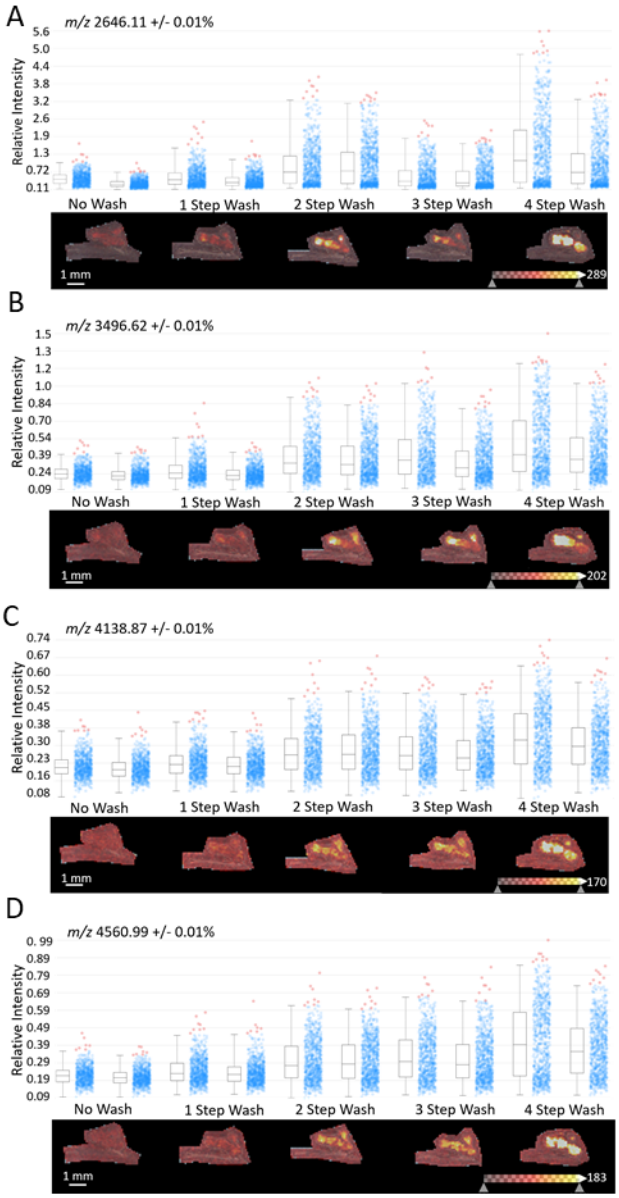



**Figure 2.** Box and whisker plots with associated ion images in reflectron mode for biological replicate 2 for **(A)** *m/z* 2646.11 **(B)** *m/z* 3496.62 **(C)** *m/z* 4138.87 **(D)** *m/z* 4560.99. Box and whisker plots and ion images were created +/- 0.010% of the m/z, and hotspot removal on the ion images is shown in the color bar (black low to white high) with the triangles representing 0% and 100%.

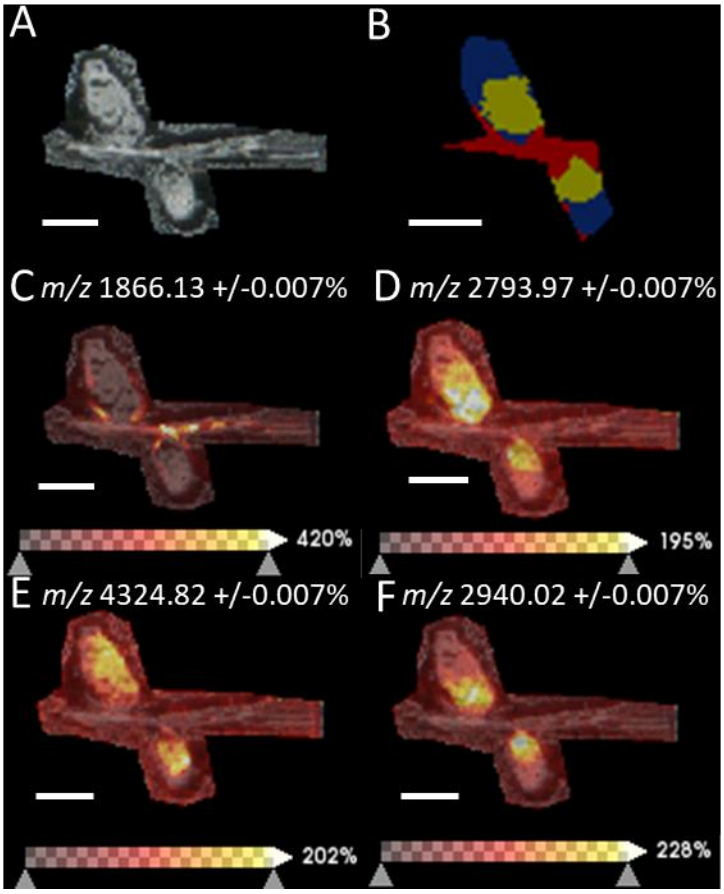



**Figure 3.** Example MALDI-MSI results from the four-step wash. **(A)** optical image, **(B)** segmentation results, **(C-F)** images for *m/z* 1866.13, 2793.97, 4324.82, and 2940.02 respectively. White bar represents 1 mm, and the color bar goes from 0% (black) to 100% (white), indicated by the two triangles, with the degree of hotspot removal indicated by the percentage to the right of the color bar.

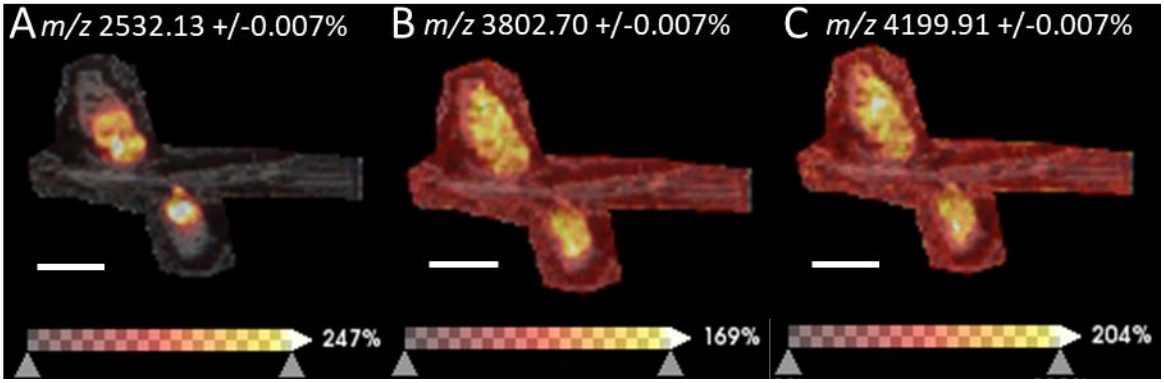



**Figure 4**. MALDI-MSI images of mature NCR peptides identified in the LC-MS/MS results. White bar represents 1 mm, and the color bar goes from 0% (black) to 100% (white), indicated by the two triangles, with the degree of hotspot removal indicated by the percentage to the right of the color bar.

**Supplemental Information**

**Supporting Table S1.** Monoisotopic peaks observed in reflectron mode with the four-step wash. Peaks where the correct monoisotopic peak was more challenging to determine are in italics.

| | | |
|---|---|---|
| 1505.859 | 2736.513 | 3912.242 |
| 1521.812 | 2742.621 | 3920.809 |
| 1543.829 | 2793.965 | 3928.314 |
| 1738.661 | 2940.018 | 3935.731 |
| 1850.119 | 2955.999 | 3944.845 |
| 1866.130 | 3102.089 | 3951.852 |
| 1888.159 | 3118.062 | *3988.155* |
| 1937.162 | 3264.151 | 4100.024 |
| 1953.158 | 3280.111 | 4111.042 |
| 1975.143 | 3387.660 | 4127.040 |
| 1991.115 | 3426.216 | 4138.993 |
| 2035.070 | 3442.172 | 4146.989 |
| 2068.180 | 3459.640 | 4156.837 |
| 2103.839 | 3466.576 | 4199.907 |
| 2233.878 | 3474.679 | *4213.874* |
| 2249.846 | 3496.702 | 4227.764 |
| 2264.968 | 3512.760 | 4238.935 |
| 2442.457 | 3523.622 | *4324.824* |
| 2458.448 | 3532.397 | 4355.993 |
| 2464.479 | 3562.493 | 4379.071 |
| 2480.413 | 3588.261 | 4412.013 |
| 2496.508 | 3604.231 | 4444.926 |
| 2510.430 | 3676.735 | 4446.244 |
| 2532.130 | 3697.781 | 4494.001 |
| 2554.122 | 3734.723 | 4543.949 |
| 2570.077 | 3748.839 | 4561.108 |
| 2607.373 | 3750.316 | 4598.088 |
| 2646.173 | 3764.690 | 4621.099 |
| 2680.201 | 3766.266 | 4734.952 |
| 2684.141 | 3786.739 | *4783.017* |
| 2720.463 | 3802.697 | *4820.364* |

**Supporting Table S2. P**eaks observed in linear mode with the four-step wash.

| | | | |
|---|---|---|---|
| 1868.01 | 3383.55 | 3931.24 | 4351.59 |
| 1939.21 | 3389.44 | 3939.27 | 4360.95 |
| 1955.05 | 3429.51 | 3947.29 | 4383.73 |
| 1977.56 | 3445.58 | 3955.71 | 4417.28 |
| 1993.22 | 3469.86 | 3967.82 | 4436.18 |
| 2070.06 | 3478.09 | 3991.25 | 4449.88 |
| 2235.82 | 3500.39 | 4000.70 | 4472.46 |
| 2252.17 | 3516.64 | 4014.21 | 4487.97 |
| 2267.48 | 3527.70 | 4023.31 | 4498.17 |
| 2444.49 | 3536.27 | 4035.76 | 4548.16 |
| 2460.96 | 3565.69 | 4094.54 | 4565.46 |
| 2482.98 | 3591.49 | 4103.71 | 4587.36 |
| 2513.13 | 3607.75 | 4114.39 | 4603.26 |
| 2535.02 | 3655.78 | 4131.21 | 4625.47 |
| 2556.66 | 3671.10 | 4142.83 | 4653.44 |
| 2572.92 | 3681.37 | 4152.00 | 4663.08 |
| 2649.18 | 3701.48 | 4161.68 | 4729.52 |
| 2686.18 | 3738.54 | 4167.69 | 4738.31 |
| 2722.80 | 3753.57 | 4204.64 | 4786.88 |
| 2738.77 | 3769.35 | 4218.12 | 4801.06 |
| 2797.17 | 3790.43 | 4231.48 | 4824.69 |
| 2942.61 | 3806.11 | 4243.58 | 4855.59 |
| 2958.97 | 3825.51 | 4255.32 | 5258.46 |
| 3105.17 | 3844.16 | 4282.32 | 5317.62 |
| 3121.14 | 3861.85 | 4310.67 | 5489.30 |
| 3267.15 | 3905.78 | 4317.86 | |
| 3283.31 | 3924.15 | 4328.25 | |

**Supporting Table 3.** SignalP predicted NCR peptide database.

| Description | Sequence | Monoisotopic Mass | # Cys | Cys in disulfide bond | Monoisotopic Mass with Disulfide bonds | [M+H]+ |
|---|---|---|---|---|---|---|
| >tr_A0A072UMS5_A0A072UMS5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g088280 PE_4 SV_1 ; MatureChain: 27-42 | ADCDSDYNCKEQLRAS | 1816.735 | 2 | 2 | 1814.721 | 1815.728 |
| >tr_A0A072ULV0_A0A072ULV0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g059755 PE_4 SV_1 ; MatureChain: 24-45 | DCPEELALVMKCINKLCELV ME | 2522.196 | 3 | 2 | 2520.181 | 2521.188 |
| >tr_A0A072UDU7_A0A072UDU7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g037795 PE_4 SV_1 ; MatureChain: 27-49 | KVLKCTPPEVVKCTCLCGKRT LY | 2581.361 | 4 | 4 | 2577.332 | 2578.339 |
| >tr_A0A072THJ9_A0A072THJ9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0115s0030 PE_4 SV_1 ; MatureChain: 23-48 | NIPGARCATDEDCPVGEKCIG GNCVE | 2649.129 | 4 | 4 | 2645.099 | 2646.107 |
| >tr_A0A072VK32_A0A072VK32_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g072780 PE_4 SV_1 ; MatureChain: 24-46 | DCPSYLCHYPKNPECVERECI CW | 2786.141 | 5 | 4 | 2782.112 | 2783.120 |
| >tr_A0A072TI02_A0A072TI02_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide _Fragment_ OS_Medicago truncatula OX_3880 GN_MTR_0149s0080 PE_4 SV_1 ; MatureChain: 28-52 | YRCTTDSDCPPNMCPPGMEPK CVRY | 2862.172 | 4 | 4 | 2858.143 | 2859.150 |
| >tr_A0A072TFT6_A0A072TFT6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0330s0030 PE_4 SV_1 ; MatureChain: 28-53 | VSKIICVLSQKPLCRNHICVCY KLNR | 3029.627 | 4 | 4 | 3025.598 | 3026.605 |
| >tr_A0A072TPM2_A0A072TPM2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g442510 PE_4 SV_1 ; MatureChain: 21-46 | NTEADCPQRFDNIVECLFGICH FYIK | 3074.408 | 3 | 2 | 3072.394 | 3073.401 |
| >tr_A0A072UFK5_A0A072UFK5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g076397 PE_4 SV_1 ; MatureChain: 29-54 | DDDCQEEIGVRKICIREVCRYF AKIH | 3137.520 | 3 | 2 | 3135.506 | 3136.513 |
| >tr_G7L4Z6_G7L4Z6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g102806 PE_4 SV_1 ; MatureChain: 30-57 | HCVIDAHCPRNMCGFHFPPRC VEGDCVC | 3144.288 | 6 | 6 | 3138.245 | 3139.252 |
| >tr_A0A072VK25_A0A072VK25_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g072660 PE_4 SV_1 ; MatureChain: 26-53 | CKTRVDCPVDMCELSQIVWCI DGSCKCL | 3146.385 | 6 | 6 | 3140.342 | 3141.349 |
| >tr_A0A072TWJ1_A0A072TWJ1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g007810 PE_4 SV_1 ; MatureChain: 27-56 | LNCTTASQCVNNRCYLHGKP SCLNGQCACV | 3199.391 | 6 | 6 | 3193.347 | 3194.354 |
| >tr_A0A072U8A7_A0A072U8A7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g033110 PE_4 SV_1 ; MatureChain: 24-51 | TQNYDCPDDMCPFPDISWCN KHNICACA | 3203.237 | 5 | 4 | 3199.208 | 3200.215 |
| >tr_A0A072UB79_A0A072UB79_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g084730 PE_4 SV_1 ; MatureChain: 26-54 | TCITDADCPYDGKCIDGFCRF NVKNNNQV | 3252.409 | 4 | 4 | 3248.380 | 3249.387 |
| >tr_A0A072UEE8_A0A072UEE8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g059325 PE_4 SV_1 ; MatureChain: 27-56 | LTKCETDANCPEISIFSPFFISA LTTVVYL | 3321.655 | 2 | 2 | 3319.640 | 3320.647 |
| >tr_A0A072UBS9_A0A072UBS9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g461840 PE_4 SV_1 ; MatureChain: 29-56 | DCPPVKKPLKMWCIRQTCFY GFGKRPDL | 3325.674 | 3 | 2 | 3323.660 | 3324.667 |
| >tr_A0A072TYZ0_A0A072TYZ0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g445340 PE_4 SV_1 ; MatureChain: 27-57 | EKCGSSPDCEIYKCIPPAKPKC VLFGCMCIT | 3362.537 | 6 | 6 | 3356.493 | 3357.500 |
| >tr_G7I3W5_G7I3W5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g075045 PE_4 SV_1 ; MatureChain: 29-55 | DEDCRRRGSNQYWVYKCINH GCEYVQR | 3376.503 | 3 | 2 | 3374.489 | 3375.496 |
| >tr_A7KHF0_A7KHF0_MEDTR Nodule-specific cysteine-rich peptide 320 OS_Medicago truncatula OX_3880 GN_MTR_2g054250 PE_2 SV_1 ; MatureChain: 28-57 | FVGGSNDECVYPDVFQCINNI CKCVSHHRT | 3383.494 | 4 | 4 | 3379.465 | 3380.472 |
| >tr_G7ZXD6_G7ZXD6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g466410 PE_4 SV_1 ; MatureChain: 30-59 | IPCATSDDCLKNMCRPPLTPR CIEHNCKCK | 3388.546 | 6 | 6 | 3382.502 | 3383.509 |
| >tr_A0A072UB12_A0A072UB12_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g064020 PE_4 SV_1 ; MatureChain: 25-54 | SIPCETTADCPVAVPPEYYKC MYKVCVLIR | 3390.619 | 4 | 4 | 3386.590 | 3387.597 |
| >tr_A0A072U1J7_A0A072U1J7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g064000 PE_4 SV_1 ; MatureChain: 25-55 | APKPCVTVKDCPKTVFPIGYK CIKNLCILVI | 3402.866 | 4 | 4 | 3398.837 | 3399.844 |
| >tr_A0A072TT06_A0A072TT06_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g064010 PE_4 SV_1 ; MatureChain: 21-49 | IPCFGTKDKCPFNLYYKVECID GFCYYPV | 3424.568 | 4 | 4 | 3420.538 | 3421.546 |
| >tr_A0A072VUR1_A0A072VUR1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g055120 PE_4 SV_1 ; MatureChain: 24-54 | SFTEYNLCAPASSPCRIDEIKP QVHGFTEVV | 3436.643 | 2 | 2 | 3434.628 | 3435.635 |
| >tr_A0A072VLD0_A0A072VLD0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g084745 PE_4 SV_1 ; MatureChain: 24-52 | QYSSCATKEECKCPDNKRPAC LWKQCYCY | 3447.463 | 6 | 6 | 3441.419 | 3442.427 |
| >tr_A7KH77_A7KH77_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g083280 PE_2 SV_1 ; MatureChain: 25-56 | DLSPCLRSGDCSKDECPSHLV PKCIGLTCYCI | 3454.552 | 6 | 6 | 3448.508 | 3449.515 |
| >tr_A7KHC7_A7KHC7_MEDTR Nodule-specific cysteine-rich peptide 244 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 33-64 | NNCTDTSDCSSNHCSYEGVSL CMNGQCICIYE | 3492.280 | 6 | 6 | 3486.236 | 3487.243 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_A0A072U0S0_A0A072U0S0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g063970 PE_4 SV_1 ; MatureChain: 25-54 | ENSCVTNDDCPEAVFFVTFRC IKNICVRIR | 3490.661 | 4 | 4 | 3486.632 | 3487.639 |
| >tr_G7IT95_G7IT95_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g066235 PE_4 SV_1 ; MatureChain: 28-58 | LLPCGTDDDCANDPCIHPEYP HCHMEQCHCV | 3494.337 | 6 | 6 | 3488.293 | 3489.301 |
| >tr_G7J0L3_G7J0L3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g014705 PE_2 SV_1 ; MatureChain: 26-55 | VLIDCKTVKDCPPSYTKIYRCI DNKCRLVL | 3498.822 | 4 | 4 | 3494.793 | 3495.800 |
| >tr_A0A072V920_A0A072V920_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g073170 PE_4 SV_1 ; MatureChain: 25-55 | SIPCKTRTQCPEKMCRLPKFV WCIDGSCVCA | 3501.634 | 6 | 6 | 3495.590 | 3496.597 |
| >tr_G7JQH4_G7JQH4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11413600 PE_4 SV_1 ; MatureChain: 25-56 | VKIPCVKDTDCPTLPCPLYSK CVDGFCKMLSI | 3515.706 | 5 | 4 | 3511.677 | 3512.685 |
| >tr_G7I794_G7I794_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g046020 PE_2 SV_1 ; MatureChain: 27-57 | FSGCMNDSDCPDLFCLPPLDM KCHELVCKCR | 3518.475 | 6 | 6 | 3512.431 | 3513.438 |
| >tr_A0A072U652_A0A072U652_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g406350 PE_4 SV_1 ; MatureChain: 23-52 | KDLPFNICEKDEDRLEFCAHD EVAKCLLNI | 3519.683 | 3 | 2 | 3517.669 | 3518.676 |
| >tr_A7KHG8_A7KHG8_MEDTR Nodule-specific cysteine-rich peptide 338 OS_Medicago truncatula OX_3880 GN_MtrunA17_Chr5g0400281 PE_2 SV_1 ; MatureChain: 45-76 | HNCTDISDCSSNHCSYEGVSL CMNGQCICIYE | 3527.332 | 6 | 6 | 3521.288 | 3522.296 |
| >tr_G7JZU8_G7JZU8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g014085 PE_4 SV_1 ; MatureChain: 27-58 | HNCTDISDCSSNHCSYEGVSL CMNGQCICIYE | 3527.332 | 6 | 6 | 3521.288 | 3522.296 |
| >tr_G7JLV0_G7JLV0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g014790 PE_4 SV_1 ; MatureChain: 33-63 | GKTHYSEIIECKNDADCPIGYK CIDEMCKYG | 3525.538 | 4 | 4 | 3521.509 | 3522.516 |
| >tr_G7I8L8_G7I8L8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g058880 PE_4 SV_1 ; MatureChain: 25-56 | EVKPCTYNIDCPLSMCPSPEIP KCVNSICECK | 3542.575 | 6 | 6 | 3536.531 | 3537.539 |
| >tr_A0A072V566_A0A072V566_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g435320 PE_4 SV_1 ; MatureChain: 25-56 | DVVNCTSVLQCFTTYCYLHG TMLCLNGQCLCV | 3543.549 | 6 | 6 | 3537.506 | 3538.513 |
| >tr_A0A072V9P7_A0A072V9P7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g073100 PE_4 SV_1 ; MatureChain: 25-56 | DSILCTTHAQCPGDMCELPQF VWCVVGFCECI | 3546.491 | 6 | 6 | 3540.448 | 3541.455 |
| >tr_A0A072VKW9_A0A072VKW9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g073510 PE_4 SV_1 ; MatureChain: 27-56 | AVHKECKTDVDCRQIWFVTK CINHECQPIL | 3555.724 | 4 | 4 | 3551.695 | 3552.702 |
| >tr_A0A072TD23_A0A072TD23_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1999s0010 PE_4 SV_1 ; MatureChain: 27-57 | KNICIDDVHCQKYKCSPGLYP TCINGWCECK | 3560.584 | 6 | 6 | 3554.540 | 3555.547 |
| >tr_G7KYC2_G7KYC2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g084820 PE_4 SV_2 ; MatureChain: 25-56 | TTSCITDDDCPKAVSFLVFKCI DNICVRVEIL | 3559.743 | 4 | 4 | 3555.714 | 3556.721 |
| >tr_A0A072TRV3_A0A072TRV3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g069245 PE_4 SV_1 ; MatureChain: 27-58 | LDCHSDHECYNQLSCIIEEICL DGSCHCPPIL | 3601.511 | 6 | 6 | 3595.467 | 3596.475 |
| >tr_A0A072TR46_A0A072TR46_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g064150 PE_4 SV_1 ; MatureChain: 20-49 | TIPCKTKVDCPQQIYYVVECL DGFCDYWRD | 3599.623 | 4 | 4 | 3595.594 | 3596.601 |
| >tr_A0A072TS14_A0A072TS14_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g064050 PE_4 SV_1 ; MatureChain: 25-54 | TIPCKTKADCPQHIYYIIECLD GFCNYWRD | 3607.639 | 4 | 4 | 3603.610 | 3604.617 |
| >tr_G7IQM1_G7IQM1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g058625 PE_4 SV_1 ; MatureChain: 26-56 | IPICQTYMDCPSDMCTRPKHA YCVSYKCYCV | 3620.522 | 6 | 6 | 3614.478 | 3615.485 |
| >tr_A0A072V478_A0A072V478_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g022740 PE_4 SV_1 ; MatureChain: 25-55 | HPCLSDHDCYIQYPKTPFGHM ECYKGSCRPI | 3624.586 | 4 | 4 | 3620.557 | 3621.565 |
| >tr_G7KSG9_G7KSG9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g080850 PE_4 SV_1 ; MatureChain: 25-56 | VRIPCVTVADCPPTILPVFYECI DKFCMLHIE | 3663.803 | 4 | 4 | 3659.774 | 3660.781 |
| >tr_A0A072U8J4_A0A072U8J4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g445020 PE_4 SV_1 ; MatureChain: 27-60 | LAGCITDADCVIKKCSSSCRIK CIDFRCLCPTGF | 3668.713 | 7 | 6 | 3662.670 | 3663.677 |
| >tr_A0A072V7Y1_A0A072V7Y1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g046630 PE_4 SV_1 ; MatureChain: 25-55 | FRDPCNFDFDCRNSNCTAPYV ATCMYEHCYC | 3687.390 | 6 | 6 | 3681.346 | 3682.353 |
| >tr_A0A072UW27_A0A072UW27_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g060437 PE_4 SV_1 ; MatureChain: 25-56 | DPIYCKSDADCPKESYPLFVK CVDNFCDFIIV | 3685.685 | 4 | 4 | 3681.656 | 3682.663 |
| >tr_A0A072UJ73_A0A072UJ73_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g026818 PE_4 SV_1 ; MatureChain: 27-57 | QIMFSDCKTDKDCPQFRRANI RCRKGQCVKL | 3686.820 | 4 | 4 | 3682.790 | 3683.798 |
| >tr_A0A072V961_A0A072V961_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g450150 PE_4 SV_1 ; MatureChain: 14-46 | YVVMCEKDSDCVDSFCVPPN VPKCRVVCKCLPK | 3702.723 | 6 | 6 | 3696.679 | 3697.686 |
| >tr_G7JQH0_G7JQH0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g060610 PE_4 SV_1 ; MatureChain: 22-52 | DTSCHSFDDCPWVAHHYRECI EGLCAYRILY | 3701.594 | 4 | 4 | 3697.565 | 3698.572 |
| >tr_A0A072TFZ5_A0A072TFZ5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide _Fragment_ OS_Medicago truncatula OX_3880 GN_MTR_0291s0020 PE_4 SV_1 ; MatureChain: 24-55 | HSFLPCQTKDDCVFDDCKFPK NPVCYLEACHC | 3704.568 | 6 | 6 | 3698.525 | 3699.532 |
| >tr_A7KH67_A7KH67_MEDTR Nodule-specific cysteine-rich peptide 21 OS_Medicago truncatula OX_3880 GN_MtrunA17_Chr7g0243321 PE_2 SV_1 ; MatureChain: 27-57 | RVCKSDKDCKDIIIYRYILKCR NGECVKIKI | 3714.992 | 4 | 4 | 3710.963 | 3711.970 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_A0A072UE17_A0A072UE17_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g459540 PE_4 SV_1 ; MatureChain: 27-59 | LTKCETDANCPEISIFSPFFYKC INNGCVLIML | 3725.767 | 4 | 4 | 3721.738 | 3722.745 |
| >tr_A0A072TZN8_A0A072TZN8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g063420 PE_4 SV_1 ; MatureChain: 27-58 | KCNSDAECRERWIMCPLETV VKCVEDECICVH | 3739.641 | 6 | 6 | 3733.598 | 3734.605 |
| >tr_A0A072UYP4_A0A072UYP4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g088265 PE_4 SV_1 ; MatureChain: 27-59 | FHCHMDYDCYDQITCIIGDVT CLEGSCDCPQDV | 3740.436 | 6 | 6 | 3734.392 | 3735.400 |
| >tr_A0A072VIU4_A0A072VIU4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g052005 PE_4 SV_1 ; MatureChain: 24-56 | KDITCTVAGDCPNFFVCPPNN FVRCIRNLCKCR | 3745.759 | 6 | 6 | 3739.715 | 3740.722 |
| >tr_A7KHC8_A7KHC8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g033260 PE_2 SV_1 ; MatureChain: 26-58 | EKIRRCFNDAHCPPDMCTLGV IPKCSRFTICIC | 3768.767 | 6 | 6 | 3762.723 | 3763.731 |
| >tr_A0A072TZS7_A0A072TZS7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g064010 PE_4 SV_1 ; MatureChain: 25-57 | DRIPCATDADCPPKILPIIHKCI NNFCKLKLYN | 3768.933 | 4 | 4 | 3764.904 | 3765.912 |
| >tr_G7JQH9_G7JQH9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g060730 PE_4 SV_1 ; MatureChain: 25-57 | ELEIRCVSDADCPLFPLPLHNR CIDDVCHLFTS | 3769.772 | 4 | 4 | 3765.743 | 3766.750 |
| >tr_A0A072UIZ7_A0A072UIZ7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g035950 PE_4 SV_1 ; MatureChain: 23-56 | TKTAIACVTNKDCLKFFTPLD NVKCVGNVCEFFL | 3780.875 | 4 | 4 | 3776.845 | 3777.853 |
| >tr_A0A072VB20_A0A072VB20_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g080180 PE_4 SV_1 ; MatureChain: 25-57 | RPCNNVDDCRKHMCTPYGQL VRCINSTCECVLD | 3784.649 | 6 | 6 | 3778.605 | 3779.612 |
| >tr_G7JHR4_G7JHR4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g100690 PE_4 SV_1 ; MatureChain: 24-56 | HSFLPCVTKDDCAYDECISPR KPTCYLETCHCL | 3789.642 | 6 | 6 | 3783.599 | 3784.606 |
| >tr_A0A072V4D2_A0A072V4D2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g015845 PE_4 SV_1 ; MatureChain: 25-56 | YVVCITNNDCPPRTRLILYKC RNRKCVSYSII | 3802.961 | 4 | 4 | 3798.932 | 3799.940 |
| >tr_A0A072U6V1_A0A072U6V1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g007940 PE_4 SV_1 ; MatureChain: 25-57 | DSFRGCNKDTDCPEKFCSSPD VVRCIYIECYCI | 3807.616 | 6 | 6 | 3801.573 | 3802.580 |
| >tr_A0A072UJQ2_A0A072UJQ2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g057680 PE_4 SV_1 ; MatureChain: 29-61 | EDRCRSILDCPQDKCFPLLTLV CTNFICDCLHV | 3811.768 | 6 | 6 | 3805.724 | 3806.732 |
| >tr_A0A072VAB0_A0A072VAB0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g073300 PE_4 SV_1 ; MatureChain: 24-56 | EDIENCETHSDCPHYMCTSPE TPWCVAYQCGCY | 3813.395 | 6 | 6 | 3807.351 | 3808.359 |
| >tr_A0A072UW93_A0A072UW93_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g063740 PE_4 SV_1 ; MatureChain: 27-59 | ALIECQIDDDCPPIKFAKYLCI NYKCRKICLGE | 3817.873 | 5 | 4 | 3813.844 | 3814.851 |
| >tr_A0A072UKL5_A0A072UKL5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g065970 PE_4 SV_1 ; MatureChain: 25-58 | KDGCKTNFDCLIKYPDHNEDI LQCIGGHCLCLTN | 3822.729 | 5 | 4 | 3818.700 | 3819.707 |
| >tr_A0A072ULJ0_A0A072ULJ0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g065960 PE_4 SV_1 ; MatureChain: 25-58 | KDGCKTNFDCLIKYPDHNEDI LQCIGGHCLCLTN | 3822.729 | 5 | 4 | 3818.700 | 3819.707 |
| >tr_A0A072UUY4_A0A072UUY4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g450180 PE_4 SV_1 ; MatureChain: 27-58 | QIRCNDAFECRRSAICNFPNK WKCNDHKCECV | 3829.693 | 6 | 6 | 3823.650 | 3824.657 |
| >tr_G7IW01_G7IW01_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g015870 PE_2 SV_1 ; MatureChain: 25-58 | DIVCITDNDCPPNTLVQGYRCI DGKCESVFLSYR | 3835.767 | 4 | 4 | 3831.738 | 3832.745 |
| >tr_A0A072TC73_A0A072TC73_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1886s0010 PE_4 SV_1 ; MatureChain: 25-58 | KDGCKTKFDCLIKYPDHNEDI LQCIGGHCLCLTN | 3836.781 | 5 | 4 | 3832.752 | 3833.759 |
| >tr_A0A072UVS0_A0A072UVS0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g053570 PE_4 SV_1 ; MatureChain: 25-57 | SSFTRCRNDVECPQYSCLRGL KMKCICFKCMCV | 3850.722 | 7 | 6 | 3844.678 | 3845.685 |
| >tr_A0A072UJP9_A0A072UJP9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g046807 PE_4 SV_1 ; MatureChain: 24-58 | DERECVTDADCQKKYPGPYE HLLKCVSGYCVGVTG | 3861.747 | 4 | 4 | 3857.717 | 3858.725 |
| >tr_A7KHA2_A7KHA2_MEDTR Nodule-specific cysteine-rich peptide 145 OS_Medicago truncatula OX_3880 GN_MtrunA17_Chr4g0029781 PE_2 SV_1 ; MatureChain: 22-54 | EADTSCHSFDDCPWVAHHYR ECIEGLCAYRILY | 3901.674 | 4 | 4 | 3897.645 | 3898.652 |
| >tr_A7KHF5_A7KHF5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g071385 PE_2 SV_1 ; MatureChain: 24-57 | GIRKKECRQDSDCPSYFCEKL TIAKCIHSTCLCK | 3905.854 | 6 | 6 | 3899.810 | 3900.817 |
| >tr_G7K947_G7K947_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g061160 PE_2 SV_1 ; MatureChain: 26-59 | ESIECRTVADCPKLISSKFVIKC IEKRCVAQFFD | 3904.971 | 4 | 4 | 3900.942 | 3901.949 |
| >tr_A0A072VJU2_A0A072VJU2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g070030 PE_4 SV_1 ; MatureChain: 25-61 | GLINGGSVPCLTSFGCPRSTCY PPSTPNCIRRICECI | 3913.822 | 6 | 6 | 3907.779 | 3908.786 |
| >tr_A0A072UKH2_A0A072UKH2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g452900 PE_4 SV_1 ; MatureChain: 27-59 | AFVKCETDDDCPKYNGFRKY ECVNNWCRLTGLH | 3923.763 | 4 | 4 | 3919.734 | 3920.742 |
| >tr_A0A072V9I4_A0A072V9I4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g072130 PE_4 SV_1 ; MatureChain: 25-61 | DLINGGSVPCLTSFGCPRSTCY PPSTPNCILRICECI | 3928.811 | 6 | 6 | 3922.767 | 3923.774 |
| >tr_A0A072V9E8_A0A072V9E8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g070670 PE_4 SV_1 ; MatureChain: 27-61 | LSGCITNEDCIKYQCSAENCM VCINFACKCKYSVF | 3929.675 | 7 | 6 | 3923.632 | 3924.639 |
| >tr_A7KH96_A7KH96_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g060590 PE_2 SV_1 ; MatureChain: 24-57 | RQTDIPCKSDDACPRVSSHHIE CVKGFCTYWKLD | 3936.816 | 4 | 4 | 3932.787 | 3933.794 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_G7JPA2_G7JPA2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g033900 PE_4 SV_1 ; MatureChain: 25-58 | EKECDTDADCRKKFAGANQH LLWCNNGYCECHTH | 3938.643 | 5 | 4 | 3934.614 | 3935.622 |
| >tr_G7KS99_G7KS99_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g068360 PE_4 SV_1 ; MatureChain: 24-57 | AVRKPECRQNSDCPPYFCIKP TVPKCIKFKCLCK | 3941.978 | 6 | 6 | 3935.934 | 3936.942 |
| >tr_G7K5Z0_G7K5Z0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g076255 PE_4 SV_1 ; MatureChain: 27-60 | YIACQSEIDCPPNYSFLFAIRCI KQKCVTVGRYL | 3940.949 | 4 | 4 | 3936.920 | 3937.928 |
| >tr_A0A072TVQ9_A0A072TVQ9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g008020 PE_4 SV_1 ; MatureChain: 27-59 | YRECNSDGDCPNYNCRVKEV GMCYFTKCYCIRL | 3944.669 | 6 | 6 | 3938.625 | 3939.632 |
| >tr_G7ISZ8_G7ISZ8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g059035 PE_4 SV_1 ; MatureChain: 24-60 | ERFGPCSSDSDCPSFLCDHDG VMKCFSNGCSCVDPSD | 3945.481 | 6 | 6 | 3939.437 | 3940.444 |
| >tr_G7KJJ7_G7KJJ7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g044735 PE_4 SV_1 ; MatureChain: 27-58 | YIQCDFDADCPEMFRHIFYLCI DKLCRQFVTL | 3944.822 | 4 | 4 | 3940.792 | 3941.800 |
| >tr_A0A072UE21_A0A072UE21_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g059420 PE_4 SV_1 ; MatureChain: 27-59 | KSICKVDDDCPQRFVMYPLM FMCIKNICRLVNE | 3950.886 | 4 | 4 | 3946.857 | 3947.865 |
| >tr_G7J0L2_G7J0L2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g014705 PE_4 SV_2 ; MatureChain: 26-59 | VLIDCKTVKDCPTSYTKIYRC EDNKCRFSFVIGL | 3956.966 | 4 | 4 | 3952.936 | 3953.944 |
| >tr_G7JZA9_G7JZA9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g048310 PE_4 SV_1 ; MatureChain: 25-60 | QECKDDGDCPTNMCLPSLVS KCINFICECTHSMSTD | 3966.604 | 6 | 6 | 3960.560 | 3961.567 |
| >tr_I3SP41_I3SP41_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0032s0170 PE_2 SV_1 ; MatureChain: 23-56 | KKTDIPCDSRNDCPQQILPRY VLCVNGLCRIYFP | 3966.972 | 4 | 4 | 3962.943 | 3963.951 |
| >tr_A0A072UU77_A0A072UU77_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g436990 PE_4 SV_1 ; MatureChain: 25-58 | GLFRCKVDIDCPQILCFDEQIA KCIARMCECDYE | 3969.772 | 6 | 6 | 3963.728 | 3964.735 |
| >tr_A0A072VU91_A0A072VU91_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g053640 PE_4 SV_1 ; MatureChain: 30-63 | TTPTPCRTDQDCPRKKKFSVT CRKGFCAEIRHVY | 3970.953 | 4 | 4 | 3966.924 | 3967.932 |
| >tr_A0A072UHA9_A0A072UHA9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g013160 PE_4 SV_1 ; MatureChain: 24-59 | APLMLKPCITDKDCPKHRGVN GKCRKGYCVGVGRNY | 3975.999 | 4 | 4 | 3971.969 | 3972.977 |
| >tr_A0A072TQ17_A0A072TQ17_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0672s0010 PE_4 SV_1 ; MatureChain: 23-56 | KQTNIPCKSAEDCPKPIYPRYV LCSYGFCRIFFP | 3982.939 | 4 | 4 | 3978.910 | 3979.917 |
| >tr_A7KHC0_A7KHC0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g070690 PE_2 SV_1 ; MatureChain: 25-58 | KSYGPCTTLQDCETHNWFEV CSCIDFECKCWSLL | 3987.674 | 6 | 6 | 3981.630 | 3982.638 |
| >tr_G7KAI9_G7KAI9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g064860 PE_4 SV_2 ; MatureChain: 27-61 | YVECETDADCQPNMCKWPFI VQCYKNVIKIGPVGP | 3986.853 | 4 | 4 | 3982.824 | 3983.831 |
| >tr_A0A072V9R1_A0A072V9R1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g073310 PE_4 SV_1 ; MatureChain: 24-56 | NEIEHCETHSDCPHYMCDSPQ TPWCMAYECWCF | 3994.441 | 6 | 6 | 3988.398 | 3989.405 |
| >tr_A7KH82_A7KH82_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g016130 PE_2 SV_1 ; MatureChain: 26-59 | YVNCETDADCPPSTRVKRFKC VKGECRWTRMSYA | 3998.847 | 4 | 4 | 3994.817 | 3995.825 |
| >tr_A7KHE4_A7KHE4_MEDTR Nodule-specific cysteine-rich peptide 314 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 26-59 | YVNCETDADCPPSTRVKRFKC VKGECRWTRMSYA | 3998.847 | 4 | 4 | 3994.817 | 3995.825 |
| >tr_A0A072UVD6_A0A072UVD6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g015665 PE_4 SV_1 ; MatureChain: 27-60 | YIVCITDNDCPENTEVRQYECI EGRCRLSRVLNP | 3999.869 | 4 | 4 | 3995.840 | 3996.848 |
| >tr_G7KEA4_G7KEA4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g072455 PE_4 SV_1 ; MatureChain: 27-60 | SYYGCETDADCPRSMNKDFY LKCVDKKCEWTAKI | 4009.781 | 4 | 4 | 4005.752 | 4006.759 |
| >tr_G7IVN8_G7IVN8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g008446 PE_4 SV_1 ; MatureChain: 27-59 | KWPSCKEAIDCGENFCISRFR VKCWRFKCFCIL | 4014.916 | 6 | 6 | 4008.872 | 4009.879 |
| >tr_A7KH98_A7KH98_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g016090 PE_2 SV_1 ; MatureChain: 25-58 | YVVCRTVDDCPPDTRDLRYR CLNGKCKSYRLSYG | 4013.912 | 4 | 4 | 4009.882 | 4010.890 |
| >tr_A0A072UGC5_A0A072UGC5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g006500 PE_4 SV_1 ; MatureChain: 23-56 | EDFPFHKCEKDEDCLEICADD QMAMCILNVCFCY | 4019.602 | 6 | 6 | 4013.558 | 4014.565 |
| >tr_A7KH94_A7KH94_MEDTR Nodule-specific cysteine-rich peptide 122 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 24-59 | TRCNRDEDCPFICTGPQIPKCV SHICFCLSSGKEAY | 4019.791 | 6 | 6 | 4013.748 | 4014.755 |
| >tr_G7J0L0_G7J0L0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g014695 PE_4 SV_1 ; MatureChain: 26-60 | LIDCKTVDDCPSSWTKIYKCID NKCRYSVVKGLII | 4019.039 | 4 | 4 | 4015.010 | 4016.017 |
| >tr_A0A072VAZ7_A0A072VAZ7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g084725 PE_4 SV_1 ; MatureChain: 21-56 | QHSLIPCKTNMDCAKDVCLN YKFPTCVGKKCYCLSA | 4023.866 | 6 | 6 | 4017.823 | 4018.830 |
| >tr_G7KA04_G7KA04_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g061640 PE_4 SV_1 ; MatureChain: 27-60 | LIECYTYFDCPKDMCIFPSIVV CVRNQCDCVEMF | 4025.737 | 6 | 6 | 4019.693 | 4020.700 |
| >tr_A0A072VEF3_A0A072VEF3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g027750 PE_4 SV_1 ; MatureChain: 23-58 | SSLCISDEDCPEALPLMFVKCV CGACEYYTQIQDEK | 4029.752 | 5 | 4 | 4025.723 | 4026.730 |
| >tr_A7KHD4_A7KHD4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g016440 PE_2 SV_1 ; MatureChain: 27-61 | SVECIANIDCPQIFMLPFVMRC INFRCQIVNSEDT | 4047.884 | 4 | 4 | 4043.855 | 4044.862 |
| >tr_G7IVN7_G7IVN7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g008445 PE_4 SV_1 ; MatureChain: 27-61 | KWPSCKEAIDCGINFCIRPFKA KCMMFTCFCVQNP | 4057.848 | 6 | 6 | 4051.805 | 4052.812 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_A7KH97_A7KH97_MEDTR Nodule-specific cysteine-rich peptide 129 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 27-61 | EMTTTTIPCTFIDDCPKMPLVVKCIDNFCNYFEIK | 4071.887 | 4 | 4 | 4067.858 | 4068.865 |
| >tr_A0A072UL49_A0A072UL49_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g063760 PE_4 SV_1 ; MatureChain: 27-60 | ELINCITDDDCPKSEFRKWVCINNICRKMCRTKV | 4072.959 | 5 | 4 | 4068.930 | 4069.938 |
| >tr_A0A072V4B1_A0A072V4B1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11420593 PE_4 SV_1 ; MatureChain: 33-68 | LPVCQTAADCPEGTRGRTYKCINNKCRYPKLLKPIQ | 4077.089 | 4 | 4 | 4073.060 | 4074.067 |
| >tr_G7KA66_G7KA66_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g062510 PE_4 SV_1 ; MatureChain: 27-61 | EHNECETDADCPKHTTIFFVMKCIDHICRCMKTSI | 4097.805 | 5 | 4 | 4093.776 | 4094.783 |
| >tr_G7K1L6_G7K1L6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g056735 PE_4 SV_1 ; MatureChain: 22-57 | ILECIFDIDCPTKKCAPPLVAKCDMYECYCRCPPNN | 4107.822 | 7 | 6 | 4101.779 | 4102.786 |
| >tr_A0A072UWX6_A0A072UWX6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g053600 PE_4 SV_1 ; MatureChain: 27-62 | IHCNDVNDCPPDISNPFVRCESNRCIYSRLEPPFGC | 4108.811 | 5 | 4 | 4104.781 | 4105.789 |
| >tr_G7I5S5_G7I5S5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g011245 PE_4 SV_2 ; MatureChain: 24-57 | YVPCITVADCPPNTWFKIYRCEKGICRYHKLWIV | 4114.060 | 4 | 4 | 4110.031 | 4111.038 |
| >tr_G7JP95_G7JP95_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g033830 PE_4 SV_1 ; MatureChain: 23-59 | NMVVLGCVSDDDCPKVPLPRFLKCIANLCCLVRKKDL | 4117.120 | 5 | 4 | 4113.091 | 4114.098 |
| >tr_A7KHD5_A7KHD5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g069545 PE_2 SV_1 ; MatureChain: 27-62 | WTCVEDSDCPANICQPPMQRMCFYGECACVRSKFCT | 4120.665 | 7 | 6 | 4114.622 | 4115.629 |
| >tr_A0A072V668_A0A072V668_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g041380 PE_4 SV_1 ; MatureChain: 25-60 | EKECISDAVCYEKYPGPFNFIMNCVDGYCKAFPKLV | 4119.895 | 4 | 4 | 4115.866 | 4116.873 |
| >tr_A7KH64_A7KH64_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g092720 PE_2 SV_1 ; MatureChain: 23-59 | AMFELTKSTIRCVTDADCPNVVKPLKPKCVDGFCEYT | 4120.980 | 4 | 4 | 4116.951 | 4117.958 |
| >tr_A7KHF6_A7KHF6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g071345 PE_2 SV_1 ; MatureChain: 24-57 | GIRRFECRQDSDCPSYFCEKLTVPKCFWSKCYCK | 4124.864 | 6 | 6 | 4118.821 | 4119.828 |
| >tr_A0A072U1S4_A0A072U1S4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g066720 PE_4 SV_1 ; MatureChain: 24-59 | KEDDIECVTDADCYEKLPALQRAVMKCIQGFCKIHI | 4124.986 | 4 | 4 | 4120.957 | 4121.964 |
| >tr_G7K731_G7K731_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g071880 PE_4 SV_1 ; MatureChain: 27-61 | ECINDIDCPQTGNLFYVFICKNRICELINKYPQNL | 4132.988 | 4 | 4 | 4128.959 | 4129.966 |
| >tr_A0A072UKQ0_A0A072UKQ0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g067130 PE_4 SV_1 ; MatureChain: 24-58 | RTPCVSDDDCLEAFLPPIWKCVHNFCEFNIRNIFN | 4154.926 | 4 | 4 | 4150.897 | 4151.904 |
| >tr_A0A072VJ77_A0A072VJ77_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g061110 PE_4 SV_1 ; MatureChain: 27-61 | YLKCKTVHDCPKSQVVYRCVGNYCRAVKIRRWNLG | 4155.137 | 4 | 4 | 4151.108 | 4152.116 |
| >tr_I3S7Z9_I3S7Z9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g028980 PE_2 SV_1 ; MatureChain: 24-61 | TTTFPCVSDDDCPVPLPPPFAKCVDGICEFFIKAQVEK | 4155.970 | 4 | 4 | 4151.941 | 4152.948 |
| >tr_A0A072ULH6_A0A072ULH6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g065820 PE_4 SV_1 ; MatureChain: 24-58 | DKRECDTNFDCQQKFSTQAEDLLWCIRGYCMSIPN | 4156.820 | 4 | 4 | 4152.791 | 4153.799 |
| >tr_A7KHC4_A7KHC4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g006240 PE_2 SV_1 ; MatureChain: 27-63 | QKRLHGCKEDRDCDNICSVHAVTKCIGNMCRCLANVK | 4159.956 | 6 | 6 | 4153.912 | 4154.919 |
| >tr_A0A072TT10_A0A072TT10_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g064070 PE_4 SV_1 ; MatureChain: 20-54 | DFPCKTKDDCAQQIDYIAECIIGFCRYFTPFEHPF | 4159.861 | 4 | 4 | 4155.832 | 4156.839 |
| >tr_A0A072V8L8_A0A072V8L8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g054830 PE_4 SV_1 ; MatureChain: 31-67 | GLAHFCNEYNPCPEHLCMPDMKVVCTAEELCLCIKLI | 4178.896 | 6 | 6 | 4172.852 | 4173.859 |
| >tr_G7LJA1_G7LJA1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g069130 PE_4 SV_1 ; MatureChain: 23-59 | GEIPYHQCKFDMECMLMKCVPGKVNVCSLGRCYCVNS | 4183.843 | 6 | 6 | 4177.799 | 4178.807 |
| >tr_A0A072U2G2_A0A072U2G2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g064030 PE_4 SV_1 ; MatureChain: 20-54 | DFPCKTKADCLQHIYYIVECIFGFCQYFKPLKHSV | 4183.023 | 4 | 4 | 4178.994 | 4180.001 |
| >tr_G7J1T4_G7J1T4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g027180 PE_4 SV_2 ; MatureChain: 23-58 | LVQNECVTDGDCRRLYPHLIPRYPMCNEGTCVCIFE | 4185.902 | 5 | 4 | 4181.873 | 4182.880 |
| >tr_G7IDP6_G7IDP6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g072096 PE_4 SV_2 ; MatureChain: 23-58 | IRTQISCVSDDDCPKVPYPLYIKCEDNFCDIWASPY | 4195.903 | 4 | 4 | 4191.874 | 4192.882 |
| >tr_G7L160_G7L160_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g027125 PE_4 SV_1 ; MatureChain: 28-65 | DAPPCLFTPECPPDMCPTDLTLKCINLTCQCTSEYDID | 4204.782 | 6 | 6 | 4198.739 | 4199.746 |
| >tr_A0A072U1G7_A0A072U1G7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g063400 PE_4 SV_1 ; MatureChain: 27-63 | KCVSDAECSGQYMCPTLTVIKCIKDECVCIHYDHDKQ | 4204.852 | 6 | 6 | 4198.809 | 4199.816 |
| >tr_A0A072U2G4_A0A072U2G4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g064100 PE_4 SV_1 ; MatureChain: 20-54 | DFPCKTKVDCPQHKKYIAECIFGFCRHFKPLEHPF | 4208.040 | 4 | 4 | 4204.011 | 4205.019 |
| >tr_A0A072UEW4_A0A072UEW4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g076435 PE_4 SV_1 ; MatureChain: 26-62 | HNPCKTVVDCPPPHPYGFVVAVKCVRRLCIYNVHLHT | 4211.131 | 4 | 4 | 4207.102 | 4208.109 |
| >tr_G7KU61_G7KU61_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g070645 PE_4 SV_1 ; MatureChain: 25-60 | KKKCASESDCYTMFPVPHFIVMTCIEKKCHITGIYY | 4213.004 | 4 | 4 | 4208.974 | 4209.982 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_A0A072UK90_A0A072UK90_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g063690 PE_4 SV_1 ; MatureChain: 23-60 | DTDPFAFCIKDSNCGQDLCTS PNEVPECRLLKCQCIKS | 4219.881 | 6 | 6 | 4213.837 | 4214.845 |
| >tr_A0A072UES3_A0A072UES3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g072275 PE_4 SV_1 ; MatureChain: 29-64 | DECQIDADCPKSGNLFYIYKCI NHKCELVAAHLRFY | 4218.978 | 4 | 4 | 4214.949 | 4215.956 |
| >tr_A0A072UBM1_A0A072UBM1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g071720 PE_4 SV_1 ; MatureChain: 25-63 | FEECKEDADCHPVCSVPGCSN ICTLPDVPTCIDNNCFCI | 4232.709 | 8 | 6 | 4226.665 | 4227.673 |
| >tr_G7KV51_G7KV51_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g071720 PE_4 SV_2 ; MatureChain: 31-69 | FEECKEDADCHPVCSVPGCSN ICTLPDVPTCIDNNCFCI | 4232.709 | 8 | 6 | 4226.665 | 4227.673 |
| >tr_G7KV52_G7KV52_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g071720 PE_4 SV_1 ; MatureChain: 27-65 | FEECKEDADCHPVCSVPGCSN ICTLPDVPTCIDNNCFCI | 4232.709 | 8 | 6 | 4226.665 | 4227.673 |
| >tr_A7KHE5_A7KHE5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g069530 PE_2 SV_1 ; MatureChain: 24-61 | DDGSFCFKDSDCPDEMCPSPL KEMCYFLQCKCGVDTIA | 4239.708 | 6 | 6 | 4233.664 | 4234.671 |
| >tr_A0A072TQE0_A0A072TQE0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g465280 PE_4 SV_1 ; MatureChain: 22-57 | ADCYKRYPRWSLLPNYCIEGS CYSDFLNSGKKYLSP | 4235.990 | 3 | 2 | 4233.976 | 4234.983 |
| >tr_A0A072TS18_A0A072TS18_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g064110 PE_4 SV_1 ; MatureChain: 20-54 | DIPCKTKADCTQHRKYFAECI FGFCRHFKHLEHPF | 4252.016 | 4 | 4 | 4247.987 | 4248.994 |
| >tr_G7LDP2_G7LDP2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g036850 PE_4 SV_1 ; MatureChain: 25-60 | SECISDTDCNVLYPMYINRRL RCIQGICHTTTARRR | 4257.059 | 4 | 4 | 4253.030 | 4254.037 |
| >tr_G7JH11_G7JH11_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g026750 PE_4 SV_1 ; MatureChain: 24-60 | SFSQMINFRGCKRDKDCPQFR GVNIRCRSGFCTPIDS | 4268.007 | 4 | 4 | 4263.977 | 4264.985 |
| >tr_A0A072UC03_A0A072UC03_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g464870 PE_4 SV_1 ; MatureChain: 27-61 | WRPDCKENNDCPTFYCATWI NTCIKFKCYCIRPWG | 4273.891 | 6 | 6 | 4267.847 | 4268.855 |
| >tr_A0A072V6P6_A0A072V6P6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g046590 PE_4 SV_1 ; MatureChain: 27-61 | FRNPCKSDSDCTKLKLKCDRH TILTCFWRHCYCIF | 4280.018 | 6 | 6 | 4273.974 | 4274.982 |
| >tr_G7J0L4_G7J0L4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g014720 PE_4 SV_1 ; MatureChain: 26-60 | VLIDCKHVRDCPKGIWRSCRY KCIDNKCVFTYWPH | 4281.083 | 5 | 4 | 4277.053 | 4278.061 |
| >tr_A0A072UGB9_A0A072UGB9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g006275 PE_4 SV_1 ; MatureChain: 23-59 | QEVLVIHECNRDRDCHISCVP PELPKCIAHMCFCFNS | 4284.927 | 6 | 6 | 4278.884 | 4279.891 |
| >tr_G7KPK3_G7KPK3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g027155 PE_4 SV_1 ; MatureChain: 23-58 | EKTHVRCITADDCPKVERPLK MKCIGNYCHYFLNNF | 4285.051 | 4 | 4 | 4281.022 | 4282.029 |
| >tr_A0A072V6Y7_A0A072V6Y7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g053590 PE_4 SV_1 ; MatureChain: 27-62 | LHCNNDNECPPSTWKPFVRC KMNRCIYSRVQPPWAC | 4291.956 | 5 | 4 | 4287.927 | 4288.935 |
| >tr_A7KH93_A7KH93_MEDTR Nodule-specific cysteine-rich peptide 121 OS_Medicago truncatula OX_3880 GN_MtrunA17_Chr2g0306741 PE_2 SV_1 ; MatureChain: 27-62 | LFECNRDFVCGNDDECVYPY AVQCIHRYCKCLKSRN | 4302.898 | 6 | 6 | 4296.855 | 4297.862 |
| >tr_A0A072VIH1_A0A072VIH1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g050260 PE_4 SV_1 ; MatureChain: 25-62 | SIITCESDEDCPTSFCIPPQIPKC RTICECITKTQHLR | 4308.017 | 6 | 6 | 4301.974 | 4302.981 |
| >tr_A7KHB7_A7KHB7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0240s0040 PE_2 SV_1 ; MatureChain: 28-64 | RECHANSHCVGKITCVLPQKP ECWNYACVCYDSNKYR | 4316.925 | 6 | 6 | 4310.882 | 4311.889 |
| >tr_G7K901_G7K901_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g059740 PE_4 SV_1 ; MatureChain: 27-64 | QRTCKTDADCPKIPSLYPTIYK CLDGICRFSEAKLLIL | 4315.224 | 4 | 4 | 4311.194 | 4312.202 |
| >tr_G7JRC1_G7JRC1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g035268 PE_4 SV_1 ; MatureChain: 18-55 | STTNTKTTIPCKFDNDCPEISY PLILMCIDDFCEYLLA | 4329.990 | 4 | 4 | 4325.961 | 4326.968 |
| >tr_A0A072UFX8_A0A072UFX8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g459510 PE_4 SV_1 ; MatureChain: 25-62 | MVSTNAYIHRCIHQDDCPKY MCEISVLPECINGFCTCV | 4334.888 | 6 | 6 | 4328.844 | 4329.851 |
| >tr_A0A072TV89_A0A072TV89_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0015s0230 PE_4 SV_1 ; MatureChain: 27-64 | TPQKKVTDVSHPCTTDKDCPK VQHGYKLRCRKGQCVHI | 4333.181 | 4 | 4 | 4329.152 | 4330.159 |
| >tr_A0A072TZ06_A0A072TZ06_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g445930 PE_4 SV_1 ; MatureChain: 23-59 | KPGSPNLRNHSCLRDKDCPQL KFHNIRCRNGFCVDIR | 4336.157 | 4 | 4 | 4332.128 | 4333.135 |
| >tr_A0A072U556_A0A072U556_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g006220 PE_4 SV_1 ; MatureChain: 23-60 | QEVLQYELFDCNEDRDCDNVI CVAGGIPKCITPFCFCF | 4345.896 | 6 | 6 | 4339.852 | 4340.859 |
| >tr_A0A072VAA4_A0A072VAA4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g084915 PE_4 SV_1 ; MatureChain: 26-61 | ESFHPCKINEHCTTYKCLLTG QPWCFMDFCLCMYFN | 4351.828 | 6 | 6 | 4345.785 | 4346.792 |
| >tr_G7KNA5_G7KNA5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g060370 PE_4 SV_1 ; MatureChain: 23-58 | EHLKIRCVTDDDCPKVEKPLY MYCGNHWCAYKLHFV | 4353.045 | 4 | 4 | 4349.016 | 4350.023 |
| >tr_A7KH99_A7KH99_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g063520 PE_2 SV_1 ; MatureChain: 23-60 | AITNIRCVSDDDCPKVIKPLV MKCIGNYCYFFMIYEGP | 4358.077 | 4 | 4 | 4354.048 | 4355.056 |
| >tr_A7KH73_A7KH73_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g042940 PE_2 SV_1 ; MatureChain: 25-62 | MLVTCEDHFDCRQNVQQVGC SFREIPQCINSICKCMKG | 4361.942 | 6 | 6 | 4355.898 | 4356.906 |
| >tr_A7KHA1_A7KHA1_MEDTR Nodule-specific cysteine-rich peptide 144 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 34-69 | AYIECEVDDDCPKPMKNSHPD TYYKCVKHRCQWAWK | 4385.957 | 4 | 4 | 4381.928 | 4382.935 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_G7K672_G7K672_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g068810 PE_2 SV_1 ; MatureChain: 27-62 | AYIECEVDDDCPKPMKNSHPDTYYKCVKHRCQWAWK | 4385.957 | 4 | 4 | 4381.928 | 4382.935 |
| >tr_G7JP15_G7JP15_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g015750 PE_4 SV_1 ; MatureChain: 27-62 | ERTCKEDFDCRMRYCVYPTIPLCDVKHCRCRRPPNL | 4415.061 | 6 | 6 | 4409.017 | 4410.024 |
| >tr_G7KU13_G7KU13_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g045520 PE_2 SV_1 ; MatureChain: 27-67 | INKCSQDSHCPKDMCKKPSKPRCVVSPKLPLSSKSGVCTCV | 4417.169 | 6 | 6 | 4411.125 | 4412.132 |
| >tr_A0A072UYE1_A0A072UYE1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g065750 PE_4 SV_1 ; MatureChain: 24-59 | YFADILCKVHEDCPQKSTHKYYCIDDECFLYYWEAP | 4434.941 | 4 | 4 | 4430.911 | 4431.919 |
| >tr_G7J277_G7J277_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g061750 PE_4 SV_1 ; MatureChain: 27-64 | FRCLRNLDCPDSMCSSAYTPRCRHRTCVCLNNDEIKIL | 4445.056 | 6 | 6 | 4439.012 | 4440.019 |
| >tr_A0A072UIP6_A0A072UIP6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g033290 PE_4 SV_1 ; MatureChain: 24-61 | GNHTYWCVTTDDCATNICRSGLTEECWVFRCICKYETK | 4447.925 | 6 | 6 | 4441.881 | 4442.888 |
| >tr_A0A072U6R1_A0A072U6R1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g006245 PE_4 SV_1 ; MatureChain: 25-63 | DDYLKYIYRCQNDGDCDQICATHGISKCVATMCFCNLNL | 4448.912 | 6 | 6 | 4442.868 | 4443.876 |
| >tr_A0A072UWB0_A0A072UWB0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g451370 PE_4 SV_1 ; MatureChain: 25-62 | NSFNSKIVFTDCKTDKDCQNHRGFNFRCRKGNCVAKIR | 4449.157 | 4 | 4 | 4445.128 | 4446.135 |
| >tr_A7KH78_A7KH78_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g055160 PE_2 SV_1 ; MatureChain: 23-60 | IKIFTEHRCRTDADCPARELPEYLKCQGGMCRLLIKKD | 4449.236 | 4 | 4 | 4445.207 | 4446.214 |
| >tr_A7KHD7_A7KHD7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11429208 PE_2 SV_1 ; MatureChain: 20-58 | NNIEDDIFCITDNDCPPNTLVQRYRCINGKCNLSFVSYG | 4453.023 | 4 | 4 | 4448.994 | 4450.001 |
| >tr_G7K8Z9_G7K8Z9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g059670 PE_4 SV_1 ; MatureChain: 27-64 | YPGCETDAECPKIYELYPLIYKCENKFCILSQVLPYIV | 4457.174 | 4 | 4 | 4453.145 | 4454.152 |
| >tr_A7KHA9_A7KHA9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0240s0070 PE_2 SV_1 ; MatureChain: 26-64 | TATFCHDDSHCVTKIKCVLPRTPQCRNEACGCYHSNKFR | 4468.032 | 6 | 6 | 4461.988 | 4462.996 |
| >tr_Q1RU35_Q1RU35_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g038715 PE_2 SV_1 ; MatureChain: 24-62 | SEFIFTKKLTSCDSSKDCRSFLCYSPKFPVCKRGICECI | 4468.121 | 6 | 6 | 4462.078 | 4463.085 |
| >tr_G7KV47_G7KV47_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g071690 PE_2 SV_1 ; MatureChain: 25-65 | YPFQECKVDADCPTVCTLPGCPDICSFPDVPTCIDNNCFCT | 4472.824 | 8 | 6 | 4466.780 | 4467.788 |
| >tr_A0A072UA23_A0A072UA23_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g461820 PE_4 SV_1 ; MatureChain: 24-61 | TELNIPCVAVDDCPKVEKPLNMWCMHQYCVYGFIKPYK | 4474.115 | 4 | 4 | 4470.086 | 4471.093 |
| >tr_G7J9F3_G7J9F3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g065710 PE_4 SV_1 ; MatureChain: 18-55 | MPCKTDKECPNTSTHKYKCINDDCFCFYIYWPLGNSLV | 4475.985 | 5 | 4 | 4471.956 | 4472.963 |
| >tr_A0A072UAP3_A0A072UAP3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g463320 PE_4 SV_1 ; MatureChain: 23-62 | REPTKIPCVSDSDCHKVKKPLLLTCIDGICQYTLEATPFD | 4476.220 | 4 | 4 | 4472.190 | 4473.198 |
| >tr_A0A072UQA7_A0A072UQA7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g059775 PE_4 SV_1 ; MatureChain: 27-63 | NFECKSTNDCPKSVLRVWRCINNYCRPVRMKKIHNTM | 4483.189 | 4 | 4 | 4479.159 | 4480.167 |
| >tr_G7KNA1_G7KNA1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g060320 PE_4 SV_1 ; MatureChain: 18-55 | SIKTKIACVTDNDCPRAIKPVVMWCINNYCHYYLYGYQ | 4485.124 | 4 | 4 | 4481.094 | 4482.102 |
| >tr_A0A072UMI8_A0A072UMI8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g066290 PE_4 SV_1 ; MatureChain: 33-71 | AECYQDSDCPEDMCSYLAKPTCIFTEYFPIFWMAVCGCD | 4493.781 | 6 | 6 | 4487.737 | 4488.744 |
| >tr_A0A072UMD4_A0A072UMD4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g065720 PE_4 SV_1 ; MatureChain: 24-63 | GGNECVTDVDCEKLYPGNKKPLICNIGYCLSLYKEPPRYM | 4522.150 | 4 | 4 | 4518.121 | 4519.128 |
| >tr_A7KH88_A7KH88_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g010490 PE_2 SV_1 ; MatureChain: 24-62 | GGKKIYCENAASCPRLMYPLVYKCLDNKCVKFMMKSRFV | 4533.251 | 4 | 4 | 4529.222 | 4530.229 |
| >tr_A7KHB9_A7KHB9_MEDTR Nodule-specific cysteine-rich peptide 206 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 26-64 | KECVTDADCENLYPGNKKPMFCNNTGYCMSLYKEPSRYM | 4541.959 | 4 | 4 | 4537.929 | 4538.937 |
| >tr_G7IX62_G7IX62_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g033965 PE_4 SV_1 ; MatureChain: 24-62 | DEEDCFVDADCVNLILCDFDEKPKCIINICQCFPWTVIH | 4545.016 | 6 | 6 | 4538.973 | 4539.980 |
| >tr_A0A072VTT4_A0A072VTT4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g052005 PE_4 SV_1 ; MatureChain: 24-63 | KDITCTVAGDCPNFFVCPPNNFVRCIRNLCKCRSLSYKQP | 4549.177 | 6 | 6 | 4543.133 | 4544.140 |
| >tr_A0A072VAL4_A0A072VAL4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g070770 PE_4 SV_1 ; MatureChain: 25-63 | EFENYCTKDSDCESYCPNPKYGKCLDNKCICQLIWMGYG | 4559.900 | 6 | 6 | 4553.857 | 4554.864 |
| >tr_G7KNS7_G7KNS7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g061820 PE_4 SV_1 ; MatureChain: 23-61 | EKLDIRCATVDDCPKVTKPVVMMCTGKFCHYFFVRKQIL | 4561.300 | 4 | 4 | 4557.271 | 4558.278 |
| >tr_A7KH84_A7KH84_MEDTR Nodule-specific cysteine-rich peptide 90 OS_Medicago truncatula OX_3880 GN_MtrunA17_Chr5g0429921 PE_2 SV_1 ; MatureChain: 30-66 | SDDECKIDGDCPISWQKFHTYKCINQKCKWVLRFHEY | 4562.106 | 4 | 4 | 4558.077 | 4559.084 |
| >tr_G7KE56_G7KE56_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g070410 PE_4 SV_1 ; MatureChain: 24-60 | SDDECKIDGDCPISWQKFHTYKCINQKCKWVLRFHEY | 4562.106 | 4 | 4 | 4558.077 | 4559.084 |
| >tr_A7KHA7_A7KHA7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g029760 PE_2 SV_1 ; MatureChain: 24-61 | EDIGHIKYCGIVDDCYKSKKPLFKIWKCVENVCVLWYK | 4562.302 | 4 | 4 | 4558.273 | 4559.280 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_A7KH86_A7KH86_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g055370 PE_2 SV_1 ; MatureChain: 27-65 | FQPCVTTADCMKTLKTDENI WYECINDFCIPFPIPKGRK | 4564.176 | 4 | 4 | 4560.147 | 4561.154 |
| >tr_I3S2V0_I3S2V0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g464450 PE_2 SV_1 ; MatureChain: 24-63 | EITNIPCVSDESCPQVIKPLVIK CIDKFCEYFMEGEYEGP | 4565.122 | 4 | 4 | 4561.093 | 4562.100 |
| >tr_A7KH83_A7KH83_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g063600 PE_2 SV_1 ; MatureChain: 26-64 | IDIFVCQTDADCPKSELSMYT WKCIDNECNLFKVMQQMV | 4588.062 | 4 | 4 | 4584.033 | 4585.040 |
| >tr_G7KJJ8_G7KJJ8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g044730 PE_4 SV_1 ; MatureChain: 27-65 | NTFLMADNIECDTDAGCPKM VHHIFYKCIDNKCKQFRRS | 4591.114 | 4 | 4 | 4587.085 | 4588.093 |
| >tr_A0A072V3F1_A0A072V3F1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g103440 PE_4 SV_1 ; MatureChain: 27-63 | KSKLYFHCFKFPFSSKFYRHIC HNNRCKIVMFMPPNV | 4593.282 | 3 | 2 | 4591.267 | 4592.274 |
| >tr_G7JXA2_G7JXA2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g057910 PE_4 SV_1 ; MatureChain: 27-65 | LKKCITFEDCPISKTRVYKCLH GECRYTIPYIPKVPKVK | 4596.460 | 4 | 4 | 4592.431 | 4593.439 |
| >tr_A7KHG7_A7KHG7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g065015 PE_2 SV_1 ; MatureChain: 25-64 | ESKLEQTCSEDFECYIKNPHVP FGHLRCFEGFCQQLNGPA | 4600.070 | 4 | 4 | 4596.041 | 4597.048 |
| >tr_A0A072UME6_A0A072UME6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g065840 PE_4 SV_1 ; MatureChain: 29-69 | CDTDVDCQKKYPGPFEHLLK CIHGYCVCFPRNPGDSSGYPI | 4601.073 | 5 | 4 | 4597.044 | 4598.051 |
| >tr_G7J0N2_G7J0N2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g031335 PE_4 SV_1 ; MatureChain: 27-67 | GIRCHDVSECPKGLYCNVGSH MECVKHQCKCIKNFEPIDLA | 4603.118 | 6 | 6 | 4597.074 | 4598.081 |
| >tr_A7KHC5_A7KHC5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g065025 PE_2 SV_1 ; MatureChain: 25-64 | VSKLAQSCSEDFECYIKNPHA PFGQLRCFEGYCQRLDKPT | 4607.149 | 4 | 4 | 4603.120 | 4604.127 |
| >tr_A0A072VI49_A0A072VI49_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g053975 PE_4 SV_1 ; MatureChain: 27-66 | SSPIRVTDVSRPCTTDKDCSKV EYGYKLRCRKGRCVHIPR | 4608.340 | 4 | 4 | 4604.311 | 4605.319 |
| >tr_A0A072TFE5_A0A072TFE5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0330s0040 PE_4 SV_1 ; MatureChain: 26-65 | TAIHECRAHSHCVARINCVLP RKPQCRNYACGCYDSNKYR | 4635.160 | 6 | 6 | 4629.117 | 4630.124 |
| >tr_A0A072U6G3_A0A072U6G3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g011210 PE_4 SV_1 ; MatureChain: 27-67 | SSRCKSDADCDCPPPFKPKCR FRSCTCSNLKVWKGPELSFI | 4635.177 | 6 | 6 | 4629.133 | 4630.141 |
| >tr_A0A072TR97_A0A072TR97_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0330s0020 PE_4 SV_1 ; MatureChain: 26-64 | ARYECREDSHCVTKVKCGLP RTPKCRNYICFCHNPNKYI | 4644.200 | 6 | 6 | 4638.156 | 4639.163 |
| >tr_G7IJE8_G7IJE8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g075085 PE_4 SV_2 ; MatureChain: 27-67 | GEREEITPCNTDFNCPSSLCTF PLKLLKPICTNHQCICQHI | 4645.171 | 6 | 6 | 4639.128 | 4640.135 |
| >tr_G7L159_G7L159_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g027120 PE_4 SV_1 ; MatureChain: 24-65 | QSIPDVLPCLFSNECPPDLCPT DLFAKCINLTCQCTAEYDLD | 4647.069 | 6 | 6 | 4641.026 | 4642.033 |
| >tr_G7K8Y8_G7K8Y8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g059440 PE_4 SV_1 ; MatureChain: 24-62 | AQNDWMKCKTDDECPKVSN PPLYFKCIDRGCRIVIKMRF | 4647.250 | 4 | 4 | 4643.221 | 4644.228 |
| >tr_G7L3W6_G7L3W6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g102160 PE_4 SV_1 ; MatureChain: 25-63 | GPIFCYNDDDCPHICSHPRVQ KCRMFLCHCEEVEEKDEK | 4651.993 | 6 | 6 | 4645.949 | 4646.956 |
| >tr_G7I616_G7I616_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g044500 PE_4 SV_1 ; MatureChain: 27-65 | VHRCIEYTDCPEDMCHLPLVV VCHDHICKCLRLIKIRSY | 4652.259 | 6 | 6 | 4646.215 | 4647.222 |
| >tr_A7KH72_A7KH72_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g042910 PE_2 SV_1 ; MatureChain: 25-63 | KRIPCKDNNDCNNNWQLLAC RFEREVPRCINSICKCMPM | 4654.155 | 6 | 6 | 4648.111 | 4649.118 |
| >tr_A0A072U8T2_A0A072U8T2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g043300 PE_4 SV_1 ; MatureChain: 27-69 | FDGHGGPPCTKVSDCPPNLVC KPGYKLGCSANYQCICYMKM SF | 4658.051 | 6 | 6 | 4652.007 | 4653.015 |
| >tr_G7INW8_G7INW8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g008910 PE_4 SV_1 ; MatureChain: 27-67 | AHIKCETDDDCPKSLNNLLAI KCIDHVCKFVSNLSQFEDLF | 4665.237 | 4 | 4 | 4661.208 | 4662.215 |
| >tr_A7KHF9_A7KHF9_MEDTR Nodule-specific cysteine-rich peptide 329 OS_Medicago truncatula OX_3880 GN_MtrunA17_Chr6g0487941 PE_2 SV_1 ; MatureChain: 34-72 | ARFECREDSHCVTRIKCVLPR KPECRNYACGCYDSNKYR | 4668.159 | 6 | 6 | 4662.116 | 4663.123 |
| >tr_A7KH69_A7KH69_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g063580 PE_2 SV_1 ; MatureChain: 26-64 | LYIIECKTDADCPISKLNMYN WRCIKSSCHLYKVIQFMV | 4668.289 | 4 | 4 | 4664.260 | 4665.267 |
| >tr_G7JM81_G7JM81_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g057120 PE_2 SV_1 ; MatureChain: 21-60 | DGIPCVFDYNCPEISYYPVRCN VNNICEYNLNVDLVEEIE | 4669.079 | 4 | 4 | 4665.050 | 4666.057 |
| >tr_G7I4S7_G7I4S7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g075500 PE_4 SV_1 ; MatureChain: 23-59 | EHYRELICKTDDNCPRRGTNK YFIHKCIDYRCQWIPR | 4669.246 | 4 | 4 | 4665.217 | 4666.224 |
| >tr_G7K5W7_G7K5W7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g076040 PE_4 SV_1 ; MatureChain: 23-63 | EEDIGGHLECVEDEDCMEESC PIFSVHKCKNSGCECDEMFR | 4680.828 | 6 | 6 | 4674.784 | 4675.792 |
| >tr_G7L6H2_G7L6H2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g065635 PE_4 SV_1 ; MatureChain: 23-63 | GNFFEFFHKCTQDSDCPSLLC RNKSELPKCIAGFMCRCPNV | 4684.107 | 6 | 6 | 4678.063 | 4679.071 |
| >tr_A7KHA5_A7KHA5_MEDTR Nodule-specific cysteine-rich peptide 159 OS_Medicago truncatula OX_3880 GN_MTR_6g091600 PE_2 SV_1 ; MatureChain: 26-66 | AILECREDSHCVTKIKCVLPRK PECRNNACTCYKGGFSFHH | 4692.221 | 6 | 6 | 4686.177 | 4687.185 |
| >tr_A0A072TIJ0_A0A072TIJ0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0050s0100 PE_4 SV_1 ; MatureChain: 23-62 | NIRINILFFSFDNTDKIGCKTSE DCPYLGKCIEDFCQFKK | 4692.252 | 4 | 4 | 4688.223 | 4689.230 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_G7JDM1_G7JDM1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g065390 PE_4 SV_1 ; MatureChain: 24-63 | SERECVTDDDCEKLYPTNEYRMMCDSGYCMNLLNEPPCNI | 4692.901 | 5 | 4 | 4688.872 | 4689.879 |
| >tr_A7KHG6_A7KHG6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g010630 PE_2 SV_1 ; MatureChain: 28-67 | YLCVTDSHCPPHMCPPGMEPRCVRRMCKCLPIGWRKYFVP | 4705.192 | 6 | 6 | 4699.148 | 4700.156 |
| >tr_A7KHD3_A7KHD3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g069100 PE_2 SV_1 ; MatureChain: 27-67 | NIMNCQSTFDCPRDMCSHIRDVICIFKKCKCAGGRYMPQVP | 4707.141 | 6 | 6 | 4701.097 | 4702.104 |
| >tr_G7IV71_G7IV71_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g062820 PE_4 SV_1 ; MatureChain: 28-71 | HACTVNADCEQSMCDPFCVGGYHFTPICVIGWCVCVGNRVAPVL | 4713.068 | 7 | 6 | 4707.024 | 4708.032 |
| >tr_G7JBD8_G7JBD8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g084910 PE_2 SV_1 ; MatureChain: 24-64 | DDVKIKCVSAIDCMDLFNLLPIVYKCINNICVYEQSSQRLI | 4718.365 | 4 | 4 | 4714.336 | 4715.343 |
| >tr_A7KH89_A7KH89_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g030060 PE_2 SV_1 ; MatureChain: 27-67 | KPFLTRPYPCNTGSDCPQNMCPPGYKPGCEDGYCNHCYKRW | 4725.003 | 6 | 6 | 4718.960 | 4719.967 |
| >tr_A0A072UAN2_A0A072UAN2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g461980 PE_4 SV_1 ; MatureChain: 21-60 | DTKQTNIPCENKRDCPQPLYPKFVTCFEGLCRMHYPLKKI | 4753.346 | 4 | 4 | 4749.317 | 4750.324 |
| >tr_G7I4L5_G7I4L5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g042200 PE_4 SV_1 ; MatureChain: 25-63 | IRIHCKDDFDCIENRLQVGCRLQREKPRCVNLVCRCLRR | 4755.428 | 6 | 6 | 4749.385 | 4750.392 |
| >tr_G7JJW3_G7JJW3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g071890 PE_4 SV_1 ; MatureChain: 23-64 | DKFVFDKNGADRCRSILDCPQDKCFPLLTLVCVNFACDCLHV | 4760.250 | 6 | 6 | 4754.206 | 4755.213 |
| >tr_A0A072U9J8_A0A072U9J8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g453200 PE_4 SV_1 ; MatureChain: 27-69 | CNPCLVTCPDDLLNRCPPGMEPICEYGVIKCYPIGKETNRVLT | 4766.253 | 6 | 6 | 4760.209 | 4761.216 |
| >tr_G7KAD9_G7KAD9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g063460 PE_4 SV_1 ; MatureChain: 26-66 | AYIIECQTDDDCPKSQLEMFAWKCVKNGCHLFGMYEDDDDP | 4771.998 | 4 | 4 | 4767.969 | 4768.976 |
| >tr_A0A072TE63_A0A072TE63_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0888s0020 PE_4 SV_1 ; MatureChain: 17-57 | LPNYYEFYHCYNHSDCQGSMCPTGSKPKCVDQVCECILIRM | 4774.037 | 6 | 6 | 4767.993 | 4769.000 |
| >tr_A7KHC6_A7KHC6_MEDTR Nodule-specific cysteine-rich peptide 239 OS_Medicago truncatula OX_3880 GN_MtrunA17_Chr0c32g0494191 PE_2 SV_1 ; MatureChain: 28-69 | EESHYMKFSICKDDTDCPTLFCVLPNVPKCIGSKCHCKLMVN | 4775.188 | 6 | 6 | 4769.144 | 4770.151 |
| >tr_A7KHE7_A7KHE7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g042850 PE_2 SV_1 ; MatureChain: 25-64 | LPISCKDHFECRRKINILRCIYRQEKPMCINSICTCVKLL | 4779.463 | 6 | 6 | 4773.420 | 4774.427 |
| >tr_G7IV70_G7IV70_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g062775 PE_4 SV_1 ; MatureChain: 28-70 | IPCNDDVDCPQTLCEQLIADFKYMIDFKSECVSRMCACTGSPV | 4787.088 | 6 | 6 | 4781.045 | 4782.052 |
| >tr_G7J0N1_G7J0N1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g031340 PE_4 SV_1 ; MatureChain: 27-67 | GMRCNHVSDCPKDTFCWLDSHMQCIKHQCKCVRIFEPIDPA | 4790.138 | 6 | 6 | 4784.095 | 4785.102 |
| >tr_G7IGL5_G7IGL5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g048885 PE_4 SV_1 ; MatureChain: 24-65 | HDCHNDLDCGDKIKCVPPRIALCINYKCYCILENDAVIPWST | 4790.224 | 6 | 6 | 4784.180 | 4785.188 |
| >tr_A0A072UKJ5_A0A072UKJ5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g065720 PE_4 SV_1 ; MatureChain: 16-56 | NECVTDVDCEKLYPGNKKPLICNIGYCLSLYKGNFFFIYYI | 4799.318 | 4 | 4 | 4795.289 | 4796.296 |
| >tr_A0A072UI64_A0A072UI64_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide-like protein OS_Medicago truncatula OX_3880 GN_MTR_4g036540 PE_4 SV_1 ; MatureChain: 31-71 | WRVDHDPYKSCHSDEDCTKIHYYCPPSKVPYCQVDRCGCGN | 4803.027 | 6 | 6 | 4796.984 | 4797.991 |
| >tr_A7KHA3_A7KHA3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g073530 PE_2 SV_1 ; MatureChain: 33-72 | LETIECETDGDCPRSMIKMWNKNYRHKCIDGKCEWIKKLP | 4810.298 | 4 | 4 | 4806.269 | 4807.276 |
| >tr_A0A072U1H1_A0A072U1H1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g063450 PE_4 SV_1 ; MatureChain: 27-67 | IRPGRECVSDYDCYIKYTILWKYNNVCTKGVCYTLLDAVHP | 4811.337 | 4 | 4 | 4807.308 | 4808.315 |
| >tr_G7J0M9_G7J0M9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g031320 PE_2 SV_1 ; MatureChain: 27-68 | GIRCRNVYDCPKATYCRAGSHRVQCIKHQCKCVRIFESIDPA | 4823.338 | 6 | 6 | 4817.294 | 4818.302 |
| >tr_A7KH80_A7KH80_MEDTR Nodule-specific cysteine-rich peptide 76 OS_Medicago truncatula OX_3880 GN_MtrunA17_Chr4g0029901 PE_2 SV_1 ; MatureChain: 25-66 | DRIPCVTNGDCPVMRLPLYMRCITYSCELFFDGPNLCAVERI | 4822.269 | 5 | 4 | 4818.240 | 4819.247 |
| >tr_G7KA73_G7KA73_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11431999 PE_4 SV_1 ; MatureChain: 22-62 | MEVGRRANVECESDKDCQEHWSEFFIIQCIDNICVPSERPL | 4824.186 | 4 | 4 | 4820.157 | 4821.164 |
| >tr_A7KHA6_A7KHA6_MEDTR Nodule-specific cysteine-rich peptide 165 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 24-63 | LKIFCIDVADCPKDLYPLLYKCIYNKCIVFTRIPFPFDWI | 4825.495 | 4 | 4 | 4821.466 | 4822.473 |
| >tr_G7IDP5_G7IDP5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g072095 PE_4 SV_2 ; MatureChain: 24-64 | YRTRIPCVSDYDCPKASYPLFIKCIYNFCEIWGSPTWDATN | 4835.232 | 4 | 4 | 4831.202 | 4832.210 |
| >tr_A0A072UX97_A0A072UX97_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g466130 PE_4 SV_1 ; MatureChain: 30-70 | YKNRCFRDSDCPKEMCNHPKIPKCVNNAYCKCVVAMYFPPK | 4839.231 | 6 | 6 | 4833.188 | 4834.195 |
| >tr_G7IW98_G7IW98_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g020980 PE_4 SV_1 ; MatureChain: 25-66 | GKDKCFRDSDCPKHMCPSSLVAKCINRLCRCRRPELQVQLNP | 4843.368 | 6 | 6 | 4837.324 | 4838.331 |
| >tr_G7IWG0_G7IWG0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g025420 PE_4 SV_1 ; MatureChain: 23-63 | DGIILCKDHFDCYENIRKLRCDFDTEKPFCISLNVCQCIKQ | 4865.292 | 6 | 6 | 4859.249 | 4860.256 |
| >tr_A7KH75_A7KH75_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g042895 PE_2 SV_1 ; MatureChain: 25-66 | RPVSCKDHYDCRRKVKIVGCIFPQEKPMCINSMCTCIREIVP | 4894.400 | 6 | 6 | 4888.356 | 4889.363 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_A0A072VH74_A0A072VH74_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g046170 PE_4 SV_1 ; MatureChain: 26-66 | YVPCRRNADCLGEKCLPPKR YWCRIITEPYEDFPMGRCDCI | 4908.271 | 6 | 6 | 4902.227 | 4903.234 |
| >tr_A0A072VHS4_A0A072VHS4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g046150 PE_4 SV_1 ; MatureChain: 27-69 | FSICQNNSDCKDQEICLPPKKH WCNKIVPVMIEETMVGNCECI | 4938.247 | 6 | 6 | 4932.203 | 4933.210 |
| >tr_G7IV77_G7IV77_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g062880 PE_4 SV_1 ; MatureChain: 27-70 | IRNKCFRPSDCPPSMYCDAGF QIGCVRKICTCLRILAPIDFVPT | 4946.417 | 6 | 6 | 4940.373 | 4941.380 |
| >tr_G7IY11_G7IY11_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g052100 PE_4 SV_1 ; MatureChain: 23-64 | VRIPRPLIDPLNCHIDIHCIYKE CRRPFKPSCLNFKCDCGKE | 4968.478 | 6 | 6 | 4962.434 | 4963.441 |
| >tr_A0A072UK35_A0A072UK35_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide-like protein OS_Medicago truncatula OX_3880 GN_MTR_4g036530 PE_4 SV_1 ; MatureChain: 31-71 | WRVDHDPYKSCHSDEDCVKI HFYCPPPKVPYCRVDRCDCFN | 4971.169 | 6 | 6 | 4965.125 | 4966.133 |
| >tr_A0A072UYL2_A0A072UYL2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g062865 PE_4 SV_1 ; MatureChain: 27-69 | GIRCVFPSDCPRTMRCLSEFHL TCKKKQCKCVKMFDPINFVT A | 4979.420 | 6 | 6 | 4973.377 | 4974.384 |
| >tr_A0A072UKS5_A0A072UKS5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g048520 PE_4 SV_1 ; MatureChain: 22-65 | LVIIDHHKPCVSDTDCAFYLDI PPTVKYCSDGLCAWYFPDNPL P | 4980.331 | 4 | 4 | 4976.301 | 4977.309 |
| >tr_A0A072ULL8_A0A072ULL8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g066270 PE_4 SV_1 ; MatureChain: 22-64 | KCYQDSDCPKDMCSFLFKPK CIFTRYFPIYLGGICGCDRKTC P | 4989.277 | 7 | 6 | 4983.233 | 4984.241 |
| >tr_A0A072TQP5_A0A072TQP5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g442760 PE_4 SV_1 ; MatureChain: 24-67 | DDYHLKCTTDYDCREGFCPE GLAPKCFVSFALARFLSEGRC LCI | 4991.267 | 6 | 6 | 4985.223 | 4986.230 |
| >tr_G7KUV1_G7KUV1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g045910 PE_4 SV_2 ; MatureChain: 20-63 | EKECACVADCIYKYPTLRDLV VKCIEGYCKAILYRKGWVGG VQD | 4997.477 | 5 | 4 | 4993.448 | 4994.455 |
| >tr_A7KH81_A7KH81_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g065050 PE_2 SV_1 ; MatureChain: 17-59 | LFLVEMGHLSIHCVSVDDCPK VEKPITMKCINNYCKYFVDHK L | 5008.464 | 4 | 4 | 5004.435 | 5005.442 |
| >tr_G7JY89_G7JY89_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g047670 PE_4 SV_1 ; MatureChain: 26-66 | YKCNIDVDCPITPSPKFKWKCI NKRCLYIRFDEIWTSDPRE | 5019.469 | 4 | 4 | 5015.440 | 5016.447 |
| >tr_A0A072U9K0_A0A072U9K0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g445080 PE_4 SV_1 ; MatureChain: 26-72 | RITHDPSTRSTVSGGFGKCVR DADCVDEVCSPGCNKRCVGF ECQCPL | 5031.243 | 7 | 6 | 5025.199 | 5026.207 |
| >tr_A0A072TVV1_A0A072TVV1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0003s0180 PE_4 SV_1 ; MatureChain: 24-67 | NDTEYTDCLQHSDCQAYACE LPFKPDCLMVEYAPQFFRLAC GCV | 5036.139 | 6 | 6 | 5030.095 | 5031.102 |
| >tr_A0A072TDM7_A0A072TDM7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0683s0010 PE_4 SV_1 ; MatureChain: 27-69 | SPVLCQRNYECYEQICLPPKK HWCNILELVRINGFYLGLCAC I | 5041.475 | 6 | 6 | 5035.432 | 5036.439 |
| >tr_A7KHH0_A7KHH0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g016470 PE_2 SV_1 ; MatureChain: 23-63 | KKKRYIECETHEDCSQVFMPP FVMRCVIHECKIFNGEHLRY | 5042.409 | 4 | 4 | 5038.380 | 5039.387 |
| >tr_A0A072TXJ7_A0A072TXJ7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g033870 PE_4 SV_1 ; MatureChain: 22-70 | SGPTPSSLHPSGAVNCLGPVPC DQCPQRCIAIGANSYICGILTC CCYYN | 5045.222 | 8 | 6 | 5039.179 | 5040.186 |
| >tr_G7ZZE8_G7ZZE8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g011830 PE_4 SV_1 ; MatureChain: 25-65 | IDPPHHITNHEIPCKYNHDCPTI LDYISICPYHYCEFWRTY | 5046.281 | 4 | 4 | 5042.252 | 5043.259 |
| >tr_G7I3P8_G7I3P8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g041915 PE_4 SV_1 ; MatureChain: 22-66 | TNIEGTMSCFHDADCVHKRC QLPQIPKCVGKKCRCRGQYQ ANPMG | 5050.330 | 6 | 6 | 5044.287 | 5045.294 |
| >tr_A0A072TR69_A0A072TR69_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g467170 PE_4 SV_1 ; MatureChain: 25-70 | EPGGHRCSTDSFCPPNMCPPG MTPKCVRFRCKCVPIGWKNL SHVLA | 5055.365 | 6 | 6 | 5049.321 | 5050.328 |
| >tr_G7KU03_G7KU03_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g045410 PE_4 SV_2 ; MatureChain: 27-71 | KPLSLNECTQDYDCPIEMCPFP FQPKCIMLKNLSIFSNSGICSCT | 5057.298 | 6 | 6 | 5051.254 | 5052.261 |
| >tr_A0A072UWG7_A0A072UWG7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g436830 PE_4 SV_1 ; MatureChain: 26-68 | GRYTTPWCVRDIDCPKEKCK HPFKPRCLTHSCVCRLWGSQD VI | 5060.442 | 6 | 6 | 5054.399 | 5055.406 |
| >tr_G7IX58_G7IX58_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g033930 PE_4 SV_1 ; MatureChain: 27-68 | YCLVDADCVTLVCNFDEKRK CLRSTCVCRKFRFTGFYYEQL H | 5071.388 | 6 | 6 | 5065.344 | 5066.352 |
| >tr_A0A072UG60_A0A072UG60_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g011500 PE_4 SV_1 ; MatureChain: 24-64 | FFVDTPCKIDEDCPQFQRPWS QIVKYYCIADQCFYYIKHIK | 5077.410 | 4 | 4 | 5073.381 | 5074.388 |
| >tr_G7I3U8_G7I3U8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g074860 PE_4 SV_1 ; MatureChain: 27-67 | YPPCETVADCPESYFRIYRCEN NFCRYREAVRRLRPPLRKK | 5093.526 | 4 | 4 | 5089.497 | 5090.504 |
| >tr_G7L162_G7L162_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g027180 PE_4 SV_1 ; MatureChain: 25-71 | SLPDAPPCLFTPECPPDMCPTD LTLKCINLSCQCTIEYDIDPDV VPS | 5108.289 | 6 | 6 | 5102.245 | 5103.252 |
| >tr_G7KEA2_G7KEA2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g072420 PE_4 SV_1 ; MatureChain: 24-67 | VHDVAHTDIPCEPDADCPKSL HEYFEMKCIDKKCEWSRKTSL IP | 5113.390 | 4 | 4 | 5109.361 | 5110.368 |
| >tr_I3SFJ9_I3SFJ9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g090485 PE_2 SV_1 ; MatureChain: 26-69 | VIFECSEDSHCVTKIKCVLPRK PECRNTQCTCYRGYKGSFTLH H | 5115.422 | 6 | 6 | 5109.378 | 5110.385 |
| >tr_A7KHB4_A7KHB4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g057160 PE_2 SV_1 ; MatureChain: 27-69 | IPCLSDDECPEMSHYSFKCNN KICEYDLGEMSDDDYYLEMS RE | 5116.050 | 4 | 4 | 5112.021 | 5113.029 |
| >tr_A0A072UTU0_A0A072UTU0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide-like protein OS_Medicago truncatula OX_3880 GN_MTR_3g433530 PE_4 SV_1 ; MatureChain: 28-74 | LDRKRPIGKKFPCKADKDCGC GNGERPECLGGQCWCFSPPSS ANKHT | 5120.387 | 6 | 6 | 5114.343 | 5115.350 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_G7K944_G7K944_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g061120 PE_4 SV_1 ; MatureChain: 29-71 | FLRCDFDLDCPPKMCYSHLYF VPMCVDNHCDCTQWKDIIPTI P | 5121.262 | 6 | 6 | 5115.218 | 5116.225 |
| >tr_A0A072TJN6_A0A072TJN6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0003s0210 PE_4 SV_1 ; MatureChain: 24-67 | IDKEYTVCSLHSDCKAYVCQL PLKPECILLEYVPHFYRLTCSC V | 5121.464 | 6 | 6 | 5115.420 | 5116.428 |
| >tr_A0A072U9D0_A0A072U9D0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g045910 PE_4 SV_1 ; MatureChain: 15-59 | EKECACVADCIYKYPTLRDLV VKCIEGYCKAILYRKVSLTTL KTC | 5145.627 | 6 | 6 | 5139.583 | 5140.590 |
| >tr_A0A072TNG5_A0A072TNG5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2125s0010 PE_4 SV_1 ; MatureChain: 25-69 | QCIYPACFKDHMCRQLKCSPG RTPKCVNYQCRCSPQALGSY HLLT | 5146.392 | 7 | 6 | 5140.348 | 5141.355 |
| >tr_G7KV48_G7KV48_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g071695 PE_4 SV_1 ; MatureChain: 25-72 | YSVYGCNDDTDCPPSCTTRGC PDSCAYPHVLRCIGKNCAENK NGIAEL | 5152.200 | 7 | 6 | 5146.157 | 5147.164 |
| >tr_G7K1I6_G7K1I6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g056360 PE_4 SV_1 ; MatureChain: 27-69 | FVKCETTDDCPKSDYIRQYEC VNNWCRLARLHEFQPKKSTL TS | 5151.446 | 4 | 4 | 5147.417 | 5148.424 |
| >tr_A0A072UJR7_A0A072UJR7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g044562 PE_4 SV_1 ; MatureChain: 27-69 | AFAGWIKCKVDEDCPNVFTY SYYKCVNELCEIFLREIPKKPY M | 5152.470 | 4 | 4 | 5148.441 | 5149.449 |
| >tr_G7IX57_G7IX57_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g033925 PE_4 SV_1 ; MatureChain: 24-69 | AKEEVNCLDDADCLEVLCVF GSKAECVVNICICVPPRFGKFD EHFR | 5159.414 | 6 | 6 | 5153.370 | 5154.378 |
| >tr_A7KHB5_A7KHB5_MEDTR Nodule-specific cysteine-rich peptide 189 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 24-69 | LNGGGKDKCFRDSDCPKHMC PSSLVAKCINRLCRCRRPELQ VQLNP | 5184.538 | 6 | 6 | 5178.494 | 5179.501 |
| >tr_G7IX60_G7IX60_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g033955 PE_4 SV_1 ; MatureChain: 25-68 | KDDCLVDADCVTLVCEFDER PQCVINTCRCRPLRFSGFYYE QLH | 5198.363 | 6 | 6 | 5192.320 | 5193.327 |
| >tr_G7KJJ5_G7KJJ5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g044700 PE_4 SV_1 ; MatureChain: 27-69 | KTFLMAEYIKCDTDADCPIVI HHSFYKCIDNLCKRFRRQKHL V | 5197.606 | 4 | 4 | 5193.576 | 5194.584 |
| >tr_G7IYS3_G7IYS3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11420609 PE_4 SV_1 ; MatureChain: 24-66 | VDYIKIFCIDVADCPKDLYPLL YKCIYNKCIVFTRIPFPFDWI | 5202.653 | 4 | 4 | 5198.624 | 5199.632 |
| >tr_A0A072UKI9_A0A072UKI9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g055095 PE_4 SV_1 ; MatureChain: 27-71 | VIECNQHSDCPKDMCQFHLKP NCILMKVRLSNFFPNFYDGIC GCD | 5221.333 | 6 | 6 | 5215.289 | 5216.296 |
| >tr_A0A072TQJ0_A0A072TQJ0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_8g468070 PE_4 SV_1 ; MatureChain: 27-70 | VHFRCFRDSDCVKLYCRPPLK SKCMYKTNCKCIAVYTQEDY GLT | 5224.470 | 6 | 6 | 5218.427 | 5219.434 |
| >tr_A0A072UUT3_A0A072UUT3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g006650 PE_4 SV_1 ; MatureChain: 26-70 | ERIYRCLDHSHCPTFMCSPGL KPKCMNPKVCKCVPVQSRKY YALT | 5229.527 | 6 | 6 | 5223.483 | 5224.490 |
| >tr_G7IWA5_G7IWA5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g021050 PE_4 SV_1 ; MatureChain: 24-70 | LNGHGKDRCFKDSDCPKYMC PSSLVAKCIKKLCSCRKPGLQI QLNPK | 5249.640 | 6 | 6 | 5243.596 | 5244.603 |
| >tr_A7KHE0_A7KHE0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g023900 PE_2 SV_1 ; MatureChain: 29-73 | QEIENGIHPCKKNEDCNHMCV MPGLPWCHENNLCFCYENAY GNTR | 5253.199 | 6 | 6 | 5247.156 | 5248.163 |
| >tr_A0A072VDG0_A0A072VDG0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g010310 PE_4 SV_1 ; MatureChain: 22-65 | HNFFLVAQNIDAIIKCETVADC PPDTEHKKYRCVKNICIYRWF L | 5253.617 | 4 | 4 | 5249.588 | 5250.595 |
| >tr_A0A072VKY3_A0A072VKY3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g061100 PE_4 SV_1 ; MatureChain: 27-70 | YVWCETVEDCFKSQYFIFDCI NNQCINVGKNPKEPRYPGIPR DQ | 5255.440 | 4 | 4 | 5251.411 | 5252.418 |
| >tr_A0A072UKX9_A0A072UKX9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g463200 PE_4 SV_1 ; MatureChain: 23-69 | AFERTETRMLTIPCTSDANCP KVISPCHTKCFDGFCGWYIEG SYEGP | 5259.336 | 5 | 4 | 5255.307 | 5256.314 |
| >tr_A0A072TEN9_A0A072TEN9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0512s0030 PE_4 SV_1 ; MatureChain: 22-68 | RKGGPPGGRTYIPCISDDDCIV AQPPYVLLCVNNFCTYFRDD DLPQR | 5286.514 | 4 | 4 | 5282.485 | 5283.492 |
| >tr_G7IMB5_G7IMB5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g104570 PE_2 SV_1 ; MatureChain: 26-70 | IPMIHPLLYKKRVVPNCQTIVD CPDNMCTHPKEVYCIGYRCY CLK | 5296.583 | 6 | 6 | 5290.539 | 5291.547 |
| >tr_A0A072UBT3_A0A072UBT3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g462060 PE_4 SV_1 ; MatureChain: 18-66 | AIERTEPMLTTYLILCVSEADC PKVVKPNYTMCAGGICWQSV QGSNQGP | 5300.514 | 4 | 4 | 5296.485 | 5297.492 |
| >tr_G7JYH7_G7JYH7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g066750 PE_4 SV_1 ; MatureChain: 31-75 | EPEPKFIECVTDADCLNSQSK MYALICEKNRCIYEFLKSMHY NLS | 5302.461 | 4 | 4 | 5298.432 | 5299.439 |
| >tr_A0A072V7Y8_A0A072V7Y8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g046680 PE_4 SV_1 ; MatureChain: 25-71 | GTGNIRQSCEFDVDCENKYCP PSHDGKCVWEEEGEEGEEEY CGCIPR | 5325.146 | 6 | 6 | 5319.102 | 5320.109 |
| >tr_A0A072VID4_A0A072VID4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g450280 PE_4 SV_1 ; MatureChain: 27-71 | HFQCHIDSECENQIKCVLPRV AKCVRYKCDCVRFDAEQDPW SART | 5325.460 | 6 | 6 | 5319.417 | 5320.424 |
| >tr_A0A072VK98_A0A072VK98_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g073670 PE_4 SV_1 ; MatureChain: 24-68 | LNFGPCNTHEDCQKYEYTYSC VMECVESVCDCWSNEIVDFIF PRN | 5329.203 | 6 | 6 | 5323.160 | 5324.167 |
| >tr_G7KT88_G7KT88_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g070090 PE_4 SV_1 ; MatureChain: 24-69 | TDVHKNVSLSFFIAARVCKSD KDCKDIIIYRYILKCRNGECVK IKI | 5344.843 | 4 | 4 | 5340.814 | 5341.821 |
| >tr_G7KM16_G7KM16_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g055700 PE_4 SV_1 ; MatureChain: 24-71 | NGTICDGDHDCSRNVCSHPQQ VWCIFITRVVPRLRRMGLCSC SSKLAP | 5357.549 | 6 | 6 | 5351.505 | 5352.512 |
| >tr_G7K939_G7K939_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g061060 PE_4 SV_1 ; MatureChain: 27-72 | FLKCDSDLDCPPKMCYSHLSF VPLCVDNHCDCIQWKFKNNIP KAFP | 5358.475 | 6 | 6 | 5352.431 | 5353.438 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_G7KDX9_G7KDX9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g040380 PE_4 SV_1 ; MatureChain: 31-76 | RGPIVTFWCQTNDACEHRFPN HGYKCINHLCQPPKISHNDLQ FIST | 5364.537 | 4 | 4 | 5360.508 | 5361.516 |
| >tr_A0A072UZJ3_A0A072UZJ3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g063090 PE_4 SV_1 ; MatureChain: 22-68 | INSDGYLECTTDYDCREEWLC PPDMEAKCFVSFALARFLSKG KCLCV | 5368.417 | 6 | 6 | 5362.374 | 5363.381 |
| >tr_A0A072UT45_A0A072UT45_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g006380 PE_4 SV_1 ; MatureChain: 24-69 | TQRIYQCIKESDCPQYMCSAG LRANCVDRGVCKCVPVWWR KFHVLT | 5389.583 | 6 | 6 | 5383.539 | 5384.547 |
| >tr_A0A072UXR0_A0A072UXR0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g065675 PE_4 SV_1 ; MatureChain: 24-69 | GGKYENIPCESREQCPNTATR RYACLNKLCYCYDNNYPNG WNPFEP | 5390.352 | 5 | 4 | 5386.323 | 5387.330 |
| >tr_G7JX80_G7JX80_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g057460 PE_4 SV_1 ; MatureChain: 21-64 | QKRRRSTECRNDSDCEKMVK CVLPRIARCIKYRCQCRNFLES FE | 5396.628 | 6 | 6 | 5390.585 | 5391.592 |
| >tr_G7KME0_G7KME0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g057520 PE_4 SV_1 ; MatureChain: 23-68 | WFKRTETGEIIWVVRCVTDTD CPKMGEPQYFKCLNGVCLEHI RELP | 5438.653 | 4 | 4 | 5434.624 | 5435.631 |
| >tr_G7IWA4_G7IWA4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g021040 PE_4 SV_1 ; MatureChain: 24-70 | LNGHGRNRCFRDSDCPKVMC PSYLVTKCFKKHCRCRKPGLQ VQLNPK | 5459.727 | 6 | 6 | 5453.684 | 5454.691 |
| >tr_A0A072VH59_A0A072VH59_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g037670 PE_4 SV_1 ; MatureChain: 27-73 | RRVSYNSFFRALPVCQTAADC PEGTRGRTYKCINNKCRYPKL LKPIQ | 5460.805 | 4 | 4 | 5456.776 | 5457.783 |
| >tr_G7L378_G7L378_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11443881 PE_4 SV_2 ; MatureChain: 25-71 | ISIYVRCASTNECYTTFKFAPL GSMRCVEGYCKHLKDFKVTQ FLYCV | 5461.629 | 5 | 4 | 5457.600 | 5458.607 |
| >tr_G7IX56_G7IX56_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g033915 PE_4 SV_1 ; MatureChain: 24-71 | AREKVNCLDDADCLEVSCLN GSNAECVGNSCVCVFVFYRE NFDEQFRR | 5465.408 | 6 | 6 | 5459.365 | 5460.372 |
| >tr_A7KHF8_A7KHF8_MEDTR Nodule-specific cysteine-rich peptide 328 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 31-78 | YREPFSSFTEGPTCKEDIDCPSI SCVNPQVPKCIMFECHCKYIP TTLK | 5483.517 | 6 | 6 | 5477.474 | 5478.481 |
| >tr_G7KV27_G7KV27_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g071310 PE_4 SV_1 ; MatureChain: 25-72 | YREPFSSFTEGPTCKEDIDCPSI SCVNPQVPKCIMFECHCKYIP TTLK | 5483.517 | 6 | 6 | 5477.474 | 5478.481 |
| >tr_A0A072TTX6_A0A072TTX6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0054s0190 PE_4 SV_1 ; MatureChain: 23-69 | VFKRTETGEIIWTLPCATDTDC PKMGEPMYFKCLNGFCLEHIR ELHD | 5486.572 | 4 | 4 | 5482.543 | 5483.550 |
| >tr_G7IX41_G7IX41_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide-like protein OS_Medicago truncatula OX_3880 GN_MTR_3g033700 PE_4 SV_2 ; MatureChain: 33-83 | KEIDAECDITSDCHDCAKGFH KLCSFGHCYCITGPPAKPEQNI LDVKDGGI | 5520.501 | 6 | 6 | 5514.457 | 5515.465 |
| >tr_A7KHG9_A7KHG9_MEDTR Nodule-specific cysteine-rich peptide 339 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 30-76 | EIDADCPQICMPPYEVRCVNH RCGWVNTDDSLFLTQEFTRSK QYIIS | 5519.550 | 4 | 4 | 5515.521 | 5516.528 |
| >tr_A0A072U9R9_A0A072U9R9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g055763 PE_4 SV_1 ; MatureChain: 27-76 | GPIRICSSDDDCIGYWCPLSIQP RSTKPICRLVESISKRSRTPVG LCTCI | 5521.758 | 6 | 6 | 5515.714 | 5516.722 |
| >tr_G7IW95_G7IW95_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g020950 PE_4 SV_1 ; MatureChain: 22-70 | QNIDSGGNRRCFRDSDCPKN MCPSYLVVKCLRSNCKCVRP GLQVRLNPN | 5552.682 | 6 | 6 | 5546.638 | 5547.646 |
| >tr_A0A072U9U0_A0A072U9U0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g056047 PE_4 SV_1 ; MatureChain: 26-72 | SNTLLAFRECVYDKDCPVMP RCNMRCRKGVCIPVRRKEYF KMNVSLL | 5580.786 | 5 | 4 | 5576.757 | 5577.764 |
| >tr_G7K5E0_G7K5E0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g026080 PE_4 SV_1 ; MatureChain: 23-67 | NFEDISIECMFSIDCPQIKSNIF RFKCIEDRCKIEFIYQRKKYEI | 5581.736 | 4 | 4 | 5577.707 | 5578.715 |
| >tr_G7KEA3_G7KEA3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g072450 PE_4 SV_1 ; MatureChain: 27-74 | LYIGCETDRDYPPLANKTFYL KCIDKKCEWTVTDSLSTRSGR MQKLSI | 5585.781 | 3 | 2 | 5583.767 | 5584.774 |
| >tr_A0A072TZE3_A0A072TZE3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g055783 PE_4 SV_1 ; MatureChain: 27-76 | RTIRICSHDYDCIGYFCPYSIQP RSTKPLCRLVGGIWKPSGEPV GLCTCI | 5600.747 | 6 | 6 | 5594.703 | 5595.710 |
| >tr_A0A072UMU2_A0A072UMU2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g088495 PE_4 SV_1 ; MatureChain: 25-78 | KNVVKNPGPCFVMGACSVPD DSKCHQFCLKSKGIDAGKCA KNNADSTCCCNNIG | 5618.524 | 8 | 6 | 5612.480 | 5613.488 |
| >tr_A0A072VCR5_A0A072VCR5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g011415 PE_4 SV_1 ; MatureChain: 26-71 | FIDCETVADCPTHWAYIYVCE KNKCRYHFKSGRVRPDHQKN RHNRV | 5618.698 | 4 | 4 | 5614.669 | 5615.676 |
| >tr_A7KH91_A7KH91_MEDTR Nodule-specific cysteine-rich peptide 118 OS_Medicago truncatula OX_3880 GN_MtrunA17_Chr4g0013951 PE_2 SV_1 ; MatureChain: 28-76 | RIMVVNPNNPCVTDADCQRY RHKLATRMVCNIGFCLMDFT HDPYAPSLP | 5621.653 | 4 | 4 | 5617.624 | 5618.631 |
| >tr_G7JM04_G7JM04_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g031525 PE_4 SV_1 ; MatureChain: 22-70 | RIMVVNPNNPCVTDADCQRY RHKLATRMVCNIGFCLMDFT HDPYAPSLP | 5621.653 | 4 | 4 | 5617.624 | 5618.631 |
| >tr_G7JM03_G7JM03_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g031520 PE_4 SV_1 ; MatureChain: 22-68 | RLVFVNPEKPCVTDADCDRY RHESAIYSDMFCKDGYCFIDY HHDPYP | 5629.472 | 4 | 4 | 5625.443 | 5626.450 |
| >tr_G7K8Y9_G7K8Y9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g059445 PE_4 SV_1 ; MatureChain: 24-71 | AQNLMKCNTDDECPKFDDKF PLSFKCINDGCRMVINDKYKH KTVQKLL | 5633.742 | 4 | 4 | 5629.713 | 5630.720 |
| >tr_A0A072UTM9_A0A072UTM9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g012380 PE_4 SV_1 ; MatureChain: 22-71 | MNAKRNYQCDPCFGHPDDMI NFCPPGTAPKCFHGLIKCVPIM RGTNRMFA | 5641.586 | 5 | 4 | 5637.557 | 5638.564 |
| >tr_G7J0H4_G7J0H4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11420593 PE_4 SV_1 ; MatureChain: 22-71 | KNIDGRVSYNSFIALPVCQTA ADCPEGTRGRTYKCINNKCRY PKLLKPIQ | 5641.889 | 4 | 4 | 5637.860 | 5638.867 |
| >tr_G7JLZ7_G7JLZ7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g031430 PE_4 SV_1 ; MatureChain: 22-70 | RIMVVNPNNPCVTDADCQRY RHKLATRMICNQGFCLMDFT HDPYAPSLP | 5650.643 | 4 | 4 | 5646.614 | 5647.621 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_A0A072TFS3_A0A072TFS3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0337s0030 PE_4 SV_1 ; MatureChain: 23-72 | AYERTEPIMHNGEPINLIPCVT VADCPRMDEPLHMTCLVGAC WPCIRSLY | 5656.638 | 5 | 4 | 5652.609 | 5653.616 |
| >tr_A0A072UE15_A0A072UE15_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide-like protein OS_Medicago truncatula OX_3880 GN_MTR_5g029495 PE_4 SV_1 ; MatureChain: 27-77 | RTPVKCSENIDCKPFCNQAFLS CCAFGQCICEACPPPLNHVPII QSNYRSK | 5671.624 | 8 | 6 | 5665.580 | 5666.588 |
| >tr_G7IZI9_G7IZI9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g034220 PE_4 SV_1 ; MatureChain: 29-78 | IHCRTNADCPRNMCKIGLPEC DQTRKECWCFLPPSVDNNNIP NVIPQVTN | 5695.649 | 6 | 6 | 5689.605 | 5690.613 |
| >tr_G7KA17_G7KA17_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g061800 PE_4 SV_2 ; MatureChain: 24-72 | AELGGPCRSDEECPQLSLRFF AIKCRENVCIYVDLDPYKPRA EKNQFLH | 5696.778 | 4 | 4 | 5692.749 | 5693.757 |
| >tr_G7K5D9_G7K5D9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g026070 PE_4 SV_1 ; MatureChain: 27-73 | EDYYYIECQRDFDCPQLNSEIF AFKCIEKLCKLEFIYQQAPFLL GQV | 5699.716 | 4 | 4 | 5695.687 | 5696.694 |
| >tr_A7KHG4_A7KHG4_MEDTR Nodule-specific cysteine-rich peptide 334 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 18-65 | VSYFSYFSTYIIECKTDNDCPIS QLKIYAWKCVKNGCHLFDVIP MMYE | 5704.659 | 4 | 4 | 5700.630 | 5701.638 |
| >tr_A0A072UA55_A0A072UA55_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g463780 PE_4 SV_1 ; MatureChain: 23-71 | WFKRTETGDNVQPSHFIPCFT AADCPMIDEPHYIECVTGFCW ALMRNLH | 5725.596 | 4 | 4 | 5721.566 | 5722.574 |
| >tr_G7IW92_G7IW92_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g020920 PE_4 SV_1 ; MatureChain: 22-72 | QNIDGNTGGNRKCFRDSDCPK FMCPSYLAVKCIGRLCRCGRP ELQVELNPK | 5728.754 | 6 | 6 | 5722.711 | 5723.718 |
| >tr_A7KHB1_A7KHB1_MEDTR Nodule-specific cysteine-rich peptide 175 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 27-77 | HIQCVIDDDCPKSLNKLLIIKCI NHVCQYVGNLPDFASQIPKST KMPYKGE | 5755.906 | 4 | 4 | 5751.876 | 5752.884 |
| >tr_A0A072TR88_A0A072TR88_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0337s0090 PE_4 SV_1 ; MatureChain: 23-72 | ALERTETTMHNVQPSHFIPCFT AADCPMIDEPHYIECVTGFCW ALMRNLH | 5774.615 | 4 | 4 | 5770.586 | 5771.593 |
| >tr_G7K8E1_G7K8E1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g037780 PE_4 SV_1 ; MatureChain: 28-80 | LKCGPLEILRCTLKGCYCEGL GEVLDCIPPDIMKCTSNGCNC YSEGWSYKHPS | 5825.632 | 8 | 6 | 5819.588 | 5820.595 |
| >tr_G7K5W8_G7K5W8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g076055 PE_4 SV_1 ; MatureChain: 23-71 | EDIGGNCECIRDEDCFKQKRD EDCHKEYCMIFYVHKCENYK CVCAGMFN | 5860.423 | 8 | 6 | 5854.379 | 5855.386 |
| >tr_A7KHG0_A7KHG0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11423521 PE_2 SV_1 ; MatureChain: 27-78 | AHFPCVTDDDCPKPVNKLRVI KCIDHICQYARNLPDFASEISE STKMPCKGE | 5860.796 | 5 | 4 | 5856.767 | 5857.774 |
| >tr_A0A072UJJ3_A0A072UJJ3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g052650 PE_4 SV_1 ; MatureChain: 24-75 | RTSVPCNSDSDCPETTLRKLW KCVDGFCDVTVKEITKSCFICP KDVISMIRD | 5880.814 | 6 | 6 | 5874.771 | 5875.778 |
| >tr_G7K1M3_G7K1M3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g056890 PE_2 SV_1 ; MatureChain: 23-71 | EVDWIYHLCDTDTDCPEHWS KFFIYKCVNHVCDSISKVTTDS KEYKNFP | 5884.662 | 4 | 4 | 5880.633 | 5881.640 |
| >tr_G7K4P4_G7K4P4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11435503 PE_4 SV_1 ; MatureChain: 27-78 | EHIQCVIDDDCPKSLNKLLIIK CINHVCQYVGNLPDFASQIPK STKMPYKGE | 5884.948 | 4 | 4 | 5880.919 | 5881.926 |
| >tr_G7K8T1_G7K8T1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g058510 PE_4 SV_1 ; MatureChain: 27-77 | HIECKNDFDCPKNMCLAPRV AWCVNNKCECVLTYGPKYST MKEKLLQKEKI | 5960.886 | 6 | 6 | 5954.842 | 5955.849 |
| >tr_A0A072V580_A0A072V580_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide-like protein OS_Medicago truncatula OX_3880 GN_MTR_3g435860 PE_4 SV_1 ; MatureChain: 31-82 | GREYCNKDSDCHREKNPDCN LYIACVAHECQCVHLQSALTE EIPIRAPLFKH | 5995.789 | 6 | 6 | 5989.745 | 5990.753 |
| >tr_I3S866_I3S866_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g450690 PE_2 SV_1 ; MatureChain: 27-78 | KLNCIDDSDCPYDMCDPGLLP RCLNGWCDCSRFQPWPMDS MSSNLREFTLPN | 5997.561 | 6 | 6 | 5991.517 | 5992.525 |
| >tr_A0A072VI94_A0A072VI94_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_1g052005 PE_4 SV_1 ; MatureChain: 24-74 | KDITCTVAGDCPNFFVCPPNN FVRCIRNLCKCRFVYLNTFLK VYTDKDFIF | 5999.930 | 6 | 6 | 5993.887 | 5994.894 |
| >tr_G7JBC9_G7JBC9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g084820 PE_2 SV_1 ; MatureChain: 24-75 | EEIIIIKCQTAKDCPDIYNLFPL VYKCIDNICVDVRLEPPYDMS ISPKSVHK | 6008.030 | 4 | 4 | 6004.001 | 6005.009 |
| >tr_A0A072TJN4_A0A072TJN4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0003s0130 PE_4 SV_1 ; MatureChain: 24-76 | DDSKPFSSLFKSPLVYCLQHS DCQAYECELPFKPDCLMVEYS PQFVALRCGCV | 6034.755 | 6 | 6 | 6028.712 | 6029.719 |
| >tr_G7K748_G7K748_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g072070 PE_4 SV_2 ; MatureChain: 27-76 | DIYCETDADCPQITDWFYVVK CVDHKCELTKKLRRLYEYQT QKSAETPYI | 6047.888 | 4 | 4 | 6043.859 | 6044.866 |
| >tr_A0A072UTT4_A0A072UTT4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_25488623 PE_4 SV_1 ; MatureChain: 25-75 | YILCKTVNDCPPNTRNLRYRC IDGKCKSHRVLYEWDESHTQ DITITPCIEE | 6067.889 | 5 | 4 | 6063.860 | 6064.868 |
| >tr_A0A072UBT5_A0A072UBT5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g084970 PE_4 SV_1 ; MatureChain: 23-75 | AEITNIPCVSKDDCPKVVKPL VIKCIDHFCEYFWLNMMGHN HVIVTTRSKSNN | 6085.995 | 4 | 4 | 6081.966 | 6082.974 |
| >tr_A0A072UU20_A0A072UU20_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide-like protein OS_Medicago truncatula OX_3880 GN_MTR_3g435850 PE_4 SV_1 ; MatureChain: 21-72 | RREYCNKDSDCHKEKIPECH MYIACVAHECQCVHLQPAFT EEIPIRAPSIKH | 6104.860 | 6 | 6 | 6098.817 | 6099.824 |
| >tr_A0A072UC74_A0A072UC74_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g084960 PE_4 SV_1 ; MatureChain: 24-75 | EITNIPCVSKDDCPKVITPLVIK CIDHFCEYFWLNMRDHNHVIE TMRSKSNN | 6144.960 | 4 | 4 | 6140.931 | 6141.938 |
| >tr_A0A072UKF8_A0A072UKF8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g065730 PE_4 SV_1 ; MatureChain: 26-78 | KECVTDADCENLYYGNKWPL ICSNIGYCLSSYEEDACRKHLH PFENEKPDMGK | 6154.740 | 5 | 4 | 6150.711 | 6151.718 |
| >tr_G7L5G4_G7L5G4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g032820 PE_4 SV_2 ; MatureChain: 22-76 | NHAISGLLPKLPFGCCTSNLDC PRHMCTHPQQPWCIFYGNRIM YRGSRLGICKCS | 6221.930 | 7 | 6 | 6215.887 | 6216.894 |
| >tr_A0A072UIG7_A0A072UIG7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g036950 PE_4 SV_1 ; MatureChain: 27-78 | RNIVDSSELYAFSDEKGLNKR LLQENRRAPCKRGKRDQMRQ ARDNPRCHNYL | 6244.167 | 2 | 2 | 6242.153 | 6243.160 |

| | | | | | | |
|---|---|---|---|---|---|---|
| >tr_G7K4P1_G7K4P1_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g095590 PE_4 SV_2 ; MatureChain: 27-79 | EVFQRVTDDGCPKPVNHLRV VKCIEHICEYGYNYRPDFASQI PESTKMPRKRE | 6277.086 | 3 | 2 | 6275.072 | 6276.079 |
| >tr_A0A072UA67_A0A072UA67_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g038620 PE_4 SV_1 ; MatureChain: 23-76 | DSQRTCSNFLECYKIYGIPLDG VWRCVKGFCELLIDFNTYKV REVAIVRGENIN | 6286.126 | 4 | 4 | 6282.097 | 6283.104 |
| >tr_A7KH70_A7KH70_MEDTR Nodule-specific cysteine-rich peptide 28 OS_Medicago truncatula OX_3880 GN_MtrunA17_Chr7g0222431 PE_2 SV_1 ; MatureChain: 22-77 | SYTDECSTDADCEYILCLFPIIK RCIHNHCKCVPMGSIEPMSTIP NGVHKFHIINN | 6348.951 | 6 | 6 | 6342.908 | 6343.915 |
| >tr_G7JAL7_G7JAL7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g069870 PE_4 SV_1 ; MatureChain: 22-77 | TKILEKHTNKCAATVGLDIYE KDKCVTDFDCVKNLWLCPID QFVRCIDETCKCILF | 6481.182 | 7 | 6 | 6475.138 | 6476.146 |
| >tr_G7K955_G7K955_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g061275 PE_4 SV_1 ; MatureChain: 24-79 | LYVCRKVAECPENFCVPPLIT KCVNYTCICDDPAYGEPIYDF VSVRTEKQKIKIKR | 6497.258 | 6 | 6 | 6491.214 | 6492.221 |
| >tr_A0A072V8H8_A0A072V8H8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g044310 PE_4 SV_1 ; MatureChain: 27-81 | YDDCYNHAECKTKIKCVLPRI AECVRFKCDCVRLNVPRTPW STRPENAHKPYKEE | 6551.179 | 6 | 6 | 6545.135 | 6546.142 |
| >tr_A7KHG3_A7KHG3_MEDTR Nodule-specific cysteine-rich peptide 333 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 24-79 | SERECVTDDDCEKLYPTNEYR MMCDSGYCMNLLNGKIIYLL CLKKKKFLIIISVLL | 6580.279 | 5 | 4 | 6576.250 | 6577.257 |
| >tr_G7K1L8_G7K1L8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g056760 PE_4 SV_1 ; MatureChain: 24-82 | AYYECSNDSACQATTKCVLP RVPRCIKYKCLCGNSNGSGNR WSTRPNRIQKGSTESNYF | 6624.082 | 6 | 6 | 6618.038 | 6619.046 |
| >tr_A0A072VBA6_A0A072VBA6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g079180 PE_4 SV_1 ; MatureChain: 24-82 | AYFECHSDSACETTVKCVLPR IPRCIKYKCLCGNGVGKRWST TPKRIEKGSTVRNGFLH | 6642.326 | 6 | 6 | 6636.282 | 6637.290 |
| >tr_A0A072V8Q0_A0A072V8Q0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g063420 PE_4 SV_1 ; MatureChain: 27-83 | RVKCIKDYGCQNKSCQFPLKP YIRTGRVTCKNDSGCELKLYQ SQRKYVEASIVNWLY | 6719.384 | 5 | 4 | 6715.355 | 6716.362 |
| >tr_A0A072UZ22_A0A072UZ22_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g071330 PE_4 SV_1 ; MatureChain: 23-80 | HYSYIILFSLFITTDDSCYNDY HCINKSWLCPSGLVVRCITRQ CKCITILNPIDFVST | 6734.264 | 6 | 6 | 6728.220 | 6729.227 |
| >tr_G7IZF0_G7IZF0_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g028440 PE_2 SV_1 ; MatureChain: 23-79 | TNFERKQISFSFFMKEYWPCV TDDDCPSDLCKKVDQIPKCVG GLCKCFPIRFGQWER | 6767.185 | 6 | 6 | 6761.141 | 6762.148 |
| >tr_A0A072VHP7_A0A072VHP7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g044330 PE_4 SV_1 ; MatureChain: 27-85 | YDDCYNHTECTNKIKCVPPRI AQCFRFKCDCIRLNNGPKTPW SATPKRVHISPTRKNDF | 6924.365 | 6 | 6 | 6918.321 | 6919.329 |
| >tr_A0A072UJ01_A0A072UJ01_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_4g046850 PE_4 SV_1 ; MatureChain: 27-87 | KFTRCFRDSDCPKTLCHSPGK AKYFCSLLSLKSNKDMGRSTN LFKNAPSIKPENVKYFKNP | 6984.527 | 4 | 4 | 6980.498 | 6981.505 |
| >tr_G7JAL3_G7JAL3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g069830 PE_2 SV_1 ; MatureChain: 23-83 | KILEKHKCVTDGVEILEKGKC FTDWECVRNSWLCPVDLVVR CIKETCKCIKILEPINVVPT | 7012.640 | 7 | 6 | 7006.597 | 7007.604 |
| >tr_G7K1L2_G7K1L2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g056710 PE_4 SV_1 ; MatureChain: 24-84 | AHVECHNDSACEKTVKCMLP RIPRCIKYQCLCGYSDDPGNR WSTRPKRIQKGSTERKGFLY | 7071.433 | 6 | 6 | 7065.389 | 7066.396 |
| >tr_G7IX29_G7IX29_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide-like protein OS_Medicago truncatula OX_3880 GN_MTR_3g033510 PE_4 SV_2 ; MatureChain: 25-89 | SDMCMKLEGKGVEGTHPCKY DSDCQPGCPPNTHGCCIHGRC WCFNSPIFADKIPGLIPQGNCFI R | 7110.146 | 9 | 6 | 7104.102 | 7105.109 |
| >tr_A0A072UJS9_A0A072UJS9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g445040 PE_4 SV_1 ; MatureChain: 26-88 | RITHETLRNMPLPGSKPIPILRG ECISDAECKHPECDNCRGICL NSRCVCMMRLGWTYTTPQN | 7156.390 | 7 | 6 | 7150.347 | 7151.354 |
| >tr_A0A072TGW4_A0A072TGW4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_0192s0050 PE_4 SV_1 ; MatureChain: 24-83 | ESNWSQRNDITKIHCIKNADC PKHMCIRPQKPKCFDSWYMS QNQNQGNVRVMRRTKKNRL | 7258.562 | 4 | 4 | 7254.533 | 7255.540 |
| >tr_A7KHD9_A7KHD9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g069500 PE_2 SV_1 ; MatureChain: 24-85 | DDNKLLLSFIEEGFLCFKDSDC PYNMCPSPLKEMCYFIKCVCG VYGPIRERRLYQSHNPMIQ | 7295.414 | 6 | 6 | 7289.370 | 7290.377 |
| >tr_A0A072UBI5_A0A072UBI5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g075430 PE_4 SV_1 ; MatureChain: 21-85 | KNVVIETRASEVLSDGVCMSL SGTFNGLCIPPFMNNRCDKSC KNKEHKYYGKCWQDLRCWC YGEC | 7455.390 | 8 | 6 | 7449.346 | 7450.353 |
| >tr_G7KA19_G7KA19_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_5g061825 PE_4 SV_2 ; MatureChain: 25-89 | AENIECEVDADCPKSQVNSFV IKCIKNLCLYTKIHILYDTISKS ESTLPQKKKKSNSTLTYFKEE | 7454.780 | 4 | 4 | 7450.751 | 7451.758 |
| >tr_A0A072TYB9_A0A072TYB9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g446540 PE_4 SV_1 ; MatureChain: 24-91 | ATETFQEPQGMTNCGETTCA KGDGVCYKSCIEEGFNRGGD CLFHNNHEDKICCCYKNESSL SPNYYLN | 7565.163 | 8 | 6 | 7559.120 | 7560.127 |
| >tr_A0A072U8H9_A0A072U8H9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_6g038320 PE_4 SV_1 ; MatureChain: 31-98 | KDKFPGNKYPIKCINGIFYDAT NALTNNVKDPRELNSVGRDIA YYMQGRSSNPGHSTSPQLNC MSSNH | 7572.624 | 2 | 2 | 7570.610 | 7571.617 |
| >tr_A0A072UYQ3_A0A072UYQ3_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g063420 PE_4 SV_1 ; MatureChain: 27-93 | RVKCIKDYGCQNKSCQFPLKP YIRTGRVTCKNDSGCELKLCK LPLFPKCVKPYFLFFSTKEGFC ACN | 7702.835 | 9 | 6 | 7696.791 | 7697.798 |
| >tr_G7KV22_G7KV22_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g071220 PE_4 SV_1 ; MatureChain: 23-90 | EESHYMKFSICKDDTDCPTLF CVLPNVPKCIGSKCHCKLMEH KHIQDKSCFICLLLQYMENES VIADI | 7862.674 | 8 | 6 | 7856.631 | 7857.638 |
| >tr_G7KZ18_G7KZ18_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11431692 PE_2 SV_1 ; MatureChain: 29-104 | MSLCNMNEDGLDACKPSVTQ PYPAKPSTECCKALTGADLQC LCSYKNSAELPLLGIDPTLAAS LPKECDLTPPSNC | 7986.699 | 8 | 6 | 7980.655 | 7981.662 |
| >tr_A0A072V8P6_A0A072V8P6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g063370 PE_4 SV_1 ; MatureChain: 27-94 | RVKYKEDYGGQNKWNEFPLK PHIRTYRIKCKDDSGCEGNNL | 8031.954 | 6 | 6 | 8025.911 | 8026.918 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | CEGRWIPKCLKPYFLFLTTKE GFCACV | | | | | |
| >tr_A0A072ULG5_A0A072ULG5_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_25492587 PE_4 SV_1 ; MatureChain: 16-86 | KECVTDADCDKLYPDIRKPL MCSIGECYSLYKGKFSLSIISK TSFSLMVYNVVTLVICLRLAY ISLLLKFL | 8073.239 | 5 | 4 | 8069.210 | 8070.217 |
| >tr_A7KHG2_A7KHG2_MEDTR Nodule-specific cysteine-rich peptide 332 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 25-95 | EECVTDADCDKLYPDIRKPLM CSIGECYSLYKGKFSLSIISKTS FSLMVYNVVTLVICLRLAYIS LLLKFL | 8074.186 | 5 | 4 | 8070.157 | 8071.165 |
| >tr_A0A072U0B7_A0A072U0B7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g066220 PE_4 SV_1 ; MatureChain: 26-95 | VPPLTIQCINYTCICDDPPYGE PEYDNNDDFVTLNREKAKIKN EEMMMRERDMMIEIETYSVA DDLDPHL | 8237.746 | 3 | 2 | 8235.732 | 8236.739 |
| >tr_A0A072TVW8_A0A072TVW8_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g007970 PE_4 SV_1 ; MatureChain: 25-96 | DVKCDTDDDCRDYLCARPTV GKCIYDYCHCIVMITIDEKLSH QSGINKVVRENGHVSIDPTIKE IKLRENIL | 8248.044 | 6 | 6 | 8242.000 | 8243.008 |
| >tr_G7JDN2_G7JDN2_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11412955 PE_4 SV_1 ; MatureChain: 16-87 | KECVTDADCENLYPGNKKPM FCNNTGYCMSLYKGKIYLSSK SNNFIFFSCYNVVTLIICLKLTY ISPNLKYL | 8280.046 | 6 | 6 | 8274.002 | 8275.010 |
| >tr_G7IW93_G7IW93_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_11409031 PE_2 SV_1 ; MatureChain: 22-94 | KNIDGRNNPTRRNWSELVGV TAEEAERKIKEEMNGVEIRVV PPGYFVTADYNTQRVRLYVD QSNKLIKTPTIG | 8371.367 | 0 | 0 | 8371.367 | 8372.374 |
| >tr_A7KHG1_A7KHG1_MEDTR Nodule-specific cysteine-rich peptide 331 OS_Medicago truncatula OX_3880 PE_2 SV_1 ; MatureChain: 27-100 | ETLSLTHPKCHHIMLPSLFITE VFQRVTDDGCPKPVNHLRVV KCIEHICEYGYNYRPDFASQIP ESTKMPRKRE | 8676.336 | 4 | 4 | 8672.307 | 8673.314 |
| >tr_G7L3W7_G7L3W7_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g102170 PE_4 SV_2 ; MatureChain: 25-100 | GKSLFSDISNFHIYHIIIIHNHSF LLVTFFYFYFNSDPMYCFNDD DCRELKCSHPRVRKCRMFLC RCEEVDKEDEK | 9302.360 | 6 | 6 | 9296.316 | 9297.324 |
| >tr_A0A072V8J4_A0A072V8J4_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g467490 PE_4 SV_1 ; MatureChain: 19-100 | NPNFNPGDPFLDVLNFLNYFI HNIFSHLNNTLSFSFFITTGRK CKQNSDCSKEICVFPWKPTCV EPYFLMILIKRYNYCTCT | 9679.675 | 6 | 6 | 9673.631 | 9674.639 |
| >tr_G7IQF6_G7IQF6_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_2g045290 PE_4 SV_2 ; MatureChain: 23-116 | LKVIIPSSTCDSDYDCLRYEEA LNVITCCNNGLCVMFWRDMV RSRFDGEVEAAIGVGESRDGT LGGGVEILTHIWEILMKSTTR VKALVKSTKK | 10438.231 | 5 | 4 | 10434.202 | 10435.209 |
| >tr_A0A072TXA9_A0A072TXA9_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_25497868 PE_4 SV_1 ; MatureChain: 30-124 | HHCSIGPCSTPWATCGSEYCIC IPMGSSNICQPSSYKDVVKITG KNHNFCQSHVECKEKGRGSF CARYPSSKVDYGRCVASISEE EDFLRMSVIV | 10465.781 | 10 | 6 | 10459.737 | 10460.745 |
| >tr_A0A072TX06_A0A072TX06_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_7g015700 PE_4 SV_1 ; MatureChain: 24-120 | ALDSKLNNNNTQVQRCIGFHC LVTGDDETDLFMDQRGPSRM LATNDNQVTPQTENSGQQSPN CQPNSLQGSSSCLASPQLNQG RPCEPLNRAYPYCK | 10652.890 | 6 | 6 | 10646.846 | 10647.853 |
| >tr_I3SB50_I3SB50_MEDTR Nodule Cysteine-Rich _NCR_ secreted peptide OS_Medicago truncatula OX_3880 GN_MTR_3g016105 PE_2 SV_1 ; MatureChain: 20-122 | LKGLSALKENVDGPECRTHS MSLILKIVQLHHALKNEIFHW KKQICNIHFAKLFYSYQSLYFI YFFNFTDLAVRGVHGLGKPK KPVKPTQKSGLGWVIERIWF | 12037.471 | 2 | 2 | 12035.456 | 12036.464 |

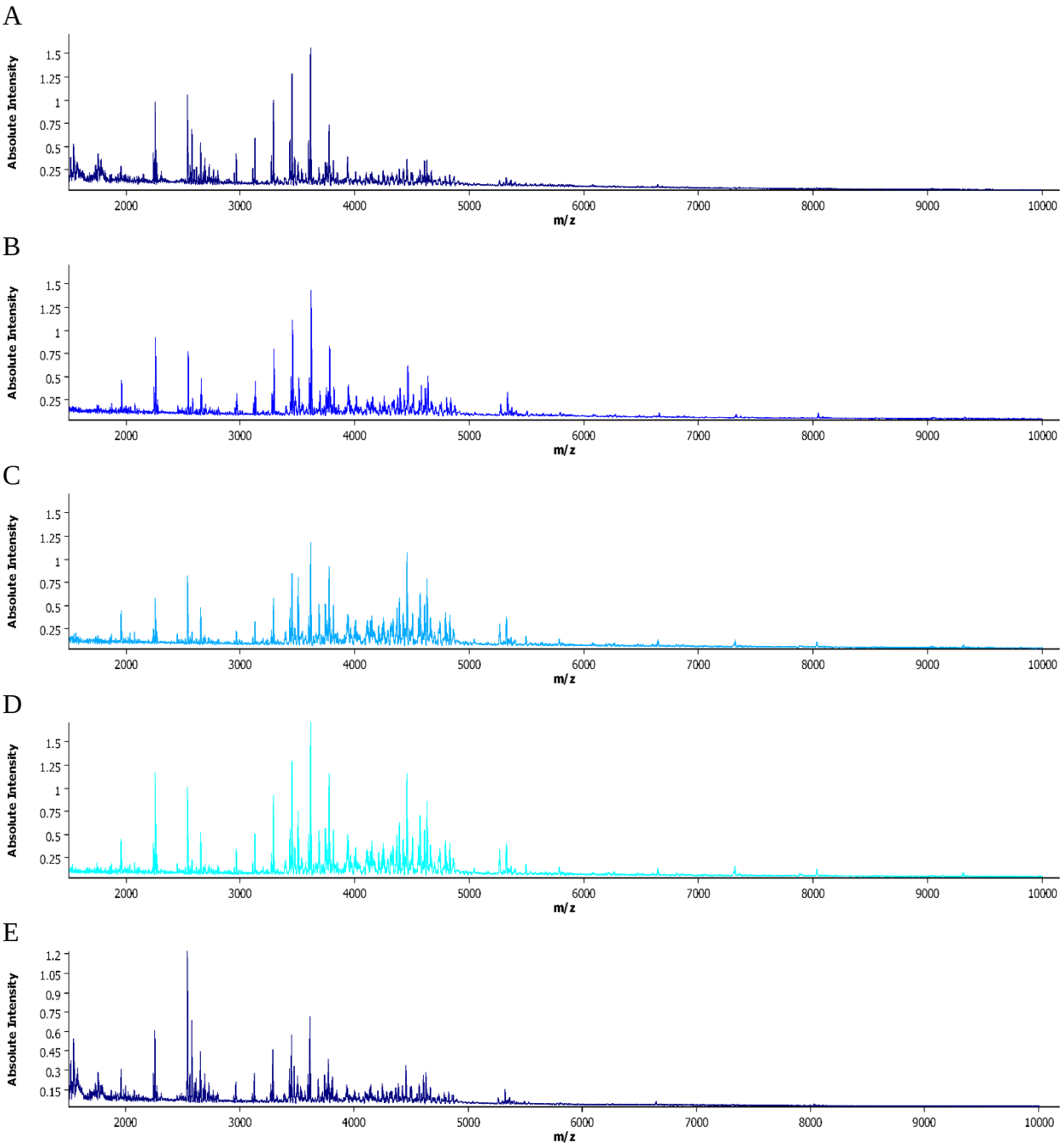
**Supporting Table S4**. LC-MS/MS identifications from wild type root nodules.

| Protein Group | Protein ID | Protein Accession | Peptide | Unique | -10lgP | Mass | Length | ppm | m/z | z | RT | Start | End | PTM | Uniprot Predicted |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 41 | 2 | A0A072TFS3\|A0A072TFS3_MEDTR | A.AYERTEPIMHNGEPINL.I | Y | 35.49 | 1982.952 | 17 | 3.5 | 661.9936 | 3 | 40.56 | 23 | 39 | | 32-72 |
| 41 | 2 | | L.IPC(+57.02)VTVADC(+57.02)PR.M | Y | 33.69 | 1286.611 | 11 | -1 | 644.3122 | 2 | 35.42 | 40 | 50 | Carbamidomethylation | |
| 41 | 2 | | L.IPC(+57.02)VTVADC(+57.02)PRMD.E | Y | 29.33 | 1532.679 | 13 | -4.2 | 767.3434 | 2 | 38.91 | 40 | 52 | Carbamidomethylation | |
| 41 | 2 | | L.IPC(+57.02)VTVADC(+57.02)PRMDEPLH.M | Y | 29.07 | 2008.917 | 17 | 0.7 | 670.6467 | 3 | 40.89 | 40 | 56 | Carbamidomethylation | |
| 41 | 2 | | L.VGAC(+57.02)WPC(+57.02)IRSLY | Y | 28.93 | 1480.696 | 12 | 1.6 | 741.3563 | 2 | 51.9 | 61 | 72 | Carbamidomethylation | |
| 41 | 2 | | G.AC(+57.02)WPC(+57.02)IRSLY | Y | 28.89 | 1324.606 | 10 | -1.6 | 663.309 | 2 | 49.77 | 63 | 72 | Carbamidomethylation | |
| 41 | 2 | | L.VGAC(+57.02)WPC(+57.02)IRS.L | Y | 28.54 | 1204.548 | 10 | 0.9 | 603.2819 | 2 | 39.6 | 61 | 70 | Carbamidomethylation | |
| 135 | 50 | A0A072U6V1\|A0A072U6V1_MEDTR | G.DSFRGC(+57.02)NKDTDC(+57.02)PEKFC(+57.02)SSPDVVR C(+57.02)YIEC(+57.02)YC(+57.02)I | Y | 38.1 | 4149.746 | 33 | -5.7 | 1038.438 | 4 | 50.04 | 25 | 57 | Carbamidomethylation | 25-57 |
| 135 | 50 | | G.DSFRGC(+57.02)NKDTDC(+57.02)PEKFC(+57.02)SSPDVVR C(+57.02)IYIE.C | Y | 33.68 | 3553.537 | 29 | -1.3 | 889.3904 | 4 | 42.95 | 25 | 53 | Carbamidomethylation | |
| 135 | 50 | | G.DSFRGC(+57.02)NKDTDC(+57.02)PEKFC(+57.02)SSPDVVR C(+57.02)IYIEC(+57.02).Y | Y | 26.48 | 3713.568 | 30 | -2.7 | 929.3967 | 4 | 42.81 | 25 | 54 | Carbamidomethylation | |
| 93 | 34 | A0A072UBM1\|A0A072UBM1_MEDTR | A.FEEC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 37.2 | 1794.665 | 14 | -2.5 | 898.3373 | 2 | 29.05 | 25 | 38 | Carbamidomethylation | |
| 93 | 34 | | E.EC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 32.75 | 1518.554 | 12 | -1.2 | 760.2832 | 2 | 24.28 | 27 | 38 | Carbamidomethylation | |
| 93 | 11 | G7KV52\|G7KV52_MEDTR | S.FEEC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 37.2 | 1794.665 | 14 | -2.5 | 898.3373 | 2 | 29.05 | 27 | 40 | Carbamidomethylation | |
| 93 | 11 | | E.EC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 32.75 | 1518.554 | 12 | -1.2 | 760.2832 | 2 | 24.28 | 29 | 40 | Carbamidomethylation | |
| 93 | 33 | *G7KV51\|G7KV51_MEDTR* | A.FEEC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 37.2 | 1794.665 | 14 | -2.5 | 898.3373 | 2 | 29.05 | 31 | 44 | Carbamidomethylation | |
| 93 | 33 | | E.EC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 32.75 | 1518.554 | 12 | -1.2 | 760.2832 | 2 | 24.28 | 33 | 44 | Carbamidomethylation | |
| 93 | 10 | G7KV53\|G7KV53_MEDTR | S.FEEC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 37.2 | 1794.665 | 14 | -2.5 | 898.3373 | 2 | 29.05 | 33 | 46 | Carbamidomethylation | |
| 93 | 10 | | E.EC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 32.75 | 1518.554 | 12 | -1.2 | 760.2832 | 2 | 24.28 | 35 | 46 | Carbamidomethylation | |
| 114 | 6 | G7JAL7\|G7JAL7_MEDTR | E.KDKC(+57.02)VTDFDC(+57.02)VK.N | Y | 37.11 | 1513.69 | 12 | -4.6 | 757.849 | 2 | 31.42 | 43 | 54 | Carbamidomethylation | |
| 114 | 6 | | E.KDKC(+57.02)VTDFDC(+57.02)VKN.L | Y | 36.37 | 1627.733 | 13 | 0.7 | 814.8746 | 2 | 30.76 | 43 | 55 | Carbamidomethylation | |
| 267 | 16 | A0A072TZN8\|A0A072TZN8_MEDTR | M.C(+57.02)PLETVVKC(+57.02)VEDEC(+57.02)IC(+57.02)VH | Y | 45.96 | 2245.984 | 18 | -0.6 | 749.6682 | 3 | 51.84 | 41 | 58 | Carbamidomethylation | |
| 267 | 16 | | L.ETVVKC(+57.02)VEDEC(+57.02)IC(+57.02)VH | Y | 27.9 | 1875.817 | 15 | -2.1 | 626.2781 | 3 | 37.71 | 44 | 58 | Carbamidomethylation | |
| 186 | 25 | A7KHC0\|A7KHC0_MEDTR | S.YGPC(+57.02)TTLQDC(+57.02)ETHN.W | Y | 34.01 | 1694.667 | 14 | -6.1 | 848.3353 | 2 | 32.87 | 27 | 40 | Carbamidomethylation | 25-58 |
| 186 | 25 | | Y.GPC(+57.02)TTLQDC(+57.02)ETHN.W | Y | 30.52 | 1531.603 | 13 | -2.8 | 766.8067 | 2 | 29.66 | 28 | 40 | Carbamidomethylation | |
| 186 | 25 | | Y.GPC(+57.02)TTLQDC(+57.02)ETH.N | Y | 29.78 | 1417.56 | 12 | -1.8 | 709.7861 | 2 | 29.78 | 28 | 39 | Carbamidomethylation | |
| 89 | 5 | A0A072U9K0\|A0A072U9K0_MEDTR | G.RITHDPSTRSTVS.G | Y | 28.71 | 1455.743 | 13 | 0.2 | 728.8789 | 2 | 23.65 | 26 | 38 | | 26-72 |
| 234 | 7 | A0A072TT06\|A0A072TT06_MEDTR | Y.YKVEC(+57.02)IDGFC(+57.02)YYPV | Y | 36.78 | 1811.79 | 14 | -0.8 | 906.9015 | 2 | 51.08 | 36 | 49 | Carbamidomethylation | 21-49 |
| 234 | 7 | | Y.KVEC(+57.02)IDGFC(+57.02)YYPV | Y | 30.21 | 1648.727 | 13 | -2.2 | 825.3687 | 2 | 48.76 | 37 | 49 | Carbamidomethylation | |
| 269 | 15 | A0A072THJ9\|A0A072THJ9_MEDTR | D.EDC(+57.02)PVGEKC(+57.02)IGGNC(+57.02)VE | Y | 34.86 | 1821.733 | 16 | -5 | 911.8693 | 2 | 32.95 | 33 | 48 | Carbamidomethylation | 23-48 |
| 269 | 15 | | D.EDC(+57.02)PVGEKC(+57.02)IGGNC(+57.02).V | Y | 32.59 | 1593.622 | 14 | -5.5 | 797.814 | 2 | 30.58 | 33 | 46 | Carbamidomethylation | |
| 499 | 24 | A7KHF0\|A7KHF0_MEDTR | G.SNDEC(+57.02)VYPDVFQC(+57.02)INNIC(+57.02)KC(+57.02)VSHHRT | Y | 48.17 | 3251.4 | 26 | 1.7 | 651.2885 | 5 | 55.23 | 32 | 57 | Carbamidomethylation | 28-57 |
| 271 | 22 | A0A072U8J4\|A0A072U8J4_MEDTR | A.GC(+57.02)ITDADC(+57.02)VIK.K | Y | 32.7 | 1250.564 | 11 | -0.9 | 626.2885 | 2 | 35.41 | 29 | 39 | Carbamidomethylation | 27-60 |
| 271 | 22 | | L.AGC(+57.02)ITDADC(+57.02)VIK.K | Y | 30.82 | 1321.601 | 12 | -5.7 | 661.8038 | 2 | 35.61 | 28 | 39 | Carbamidomethylation | |
| 104 | 4 | A0A072TTX6\|A0A072TTX6_MEDTR | T.LPC(+57.02)ATDTDC(+57.02)PK.M | Y | 26.88 | 1276.543 | 11 | -3.3 | 639.2766 | 2 | 29.06 | 36 | 46 | Carbamidomethylation | 23-69 |
| 502 | 59 | A0A072TIU0\|A0A072TIU0_MEDTR | K.IGC(+57.02)KTSEDC(+57.02)PVLGKC(+57.02)IEDFC(+57.02)QFKK | Y | 41.56 | 2982.338 | 24 | 0.8 | 597.4754 | 5 | 43.76 | 39 | 62 | Carbamidomethylation | 23-62 |
| 182 | 37 | A0A072U1G7\|A0A072U1G7_MEDTR | G.KC(+57.02)VSDAEC(+57.02)SGQYMC(+57.02)PTLTVIKC(+57.02)IKDEC(+57.02)VC(+57.02)IHYDHDKQ | Y | 40.8 | 4546.981 | 37 | -4.8 | 910.3992 | 5 | 49.09 | 27 | 63 | Carbamidomethylation | |
| 182 | 37 | | G.KC(+57.02)VSDAEC(+57.02)SGQYM(+15.99)C(+57.02)PTLTVIKC(+57.02)IKDEC(+57.02)VC(+57.02)IHYDHDKQ | Y | 33.63 | 4562.977 | 37 | 5.4 | 761.5074 | 6 | 46.1 | 27 | 63 | Carbamidomethylation; Oxidation (M) | |
| 519 | 14 | A0A072UKX9\|A0A072UKX9_MEDTR | T.IPC(+57.02)TSDANC(+57.02)PK.V | Y | 30.61 | 1261.543 | 11 | 1.4 | 631.7797 | 2 | 26.72 | 34 | 44 | Carbamidomethylation | 23-69 |
| 546 | 42 | A0A072UAN2\|A0A072UAN2_MEDTR | F.VTC(+57.02)FEGLC(+57.02)R.M | Y | 25.39 | 1140.506 | 9 | -1.9 | 571.259 | 2 | 35.86 | 44 | 52 | Carbamidomethylation | 21-60 |
| 548 | 44 | G7IY11\|G7IY11_MEDTR | R.IPRPLIDPLN.C | Y | 24.73 | 1146.676 | 10 | 0.9 | 574.3458 | 2 | 45.58 | 25 | 34 | | 23-64 |

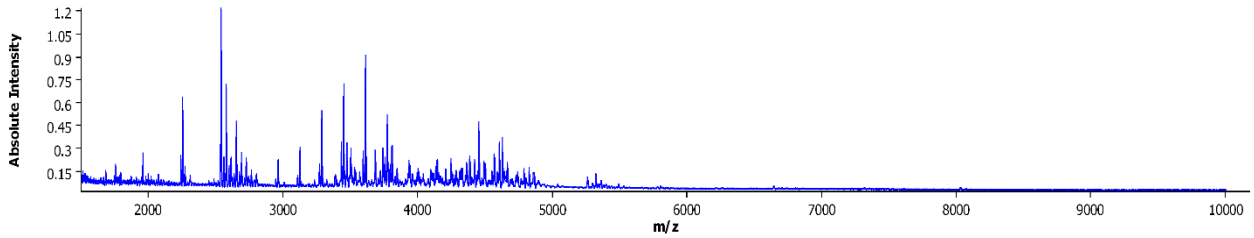
**Supporting Table S5**. LC-MS/MS Identifications from 35S:MtCLE13 root nodules.

| Protein Group | Protein ID | Protein Accession | Peptide | Unique | -10lgP | Mass | Length | ppm | m/z | z | RT | Start | End | PTM | Uniprot Predicted |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 43 | 2 | A0A072TFS3\|A0A072TFS3_MEDTR | L.IPC(+57.02)VTVADC(+57.02)PRMDEPLHMTC(+57.02)L<br>VGAC(+57.02)WPC(+57.02)IRSLY | Y | 36.82 | 3976.805 | 33 | 1.5 | 995.21 | 4 | 63.83 | 40 | 72 | Carbamidomethylation | 23-72 |
| 43 | 2 | | A.AYERTEPIMHNGEPINLI.I | Y | 35.78 | 1982.952 | 17 | 1.6 | 661.9923 | 3 | 41.08 | 23 | 39 | | |
| 43 | 2 | | L.IPC(+57.02)VTVADC(+57.02)PR.M | Y | 35.35 | 1286.611 | 11 | -1 | 644.3122 | 2 | 35.77 | 40 | 50 | Carbamidomethylation | |
| 43 | 2 | | L.IPC(+57.02)VTVADC(+57.02)PRMD.E | Y | 35.13 | 1532.679 | 13 | -1.4 | 767.3455 | 2 | 39.29 | 40 | 52 | Carbamidomethylation | |
| 43 | 2 | | L.IPC(+57.02)VTVADC(+57.02)PRMDEPLH.M | Y | 34.75 | 2008.917 | 17 | 0.6 | 670.6466 | 3 | 41.3 | 40 | 56 | Carbamidomethylation | |
| 43 | 2 | | L.VGAC(+57.02)WPC(+57.02)IRSL.L | Y | 31.42 | 1204.548 | 10 | -1.5 | 603.2804 | 2 | 39.99 | 61 | 70 | Carbamidomethylation | |
| 43 | 2 | | L.IPC(+57.02)VTVADC(+57.02)P(+15.99)RMDEPLH.M | Y | 30.02 | 2024.912 | 17 | -1.7 | 675.9767 | 3 | 37.6 | 40 | 56 | Carbamidomethylation; Hydroxylation | |
| 43 | 2 | | L.VGAC(+57.02)WPC(+57.02)IRSLY | Y | 29.22 | 1480.696 | 12 | -3.1 | 741.3527 | 2 | 51.74 | 61 | 72 | Carbamidomethylation | |
| 43 | 2 | | L.IPC(+57.02)VTVADC(+57.02)P(+15.99)RMD.E | Y | 29.2 | 1548.674 | 13 | -4 | 775.3409 | 2 | 35.73 | 40 | 52 | Carbamidomethylation; Hydroxylation | |
| 43 | 2 | | A.AYERTEPIMHN.G | Y | 28.49 | 1359.624 | 11 | 2.9 | 680.8214 | 2 | 29.88 | 23 | 33 | | |
| 43 | 2 | | A.AYERTEPIM(+15.99)H.N | Y | 26.4 | 1261.576 | 10 | -1.8 | 631.7942 | 2 | 26.45 | 23 | 32 | Oxidation (M) | |
| 43 | 2 | | G.ACI(+57.02)WPC(+57.02)IRSLY | Y | 26.25 | 1324.606 | 10 | -2.2 | 663.3086 | 2 | 50.11 | 63 | 72 | Carbamidomethylation | |
| 43 | 2 | | A.AYERTEPIM(+15.99)HN.G | Y | 25.72 | 1375.619 | 11 | 2.4 | 688.8184 | 2 | 26.19 | 23 | 33 | Oxidation (M) | |
| 43 | 2 | | R.TEPIM(+15.99)HNGEPINL.I | Y | 24.36 | 1479.703 | 13 | 0.9 | 740.8593 | 2 | 38.68 | 27 | 39 | Oxidation (M) | |
| 43 | 2 | | M.HNGEPINL.I | Y | 23.49 | 892.4402 | 8 | -4.9 | 447.2252 | 2 | 35.79 | 32 | 39 | | |
| 43 | 2 | | AAYERTEPIM(+15.99)H | Y | 22.97 | 1124.517 | 9 | -1.9 | 563.2648 | 2 | 29.1 | 23 | 31 | Oxidation (M) | |
| 114 | 11 | G7KV52\|G7KV52_MEDTR | S.FEEC(+57.02)KEDADC(+57.02)HPVC(+57.02)LS | Y | 48.13 | 1794.665 | 14 | -1 | 898.3387 | 2 | 29.46 | 27 | 40 | Carbamidomethylation | |
| 114 | 11 | | E.EC(+57.02)KEDADC(+57.02)HPVC(+57.02)LS | Y | 34.49 | 1518.554 | 12 | -2.5 | 760.2822 | 2 | 24.58 | 29 | 40 | Carbamidomethylation | |
| 114 | 11 | | E.EC(+57.02)KEDADC(+57.02)HPV.C | Y | 29.36 | 1358.523 | 11 | -1.9 | 680.2675 | 2 | 24.51 | 29 | 39 | Carbamidomethylation | |
| 114 | 34 | A0A072UBM1\|A0A072UBM1_MEDTR | A.FEEC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 48.13 | 1794.665 | 14 | -1 | 898.3387 | 2 | 29.46 | 25 | 38 | Carbamidomethylation | |
| 114 | 34 | | E.EC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 34.49 | 1518.554 | 12 | -2.5 | 760.2822 | 2 | 24.58 | 27 | 38 | Carbamidomethylation | |
| 114 | 34 | | E.EC(+57.02)KEDADC(+57.02)HPV.C | Y | 29.36 | 1358.523 | 11 | -1.9 | 680.2675 | 2 | 24.51 | 27 | 37 | Carbamidomethylation | |
| 114 | 10 | G7KV53\|G7KV53_MEDTR | S.FEEC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 48.13 | 1794.665 | 14 | -1 | 898.3387 | 2 | 29.46 | 33 | 46 | Carbamidomethylation | |
| 114 | 10 | | E.EC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 34.49 | 1518.554 | 12 | -2.5 | 760.2822 | 2 | 24.58 | 35 | 46 | Carbamidomethylation | |
| 114 | 10 | | E.EC(+57.02)KEDADC(+57.02)HPV.C | Y | 29.36 | 1358.523 | 11 | -1.9 | 680.2675 | 2 | 24.51 | 35 | 45 | Carbamidomethylation | |
| 114 | 33 | *G7KV51\|G7KV51_MEDTR* | A.FEEC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 48.13 | 1794.665 | 14 | -1 | 898.3387 | 2 | 29.46 | 31 | 44 | Carbamidomethylation | |
| 114 | 33 | | E.EC(+57.02)KEDADC(+57.02)HPVC(+57.02).S | Y | 34.49 | 1518.554 | 12 | -2.5 | 760.2822 | 2 | 24.58 | 33 | 44 | Carbamidomethylation | |
| 114 | 33 | | E.EC(+57.02)KEDADC(+57.02)HPV.C | Y | 29.36 | 1358.523 | 11 | -1.9 | 680.2675 | 2 | 24.51 | 33 | 43 | Carbamidomethylation | |
| 219 | 7 | A0A072T706\|A0A072T706_MEDTR | Y.YVEC(+57.02)IDGFC(+57.02)YYPV | Y | 35.72 | 1811.79 | 14 | 2.5 | 906.9044 | 2 | 51.3 | 36 | 49 | Carbamidomethylation | 21-49 |
| 219 | 7 | | Y.KVEC(+57.02)IDGFC(+57.02)YYPV | Y | 34.45 | 1648.727 | 13 | -1.5 | 825.3693 | 2 | 49.01 | 37 | 49 | Carbamidomethylation | |
| 112 | 5 | A0A072U9K0\|A0A072U9K0_MEDTR | S.GGFGIK(+57.02)VRDADC(+57.02)VDEVC(+57.02)SPGC<br>(+57.02)NKRC(+57.02)WGFEC(+57.02)QC(+57.02)PL | Y | 29.88 | 3992.661 | 34 | 8.5 | 799.5463 | 5 | 40.2 | 39 | 72 | Carbamidomethylation | 26-72 |
| 112 | 5 | | G.RITHDPSTRSTVS.G | Y | 29.7 | 1455.743 | 13 | 3.1 | 728.881 | 2 | 23.89 | 26 | 38 | | |
| 112 | 5 | | F.GKC(+57.02)VRDADC(+57.02)VDE.V | Y | 26.96 | 1422.587 | 12 | -0.4 | 712.3004 | 2 | 25.39 | 42 | 53 | Carbamidomethylation | |
| 112 | 5 | | G.RITHDPSTR.S | Y | 23.56 | 1081.563 | 9 | 0.7 | 541.7891 | 2 | 20.01 | 26 | 34 | | |
| 276 | 22 | A0A072U8J4\|A0A072U8J4_MEDTR | A.LAGC(+57.02)ITDADC(+57.02)VIKKC(+57.02).S | Y | 36.46 | 1722.81 | 15 | -3.1 | 575.2756 | 3 | 36.69 | 27 | 41 | Carbamidomethylation | 27-60 |
| 276 | 22 | | A.GC(+57.02)ITDADC(+57.02)VIK.K | Y | 33.55 | 1250.564 | 11 | -0.2 | 626.2889 | 2 | 35.72 | 29 | 39 | Carbamidomethylation | |
| 276 | 22 | | L.AGC(+57.02)ITDADC(+57.02)VIK.K | Y | 29.91 | 1321.601 | 12 | 0.6 | 661.808 | 2 | 36 | 28 | 39 | Carbamidomethylation | |
| 169 | 6 | G7JAL7\|G7JAL7_MEDTR | E.KDKC(+57.02)VTDFDC(+57.02)VKN.L | Y | 35.92 | 1627.733 | 13 | -4.1 | 814.8707 | 2 | 31.13 | 43 | 55 | Carbamidomethylation | |
| 169 | 6 | | E.KDKC(+57.02)VTDFDC(+57.02)VKC.N<br>W.LC(+57.02)PIDQPVRC(+57.02)IDETC(+57.02)KC(+57.02)<br>2)I.L | Y | 34.76 | 1513.69 | 12 | 1.4 | 505.5714 | 3 | 31.79 | 43 | 54 | Carbamidomethylation | |
| 169 | 6 | | P.RYPGIPRDQ | Y | 25.18 | 2326.058 | 18 | -2.5 | 776.358 | 3 | 60.96 | 58 | 75 | Carbamidomethylation | |
| 196 | 18 | A0A072VKY3\|A0A072VKY3_MEDTR | D.EDC(+57.02)PVGEKC(+57.02)GGNC(+57.02)VE | Y | 26.78 | 1100.573 | 9 | 0.5 | 551.2939 | 2 | 30.26 | 62 | 70 | | 27-70 |
| 273 | 15 | A0A072TH93\|A0A072TH93_MEDTR | D.EDC(+57.02)PVGEKC(+57.02)GGNC(+57.02).V<br>N.IPGARC(+57.02)AITDEDC(+57.02)PVGEKC(+57.02)IGG | Y | 32.97 | 1821.733 | 16 | -3 | 911.8712 | 2 | 33.28 | 33 | 48 | Carbamidomethylation | 23-48 |
| 273 | 15 | | D.EDC(+57.02)PVGEKC(+57.02)GGNC(+57.02).V | Y | 31.22 | 1593.622 | 14 | -2 | 797.8168 | 2 | 30.95 | 33 | 46 | Carbamidomethylation | |
| 273 | 15 | | NC(+57.02)VE | Y | 30.64 | 2763.172 | 25 | 3.4 | 922.0678 | 3 | 35.01 | 24 | 48 | Carbamidomethylation | |
| 173 | 4 | A0A072TTX6\|A0A072TTX6_MEDTR | T.LPC(+57.02)ATDTDC(+57.02)PK.M | Y | 34.14 | 1276.543 | 11 | 1.5 | 639.2796 | 2 | 29.39 | 36 | 46 | Carbamidomethylation | 23-69 |
| 173 | 4 | | F.KC(+57.02)LNGFC(+57.02)LE.H | Y | 26.41 | 1139.51 | 9 | 8.3 | 570.7672 | 2 | 40.06 | 54 | 62 | Carbamidomethylation | |
| 289 | 25 | A7KHC0\|A7KHC0_MEDTR | Y.GPC(+57.02)TTLQDC(+57.02)ETHN.W | Y | 39.43 | 1531.603 | 13 | 0.6 | 766.8093 | 2 | 30.04 | 28 | 40 | Carbamidomethylation | 25-58 |
| 289 | 25 | | Y.GPC(+57.02)TTLQDC(+57.02)ETH.N | Y | 31.08 | 1417.56 | 12 | 1.9 | 709.7887 | 2 | 30.2 | 28 | 39 | Carbamidomethylation | |
| 209 | 37 | A0A072U1G7\|A0A072U1G7_MEDTR | G.KC(+57.02)VSDAEC(+57.02)SGQYMC(+57.02)PTLTVIK<br>C(+57.02)IKDEC(+57.02)VC(+57.02)HYDHDKQ | Y | 42.46 | 4546.981 | 37 | 2 | 910.4054 | 5 | 49.25 | 27 | 63 | Carbamidomethylation | |
| 209 | 37 | | G.KC(+57.02)VSDAEC(+57.02)SGQ.Y | Y | 24.8 | 1239.486 | 11 | -2.7 | 620.7486 | 2 | 22.45 | 27 | 37 | Carbamidomethylation | |
| 690 | 24 | A7KHF0\|A7KHF0_MEDTR | G.SNDEC(+57.02)VYPDVFQC(+57.02)NNWC(+57.02)KC(+<br>57.02)VSHHRT | Y | 52.82 | 3251.4 | 26 | 3.2 | 651.2894 | 5 | 54.99 | 32 | 57 | Carbamidomethylation | 28-57 |
| 691 | 21 | G7IZA9\|G7IZA9_MEDTR | G.QEC(+57.02)KDDGDC(+57.02)PTNMC(+57.02)LPSLVSK<br>C(+57.02)INFIC(+57.02)EC(+57.02)THSM.S | Y | 51.15 | 4005.626 | 33 | 6.5 | 1002.42 | 4 | 64.78 | 25 | 57 | Carbamidomethylation | 25-60 |
| 692 | 40 | G7JHR4\|G7JHR4_MEDTR | S.HSFIPC(+57.02)VTKDDC(+57.02)AYDEC(+57.02)SPRKP<br>TC(+57.02)YLETC(+57.02)HCI(+57.02)L | Y | 48.55 | 4131.772 | 33 | -4 | 827.3583 | 5 | 44.93 | 24 | 56 | Carbamidomethylation | |
| 401 | 17 | A0A072UTM9\|A0A072UTM9_MEDTR | A.KRNYQC(+57.02)DPC(+57.02)FGHPD.D<br>M.C(+57.02)PLETVVKC(+57.02)VEDEC(+57.02)CC(+57.02<br>)VH | Y | 37.25 | 1792.741 | 14 | -0.6 | 598.5872 | 3 | 29.96 | 25 | 38 | Carbamidomethylation | |
| 697 | 16 | A0A072TZN8\|A0A072TZN8_MEDTR | C.EKDKDC(+57.02)PKPPR.I | Y | 38.37 | 2245.984 | 18 | -0.4 | 749.6683 | 3 | 52.2 | 41 | 58 | Carbamidomethylation | |
| 416 | 35 | G7IR68\|G7IR68_MEDTR | C.DC(+57.02)PKPPR.I | Y | 25.78 | 1368.682 | 11 | 3.5 | 457.2361 | 3 | 20.93 | 37 | 47 | Carbamidomethylation | |
| 416 | 35 | | K.IGC(+57.02)KTSEDC(+57.02)PYLGKC(+57.02)IEDFC(+5<br>7.02)QFKK | Y | 23.38 | 868.4225 | 7 | 3.5 | 435.22 | 2 | 22.94 | 41 | 47 | Carbamidomethylation | |
| 698 | 59 | A0A072TI00\|A0A072TI00_MEDTR | F.VTC(+57.02)FEGLC(+57.02)R.M | Y | 36.64 | 2982.338 | 24 | 2.2 | 597.4763 | 5 | 44.15 | 39 | 62 | Carbamidomethylation | 23-62 |
| 389 | 42 | A0A072UAN2\|A0A072UAN2_MEDTR | T.IPC(+57.02)TSDANC(+57.02)PK.V | Y | 33.99 | 1261.543 | 11 | 0 | 631.7788 | 2 | 27.15 | 34 | 44 | Carbamidomethylation | 21-60 |
| 499 | 14 | A0A072UKX9\|A0A072UKX9_MEDTR | T.IPC(+57.02)TSDANC(+57.02)PK.V | Y | 33.99 | 1261.543 | 11 | 0 | 631.7788 | 2 | 27.15 | 34 | 44 | Carbamidomethylation | 23-69 |
| 714 | 38 | G7KV47\|G7KV47_MEDTR | Q.EC(+57.02)KVDADC(+57.02)PTVC(+57.02).T<br>C.MLVTC(+57.02)EDHDC(+57.02)RQLNVQQVGC(+57.02)<br>SFREIPQC(+57.02)NSIC(+57.02)KC(+57.02)MKG | Y | 31.14 | 1452.568 | 12 | -8 | 727.2856 | 2 | 30.89 | 29 | 40 | Carbamidomethylation | |
| 176 | 351 | A7KH73\|A7KH73_MEDTR | | Y | 22.65 | 4704.071 | 38 | -7.5 | 941.8145 | 5 | 52.04 | 25 | 62 | Carbamidomethylation | |

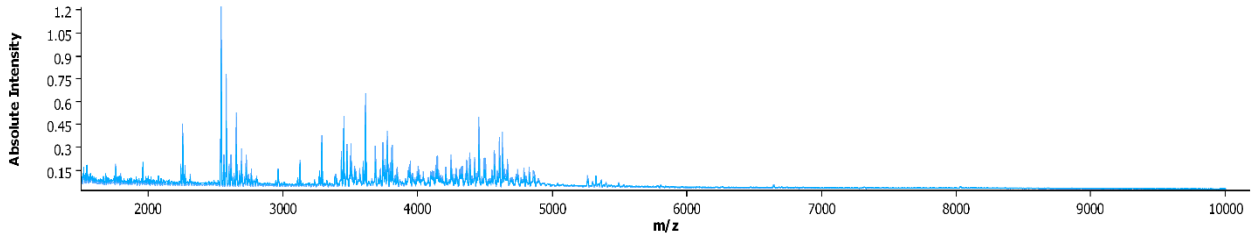
**Supporting Figure S1.** Linear mode spectra averaged over roots and root nodules for performed washes, **(A)** control no wash for comparison to two-step **(B)**, three-step **(C)**, and four-step **(D)** washes and **(E)** control for comparison to single step washes of ethanol **(F)**, chloroform **(G)**, and hexanes **(H)**.



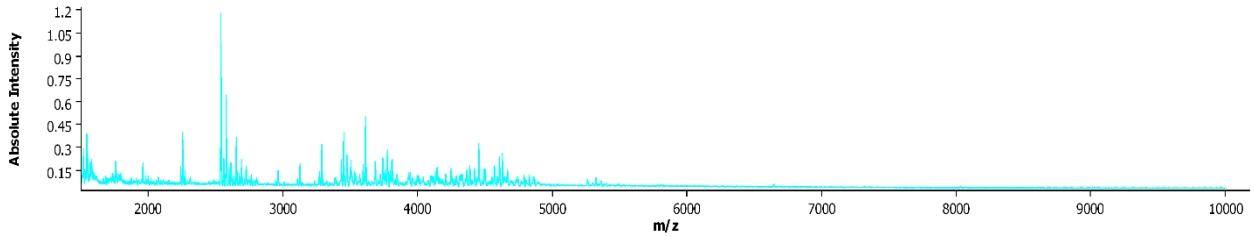
F


Absolute Intensity
1.2
1.05
0.9
0.75
0.6
0.45
0.3
0.15
2000
3000
4000
5000
6000
7000
8000
9000
10000
m/z

G

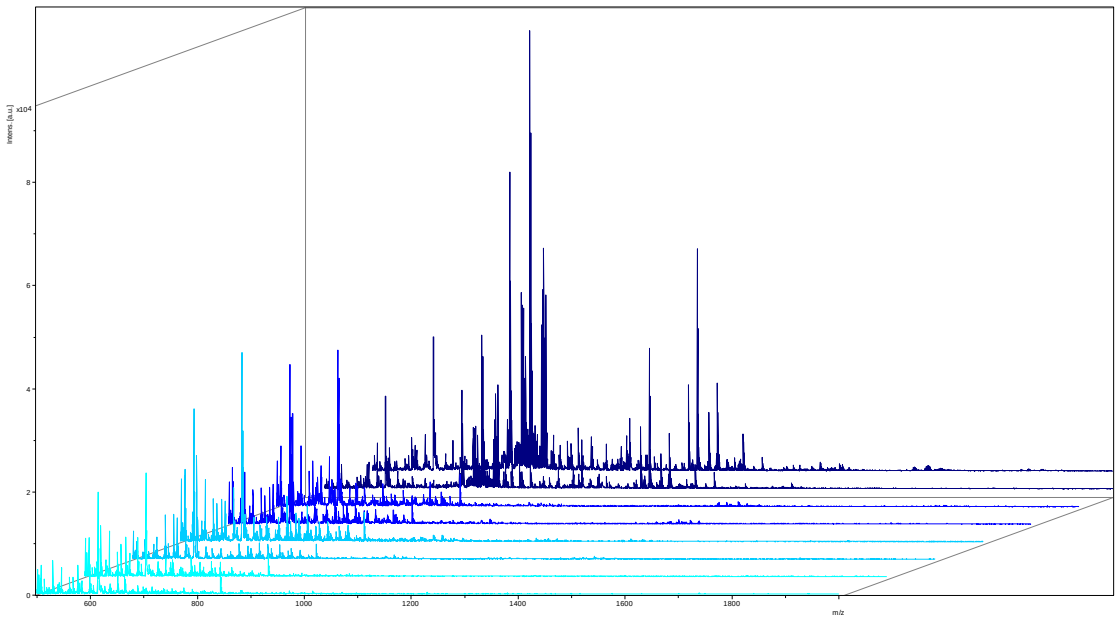

Absolute Intensity
1.2
1.05
0.9
0.75
0.6
0.45
0.3
0.15
2000
3000
4000
5000
6000
7000
8000
9000
10000
m/z

H

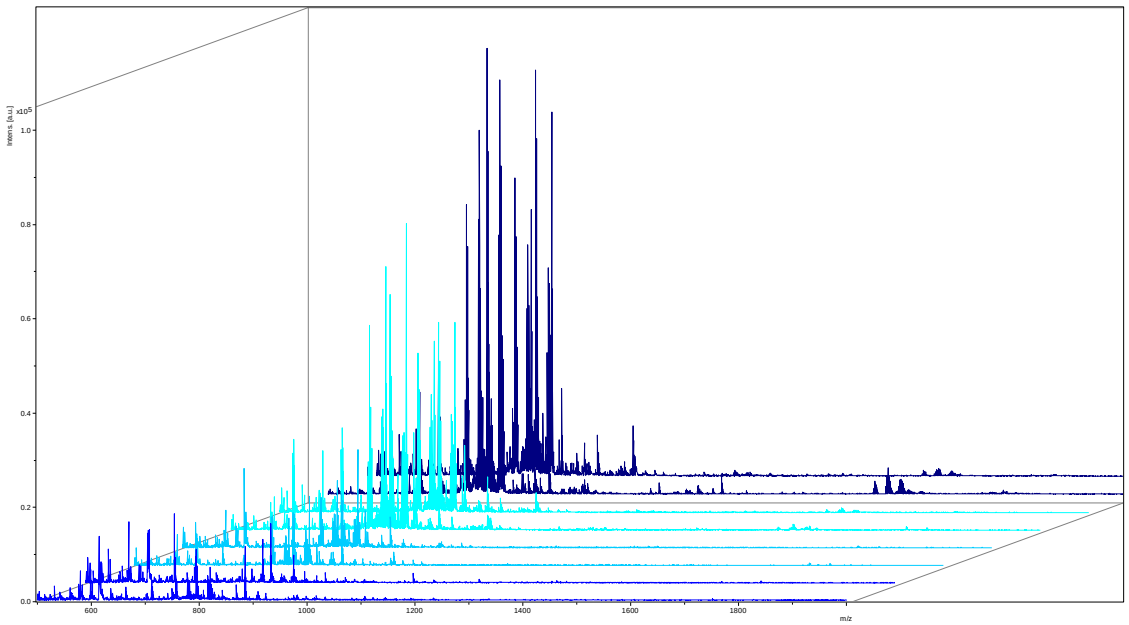

Absolute Intensity
1.2
1.05
0.9
0.75
0.6
0.45
0.3
0.15
2000
3000
4000
5000
6000
7000
8000
9000
10000
m/z

**Supporting Figure S2.** Profiling spectra from 500-2000 *m/z* showing lipid content after washing. In **(A)**, the two-step (bright blue), three-step (light blue), and four-step (teal) washes are compared to a no wash control (dark blue). In **(B)**, the single step washes of ethanol (bright blue), chloroform (light blue), and hexanes (teal) are compared to a no wash control (dark blue).

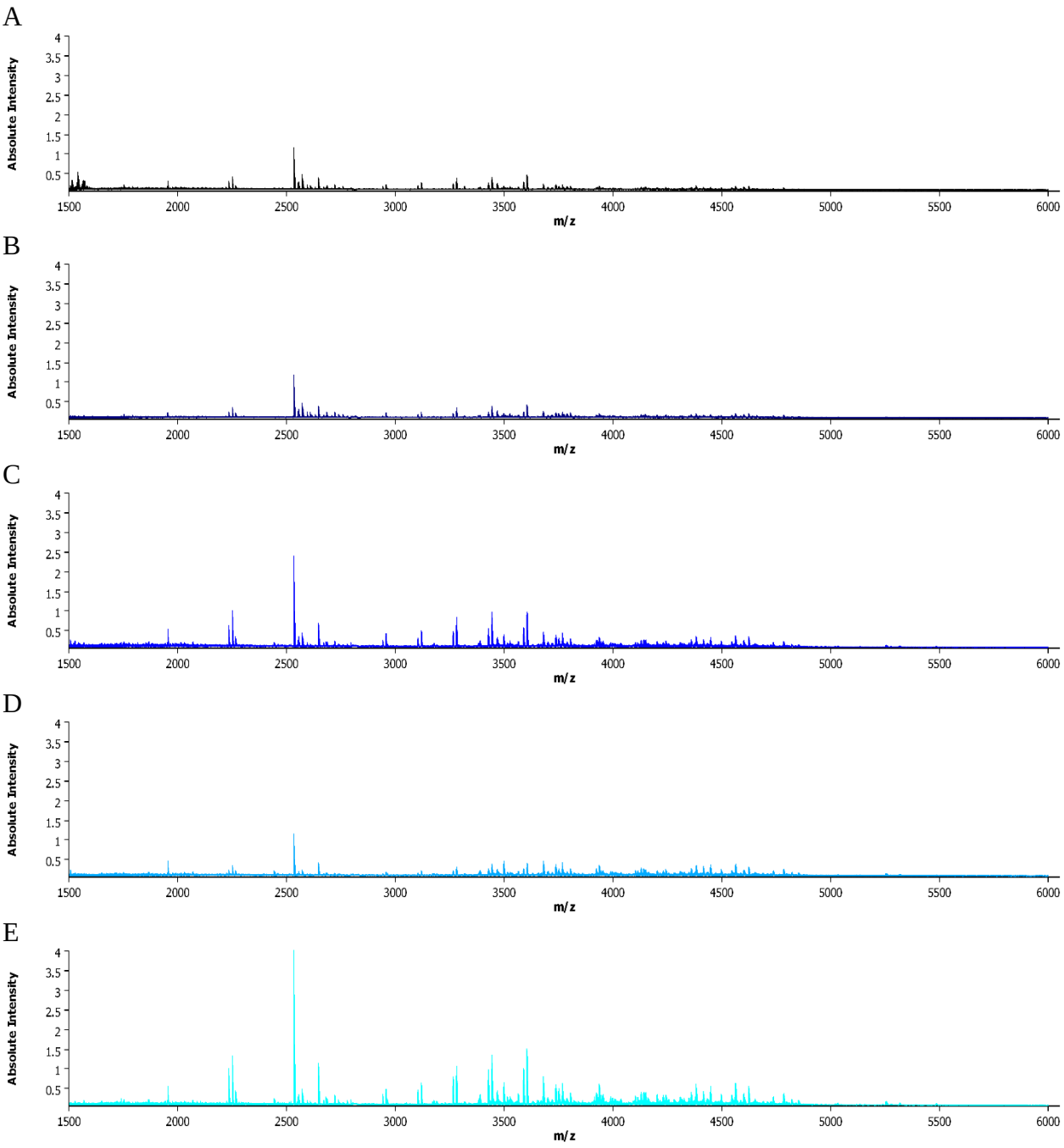
A



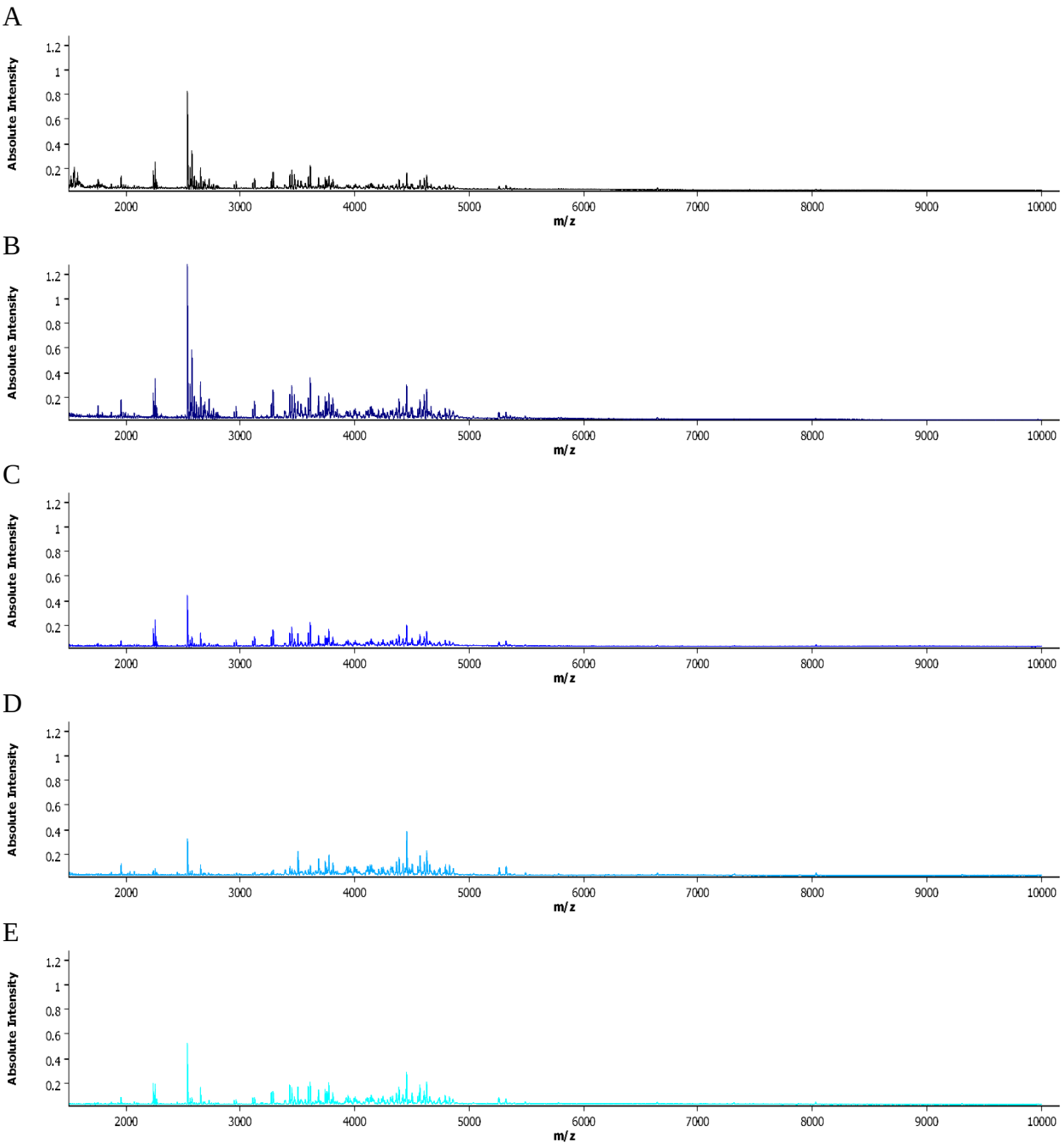

B



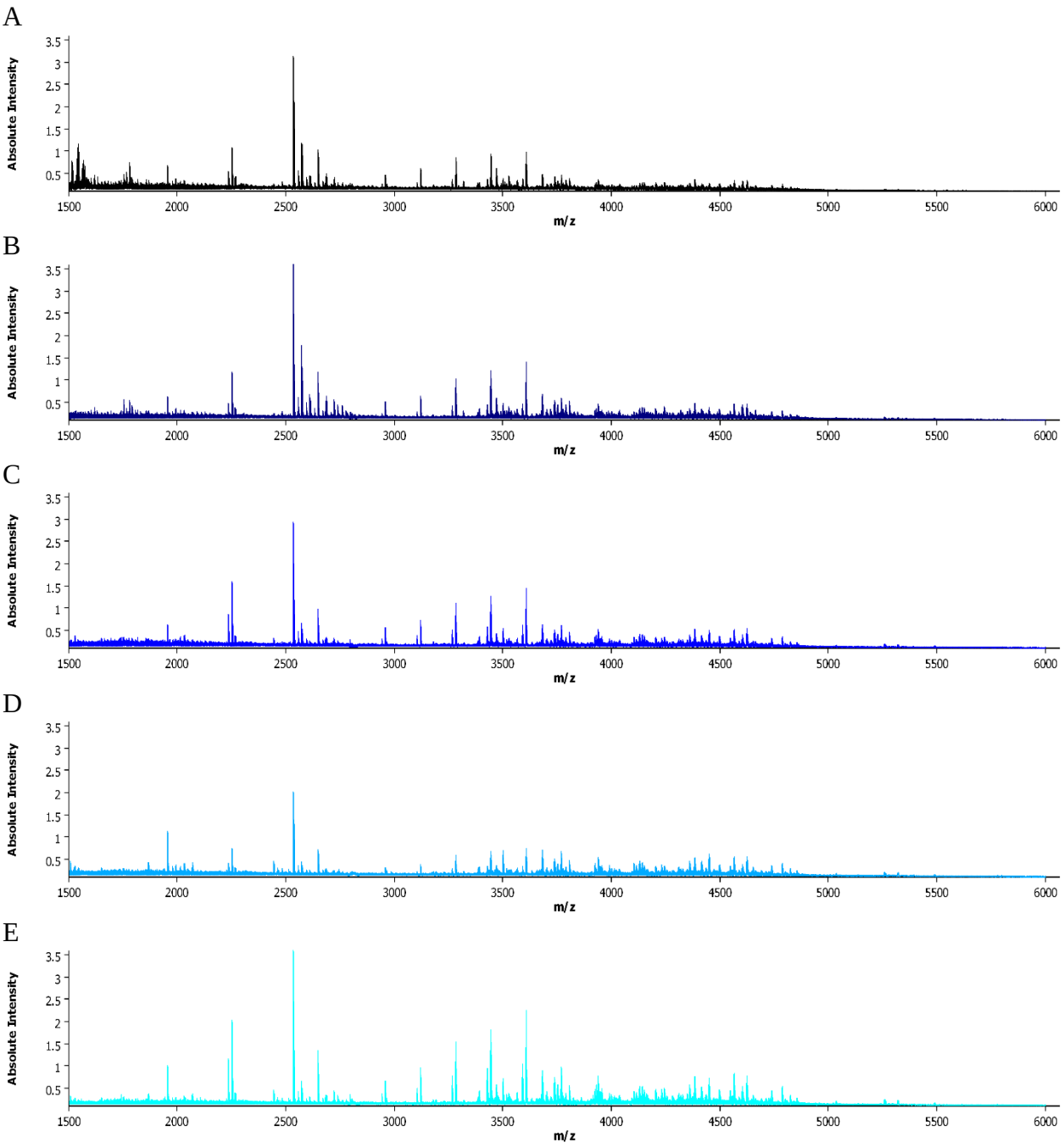
**Supporting Figure S3.** Reflectron spectra for biological replicate 2 showing no wash **(A)**, ethanol wash **(B)**, two-step wash **(C)**, three-step wash **(D)**, and four-step wash **(E)**.



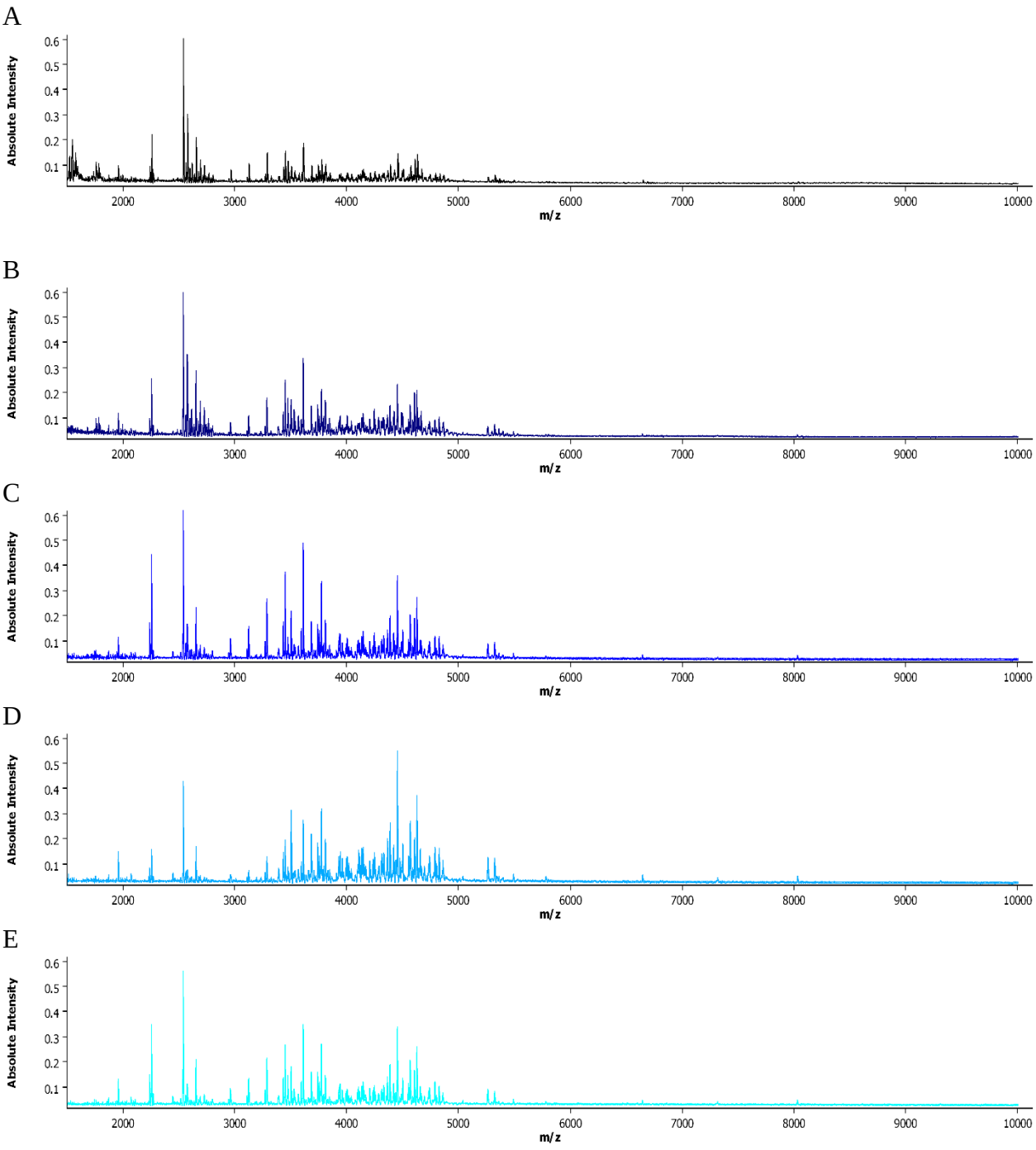
**Supporting Figure S4.** Linear spectra for biological replicate 2 showing no wash **(A)**, ethanol wash **(B)**, two-step wash **(C)**, three-step wash **(D)**, and four-step wash **(E)**.



**Supporting Figure S5.** Reflectron spectra for biological replicate 3 showing no wash **(A)**, ethanol wash **(B)**, two-step wash **(C)**, three-step wash **(D)**, and four-step wash **(E)**.

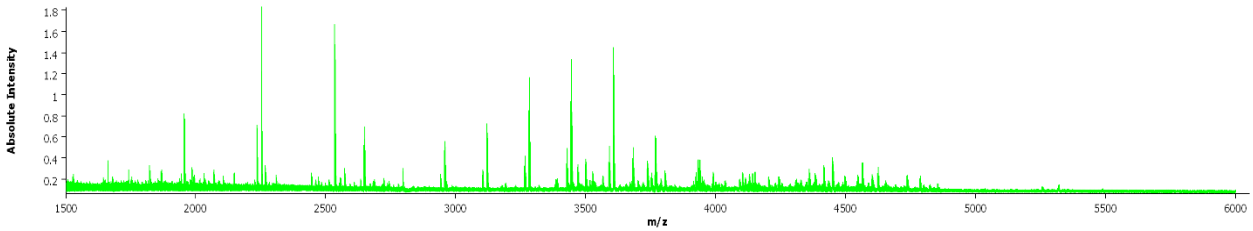


**Supporting Figure S5.** Linear spectra for biological replicate 3 showing no wash **(A)**, ethanol wash **(B)**, two-step wash **(C)**, three-step wash **(D)**, and four-step wash **(E)**.

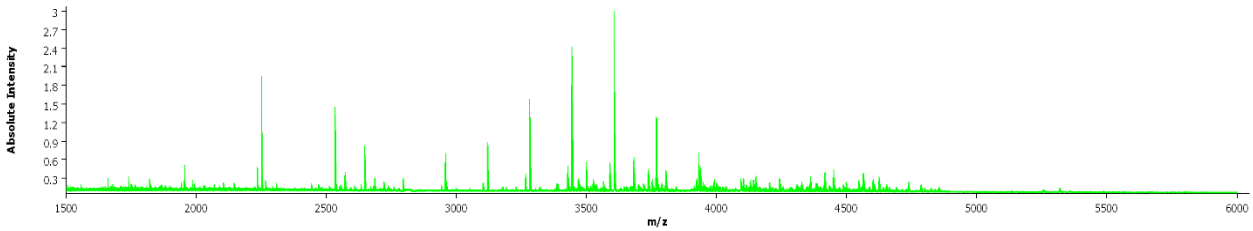


**Supporting Figure S6.** Spectra averaged over nodule and root for root nodules treated with Denator. The Two biological replicates are shown in **(A)** and **(B)**.

A




B



# Chapter 6

# Extraction Optimization for Combined Metabolomics, Peptidomics, and Proteomics Analysis of Gut Microbiota Samples

Adapted from: Keller C, Wei P, Wancewicz B, Cross T-W L, Rey FE, Li L. Extraction Optimization for Combined Metabolomics, Peptidomics, and Proteomics Analysis of Gut Microbiota Samples. *Anal Bioanal Chem*. In Preparation.

Keywords: LC-MS; Metabolomics; Microbiome; Infection; Metagenomics

**Abstract**

Multiomic studies are increasingly used to gain a deeper understanding of molecular processes occurring in a biological system, such as the complex microbial communities (*i.e.*, microbiota) that reside the distal gut. While a combination of metabolomics and proteomics is more commonly used, multiomics studies including peptidomcis characterization are less frequently undertaken. Here, we investigated three different extraction methods, chosen for their previous use in extracting metabolites, peptides, and proteins, and compared their ability to perform metabolomic, peptidomic, and proteomic analysis of mouse cecum content. The methanol/ chloroform/ water extraction performed the best for metabolomic and peptidomic analysis as it detected the largest number of small molecules and identified the largest number of peptides, but the acidified methanol extraction performed best for proteomics analysis as it had the highest number of protein identifications. The methanol/ chloroform/ water extraction was further analyzed by identifying metabolites with MS/MS analysis and by gene ontology analysis for the peptide and protein results to provide a multiomics analysis of the gut microbiota.

**Introduction**

The gastrointestinal tract is host to large and dynamic communities of microbes, containing ~$10^8$ to $10^{10}$ organisms per gram in the illium and stool,[1] and encompassing the 3 domains of life: bacterial, archaea and eukaryotes.[2, 3] The gut microbiota has various roles in human health, including nutrition and immune function modulation.[4, 5] Furthermore, the disruption of the gut microbiota has been connected to various inflammatory chronic conditions such as inflammatory bowel disease,[6] obesity[7, 8] and cardiovascular disease.[9, 10] Due to the importance that the microbiota plays in human health, there is an abundance of studies on various aspects of the microbiota.

Research includes various aspects of the microbiome such as the metagenomic[11] and proteomic content,[12] but also looks at how the microbiome influences other systems, including blood and cerebrospinal fluid metabolites.[13, 14]

While studying individual "omics" results in a depth of information about the genomic, proteomic or metabolomic content of the desired biological specimen, the combination of multiple of these analyses can determine different connections between molecular classes in a biological system. Consequently, multiomic approaches are used to study the microbiota, as well as in other fields.[15-18] A combination of metagenomic and metabolomic approaches has been used, for example, to study the microbiota's response to infection,[19] the effect of arsenic on the gut microbiome,[20] and the effect of trichloroacetamide exposure on the gut microbiome and urinary metabolites.[21] Other combinations of metagenomics, metatranscriptomics, metabolomics, and (meta)proteomics can also be employed.[22] Although large amounts of data created in multiomic techniques can be challenging and time consuming to analyze, advances in software are arising to simplify the task.[23, 24]

While the field of peptidomics is more recent than the fields of metabolomics and proteomics, it has nevertheless risen to importance.[25, 26] Well-studied signaling peptides in the brain, *i.e.* neuropeptides, are an important class of molecules regulating a wide variety of processes.[27, 28] Similarly, peptide hormones in the endocrine system are also important endogenous peptides that have been studied due to their role in regulating metabolism.[29, 30] In the gut, bioactive peptides, derived from digestion of proteins in the intestines, has resulted in a number of peptides with roles in health.[31] For example, bioactive peptides present in the gut that inhibit the angiotensin I-converting enzyme (ACE) can reduce hypertension and improve cardiovascular health.[32, 33]

Although multiomic studies with a combination of genomics, metabolomics, and proteomics is common, peptidomics is much less frequently combined with other "omic" techniques. Therefore, we set out to identify an optimal extraction method ideal for combining peptidomics with metabolomics and proteomics for a multiomics approach to analyze the gut microbiota. Here, we tested three different extraction methods for combined metabolomics, peptidomics, and proteomics. The three extractions chosen have been previously used in various biological systems to achieve good results in one or more of the metabolomics, peptidomics, and proteomics fields and comprise of a variety of solvent systems. The chloroform/ methanol/ water extraction is a common small molecule extraction that can be used for metabolomics and proteomics multiomics studies,[34] and has previously been used to study cecal metabolomics.[19] An acidified methanol extraction is commonly used in neuropeptide extractions[29, 35] but has also been applied for metabolomics.[36] Finally, the 40/40/20 acetonitrile/ methanol/ water extraction adds acetonitrile to the extraction solvent, which has been shown to be beneficial for certain metabolites[37] and has been shown to work well in metabolomics studies.[38] The small molecules, peptides, and proteins detected for each extraction were compared and used to determine an optimal extraction for metabolomic, peptidomic, and proteomic analysis. While the acidified methanol extraction performed best for the proteomics experiments, the methanol/ chloroform/ water extraction performed the best in terms of the number of small molecules and peptides detected and was further evaluated with gene ontology analysis.

## Materials and Methods

### Cecum Collection

All animal procedures were approved by the Institutional Animal Care and Use Committee at the University of Wisconsin-Madison. Germ-free (n=1) and conventionally-raised (n=1) male C57BL/6 mice were used and fed LabDiet #5021 (Purina Mills, Inc. Richmond, IN). At 21 weeks of age, mice were euthanized via $CO_2$ inhalation and exsanguination. Cecal content was collected, snap frozen in liquid nitrogen immediately, and stored in the -80°C until analysis.

Sample Preparation

The two cecum samples were combined and separated into three approximately equal aliquots. On one aliquot, a methanol/ chloroform/ water extraction was performed in a PTFE tube by adding in order, 3 parts methanol, 1 part chloroform, and 4 parts Milli-Q ultrapure water (total volute 4.0 mL). The tube was vortexed and centrifuged at 3200 x g and 4$^{o}$C for 15 minutes. The upper aqueous layer was removed, and 4 parts methanol were added to the tube and vortexed. The tube was centrifuged again at 1500 x g and 4$^{o}$C for 5 minutes. The supernatant was removed (organic fraction). The aqueous fraction, organic fraction, and pellet were dried down and saved in the -80$^{o}$C until further processing. The second aliquot was extracted with acidified methanol (methanol/ water/ acetic acid 90/9/1 v/v/v) and the third aliquot with 40% methanol 40% acetonitrile 20% water. Both of these two extractions were probe sonicated for 3 cycles (8s on 15 s off) at 4$^{o}$C and centrifuged at 15,000 x g and 4$^{o}$C for 15 minutes. The supernatant and pellet were separated and then dried down in a speed vac and saved in the -80$^{o}$C until further processing.

A 3 kDa molecular weight cut-off (MWCO) filtration was performed on the aqueous and organic fraction of the methanol/ chloroform/ water extraction, the acidified methanol (AcMeOH) extraction, and the methanol/ acetonitrile/ water extraction. The Amicon Ultra (Millipore) MWCO device was first rinsed with 0.2 mL 0.1 M NaOH, and 0.5 mL 50% methanol. Both rinses were centrifuged at 14,000 x g until the rinse was through the membrane. The sample was then loaded

into the device and centrifuged through the device at 14,000 x g. A final rinse of 0.1 mL 50% methanol was added to the device and the MWCO was centrifuged at 14,000 x g. The content below 3 kDa was split into two aliquots (one metabolomics and one peptidomics) for each of the four samples and dried down in a speed vac. The MWCO was rinsed with 0.4 mL 50% methanol and equilibrated for 5 mins, and then flipped over and centrifuged at 14,000 x g for 2 mins to collect content above 3 kDa from the device. A 30kDa MWCO was performed on the above 3 kDa fraction to separate the extract into peptidomics (3 kDa to 30kDa) and proteomics (above 30kDa) fractions.

The below 3 kDa and 3-30 kDa peptidomics fractions were combined. Sep-Pak C18 was used for peptide desalting, and then peptide samples were dried down in a speed vac and saved at -80$^{o}$C until LC-MS/MS analysis. The proteomics fractions from the supernatant contents above 30 kDa and from the pellets were combined. The aqueous and organic fraction above 30 kDa from the methanol/ chloroform/ water extraction method were combined together with the pellet. The protein mixture samples were dissolved in 1 mL ice-cold PBS, and debris from the pellet was removed with a low centrifuge speed (300 x g, 4$^{o}$C for 5 mins).[39] The supernatant was carefully collected, and the pellets were washed another two times with PBS and all the supernatant obtained from each time was combined. The supernatant was then centrifuged at 20,000 x g for 10 min to pellet the bacterial cells and host cells. The pellet was then lysed with 8 M urea Lysis buffer with sonication (On 8 sec, Off 15 sec, 3 cycles). The total protein concentration of each pellet was determined by BCA assay and then digested with Trypsin/Lys-C mixture overnight. Then, the digested proteins were desalted, dried down in a speed vac, and saved in the -80$^{o}$C until LC-MS/MS analysis.

Metabolomics Data Acquisition and Analysis

The aqueous fraction was resuspended at 10 mg/mL in optima grade water with 0.1% formic acid while the other three metabolomic samples (organic fraction, AcMeOH sample, MeOH/AcN/water sample) were resuspended at 10 mg/mL in optima grade methanol with 0.1% formic acid. Any samples that were cloudy were centrifuged briefly and the supernatant used for the analysis. LC-MS/MS analysis was performed with a Dionex Ultimate 3000 UHPLC system connected to a Q Exactive mass spectrometer (Thermo Scientific). Separation occurred on a Cortecs C18 column (2.1 mm internal diameter x 100 mm length, 1.6 μm particle size; Waters), equipped with a corresponding guard column with a column temperature of 35°C, and mobile phases of optima grade water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). A 35 minute gradient at a flow rate of 0.3 mL/minutes with the following conditions was used for separation: 0–5 min, 1% B; 5–10 min, linear gradient from 1–3% B; 10–18 min, linear gradient from 3–40% B; 18–22 min, linear gradient from 40–80% B; 22–27 min, column cleaning at 95% B; and 27–35 min, re-equilibration at 1% B. A top 5 data dependent acquisition method was used for MS/MS of small molecules in the extractions. The full MS settings were 70,000 resolution, 1e6 AGC, 100 ms max inject time, 100-1500 *m/z*. The MS/MS settings were 35,000 resolution, 1e5 AGC, 100 ms max inject time, 1.0 *m/z* isolation window, and 30 dynamic exclusion. Three technical replicates were run for each extraction, and each technical replicate used a different HCD collision energy (25, 30, 40 respectively). Samples were run in both positive and negative mode.

Compound Discoverer software was used to analyze the LC-MS/MS data for each extraction in both positive and negative ion mode (the aqueous and organic fractions were analyzed together). Individual runs were aligned with an adaptive curve model with a maximum shift of 1 minute and 5 ppm tolerance. Unknown compounds were detected with a 5 ppm mass tolerance, 30% intensity tolerance, 3 S/N ratio, and 1,000,000 minimum peak intensity. Unknown

compounds were grouped with a 5 ppm mass tolerance and 0.1-minute retention time tolerance. A fill gaps step was used with 5 ppm mass error and 0.1 retention time error. Constant sum normalization and marking of background compounds were used. The MS/MS spectra were searched in the mzCloud library against all activation types and activation energies and matches were manually validated by ensuring that all the major fragment ions in the database spectra matched the experimental spectra. MS/MS spectra were also searched against the MassBank of North America MS/MS database for additional identifications.

Peptidomics Acquisition and Analysis

Peptide samples were resuspended in optima grade water with 3% acetonitrile and 0.1% formic acid. LC-MS/MS was performed on an Ultimate 3000 UPLC system coupled with the Orbitrap Fusion™ Tribrid™ Mass Spectrometer. A 75 μm × 16 cm homemade column packed with 1.7 μm, 150 Å, BEH C18 material obtained from a Waters (Milford, MA) UPLC column (part no. 186004661) was used for label-free peptide separation at a flow rate of 0.3 μl/min. Mobile phase A was 0.1% formic acid in optima water and mobile phase B was 0.1% formic acid in optima acetonitrile. The 145 min optimized gradient used was as follows: 0-18.33 min, 3% solvent B; 18.33-30 min, 3-10% B; 30-50 min, 10-20% B; 50-108 min, 20-75% B; 108-118 min, 75% B; 118-118.5 min 75%-95% B; 118.5-128 min, 95% B; 128-128.5 min, 95%-3% B; 128.5-145 min, 3% B. Full MS scans were acquired from *m/z* 300 to 1500 at a resolution of 60 K, automatic gain control (AGC) at $2 \times 10^5$, and maximum injection time (IT) of 100 ms. The top 20 precursors were then selected for higher-energy C-trap dissociation tandem mass spectrometry (HCD MS2) analysis with an isolation window of 1 *m/z*, a HCD collision energy (NCE) of 30, a resolving power of 15 k, an AGC target of $5 \times 10^4$, a maximum injection time of 100 ms, and a lower mass limit of 120 *m/z*.

The .raw data files from the Orbitrap MS analysis were searched against a combined database which included food, 93 strains of bacteria, and mouse proteome from Uniport with PEAKS STUDIO 8.5 software. A precursor tolerance of 10 ppm and a fragment mass tolerance of 0.02 Da were allowed. Acetylation (N-term), amidation, oxidation (M), pyro-Glu from E, pyro-Glu from Q, sulfation (STY), were set as rare dynamic modifications and allowing three maximum variable PTM per peptide. Parameters for confident peptide identification were Ascore (PTM site confidence) higher than 20, FDR lower than 1%, and the presence of at least one unique peptide.

Proteomics Acquisition and Analysis

A Dionex UltiMate 3000 nanoLC system coupled with a Q Exactive HF Orbitrap MS was used for Ultra-performance LC-MS analysis. Homemade column and mobile phases were the same as mentioned above for the peptidomic analysis. The optimized gradient used was as follows: 0-16 min, 3% solvent B; 16-20 min, 3-25% B; 20-30 min, 25-45% B; 30-50 min, 45-70% B; 50-56 min, 70-95% B; 56-60 min 95% B; 60-60.5 min, 95-3% B; 60.5-70 min, 3% B. Full MS scans were acquired from *m/z* 300 to 1500 at a resolution of 60 K, AGC at $1 \times 10^6$, and maximum injection time (IT) of 100 ms. The top 15 precursors were then selected for higher-energy C-trap dissociation tandem mass spectrometry (HCD MS2) analysis with an isolation window of 1.4 *m/z*, a normalized collision energy (NCE) of 30, a resolving power of 15 K, an AGC target of $1 \times 10^5$, a maximum injection time of 100 ms, and a lower mass limit of 120 *m/z*. PEAKS software was used for protein identification. The parameters were the same as used above for peptidomics analysis, except trypsin with D&P enzyme was selected for this bottom-up proteomics study. Non-specific cleavage at both ends of the peptide was allowed and the maximum missed cleavages per peptide was set at two.

## Results

Two cecal content samples were combined and split into three approximately equal aliquots to test three different extraction protocols. The sample preparation workflow is provided in **Supplemental Figure S1** for the three extractions. The three extractions tested were a methanol/ chloroform/ water (MeOH/$CHCl_3$/$H_2O$) extraction, an acidified methanol (AcMeOH) extraction, and a methanol, acetonitrile, water (MeOH/AcN/$H_2O$) extraction. The methanol/ chloroform/ water extraction results in two liquid fractions, and aqueous fraction and an organic fraction, as well as a pellet. The other two extractions resulted in a supernatant and a pellet. The four liquid portions were processed with a 3 kDa molecular weight cut-off followed by a 30 kDa molecular weight cut-off. The <3 kDa portion was split for LC-MS metabolomics analysis and nano-LC-MS peptidomics analysis. The 3<x<30 kDa portion was saved for nano-LC-MS peptidomics analysis as well. The >30 kDa fraction was combined with the pellet for bottom up nano-LC-MS proteomics analysis.

Metabolomics

Metabolomics data was analyzed in Compound Discoverer 2.0 to detect unknown compounds and perform tandem MS (MS/MS) matching of experimental MS/MS spectra to the mzCloud high resolution/accurate mass spectral database. Each of the three extractions were analyzed separately in the software with the same parameters to test how many compounds were detected in each extraction. The aqueous and organic fractions for the methanol/ chloroform/ water extraction were combined in the software. **Figure 1** shows the results for the metabolomics analysis. In both positive and negative mode, the methanol/ chloroform/ water extraction detected about 2-times more compounds than the other extractions in Compound Discoverer. Compound monoisotopic molecular weights were processed in METLIN with a 5 ppm error to approximate

how many of the *m/z* potentially matched to known small molecules. The methanol/ chloroform/ water extraction still out-performed the other two extractions in both positive and negative mode in the METLIN analysis. In order to compare the detected *m/z* across the three extractions, the METLIN results were used to create Venn diagrams. To be consistent, the lowest METLIN identification number was used for the *m/z* that had multiple accurate mass matches to the METLIN database. In the Venn diagrams (**Figure 1(B)**, positive data, **Figure 1(D)**, negative data), a majority of the compounds are detected with the methanol, chloroform water extraction, with only a small percentage unique to one of the other two extractions. Thus, the methanol/ chloroform/ water extraction performs the best for metabolomics analysis.

Peptidomics

Peptidomic analysis was performed on a nano-LC QE-HF system and the data analyzed in PEAKS 8.5 software. **Figure 2** shows the results of the peptidomics experiments by showing the peptide sequences detected and by comparing the protein accession numbers detected. In **Figure 2(A)**, the number of peptide sequences is shown with the number of peptide sequences shared in both technical replicates represented with the diagonal lines. The organic fraction of the methanol/ chloroform/ water extraction had the most peptides identified, with the aqueous fraction in second. The other two extractions showed low peptide identifications. When comparing the proteins detected from the peptide sequences that were shared between both technical replicates for the organic and aqueous fractions, most of the aqueous proteins are also in the organic fraction, but there are proteins only in the aqueous fraction. To compare the three extractions, the combination of unique proteins from the aqueous and organic fraction were taken, along with the proteins in both technical replicates for the other two extractions. The Venn diagram shows only 2 proteins that were not detected by the methanol/ chloroform/ water extraction and over 100 proteins that

were only detected in the methanol/ chloroform/ water extraction. Thus, the methanol/ chloroform/ water extraction performs best for the peptidomics analysis as it identifies the most proteins and covers almost all of the proteins identified in the other two extractions.

To look further at the differences between the peptides detected in the aqueous and organic fractions, peptide sequences found in each fraction from the proteins present in both technical replicates were compared with respect to their length and isoelectric point. The isoelectric point was calculated using the Peptide Property calculator (GeneScript, online tool) as it allowed the inclusion of certain post translational modifications (PTMs), for example acetylation. PTMs that were not present in this online tool were excluded from the isoelectric point analysis, but this was minority for each fraction (11/198 for organic fraction 2/89 for aqueous fraction). **Figure 3** shows the results of the length and isoelectric point comparison. The length distribution in the aqueous and organic fractions seemed to be similar as the organic fraction had approximately 2-fold more peptides sequences at each length, and overall the organic fraction has about 2 times more sequences. The isoelectric point distribution, however, does appear to be different between the two fractions. While the number of peptides with very low pI's is approximately the same for both fractions, the organic fraction has more peptide sequences with pI's above 4. The organic fraction potentially extracted a larger number of peptides than the aqueous fraction due to the preference of the peptides for the methanol/ chloroform solvent.

Proteomics

Bottom-up proteomics was performed on the pellets and the >30 kDa content from the molecular weight cut-off step. The results of this analysis are shown in **Figure 4**. The acidified methanol extraction had noticeably more proteins than the other two extractions. While only proteins detected in both technical replicates were counted as identifications, it is worth noting that

many proteins were only in 1 technical replicate, rather than both technical replicates (as shown by the diagonal lines). Potentially adding in a third technical replicate in the experimental analysis could increase the number of identifications by increasing the identifications seen in multiple technical replicates. The comparison of the three extractions shows that each extraction has hundreds of identifications unique to that extraction, with the acidified methanol extraction having over a thousand identifications unique to it. For searching, a combined multi-organism database was used containing the mouse genome, the genomes of approximately 90 bacterial strains for a model microbiota,[19] and potential proteins from the food the mice were fed. While conventionally raised mice very likely have a more diverse microbial community,[40] here a more focused database was used because increasing the database size to a more comprehensive one would make processing the MS/MS data significantly more challenging. For studies highly interested in the microbial species present, a targeted microbiome database without the mouse and food components could be used instead. **Figure 4(C)** shows the number of proteins that match to each of the different potential sources of the protein, namely the mouse, the microbiome, and the food. The methanol/ acetonitrile/ water extraction has the most proteins matching to both the mouse database and the food proteins. Where the acidified methanol extraction gets a vast majority of its protein identifications from is from microbiome proteins, as it has a much higher number of microbiome identifications compared to the other two extractions.

Combined Metabolomics, Peptidomics, Proteomics

**Figure 5** combines the results of the three different omics by total identifications and by a source comparison for the combined peptide and protein results. The total identifications were calculated with the total number of hits to the METLIN database from the positive mode LC-MS data for the metabolomics analysis, the identified proteins from the peptidomics study shared

between both technical replicates and, the protein number from the proteomics study shared between both technical replicates. When all three omics are combined, the methanol/ chloroform/ water and the acidified methanol extractions have similar number of total identifications. When the peptide and protein results are combined, and the source of the protein identification investigated, the methanol/ chloroform/ water, and methanol/ acetonitrile/ water extractions have similar numbers of mouse proteins. However, the acidified methanol extraction still has the most identifications due to the large number of microbial protein identifications. For microbiome studies focusing on the host response, the methanol/ chloroform/ water performs well for metabolomics, peptidomics, and proteomics, but if microbial proteins are desired, the acidified methanol extraction preforms best.

**Discussion**

In order to get a good representation of metabolites, peptides, and proteins, the methanol/ chloroform/ water extraction was chosen for further investigation into the compounds detected. For metabolomics identification, the mzCloud and MassBank small molecule MS/MS databases were used for spectral matching of fragment ions. Potential matches to either database were manually inspected for verification. **Supplemental Figure S2** compares the total small molecule putative identifications in the methanol/ chloroform/ water extraction with the two MS/MS databases with both databases in positive and negative modes, as well as comparing the overlap between the positive and negative putative identifications. **Supplemental Table S1** provides all the small molecule identifications for the methanol/ chloroform/ water extraction with their molecular weight, retention time, and which database and polarity they were identified from. In positive mode, both databases provided a similar number of identifications and resulted in

complementary coverage for a total of 57 identifications. In negative mode, almost all the results came from the MSDial database. Overall, the positive and negative mode identifications results showed complementary coverage as many identifications were made only in one of the polarities. To verified putatively identified compounds, standards could be obtained, and retention times and fragmentation patterns compared to the experimental data. More identifications are potentially possible using other databases (*i.e.*, METLIN) or using *in silico* fragmentation software. However, these were not utilized here due to the time involved in metabolomics identification and the lack of a biological experiment that would show upregulation or downregulation of certain *m/z* in a biological condition.

The identifications for the peptidomics results for the methanol/ chloroform/ water extraction are in **Supplemental Table S2** (aqueous fraction) and **Supplemental Table S3** (organic fraction), and the proteomics results for the methanol/ chloroform/ water extraction are provided in **Supplemental Table S4**. Gene Ontology analysis was conducted on the protein and peptide results from the methanol/ chloroform/ water extraction using DAVID Bioinformatics Resources 6.8.[41, 42] The gene ontology results of the biological processes for the detected peptides and proteins are shown in **Figure 6**. The top biological process for both the protein and peptide results was proteolysis. Other shared biological process include digestion, chromatin silencing, regulation of systemic arterial blood pressure by renin-angiotensin, and metabolic process. Overall, the many shared biological processes in the protein and peptide gene ontology results indicate that the protein and peptide result agree well with each other. By integrating the peptide results with the protein results, a more comprehensive understanding of the biological processes can be achieved for a biological question of interest.

To look further at the peptidomcis results, the peptides belonging to the biological process of digestion, a key function of the digestive track that the cecum is a part of, were investigated. **Table 1** shows the peptides that fell under the biological process of digestion. Peptide sequences from the enzymes chymotrypsinogen B1(Ctrb1), serine protease, and trypsin were detected. Further analysis would be necessary to discover the potential role that these peptides play. Furthermore, an *in silico* study reported the potential for endogenous proteins in the gut to be digested into bioactive peptides.[43] Experimentally, peptides were observed from three proteins, Mucin-13, Chymotrypsinogen B, and pancreatic triacylglycerol lipase, that have predicted potential for bioactive peptide release after intestinal digestion of the gut endogenous protein. The predicted activity for peptides from these three proteins is ACE inhibition, which can prevent hypertension.[33]

Multiomic analysis can provide a greater understanding by studying not just the one subclass of molecules, protein changes, for example, but also changes in the metagenomic, or metabolome in order to better understand the relationships between various biological systems. Here, the ability of three extractions for combined metabolomics, peptidomics, and proteomics analysis was compared. The methanol/ chloroform/ water extraction method enabled a more comprehensive view of all three omics in the mouse host system and preformed particularly strong in metabolomics and peptidomics analysis. By including peptidomics in the multiomics experiments, a deeper understanding of the role of peptides could be obtained, for example, by characterization and discovery of bioactive peptides and their role in various pathways.

**Acknowledgements**

This study was supported in part by grant funding from the NIH (R01DK071801, RF1AG052324, to LL) and a grant from a Transatlantic Networks of Excellence Award from the Leducq Foundation (to FER). The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531 to LL) and Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. T.-W. L. C. is supported by the National Institutes of Health, under Ruth L. Kirschstein National Research Service Award T32 HL 007936 from the National Heart Lung and Blood Institute to the University of Wisconsin-Madison Cardiovascular Research Center. LL acknowledges a Vilas Distinguished Achievement Professorship and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

**References**

1. Sender, R.; Fuchs, S.; Milo, R., Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol* **2016,** *14* (8), e1002533.

2. Scanlan, P. D.; Marchesi, J. R., Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *Isme j* **2008,** *2* (12), 1183-93.

3. Hoffmann, C.; Dollive, S.; Grunberg, S.; Chen, J.; Li, H.; Wu, G. D.; Lewis, J. D.; Bushman, F. D., Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One* **2013,** *8* (6), e66019.

4. Sharon, G.; Garg, N.; Debelius, J.; Knight, R.; Dorrestein, P. C.; Mazmanian, S. K., Specialized metabolites from the microbiome in health and disease. *Cell Metab* **2014,** *20* (5), 719-730.

5. Kau, A. L.; Ahern, P. P.; Griffin, N. W.; Goodman, A. L.; Gordon, J. I., Human nutrition, the gut microbiome and the immune system. *Nature* **2011,** *474* (7351), 327-336.

6. Kostic, A. D.; Xavier, R. J.; Gevers, D., The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* **2014,** *146* (6), 1489-1499.

7. Turnbaugh, P. J.; Backhed, F.; Fulton, L.; Gordon, J. I., Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* **2008,** *3* (4), 213-23.

8. Rosenbaum, M.; Knight, R.; Leibel, R. L., The gut microbiota in human energy homeostasis and obesity. *Trends Endocrinol Metab* **2015,** *26* (9), 493-501.

9. Jie, Z.; Xia, H.; Zhong, S. L.; Feng, Q.; Li, S.; Liang, S.; Zhong, H.; Liu, Z.; Gao, Y.; Zhao, H.; Zhang, D.; Su, Z.; Fang, Z.; Lan, Z.; Li, J.; Xiao, L.; Li, R.; Li, X.; Li, F.; Ren, H.; Huang, Y.; Peng, Y.; Li, G.; Wen, B.; Dong, B.; Chen, J. Y.; Geng, Q. S.; Zhang, Z. W.; Yang, H.; Wang, J.; Zhang, X.; Madsen, L.; Brix, S.; Ning, G.; Xu, X.; Liu, X.; Hou, Y.; Jia, H.; He, K.; Kristiansen, K., The gut microbiome in atherosclerotic cardiovascular disease. *Nat Commun* **2017,** *8* (1), 845.

10. Tang, W. H.; Kitai, T.; Hazen, S. L., Gut Microbiota in Cardiovascular Health and Disease. *Circ Res* **2017,** *120* (7), 1183-1196.

11. Turnbaugh, P. J.; Hamady, M.; Yatsunenko, T.; Cantarel, B. L.; Duncan, A.; Ley, R. E.; Sogin, M. L.; Jones, W. J.; Roe, B. A.; Affourtit, J. P.; Egholm, M.; Henrissat, B.; Heath, A. C.; Knight, R.; Gordon, J. I., A core gut microbiome in obese and lean twins. *Nature* **2009,** *457* (7228), 480-U7.

12. Verberkmoes, N. C.; Russell, A. L.; Shah, M.; Godzik, A.; Rosenquist, M.; Halfvarson, J.; Lefsrud, M. G.; Apajalahti, J.; Tysk, C.; Hettich, R. L.; Jansson, J. K., Shotgun metaproteomics of the human distal gut microbiota. *Isme Journal* **2009,** *3* (2), 179-189.

13. Wikoff, W. R.; Anfora, A. T.; Liu, J.; Schultz, P. G.; Lesley, S. A.; Peters, E. C.; Siuzdak, G., Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (10), 3698-3703.

14. Vogt, N. M.; Romano, K. A.; Darst, B. F.; Engelman, C. D.; Johnson, S. C.; Carlsson, C. M.; Asthana, S.; Blennow, K.; Zetterberg, H.; Bendlin, B. B.; Rey, F. E., The gut microbiota-derived metabolite trimethylamine N-oxide is elevated in Alzheimer's disease. *Alzheimer's Research & Therapy* **2018,** *10* (1), 124.

15. Roume, H.; Muller, E. E.; Cordes, T.; Renaut, J.; Hiller, K.; Wilmes, P., A biomolecular isolation framework for eco-systems biology. *Isme j* **2013,** *7* (1), 110-21.

16. Wettersten, H. I.; Hakimi, A. A.; Morin, D.; Bianchi, C.; Johnstone, M.; Donohoe, D. R.; Trott, J. F.; Abu Aboud, O.; Stirdivant, S.; Neri, B.; Wolfert, R.; Stewart, B.; Perego, R.; Hsieh, J. J.; Weiss, R. H., Grade-Dependent Metabolic Reprogramming in Kidney Cancer Revealed by Combined Proteomics and Metabolomics Analysis. *Cancer Research* **2015,** *75* (12), 2541-2552.

17. Weckwerth, W.; Wenzel, K.; Fiehn, O., Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks. *Proteomics* **2004,** *4* (1), 78-83.

18. Mostafa, I.; Zhu, N.; Yoo, M. J.; Balmant, K. M.; Misra, B. B.; Dufresne, C.; Abou-Hashem, M.; Chen, S.; El-Domiaty, M., New nodes and edges in the glucosinolate molecular network revealed by proteomics and metabolomics of Arabidopsis myb28/29 and cyp79B2/B3 glucosinolate mutants. *J Proteomics* **2016,** *138*, 1-19.

19. Bratburd, J. R.; Keller, C.; Vivas, E.; Gemperline, E.; Li, L.; Rey, F. E.; Currie, C. R., Gut Microbial and Metabolic Responses to Salmonella enterica Serovar Typhimurium and Candida albicans. *MBio* **2018,** *9* (6).

20. Lu, K.; Abo, R. P.; Schlieper, K. A.; Graffam, M. E.; Levine, S.; Wishnok, J. S.; Swenberg, J. A.; Tannenbaum, S. R.; Fox, J. G., Arsenic exposure perturbs the gut microbiome and its metabolic profile in mice: an integrated metagenomics and metabolomics analysis. *Environ Health Perspect* **2014,** *122* (3), 284-91.

21. Zhang, Y.; Zhao, F.; Deng, Y.; Zhao, Y.; Ren, H., Metagenomic and metabolomic analysis of the toxic effects of trichloroacetamide-induced gut microbiome and urine metabolome perturbations in mice. *J Proteome Res* **2015,** *14* (4), 1752-61.

22. Kaiser, B. L. D.; Li, J.; Sanford, J. A.; Kim, Y. M.; Kronewitter, S. R.; Jones, M. B.; Peterson, C. T.; Peterson, S. N.; Frank, B. C.; Purvine, S. O.; Brown, J. N.; Metz, T. O.; Smith, R. D.; Heffron, F.; Adkins, J. N., A Multi-Omic View of Host-Pathogen-Commensal Interplay in Salmonella-Mediated Intestinal Infection. *Plos One* **2013,** *8* (6), 13.

23. Fan, J.; Saha, S.; Barker, G.; Heesom, K. J.; Ghali, F.; Jones, A. R.; Matthews, D. A.; Bessant, C., Galaxy Integrated Omics: Web-based Standards-Compliant Workflows for Proteomics Informed by Transcriptomics. *Mol Cell Proteomics* **2015,** *14* (11), 3087-93.

24. Caspi, R.; Billington, R.; Fulcher, C. A.; Keseler, I. M.; Kothari, A.; Krummenacker, M.; Latendresse, M.; Midford, P. E.; Ong, Q.; Ong, W. K.; Paley, S.; Subhraveti, P.; Karp, P. D., The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res* **2018,** *46* (D1), D633-d639.

25. Clynen, E.; Baggerman, G.; Veelaert, D.; Cerstiaens, A.; Van der Horst, D.; Harthoorn, L.; Derua, R.; Waelkens, E.; De Loof, A.; Schoofs, L., Peptidomics of the pars intercerebralis-corpus cardiacum complex of the migratory locust, Locusta migratoria. *Eur J Biochem* **2001,** *268* (7), 1929-39.

26. Verhaert, P.; Uttenweiler-Joseph, S.; de Vries, M.; Loboda, A.; Ens, W.; Standing, K. G., Matrix-assisted laser desorption/ionization quadrupole time-of-flight mass spectrometry: an elegant tool for peptidomics. *Proteomics* **2001,** *1* (1), 118-31.

27. An, S.; Harang, R.; Meeker, K.; Granados-Fuentes, D.; Tsai, C. A.; Mazuski, C.; Kim, J.; Doyle, F. J., 3rd; Petzold, L. R.; Herzog, E. D., A neuropeptide speeds circadian entrainment by reducing intercellular synchrony. *Proc Natl Acad Sci U S A* **2013,** *110* (46), E4355-61.

28. Schank, J. R.; Ryabinin, A. E.; Giardino, W. J.; Ciccocioppo, R.; Heilig, M., Stress-related neuropeptides and addictive behaviors: beyond the usual suspects. *Neuron* **2012,** *76* (1), 192-208.

29. Yu, Q.; Canales, A.; Glover, M. S.; Das, R.; Shi, X.; Liu, Y.; Keller, M. P.; Attie, A. D.; Li, L., Targeted Mass Spectrometry Approach Enabled Discovery of O-Glycosylated Insulin and Related Signaling Peptides in Mouse and Human Pancreatic Islets. *Anal Chem* **2017,** *89* (17), 9184-9191.

30. Barreto, S. G.; Carati, C. J.; Toouli, J.; Saccone, G. T., The islet-acinar axis of the pancreas: more than just insulin. *Am J Physiol Gastrointest Liver Physiol* **2010,** *299* (1), G10-22.

31. Moller, N. P.; Scholz-Ahrens, K. E.; Roos, N.; Schrezenmeir, J., Bioactive peptides and proteins from foods: indication for health effects. *Eur J Nutr* **2008,** *47* (4), 171-82.

32. Yamamoto, N.; Akino, A.; Takano, T., Antihypertensive effect of the peptides derived from casein by an extracellular proteinase from Lactobacillus helveticus CP790. *J Dairy Sci* **1994,** *77* (4), 917-22.

33. Dave, L. A.; Hayes, M.; Montoya, C. A.; Rutherfurd, S. M.; Moughan, P. J., Human gut endogenous proteins as a potential source of angiotensin-I-converting enzyme (ACE-I)-, renin inhibitory and antioxidant peptides. *Peptides* **2016,** *76*, 30-44.

34. Sapcariu, S. C.; Kanashova, T.; Weindl, D.; Ghelfi, J.; Dittmar, G.; Hiller, K., Simultaneous extraction of proteins and metabolites from cells in culture. *MethodsX* **2014,** *1*, 74-80.

35. Ye, H.; Wang, J.; Tian, Z.; Ma, F.; Dowell, J.; Bremer, Q.; Lu, G.; Baldo, B.; Li, L., Quantitative Mass Spectrometry Reveals Food Intake-Induced Neuropeptide Level Changes in Rat Brain: Functional Assessment of Selected Neuropeptides as Feeding Regulators. *Mol Cell Proteomics* **2017,** *16* (11), 1922-1937.

36. Cao, Q. J. W.; Ouyang, C. Z.; Zhong, X. F.; Li, L. J., Profiling of small molecule metabolites and neurotransmitters in crustacean hemolymph and neuronal tissues using reversed-phase LC-MS/MS. *Electrophoresis* **2018,** *39* (9-10), 1241-1248.

37. Rabinowitz, J. D.; Kimball, E., Acidic acetonitrile for cellular metabolome extraction from Escherichia coli. *Anal Chem* **2007,** *79* (16), 6167-73.

38. Bazurto, J. V.; Dearth, S. P.; Tague, E. D.; Campagna, S. R.; Downs, D. M., Untargeted metabolomics confirms and extends the understanding of the impact of aminoimidazole carboxamide ribotide (AICAR) in the metabolic network of Salmonella enterica. *Microb Cell* **2017,** *5* (2), 74-87.

39. Zhang, X.; Ning, Z.; Mayne, J.; Moore, J. I.; Li, J.; Butcher, J.; Deeke, S. A.; Chen, R.; Chiang, C. K.; Wen, M.; Mack, D.; Stintzi, A.; Figeys, D., MetaPro-IQ: a universal metaproteomic approach to studying human and mouse gut microbiota. *Microbiome* **2016,** *4* (1), 31.

40. Xiao, L.; Feng, Q.; Liang, S.; Sonne, S. B.; Xia, Z.; Qiu, X.; Li, X.; Long, H.; Zhang, J.; Zhang, D.; Liu, C.; Fang, Z.; Chou, J.; Glanville, J.; Hao, Q.; Kotowska, D.; Colding, C.; Licht, T. R.; Wu, D.; Yu, J.; Sung, J. J.; Liang, Q.; Li, J.; Jia, H.; Lan, Z.; Tremaroli, V.; Dworzynski, P.; Nielsen, H. B.; Backhed, F.; Dore, J.; Le Chatelier, E.; Ehrlich, S. D.; Lin, J. C.; Arumugam, M.; Wang, J.; Madsen, L.; Kristiansen, K., A catalog of the mouse gut metagenome. *Nat Biotechnol* **2015,** *33* (10), 1103-8.

41. Huang da, W.; Sherman, B. T.; Lempicki, R. A., Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **2009,** *37* (1), 1-13.

42. Huang da, W.; Sherman, B. T.; Lempicki, R. A., Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **2009,** *4* (1), 44-57.

43. Dave, L. A.; Montoya, C. A.; Rutherfurd, S. M.; Moughan, P. J., Gastrointestinal endogenous proteins as a source of bioactive peptides--an in silico study. *PLoS One* **2014,** *9* (6), e98922.

**Tables**

**Table 1.** Peptide Sequences from the Digestion biological process in the peptide gene ontology results.

| Identification Name | Uniptrot ID | Sequence | Mass (Da) |
|---|---|---|---|
| chymotrypsinogen B1(Ctrb1) | Q9CR35 | A.GEFDQGSDEENVQVLK.I | 1792.8115 |
| | | K.IAQVFKNPK.F | 1043.6127 |
| protease, serine 2(Prss2) | P07146 | I.NVLEGNEQFVDSAK.I | 1548.7419 |
| trypsin 4(Try4)/ trypsin 5(Try5) | Q9R0T7/ Q9QUK9 | R.TLNNDIM(+15.99)LIK.L | 1189.6377 |

**Figures**

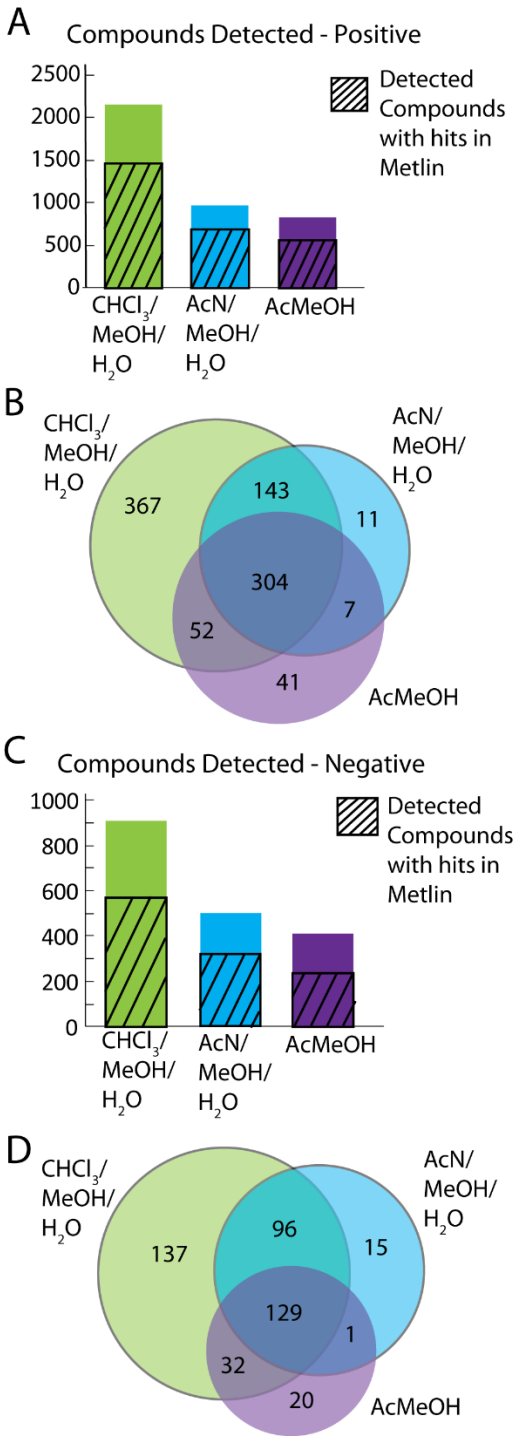



**Figure 1.** Metabolomics results for the three extractions. The number of compounds detected is shown for positive **(A)** and negative **(C)** mode. The *m/z* with hits in METLIN were compared between the three extractions in Venn Diagrams (**B**, positive and **D**, negative mode).

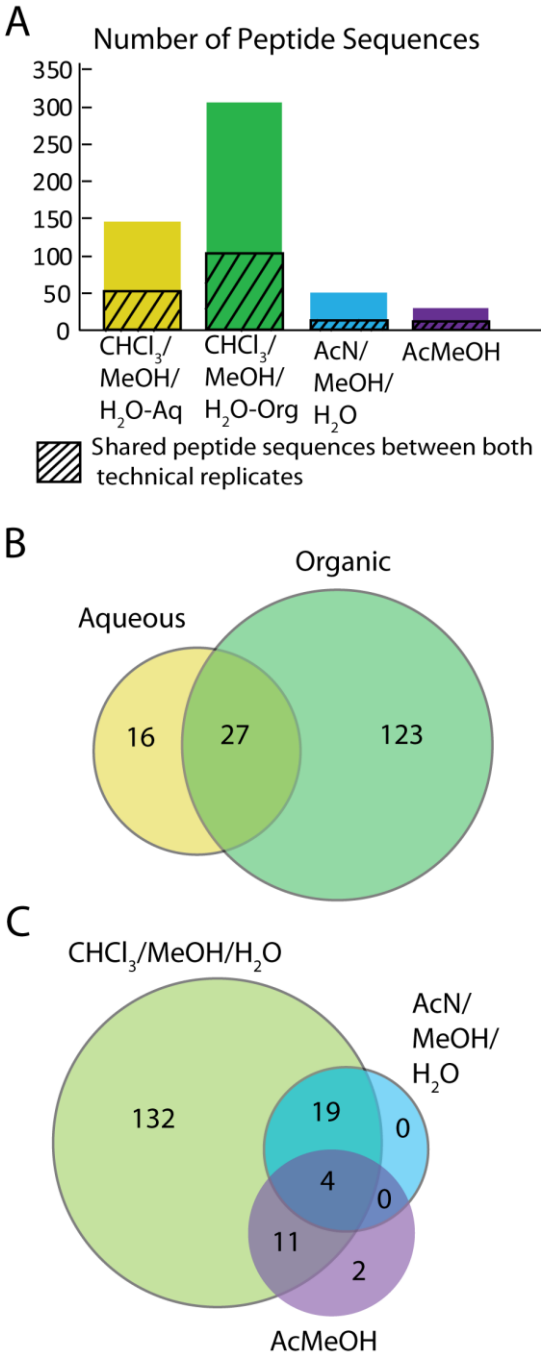



**Figure 2.** Peptidomics results from the extraction test. In **(A)** the number of peptide sequences are shown. **(B)** compares the protein accession numbers in the aqueous and organic fractions of the methanol/ chloroform/ water extraction. **(C)** compares the protein accession numbers for the three extractions.

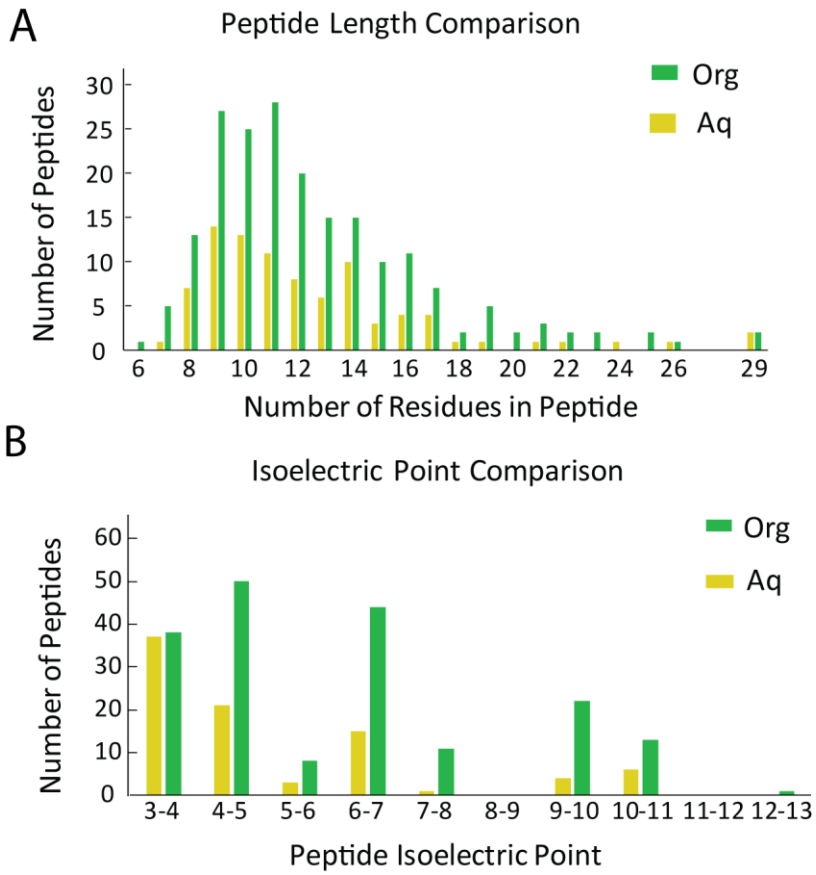



**Figure 3.** Comparison of peptide length **(A)** and isoelectric point **(B)** between peptide sequences detected in organic and aqueous fractions.

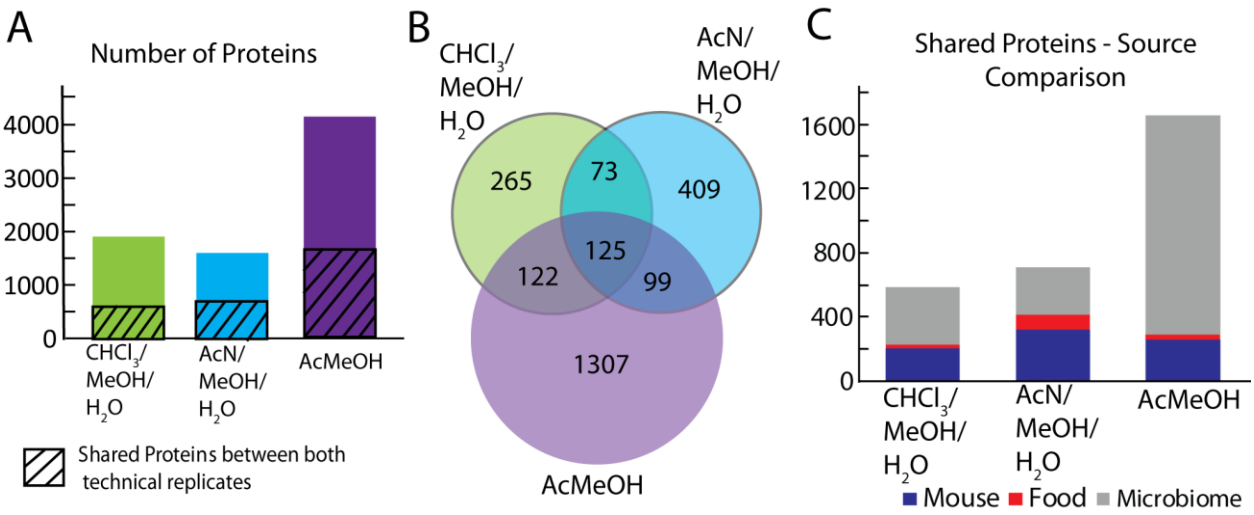



**Figure 4.** Proteomics results from extraction optimization. **(A)** shows the detected proteins with the diagonal lines representing the number of shared proteins in both technical replicates. **(B)** compares the shared proteins in both technical replicates for the three extractions. **(C)** shows the breakdown of the shared proteins to the source of the protein ID for each extraction.

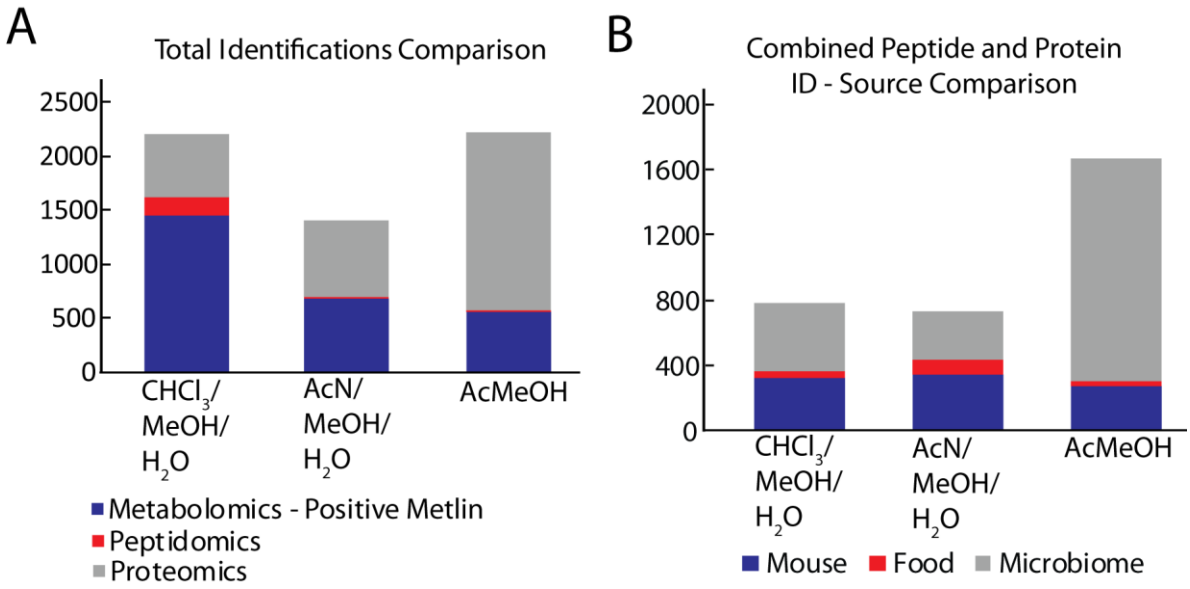



**Figure 5.** Combined results of the three omics analysis. **(A)** shows the combined identifications of metabolomics, peptidomics, and proteomics results. **(B)** compares the protein source for the combined peptide and protein results.

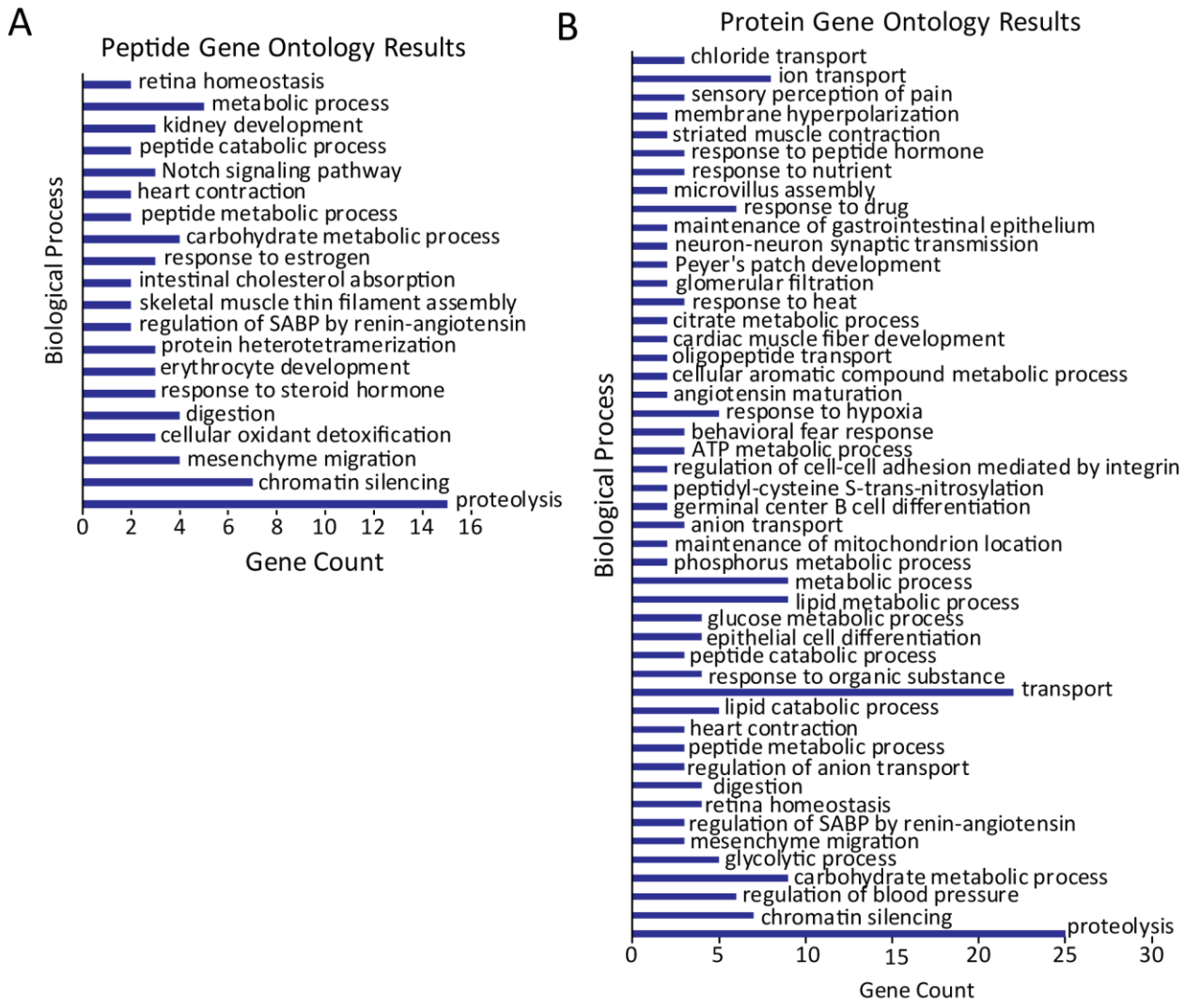



**Figure 6.** DAVID gene ontology results showing the biological processes associated with detected peptide **(A)** and protein **(B).** SABP stands for systemic arterial blood pressure.

## Supplemental Information

## Supplemental Tables

**Supplemental Table S1.** Metabolite Identifications from mzCloud and MSDial database search.

| Name | Molecular Weight | RT [min] | ID Source | Polarity |
|---|---|---|---|---|
| (2S,3S,4S,5R,6R)-6-(3-benzoyloxy-2-hydroxypropoxy)-3,4,5-trihydroxyoxane-2-carboxylic acid | 372.10556 | 14.149 | MSDial | negative |
| (2S,3S,4S,5R,6R)-6-[[(3S,4S,6aR,6bS,8aR,9R,12aS,14bR)-9-hydroxy-4-(hydroxymethyl)-4,6a,6b,8a,11,11,14b-heptamethyl-1,2,3,4a,5,6,7,8,9,10,12,12a,14,14a-tetradecahydropicen-3-yl]oxy]-5-[(2S,3R,4S,5R,6R)-4,5-dihydroxy-6-(hydroxymethyl)-3-[(2S,3R,4R,5R,6S)-3,4,5-trihydroxy-6-methyloxan-2-yl]oxyoxan-2-yl]oxy-3,4-dihydroxyoxane-2-carboxylic acid | 942.51604 | 19.799 | MSDial | negative |
| (2S,3S,4S,5R,6R)-6-[[(3S,6aR,6bS,8aS,14bR)-8a-carboxy-4,4,6a,6b,11,11,14b-heptamethyl-1,2,3,4a,5,6,7,8,9,10,12,12a,14,14a-tetradecahydropicen-3-yl]oxy]-3,4,5-trihydroxyoxane-2-carboxylic acid | 632.39231 | 21.18 | MSDial | negative |
| (2S,3S,4S,5R,6S)-3,4,5-trihydroxy-6-(5-hydroxy-4-oxo-2-phenylchromen-7-yl)oxyoxane-2-carboxylic acid | 430.08983 | 14.227 | MSDial | negative |
| (S)-Equol | 242.09391 | 17.92 | Both | positive |
| 2'-Deoxyguanosine | 267.09655 | 3.064 | mzCloud | positive |
| 2-Hydroxycinnamic acid | 164.04727 | 1.542 | mzCloud | positive |
| 3,4-Dihydroxybenzoic acid | 154.02527 | 7.962 | MSDial | negative |
| 4-Dodecylbenzenesulfonic acid | 326.19163 | 23.014 | mzCloud | negative |
| 4-Guanidinobutanoate | 145.08509 | 1.074 | MSDial | positive |
| 4-Pyridoxate | 183.05205 | 1.515 | MSDial | negative |
| 5,7-dihydroxy-2-(4-hydroxyphenyl)-6-[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]-8-(3,4,5-trihydroxyoxan-2-yl)chromen-4-one | 564.14785 | 14.346 | MSDial | negative |
| 5-Hydroxyindole | 133.05278 | 13.881 | mzCloud | positive |
| 5-Hydroxyindole-3-acetic acid | 191.05802 | 13.707 | mzCloud | positive |
| 6-Hydroxypicolinic acid | 139.02688 | 1.229 | mzCloud | positive |
| 7-Methylguanine | 165.06502 | 1.331 | mzCloud | positive |
| Acetylcarnitine | 203.11575 | 1.109 | MSDial | positive |
| Acetyl-β-methylcholine | 159.12588 | 1.029 | mzCloud | positive |
| Allantoin | 158.04264 | 0.797 | MSDial | negative |
| apigenin-7-O-glucuronide | 446.08489 | 15.154 | MSDial | negative |
| Asparagine | 132.05349 | 0.768 | mzCloud | positive |
| Asparagine | 132.05194 | 0.753 | MSDial | negative |
| Asp-Glu | 262.07994 | 0.854 | MSDial | positive |
| Asp-Phe | 280.10576 | 11.495 | MSDial | negative |
| AZELAIC ACID | 188.1038 | 15.766 | MSDial | negative |
| Cholic Acid | 425.31351 | 19.938 | MSDial | positive |
| Choline | 103.10008 | 0.763 | Both | positive |
| Citric acid | 192.02597 | 1.227 | MSDial | negative |
| Cytosine | 111.03234 | 1.23 | MSDial | positive |
| D-(-)-quinic acid | 192.06242 | 0.819 | MSDial | negative |
| D-(+)-Malic acid | 134.02 | 0.933 | MSDial | negative |
| Daidzein | 254.05753 | 15.262 | Both | positive |
| Deoxycholate | 392.29272 | 20.579 | MSDial | negative |
| D-Gluconic acid | 196.05725 | 0.776 | MSDial | negative |
| D-GLUCURONIC ACID | 194.04164 | 0.777 | MSDial | negative |
| DL-Carnitine | 161.10507 | 0.812 | Both | positive |
| DL-Tryptophan | 204.08974 | 8.472 | Both | positive |
| Docosahexanoic acid | 328.24017 | 23.464 | MSDial | negative |
| Eicosapentaenoic acid | 302.22427 | 23.294 | mzCloud | positive |
| Erucamide | 337.33411 | 25.411 | MSDial | positive |
| Ethoxyquin | 217.14641 | 17.848 | Both | positive |
| Ferulic acid | 194.05773 | 14.724 | MSDial | positive |
| Ferulic acid | 194.05697 | 14.74 | MSDial | negative |
| Genistein | 270.05239 | 17.836 | MSDial | positive |

| | | | | |
|---|---|---|---|---|
| Genistein | 270.0528 | 17.864 | Both | negative |
| Genistein 4'-O-glucuronide | 446.08442 | 15.135 | mzCloud | positive |
| Glutamine | 146.0677 | 0.76 | MSDial | negative |
| Glycitein | 284.06813 | 15.417 | Both | positive |
| Glycitein | 284.06843 | 16.848 | MSDial | negative |
| Guanine | 151.04942 | 2.624 | mzCloud | positive |
| Guanosine | 283.09151 | 2.631 | MSDial | positive |
| GUANOSINE | 283.0916 | 2.633 | MSDial | negative |
| Hexamethoxymethyl melamine | 390.2221 | 17.947 | mzCloud | positive |
| Histidine | 155.06805 | 0.744 | MSDial | negative |
| Hypoxanthine | 136.03852 | 1.226 | mzCloud | positive |
| Indole-3-acrylic acid | 187.06319 | 8.339 | mzCloud | positive |
| Indole-3-carbinol | 147.06826 | 15.22 | MSDial | positive |
| Indole-3-lactic acid | 205.07378 | 14.163 | Both | positive |
| L-(+)-Arginine | 174.11157 | 0.761 | Both | positive |
| L-ASPARTATE | 133.03592 | 0.765 | MSDial | negative |
| L-Glutamic acid | 147.05309 | 0.779 | Both | positive |
| L-Glutamic acid | 147.05167 | 0.77 | MSDial | negative |
| L-Histidine | 155.06941 | 0.757 | Both | positive |
| Linoleic acid | 280.23993 | 22.198 | MSDial | positive |
| L-Lysine | 146.1055 | 0.718 | mzCloud | positive |
| L-Phenylalanine | 165.07888 | 3.276 | Both | positive |
| L-Pyroglutamic Acid | 129.04265 | 1.292 | mzCloud | positive |
| L-Tryptophane | 204.08902 | 8.473 | MSDial | negative |
| L-Tyrosine | 181.07277 | 1.562 | MSDial | negative |
| Myristyl sulfate | 294.18651 | 23.021 | mzCloud | negative |
| N,N'-Dicyclohexylurea | 224.18862 | 19.593 | mzCloud | positive |
| N6,N6,N6-Trimethyl-L-lysine | 188.15245 | 0.736 | mzCloud | positive |
| N-Acetyl-DL-aspartic acid | 175.04689 | 1.101 | MSDial | negative |
| N-Acetyl-DL-glutamic acid | 189.06266 | 1.411 | MSDial | negative |
| N-Acetylhistidine | 197.08002 | 0.852 | MSDial | positive |
| N-Acetylleucine | 173.10398 | 13.388 | MSDial | negative |
| N-Acetylmethionine | 191.06148 | 8.801 | MSDial | positive |
| N-Acetylneuraminic acid | 309.10597 | 0.799 | MSDial | negative |
| N-alpha-acetyl-L-lysine | 188.11598 | 0.826 | MSDial | positive |
| naringenin | 272.06816 | 17.829 | MSDial | negative |
| Nicotinamide | 122.04817 | 1.166 | MSDial | positive |
| O-Acetyl-L-homoserine | 161.06871 | 1.319 | MSDial | positive |
| oroxindin | 460.10071 | 14.389 | MSDial | negative |
| Pantothenic acid | 219.11067 | 6.366 | MSDial | positive |
| Pantothenic acid | 219.10994 | 6.347 | MSDial | negative |
| Phenylalanine | 165.07772 | 3.395 | MSDial | negative |
| Proline | 115.06354 | 0.813 | MSDial | positive |
| Saccharopine | 276.13198 | 1.114 | mzCloud | positive |
| Serine | 105.04259 | 0.773 | MSDial | positive |
| Sinapic acid | 224.06822 | 14.846 | MSDial | positive |
| Sinapic acid | 224.0679 | 14.86 | MSDial | negative |
| Spermidine | 145.02254 | 0.69 | MSDial | positive |
| Stachyose | 666.22039 | 0.846 | MSDial | negative |
| Succinic acid | 118.02503 | 1.479 | MSDial | negative |
| Taurocholate | 515.29112 | 17.951 | MSDial | positive |
| Threonic acid | 136.03614 | 0.786 | MSDial | negative |
| Threonine | 119.05662 | 0.781 | MSDial | negative |
| Thymine | 126.04305 | 1.92 | MSDial | positive |
| trans-4-Hydroxy-3-methoxycinnamate | 194.05691 | 13.058 | MSDial | negative |
| Tricin | 330.07354 | 18.027 | MSDial | positive |
| Tricin | 330.07395 | 18.053 | MSDial | negative |
| Tyrosine | 181.07381 | 1.326 | MSDial | positive |
| Urocanic acid | 138.04291 | 1.106 | mzCloud | positive |

**Supplemental Table S2.** Peptide Identifications from Aqueous Fraction. L stands for length and U for Unique.

| Peptide | U | -10lgP | Mass | L | ppm | m/z | z | RT | PTM | AScore | Protein Accession |
|---|---|---|---|---|---|---|---|---|---|---|---|
| D.LTPEQQADVISR.T | Y | 32.02 | 1355.7045 | 12 | -1.2 | 678.8587 | 2 | 53.15 | | | CUI\|CUI_0589 |
| K.ATGNTPDGR.R | Y | 25.41 | 887.4097 | 9 | -0.1 | 444.7121 | 2 | 5.22 | | | CUI\|CUI_0589 |
| R.ATDFVVPGPGK.V | Y | 23.85 | 1086.5709 | 11 | 2.3 | 544.294 | 2 | 67.85 | | | O88844\|IDHC_MOUSE |
| R.QNGVLNSWTDQDSK.D | Y | 38.37 | 1590.7274 | 14 | 0.6 | 796.3715 | 2 | 70.57 | | | P01837\|IGKC_MOUSE |
| I.SLLDTSNFNNQLDQTPR.V | Y | 38.61 | 1961.9442 | 17 | -2.8 | 981.9767 | 2 | 96.99 | | | P02858\|GLYG4_SOYBN |
| S.NFNNQLDQTPR.V | Y | 32.23 | 1345.6375 | 11 | 0.8 | 673.8265 | 2 | 47.6 | | | P02858\|GLYG4_SOYBN |
| E.DDEDEQIPSHPP.R | Y | 29.18 | 1377.5684 | 12 | -2 | 689.7901 | 2 | 48.15 | | | P02858\|GLYG4_SOYBN |
| D.E(-18.01)DDEDEQIPSHPP.R | Y | 28.12 | 1488.6005 | 13 | -3.4 | 745.305 | 2 | 55.32 | Pyro-glu from E | E1:Pyro-glu from E:1000.00 | P02858\|GLYG4_SOYBN |
| E.DEDDEDEQIPSHPP.R | Y | 25.72 | 1621.6379 | 14 | -1.1 | 811.8254 | 2 | 49.68 | | | P02858\|GLYG4_SOYBN |
| D.EDDEDEQIPSHPP.R | Y | 25.14 | 1506.611 | 13 | -1 | 754.312 | 2 | 48.62 | | | P02858\|GLYG4_SOYBN |
| D.E(-18.01)DEDEDKPRP.S | Y | 20.9 | 1210.5101 | 10 | -0.8 | 606.2618 | 2 | 13.71 | Pyro-glu from E | E1:Pyro-glu from E:1000.00 | P02858\|GLYG4_SOYBN |
| D.EDEQIPSHPP.R | Y | 19.54 | 1147.5145 | 10 | -1.4 | 574.7637 | 2 | 42.73 | | | P02858\|GLYG4_SOYBN |
| I.DTNSLENQLDQMPR.R | Y | 35.79 | 1659.7522 | 14 | 1.2 | 830.8843 | 2 | 96.87 | | | P04776\|GLYG1_SOYBN |
| V.SIIDTNSLENQLDQMPR.R | Y | 30.57 | 1972.9524 | 17 | 0.4 | 987.4839 | 2 | 97.02 | | | P04776\|GLYG1_SOYBN |
| R.HNIGQTSSPDIYNPQ.A | Y | 24.73 | 1669.7695 | 15 | -1.6 | 835.8907 | 2 | 58.37 | | | P04776\|GLYG1_SOYBN |
| L.SVIKPPTDEQQQRPQE.E | Y | 19.31 | 1878.9435 | 16 | -0.5 | 627.3215 | 3 | 40.14 | | | P04776\|GLYG1_SOYBN |
| N.SLENQLDQMPR.R | Y | 18.54 | 1329.6346 | 11 | 0.4 | 665.8249 | 2 | 85.3 | | | P04776\|GLYG1_SOYBN |
| I.NVLEGNEQFVDSAK.I | Y | 31.03 | 1548.7419 | 14 | -0.4 | 775.3779 | 2 | 74.93 | | | P07146\|TRY2_MOUSE |
| K.LASPVTLNAR.V | Y | 24.77 | 1040.5978 | 10 | -0.4 | 521.306 | 2 | 67.4 | | | P07146\|TRY2_MOUSE |
| R.TLGPANLPLAQR.Q | Y | 23.36 | 1249.7142 | 12 | 0.2 | 625.8645 | 2 | 96.87 | | | P09470\|ACE_MOUSE |
| K.LNGYTDAGDSWR.S | Y | 17.92 | 1353.595 | 12 | 1.4 | 677.8057 | 2 | 61.7 | | | P09470\|ACE_MOUSE |
| M.A(+42.01)SADWGYGSENGPDQWSK.L | Y | 47.89 | 1995.8235 | 18 | -4.4 | 998.9147 | 2 | 93.09 | Acetylation (N-term) | A1:Acetylation (N-term):1000.00 | P13634\|CAH1_MOUSE |
| A.DWGYGSENGPDQWSK.L | Y | 34.41 | 1724.7067 | 15 | -2.4 | 863.3585 | 2 | 82.96 | | | P13634\|CAH1_MOUSE |
| K.VGPANPSLQK.V | Y | 28.58 | 1009.5556 | 10 | 0.8 | 505.7855 | 2 | 27.78 | | | P13634\|CAH1_MOUSE |
| V.LSNHRPPQPL.K | Y | 25.24 | 1157.6305 | 10 | 0.2 | 579.8226 | 2 | 45.21 | | | P13634\|CAH1_MOUSE |
| K.GGPLADSYR.L | Y | 25.08 | 934.4508 | 9 | 0.4 | 468.2329 | 2 | 33.95 | | | P13634\|CAH1_MOUSE |
| K.TSEANHDSSLKPL.S | Y | 20.99 | 1397.6786 | 13 | -1.4 | 699.8456 | 2 | 37.79 | | | P13634\|CAH1_MOUSE |
| E.EDEDEQPR.P | Y | 25.04 | 1016.4047 | 8 | -0.1 | 509.2095 | 2 | 5.93 | | | P13916\|GLCA_SOYBN |
| E.EDEDEQPRPI.P | Y | 21.66 | 1226.5415 | 10 | 3.7 | 614.2803 | 2 | 27.85 | | | P13916\|GLCA_SOYBN |
| E.EDEDEQPRPIP.F | Y | 18.72 | 1323.5942 | 11 | 2.9 | 662.8063 | 2 | 39.02 | | | P13916\|GLCA_SOYBN |
| E.EDEDEQPRP.I | Y | 18.66 | 1113.4574 | 9 | -1.1 | 557.7354 | 2 | 12.12 | | | P13916\|GLCA_SOYBN |
| V.SVEVDSTPGVDLAK.I | Y | 27.49 | 1415.7144 | 14 | -1 | 708.8638 | 2 | 68.27 | | | P19001\|K1C19_MOUSE |
| R.ALEQANGELEVK.I | Y | 23.15 | 1299.667 | 12 | -3.6 | 650.8384 | 2 | 50.68 | | | P19001\|K1C19_MOUSE |
| M.S(+42.01)DKPDMAEIEK.F | Y | 27.86 | 1303.5966 | 11 | 2 | 652.8069 | 2 | 48.15 | Acetylation (N-term) | S1:Acetylation (N-term):1000.00 | P20065\|TYB4_MOUSE |
| M.S(+42.01)DKPDMAEIEKFDK.S | Y | 27.57 | 1693.7869 | 14 | -0.7 | 565.6025 | 3 | 96.89 | Acetylation (N-term) | S1:Acetylation (N-term):1000.00 | P20065\|TYB4_MOUSE |
| M.S(+42.01)DKPDMAEIEK.F | Y | 27.86 | 1303.5966 | 11 | 2 | 652.8069 | 2 | 48.15 | Acetylation (N-term) | S1:Acetylation (N-term):1000.00 | P20065-2\|TYB4_MOUSE |
| M.S(+42.01)DKPDMAEIEKFDK.S | Y | 27.57 | 1693.7869 | 14 | -0.7 | 565.6025 | 3 | 96.89 | Acetylation (N-term) | S1:Acetylation (N-term):1000.00 | P20065-2\|TYB4_MOUSE |
| R.VGAGAPVYLAAVLEYLTAEILELAGNAAR.D | Y | 37.8 | 2914.5803 | 29 | -5.9 | 1458.2888 | 2 | 98.26 | | | P22752\|H2A1_MOUSE |
| R.VGAGAPVYLAAVLEYLTAEILELAGNAAR.D | Y | 37.8 | 2914.5803 | 29 | -5.9 | 1458.2888 | 2 | 98.26 | | | P27661\|H2AX_MOUSE |
| M.S(+42.01)ETAPAAPAAPAPVEK.T | Y | 31.05 | 1547.7831 | 16 | -1.1 | 774.8979 | 2 | 50.77 | Acetylation (N-term) | S1:Acetylation (N-term):1000.00 | P43277\|H13_MOUSE |
| K.DSYVGDEAQSK.R | Y | 29.81 | 1197.5149 | 11 | 0.2 | 599.7648 | 2 | 16.24 | | | P60710\|ACTB_MOUSE |
| V.LLTEAPLNPK.A | Y | 24.93 | 1094.6335 | 10 | -1 | 548.3235 | 2 | 68.02 | | | P60710\|ACTB_MOUSE |
| K.EITALAPSTMK.I | Y | 24.32 | 1160.6111 | 11 | -1.6 | 581.3119 | 2 | 68.84 | | | P60710\|ACTB_MOUSE |
| K.AGFAGDDAPR.A | Y | 22.75 | 975.441 | 10 | -0.3 | 488.7276 | 2 | 25.54 | | | P60710\|ACTB_MOUSE |
| A.GFAGDDAPR.A | Y | 22.59 | 904.4039 | 9 | -1.7 | 453.2084 | 2 | 21.12 | | | P60710\|ACTB_MOUSE |
| V.LSGGTTMYPGIADR.M | Y | 20.55 | 1437.6921 | 14 | -2.3 | 719.8517 | 2 | 69.5 | | | P60710\|ACTB_MOUSE |
| K.MTQIMFETFNTPAMYVAIQAVLSLYASGR.T | Y | 18.53 | 3252.6021 | 29 | 4.8 | 1085.2131 | 3 | 116.11 | | | P60710\|ACTB_MOUSE |
| K.DSYVGDEAQSK.R | Y | 29.81 | 1197.5149 | 11 | 0.2 | 599.7648 | 2 | 16.24 | | | P63260\|ACTG_MOUSE |
| V.LLTEAPLNPK.A | Y | 24.93 | 1094.6335 | 10 | -1 | 548.3235 | 2 | 68.02 | | | P63260\|ACTG_MOUSE |
| K.EITALAPSTMK.I | Y | 24.32 | 1160.6111 | 11 | -1.6 | 581.3119 | 2 | 68.84 | | | P63260\|ACTG_MOUSE |
| K.AGFAGDDAPR.A | Y | 22.75 | 975.441 | 10 | -0.3 | 488.7276 | 2 | 25.54 | | | P63260\|ACTG_MOUSE |
| A.GFAGDDAPR.A | Y | 22.59 | 904.4039 | 9 | -1.7 | 453.2084 | 2 | 21.12 | | | P63260\|ACTG_MOUSE |
| V.LSGGTTMYPGIADR.M | Y | 20.55 | 1437.6921 | 14 | -2.3 | 719.8517 | 2 | 69.5 | | | P63260\|ACTG_MOUSE |
| K.MTQIMFETFNTPAMYVAIQAVLSLYASGR.T | Y | 18.53 | 3252.6021 | 29 | 4.8 | 1085.2131 | 3 | 116.11 | | | P63260\|ACTG_MOUSE |
| V.DALASSHPLSSPADEIK.T | Y | 30.43 | 1736.858 | 17 | -1.9 | 869.4346 | 2 | 73.92 | | | P97449\|AMPN_MOUSE |
| R.SEYMEGDVK.K | Y | 23.77 | 1056.4434 | 9 | -1.6 | 529.2281 | 2 | 26.97 | | | P97449\|AMPN_MOUSE |
| R.SEVYGPMK.R | Y | 18.94 | 909.4266 | 8 | -1 | 455.7201 | 2 | 30.85 | | | P97449\|AMPN_MOUSE |
| R.ALEQALEK.T | Y | 18.47 | 900.4916 | 8 | 0.5 | 451.2533 | 2 | 35.75 | | | P97449\|AMPN_MOUSE |
| R.TLDGTPAPNIDK.T | Y | 18.2 | 1240.6299 | 12 | 1.5 | 621.3231 | 2 | 47.61 | | | P97449\|AMPN_MOUSE |
| T.LLDQNPDIR.Q | Y | 21.24 | 1082.572 | 9 | 0.4 | 542.2935 | 2 | 47.9 | | | Q61847\|MEP1B_MOUSE |
| T.LLDQNPDIR.Q | Y | 21.24 | 1082.572 | 9 | 0.4 | 542.2935 | 2 | 47.9 | | | Q61847-2\|MEP1B_MOUSE |
| P.GIGRPPPPP.F | Y | 20.9 | 886.5024 | 9 | 0 | 444.2585 | 2 | 38.84 | | | Q61900\|SMR3A_MOUSE |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| G.PGIGRPPPPP.F | Y | 19.87 | 983.5552 | 10 | -0.3 | 492.7847 | 2 | 47.29 | | | Q61900\|SMR3A_MOUSE |
| R.LQEENQVITPR.L | Y | 28.85 | 1325.6938 | 11 | 1.1 | 663.8549 | 2 | 42.62 | | | Q62468\|VILI_MOUSE |
| V.NEFDSDVDVGDLQK.V | Y | 33.93 | 1579.7002 | 14 | -2.6 | 790.8553 | 2 | 81.73 | | | Q6P8U6\|LIPP_MOUSE |
| V.LVYTDPGDINDGK.S | Y | 28.62 | 1405.6725 | 13 | -0.7 | 703.843 | 2 | 58.08 | | | Q7M758\|NALDL_MOUSE |
| K.SSPDPLQVR.M | Y | 22.47 | 997.5192 | 9 | -1.1 | 499.7664 | 2 | 48.04 | | | Q7M758\|NALDL_MOUSE |
| V.YTDPGDINDGK.S | Y | 20.03 | 1193.52 | 11 | 0.7 | 597.7677 | 2 | 25.32 | | | Q7M758\|NALDL_MOUSE |
| R.VGAGAPVYLAAVLEYLT AEILELAGNAAR.D | Y | 37.8 | 2914.5803 | 29 | -5.9 | 1458.288 8 | 2 | 98.26 | | | Q8BFU2\|H2A3_MOUSE |
| R.VGAGAPVYLAAVLEYLT AEILELAGNAAR.D | Y | 37.8 | 2914.5803 | 29 | -5.9 | 1458.288 8 | 2 | 98.26 | | | Q8CGP5\|H2A1F_MOUSE |
| R.VGAGAPVYLAAVLEYLT AEILELAGNAAR.D | Y | 37.8 | 2914.5803 | 29 | -5.9 | 1458.288 8 | 2 | 98.26 | | | Q8CGP6\|H2A1H_MOUSE |
| R.VGAGAPVYLAAVLEYLT AEILELAGNAAR.D | Y | 37.8 | 2914.5803 | 29 | -5.9 | 1458.288 8 | 2 | 98.26 | | | Q8CGP7\|H2A1K_MOUSE |
| R.VGAGAPVYLAAVLEYLT AEILELAGNAAR.D | Y | 37.8 | 2914.5803 | 29 | -5.9 | 1458.288 8 | 2 | 98.26 | | | Q8R1M2\|H2AJ_MOUSE |
| R.QNGVLNSWTDQDSK.D | Y | 38.37 | 1590.7274 | 14 | 0.6 | 796.3715 | 2 | 70.57 | | | tr\|A0A075B5P2\|A0A075B5P2_ MOUSE |
| R.ATDFVVPGPGK.V | Y | 23.85 | 1086.5709 | 11 | 2.3 | 544.294 | 2 | 67.85 | | | tr\|A0A087WPT4\|A0A087WPT4 _MOUSE |
| R.ATDFVVPGPGK.V | Y | 23.85 | 1086.5709 | 11 | 2.3 | 544.294 | 2 | 67.85 | | | tr\|A0A087WRS9\|A0A087WRS 9_MOUSE |
| R.ASGHSYTK.Y | Y | 18.19 | 849.3981 | 8 | -1 | 425.7059 | 2 | 2.48 | | | tr\|B2RS76\|B2RS76_MOUSE |
| R.INTDQYYLAR.F | Y | 29.81 | 1255.6196 | 10 | -1.3 | 628.8163 | 2 | 67.04 | | | tr\|B5THE2\|B5THE2_MOUSE |
| K.ASTYSNAFPSTPVNK.L | Y | 29.48 | 1582.7627 | 15 | 0.7 | 792.3892 | 2 | 67.15 | | | tr\|B5THE2\|B5THE2_MOUSE |
| R.GVENDVFIR.Y | Y | 23.99 | 1047.5349 | 9 | 1.7 | 524.7756 | 2 | 75.34 | | | tr\|B5THE2\|B5THE2_MOUSE |
| Y.DVQYSDIDYMER.Q | Y | 23.28 | 1532.6453 | 12 | -1.4 | 767.3289 | 2 | 80.98 | | | tr\|B5THE2\|B5THE2_MOUSE |
| L.NIPSAPLSTPEGR.L | Y | 23.28 | 1337.6938 | 13 | -1.3 | 669.8533 | 2 | 64.19 | | | tr\|B5THE2\|B5THE2_MOUSE |
| R.NHNGQGYK.D | Y | 18.46 | 916.4151 | 8 | 0.2 | 459.2149 | 2 | 2.08 | | | tr\|B5THE2\|B5THE2_MOUSE |
| K.ELHNNGQK.L | Y | 17.84 | 938.457 | 8 | 2.2 | 470.2368 | 2 | 1.88 | | | tr\|B5THE2\|B5THE2_MOUSE |
| P.EPEPEPEPEPEPEPEPV.P | Y | 24.37 | 1925.8417 | 17 | 2.5 | 963.9305 | 2 | 67.32 | | | tr\|E1AZ71\|E1AZ71_MOUSE |
| P.EPEPEPEPEPEPVP.E | Y | 21.98 | 1570.7039 | 14 | -1 | 786.3584 | 2 | 59.5 | | | tr\|E1AZ71\|E1AZ71_MOUSE |
| P.EPEPEPEPEPEPEPVP.E | Y | 21.4 | 1796.7992 | 16 | -3.1 | 899.4041 | 2 | 65.89 | | | tr\|E1AZ71\|E1AZ71_MOUSE |
| P.EPEPEPEPEPEPVPEE.A | Y | 17.92 | 1828.7889 | 16 | -4.1 | 915.398 | 2 | 52.74 | | | tr\|E1AZ71\|E1AZ71_MOUSE |
| R.IIGVDINEEK.F | Y | 20.93 | 1128.6027 | 10 | -1.3 | 565.3079 | 2 | 68.99 | | | tr\|E9Q4Y1\|E9Q4Y1_MOUSE |
| R.IIGVDINEEK.F | Y | 20.93 | 1128.6027 | 10 | -1.3 | 565.3079 | 2 | 68.99 | | | tr\|E9Q5Z6\|E9Q5Z6_MOUSE |
| R.VDMYLPEDK.I | Y | 26.09 | 1108.511 | 9 | -1.2 | 555.2621 | 2 | 68.63 | | | tr\|F8VQM5\|F8VQM5_MOUSE |
| R.EIYDFYNEK.M | Y | 23.27 | 1219.5397 | 9 | 0.6 | 610.7775 | 2 | 79.57 | | | tr\|F8VQM5\|F8VQM5_MOUSE |
| K.TLPQFVEK.I | Y | 20.81 | 960.528 | 8 | 1.8 | 481.2722 | 2 | 69.15 | | | tr\|F8VQM5\|F8VQM5_MOUSE |
| E.EDEDEQPR.P | Y | 25.04 | 1016.4047 | 8 | -0.1 | 509.2095 | 2 | 5.93 | | | tr\|I1NGH2\|I1NGH2_SOYBN |
| E.EDEDEQPRPI.P | Y | 21.66 | 1226.5415 | 10 | 3.7 | 614.2803 | 2 | 27.85 | | | tr\|I1NGH2\|I1NGH2_SOYBN |
| E.EDEDEQPRPIP.F | Y | 18.72 | 1323.5942 | 11 | 2.9 | 662.8063 | 2 | 39.02 | | | tr\|I1NGH2\|I1NGH2_SOYBN |
| E.EDEDEQPRP.I | Y | 18.66 | 1113.4574 | 9 | -1.1 | 557.7354 | 2 | 12.12 | | | tr\|I1NGH2\|I1NGH2_SOYBN |
| R.IIGVDINEEK.F | Y | 20.93 | 1128.6027 | 10 | -1.3 | 565.3079 | 2 | 68.99 | | | tr\|Q3UQ40\|Q3UQ40_MOUSE |
| Q.AEPVPDESPR.S | Y | 23.59 | 1095.5197 | 10 | 1 | 548.7676 | 2 | 23.5 | | | tr\|Q3UW98\|Q3UW98_MOUSE |
| K.QAPAIYLR.D | N | 21.98 | 930.5287 | 8 | 0.5 | 466.2719 | 2 | 67.94 | | | tr\|Q3UW98\|Q3UW98_MOUSE |
| K.VAPPAFEGR.D | Y | 21.12 | 942.4922 | 9 | 2.4 | 472.2545 | 2 | 44.91 | | | tr\|Q3UW98\|Q3UW98_MOUSE |
| T.WTAPGDDYDVGR.V | N | 18.04 | 1350.584 | 12 | -1.2 | 676.2985 | 2 | 66.41 | | | tr\|Q3UW98\|Q3UW98_MOUSE |
| E.EDEDEQPR.P | Y | 25.04 | 1016.4047 | 8 | -0.1 | 509.2095 | 2 | 5.93 | | | tr\|Q94LX2\|Q94LX2_SOYBN |
| E.EDEDEQPRPI.P | Y | 21.66 | 1226.5415 | 10 | 3.7 | 614.2803 | 2 | 27.85 | | | tr\|Q94LX2\|Q94LX2_SOYBN |
| E.EDEDEQPRPIP.F | Y | 18.72 | 1323.5942 | 11 | 2.9 | 662.8063 | 2 | 39.02 | | | tr\|Q94LX2\|Q94LX2_SOYBN |
| E.EDEDEQPRP.I | Y | 18.66 | 1113.4574 | 9 | -1.1 | 557.7354 | 2 | 12.12 | | | tr\|Q94LX2\|Q94LX2_SOYBN |
| I.SLLDTSNFNNQLDQTPR.V | Y | 38.61 | 1961.9442 | 17 | -2.8 | 981.9767 | 2 | 96.99 | | | tr\|Q9SB11\|Q9SB11_SOYBN |
| S.NFNNQLDQTPR.V | Y | 32.23 | 1345.6375 | 11 | 0.8 | 673.8265 | 2 | 47.6 | | | tr\|Q9SB11\|Q9SB11_SOYBN |
| E.DDEDEQIPSHPP.R | Y | 29.18 | 1377.5684 | 12 | -2 | 689.7901 | 2 | 48.15 | | | tr\|Q9SB11\|Q9SB11_SOYBN |
| D.E(-18.01)DDEDEQIPSHPP.R | Y | 28.12 | 1488.6005 | 13 | -3.4 | 745.305 | 2 | 55.32 | Pyro-glu from E | E1:Pyro-glu from E:1000.00 | tr\|Q9SB11\|Q9SB11_SOYBN |
| E.DEDDEDEQIPSHPP.R | Y | 25.72 | 1621.6379 | 14 | -1.1 | 811.8254 | 2 | 49.68 | | | tr\|Q9SB11\|Q9SB11_SOYBN |
| D.EDDEDEQIPSHPP.R | Y | 25.14 | 1506.611 | 13 | -1 | 754.312 | 2 | 48.62 | | | tr\|Q9SB11\|Q9SB11_SOYBN |
| D.E(-18.01)DEDEDKPRP.S | Y | 20.9 | 1210.5101 | 10 | -0.8 | 606.2618 | 2 | 13.71 | Pyro-glu from E | E1:Pyro-glu from E:1000.00 | tr\|Q9SB11\|Q9SB11_SOYBN |
| D.EDEQIPSHPP.R | Y | 19.54 | 1147.5145 | 10 | -1.4 | 574.7637 | 2 | 42.73 | | | tr\|Q9SB11\|Q9SB11_SOYBN |

**Supplemental Table S3.** Peptide Identifications from Organic Fraction. L stands for length and U for Unique.

| z | U | -10lgP | Mass | L | ppm | m/z | z | RT | PTM | AScore | Protein Accession |
|---|---|---|---|---|---|---|---|---|---|---|---|
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | Amuc\|Amuc_1249 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACCELL\|BACCELL_00669 |
| S.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACCELL\|BACCELL_04197 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACCELL\|BACCELL_05063 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACEGG\|BACEGG_02321 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACEGG\|BACEGG_02376 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACFIN\|BACFIN_07158 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACINT\|BACINT_02306 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACINT\|BACINT_02479 |
| T.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACINT\|BACINT_02628 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACINT\|BACINT_03052 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACOVA\|BACOVA_00511 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | BACUNI\|BACUNI_00827 |
| R.SFDSIDNAPEEK.E | Y | 22.98 | 1350.5939 | 12 | 1 | 676.3049 | 2 | 45.78 | | | BACUNI\|BACUNI_00827 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACUNI\|BACUNI_02103 |
| S.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BACUNI\|BACUNI_02154 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | BLAHAN\|BLAHAN_06076 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | BLAHAN\|BLAHAN_06076 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | BXY\|BXY_03170 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | CLOHIR\|CLOHIR_02278 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | CLOHIR\|CLOHIR_02278 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | CLOHYLEM\|CLOHYLEM_05830 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | CLOHYLEM\|CLOHYLEM_05830 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | CLONEX\|CLONEX_00924 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | CLONEX\|CLONEX_00924 |
| D.LTPEQQADVISR.T | Y | 30.83 | 1355.7045 | 12 | 0.1 | 678.8596 | 2 | 53.01 | | | CUI\|CUI_0589 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | CUU\|CUU_2247 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | CUU\|CUU_3020 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | CUY\|CUY_3315 |
| R.SFDSIDNAPEEK.E | Y | 22.98 | 1350.5939 | 12 | 1 | 676.3049 | 2 | 45.78 | | | CUY\|CUY_3315 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | CUY\|CUY_4852 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | CW1\|CW1_3301 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | CW1\|CW1_3447 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | CW3\|CW3_3684 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | CW3\|CW3_3735 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | DORFOR\|DORFOR_01085 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | DORFOR\|DORFOR_01085 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | ECD227\|ECD227_01969 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | ENTCAN\|ENTCAN_06841 |
| R.AADDAAGLSISEK.M | Y | 21.86 | 1246.604 | 13 | 1.9 | 624.3105 | 2 | 45.99 | | | EUBELI\|EUBELI_00264 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | EUBELI\|EUBELI_00288 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | EUBELI\|EUBELI_00288 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | HMPREF1057\|HMPREF1057_00715 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | HMPREF1057\|HMPREF1057_02459 |
| R.SFDSIDNAPEEK.E | Y | 22.98 | 1350.5939 | 12 | 1 | 676.3049 | 2 | 45.78 | | | HMPREF1057\|HMPREF1057_02459 |
| R.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | HMPREF1059\|HMPREF1059_00643 |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | HMPREF1069\|HMPREF1069_01153 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | HMPREF9457\|HMPREF9457_01089 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | HMPREF9457\|HMPREF9457_01089 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | HMPREF9473\|HMPREF9473_04269 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | HMPREF9473\|HMPREF9473_04269 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | HMPREF9474\|HMPREF9474_00175 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | HMPREF9474\|HMPREF9474_00175 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | HMPREF9475\|HMPREF9475_00654 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | HMPREF9475\|HMPREF9475_00654 |
| K.LAPEFAK.R | Y | 22.97 | 774.4276 | 7 | -0.3 | 388.2209 | 2 | 37.44 | | | O08709\|PRDX6_MOUSE |
| R.IPYTPGEIPK.F | Y | 27.35 | 1113.6069 | 10 | 1.8 | 557.8118 | 2 | 68.18 | | | O70475\|UGDH_MOUSE |
| R.VLIGGDETPEGQK.A | Y | 24.93 | 1341.6776 | 13 | -0.9 | 671.8455 | 2 | 43.22 | | | O70475\|UGDH_MOUSE |
| L.M(+15.99)IGAPDMAFPR.M | Y | 23.61 | 1220.5681 | 11 | -0.9 | 611.2908 | 2 | 96.83 | Oxidation (M) | M1:Oxidation (M):90.32 | P00397\|COX1_MOUSE |
| I.GAPDMAFPR.M | Y | 21.56 | 960.4487 | 9 | -1 | 481.2312 | 2 | 65.65 | | | P00397\|COX1_MOUSE |
| K.DFPIANGDR.Q | Y | 22.78 | 1003.4723 | 9 | -0.3 | 502.7433 | 2 | 47.57 | | | P00920\|CAH2_MOUSE |
| K.IGGHGAEYGAEALER.M | Y | 36.3 | 1528.7269 | 15 | 0 | 765.3707 | 2 | 57.64 | | | P01942\|HBA_MOUSE |
| V.GHGYHTIEDEALYNR.L | Y | 29.69 | 1773.807 | 15 | -1.7 | 592.2753 | 3 | 96.73 | | | P03958\|ADA_MOUSE |
| K.STLDTDYQMTK.K | Y | 28.44 | 1301.5809 | 11 | 0.1 | 651.7978 | 2 | 46.37 | | | P03958\|ADA_MOUSE |
| K.EGVVYVEVR.Y | Y | 22.81 | 1048.5553 | 9 | 1 | 525.2855 | 2 | 65.04 | | | P03958\|ADA_MOUSE |
| R.TVHAGEVGSPEVV.R | Y | 22.46 | 1279.6407 | 13 | -0.4 | 640.8274 | 2 | 56.25 | | | P03958\|ADA_MOUSE |
| M.A(+42.01)QTPAFNKPK.V | Y | 22.01 | 1142.6084 | 10 | 0.7 | 572.3119 | 2 | 35.72 | Acetylation (N-term) | A1:Acetylation (N-term):1000.00 | P03958\|ADA_MOUSE |
| R.HNIGQTSSPDIYNPQ.A | Y | 35.15 | 1669.7695 | 15 | -0.9 | 835.8913 | 2 | 56.79 | | | P04776\|GLYG1_SOYBN |
| V.SIIDTNSLENQLDQMPR.R | Y | 34.68 | 1972.9524 | 17 | -1.2 | 987.4823 | 2 | 96.95 | | | P04776\|GLYG1_SOYBN |
| R.RPSYTNGPQEIYIQQGK.G | Y | 27.38 | 1977.9908 | 17 | -0.1 | 660.3375 | 3 | 79.49 | | | P04776\|GLYG1_SOYBN |
| R.SQSDNFEYVSFK.T | Y | 25.82 | 1449.6412 | 12 | 0.7 | 725.8284 | 2 | 96.57 | | | P04776\|GLYG1_SOYBN |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| R.PSYTNGPQEIYIQQGK.G | Y | 24.37 | 1821.8896 | 16 | -0.3 | 911.9518 | 2 | 69.28 | | | P04776\|GLYG1_SOYBN |
| R.RFYLAGNQEQE.F | Y | 24.22 | 1353.6313 | 11 | -1.9 | 677.8217 | 2 | 59.54 | | | P04776\|GLYG1_SOYBN |
| R.HNIGQTSSPDIYNPQAGSVTT.A | Y | 24.19 | 2186.0239 | 21 | -0.1 | 1094.0192 | 2 | 71.92 | | | P04776\|GLYG1_SOYBN |
| V.SIIDTNSLENQLDQM(+15.99)PR.R | Y | 23.11 | 1988.9473 | 17 | 1.4 | 995.4823 | 2 | 96.89 | Oxidation (M) | M15:Oxidation (M):1000.00 | P04776\|GLYG1_SOYBN |
| R.HNIGQTSSPDIYNPQAGSV.T | Y | 22.04 | 1983.9286 | 19 | -0.4 | 992.9711 | 2 | 76.07 | | | P04776\|GLYG1_SOYBN |
| I.DTNSLENQLDQM(+15.99)PR.R | Y | 21.79 | 1675.7472 | 14 | 0.1 | 838.881 | 2 | 63.58 | Oxidation (M) | M12:Oxidation (M):1000.00 | P04776\|GLYG1_SOYBN |
| K.GILAADESTGSIAK.R | Y | 21.92 | 1331.6932 | 14 | 0.7 | 666.8544 | 2 | 64.62 | | | P05064\|ALDOA_MOUSE |
| I.NVLEGNEQFVDSAK.I | Y | 37.27 | 1548.7419 | 14 | -1.3 | 775.3773 | 2 | 72.21 | | | P07146\|TRY2_MOUSE |
| K.TPAQYDASELK.A | Y | 20.82 | 1221.5876 | 11 | 0.8 | 611.8016 | 2 | 37.65 | | | P07356\|ANXA2_MOUSE |
| R.TLGPANLPLAQR.Q | Y | 29.33 | 1249.7142 | 12 | 0 | 625.8644 | 2 | 96.24 | | | P09470\|ACE_MOUSE |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | P0CG89\|H4_SOYBN |
| M.A(+42.01)SADWGYGSENGPDQWSK.L | Y | 47.61 | 1995.8235 | 18 | 2.3 | 998.9213 | 2 | 91.33 | Acetylation (N-term) | A1:Acetylation (N-term):1000.00 | P13634\|CAH1_MOUSE |
| K.TSEANHDSSLKPL.S | Y | 32.31 | 1397.6786 | 13 | 1.1 | 699.8474 | 2 | 35.57 | | | P13634\|CAH1_MOUSE |
| K.LYPIANGNNQSPIDIK.T | Y | 32.09 | 1755.9155 | 16 | 1 | 878.9659 | 2 | 96.88 | | | P13634\|CAH1_MOUSE |
| V.LSNHRPPQPL.K | Y | 30.15 | 1157.6305 | 10 | 2.6 | 579.824 | 2 | 35.71 | | | P13634\|CAH1_MOUSE |
| K.VGPANPSLQK.V | Y | 28.07 | 1009.5556 | 10 | -0.2 | 505.785 | 2 | 26.96 | | | P13634\|CAH1_MOUSE |
| A.NHDSSLKPL.S | Y | 22.29 | 1009.5192 | 9 | 0.5 | 505.7672 | 2 | 31.06 | | | P13634\|CAH1_MOUSE |
| K.KPADITDVVKPI.D | Y | 28.57 | 1294.7496 | 12 | 0 | 648.3821 | 2 | 96.96 | | | P15947\|KLK1_MOUSE |
| K.NNFLEDEPSAQHR.L | Y | 27.14 | 1555.7015 | 13 | 0.3 | 778.8583 | 2 | 54.52 | | | P15947\|KLK1_MOUSE |
| R.QYDISNPQKPR.L | Y | 26.96 | 1344.6786 | 11 | -1.5 | 673.3456 | 2 | 28.92 | | | P17563\|SBP1_MOUSE |
| K.VIEASEIQAK.C | Y | 23.46 | 1086.592 | 10 | 2.9 | 544.3049 | 2 | 37.37 | | | P17563\|SBP1_MOUSE |
| R.LIEDDFHNLR.L | Y | 26.91 | 1270.6306 | 10 | 1.5 | 636.3235 | 2 | 91.25 | | | P19467\|MUC13_MOUSE |
| R.TGVPSQTPNPYANQR.S | Y | 21.85 | 1628.7906 | 15 | -1.5 | 815.4014 | 2 | 41.48 | | | P19467\|MUC13_MOUSE |
| R.VGAGAPVYLAAVLEYLTAEILELAGNAAR.D | Y | 22.37 | 2914.5803 | 29 | 3 | 972.5369 | 3 | 116.29 | | | P22752\|H2A1_MOUSE |
| R.VGAGAPVYLAAVLEYLTAEILELAGNAAR.D | Y | 22.37 | 2914.5803 | 29 | 3 | 972.5369 | 3 | 116.29 | | | P27661\|H2AX_MOUSE |
| K.DSYVGDEAQSK.R | Y | 28.82 | 1197.5149 | 11 | 1.3 | 599.7655 | 2 | 16.94 | | | P62737\|ACTA_MOUSE |
| T.LLTEAPLNPK.A | Y | 25.21 | 1094.6335 | 10 | 0.2 | 548.3242 | 2 | 65.66 | | | P62737\|ACTA_MOUSE |
| K.EITALAPSTMK.I | Y | 22.53 | 1160.6111 | 11 | -1.1 | 581.3122 | 2 | 66.46 | | | P62737\|ACTA_MOUSE |
| A.GFAGDDAPR.A | Y | 22.19 | 904.4039 | 9 | 2 | 453.2101 | 2 | 21.01 | | | P62737\|ACTA_MOUSE |
| R.AVFPSIVGRPR.H | Y | 21.87 | 1197.6981 | 11 | -1.2 | 400.2395 | 3 | 96.99 | | | P62737\|ACTA_MOUSE |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | P62806\|H4_MOUSE |
| K.DSYVGDEAQSK.R | Y | 28.82 | 1197.5149 | 11 | 1.3 | 599.7655 | 2 | 16.94 | | | P63268\|ACTH_MOUSE |
| T.LLTEAPLNPK.A | Y | 25.21 | 1094.6335 | 10 | 0.2 | 548.3242 | 2 | 65.66 | | | P63268\|ACTH_MOUSE |
| K.EITALAPSTMK.I | Y | 22.53 | 1160.6111 | 11 | -1.1 | 581.3122 | 2 | 66.46 | | | P63268\|ACTH_MOUSE |
| A.GFAGDDAPR.A | Y | 22.19 | 904.4039 | 9 | 2 | 453.2101 | 2 | 21.01 | | | P63268\|ACTH_MOUSE |
| R.AVFPSIVGRPR.H | Y | 21.87 | 1197.6981 | 11 | -1.2 | 400.2395 | 3 | 96.99 | | | P63268\|ACTH_MOUSE |
| K.DSYVGDEAQSK.R | Y | 28.82 | 1197.5149 | 11 | 1.3 | 599.7655 | 2 | 16.94 | | | P68033\|ACTC_MOUSE |
| T.LLTEAPLNPK.A | Y | 25.21 | 1094.6335 | 10 | 0.2 | 548.3242 | 2 | 65.66 | | | P68033\|ACTC_MOUSE |
| K.EITALAPSTMK.I | Y | 22.53 | 1160.6111 | 11 | -1.1 | 581.3122 | 2 | 66.46 | | | P68033\|ACTC_MOUSE |
| A.GFAGDDAPR.A | Y | 22.19 | 904.4039 | 9 | 2 | 453.2101 | 2 | 21.01 | | | P68033\|ACTC_MOUSE |
| R.AVFPSIVGRPR.H | Y | 21.87 | 1197.6981 | 11 | -1.2 | 400.2395 | 3 | 96.99 | | | P68033\|ACTC_MOUSE |
| K.DSYVGDEAQSK.R | Y | 28.82 | 1197.5149 | 11 | 1.3 | 599.7655 | 2 | 16.94 | | | P68134\|ACTS_MOUSE |
| T.LLTEAPLNPK.A | Y | 25.21 | 1094.6335 | 10 | 0.2 | 548.3242 | 2 | 65.66 | | | P68134\|ACTS_MOUSE |
| K.EITALAPSTMK.I | Y | 22.53 | 1160.6111 | 11 | -1.1 | 581.3122 | 2 | 66.46 | | | P68134\|ACTS_MOUSE |
| A.GFAGDDAPR.A | Y | 22.19 | 904.4039 | 9 | 2 | 453.2101 | 2 | 21.01 | | | P68134\|ACTS_MOUSE |
| R.AVFPSIVGRPR.H | Y | 21.87 | 1197.6981 | 11 | -1.2 | 400.2395 | 3 | 96.99 | | | P68134\|ACTS_MOUSE |
| I.TEDDIIQNQSVLGK.Y | Y | 24.79 | 1558.7838 | 14 | 3.3 | 780.4017 | 2 | 73.2 | | | P70412\|CUZD1_MOUSE |
| K.AASGFNATEDAQTLRK.A | Y | 29.82 | 1678.8274 | 16 | -0.1 | 840.4209 | 2 | 54.76 | | | P97429\|ANXA4_MOUSE |
| K.AASGFNATEDAQTLR.K | Y | 28.85 | 1550.7324 | 15 | -1.4 | 776.3724 | 2 | 61.75 | | | P97429\|ANXA4_MOUSE |
| K.GLGTDDNTLIR.V | Y | 21.3 | 1173.599 | 11 | 1.9 | 587.8079 | 2 | 64.84 | | | P97429\|ANXA4_MOUSE |
| K.VVATTQMQAADAR.K | Y | 27.37 | 1360.6769 | 13 | 0.1 | 681.3458 | 2 | 32.6 | | | P97449\|AMPN_MOUSE |
| K.ADNSATGFGTGTR.A | Y | 23.27 | 1253.5636 | 13 | 0.8 | 627.7896 | 2 | 23.21 | | | P97449\|AMPN_MOUSE |
| R.SEYMEGDVKK.V | Y | 21.37 | 1184.5383 | 10 | 1.6 | 593.2774 | 2 | 18.74 | | | P97449\|AMPN_MOUSE |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | PARMER\|PARMER_00510 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | PARMER\|PARMER_01340 |
| R.SFDSIDNAPEEK.E | Y | 22.98 | 1350.5939 | 12 | 1 | 676.3049 | 2 | 45.78 | | | PARMER\|PARMER_01340 |
| T.LLDQNPDIRQ.R | Y | 25.54 | 1210.6306 | 10 | -1 | 606.322 | 2 | 44.74 | | | Q61847\|MEP1B_MOUSE |
| R.ISEFEDVIGQR.M | Y | 25.13 | 1291.6407 | 11 | -1.5 | 646.8267 | 2 | 96.54 | | | Q61847\|MEP1B_MOUSE |
| T.LLDQNPDIRQ.R | Y | 25.54 | 1210.6306 | 10 | -1 | 606.322 | 2 | 44.74 | | | Q61847-2\|MEP1B_MOUSE |
| R.ISEFEDVIGQR.M | Y | 25.13 | 1291.6407 | 11 | -1.5 | 646.8267 | 2 | 96.54 | | | Q61847-2\|MEP1B_MOUSE |
| F.GPGIGRPPPPP.F | Y | 22.84 | 1040.5767 | 11 | 3 | 521.2972 | 2 | 49.04 | | | Q61900\|SMR3A_MOUSE |
| P.GIGRPPPPP.F | Y | 21.43 | 886.5024 | 9 | -2 | 444.2576 | 2 | 37.21 | | | Q61900\|SMR3A_MOUSE |
| R.LQEENQVITPR.L | Y | 22.54 | 1325.6938 | 11 | 0 | 663.8542 | 2 | 41.63 | | | Q62468\|VILI_MOUSE |
| R.QYDISNPQKPR.L | Y | 26.96 | 1344.6786 | 11 | -1.5 | 673.3456 | 2 | 28.92 | | | Q63836\|SBP2_MOUSE |
| K.VIEASEIQAK.C | Y | 23.46 | 1086.592 | 10 | 2.9 | 544.3049 | 2 | 37.37 | | | Q63836\|SBP2_MOUSE |
| K.INFYM(+15.99)YQGPAPR.I | Y | 29.58 | 1471.6918 | 12 | -2.3 | 736.8515 | 2 | 93.88 | Oxidation (M) | M5:Oxidation (M):1000.00 | Q6DYE8\|ENPP3_MOUSE |
| V.NEFDSDVDVGDLQKVK.F | Y | 37.31 | 1806.8635 | 16 | 6.9 | 904.4453 | 2 | 96.77 | | | Q6P8U6\|LIPP_MOUSE |
| R.TTYTQATQNVR.V | Y | 32.65 | 1281.6313 | 11 | -1.2 | 641.8222 | 2 | 26.39 | | | Q6P8U6\|LIPP_MOUSE |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| V.RLDPTDAQFV.D | Y | 25.32 | 1160.5825 | 10 | 1 | 581.2991 | 2 | 85.74 | | | Q6P8U6\|LIPP_MOUSE |
| K.TAGSQEAPNLPR.K | Y | 25.06 | 1239.6207 | 12 | -1.2 | 620.8169 | 2 | 35.67 | | | Q6T3U4\|NPCL1_MOUSE |
| K.AQEINSGSEAWAEVQR.Q | Y | 30.97 | 1773.8281 | 16 | 2.4 | 887.9235 | 2 | 72.34 | | | Q7M758\|NALDL_MOUSE |
| I.NSGSEAWAEVQR.Q | Y | 30.68 | 1332.6058 | 12 | -1.1 | 667.3094 | 2 | 52.83 | | | Q7M758\|NALDL_MOUSE |
| K.FLDPGFSSHQAVAR.T | Y | 29.59 | 1530.7578 | 14 | 0.5 | 511.2601 | 3 | 91.01 | | | Q7M758\|NALDL_MOUSE |
| A.SVDTFPGLANAYAK.A | Y | 27.46 | 1452.7249 | 14 | 0.5 | 727.3701 | 2 | 96.94 | | | Q7M758\|NALDL_MOUSE |
| V.LVYTDPGDINDGK.S | Y | 24.4 | 1405.6725 | 13 | 0.6 | 703.8439 | 2 | 57.79 | | | Q7M758\|NALDL_MOUSE |
| K.SSPDPLQVR.M | Y | 23.89 | 997.5192 | 9 | 2.3 | 499.768 | 2 | 45.4 | | | Q7M758\|NALDL_MOUSE |
| A.ILDSVM(+15.99)GQLDASR.I | Y | 23.36 | 1419.7028 | 13 | 1.7 | 710.8599 | 2 | 68.84 | Oxidation (M) | M6:Oxidation (M):1000.00 | Q7M758\|NALDL_MOUSE |
| T.SEPAPDKEELDQLAK.S | Y | 31.5 | 1668.8206 | 15 | 2.2 | 835.4194 | 2 | 65.65 | | | Q7TPZ8\|CBPA1_MOUSE |
| R.VGAGAPVYLAAVLEYLTAEILELAGNAAR.D | Y | 22.37 | 2914.5803 | 29 | 3 | 972.5369 | 3 | 116.29 | | | Q8BFU2\|H2A3_MOUSE |
| R.VGAGAPVYLAAVLEYLTAEILELAGNAAR.D | Y | 22.37 | 2914.5803 | 29 | 3 | 972.5369 | 3 | 116.29 | | | Q8CGP5\|H2A1F_MOUSE |
| R.VGAGAPVYLAAVLEYLTAEILELAGNAAR.D | Y | 22.37 | 2914.5803 | 29 | 3 | 972.5369 | 3 | 116.29 | | | Q8CGP6\|H2A1H_MOUSE |
| R.VGAGAPVYLAAVLEYLTAEILELAGNAAR.D | Y | 22.37 | 2914.5803 | 29 | 3 | 972.5369 | 3 | 116.29 | | | Q8CGP7\|H2A1K_MOUSE |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | Q8K0C9\|GMDS_MOUSE |
| V.MTGPPVFNPR.V | Y | 21.23 | 1114.5593 | 10 | 2.6 | 558.2884 | 2 | 71.4 | | | Q8K419\|LEG4_MOUSE |
| V.GEIMSLSAATPK.H | Y | 22.56 | 1203.6169 | 12 | 1.7 | 602.8168 | 2 | 74.23 | | | Q8R0I0\|ACE2_MOUSE |
| R.VGAGAPVYLAAVLEYLTAEILELAGNAAR.D | Y | 22.37 | 2914.5803 | 29 | 3 | 972.5369 | 3 | 116.29 | | | Q8R1M2\|H2AJ_MOUSE |
| R.NPSDNGIGAPTSGDGPAGPYTR.Q | Y | 26.19 | 2099.9507 | 22 | -1.5 | 1050.9811 | 2 | 59.65 | | | Q91XA9\|CHIA_MOUSE |
| K.ANEWLGYDNIK.S | Y | 23.22 | 1321.6302 | 11 | 1.6 | 661.8235 | 2 | 96.86 | | | Q91XA9\|CHIA_MOUSE |
| F.ILRNPSDNGIGAPT.S | Y | 22.32 | 1423.7419 | 14 | 0.2 | 712.8784 | 2 | 55.64 | | | Q91XA9\|CHIA_MOUSE |
| I.VGFPEYGHT.F | Y | 21.07 | 1005.4556 | 9 | -0.1 | 503.735 | 2 | 64.45 | | | Q91XA9\|CHIA_MOUSE |
| I.NAGDLFVHPK.W | Y | 28.88 | 1096.5665 | 10 | 1.1 | 549.2911 | 2 | 64.1 | | | Q9CQ52\|CEL3B_MOUSE |
| A.GEFDQGSDEENVQVLK.I | Y | 31.11 | 1792.8115 | 16 | 5.2 | 897.4177 | 2 | 70.61 | | | Q9CR35\|CTRB1_MOUSE |
| K.IAQVFKNPK.F | Y | 23.91 | 1043.6127 | 9 | 1.1 | 522.8142 | 2 | 35.77 | | | Q9CR35\|CTRB1_MOUSE |
| M.A(+42.01)QAPTVIVTQPGFVR.A | Y | 36.96 | 1624.8937 | 15 | 1.8 | 813.4556 | 2 | 96.92 | Acetylation (N-term) | A1:Acetylation (N-term):1000.00 | Q9JI48\|PLAC8_MOUSE |
| M.T(+42.01)ASSVLLHTGQK.M | Y | 22.89 | 1282.6881 | 12 | 0.5 | 642.3516 | 2 | 66.71 | Acetylation (N-term) | T1:Acetylation (N-term):1000.00 | Q9JII6\|AK1A1_MOUSE |
| R.LTAFEEAIPK.S | Y | 26.03 | 1117.6019 | 10 | 1.2 | 559.8089 | 2 | 96.89 | | | Q9Z2W0\|DNPEP_MOUSE |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | RUMGNA\|RUMGNA_01593 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | RUMGNA\|RUMGNA_01593 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | RUMHYD\|RUMHYD_01594 |
| K.TTLTAAITK.T | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | RUMHYD\|RUMHYD_01594 |
| I.GHVDHGK.T | N | 23.23 | 748.3616 | 7 | -0.9 | 375.1877 | 2 | 5.05 | | | RUMLAC\|RUMLAC_02673 |
| K.TTLTAAITK.V | Y | 22.54 | 918.5386 | 9 | -0.6 | 460.2763 | 2 | 47.32 | | | RUMLAC\|RUMLAC_02673 |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|A0A096UJU8\|A0A096UJU8_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|A0A096UWG0\|A0A096UWG0_WHEAT |
| K.ETTPFYPR.S | Y | 23.31 | 1009.4869 | 8 | 0.5 | 505.7509 | 2 | 50.83 | | | tr\|A0A0F6B3F9\|A0A0F6B3F9_SALT1 |
| K.GILAADESTGSIAK.R | Y | 21.92 | 1331.6932 | 14 | 0.7 | 666.8544 | 2 | 64.62 | | | tr\|A6ZI44\|A6ZI44_MOUSE |
| K.GILAADESTGSIAK.R | Y | 21.92 | 1331.6932 | 14 | 0.7 | 666.8544 | 2 | 64.62 | | | tr\|A6ZI47\|A6ZI47_MOUSE |
| K.VNADEVGGEALGR.L | Y | 25.74 | 1285.6262 | 13 | 0.5 | 643.8207 | 2 | 50.82 | | | tr\|A8DUK4\|A8DUK4_MOUSE |
| K.AAVSGLWGK.V | Y | 22.4 | 887.4865 | 9 | 0.7 | 444.7508 | 2 | 67.16 | | | tr\|A8DUK4\|A8DUK4_MOUSE |
| K.TPAQYDASELK.A | Y | 20.82 | 1221.5876 | 11 | 0.8 | 611.8016 | 2 | 37.65 | | | tr\|B0V2N7\|B0V2N7_MOUSE |
| K.TPAQYDASELK.A | Y | 20.82 | 1221.5876 | 11 | 0.8 | 611.8016 | 2 | 37.65 | | | tr\|B0V2N8\|B0V2N8_MOUSE |
| M.T(+42.01)ASSVLLHTGQK.M | Y | 22.89 | 1282.6881 | 12 | 0.5 | 642.3516 | 2 | 66.71 | Acetylation (N-term) | T1:Acetylation (N-term):1000.00 | tr\|B1AXW3\|B1AXW3_MOUSE |
| V.VNIDGYVYTWAK.D | Y | 26.94 | 1427.7085 | 12 | 0.2 | 714.8617 | 2 | 97 | | | tr\|B2RS76\|B2RS76_MOUSE |
| V.ANDNPDLVSKR.V | Y | 26.66 | 1227.6207 | 11 | 1.3 | 614.8184 | 2 | 19.27 | | | tr\|B2RS76\|B2RS76_MOUSE |
| I.TEDDIIQNQSVLGK.Y | Y | 24.79 | 1558.7838 | 14 | 3.3 | 780.4017 | 2 | 73.2 | | | tr\|B2RU69\|B2RU69_MOUSE |
| Y.SNAFPSTPVNK.L | Y | 26.67 | 1160.5825 | 11 | 1.8 | 581.2996 | 2 | 44.07 | | | tr\|B5THE2\|B5THE2_MOUSE |
| K.ELHNNGQK.L | Y | 22.86 | 938.457 | 8 | 3.3 | 470.2373 | 2 | 5.77 | | | tr\|B5THE2\|B5THE2_MOUSE |
| L.NIPSAPLSTPEGR.L | Y | 22.1 | 1337.6938 | 13 | 0.3 | 669.8544 | 2 | 62.82 | | | tr\|B5THE2\|B5THE2_MOUSE |
| K.GILAADESTGSIAK.R | Y | 21.92 | 1331.6932 | 14 | 0.7 | 666.8544 | 2 | 64.62 | | | tr\|D3YV98\|D3YV98_MOUSE |
| K.GILAADESTGSIAK.R | Y | 21.92 | 1331.6932 | 14 | 0.7 | 666.8544 | 2 | 64.62 | | | tr\|D3YWI1\|D3YWI1_MOUSE |
| K.AASGFNATEDAQTLRK.A | Y | 29.82 | 1678.8274 | 16 | -0.1 | 840.4209 | 2 | 54.76 | | | tr\|D3Z0S1\|D3Z0S1_MOUSE |
| K.AASGFNATEDAQTLR.K | Y | 28.85 | 1550.7324 | 15 | -1.4 | 776.3724 | 2 | 61.75 | | | tr\|D3Z0S1\|D3Z0S1_MOUSE |
| K.GLGTDDNTLIR.V | Y | 21.3 | 1173.599 | 11 | 1.9 | 587.8079 | 2 | 64.84 | | | tr\|D3Z0S1\|D3Z0S1_MOUSE |
| K.LAPEFAK.R | Y | 22.97 | 774.4276 | 7 | -0.3 | 388.2209 | 2 | 37.44 | | | tr\|D3Z0Y2\|D3Z0Y2_MOUSE |
| R.IPYTPGEIPK.F | Y | 27.35 | 1113.6069 | 10 | 1.8 | 557.8118 | 2 | 68.18 | | | tr\|D3Z3F7\|D3Z3F7_MOUSE |
| R.VLIGGDETPEGQK.A | Y | 24.93 | 1341.6776 | 13 | -0.9 | 671.8455 | 2 | 43.22 | | | tr\|D3Z3F7\|D3Z3F7_MOUSE |
| K.GILAADESTGSIAK.R | Y | 21.92 | 1331.6932 | 14 | 0.7 | 666.8544 | 2 | 64.62 | | | tr\|D3Z510\|D3Z510_MOUSE |
| R.IVSGQDAQLGR.W | Y | 23.58 | 1142.6044 | 11 | 1.6 | 572.3104 | 2 | 29.29 | | | tr\|D6RGT5\|D6RGT5_MOUSE |
| P.EPEPEPEPEPEPVPEE.A | Y | 26.09 | 1828.7889 | 16 | -3.9 | 915.3982 | 2 | 53.34 | | | tr\|E1AZ71\|E1AZ71_MOUSE |
| E.PEPEPEPEPEPEPEPV.P | Y | 21.57 | 1796.7992 | 16 | 1.3 | 899.4081 | 2 | 65.07 | | | tr\|E1AZ71\|E1AZ71_MOUSE |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| K.VNADEVGGEALGR.L | Y | 25.74 | 1285.6262 | 13 | 0.5 | 643.8207 | 2 | 50.82 | | | tr\|E9Q223\|E9Q223_MOUSE |
| K.AAVSGLWGK.V | Y | 22.4 | 887.4865 | 9 | 0.7 | 444.7508 | 2 | 67.16 | | | tr\|E9Q223\|E9Q223_MOUSE |
| R.IVSGQDAQLGR.W | Y | 23.58 | 1142.6044 | 11 | 1.6 | 572.3104 | 2 | 29.29 | | | tr\|E9Q409\|E9Q409_MOUSE |
| I.AGLAPSDAQLSLEAPK.I | Y | 32.21 | 1566.8253 | 16 | -0.3 | 784.4197 | 2 | 94.49 | | | tr\|E9Q4Y1\|E9Q4Y1_MOUSE |
| I.GVDINEEKFPR.A | Y | 21.13 | 1302.6567 | 11 | -0.7 | 435.2259 | 3 | 64.2 | | | tr\|E9Q4Y1\|E9Q4Y1_MOUSE |
| I.AGLAPSDAQLSLEAPK.I | Y | 32.21 | 1566.8253 | 16 | -0.3 | 784.4197 | 2 | 94.49 | | | tr\|E9Q5Z6\|E9Q5Z6_MOUSE |
| I.GVDINEEKFPR.A | Y | 21.13 | 1302.6567 | 11 | -0.7 | 435.2259 | 3 | 64.2 | | | tr\|E9Q5Z6\|E9Q5Z6_MOUSE |
| AASGFNATEDAQTLRK.A | Y | 29.82 | 1678.8274 | 16 | -0.1 | 840.4209 | 2 | 54.76 | | | tr\|F7ANV6\|F7ANV6_MOUSE |
| AASGFNATEDAQTLR.K | Y | 28.85 | 1550.7324 | 15 | -1.4 | 776.3724 | 2 | 61.75 | | | tr\|F7ANV6\|F7ANV6_MOUSE |
| K.GLGTDDNTLIR.V | Y | 21.3 | 1173.599 | 11 | 1.9 | 587.8079 | 2 | 64.84 | | | tr\|F7ANV6\|F7ANV6_MOUSE |
| K.KAAEALDAAK.K | Y | 23.43 | 986.5396 | 10 | 0.9 | 494.2775 | 2 | 11.54 | | | tr\|F8VQM0\|F8VQM0_MOUSE |
| T.AAESSGSSYR.Q | Y | 22.85 | 1013.4413 | 10 | 1.1 | 507.7285 | 2 | 7.19 | | | tr\|F8VQM0\|F8VQM0_MOUSE |
| K.AAEALDAAKKLQPIQTSAK.N | Y | 22.18 | 1953.0894 | 19 | -2 | 652.0358 | 3 | 95.58 | | | tr\|F8VQM0\|F8VQM0_MOUSE |
| R.ANGAEYQTYVR.G | Y | 27.28 | 1270.5942 | 11 | -0.2 | 636.3043 | 2 | 41.5 | | | tr\|F8VQM5\|F8VQM5_MOUSE |
| R.VDMYLPEDK.I | Y | 25.98 | 1108.511 | 9 | 0.9 | 555.2632 | 2 | 66.29 | | | tr\|F8VQM5\|F8VQM5_MOUSE |
| A.SVTNDNSGLK.A | Y | 24.86 | 1033.504 | 10 | 0.6 | 517.7596 | 2 | 11.8 | | | tr\|F8VQM5\|F8VQM5_MOUSE |
| R.QLDFTIGER.F | Y | 22.76 | 1077.5454 | 9 | 0.9 | 539.7805 | 2 | 95.78 | | | tr\|F8VQM5\|F8VQM5_MOUSE |
| K.LIVAADDNQTAK.G | Y | 21.27 | 1257.6565 | 12 | -1.6 | 629.8345 | 2 | 33.42 | | | tr\|F8VQM5\|F8VQM5_MOUSE |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|K7KA53\|K7KA53_SOYBN |
| I.SLLDTSNFNNQLDQNPR.V | Y | 41.8 | 1974.9395 | 17 | 1.8 | 988.4788 | 2 | 96.86 | | | tr\|K7LZA4\|K7LZA4_SOYBN |
| R.SQQQLQDSHQK.I | Y | 30 | 1325.6323 | 11 | 1.2 | 663.8242 | 2 | 6.95 | | | tr\|K7LZA4\|K7LZA4_SOYBN |
| Q.QQLQDSHQ.K | N | 23.11 | 982.4468 | 8 | 1.7 | 492.2315 | 2 | 7.31 | | | tr\|K7LZA4\|K7LZA4_SOYBN |
| K.LHENIARPS.R | N | 22.38 | 1035.5461 | 9 | 0 | 518.7803 | 2 | 22.94 | | | tr\|K7LZA4\|K7LZA4_SOYBN |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|K7LZT4\|K7LZT4_SOYBN |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|K7M9J2\|K7M9J2_SOYBN |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|K7MQA4\|K7MQA4_SOYBN |
| R.LTAFEEAIPK.S | Y | 26.03 | 1117.6019 | 10 | 1.2 | 559.8089 | 2 | 96.89 | | | tr\|Q3TVK3\|Q3TVK3_MOUSE |
| A.M(+15.99)FTPQAEPVPDESPR.S | Y | 34.45 | 1715.7825 | 15 | -3.2 | 858.8958 | 2 | 61.8 | Oxidation (M) | M1:Oxidation (M):1000.00 | tr\|Q3UW98\|Q3UW98_MOUSE |
| I.SLTWTAPGDDYDVGR.V | N | 23.73 | 1651.7478 | 15 | 1.8 | 826.8826 | 2 | 96.84 | | | tr\|Q3UW98\|Q3UW98_MOUSE |
| A.MFTPQAEPVPDESPR.S | Y | 23.23 | 1699.7875 | 15 | 2.6 | 850.9032 | 2 | 71.39 | | | tr\|Q3UW98\|Q3UW98_MOUSE |
| Q.AEPVPDESPR.S | Y | 20.82 | 1095.5197 | 10 | 3.2 | 548.7689 | 2 | 23.75 | | | tr\|Q3UW98\|Q3UW98_MOUSE |
| K.LAPEFAK.R | Y | 22.97 | 774.4276 | 7 | -0.3 | 388.2209 | 2 | 37.44 | | | tr\|Q6GT24\|Q6GT24_MOUSE |
| R.LTAFEEAIPK.S | Y | 26.03 | 1117.6019 | 10 | 1.2 | 559.8089 | 2 | 96.89 | | | tr\|Q8BPW9\|Q8BPW9_MOUSE |
| K.IGGHGAEYGAEALER.M | Y | 36.3 | 1528.7269 | 15 | 0 | 765.3707 | 2 | 57.64 | | | tr\|Q91VB8\|Q91VB8_MOUSE |
| P.EPEPEPEPEPEPVPEE.A | Y | 26.09 | 1828.7889 | 16 | -3.9 | 915.3982 | 2 | 53.34 | | | tr\|Q91WA8\|Q91WA8_MOUSE |
| E.PEPEPEPEPEPEPEPV.P | Y | 21.57 | 1796.7992 | 16 | 1.3 | 899.4081 | 2 | 65.07 | | | tr\|Q91WA8\|Q91WA8_MOUSE |
| R.TLNNDIM(+15.99)LIK.L | Y | 22.91 | 1189.6377 | 10 | 0.5 | 595.8264 | 2 | 96.26 | Oxidation (M) | M7:Oxidation (M):1000.00 | tr\|Q9QUK9\|Q9QUK9_MOUSE |
| R.TLNNDIM(+15.99)LIK.L | Y | 22.91 | 1189.6377 | 10 | 0.5 | 595.8264 | 2 | 96.26 | Oxidation (M) | M7:Oxidation (M):1000.00 | tr\|Q9R0T7\|Q9R0T7_MOUSE |
| I.SLLDTSNFNNQLDQTPR.V | Y | 30.23 | 1961.9442 | 17 | 1.4 | 981.9807 | 2 | 96.92 | | | tr\|Q9SB11\|Q9SB11_SOYBN |
| K.QQLQDSHQ.K | N | 23.11 | 982.4468 | 8 | 1.7 | 492.2315 | 2 | 7.31 | | | tr\|Q9SB11\|Q9SB11_SOYBN |
| K.LHENIARPS.R | N | 22.38 | 1035.5461 | 9 | 0 | 518.7803 | 2 | 22.94 | | | tr\|Q9SB11\|Q9SB11_SOYBN |
| H.QQEEEEEGGSVLSGFSK.H | Y | 21.4 | 1838.8169 | 17 | 3.4 | 920.4188 | 2 | 76.97 | | | tr\|Q9SB11\|Q9SB11_SOYBN |
| R.NGLHLPSYSPYPR.M | Y | 21.37 | 1499.7521 | 13 | -0.1 | 500.9246 | 3 | 96.96 | | | tr\|Q9SB11\|Q9SB11_SOYBN |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W4ZQV5\|W4ZQV5_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W4ZVZ3\|W4ZVZ3_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W5APN0\|W5APN0_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W5AWL4\|W5AWL4_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W5B8I7\|W5B8I7_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W5B954\|W5B954_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W5BDU9\|W5BDU9_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W5BPM3\|W5BPM3_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W5FBB1\|W5FBB1_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W5GDX3\|W5GDX3_WHEAT |
| R.DNIQGITKPAIR.R | Y | 25.4 | 1324.7462 | 12 | 0.3 | 663.3806 | 2 | 59.52 | | | tr\|W5GE89\|W5GE89_WHEAT |
| K.TAGSQEAPNLPR.K | Y | 25.06 | 1239.6207 | 12 | -1.2 | 620.8169 | 2 | 35.67 | | | tr\|Z4YJC9\|Z4YJC9_MOUSE |

**Supplemental Table 4.** Protein Identifications from methanol/ chloroform/ water extraction.

| Protein Group | Protein ID | Accession | -10lgP | Coverage (%) | #Peptides | # Unique | PTM | Avg. Mass | Description |
|---|---|---|---|---|---|---|---|---|---|
| 61 | 68201 | Amuc\|Amuc_1048 | 45.54 | 6 | 5 | 1 | | 43431 | elongation factor Tu |
| 156 | 5893 | ANACOL\|ANACOL_03730 | 31.17 | 15 | 2 | 2 | | 14879 | hypothetical protein |
| 199 | 54005 | B216\|B216_01337 | 31.44 | 10 | 1 | 1 | | 11550 | 30S ribosomal protein S10 |
| 199 | 54015 | B217\|B217_03557 | 31.44 | 10 | 1 | 1 | | 11550 | 30S ribosomal protein S10 |
| 302 | 93804 | BACCAC\|BACCAC_00202 | 29.53 | 5 | 1 | 1 | | 21913 | hypothetical protein |
| 162 | 67854 | BACCAC\|BACCAC_00868 | 40.67 | 2 | 2 | 1 | | 99571 | hypothetical protein |
| 82 | 68416 | BACCELL\|BACCELL_00889 | 43.34 | 21 | 3 | 3 | | 10534 | hypothetical protein |
| 201 | 40731 | BACCELL\|BACCELL_02692 | 34.45 | 16 | 2 | 2 | | 15751 | hypothetical protein |
| 302 | 117828 | BACCELL\|BACCELL_02704 | 29.53 | 5 | 1 | 1 | | 21869 | hypothetical protein |
| 120 | 24605 | BACCELL\|BACCELL_05494 | 45.74 | 1 | 3 | 1 | | 129934 | hypothetical protein |
| 201 | 40721 | BACCOPRO\|BACCOPRO_00303 | 34.45 | 16 | 2 | 2 | | 15708 | hypothetical protein |
| 162 | 67910 | BACCOPRO\|BACCOPRO_00810 | 40.67 | 2 | 2 | 1 | | 99788 | hypothetical protein |
| 82 | 68444 | BACCOPRO\|BACCOPRO_03489 | 43.34 | 21 | 3 | 3 | | 10561 | hypothetical protein |
| 321 | 68587 | BACDOR\|BACDOR_00205 | 26.23 | 4 | 1 | 1 | | 32863 | hypothetical protein |
| 321 | 68588 | BACEGG\|BACEGG_00076 | 26.23 | 4 | 1 | 1 | | 32632 | hypothetical protein |
| 82 | 68418 | BACEGG\|BACEGG_00173 | 43.34 | 21 | 3 | 3 | | 10662 | hypothetical protein |
| 120 | 24607 | BACEGG\|BACEGG_01855 | 45.74 | 1 | 3 | 1 | | 130009 | hypothetical protein |
| 302 | 117833 | BACEGG\|BACEGG_02938 | 29.53 | 5 | 1 | 1 | | 21886 | hypothetical protein |
| 182 | 67777 | BACEGG\|BACEGG_02940 | 39.09 | 3 | 2 | 1 | | 77583 | hypothetical protein |
| 201 | 40709 | BACEGG\|BACEGG_02960 | 34.45 | 16 | 2 | 2 | | 15808 | hypothetical protein |
| 302 | 93802 | BACFIN\|BACFIN_04868 | 29.53 | 5 | 1 | 1 | | 21948 | 50S ribosomal protein L3 |
| 321 | 68586 | BACFIN\|BACFIN_05433 | 26.23 | 4 | 1 | 1 | | 32774 | malate dehydrogenase NAD-dependent |
| 82 | 68427 | BACFIN\|BACFIN_06438 | 43.34 | 21 | 3 | 3 | | 10676 | ribosomal protein S18 |
| 201 | 40715 | BACFIN\|BACFIN_07794 | 34.45 | 16 | 2 | 2 | | 15868 | ribosomal protein L11 |
| 82 | 68415 | BACINT\|BACINT_00313 | 43.34 | 21 | 3 | 3 | | 10534 | hypothetical protein |
| 201 | 40712 | BACINT\|BACINT_01760 | 34.45 | 16 | 2 | 2 | | 15821 | hypothetical protein |
| 302 | 117827 | BACINT\|BACINT_01771 | 29.53 | 5 | 1 | 1 | | 21870 | hypothetical protein |
| 120 | 24604 | BACINT\|BACINT_03177 | 45.74 | 1 | 3 | 1 | | 130030 | hypothetical protein |
| 302 | 93788 | BACOVA\|BACOVA_01027 | 29.53 | 5 | 1 | 1 | | 21829 | hypothetical protein |
| 201 | 40728 | BACOVA\|BACOVA_01040 | 34.45 | 16 | 2 | 2 | | 15836 | hypothetical protein |
| 49 | 24306 | BACOVA\|BACOVA_01044 | 53.38 | 12 | 5 | 3 | | 43648 | hypothetical protein |
| 162 | 67906 | BACOVA\|BACOVA_02268 | 40.67 | 2 | 2 | 1 | | 99731 | hypothetical protein |
| 82 | 68432 | BACOVA\|BACOVA_03184 | 43.34 | 21 | 3 | 3 | | 10704 | hypothetical protein |
| 217 | 93504 | BACPLE\|BACPLE_02299 | 35.43 | 5 | 2 | 2 | | 43862 | hypothetical protein |
| 82 | 68440 | BACPLE\|BACPLE_03080 | 43.34 | 21 | 3 | 3 | | 10561 | hypothetical protein |
| 49 | 54334 | BACPLE\|BACPLE_03647 | 53.38 | 12 | 5 | 3 | | 43557 | hypothetical protein |
| 201 | 40730 | BACPLE\|BACPLE_03651 | 34.45 | 16 | 2 | 2 | | 15712 | hypothetical protein |
| 302 | 117831 | BACSTE\|BACSTE_02756 | 29.53 | 5 | 1 | 1 | | 21902 | hypothetical protein |
| 321 | 68567 | BACSTE\|BACSTE_03693 | 26.23 | 4 | 1 | 1 | | 32738 | hypothetical protein |
| 302 | 117829 | BACUNI\|BACUNI_00812 | 29.53 | 5 | 1 | 1 | | 21917 | hypothetical protein |
| 201 | 40716 | BACUNI\|BACUNI_00823 | 34.45 | 16 | 2 | 2 | | 15867 | hypothetical protein |
| 321 | 68592 | BACUNI\|BACUNI_03132 | 26.23 | 4 | 1 | 1 | | 32782 | hypothetical protein |
| 82 | 68420 | BACUNI\|BACUNI_03359 | 43.34 | 21 | 3 | 3 | | 10662 | hypothetical protein |
| 199 | 54002 | BAD\|BAD_0320 | 31.44 | 10 | 1 | 1 | | 11516 | 30S ribosomal protein S10 |
| 199 | 53998 | BBB\|BBB_1541 | 31.44 | 10 | 1 | 1 | | 11550 | 30S ribosomal protein S10 |
| 199 | 54003 | BBIF\|BBIF_1505 | 31.44 | 10 | 1 | 1 | | 11550 | 30S ribosomal protein S10 |
| 199 | 53996 | BBNG\|BBNG_01539 | 31.44 | 10 | 1 | 1 | | 11550 | 30S ribosomal protein S10 |
| 199 | 54000 | BBPR\|BBPR_1558 | 31.44 | 10 | 1 | 1 | | 11550 | 30S ribosomal protein S10 |
| 321 | 68582 | BDI\|BDI_0389 | 26.23 | 4 | 1 | 1 | | 32848 | malate dehydrogenase |
| 302 | 117840 | BDI\|BDI_2380 | 29.53 | 5 | 1 | 1 | | 22134 | 50S ribosomal protein L3 |
| 82 | 68414 | BDI\|BDI_3968 | 43.34 | 21 | 3 | 3 | | 10487 | 30S ribosomal protein S18 |
| 199 | 54006 | BDP\|BDP_0429 | 31.44 | 10 | 1 | 1 | | 11550 | 30S ribosomal protein S10 |
| 199 | 54017 | BIFADO\|BIFADO_00571 | 31.44 | 10 | 1 | 1 | | 11550 | hypothetical protein |
| 199 | 54011 | BIFANG\|BIFANG_03539 | 31.44 | 10 | 1 | 1 | | 11554 | hypothetical protein |
| 199 | 54019 | BIFDEN\|BIFDEN_00349 | 31.44 | 10 | 1 | 1 | | 11550 | hypothetical protein |
| 199 | 54009 | BIFPSEUDO\|BIFPSEUDO_02628 | 31.44 | 10 | 1 | 1 | | 11550 | hypothetical protein |
| 220 | 42922 | BLAHAN\|BLAHAN_04523 | 28.51 | 5 | 1 | 1 | | 23781 | formate--tetrahydrofolate ligase |
| 236 | 5589 | BLAHAN\|BLAHAN_04933 | 30.31 | 1 | 1 | 1 | | 76888 | translation elongation factor G |
| 94 | 67756 | BLAHAN\|BLAHAN_04936 | 50.83 | 5 | 4 | 4 | | 59448 | phosphoenolpyruvate carboxykinase |
| 220 | 42921 | BLAHAN\|BLAHAN_05779 | 28.51 | 9 | 1 | 1 | | 12335 | formate--tetrahydrofolate ligase |
| 156 | 54311 | BLAHAN\|BLAHAN_05968 | 31.17 | 15 | 2 | 2 | | 14808 | ribosomal protein L11 |
| 274 | 117855 | BLAHAN\|BLAHAN_06361 | 21.88 | 6 | 1 | 1 | | 15339 | ribosomal protein L15 |
| 220 | 42936 | BRYFOR\|BRYFOR_06351 | 28.51 | 2 | 1 | 1 | | 60240 | formate--tetrahydrofolate ligase |
| 315 | 93222 | BRYFOR\|BRYFOR_06378 | 22.58 | 10 | 1 | 1 | | 13097 | ribosomal protein L19 |
| 220 | 42937 | BRYFOR\|BRYFOR_07447 | 28.51 | 2 | 1 | 1 | | 60597 | formate--tetrahydrofolate ligase |
| 236 | 54055 | BRYFOR\|BRYFOR_07805 | 30.31 | 1 | 1 | 1 | | 77063 | putative translation elongation factor G |
| 226 | 117792 | BRYFOR\|BRYFOR_07955 | 35.04 | 5 | 1 | 1 | | 22588 | 50S ribosomal protein L4 |
| 137 | 304 | BRYFOR\|BRYFOR_09074 | 41.26 | 3 | 2 | 1 | Oxidation (M) | 75905 | translation elongation factor G |
| 321 | 68583 | BSEG\|BSEG_02576 | 26.23 | 4 | 1 | 1 | | 32863 | malate dehydrogenase |
| 162 | 117718 | BT\|BT_0644 | 40.67 | 2 | 2 | 1 | | 99684 | pyruvate phosphate dikinase |
| 302 | 117807 | BT\|BT_2727 | 29.53 | 5 | 1 | 1 | | 21903 | 50S ribosomal protein L3 |
| 321 | 68685 | BT\|BT_3911 | 26.23 | 4 | 1 | 1 | | 32763 | malate dehydrogenase |
| 321 | 68571 | BVU\|BVU_0256 | 26.23 | 4 | 1 | 1 | | 32847 | malate dehydrogenase |
| 162 | 67903 | BXY\|BXY_02130 | 40.67 | 2 | 2 | 1 | | 99673 | pyruvate phosphate dikinase |
| 82 | 68433 | BXY\|BXY_11220 | 43.34 | 21 | 3 | 3 | | 10704 | SSU ribosomal protein S18P |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 302 | 117803 | BXY\|BXY_18520 | 29.53 | 5 | 1 | 1 | | 21874 | LSU ribosomal protein L3P |
| 201 | 40726 | BXY\|BXY_18630 | 34.45 | 16 | 2 | 2 | | 15836 | LSU ribosomal protein L11P |
| 49 | 18416 | BXY\|BXY_18660 | 53.38 | 12 | 5 | 3 | | 43662 | translation elongation factor 1A (EF-1A/EF-Tu) |
| 185 | 67876 | CIT292\|CIT292_07476 | 35.62 | 6 | 2 | 1 | | 48549 | NADP-specific glutamate dehydrogenase |
| 236 | 54084 | CLOBOL\|CLOBOL_01950 | 30.31 | 1 | 1 | 1 | | 77531 | hypothetical protein |
| 78 | 41561 | CLOBOL\|CLOBOL_03031 | 54.11 | 9 | 4 | 4 | | 35997 | hypothetical protein |
| 274 | 117853 | CLOBOL\|CLOBOL_05944 | 21.88 | 6 | 1 | 1 | | 15477 | hypothetical protein |
| 76 | 41809 | CLOBOL\|CLOBOL_05950 | 49.41 | 25 | 5 | 1 | | 20369 | hypothetical protein |
| 220 | 42688 | CLOBOL\|CLOBOL_07321 | 28.51 | 2 | 1 | 1 | | 59781 | hypothetical protein |
| 156 | 5897 | CLOHIR\|CLOHIR_01566 | 31.17 | 15 | 2 | 2 | | 14786 | hypothetical protein |
| 54 | 243 | CLOHYLEM\|CLOHYLEM_05194 | 61.85 | 7 | 4 | 1 | | 62571 | hypothetical protein |
| 244 | 68088 | CLOHYLEM\|CLOHYLEM_05797 | 38.62 | 3 | 1 | 1 | | 56286 | hypothetical protein |
| 121 | 187 | CLOHYLEM\|CLOHYLEM_05831 | 38.36 | 3 | 2 | 1 | Oxidation (M) | 78082 | hypothetical protein |
| 156 | 70741 | CLOHYLEM\|CLOHYLEM_05892 | 31.17 | 15 | 2 | 2 | | 14898 | hypothetical protein |
| 68 | 67664 | CLOHYLEM\|CLOHYLEM_06153 | 61.98 | 10 | 4 | 1 | | 41300 | hypothetical protein |
| 220 | 42338 | CLOHYLEM\|CLOHYLEM_06355 | 28.51 | 2 | 1 | 1 | | 59582 | hypothetical protein |
| 167 | 67893 | CLOHYLEM\|CLOHYLEM_06525 | 38.97 | 5 | 2 | 2 | | 45729 | hypothetical protein |
| 274 | 68168 | CLOHYLEM\|CLOHYLEM_07670 | 21.88 | 6 | 1 | 1 | | 15340 | hypothetical protein |
| 156 | 5898 | CLOLEP\|CLOLEP_02932 | 31.17 | 15 | 2 | 2 | | 14928 | hypothetical protein |
| 220 | 42935 | CLOLEP\|CLOLEP_03336 | 28.51 | 2 | 1 | 1 | | 60293 | hypothetical protein |
| 195 | 117974 | CLONEX\|CLONEX_00038 | 28.59 | 14 | 2 | 2 | | 7300 | hypothetical protein |
| 244 | 69675 | CLONEX\|CLONEX_00798 | 38.62 | 3 | 1 | 1 | | 57316 | hypothetical protein |
| 274 | 19022 | CLONEX\|CLONEX_00948 | 21.88 | 6 | 1 | 1 | | 15454 | hypothetical protein |
| 200 | 41089 | CLONEX\|CLONEX_01120 | 37.75 | 16 | 2 | 2 | | 14516 | hypothetical protein |
| 156 | 54315 | CLONEX\|CLONEX_01920 | 31.17 | 15 | 2 | 2 | | 14794 | hypothetical protein |
| 315 | 93221 | CLONEX\|CLONEX_03041 | 22.58 | 10 | 1 | 1 | | 13015 | hypothetical protein |
| 210 | 68025 | CLONEX\|CLONEX_03175 | 31.4 | 1 | 1 | 1 | | 76504 | hypothetical protein |
| 94 | 43382 | CLONEX\|CLONEX_03201 | 50.83 | 5 | 4 | 4 | | 60092 | hypothetical protein |
| 290 | 117812 | CLONEX\|CLONEX_03523 | 21.31 | 2 | 1 | 1 | | 50782 | hypothetical protein |
| 69 | 67643 | CLONEX\|CLONEX_03627 | 56.86 | 10 | 4 | 1 | | 41212 | hypothetical protein |
| 114 | 40953 | CLOSCI\|CLOSCI_01585 | 43.98 | 1 | 3 | 1 | | 127484 | hypothetical protein |
| 274 | 68167 | CLOSCI\|CLOSCI_02206 | 21.88 | 6 | 1 | 1 | | 15365 | hypothetical protein |
| 51 | 43855 | CLOSCI\|CLOSCI_02348 | 57.21 | 11 | 4 | 2 | | 32746 | hypothetical protein |
| 55 | 247 | CLOSCI\|CLOSCI_03565 | 51.36 | 6 | 4 | 1 | Oxidation (M) | 62798 | hypothetical protein |
| 274 | 117854 | CLOSTASPAR\|CLOSTASPAR_00971 | 21.88 | 6 | 1 | 1 | | 15445 | hypothetical protein |
| 76 | 41752 | CLOSTASPAR\|CLOSTASPAR_00977 | 49.41 | 25 | 5 | 1 | | 20351 | hypothetical protein |
| 78 | 41567 | CLOSTASPAR\|CLOSTASPAR_01821 | 54.11 | 9 | 4 | 4 | | 35751 | hypothetical protein |
| 94 | 43379 | CLOSTASPAR\|CLOSTASPAR_02626 | 50.83 | 5 | 4 | 4 | | 59406 | hypothetical protein |
| 244 | 68895 | CLOSTASPAR\|CLOSTASPAR_02873 | 38.62 | 3 | 1 | 1 | | 54487 | hypothetical protein |
| 290 | 117848 | CLOSTASPAR\|CLOSTASPAR_05219 | 21.31 | 5 | 1 | 1 | | 21856 | hypothetical protein |
| 236 | 54085 | CLOSTASPAR\|CLOSTASPAR_05682 | 30.31 | 1 | 1 | 1 | | 77568 | hypothetical protein |
| 156 | 5899 | CLOSTASPAR\|CLOSTASPAR_06716 | 31.17 | 15 | 2 | 2 | | 14779 | hypothetical protein |
| 274 | 117862 | CLOSTHATH\|CLOSTHATH_01324 | 21.88 | 6 | 1 | 1 | | 17137 | ribosomal protein L15 |
| 236 | 54068 | CLOSTHATH\|CLOSTHATH_04938 | 30.31 | 1 | 1 | 1 | | 78232 | translation elongation factor G |
| 220 | 42928 | COLAER\|COLAER_01445 | 28.51 | 2 | 1 | 1 | | 59119 | hypothetical protein |
| 220 | 42931 | COLINT\|COLINT_01991 | 28.51 | 2 | 1 | 1 | | 59010 | hypothetical protein |
| 220 | 42927 | COLSTE\|COLSTE_02400 | 28.51 | 2 | 1 | 1 | | 59072 | hypothetical protein |
| 200 | 41090 | COPCOM\|COPCOM_00693 | 37.75 | 16 | 2 | 2 | | 14539 | hypothetical protein |
| 274 | 93279 | COPCOM\|COPCOM_00801 | 21.88 | 6 | 1 | 1 | | 15390 | hypothetical protein |
| 233 | 40681 | COPCOM\|COPCOM_01471 | 33.74 | 4 | 1 | 1 | | 37227 | hypothetical protein |
| 119 | 117721 | COPCOM\|COPCOM_01619 | 38.83 | 5 | 3 | 3 | | 55164 | hypothetical protein |
| 315 | 93258 | COPCOM\|COPCOM_02935 | 22.58 | 10 | 1 | 1 | | 13066 | hypothetical protein |
| 80 | 67680 | COPEUT\|COPEUT_01183 | 47.16 | 9 | 3 | 1 | | 35338 | hypothetical protein |
| 274 | 117860 | COPEUT\|COPEUT_02174 | 21.88 | 6 | 1 | 1 | | 15480 | hypothetical protein |
| 321 | 68580 | CUU\|CUU_1681 | 26.23 | 4 | 1 | 1 | | 32847 | malate dehydrogenase NAD-dependent |
| 201 | 40717 | CUU\|CUU_2118 | 34.45 | 16 | 2 | 2 | | 15797 | ribosomal protein L11 |
| 82 | 68439 | CUU\|CUU_3885 | 43.34 | 21 | 3 | 3 | | 10561 | ribosomal protein S18 |
| 82 | 68435 | CUY\|CUY_1523 | 43.34 | 21 | 3 | 3 | | 10676 | ribosomal protein S18 |
| 302 | 117834 | CUY\|CUY_3299 | 29.53 | 7 | 1 | 1 | | 15233 | 50S ribosomal protein L3 |
| 201 | 40727 | CUY\|CUY_3311 | 34.45 | 16 | 2 | 2 | | 15836 | ribosomal protein L11 |
| 49 | 56 | CUY\|CUY_3315 | 53.38 | 12 | 5 | 3 | | 43662 | translation elongation factor Tu |
| 162 | 67907 | CUY\|CUY_4072 | 40.67 | 2 | 2 | 1 | | 99659 | pyruvate phosphate dikinase |
| 49 | 41234 | CW1\|CW1_0481 | 53.38 | 12 | 5 | 3 | | 43662 | translation elongation factor Tu |
| 201 | 40711 | CW1\|CW1_0485 | 34.45 | 16 | 2 | 2 | | 15836 | ribosomal protein L11 |
| 302 | 117806 | CW1\|CW1_0496 | 29.53 | 5 | 1 | 1 | | 21874 | 50S ribosomal protein L3 |
| 82 | 68429 | CW1\|CW1_0560 | 43.34 | 21 | 3 | 3 | | 10704 | ribosomal protein S18 |
| 162 | 67904 | CW1\|CW1_3565 | 40.67 | 2 | 2 | 1 | | 99719 | pyruvate phosphate dikinase |
| 82 | 68430 | CW3\|CW3_1327 | 43.34 | 21 | 3 | 3 | | 10704 | ribosomal protein S18 |
| 49 | 24307 | CW3\|CW3_3618 | 53.38 | 12 | 5 | 3 | | 43662 | translation elongation factor Tu |
| 201 | 40722 | CW3\|CW3_3622 | 34.45 | 16 | 2 | 2 | | 15836 | ribosomal protein L11 |
| 162 | 67908 | CW3\|CW3_4066 | 40.67 | 2 | 2 | 1 | | 99719 | pyruvate phosphate dikinase |
| 302 | 117804 | CW3\|CW3_4420 | 29.53 | 5 | 1 | 1 | | 21874 | 50S ribosomal protein L3 |
| 113 | 18530 | DESPIG\|DESPIG_00100 | 49.92 | 10 | 4 | 2 | | 43452 | hypothetical protein |
| 51 | 67754 | DORFOR\|DORFOR_01112 | 57.21 | 12 | 4 | 2 | | 31589 | hypothetical protein |
| 156 | 54318 | DORFOR\|DORFOR_01208 | 31.17 | 13 | 2 | 2 | | 16695 | hypothetical protein |
| 244 | 68093 | DORFOR\|DORFOR_01285 | 38.62 | 3 | 1 | 1 | | 55838 | hypothetical protein |
| 274 | 68659 | DORFOR\|DORFOR_01340 | 21.88 | 6 | 1 | 1 | | 15394 | hypothetical protein |
| 89 | 24 | DORFOR\|DORFOR_01431 | 51.44 | 8 | 4 | 1 | | 36196 | hypothetical protein |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 94 | 67951 | DORFOR\|DORFOR_01796 | 50.83 | 5 | 4 | 4 | | 59354 | hypothetical protein |
| 195 | 117973 | DORFOR\|DORFOR_02174 | 28.59 | 14 | 2 | 2 | | 7190 | hypothetical protein |
| 114 | 40958 | DORFOR\|DORFOR_02415 | 43.98 | 1 | 3 | 1 | | 128131 | hypothetical protein |
| 129 | 68120 | DORFOR\|DORFOR_02813 | 41.8 | 7 | 3 | 2 | | 49611 | hypothetical protein |
| 51 | 42990 | DORLON\|DORLON_00078 | 57.21 | 12 | 4 | 2 | | 30922 | hypothetical protein |
| 195 | 117972 | DORLON\|DORLON_01454 | 28.59 | 14 | 2 | 2 | | 7210 | hypothetical protein |
| 210 | 69294 | DORLON\|DORLON_01862 | 31.4 | 1 | 1 | 1 | | 76462 | hypothetical protein |
| 274 | 94516 | DORLON\|DORLON_02208 | 21.88 | 6 | 1 | 1 | | 15374 | hypothetical protein |
| 200 | 41082 | DORLON\|DORLON_02544 | 37.75 | 16 | 2 | 2 | | 14415 | hypothetical protein |
| 185 | 40838 | ECD227\|ECD227_01179 | 35.62 | 6 | 2 | 1 | | 48511 | glutamate dehydrogenase |
| 185 | 53935 | ECDH10B\|ECDH10B_1899 | 35.62 | 6 | 2 | 1 | | 48581 | glutamate dehydrogenase |
| 185 | 67808 | EFER\|EFER_1304 | 35.62 | 6 | 2 | 1 | | 48555 | glutamate dehydrogenase |
| 185 | 40866 | ENTCAN\|ENTCAN_05740 | 35.62 | 6 | 2 | 1 | | 48398 | glutamate dehydrogenase |
| 185 | 67807 | ERIG\|ERIG_01926 | 35.62 | 6 | 2 | 1 | | 48555 | glutamate/Leucine/Phenylalanine/Valine dehydrogenase |
| 202 | 41055 | EUBDOL\|EUBDOL_00065 | 28.57 | 3 | 1 | 1 | | 29964 | hypothetical protein |
| 169 | 54198 | EUBDOL\|EUBDOL_02012 | 39.97 | 7 | 2 | 2 | | 31499 | hypothetical protein |
| 156 | 24423 | EUBELI\|EUBELI_00279 | 31.17 | 13 | 2 | 2 | | 16737 | 50S ribosomal protein L11 |
| 34 | 18412 | EUBELI\|EUBELI_00288 | 59 | 13 | 7 | 1 | | 43997 | elongation factor Tu |
| 195 | 117975 | EUBELI\|EUBELI_00506 | 28.59 | 11 | 2 | 2 | | 9497 | cold shock protein |
| 101 | 40883 | EUBELI\|EUBELI_01617 | 55.97 | 8 | 4 | 1 | | 48972 | glutamate dehydrogenase |
| 305 | 71345 | EUBELI\|EUBELI_01623 | 26.25 | 2 | 1 | 1 | | 64656 | phosphomannomutase |
| 200 | 41092 | EUBHAL\|EUBHAL_00646 | 37.75 | 16 | 2 | 2 | | 14610 | hypothetical protein |
| 156 | 54340 | EUBHAL\|EUBHAL_01120 | 31.17 | 15 | 2 | 2 | | 14961 | hypothetical protein |
| 274 | 117851 | EUBVEN\|EUBVEN_00870 | 21.88 | 6 | 1 | 1 | | 15445 | hypothetical protein |
| 76 | 41812 | EUBVEN\|EUBVEN_00877 | 49.41 | 27 | 5 | 1 | | 18730 | hypothetical protein |
| 104 | 42297 | EUBVEN\|EUBVEN_00885 | 47.58 | 12 | 3 | 1 | | 30321 | hypothetical protein |
| 220 | 42686 | FAEPRAA2165\|FAEPRAA2165_01688 | 28.51 | 2 | 1 | 1 | | 62243 | formate--tetrahydrofolate ligase |
| 220 | 42685 | FAEPRAM212\|FAEPRAM212_03554 | 28.51 | 2 | 1 | 1 | | 61243 | hypothetical protein |
| 94 | 67954 | FP2\|FP2_13680 | 50.83 | 5 | 4 | 4 | | 59289 | phosphoenolpyruvate carboxykinase (ATP) |
| 220 | 42684 | FPR\|FPR_03960 | 28.51 | 2 | 1 | 1 | | 61243 | Formate-tetrahydrofolate ligase |
| 94 | 67949 | FPR\|FPR_10380 | 50.83 | 5 | 4 | 4 | | 59146 | phosphoenolpyruvate carboxykinase (ATP) |
| 236 | 54090 | FVAG\|FVAG_00896 | 30.31 | 1 | 1 | 1 | | 76724 | translation elongation factor G |
| 290 | 117822 | FVAG\|FVAG_02380 | 21.31 | 2 | 1 | 1 | | 56942 | PTS system protein |
| 315 | 93223 | Ga0056060\|Ga0056060_00078 | 22.58 | 10 | 1 | 1 | | 13097 | ribosomal protein L19 bacterial type |
| 220 | 42934 | Ga0056060\|Ga0056060_00103 | 28.51 | 2 | 1 | 1 | | 59809 | Formate-tetrahydrofolate ligase (EC 6.3.4.3) |
| 236 | 54054 | Ga0056060\|Ga0056060_00936 | 30.31 | 1 | 1 | 1 | | 77063 | small GTP-binding protein domain |
| 220 | 6666 | Ga0056060\|Ga0056060_02051 | 28.51 | 2 | 1 | 1 | | 59883 | Formate-tetrahydrofolate ligase (EC 6.3.4.3) |
| 226 | 117793 | Ga0056060\|Ga0056060_02724 | 35.04 | 5 | 1 | 1 | | 22588 | 50S ribosomal protein L4 bacterial/organelle |
| 137 | 67749 | Ga0056060\|Ga0056060_04116 | 41.26 | 3 | 2 | 1 | Oxidation (M) | 77746 | translation elongation factor EF-G |
| 220 | 42932 | Ga0062113\|Ga0062113_01430 | 28.51 | 2 | 1 | 1 | | 59039 | Formate-tetrahydrofolate ligase (EC 6.3.4.3) |
| 140 | 67758 | Ga0062113\|Ga0062113_02010 | 41.87 | 9 | 2 | 1 | | 20444 | Ribosomal protein L5 |
| 199 | 53999 | HMPREF0168\|HMPREF0168_1504 | 31.44 | 10 | 1 | 1 | | 11550 | 30S ribosomal protein S10 |
| 156 | 71557 | HMPREF0372\|HMPREF0372_00855 | 31.17 | 15 | 2 | 2 | | 14764 | ribosomal protein L11 |
| 140 | 67759 | HMPREF0372\|HMPREF0372_03688 | 41.87 | 9 | 2 | 1 | | 20444 | ribosomal protein L5 |
| 220 | 42943 | HMPREF0372\|HMPREF0372_03978 | 28.51 | 2 | 1 | 1 | | 69807 | formate--tetrahydrofolate ligase |
| 199 | 53997 | HMPREF0534\|HMPREF0534_0210 | 31.44 | 10 | 1 | 1 | | 11702 | ribosomal protein S10 |
| 172 | 53940 | HMPREF0534\|HMPREF0534_1678 | 37.57 | 4 | 2 | 1 | | 42965 | phosphoglycerate kinase |
| 172 | 53941 | HMPREF0535\|HMPREF0535_0370 | 37.57 | 4 | 2 | 1 | | 42960 | phosphoglycerate kinase |
| 199 | 54001 | HMPREF0535\|HMPREF0535_0765 | 31.44 | 10 | 1 | 1 | | 11777 | ribosomal protein S10 |
| 172 | 53944 | HMPREF0536\|HMPREF0536_10423 | 37.57 | 4 | 2 | 1 | | 42960 | phosphoglycerate kinase |
| 199 | 54004 | HMPREF0536\|HMPREF0536_11469 | 31.44 | 10 | 1 | 1 | | 11777 | 30S ribosomal protein S10 |
| 199 | 54012 | HMPREF0538\|HMPREF0538_20591 | 31.44 | 10 | 1 | 1 | | 11702 | 30S ribosomal protein S10 |
| 172 | 53946 | HMPREF0538\|HMPREF0538_21607 | 37.57 | 4 | 2 | 1 | | 42965 | phosphoglycerate kinase |
| 120 | 42044 | HMPREF1016\|HMPREF1016_00015 | 45.74 | 1 | 3 | 1 | | 130039 | ferredoxin oxidoreductase |
| 321 | 68566 | HMPREF1016\|HMPREF1016_01914 | 26.23 | 4 | 1 | 1 | | 32632 | malate dehydrogenase |
| 302 | 93801 | HMPREF1016\|HMPREF1016_02588 | 29.53 | 5 | 1 | 1 | | 20891 | 50S ribosomal protein L3 |
| 162 | 117716 | HMPREF1017\|HMPREF1017_01162 | 40.67 | 2 | 2 | 1 | | 99803 | pyruvate phosphate dikinase |
| 302 | 93790 | HMPREF1017\|HMPREF1017_04982 | 29.53 | 5 | 1 | 1 | | 21843 | 50S ribosomal protein L3 |
| 302 | 93803 | HMPREF1057\|HMPREF1057_01845 | 29.53 | 5 | 1 | 1 | | 21976 | 50S ribosomal protein L3 |
| 201 | 40720 | HMPREF1057\|HMPREF1057_02462 | 34.45 | 16 | 2 | 2 | | 15882 | 50S ribosomal protein L11 |
| 321 | 68577 | HMPREF1057\|HMPREF1057_02815 | 26.23 | 4 | 1 | 1 | | 32774 | malate dehydrogenase |
| 82 | 68431 | HMPREF1057\|HMPREF1057_03907 | 43.34 | 21 | 3 | 3 | | 10676 | 30S ribosomal protein S18 |
| 321 | 68573 | HMPREF1058\|HMPREF1058_03596 | 26.23 | 4 | 1 | 1 | | 32847 | malate dehydrogenase |
| 321 | 68581 | HMPREF1059\|HMPREF1059_01913 | 26.23 | 4 | 1 | 1 | | 32848 | malate dehydrogenase |
| 82 | 68411 | HMPREF1059\|HMPREF1059_02083 | 43.34 | 21 | 3 | 3 | | 10487 | 30S ribosomal protein S18 |
| 120 | 7425 | HMPREF1059\|HMPREF1059_02475 | 45.74 | 1 | 3 | 1 | | 129377 | pyruvate:ferredoxin (flavodoxin) oxidoreductase |
| 302 | 117838 | HMPREF1059\|HMPREF1059_03078 | 29.53 | 5 | 1 | 1 | | 22134 | 50S ribosomal protein L3 |
| 162 | 117717 | HMPREF1060\|HMPREF1060_00821 | 40.67 | 2 | 2 | 1 | | 99737 | pyruvate phosphate dikinase |
| 302 | 68708 | HMPREF1060\|HMPREF1060_02306 | 29.53 | 5 | 1 | 1 | | 22176 | 50S ribosomal protein L3 |
| 321 | 68579 | HMPREF1060\|HMPREF1060_02613 | 26.23 | 4 | 1 | 1 | | 32910 | malate dehydrogenase |
| 162 | 67853 | HMPREF1061\|HMPREF1061_01543 | 40.67 | 2 | 2 | 1 | | 99642 | pyruvate phosphate dikinase |
| 302 | 93805 | HMPREF1061\|HMPREF1061_02878 | 29.53 | 5 | 1 | 1 | | 21913 | 50S ribosomal protein L3 |
| 201 | 40724 | HMPREF1061\|HMPREF1061_02889 | 34.45 | 16 | 2 | 2 | | 15850 | 50S ribosomal protein L11 |
| 82 | 68436 | HMPREF1061\|HMPREF1061_04079 | 43.34 | 21 | 3 | 3 | | 10676 | 30S ribosomal protein S18 |
| 120 | 24606 | HMPREF1062\|HMPREF1062_02891 | 45.74 | 1 | 3 | 1 | | 129960 | pyruvate:ferredoxin (flavodoxin) oxidoreductase |
| 82 | 68412 | HMPREF1062\|HMPREF1062_04958 | 43.34 | 21 | 3 | 3 | | 10534 | 30S ribosomal protein S18 |
| 201 | 40732 | HMPREF1062\|HMPREF1062_05872 | 34.45 | 16 | 2 | 2 | | 15751 | 50S ribosomal protein L11 |
| 302 | 117824 | HMPREF1062\|HMPREF1062_05883 | 29.53 | 5 | 1 | 1 | | 21869 | 50S ribosomal protein L3 |
| 82 | 68442 | HMPREF1063\|HMPREF1063_01429 | 43.34 | 21 | 3 | 3 | | 10561 | 30S ribosomal protein S18 |
| 321 | 68584 | HMPREF1063\|HMPREF1063_02332 | 26.23 | 4 | 1 | 1 | | 32863 | malate dehydrogenase |
| 201 | 40725 | HMPREF1063\|HMPREF1063_03179 | 34.45 | 16 | 2 | 2 | | 15797 | 50S ribosomal protein L11 |
| 201 | 40710 | HMPREF1064\|HMPREF1064_01947 | 34.45 | 16 | 2 | 2 | | 15797 | 50S ribosomal protein L11 |
| 82 | 68443 | HMPREF1064\|HMPREF1064_02323 | 43.34 | 21 | 3 | 3 | | 10561 | 30S ribosomal protein S18 |
| 321 | 68572 | HMPREF1064\|HMPREF1064_04455 | 26.23 | 4 | 1 | 1 | | 32863 | malate dehydrogenase |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 82 | 68441 | HMPREF1065\|HMPREF1065_01766 | 43.34 | 21 | 3 | 3 | | 10561 | 30S ribosomal protein S18 |
| 201 | 40723 | HMPREF1065\|HMPREF1065_03150 | 34.45 | 16 | 2 | 2 | | 15797 | 50S ribosomal protein L11 |
| 321 | 68576 | HMPREF1065\|HMPREF1065_03498 | 26.23 | 4 | 1 | 1 | | 32863 | malate dehydrogenase |
| 162 | 67902 | HMPREF1069\|HMPREF1069_00261 | 40.67 | 2 | 2 | 1 | | 99673 | pyruvate phosphate dikinase |
| 82 | 68428 | HMPREF1069\|HMPREF1069_03965 | 43.34 | 21 | 3 | 3 | | 10704 | 30S ribosomal protein S18 |
| 49 | 41237 | HMPREF1069\|HMPREF1069_06233 | 53.38 | 12 | 5 | 3 | | 43648 | elongation factor Tu |
| 201 | 40714 | HMPREF1069\|HMPREF1069_06236 | 34.45 | 16 | 2 | 2 | | 15836 | 50S ribosomal protein L11 |
| 302 | 93789 | HMPREF1069\|HMPREF1069_06247 | 29.53 | 5 | 1 | 1 | | 21829 | 50S ribosomal protein L3 |
| 302 | 117808 | HMPREF1070\|HMPREF1070_00052 | 29.53 | 5 | 1 | 1 | | 21874 | 50S ribosomal protein L3 |
| 201 | 40729 | HMPREF1070\|HMPREF1070_00063 | 34.45 | 16 | 2 | 2 | | 15836 | 50S ribosomal protein L11 |
| 162 | 67905 | HMPREF1070\|HMPREF1070_02904 | 40.67 | 2 | 2 | 1 | | 99673 | pyruvate phosphate dikinase |
| 82 | 68426 | HMPREF1070\|HMPREF1070_03731 | 43.34 | 21 | 3 | 3 | | 10704 | 30S ribosomal protein S18 |
| 82 | 68421 | HMPREF1072\|HMPREF1072_00773 | 43.34 | 21 | 3 | 3 | | 10662 | 30S ribosomal protein S18 |
| 321 | 68585 | HMPREF1072\|HMPREF1072_00885 | 26.23 | 4 | 1 | 1 | | 32782 | malate dehydrogenase |
| 302 | 117825 | HMPREF1072\|HMPREF1072_01016 | 29.53 | 5 | 1 | 1 | | 21917 | 50S ribosomal protein L3 |
| 201 | 40718 | HMPREF1072\|HMPREF1072_01027 | 34.45 | 16 | 2 | 2 | | 15867 | 50S ribosomal protein L11 |
| 321 | 68589 | HMPREF1073\|HMPREF1073_00071 | 26.23 | 4 | 1 | 1 | | 32782 | malate dehydrogenase |
| 82 | 68419 | HMPREF1073\|HMPREF1073_00186 | 43.34 | 21 | 3 | 3 | | 10662 | 30S ribosomal protein S18 |
| 201 | 40713 | HMPREF1073\|HMPREF1073_03528 | 34.45 | 16 | 2 | 2 | | 15867 | 50S ribosomal protein L11 |
| 302 | 117830 | HMPREF1073\|HMPREF1073_03539 | 29.53 | 5 | 1 | 1 | | 21917 | 50S ribosomal protein L3 |
| 162 | 55786 | HMPREF1074\|HMPREF1074_01882 | 40.67 | 2 | 2 | 1 | | 99673 | pyruvate phosphate dikinase |
| 302 | 117805 | HMPREF1074\|HMPREF1074_02162 | 29.53 | 5 | 1 | 1 | | 21874 | 50S ribosomal protein L3 |
| 201 | 40719 | HMPREF1074\|HMPREF1074_02173 | 34.45 | 16 | 2 | 2 | | 15836 | 50S ribosomal protein L11 |
| 49 | 67615 | HMPREF1074\|HMPREF1074_02176 | 53.38 | 12 | 5 | 3 | | 43662 | elongation factor Tu |
| 82 | 68434 | HMPREF1074\|HMPREF1074_03552 | 43.34 | 21 | 3 | 3 | | 10676 | 30S ribosomal protein S18 |
| 302 | 117839 | HMPREF1075\|HMPREF1075_00068 | 29.53 | 5 | 1 | 1 | | 22134 | 50S ribosomal protein L3 |
| 82 | 68413 | HMPREF1075\|HMPREF1075_00648 | 43.34 | 21 | 3 | 3 | | 10487 | 30S ribosomal protein S18 |
| 321 | 68574 | HMPREF1075\|HMPREF1075_00808 | 26.23 | 4 | 1 | 1 | | 32848 | malate dehydrogenase |
| 120 | 42048 | HMPREF1075\|HMPREF1075_02324 | 45.74 | 1 | 3 | 1 | | 129315 | pyruvate:ferredoxin (flavodoxin) oxidoreductase |
| 82 | 68425 | HMPREF1077\|HMPREF1077_00467 | 43.34 | 21 | 3 | 3 | | 10530 | 30S ribosomal protein S18 |
| 321 | 68575 | HMPREF1077\|HMPREF1077_00863 | 26.23 | 4 | 1 | 1 | | 32921 | malate dehydrogenase |
| 302 | 117823 | HMPREF1077\|HMPREF1077_02234 | 29.53 | 5 | 1 | 1 | | 22176 | 50S ribosomal protein L3 |
| 162 | 117719 | HMPREF1077\|HMPREF1077_02986 | 40.67 | 2 | 2 | 1 | | 99650 | pyruvate phosphate dikinase |
| 162 | 67901 | HMPREF1078\|HMPREF1078_00713 | 40.67 | 2 | 2 | 1 | | 99765 | pyruvate phosphate dikinase |
| 82 | 68417 | HMPREF1078\|HMPREF1078_02219 | 43.34 | 21 | 3 | 3 | | 10487 | 30S ribosomal protein S18 |
| 321 | 68578 | HMPREF1078\|HMPREF1078_02803 | 26.23 | 4 | 1 | 1 | | 32910 | malate dehydrogenase |
| 302 | 117826 | HMPREF1078\|HMPREF1078_03465 | 29.53 | 5 | 1 | 1 | | 22176 | 50S ribosomal protein L3 |
| 199 | 54008 | HMPREF9003\|HMPREF9003_1005 | 31.44 | 10 | 1 | 1 | | 11550 | ribosomal protein S10 |
| 89 | 26 | HMPREF9457\|HMPREF9457_00280 | 51.44 | 8 | 4 | 1 | | 36196 | glyceraldehyde-3-phosphate dehydrogenase |
| 114 | 40957 | HMPREF9457\|HMPREF9457_00615 | 43.98 | 1 | 3 | 1 | | 128116 | pyruvate:ferredoxin oxidoreductase |
| 129 | 68119 | HMPREF9457\|HMPREF9457_00785 | 41.8 | 7 | 3 | 2 | | 49639 | glucose-6-phosphate isomerase |
| 156 | 24420 | HMPREF9457\|HMPREF9457_01377 | 31.17 | 15 | 2 | 2 | | 14753 | 50S ribosomal protein L11 |
| 51 | 67753 | HMPREF9457\|HMPREF9457_01397 | 57.21 | 12 | 4 | 2 | | 30953 | fructose-1 6-bisphosphate aldolase class II |
| 274 | 68660 | HMPREF9457\|HMPREF9457_01646 | 21.88 | 6 | 1 | 1 | | 15364 | 50S ribosomal protein L15 |
| 94 | 67950 | HMPREF9457\|HMPREF9457_01947 | 50.83 | 5 | 4 | 4 | | 59354 | phosphoenolpyruvate carboxykinase |
| 195 | 117971 | HMPREF9457\|HMPREF9457_03631 | 28.59 | 14 | 2 | 2 | | 7190 | cold shock-like protein cspC |
| 156 | 54312 | HMPREF9473\|HMPREF9473_00630 | 31.17 | 15 | 2 | 2 | | 14818 | 50S ribosomal protein L11 |
| 244 | 68929 | HMPREF9473\|HMPREF9473_02400 | 38.62 | 3 | 1 | 1 | | 55459 | hypothetical protein |
| 265 | 117800 | HMPREF9473\|HMPREF9473_02510 | 26.11 | 2 | 1 | 1 | | 54265 | hypothetical protein |
| 27 | 244 | HMPREF9473\|HMPREF9473_02771 | 65.85 | 9 | 6 | 2 | Oxidation (M) | 62893 | hypothetical protein |
| 274 | 117857 | HMPREF9473\|HMPREF9473_04231 | 21.88 | 6 | 1 | 1 | | 15473 | 50S ribosomal protein L15 |
| 42 | 40691 | HMPREF9473\|HMPREF9473_04739 | 56.92 | 6 | 6 | 2 | | 77041 | formate acetyltransferase |
| 274 | 117852 | HMPREF9474\|HMPREF9474_00421 | 21.88 | 6 | 1 | 1 | | 15456 | 50S ribosomal protein L15 |
| 156 | 5892 | HMPREF9474\|HMPREF9474_03347 | 31.17 | 15 | 2 | 2 | | 14722 | hypothetical protein |
| 126 | 40584 | HMPREF9474\|HMPREF9474_03398 | 53.89 | 12 | 3 | 3 | | 42095 | hypothetical protein |
| 220 | 42930 | HMPREF9474\|HMPREF9474_03417 | 28.51 | 2 | 1 | 1 | | 59528 | hypothetical protein |
| 194 | 42087 | HMPREF9474\|HMPREF9474_04464 | 29.9 | 19 | 2 | 2 | | 10239 | chaperonin |
| 274 | 117858 | HMPREF9475\|HMPREF9475_00358 | 21.88 | 6 | 1 | 1 | | 15456 | 50S ribosomal protein L15 |
| 194 | 42089 | HMPREF9475\|HMPREF9475_02876 | 29.9 | 19 | 2 | 2 | | 10239 | chaperonin |
| 126 | 40585 | HMPREF9475\|HMPREF9475_02938 | 53.89 | 12 | 3 | 3 | | 42095 | butyryl-CoA dehydrogenase |
| 156 | 5891 | HMPREF9475\|HMPREF9475_02989 | 31.17 | 15 | 2 | 2 | | 14722 | ribosomal protein L11 |
| 220 | 42929 | HMPREF9475\|HMPREF9475_04045 | 28.51 | 2 | 1 | 1 | | 59528 | formate-tetrahydrofolate ligase |
| 172 | 53945 | LAR\|LAR_0382 | 37.57 | 4 | 2 | 1 | | 42960 | phosphoglycerate kinase |
| 199 | 54014 | LAR\|LAR_1395 | 31.44 | 10 | 1 | 1 | | 11777 | 30S ribosomal protein S10 |
| 172 | 53943 | LRATCC53608\|LRATCC53608_00377 | 37.57 | 4 | 2 | 1 | | 42961 | phosphoglycerate kinase |
| 199 | 54013 | LRATCC53608\|LRATCC53608_01301 | 31.44 | 10 | 1 | 1 | | 11777 | 30S ribosomal protein S10 |
| 199 | 54018 | Lreu\|Lreu_1484 | 31.44 | 10 | 1 | 1 | | 11777 | 30S ribosomal protein S10 |
| 172 | 53948 | Lreu23DRAFT\|Lreu23DRAFT_3443 | 37.57 | 4 | 2 | 1 | | 42961 | Phosphoglycerate kinase |
| 199 | 54016 | Lreu23DRAFT\|Lreu23DRAFT_4406 | 31.44 | 10 | 1 | 1 | | 11777 | ribosomal protein S10 |
| 199 | 54010 | LRI\|LRI_0490 | 31.44 | 10 | 1 | 1 | | 11702 | 30S ribosomal protein S10 |
| 172 | 53942 | LRI\|LRI_1546 | 37.57 | 4 | 2 | 1 | | 42961 | Phosphoglycerate kinase |
| 274 | 117856 | MITSMUL\|MITSMUL_05081 | 21.88 | 6 | 1 | 1 | | 15636 | ribosomal protein L15 |
| 172 | 53947 | N134\|N134_02060 | 37.57 | 4 | 2 | 1 | | 42975 | phosphoglycerate kinase |
| 199 | 54007 | N134\|N134_08140 | 31.44 | 10 | 1 | 1 | | 11777 | 30S ribosomal protein S10 |
| 157 | 42349 | O08638\|MYH11_MOUSE | 31.31 | 1 | 2 | 1 | | 227026 | Myosin-11 OS=Mus musculus GN=Myh11 PE=1 SV=1 |
| 157 | 42347 | O08638-2\|MYH11_MOUSE | 31.31 | 1 | 2 | 1 | | 223264 | Isoform 2 of Myosin-11 OS=Mus musculus GN=Myh11 |
| 201 | 128 | O08709\|PRDX6_MOUSE | 37.36 | 6 | 2 | 2 | | 24871 | Peroxiredoxin-6 OS=Mus musculus GN=Prdx6 PE=1 SV=3 |
| 180 | 5441 | O70475\|UGDH_MOUSE | 51.15 | 5 | 2 | 2 | | 54832 | UDP-glucose 6-dehydrogenase OS=Mus musculus GN=Ugdh PE=1 SV=1 |
| 165 | 40503 | O70570\|PIGR_MOUSE | 27.99 | 1 | 2 | 2 | | 84999 | Polymeric immunoglobulin receptor OS=Mus musculus GN=Pigr PE=1 SV=1 |
| 65 | 40453 | O88310\|ITL1A_MOUSE | 74.99 | 15 | 5 | 5 | | 34953 | Intelectin-1a OS=Mus musculus GN=Itln1 PE=1 SV=1 |
| 90 | 40796 | O88329\|MYO1A_MOUSE | 53.29 | 4 | 4 | 3 | | 118695 | Unconventional myosin-Ia OS=Mus musculus GN=Myo1a PE=2 SV=2 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 185 | 40862 | OOC\|OOC_01465 | 35.62 | 6 | 2 | 1 | | 48196 | glutamate dehydrogenase |
| 143 | 580 | P00397\|COX1_MOUSE | 36.96 | 2 | 2 | 2 | | 56910 | Cytochrome c oxidase subunit 1 OS=Mus musculus GN=Mtco1 PE=3 SV=2 |
| 12 | 5428 | P00688\|AMYP_MOUSE | 101.86 | 21 | 14 | 14 | Oxidation (M); Pyro-glu from Q | 57318 | Pancreatic alpha-amylase OS=Mus musculus GN=Amy2 PE=1 SV=2 |
| 148 | 5661 | P01592\|IGJ_MOUSE | 36.81 | 12 | 2 | 2 | | 18014 | Immunoglobulin J chain OS=Mus musculus GN=Igj PE=2 SV=4 |
| 226 | 41497 | P01631\|KV2A7_MOUSE | 26.41 | 18 | 2 | 2 | | 12273 | Ig kappa chain V-II region 26-10 OS=Mus musculus PE=1 SV=1 |
| 131 | 71299 | P01644\|KV5AB_MOUSE | 52.7 | 15 | 2 | 2 | | 11910 | Ig kappa chain V-V region HP R16.7 OS=Mus musculus PE=1 SV=1 |
| 131 | 71296 | P01645\|KV5AC_MOUSE | 52.7 | 15 | 2 | 2 | | 11954 | Ig kappa chain V-V region HP 93G7 OS=Mus musculus PE=1 SV=1 |
| 131 | 71298 | P01646\|KV5AD_MOUSE | 52.7 | 15 | 2 | 2 | | 11989 | Ig kappa chain V-V region HP 123E6 OS=Mus musculus PE=1 SV=1 |
| 139 | 40854 | P01756\|HVM12_MOUSE | 55.44 | 16 | 2 | 1 | | 12983 | Ig heavy chain V region MOPC 104E OS=Mus musculus PE=1 SV=1 |
| 139 | 40846 | P01757\|HVM13_MOUSE | 55.44 | 16 | 2 | 1 | | 13025 | Ig heavy chain V region J558 OS=Mus musculus PE=1 SV=1 |
| 139 | 40849 | P01758\|HVM14_MOUSE | 55.44 | 16 | 2 | 1 | | 12972 | Ig heavy chain V region 108A OS=Mus musculus GN=Igh-VJ558 PE=4 SV=1 |
| 73 | 41 | P01837\|IGKC_MOUSE | 63.68 | 24 | 4 | 4 | Pyro-glu from Q | 11778 | Ig kappa chain C region OS=Mus musculus PE=1 SV=1 |
| 100 | 24458 | P02535\|K1C10_MOUSE | 55.03 | 7 | 4 | 4 | | 57770 | Keratin type I cytoskeletal 10 OS=Mus musculus GN=Krt10 PE=1 SV=3 |
| 100 | 24456 | P02535-2\|K1C10_MOUSE | 55.03 | 7 | 4 | 4 | | 57060 | Isoform 2 of Keratin type I cytoskeletal 10 OS=Mus musculus GN=Krt10 |
| 100 | 24450 | P02535-3\|K1C10_MOUSE | 55.03 | 8 | 4 | 4 | | 49501 | Isoform 3 of Keratin type I cytoskeletal 10 OS=Mus musculus GN=Krt10 |
| 103 | 14615 | P03958\|ADA_MOUSE | 48.11 | 7 | 2 | 2 | | 39992 | Adenosine deaminase OS=Mus musculus GN=Ada PE=1 SV=3 |
| 32 | 68 | P04347\|GLYG5_SOYBN | 83.42 | 14 | 8 | 1 | | 57956 | Glycinin OS=Glycine max PE=1 SV=1 |
| 10 | 7 | P04405\|GLYG2_SOYBN | 100.32 | 12 | 12 | 3 | Oxidation (M) | 54391 | Glycinin G2 OS=Glycine max GN=Gy2 PE=1 SV=2 |
| 3 | 3 | P04776\|GLYG1_SOYBN | 113.37 | 24 | 19 | 9 | Oxidation (M) | 55706 | Glycinin G1 OS=Glycine max GN=GY1 PE=1 SV=2 |
| 131 | 71328 | P04946\|KV5AM_MOUSE | 52.7 | 15 | 2 | 2 | | 11892 | Ig kappa chain V-V region NQ5-89.4 OS=Mus musculus PE=2 SV=1 |
| 101 | 700 | P05064\|ALDOA_MOUSE | 52.76 | 7 | 4 | 4 | | 39356 | Fructose-bisphosphate aldolase A OS=Mus musculus GN=Aldoa PE=1 SV=2 |
| 47 | 18900 | P05208\|CEL2A_MOUSE | 60.33 | 12 | 5 | 5 | | 28914 | Chymotrypsin-like elastase family member 2A OS=Mus musculus GN=Cela2a PE=2 SV=1 |
| 139 | 40855 | P06330\|HVM51_MOUSE | 55.44 | 16 | 2 | 1 | | 12934 | Ig heavy chain V region AC38 205.12 OS=Mus musculus PE=1 SV=1 |
| 17 | 11 | P07146\|TRY2_MOUSE | 95.82 | 32 | 11 | 9 | Oxidation (M) | 26204 | Anionic trypsin-2 OS=Mus musculus GN=Prss2 PE=2 SV=1 |
| 69 | 40735 | P07310\|KCRM_MOUSE | 50.16 | 6 | 3 | 3 | | 43045 | Creatine kinase M-type OS=Mus musculus GN=Ckm PE=2 SV=1 |
| 48 | 1546 | P07724\|ALBU_MOUSE | 66.73 | 9 | 5 | 5 | | 68693 | Serum albumin OS=Mus musculus GN=Alb PE=1 SV=3 |
| 204 | 18572 | P07759\|SPA3K_MOUSE | 42.6 | 10 | 2 | 2 | | 46880 | Serine protease inhibitor A3K OS=Mus musculus GN=Serpina3k PE=1 SV=2 |
| 84 | 41491 | P08228\|SODC_MOUSE | 62.98 | 23 | 4 | 4 | | 15943 | Superoxide dismutase [Cu-Zn] OS=Mus musculus GN=Sod1 PE=1 SV=2 |
| 40 | 96 | P09470\|ACE_MOUSE | 86.65 | 5 | 7 | 7 | | 150918 | Angiotensin-converting enzyme OS=Mus musculus GN=Ace PE=1 SV=3 |
| 136 | 239 | P10126\|EF1A1_MOUSE | 37.76 | 5 | 2 | 2 | | 50114 | Elongation factor 1-alpha 1 OS=Mus musculus GN=Eef1a1 PE=1 SV=3 |
| 28 | 144 | P11828\|GLYG3_SOYBN | 79.97 | 11 | 7 | 5 | Oxidation (M) | 54242 | Glycinin G3 OS=Glycine max GN=GY3 PE=3 SV=1 |
| 67 | 40471 | P13541\|MYH3_MOUSE | 48.68 | 2 | 5 | 2 | | 223789 | Myosin-3 OS=Mus musculus GN=Myh3 PE=2 SV=2 |
| 22 | 4 | P13634\|CAH1_MOUSE | 115.43 | 36 | 9 | 9 | Acetylation (N-term) | 28331 | Carbonic anhydrase 1 OS=Mus musculus GN=Ca1 PE=2 SV=4 |
| 14 | 18424 | P15947\|KLK1_MOUSE | 77.21 | 20 | 8 | 7 | Oxidation (M) | 28775 | Kallikrein-1 OS=Mus musculus GN=Klk1 PE=1 SV=3 |
| 102 | 18497 | P15949\|K1KB9_MOUSE | 36.42 | 8 | 2 | 1 | | 28900 | Kallikrein 1-related peptidase b9 OS=Mus musculus GN=Klk1b9 PE=2 SV=1 |
| 14 | 18582 | P16406\|AMPE_MOUSE | 91.84 | 12 | 12 | 12 | Oxidation (M) | 107956 | Glutamyl aminopeptidase OS=Mus musculus GN=Enpep PE=1 SV=1 |
| 108 | 252 | P16858\|G3P_MOUSE | 41.29 | 8 | 3 | 2 | | 35810 | Glyceraldehyde-3-phosphate dehydrogenase OS=Mus musculus GN=Gapdh PE=1 SV=2 |
| 21 | 5450 | P17563\|SBP1_MOUSE | 87.98 | 16 | 10 | 10 | Pyro-glu from Q | 52514 | Selenium-binding protein 1 OS=Mus musculus GN=Selenbp1 PE=1 SV=2 |
| 81 | 40836 | P17751\|TPIS_MOUSE | 50.12 | 10 | 4 | 4 | | 32192 | Triosephosphate isomerase OS=Mus musculus GN=Tpi1 PE=1 SV=4 |
| 26 | 40455 | P17892\|LIPR2_MOUSE | 62.72 | 6 | 3 | 3 | | 52585 | Pancreatic lipase-related protein 2 OS=Mus musculus GN=Pnliprp2 PE=2 SV=1 |
| 28 | 40472 | P18761\|CAH6_MOUSE | 52.69 | 11 | 4 | 4 | | 36495 | Carbonic anhydrase 6 OS=Mus musculus GN=Ca6 PE=1 SV=3 |
| 125 | 9 | P19467\|MUC13_MOUSE | 38.57 | 2 | 2 | 2 | | 58701 | Mucin-13 OS=Mus musculus GN=Muc13 PE=2 SV=1 |
| 110 | 7118 | P21107\|TPM3_MOUSE | 58.88 | 12 | 4 | 1 | | 32994 | Tropomyosin alpha-3 chain OS=Mus musculus GN=Tpm3 PE=2 SV=3 |
| 112 | 16 | P22752\|H2A1_MOUSE | 46.64 | 22 | 1 | 1 | | 14135 | Histone H2A type 1 OS=Mus musculus GN=Hist1h2ab PE=1 SV=3 |
| 112 | 18 | P27661\|H2AX_MOUSE | 46.64 | 20 | 1 | 1 | | 15143 | Histone H2AX OS=Mus musculus GN=H2afx PE=1 SV=2 |
| 52 | 40781 | P28825\|MEP1A_MOUSE | 63.2 | 5 | 4 | 4 | | 84231 | Meprin A subunit alpha OS=Mus musculus GN=Mep1a PE=1 SV=4 |
| 50 | 18552 | P28843\|DPP4_MOUSE | 61.24 | 6 | 5 | 5 | | 87437 | Dipeptidyl peptidase 4 OS=Mus musculus GN=Dpp4 PE=1 SV=3 |

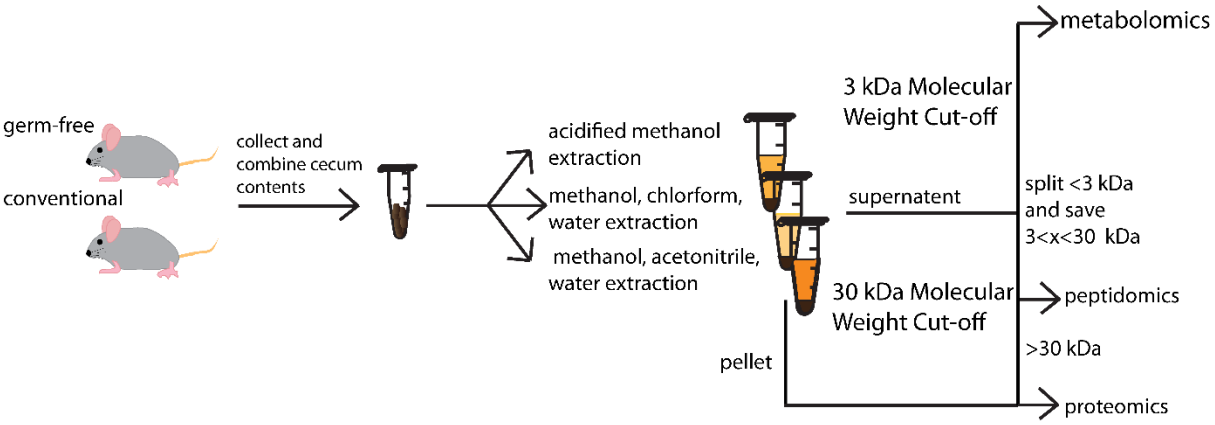
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 78 | 481 | P30275\|KCRU_MOUSE | 64.22 | 11 | 5 | 5 | | 47004 | Creatine kinase U-type mitochondrial OS=Mus musculus GN=Ckmt1 PE=1 SV=1 |
| 75 | 40450 | P31428\|DPEP1_MOUSE | 63.73 | 6 | 3 | 3 | | 45722 | Dipeptidase 1 OS=Mus musculus GN=Dpep1 PE=1 SV=2 |
| 139 | 62 | P31809\|CEAM1_MOUSE | 38.51 | 4 | 2 | 2 | | 57016 | Carcinoembryonic antigen-related cell adhesion molecule 1 OS=Mus musculus GN=Ceacam1 PE=1 SV=1 |
| 139 | 60 | P31809-2\|CEAM1_MOUSE | 38.51 | 5 | 2 | 2 | | 50057 | Isoform Short of Carcinoembryonic antigen-related cell adhesion molecule 1 OS=Mus musculus GN=Ceacam1 |
| 211 | 54233 | P32848\|PRVA_MOUSE | 31.05 | 7 | 1 | 1 | | 11931 | Parvalbumin alpha OS=Mus musculus GN=Pvalb PE=1 SV=3 |
| 61 | 18651 | P51881\|ADT2_MOUSE | 55.5 | 10 | 4 | 4 | | 32931 | ADP/ATP translocase 2 OS=Mus musculus GN=Slc25a5 PE=1 SV=3 |
| 197 | 24333 | P55050\|FABPI_MOUSE | 45.17 | 13 | 2 | 2 | | 15126 | Fatty acid-binding protein intestinal OS=Mus musculus GN=Fabp2 PE=2 SV=2 |
| 99 | 40645 | P56480\|ATPB_MOUSE | 52.18 | 4 | 2 | 2 | | 56301 | ATP synthase subunit beta mitochondrial OS=Mus musculus GN=Atp5b PE=1 SV=2 |
| 373 | 24328 | P58242\|ASM3B_MOUSE | 44.56 | 4 | 1 | 1 | | 51600 | Acid sphingomyelinase-like phosphodiesterase 3b OS=Mus musculus GN=Smpdl3b PE=1 SV=1 |
| 135 | 40794 | P62204\|CALM_MOUSE | 40.76 | 17 | 3 | 3 | | 16838 | Calmodulin OS=Mus musculus GN=Calm1 PE=1 SV=2 |
| 8 | 47 | P62737\|ACTA_MOUSE | 94.34 | 28 | 16 | 16 | Oxidation (M) | 42009 | Actin aortic smooth muscle OS=Mus musculus GN=Acta2 PE=1 SV=1 |
| 140 | 18444 | P62806\|H4_MOUSE | 32.55 | 8 | 1 | 1 | | 11367 | Histone H4 OS=Mus musculus GN=Hist1h4a PE=1 SV=2 |
| 8 | 46 | P63268\|ACTH_MOUSE | 94.34 | 28 | 16 | 16 | Oxidation (M) | 41877 | Actin gamma-enteric smooth muscle OS=Mus musculus GN=Actg2 PE=1 SV=1 |
| 8 | 48 | P68033\|ACTC_MOUSE | 94.34 | 28 | 16 | 16 | Oxidation (M) | 42019 | Actin alpha cardiac muscle 1 OS=Mus musculus GN=Actc1 PE=1 SV=1 |
| 22 | 150 | P70412\|CUZD1_MOUSE | 84.37 | 13 | 8 | 8 | | 68082 | CUB and zona pellucida-like domain-containing protein 1 OS=Mus musculus GN=Cuzd1 PE=2 SV=2 |
| 48 | 5454 | P97429\|ANXA4_MOUSE | 56.19 | 14 | 5 | 5 | | 35916 | Annexin A4 OS=Mus musculus GN=Anxa4 PE=2 SV=4 |
| 7 | 5 | P97449\|AMPN_MOUSE | 120.45 | 17 | 18 | 18 | Oxidation (M) | 109651 | Aminopeptidase N OS=Mus musculus GN=Anpep PE=1 SV=4 |
| 82 | 68410 | PARMER\|PARMER_01319 | 43.34 | 21 | 3 | 3 | | 10487 | hypothetical protein |
| 321 | 68591 | PARMER\|PARMER_03587 | 26.23 | 4 | 1 | 1 | | 32910 | hypothetical protein |
| 302 | 117832 | PARMER\|PARMER_03735 | 29.53 | 5 | 1 | 1 | | 22176 | hypothetical protein |
| 162 | 67900 | PARMER\|PARMER_04399 | 40.67 | 2 | 2 | 1 | | 99680 | hypothetical protein |
| 321 | 68590 | PRABACTJOHN\|PRABACTJOHN_03033 | 26.23 | 4 | 1 | 1 | | 32921 | hypothetical protein |
| 185 | 40566 | PROSTU\|PROSTU_04260 | 35.62 | 6 | 2 | 1 | | 47981 | hypothetical protein |
| 185 | 40863 | PROVRETT\|PROVRETT_08775 | 35.62 | 6 | 2 | 1 | | 48196 | NADP-specific glutamate dehydrogenase |
| 185 | 40567 | PROVRUST\|PROVRUST_05924 | 35.62 | 6 | 2 | 1 | | 48216 | NADP-specific glutamate dehydrogenase |
| 52 | 116 | Q00897\|A1AT4_MOUSE | 58.13 | 10 | 7 | 3 | | 45998 | Alpha-1-antitrypsin 1-4 OS=Mus musculus GN=Serpina1d PE=2 SV=1 |
| 314 | 40663 | Q02013\|AQP1_MOUSE | 28.14 | 3 | 1 | 1 | | 28793 | Aquaporin-1 OS=Mus musculus GN=Aqp1 PE=1 SV=3 |
| 66 | 40451 | Q02566\|MYH6_MOUSE | 64.53 | 2 | 4 | 2 | | 223563 | Myosin-6 OS=Mus musculus GN=Myh6 PE=1 SV=2 |
| 265 | 117798 | Q06679\|UTP4_YEAST | 26.11 | 1 | 1 | 1 | | 87801 | U3 small nucleolar RNA-associated protein 4 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=UTP4 PE=1 SV=1 |
| 80 | 68650 | Q11136\|PEPD_MOUSE | 67.54 | 9 | 4 | 4 | Acetylation (N-term) | 55029 | Xaa-Pro dipeptidase OS=Mus musculus GN=Pepd PE=2 SV=3 |
| 130 | 93085 | Q3TTY0\|PLB1_MOUSE | 41.16 | 1 | 2 | 2 | | 164540 | Phospholipase B1 membrane-associated OS=Mus musculus GN=Plb1 PE=2 SV=2 |
| 130 | 93087 | Q3TTY0-2\|PLB1_MOUSE | 41.16 | 2 | 2 | 2 | | 86140 | Isoform 2 of Phospholipase B1 membrane-associated OS=Mus musculus GN=Plb1 |
| 130 | 93083 | Q3TTY0-3\|PLB1_MOUSE | 41.16 | 2 | 2 | 2 | | 102708 | Isoform 3 of Phospholipase B1 membrane-associated OS=Mus musculus GN=Plb1 |
| 174 | 71101 | Q5SYD0\|MYO1D_MOUSE | 35.55 | 2 | 2 | 1 | | 116081 | Unconventional myosin-Id OS=Mus musculus GN=Myo1d PE=1 SV=1 |
| 174 | 71100 | Q5SYD0-2\|MYO1D_MOUSE | 35.55 | 2 | 2 | 1 | | 109278 | Isoform 2 of Unconventional myosin-Id OS=Mus musculus GN=Myo1d |
| 137 | 18524 | Q60854\|SPB6_MOUSE | 50.58 | 14 | 3 | 3 | Pyro-glu from Q | 42599 | Serpin B6 OS=Mus musculus GN=Serpinb6 PE=1 SV=1 |
| 43 | 40539 | Q60930\|VDAC2_MOUSE | 66.62 | 11 | 5 | 4 | | 31733 | Voltage-dependent anion-selective channel protein 2 OS=Mus musculus GN=Vdac2 PE=1 SV=2 |
| 111 | 40474 | Q60931\|VDAC3_MOUSE | 51.34 | 10 | 3 | 2 | | 30753 | Voltage-dependent anion-selective channel protein 3 OS=Mus musculus GN=Vdac3 PE=1 SV=1 |
| 127 | 30060 | Q60932\|VDAC1_MOUSE | 46.51 | 9 | 2 | 2 | | 32351 | Voltage-dependent anion-selective channel protein 1 OS=Mus musculus GN=Vdac1 PE=1 SV=3 |
| 127 | 30059 | Q60932-2\|VDAC1_MOUSE | 46.51 | 10 | 2 | 2 | | 30756 | Isoform Mt-VDAC1 of Voltage-dependent anion-selective channel protein 1 OS=Mus musculus GN=Vdac1 |
| 2 | 504 | Q60997\|DMBT1_MOUSE | 147.57 | 8 | 26 | 26 | Oxidation (M) | 226814 | Deleted in malignant brain tumors 1 protein OS=Mus musculus GN=Dmbt1 PE=1 SV=2 |
| 2 | 503 | Q60997-2\|DMBT1_MOUSE | 147.57 | 8 | 26 | 26 | Oxidation (M) | 221163 | Isoform 2 of Deleted in malignant brain tumors 1 protein OS=Mus musculus GN=Dmbt1 |
| 187 | 40671 | Q61207\|SAP_MOUSE | 32.73 | 2 | 1 | 1 | | 61422 | Prosaposin OS=Mus musculus GN=Psap PE=1 SV=2 |
| 98 | 15029 | Q61391\|NEP_MOUSE | 58.85 | 4 | 3 | 3 | | 85702 | Neprilysin OS=Mus musculus GN=Mme PE=1 SV=3 |
| 20 | 136 | Q61847\|MEP1B_MOUSE | 60.66 | 11 | 9 | 9 | Oxidation (M) | 79500 | Meprin A subunit beta OS=Mus musculus GN=Mep1b PE=1 SV=2 |
| 20 | 137 | Q61847-2\|MEP1B_MOUSE | 60.66 | 10 | 9 | 9 | Oxidation (M) | 80348 | Isoform 2 of Meprin A subunit beta OS=Mus musculus GN=Mep1b |
| 133 | 328 | Q61900\|SMR3A_MOUSE | 31.73 | 11 | 1 | 1 | | 15544 | Submaxillary gland androgen-regulated protein 3A OS=Mus musculus GN=Smr3a PE=2 SV=1 |
| 71 | 95 | Q62468\|VILI_MOUSE | 71.79 | 6 | 5 | 5 | | 92775 | Villin-1 OS=Mus musculus GN=Vil1 PE=1 SV=3 |
| 70 | 146 | Q6DYE8\|ENPP3_MOUSE | 44.98 | 6 | 5 | 5 | | 98662 | Ectonucleotide pyrophosphatase/phosphodiesterase family member 3 OS=Mus musculus GN=Enpp3 PE=1 SV=2 |
| 208 | 40736 | Q6IFZ9\|K2C74_MOUSE | 32.72 | 4 | 2 | 2 | | 54747 | Keratin type II cytoskeletal 74 OS=Mus musculus GN=Krt74 PE=3 SV=1 |
| 208 | 18930 | Q6NXH9\|K2C73_MOUSE | 32.72 | 4 | 2 | 2 | | 58911 | Keratin type II cytoskeletal 73 OS=Mus musculus GN=Krt73 PE=1 SV=1 |
| 24 | 84 | Q6P8U6\|LIPP_MOUSE | 85.25 | 18 | 8 | 8 | Oxidation (M); Amidatio | 51428 | Pancreatic triacylglycerol lipase OS=Mus musculus GN=Pnlip PE=1 SV=1 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | n; Acetylation (K) | | |
| 62 | 86 | Q6Q473\|CLA4A_MOUSE | 60.93 | 6 | 6 | 2 | | 101872 | Calcium-activated chloride channel regulator 4A OS=Mus musculus GN=Clca4a PE=1 SV=2 |
| 183 | 68455 | Q6ZQM8\|UD17C_MOUSE | 43.61 | 4 | 2 | 2 | | 59758 | UDP-glucuronosyltransferase 1-7C OS=Mus musculus GN=Ugt1a7c PE=2 SV=1 |
| 13 | 50 | Q7M758\|NALDL_MOUSE | 93.39 | 10 | 12 | 12 | | 80513 | N-acetylated-alpha-linked acidic dipeptidase-like protein OS=Mus musculus GN=Naaladl1 PE=2 SV=1 |
| 15 | 18425 | Q7TPZ8\|CBPA1_MOUSE | 83.67 | 16 | 9 | 9 | | 47385 | Carboxypeptidase A1 OS=Mus musculus GN=Cpa1 PE=2 SV=1 |
| 20 | 40456 | Q80Z19\|MUC2_MOUSE | 74.3 | 3 | 9 | 9 | | 293435 | Mucin-2 (Fragments) OS=Mus musculus GN=Muc2 PE=1 SV=2 |
| 112 | 15 | Q8BFU2\|H2A3_MOUSE | 46.64 | 22 | 1 | 1 | | 14121 | Histone H2A type 3 OS=Mus musculus GN=Hist3h2a PE=1 SV=3 |
| 377 | 18551 | Q8BHH8\|CLRN3_MOUSE | 31.62 | 7 | 1 | 1 | | 25350 | Clarin-3 OS=Mus musculus GN=Clrn3 PE=2 SV=1 |
| 112 | 17 | Q8CGP5\|H2A1F_MOUSE | 46.64 | 22 | 1 | 1 | | 14162 | Histone H2A type 1-F OS=Mus musculus GN=Hist1h2af PE=1 SV=3 |
| 112 | 12 | Q8CGP6\|H2A1H_MOUSE | 46.64 | 23 | 1 | 1 | | 13950 | Histone H2A type 1-H OS=Mus musculus GN=Hist1h2ah PE=1 SV=3 |
| 112 | 14 | Q8CGP7\|H2A1K_MOUSE | 46.64 | 22 | 1 | 1 | | 14150 | Histone H2A type 1-K OS=Mus musculus GN=Hist1h2ak PE=1 SV=3 |
| 34 | 40445 | Q8K0C5\|ZG16_MOUSE | 88.95 | 19 | 6 | 6 | | 18210 | Zymogen granule membrane protein 16 OS=Mus musculus GN=Zg16 PE=2 SV=1 |
| 12 | 18561 | Q8K419\|LEG4_MOUSE | 100.93 | 29 | 11 | 11 | Oxidation (M) | 36372 | Galectin-4 OS=Mus musculus GN=Lgals4 PE=1 SV=2 |
| 19 | 423 | Q8R0I0\|ACE2_MOUSE | 91.27 | 11 | 10 | 10 | | 92368 | Angiotensin-converting enzyme 2 OS=Mus musculus GN=Ace2 PE=1 SV=1 |
| 112 | 13 | Q8R1M2\|H2AJ_MOUSE | 46.64 | 22 | 1 | 1 | | 14045 | Histone H2A.J OS=Mus musculus GN=H2afj PE=1 SV=1 |
| 260 | 40548 | Q8R3G9\|TSN8_MOUSE | 36.91 | 4 | 1 | 1 | | 25582 | Tetraspanin-8 OS=Mus musculus GN=Tspan8 PE=2 SV=1 |
| 74 | 40468 | Q8R429\|AT2A1_MOUSE | 51.14 | 3 | 3 | 3 | | 109425 | Sarcoplasmic/endoplasmic reticulum calcium ATPase 1 OS=Mus musculus GN=Atp2a1 PE=1 SV=1 |
| 46 | 40867 | Q8VDN2\|AT1A1_MOUSE | 66.47 | 6 | 5 | 5 | | 112982 | Sodium/potassium-transporting ATPase subunit alpha-1 OS=Mus musculus GN=Atp1a1 PE=1 SV=1 |
| 93 | 661 | Q91WL0\|ES8L3_MOUSE | 58.93 | 6 | 4 | 4 | | 68216 | Epidermal growth factor receptor kinase substrate 8-like protein 3 OS=Mus musculus GN=Eps8l3 PE=2 SV=1 |
| 168 | 40465 | Q91WV7\|SLC31_MOUSE | 43.12 | 4 | 2 | 2 | Oxidation (M) | 78118 | Neutral and basic amino acid transport protein rBAT OS=Mus musculus GN=Slc3a1 PE=1 SV=1 |
| 60 | 40448 | Q91X79\|CELA1_MOUSE | 62.9 | 17 | 5 | 4 | | 28901 | Chymotrypsin-like elastase family member 1 OS=Mus musculus GN=Cela1 PE=2 SV=1 |
| 5 | 5491 | Q91XA9\|CHIA_MOUSE | 151.05 | 20 | 17 | 17 | | 52003 | Acidic mammalian chitinase OS=Mus musculus GN=Chia PE=1 SV=2 |
| 198 | 40446 | Q99JG3\|ANX13_MOUSE | 50.23 | 6 | 2 | 2 | | 35922 | Annexin A13 OS=Mus musculus GN=Anxa13 PE=2 SV=3 |
| 18 | 5439 | Q9CQ52\|CEL3B_MOUSE | 74.66 | 14 | 7 | 6 | | 28905 | Chymotrypsin-like elastase family member 3B OS=Mus musculus GN=Cela3b PE=2 SV=1 |
| 342 | 18512 | Q9CQQ7\|AT5F1_MOUSE | 26.9 | 4 | 1 | 1 | | 28949 | ATP synthase F(0) complex subunit B1 mitochondrial OS=Mus musculus GN=Atp5f1 PE=1 SV=1 |
| 144 | 93278 | Q9CZU6\|CISY_MOUSE | 26.07 | 2 | 1 | 1 | | 51737 | Citrate synthase mitochondrial OS=Mus musculus GN=Cs PE=1 SV=1 |
| 91 | 10765 | Q9D733\|GP2_MOUSE | 41.91 | 8 | 4 | 4 | | 59154 | Pancreatic secretory granule membrane major glycoprotein GP2 OS=Mus musculus GN=Gp2 PE=2 SV=3 |
| 41 | 40642 | Q9D7Z6\|CLCA1_MOUSE | 73.8 | 7 | 6 | 6 | | 100071 | Calcium-activated chloride channel regulator 1 OS=Mus musculus GN=Clca1 PE=1 SV=2 |
| 73 | 40447 | Q9JHE3\|ASAH2_MOUSE | 63.15 | 5 | 3 | 3 | | 83509 | Neutral ceramidase OS=Mus musculus GN=Asah2 PE=1 SV=1 |
| 136 | 40638 | Q9JIP7\|S15A1_MOUSE | 48.5 | 5 | 3 | 3 | | 78560 | Solute carrier family 15 member 1 OS=Mus musculus GN=Slc15a1 PE=1 SV=2 |
| 125 | 18854 | Q9JKA5\|GPA33_MOUSE | 59.76 | 11 | 3 | 3 | | 35692 | Cell surface A33 antigen OS=Mus musculus GN=Gpa33 PE=2 SV=2 |
| 87 | 40459 | Q9JLB4\|CUBN_MOUSE | 80.13 | 1 | 4 | 4 | | 399099 | Cubilin OS=Mus musculus GN=Cubn PE=1 SV=3 |
| 208 | 40694 | Q9R0H5\|K2C71_MOUSE | 32.72 | 4 | 2 | 2 | | 57383 | Keratin type II cytoskeletal 71 OS=Mus musculus GN=Krt71 PE=1 SV=1 |
| 13 | 24341 | Q9R100\|CAD17_MOUSE | 95.74 | 15 | 14 | 14 | | 91645 | Cadherin-17 OS=Mus musculus GN=Cdh17 PE=1 SV=1 |
| 63 | 40529 | Q9WUB3\|PYGM_MOUSE | 67.81 | 6 | 6 | 4 | | 97286 | Glycogen phosphorylase muscle form OS=Mus musculus GN=Pygm PE=1 SV=3 |
| 335 | 676 | Q9WVC8\|S26A3_MOUSE | 21.43 | 1 | 1 | 1 | | 83590 | Chloride anion exchanger OS=Mus musculus GN=Slc26a3 PE=1 SV=1 |
| 111 | 18507 | Q9Z2W0\|DNPEP_MOUSE | 55.07 | 4 | 2 | 2 | | 52207 | Aspartyl aminopeptidase OS=Mus musculus GN=Dnpep PE=2 SV=2 |
| 274 | 117861 | RO1\|RO1_02120 | 21.88 | 6 | 1 | 1 | | 15324 | LSU ribosomal protein L15P |
| 156 | 5894 | RO1\|RO1_20200 | 31.17 | 15 | 2 | 2 | | 14819 | LSU ribosomal protein L11P |
| 315 | 93226 | RO1\|RO1_22900 | 22.58 | 9 | 1 | 1 | | 13321 | ribosomal protein L19 bacterial type |
| 156 | 54339 | ROI\|ROI_07570 | 31.17 | 15 | 2 | 2 | | 14819 | LSU ribosomal protein L11P |
| 315 | 93227 | ROI\|ROI_10680 | 22.58 | 9 | 1 | 1 | | 13321 | ribosomal protein L19 bacterial type |
| 274 | 117850 | ROI\|ROI_14140 | 21.88 | 6 | 1 | 1 | | 15324 | LSU ribosomal protein L15P |
| 156 | 24422 | ROSINTL182\|ROSINTL182_05591 | 31.17 | 15 | 2 | 2 | | 14819 | ribosomal protein L11 |
| 200 | 41087 | ROSINTL182\|ROSINTL182_06238 | 37.75 | 16 | 2 | 2 | | 14550 | ribosomal protein S9 |
| 274 | 19023 | ROSINTL182\|ROSINTL182_07785 | 21.88 | 6 | 1 | 1 | | 15324 | ribosomal protein L15 |
| 315 | 93225 | ROSINTL182\|ROSINTL182_08453 | 22.58 | 9 | 1 | 1 | | 13321 | ribosomal protein L19 |
| 200 | 41084 | RUMGNA\|RUMGNA_00791 | 37.75 | 16 | 2 | 2 | | 14452 | hypothetical protein |
| 94 | 67946 | RUMGNA\|RUMGNA_01556 | 50.83 | 5 | 4 | 4 | | 59421 | hypothetical protein |
| 156 | 24421 | RUMGNA\|RUMGNA_01561 | 31.17 | 15 | 2 | 2 | | 14868 | hypothetical protein |
| 117 | 40927 | RUMGNA\|RUMGNA_01567 | 58.46 | 2 | 3 | 1 | | 136234 | hypothetical protein |
| 93 | 263 | RUMGNA\|RUMGNA_01756 | 46.22 | 8 | 4 | 1 | | 35969 | hypothetical protein |
| 274 | 93480 | RUMGNA\|RUMGNA_02663 | 21.88 | 6 | 1 | 1 | | 15443 | hypothetical protein |
| 210 | 31035 | RUMGNA\|RUMGNA_03296 | 31.4 | 1 | 1 | 1 | | 76545 | hypothetical protein |
| 308 | 70209 | RUMHYD\|RUMHYD_00687 | 28.53 | 1 | 1 | 1 | | 97229 | hypothetical protein |
| 156 | 54308 | RUMHYD\|RUMHYD_01568 | 31.17 | 15 | 2 | 2 | | 14692 | hypothetical protein |
| 274 | 117859 | RUMHYD\|RUMHYD_02798 | 21.88 | 6 | 1 | 1 | | 15377 | hypothetical protein |

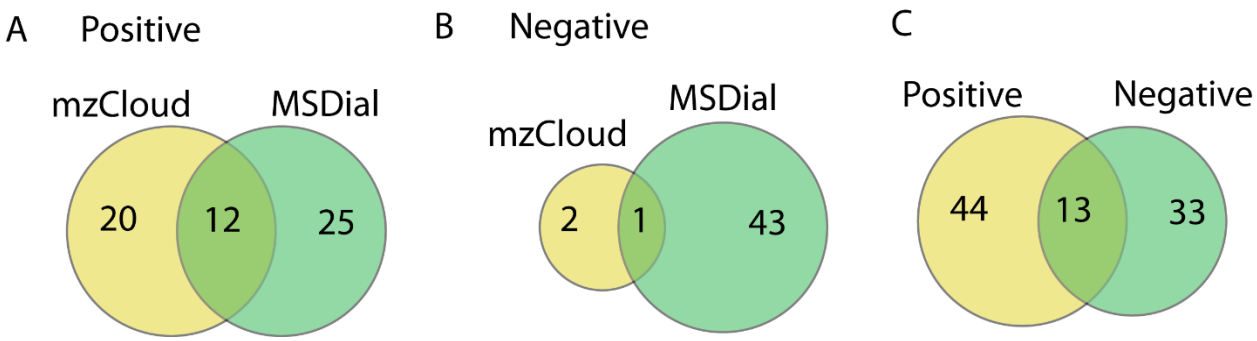
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 174 | 41071 | RUMHYD\|RUMHYD_03721 | 32.47 | 3 | 2 | 1 | | 49256 | hypothetical protein |
| 89 | 21 | RUMHYD\|RUMHYD_03911 | 51.44 | 8 | 4 | 1 | | 33671 | hypothetical protein |
| 220 | 42926 | RUMHYD\|RUMHYD_03927 | 28.51 | 2 | 1 | 1 | | 57423 | hypothetical protein |
| 200 | 41085 | RUMLAC\|RUMLAC_00318 | 37.75 | 16 | 2 | 2 | | 14392 | hypothetical protein |
| 51 | 67708 | RUMLAC\|RUMLAC_01363 | 57.21 | 12 | 4 | 2 | | 31643 | hypothetical protein |
| 114 | 31429 | RUMLAC\|RUMLAC_01487 | 43.98 | 1 | 3 | 1 | | 128904 | hypothetical protein |
| 274 | 93479 | RUMLAC\|RUMLAC_02570 | 21.88 | 6 | 1 | 1 | | 15490 | hypothetical protein |
| 156 | 54313 | RUMLAC\|RUMLAC_02706 | 31.17 | 15 | 2 | 2 | | 14906 | hypothetical protein |
| 89 | 36 | RUMOBE\|RUMOBE_00829 | 51.44 | 8 | 4 | 1 | | 36156 | hypothetical protein |
| 236 | 54103 | RUMOBE\|RUMOBE_01150 | 30.31 | 1 | 1 | 1 | | 76919 | hypothetical protein |
| 220 | 42933 | RUMOBE\|RUMOBE_03002 | 28.51 | 2 | 1 | 1 | | 60185 | hypothetical protein |
| 274 | 70550 | RUMOBE\|RUMOBE_03290 | 21.88 | 6 | 1 | 1 | | 15631 | hypothetical protein |
| 220 | 42925 | RUMOBE\|RUMOBE_04041 | 28.51 | 3 | 1 | 1 | | 45850 | hypothetical protein |
| 89 | 267 | RUMTOR\|RUMTOR_00717 | 51.44 | 8 | 4 | 1 | | 35968 | hypothetical protein |
| 274 | 93475 | RUMTOR\|RUMTOR_02480 | 21.88 | 6 | 1 | 1 | | 15522 | hypothetical protein |
| 200 | 41083 | RUMTOR\|RUMTOR_02559 | 37.75 | 16 | 2 | 2 | | 14374 | hypothetical protein |
| 185 | 53934 | S70\|S70_09995 | 35.62 | 6 | 2 | 1 | | 47981 | glutamate dehydrogenase |
| 226 | 41501 | tr\|A0A075B5K8\|A0A075B5K8_MOUSE | 26.41 | 17 | 2 | 2 | | 13144 | Protein Igkv1-99 OS=Mus musculus GN=Igkv1-99 PE=4 SV=4 |
| 131 | 71301 | tr\|A0A075B5L1\|A0A075B5L1_MOUSE | 52.7 | 14 | 2 | 2 | | 12673 | Protein Igkv10-94 OS=Mus musculus GN=Igkv10-94 PE=4 SV=4 |
| 73 | 42 | tr\|A0A075B5P2\|A0A075B5P2_MOUSE | 63.68 | 23 | 4 | 4 | Pyro-glu from Q | 11897 | Protein Igkc OS=Mus musculus GN=Igkc PE=1 SV=1 |
| 139 | 40842 | tr\|A0A075B5U4\|A0A075B5U4_MOUSE | 55.44 | 19 | 2 | 1 | | 11295 | Protein Ighv1-18 OS=Mus musculus GN=Ighv1-18 PE=4 SV=1 |
| 139 | 40850 | tr\|A0A075B5U7\|A0A075B5U7_MOUSE | 55.44 | 16 | 2 | 1 | | 13022 | Protein Ighv1-22 (Fragment) OS=Mus musculus GN=Ighv1-22 PE=4 SV=1 |
| 139 | 40851 | tr\|A0A075B5V0\|A0A075B5V0_MOUSE | 55.44 | 16 | 2 | 1 | | 12911 | MCG114299 (Fragment) OS=Mus musculus GN=Ighv1-26 PE=4 SV=1 |
| 139 | 40840 | tr\|A0A075B5V1\|A0A075B5V1_MOUSE | 55.44 | 19 | 2 | 1 | | 10966 | Protein Ighv1-31 OS=Mus musculus GN=Ighv1-31 PE=4 SV=1 |
| 139 | 40852 | tr\|A0A075B5V2\|A0A075B5V2_MOUSE | 55.44 | 16 | 2 | 1 | | 13032 | Protein Ighv1-34 (Fragment) OS=Mus musculus GN=Ighv1-34 PE=4 SV=1 |
| 139 | 40844 | tr\|A0A075B5V4\|A0A075B5V4_MOUSE | 55.44 | 16 | 2 | 1 | | 13069 | Protein Ighv1-37 OS=Mus musculus GN=Ighv1-37 PE=4 SV=1 |
| 139 | 40841 | tr\|A0A075B5V6\|A0A075B5V6_MOUSE | 55.44 | 19 | 2 | 1 | | 10776 | Protein Ighv1-42 OS=Mus musculus GN=Ighv1-42 PE=4 SV=1 |
| 171 | 40901 | tr\|A0A075B5V8\|A0A075B5V8_MOUSE | 45.4 | 19 | 2 | 2 | Pyro-glu from Q | 11083 | Protein Ighv1-47 OS=Mus musculus GN=Ighv1-47 PE=4 SV=1 |
| 188 | 40858 | tr\|A0A075B5X5\|A0A075B5X5_MOUSE | 51.81 | 16 | 2 | 1 | Pyro-glu from Q | 12922 | Protein Ighv1-66 OS=Mus musculus GN=Ighv1-66 PE=4 SV=1 |
| 188 | 40857 | tr\|A0A075B5Y2\|A0A075B5Y2_MOUSE | 51.81 | 16 | 2 | 1 | Pyro-glu from Q | 12957 | Protein Ighv1-75 OS=Mus musculus GN=Ighv1-75 PE=4 SV=1 |
| 171 | 40921 | tr\|A0A075B5Y3\|A0A075B5Y3_MOUSE | 45.4 | 16 | 2 | 2 | Pyro-glu from Q | 12763 | Protein Ighv1-80 OS=Mus musculus GN=Ighv1-80 PE=4 SV=1 |
| 188 | 40833 | tr\|A0A075B5Y6\|A0A075B5Y6_MOUSE | 51.81 | 16 | 2 | 1 | Pyro-glu from Q | 13012 | Protein Ighv1-85 OS=Mus musculus GN=Ighv1-85 PE=4 SV=1 |
| 171 | 40920 | tr\|A0A075B680\|A0A075B680_MOUSE | 45.4 | 19 | 2 | 2 | Pyro-glu from Q | 11225 | Protein Ighv1-62-2 OS=Mus musculus GN=Ighv1-62-2 PE=4 SV=1 |
| 45 | 53 | tr\|A0A075B6A3\|A0A075B6A3_MOUSE | 56.17 | 11 | 5 | 5 | | 36838 | Protein Igha (Fragment) OS=Mus musculus GN=Igha PE=1 SV=1 |
| 108 | 255 | tr\|A0A0A0MQF6\|A0A0A0MQF6_MOUSE | 41.29 | 7 | 3 | 2 | | 38653 | Glyceraldehyde-3-phosphate dehydrogenase OS=Mus musculus GN=Gapdh PE=1 SV=1 |
| 139 | 40853 | tr\|A0A0A6YW32\|A0A0A6YW32_MOUSE | 55.44 | 16 | 2 | 1 | | 13126 | Protein Ighv1-37 (Fragment) OS=Mus musculus GN=Ighv1-37 PE=4 SV=1 |
| 139 | 40847 | tr\|A0A0A6YWK5\|A0A0A6YWK5_MOUSE | 55.44 | 16 | 2 | 1 | | 13115 | Protein Ighv1-31 (Fragment) OS=Mus musculus GN=Ighv1-31 PE=4 SV=1 |
| 139 | 40848 | tr\|A0A0A6YXN4\|A0A0A6YXN4_MOUSE | 55.44 | 16 | 2 | 1 | | 12913 | Protein Ighv1-18 (Fragment) OS=Mus musculus GN=Ighv1-18 PE=4 SV=1 |
| 45 | 54 | tr\|A0A0A6YXW6\|A0A0A6YXW6_MOUSE | 56.17 | 10 | 5 | 5 | | 42078 | Protein Igha (Fragment) OS=Mus musculus GN=Igha PE=1 SV=1 |
| 171 | 40922 | tr\|A0A0A6YXZ4\|A0A0A6YXZ4_MOUSE | 45.4 | 16 | 2 | 2 | Pyro-glu from Q | 13359 | Protein Ighv1-62-2 (Fragment) OS=Mus musculus GN=Ighv1-62-2 PE=4 SV=1 |
| 139 | 40845 | tr\|A0A0A6YY38\|A0A0A6YY38_MOUSE | 55.44 | 16 | 2 | 1 | | 12933 | Protein Ighv1-42 (Fragment) OS=Mus musculus GN=Ighv1-42 PE=4 SV=1 |
| 171 | 40902 | tr\|A0A0A6YY41\|A0A0A6YY41_MOUSE | 45.4 | 16 | 2 | 2 | Pyro-glu from Q | 13108 | Protein Ighv1-47 (Fragment) OS=Mus musculus GN=Ighv1-47 PE=4 SV=1 |
| 226 | 41500 | tr\|A0A0B4J1I0\|A0A0B4J1I0_MOUSE | 26.41 | 17 | 2 | 2 | | 13176 | Protein Igkv1-110 OS=Mus musculus GN=Igkv1-110 PE=1 SV=2 |
| 188 | 40856 | tr\|A0A0B4J1J7\|A0A0B4J1J7_MOUSE | 51.81 | 19 | 2 | 1 | Pyro-glu from Q | 10650 | Protein Ighv1-82 OS=Mus musculus GN=Ighv1-82 PE=4 SV=1 |
| 185 | 40839 | tr\|A0A0F6B0M9\|A0A0F6B0M9_SALT1 | 35.62 | 6 | 2 | 1 | | 48574 | Glutamate dehydrogenase OS=Salmonella typhimurium (strain 14028s / SGSC 2262) GN=gdhA PE=4 SV=1 |
| 100 | 24457 | tr\|A2A513\|A2A513_MOUSE | 55.03 | 7 | 4 | 4 | | 57041 | Keratin type I cytoskeletal 10 OS=Mus musculus GN=Krt10 PE=1 SV=1 |
| 13 | 24340 | tr\|A2AKS7\|A2AKS7_MOUSE | 95.74 | 17 | 14 | 14 | | 80655 | Cadherin 17 isoform CRA_a OS=Mus musculus GN=Cdh17 PE=1 SV=1 |
| 101 | 702 | tr\|A6ZI44\|A6ZI44_MOUSE | 52.76 | 6 | 4 | 4 | | 45120 | Fructose-bisphosphate aldolase OS=Mus musculus GN=Aldoa PE=1 SV=1 |
| 92 | 18415 | tr\|A8DUK4\|A8DUK4_MOUSE | 58.81 | 24 | 3 | 3 | | 15748 | Beta-globin OS=Mus musculus GN=Hbbt1 PE=1 SV=1 |
| 28 | 40473 | tr\|B1ARR4\|B1ARR4_MOUSE | 52.69 | 13 | 4 | 4 | | 30347 | Carbonic anhydrase 6 OS=Mus musculus GN=Car6 PE=1 SV=1 |
| 124 | 18484 | tr\|B1AVD1\|B1AVD1_MOUSE | 57.18 | 5 | 3 | 3 | | 76434 | Protein Xpnpep2 OS=Mus musculus GN=Xpnpep2 PE=1 SV=1 |
| 124 | 18485 | tr\|B1AVD2\|B1AVD2_MOUSE | 57.18 | 5 | 3 | 3 | | 83733 | Protein Xpnpep2 OS=Mus musculus GN=Xpnpep2 PE=1 SV=1 |
| 16 | 40 | tr\|B2RS76\|B2RS76_MOUSE | 97.74 | 14 | 11 | 11 | Oxidation (M) | 47574 | Carboxypeptidase B1 (Tissue) OS=Mus musculus GN=Cpb1 PE=1 SV=1 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 22 | 151 | tr\|B2RU69\|B2RU69_MOUSE | 84.37 | 13 | 8 | 8 | | 68195 | CUB and zona pellucida-like domain-containing protein 1 OS=Mus musculus GN=Cuzd1 PE=1 SV=1 |
| 5 | 43 | tr\|B5THE2\|B5THE2_MOUSE | 106.59 | 9 | 22 | 20 | Amidation | 208548 | Maltase-glucoamylase OS=Mus musculus GN=Mgam PE=1 SV=1 |
| 101 | 699 | tr\|D3YWI1\|D3YWI1_MOUSE | 52.76 | 10 | 4 | 4 | | 25854 | Fructose-bisphosphate aldolase (Fragment) OS=Mus musculus GN=Aldoa PE=1 SV=5 |
| 108 | 254 | tr\|D3YYI5\|D3YYI5_MOUSE | 41.29 | 8 | 3 | 2 | | 36304 | Glyceraldehyde-3-phosphate dehydrogenase OS=Mus musculus GN=Gm7293 PE=3 SV=1 |
| 59 | 40534 | tr\|D3YYJ7\|D3YYJ7_MOUSE | 66.88 | 22 | 5 | 5 | | 24549 | MCG120048 OS=Mus musculus GN=Mptx2 PE=1 SV=1 |
| 201 | 127 | tr\|D3Z0Y2\|D3Z0Y2_MOUSE | 37.36 | 7 | 2 | 2 | | 22494 | Peroxiredoxin-6 OS=Mus musculus GN=Prdx6 PE=1 SV=1 |
| 180 | 5440 | tr\|D3Z3F7\|D3Z3F7_MOUSE | 51.15 | 5 | 2 | 2 | | 47411 | UDP-glucose 6-dehydrogenase OS=Mus musculus GN=Ugdh PE=1 SV=1 |
| 148 | 5660 | tr\|E9PWK4\|E9PWK4_MOUSE | 36.81 | 16 | 2 | 2 | | 13651 | Protein Jchain (Fragment) OS=Mus musculus GN=Jchain PE=1 SV=5 |
| 187 | 40668 | tr\|E9PZ00\|E9PZ00_MOUSE | 32.73 | 2 | 1 | 1 | | 60672 | Prosaposin OS=Mus musculus GN=Psap PE=1 SV=1 |
| 1 | 18626 | tr\|E9Q0B5\|E9Q0B5_MOUSE | 147.86 | 11 | 38 | 38 | Pyro-glu from Q | 275224 | Protein Fcgbp OS=Mus musculus GN=Fcgbp PE=1 SV=1 |
| 153 | 122 | tr\|E9Q4Y1\|E9Q4Y1_MOUSE | 27.9 | 2 | 1 | 1 | | 39110 | Protein Adh6a OS=Mus musculus GN=Adh6a PE=1 SV=1 |
| 153 | 123 | tr\|E9Q5Z6\|E9Q5Z6_MOUSE | 27.9 | 2 | 1 | 1 | | 39955 | Protein Adh6a OS=Mus musculus GN=Adh6a PE=1 SV=1 |
| 53 | 5447 | tr\|E9Q616\|E9Q616_MOUSE | 66.79 | 1 | 3 | 3 | | 604264 | Protein Ahnak OS=Mus musculus GN=Ahnak PE=1 SV=1 |
| 97 | 18624 | tr\|E9Q7P9\|E9Q7P9_MOUSE | 54.94 | 3 | 4 | 4 | | 142633 | Protein Cdhr2 OS=Mus musculus GN=Cdhr2 PE=1 SV=1 |
| 110 | 40462 | tr\|E9Q7Q3\|E9Q7Q3_MOUSE | 58.88 | 14 | 4 | 1 | | 28734 | Tropomyosin alpha-3 chain OS=Mus musculus GN=Tpm3 PE=1 SV=1 |
| 1 | 18627 | tr\|E9Q9C6\|E9Q9C6_MOUSE | 147.86 | 11 | 38 | 38 | Pyro-glu from Q | 275239 | Protein Fcgbp OS=Mus musculus GN=Fcgbp PE=1 SV=1 |
| 157 | 42348 | tr\|E9QPE7\|E9QPE7_MOUSE | 31.31 | 1 | 2 | 1 | | 223355 | Myosin-11 OS=Mus musculus GN=Myh11 PE=1 SV=1 |
| 2 | 500 | tr\|E9QPG8\|E9QPG8_MOUSE | 147.57 | 8 | 26 | 26 | Oxidation (M) | 211991 | Deleted in malignant brain tumors 1 protein OS=Mus musculus GN=Dmbt1 PE=1 SV=1 |
| 127 | 30058 | tr\|F2Z471\|F2Z471_MOUSE | 46.51 | 11 | 2 | 2 | | 28157 | Voltage-dependent anion-selective channel protein 1 OS=Mus musculus GN=Vdac1 PE=1 SV=1 |
| 198 | 40449 | tr\|F6ZSM3\|F6ZSM3_MOUSE | 50.23 | 7 | 2 | 2 | | 34078 | Annexin OS=Mus musculus GN=Anxa13 PE=1 SV=2 |
| 41 | 18401 | tr\|F8VPQ6\|F8VPQ6_MOUSE | 71.5 | 8 | 4 | 3 | | 59532 | Alkaline phosphatase OS=Mus musculus GN=Alpi PE=1 SV=1 |
| 26 | 40444 | tr\|F8VPT3\|F8VPT3_MOUSE | 80.57 | 4 | 9 | 9 | | 217800 | Protein Lct OS=Mus musculus GN=Lct PE=1 SV=1 |
| 40 | 5460 | tr\|F8VQM0\|F8VQM0_MOUSE | 57.12 | 4 | 3 | 2 | Oxidation (M) | 60291 | Alkaline phosphatase OS=Mus musculus GN=Akp3 PE=1 SV=1 |
| 9 | 81 | tr\|F8VQM5\|F8VQM5_MOUSE | 96.89 | 9 | 18 | 17 | | 208983 | Platelet-derived growth factor subunit B OS=Mus musculus GN=Sis PE=1 SV=1 |
| 137 | 18525 | tr\|F8WIV2\|F8WIV2_MOUSE | 50.58 | 13 | 3 | 3 | Pyro-glu from Q | 44774 | Serpin B6 OS=Mus musculus GN=Serpinb6a PE=1 SV=1 |
| 43 | 40538 | tr\|G3UX26\|G3UX26_MOUSE | 66.62 | 11 | 5 | 4 | | 30446 | Voltage-dependent anion-selective channel protein 2 (Fragment) OS=Mus musculus GN=Vdac2 PE=1 SV=1 |
| 62 | 87 | tr\|G3X8Z1\|G3X8Z1_MOUSE | 60.93 | 6 | 6 | 2 | | 102077 | Calcium-activated chloride channel regulator 4A OS=Mus musculus GN=Clca4a PE=1 SV=1 |
| 207 | 54071 | tr\|I1KAY5\|I1KAY5_SOYBN | 32.47 | 12 | 2 | 2 | | 13780 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 207 | 41064 | tr\|I1KLN6\|I1KLN6_SOYBN | 32.47 | 3 | 2 | 2 | | 59713 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 207 | 24817 | tr\|I1L0E5\|I1L0E5_SOYBN | 32.47 | 10 | 2 | 2 | | 17726 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 207 | 41061 | tr\|I1L8T7\|I1L8T7_SOYBN | 32.47 | 5 | 2 | 2 | | 34307 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 207 | 41062 | tr\|I1LYE4\|I1LYE4_SOYBN | 32.47 | 4 | 2 | 2 | | 42685 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 207 | 41063 | tr\|I1M077\|I1M077_SOYBN | 32.47 | 3 | 2 | 2 | | 51222 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 207 | 54072 | tr\|I1MG31\|I1MG31_SOYBN | 32.47 | 10 | 2 | 2 | | 17696 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 207 | 54073 | tr\|I1MS00\|I1MS00_SOYBN | 32.47 | 4 | 2 | 2 | | 42655 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 207 | 54069 | tr\|I1N7J3\|I1N7J3_SOYBN | 32.47 | 20 | 2 | 2 | | 8436 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 207 | 54070 | tr\|I1NAQ2\|I1NAQ2_SOYBN | 32.47 | 13 | 2 | 2 | | 13067 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 111 | 40475 | tr\|J3QMG3\|J3QMG3_MOUSE | 51.34 | 10 | 3 | 2 | | 30852 | Voltage-dependent anion-selective channel protein 3 OS=Mus musculus GN=Vdac3 PE=1 SV=1 |
| 187 | 40666 | tr\|J3QPG5\|J3QPG5_MOUSE | 32.73 | 2 | 1 | 1 | | 61363 | Prosaposin OS=Mus musculus GN=Psap PE=1 SV=1 |
| 187 | 40670 | tr\|K3W4L3\|K3W4L3_MOUSE | 32.73 | 2 | 1 | 1 | | 61292 | Prosaposin OS=Mus musculus GN=Psap PE=1 SV=1 |
| 207 | 41060 | tr\|K7KG22\|K7KG22_SOYBN | 32.47 | 10 | 2 | 2 | | 17320 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 25 | 67 | tr\|K7LZA4\|K7LZA4_SOYBN | 94.04 | 17 | 9 | 1 | | 55371 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 207 | 117794 | tr\|K7MD39\|K7MD39_SOYBN | 32.47 | 9 | 2 | 2 | | 18799 | Uncharacterized protein OS=Glycine max PE=4 SV=1 |
| 139 | 40843 | tr\|L7N238\|L7N238_MOUSE | 55.44 | 17 | 2 | 1 | | 12143 | Protein Ighv1-83 OS=Mus musculus GN=Ighv1-83 PE=4 SV=1 |
| 111 | 18508 | tr\|Q3TVK3\|Q3TVK3_MOUSE | 55.07 | 4 | 2 | 2 | | 52466 | Aspartyl aminopeptidase OS=Mus musculus GN=Dnpep PE=1 SV=1 |
| 135 | 40795 | tr\|Q3UKW2\|Q3UKW2_MOUSE | 40.76 | 13 | 3 | 3 | | 21560 | Calmodulin OS=Mus musculus GN=Calm1 PE=1 SV=1 |
| 82 | 6 | tr\|Q3UW98\|Q3UW98_MOUSE | 61.92 | 5 | 5 | 1 | Oxidation (M) | 102165 | Chloride channel calcium activated 7 OS=Mus musculus GN=Clca4b PE=1 SV=1 |
| 26 | 40454 | tr\|Q4VBW7\|Q4VBW7_MOUSE | 62.72 | 6 | 3 | 3 | | 54017 | Triacylglycerol lipase OS=Mus musculus GN=Pnliprp2 PE=1 SV=1 |
| 201 | 129 | tr\|Q6GT24\|Q6GT24_MOUSE | 37.36 | 6 | 2 | 2 | | 24827 | Peroxiredoxin 6 OS=Mus musculus GN=Prdx6 PE=1 SV=1 |
| 144 | 94614 | tr\|Q80X68\|Q80X68_MOUSE | 26.07 | 2 | 1 | 1 | | 52325 | Citrate synthase OS=Mus musculus GN=Csl PE=1 SV=1 |
| 187 | 40669 | tr\|Q8BFQ1\|Q8BFQ1_MOUSE | 32.73 | 2 | 1 | 1 | | 61051 | Prosaposin OS=Mus musculus GN=Psap PE=1 SV=1 |
| 201 | 260 | tr\|Q8BG37\|Q8BG37_MOUSE | 37.36 | 6 | 2 | 2 | | 24996 | MCG48959 OS=Mus musculus GN=Prdx6b PE=1 SV=1 |
| 111 | 18506 | tr\|Q8BPW9\|Q8BPW9_MOUSE | 55.07 | 4 | 2 | 2 | | 51662 | Aspartyl aminopeptidase OS=Mus musculus GN=Dnpep PE=1 SV=1 |
| 189 | 43339 | tr\|Q91XF2\|Q91XF2_MOUSE | 34.33 | 6 | 2 | 2 | | 20611 | Protein Tm4sf5 OS=Mus musculus GN=Tm4sf5 PE=2 SV=1 |
| 139 | 61 | tr\|Q925P3\|Q925P3_MOUSE | 38.51 | 4 | 2 | 2 | | 56986 | CEA-related cell adhesion molecule 1 OS=Mus musculus GN=Ceacam1 PE=1 SV=1 |
| 218 | 18805 | tr\|Q9ER05\|Q9ER05_MOUSE | 35.31 | 8 | 2 | 2 | | 28135 | Chymopasin OS=Mus musculus GN=Ctrl PE=1 SV=1 |
| 79 | 100 | tr\|Q9QUK9\|Q9QUK9_MOUSE | 72.22 | 18 | 4 | 2 | | 26277 | MCG15083 OS=Mus musculus GN=Try5 PE=1 SV=1 |
| 79 | 99 | tr\|Q9R0T7\|Q9R0T7_MOUSE | 72.22 | 18 | 4 | 2 | | 26274 | MCG15085 OS=Mus musculus GN=Try4 PE=1 SV=1 |
| 4 | 1 | tr\|Q9SB11\|Q9SB11_SOYBN | 126.45 | 21 | 14 | 9 | | 63797 | Glycinin OS=Glycine max GN=A5A4B3 PE=2 SV=1 |

## Supplemental Figures




**Supplemental Figure S1.** Diagram showing experimental workflow.




**Supplemental Figure S2.** Comparison of mzCloud and MSDial spectral matching databases for metabolomics identification.

# Chapter 7

# Gut Microbial and Metabolic Responses to *Salmonella enterica* Typhimurium and *Candida albicans*

Adapted from: Bratburd JR, Keller C, Vivas E, Gemperline E, Li L, Rey FE, Currie CR. Gut Microbial and Metabolic Responses to Salmonella enterica Serovar Typhimurium and Candida albicans. *MBio* **2018,** *9* (6)

Keywords: LC-MS; Metabolomics; Microbiome; Infection; Metagenomics

**Abstract**

The gut microbiota confers resistance to pathogens of the intestinal ecosystem, yet the dynamics of pathogen-microbiome interactions and the metabolites involved in this process remain largely unknown. Here, we use gnotobiotic mice infected with the virulent pathogen *Salmonella enterica* Typhimurium or the opportunistic pathogen *Candida albicans* in combination with metagenomics and discovery metabolomics to identify changes in the community and metabolome during infection. To isolate the role of the microbiota in response to pathogens, we compared mice monocolonized with the pathogen, uninfected mice "humanized" with a synthetic human microbiome, or infected humanized mice. In *Salmonella* infected mice, by three days into infection, microbiome community structure and function changed substantially, with a rise Enterobacteriaceae strains and a reduction in biosynthetic gene cluster potential. In contrast, *Candida* infected mice had few microbiome changes. The LC-MS metabolomic fingerprint of the cecum differed between mice monocolonized with either pathogen and humanized infected mice. Specifically, we identified an increase in glutathione disulfide, glutathione cysteine disulfide, inosine monophosphate, and hydroxybutyrylcarnitine in mice infected with *Salmonella* in contrast to uninfected mice and mice monocolonized with *Salmonella*. These metabolites potentially play a role in pathogen-induced oxidative stress. These results provide insight into how the microbiota community members interact with each other and with pathogens on a metabolic level.

**Importance**

The gut microbiota is increasingly recognized for playing a critical role in human health and disease, especially in conferring resistance to both virulent pathogens such as *Salmonella*, which infects 1.2 million people in the United States every year,[1] and opportunistic pathogens like

*Candida*, that causes an estimated 46,000 cases of invasive candidiasis each year in the United States.[2] Using a gnotobiotic mouse model, we investigate potential changes in gut microbial community structure and function during infection using metagenomics and metabolomics. We observe that changes in the community and in biosynthetic gene cluster potential occur within three days for the virulent *Salmonella enterica* Typhimurium, but there are minimal changes with a poorly colonizing *Candida albicans*. In addition, the metabolome shifts depending on infection status including changes in glutathione metabolites in response to *Salmonell*a infection, potentially in response to host oxidative stress.

**Introduction**

Symbiotic microbes help shape the biology of plants and animals.[3] In humans, gut microbes modulate nutrition, immune function, and are correlated with an increasing number of metabolic and neurological health and disease states.[4, 5] The human gastrointestinal tract harbors the largest fraction of microbial life in the body, estimated to range from $10^8$ to $10^{10}$ bacteria per gram in the ileum and stool, respectively.[6] Bacteria are the dominant taxa in the human gut microbiome, with the most abundant lineages belonging to the phyla Bacteroidetes and Firmicutes. Nevertheless, these communities are highly diverse and include viruses, archaea, fungi, and protists,[7-10] and all combined contain 150 times as many genes as the human genome.[11] In a healthy state, the human gut microbiome is relatively stable over time.[12, 13] Major disruption of the gut microbiome is associated with infections by a number of serious human pathogens, such as *Clostridium difficile*, vancomycin resistant Enterococcus (VRE), and *Salmonella enterica*.[14-16]

Preventing exogenous microbes from colonizing the human intestine is critical to the host maintaining a stable and healthy gut microbiome. The role of the microbiome in preventing

pathogens from invading the gut has been recognized since the 1950s, when pre-treatment with antibiotics was shown to drop infectious dose of *Salmonella enterica* 100,000-fold.[17] Gut microbes confer colonization resistance by outcompeting pathogens for nutrients, priming the host immune system, and directly targeting other microbes with metabolites.[18] Several examples of metabolites produced or modified by the microbiota that inhibit pathogens include short chain fatty acids, secondary bile acids, and modified compounds from the diet.[19-21] In addition, some members of the microbiota can create compounds to respond selectively to pathogen infection.[22] The gut microbiota has the potential to make a wide variety of novel natural products, and many of these large biosynthetic gene clusters encoding natural products are found in relatively small genomes, indicative of an ecological role for the products.[23]

Experiments using gnotobiotic mice with and without human microbiota, in combination with metagenomic and metabolomic approaches, can provide insight on the structure and function of the gut microbiota during pathogen invasion. Gnotobiotic mice are a mammalian model system in which defined microbiomes can be used in a controlled environment. Various metabolomics techniques, including nuclear magnetic resonance and chromatography-mass spectrometry, have been used for large-scale characterization of metabolite changes as a result of microbiome colonization, illustrating the impact of the microbiota on not only intestinal metabolism, but also on global systems.[24, 25] Furthermore, liquid chromatography-mass spectrometry (LC-MS) can help characterize metabolite changes due to disturbances in the microbiome[26, 27] and to screen for novel secondary metabolites and natural products in bacterial systems.[28, 29]

Here we examine colonization resistance in the humanized (HUM) mouse model. Specifically, we perform experimental infection with *Salmonella enterica* Typhimurium and *Candida albicans* in HUM mice and in germ-free (GF) mice. *Salmonella enterica* Typhimurium

is a disruptive pathogen that causes massive inflammation to outcompete the native microbiota in mice and human models.[30-32] *Candida albicans* can cause low-grade inflammation, but in contrast to *Salmonella enterica* Typhimurium is considered a commensal and occasional opportunistic pathogen in the GI tract.[33-36] Nevertheless, *C. albicans* has been shown to colonize GF and antibiotic-treated adult mice,[35, 37, 38] which appear otherwise resistant, suggesting that gut microbiota play a role in preventing *Candida* colonization in mice and humans. In this study we investigate how these pathogens alter the structure of the human gut microbiome, the biosynthetic gene cluster potential, and the metabolites produced in a healthy or infected state. We cross the presence and absence of the microbiome with the presence and absence of pathogen infection, using either *S. enterica* Typhimurium or *C. albicans*. To characterize strain level diversity that is not resolvable with 16S rRNA gene sequencing, we use shotgun metagenomics on fecal samples over three days of infection. We also identify the capacity of community members to produce novel antimicrobials through the biosynthetic gene clusters embedded in bacterial genomes. Further, we characterize metabolites using LC-MS for relative quantification and discovery metabolomics in the host cecum during infection, and validate the identifications of several specific metabolites with commercial standards.

**Experimental Methods**

Human gut microbiota and pathogens

For our synthetic human microbiome gut community, we used a collection of previously obtained isolates cultured from human fecal samples and maintained in long term storage in the Rey-lab at the University of Wisconsin-Madison. Bacterial isolates (**Supplemental Table 1**) were grown from glycerol stock on Mega Medium,[39] which was filter sterilized and held in a Coy anaerobic

chamber (5% $H_2$, 20% $CO_2$, and 75% $N_2$). An even mix from each bacterial culture was inoculated into each anaerobic tube. From stock cultures, *Salmonella enterica* Typhimurium ATCC 14028 was grown aerobically overnight in LB broth at 37°C, while *Candida albicans* K1 was grown on Sabourand dextrose agar (SDA).

Gnotobiotic Mice and Experimental Infections

The University of Wisconsin-Madison Animal Care and Use Committee approved protocols used in mice experiments. GF male C57BL/6J mice were maintained in gnotobiotic isolators until 8–12 weeks old with 12h light cycle and sterilized food and water *ad libitum*. These GF mice were then randomly assigned to 1 of 5 treatment groups and moved to out-of-the-isolator gnotobiotic cages in autoclaved filter top cages and subsequently gavaged in a biosafety cabinet using aseptic technique.[40] Mice were house 3 per cage, with a total of 6 mice per group.

To humanize mice, GF mice were colonized via oral gavage with 0.2 mL mixed bacterial culture as shown in Supplemental Table 3. All HUM mice were given the same inoculum, where bacteria were mixed with roughly similar proportions. Prior to infection, HUM mice were given 2 weeks to allow stabilization of the community. For mouse infections, mice were inoculated via oral gavage with 0.2 mL of overnight culture of *Salmonella enterica* Typhimurium ATCC 14028 or *Candida albicans* K1. Humanization and infection treatments were performed in a biosafety cabinet using aseptic technique.[40] Mice were sacrificed 3 days post infection or earlier depending on symptom severity and weight loss. Cecal contents were collected, flash frozen and stored at -80°C until processing. We selected cecum contents for LC-MS due to their high microbial loads and proximity to the distal ileum where Salmonella localizes to.[30, 41]

*Salmonella* and *Candida* quantification was performed by serial dilutions of fecal samples in phosphate buffered saline, followed by plating for quantification for *Salmonella* on Xylose lysine

deoxycholate (XLD) agar, and for *Candida* on SDA with chloramphenicol and gentamicin. Fecal samples from uninfected mice showed no growth on the SDA plates, as well as no growth of black colonies on XLD plates indicating no colonies capable of metabolizing thiosulfate into hydrogen sulfide as *Salmonella* does.

Metagenomics

To characterize the gut microbiome of HUM mice, we conducted metagenomics using Illumina MiSeq. Genomic DNA was extracted from fecal pellets following the Turnbaugh et al. protocol.[42] Briefly, the protocol is as follows: to each frozen fecal pellet, we added 500 μL of extraction buffer (200 mM Tris, 200 nM NaCl, 20 mM EDTA), 210 μL 20% SDS, 500 μL phenol:chloroform, 500 μL 0.1 mm zirconia/silica beads, and one 3.2 mm stainless steel bead. Cells were beaten for 3 minutes at room temperature. To remove contaminants, the Wizard SV Gel and PCR Clean-up kit was used. DNA library preparation and sequencing was done at the University of Wisconsin-Madison Biotechnology Center. Samples were prepared with the TruSeq Nano DNA LT Library Prep Kit (Illumina Inc., San Diego, California, USA) with minor modifications. After shearing samples with a Covaris M220 Ultrasonicator (Covaris Inc, Woburn, MA, USA), samples were size selected for an average insert size of 550 bp using SPRI bead based size exclusion, then libraries were standardized to 2 nM. Sequencing was done using single ends on the Illumina MiSeq Sequencer with a 50 bp (v2) sequencing cartridge.

Metagenomic data was pre-processed using BBDuk (https://sourceforge.net/projects/bbmap/) to trim adapters, remove phi-X contamination and quality trim reads to Q10. We analyzed the reads using the COPROseq (Community profiling by sequencing) pipeline,[43] which mapped reads to reference genomes using Bowtie version 1.0,[44] and normalized reads based on genome length. We also compared read mapping using Burrows-

Wheeler Alignment tool to verify reads mapped consistently.[45] Reference genomes were obtained from NCBI. Diversity was analyzed using the vegan package in R with a Kruskal-Wallis test. Biosynthetic gene clusters were identified using antiSMASH 4.0.[46] Gene clusters were then grouped by similarity using BiG-SCAPE (Navarro-Muñoz, Yeong, Medema *et al.*, in preparation: https://git.wageningenur.nl/medema-group/BiG-SCAPE). Data was analyzed and figures produced in R. Statistical testing was done using a Wilcoxon rank sum test (Mann Whitney U test) with a Benjamini-Hochberg correction.

Metabolomics

All chemicals were obtained from Fisher Scientific unless otherwise noted. Mouse cecum samples were placed in 10 mL PTFE tubes for extraction with a methanol/chloroform/water extraction. Three parts methanol, 1 part chloroform, and 4 parts water (Milli-Q system, Millipore, Billerica, MA) were added, in order, to each sample (total volume 4 mL) and centrifuged for 20 minutes at 4575 x g at 4°C. The aqueous fraction was removed and 4 parts methanol were then added. After brief vortexing, samples were centrifuged for 5 minutes at 1500 x g and 4°C. The organic layer was removed. Samples were dried in a SpeedVac and stored at -80°C. To clean up the sample, the aqueous fraction was further processed with a 3 kDa molecular weight cutoff (MWCO) (Amicon Ultra, Millipore). The MWCO device was rinsed with 0.2 mL 0.1 M NaOH and 0.5 mL 50/50 methanol/water. The sample was loaded in 0.5 mL 50/50 methanol/water and rinsed with 0.1 mL 50/50 methanol/water. All centrifuges occurred at 14,000 x g until the rinse or sample was through the device. The MWCO flow-through was dried with a SpeedVac and stored at -80°C until analysis.

Aqueous samples were resuspended in optima grade water at a concentration of 10 mg/mL. A Dionex Ultimate 3000 UHPLC system (Thermo Scientific, Waltham, MA, USA) and a Cortecs

C18 column (2.1 mm internal diameter x 100 mm length, 1.6 μm particle size; Waters, Milford, MA, USA), equipped with a corresponding guard column were used to separate the samples. The column temperature was 35°C, and the mobile phases were optima grade water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). The separation occurred with a 35 minute gradient at a flow rate of 0.3 mL/minutes with the following conditions: 0–5 min, 1% B; 5–10 min, linear gradient from 1–3% B; 10–18 min, linear gradient from 3–40% B; 18–22 min, linear gradient from 40–80% B; 22–27 min, column cleaning at 95% B; and 27–35 min, re-equilibration at 1% B. The injection volume was 3 μL and the samples were kept at 10°C during analysis. Metabolite MS data was acquired on a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific, Waltham, MA, USA), which was equipped with an ESI source and operated in positive ion mode with a scan range of *m/z* 200–1700. The MS parameters were as follows: 70,000 resolution, 1 E6 AGC, and 100 ms maximum injection time.

Metabolomics Data Analysis

Relative quantification of the metabolomics data for the different sample types was performed with Compound Discoverer software (Thermo Scientific, Waltham, MA, USA). Spectra underwent retention time alignment (adaptive curve 5 ppm, 1 minute tolerances), detection of unknown compounds (5 ppm, 30 Intensity threshold, 3 S/N threshold, 1,000,000 minimum peak intensity), and grouping of unknown compounds (5 ppm, 0.05 retention time tolerance). The Compound Discoverer workflow also included fill gaps, mark background, predict compositions, ChemSpider search, normalize areas (constant sum), merge features, and differential analysis. To isolate metabolites of interest, *m/z* values detected in the blanks or in more than 4 of 12 replicates in either of the germ-free infected conditions were removed. Additionally, *m/z* were selected if they showed 1.5-fold upregulation in 8 of 12 replicates of the infected humanized group, with the

ratios being calculated from the control with the highest normalized area. MetaboAnalyst[47, 48] was used for further statistical analysis after exporting *m/z* values, retention time and normalized areas from Compound Discoverer. Data was filtered with an interquantile range (IQR) estimate and log transformed. Heatmaps were produced using Pearson and Ward clustering.

Compound Identification

MS/MS spectra for the compounds on the target lists for both infections were collected on the Dionex UltiMate 3000 UHPLC and Q-Exactive instrument described above. The injection volume was 20 μL. An inclusion list was used for the targets with a retention time window of +/- 0.7 min. All charge states and salt adducts observed in the Compound Discoverer analysis were included in the inclusion list. The $MS^2$ parameters were as follows: 70,000 resolution, 5 E5 AGC, 100 ms maximum injection time, 1.0 *m/z* isolation window, and 30 NCE. MetFrag *in silico* fragmentation prediction software was used to aid in metabolite identification.[49] Target molecules were searched against KEGG and PubChem databases with a 5 ppm error. Candidate molecules from the databases were then processed against the MS/MS spectra of the target molecule with 5 ppm and 0.01 m/zabs settings. The top results of the *in silico* fragmentation were analyzed for putative identification. Putative identifications were then verified by comparing the experimental MS/MS to the MS/MS of the commercial standard.

## Results

Infection Severity in Mice with and without Microbiota

Germ-free mice, 8-12 weeks-old, were kept germ-free or colonized via oral gavage with a synthetic human community for two weeks, and then infected with *Salmonella enterica* Typhimurium or *Candida albicans* (**Figure 1A**). All infected mice showed presence of pathogens

in fecal samples by growth on selective media. Prior to infection, the mice weighed on average 29.8 g ± 2.3 (mean ± SD). GF mice infected with *Salmonella* (n=6), henceforth referred to as monocolonized *Salmonella* mice, lost an average body mass of 2.0 g ± 1.4 or 6.8% ± 4.7 within 12 hours post-infection. Due to severity of symptoms, three monocolonized *Salmonella* mice were sacrificed 12 hours post-infection, and the remaining monocolonized *Salmonella* mice and one HUM mouse infected with *Salmonella* were sacrificed within 24 hours of infection. HUM mice infected with *Salmonella* surviving 3 days into infection (n=5) lost an average of 4.2 g ± 0.6 or 14.3% ± 1.7, a significant loss in comparison to weight change from both the monocolonized and HUM mice infected with *Candida* (Mann Whitney U test, Bonferroni corrected, $p<0.05$). The monocolonized *Candida* mice (n=6) gained on average 0.2 g ± 0.3 or 0.8% ± 1.1 of their original weight, and the *Candida* infected HUM mice (n=6) gained on average 0.7 g ± 0.5 or 2.0% ± 1.8 of their original weight. There was no statistically significant difference in the change in weight for the monocolonized *Candida* mice compared to the HUM mice infected with *Candida* by the endpoint of the experiment, three days of infection.

Microbial Community Shifts in Response to Infection

We conducted Illumina based metagenomic sequencing on DNA from fecal pellets collected throughout infection. Each sample had on average 407,535 reads (SD = 63,381), ranging from 295,235 to 523,271 reads. The average number of reads with at least one reported alignment was 385,882 ± 96,477, or 95% of reads per sample. Prior to infection, the most abundant strains, making up over half of the relative abundance in the metagenomes from all groups, were *Bacteroides cellulosilyticus* DSM14838, *Subdoligranulum variabile*, *Bacteroides cellulosilyticus* WH2, *Akkermansia muciniphlia*, and *Clostridium bolteae* with an average relative abundance of 15.1%, 14.1%, 9.1%, 7.8% and 6.5%, respectively (**Figure 2A**). By day three in the *Salmonella*

infected HUM mice, most of the communities were dominated by *Salmonella* and other various Enterobacteriaceae strains from the original inoculum. Furthermore, diversity significantly decreased in *Salmonella* (**Supplemental Figure 1**). Prior to infection, these strains (*C. youngae*, *P. penneri*, *E. cancerogenus*, and *E. fergusonii*) in total represented an average relative abundance of 0.2%. In the metagenomes from two mice we observed an increase in the reads mapping to *Enterobacter cancerogenus,* up to 26.4% and 26.6% of the community, along with a smaller increase in *Proteus penneri*. One mouse had an increase in *Escherichia fergusonii* to 22.9% of the metagenome, while it remained below 1% of the metagenome in all the other mice. In another mouse, *Citrobacter youngae* reads increased to 15.2%, while in other mice *C. youngae* reads remained below 7.9%.After excluding *Salmonella* reads, we continued to observe a large shift in the relative abundance of community members. Using principal component analysis, we show large separation of the HUM *Salmonella* microbiome communities, three days post-infection, from a tight cluster of all other time points and treatments, with the first component explaining 31.4% of the variation (**Figure 2B**).

In all *Candida* infected HUM mice, less than 1% of reads mapped to the *Candida albicans* SC5314 reference genome. The metagenome of this group was not significantly different from uninfected HUM mice. The community structure remained fairly consistent over the infection period, although there was some variation in strain relative abundance over time (**Figure 2**). The largest change in any individual strain's relative abundance was an 8.4% increase in *Subdoligranulum variabile* in one mouse from 1 day post-infection to 3 days post-infection.

Prevalence of Biosynthetic Gene Clusters Within Genomes and Metagenomes

In total, from the genomes of the human microbiome used in this study, using antiSMASH 4.0,[46] we detected 1,081 biosynthetic gene clusters. Of these clusters, when grouped together using

BiG-SCAPE with a cutoff distance of 30 calculated based on a weighted combination of Jaccard, domain sequence similarity, and adjacency index, we identified 128 cluster nodes in 51 groups. The remaining 953 BGCs did not form any groupings with each other. Based on antiSMASH predicted classifications, most clusters were classified as other, which included putative clusters (486), fatty acids (117), fatty acid-saccharide combined clusters (22), aryl polyenes (14), siderophores (4), and resorcinol (3). Another large category was saccharides (345), followed by 62 ribosomally synthesized and post-translationally modified peptides (RiPPs), a group that includes bacteriocins, sactipeptides, lantipeptides, and thiopeptides. We also found 20 non-ribosomally synthesized peptide clusters and one hybrid polyketide-NRPS cluster in *Desulfovibrio piger* (**Supplemental Table 2**).

We found significant differences in the percentages of total metagenomic reads mapping to BGCs in *Salmonella* infected HUM mice prior to infection versus 3 days post infection (Wilcoxon $p<0.05$, corrected with Benjamini-Hochberg), excluding reads mapping to BGCs from *Salmonella* itself. Saccharides, lantipeptides, aryl polyenes, sactipeptides, fatty acids, fatty acid-saccharide, terpenes, and putative clusters were significantly reduced, while thiopeptides significantly increased 3 days post-infection (**Figure 3**). The majority of non-*Salmonella* reads mapping to thiopeptide clusters mapped to *Citrobacter youngae*, *Enterobacter cancerogenus*, *Proteus penneri*, and *Escherichia fergusonii*, consistent with the overall increase relative abundance in Enterobacteriaceae described above.

Differential Metabolomics during Infection and Novel Metabolite Potential

Analysis of the LC-MS results with Compound Discoverer (Thermo Fisher Scientific) resulted in the grouping of 8,613 merged features (chromatographic peaks) into 8,259 putative compounds. The compounds detected from the cecum samples of one or more mice from each treatment group

totaled 3,254 for the monocolonized *Candida* mice, 3,696 compounds for the monocolonized *Salmonella* mice, 3,349 compounds for the uninfected HUM mice, 2,924 compounds for the HUM mice infected with *Candida*, and 2,815 compounds for the HUM mice infected with *Salmonella.*

LC-MS *m/z* values and relative intensities from cecum contents showed separation of samples with principal component analysis (PCA). Two components were able to explain 67.7% of the variance (**Supplemental Figure 2**). Using partial least squares discriminant analysis (PLS-DA) we observed distinct separation of all groups with two components (R2=0.70799, Q2=0.66183 for component 1 and R2=0.85972 and Q2=0.81188 for component 2; **Figure 4A**). Using permutation testing of the PLS-DA, we obtained statistical significance ($p<0.001$) for 1000 permutations. The outliers in the *Salmonella* infected HUM mice group were from two technical replicates of one sample that had to be sacrificed 24 hours into infection. We also found distinct patterns for different groups of metabolites (**Figure 4B**), which indicate similar patterns between uninfected HUM and *Candida* infected HUM mice compared to monocolonized infected mice and HUM *Salmonella* mice. Additionally, we identified numerous features overrepresented in the monocolonized groups compared to the HUM groups (**Supplemental Figure 3**).

To examine metabolites potentially produced by the microbiome in response to infection, we looked for metabolites that were typically not found in pathogen monocolonized mice (absent in at least 8 of 12 samples, representing 6 biological replicates with 2 technical replicates each), and were at least 1.5 fold higher abundance in infected HUM mice compared to the highest normalized area of the controls (HUM mice with no infection). Using these guidelines, we narrowed our metabolites of interest to 31 out of 8,085 features detected overall. We detected 22 features in higher abundance in HUM *Salmonella* infected mice. In HUM *Candida* infected mice,

we found 10 features of interest based on the above criteria. One metabolite (*m/z* 347.0626, retention time 1.05 min) appeared to be shared between the lists, and also had matching tandem MS fragmentation from both infection groups. This metabolite had similar MS/MS to 3'Adenosine monophosphate and 2'Adenosine monophosphate standards, but the experimental retention time did not match that of the standards (1.37 minutes for 3'Adenosine monophosphate and 2.22 minutes for 2'Adenosine monophosphate). From the 31 selected compounds of interest, only 6 from HUM *Salmonella* and 4 from *Candida* infection had putative identifications based upon accurate mass matching to KEGG, HMDB, or AntiBase, leaving a remaining total of 21 potentially novel compounds (**Supplemental Table 3**). *In silico* fragmentation with MetFrag[49] was performed using MS/MS spectra obtained on the targets. If the top peaks in the experimental MS/MS were explained by the *in silico* fragmentation, then standards were obtained to confirm the identification. Using this procedure, we identified glutathione disulfide, glutathione cysteine disulfide, inosine 5' monophosphate, and hydroxybutyrylcarnitine as compounds upregulated from the HUM *Salmonella* group (**Supplemental Figure 4**). Although the *in silico* fragmentation approach worked well for the targets with KEGG matches, the increasing number of compounds in the more inclusive databases made it difficult to find putative identifications with MS/MS for targets that did not have matches to the KEGG databases.

**Discussion**

Understanding how the microbial communities change in response to perturbation is crucial for health, not only because the microbiota can protect the host against pathogenic microbes but also because changes in the gut microbiota have been associated with multiple health conditions.[50] Increasingly it has been recognized that pathogenicity and virulence can depends on

the context of specific microbe-microbe interactions or the whole community, indicating the importance of studying pathogen-microbiome interactions.[51, 52] In this study, we compare how two pathogenic perturbations affect the structure and function of human gut microbiota in a gnotobiotic mouse model. We find that during infection with *Salmonella*, the structure and functional capacity of the microbiota changes. Corresponding to these changes, we see significant changes in metabolites before versus during infection that vary with and without the human microbiota.

Our infection experiments revealed significant differences among treatments as measured by weight loss. *Candida* infected mice had weights that remained around their baseline starting weight. While we did isolate colony forming units of *Candida* from mice feces using media with antibiotics, indicating that viable yeast cells passed through the host, reads mapping to *Candida* from the metagenomic data were at or below the limit of detection, suggesting that *Candida* did not readily colonize these mice. Alternatively, the lack of fungal DNA may be influenced by our DNA extraction method.[53] In contrast, *Salmonella* infected mice lost significantly more weight than *Candida* infected mice by three days into infection, regardless of microbiome presence or absence. GF mice infected with *Salmonella* were moribund within 24 hours, while HUM mice infected with *Salmonella* were able to survive until the end of the three days, with the exception of one mouse, indicative of the protective effects of the microbiota against *Salmonella*.

*Salmonella* infection perturbed the microbiota and led to an increase in the relative abundance of different Enterobacteriaceae, whereas *Candida* did not. Prior to infection, the microbiota contained similar dominant taxa including Bacteroidetes and Firmicutes with relatively few Gammaproteobacteria. During *Salmonella* infection in humanized mice, the metagenomic data indicated an increase in the relative abundance of Enterobacteriaceae (including strains besides *Salmonella*). This result is consistent with previous work examining changes in gut microbial

communities during *Salmonella* infection,[30, 54, 55] and resembles increases in Enterobacteriaceae during antibiotic treatment,[15] both of which may ultimately be driven by the oxygenation of the gut.[56] These changes may represent a bloom of closely related strains or a reduction in the size of the bacterial community overall. Although Enterobacteriaceae increased in the samples, which particular strains increased appeared stochastic. Some of the variation may be due to read mapping of conserved genes to closely related strains; however, we saw similar results using different read mapping programs (Bowtie and Burrows-Wheeler Algorithm) and using parameters to exclude non-uniquely mapping reads. Given that these strains may compete with *Salmonella* over electron acceptors and trace elements, further investigation on these dynamic interactions is warranted.[57, 58] The stochasticity may also reflect the general instability of the community. While *Salmonella* dramatically perturbs the community, *Candida* did not seem to readily colonize the mice, and although some changes occurred in the microbial communities, these fluctuations are within the range of natural variation.

The synthetic human microbiome used in this study contained many biosynthetic gene clusters and the potential functional capacity changed with infection treatment. In our input strains we found potential for unknown biosynthetic gene clusters, including RiPPs, NRPS clusters, and many putative clusters. This fits with previous observations; biosynthetic gene clusters are common in human gut microbiota and anaerobic bacteria.[23, 59] Metagenomic analysis indicated a decrease in most cluster types during *Salmonella* infection, which likely reflects a drop in community diversity. One exception was the increase in reads mapping to gene clusters involved in thiopeptide biosynthesis, which was increased even after removing reads mapping to *Salmonella*'s own thiopeptide biosynthesis cluster. Thiopeptides are a class of peptide antibiotics that target gram-positive bacteria.[60] Since *Salmonella* is gram-negative and has one putative

thiopeptide BGC of its own, it seems unlikely that these thiopeptide clusters, if produced, would target *Salmonella*. Other possibilities are that if produced, these secondary metabolites encoded by clusters might add to the community instability, or that these genes are not transcribed or translated. Alternatively, this result may suggest that the pathogen-induced disruption in the microbiome helps diminish members that would have been capable of producing BGC products. Further research will be needed to characterize what role, if any, these BGCs play during infection.

Our discovery metabolomics showed differences in the metabolites present in the mouse cecum based on presence of microbiome as well as infection. For example, the metabolomes of *Salmonella*-infected, *Candida*-infected, and uninfected mice ceca grouped separately on PLS-DA analysis, suggesting distinct metabolic responses between a virulent bacterial pathogen and opportunistic fungal pathogen. The changes in overall metabolites based on gut microbiota supports previous research comparing germ-free and colonized mice, and mice with different gut microbiome donors.[61] We found more putative metabolites of interest (based on higher abundance in HUM infected mice and generally absent in GF mice) from *Salmonella* infected mice than *Candida* infected mice. Previous studies investigating global metabolomics in *Salmonella* infections have focused on the hosts with conventional mouse microbiota, finding disruptions in host hormone pathways,[62] changes in common microbial metabolites including trimethylamine N-oxide (TMAO) and hippurate,[63] and changes in sugar moieties.[54] Our study differed from these previous studies in that we used gnotobiotic mice to specifically focus on metabolites produced when human-associated gut microbiota strains were exposed to pathogens. While using native microbiota to look for pathogen interactions is valuable especially in an ecological context, the humanized mouse model enables exploration of potentially distinct chemical interactions between human microbiota strains and human pathogens.[64] Furthermore, human gut microbiota extracts

have been previously shown to inhibit virulence of *Salmonella in vitro*.[51] Mice monocolonized with pathogens serve as key controls that allowed us to focus on compounds apparently made by the microbiota during infection rather than overall host changes. Nevertheless, the possibility exists that we may detect metabolites made by *Salmonella* in response to gut microbiota in our experiments or metabolites that differ due to GF mice exhibiting colitis rather than the typical systemic Typhoid-like infection.[65] In addition, we scanned for molecular features with an *m/z* greater than 200, to avoid discovery of smaller commonly made microbial metabolites. In our metabolites of interest from humanized infection conditions, we had many molecular features that were not identified with KEGG, HMDB, or AntiBase, potentially indicating novel metabolites. One drawback in studying these metabolite interactions *in vivo* is the challenge in isolating individual novel molecules from a complex mixture, even in a well-described community with full genomes,[66] as we were unable to match known and predicted metabolites to a majority of our target *m/z* values. Although work is being done to increase MS/MS databases for natural products,[67] identifying natural products is still challenging, as many natural product databases, including AntiBase, are not MS compatible.

We were able to identify a few metabolites specific to the humanized *Salmonella* infected group, including two metabolites in the glutathione pathway. In particular, we identified glutathione disulfide and glutathione cysteine disulfide in higher abundance in humanized *Salmonella* infected mice. *Salmonella* infection triggers vast amounts of oxidative stress,[68] and glutathione metabolism is important for protection against oxidative stress.[69] Changes in gene encoding antioxidant proteins have also been identified in humans exposed to *Salmonella enterica* Typhi.[70] Further, glutathione cysteine disulfide has been shown to reduce colonic lesions in a mouse model of colitis.[71] Previous work indicates that germ-free mice have a disrupted glutathione

metabolism relative to conventional mice.[72] It remains to be seen whether experimentally manipulating glutathione metabolite amounts affect *Salmonella* infections *in vivo*, and to what extent different gut microbes contribute to glutathione pool. In contrast to the hypothesis that microbes may make specific metabolites that inhibit certain pathogens, this evidence suggests more generalized responses to certain kinds of dysbiosis, such as oxidative stress.[73] The possibility of microbial metabolites with specific responses to pathogens cannot be eliminated, however, many metabolites remain unidentified, and the roles of those identified are unclear. Further characterization of microbial metabolites made during infection is necessary to identify these responses.

Colonization resistance conferred by the microbiota helps the host resist a variety of pathogens, including *Salmonella*. Understanding the complex interactions between the host, microbiota, and pathogens will enable better microbiome based-therapies, from fecal microbiota transplants to microbiota-derived compounds.[74, 75] Combining gnotobiotic mice with genomics and metabolomics has allowed us to interrogate changes in community composition and function during infection in an unbiased manner, and demonstrates distinct metabolic responses to a virulent or opportunistic pathogen.

**Acknowledgements**

This work was supported by T32AI55397 and U19AI109673 from the National Institute of Health, and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin–Madison with funding from the Wisconsin Alumni Research Foundation. Sequencing was provided by the University of Wisconsin Biotech Center. This work was supported in part by grants NIH DK108259 (to F.E.R.). L. L. acknowledges a Vilas Distinguished Achievement

Professorship, NIH DK071801, and S10RR029531. The Q Exactive instrument was purchased through a NIH shared instrument grant (NCRR S10RR029531). We would like to thank Kimberley Romano for assistance in mouse experimental procedures, Marc Chevrette for advice on secondary metabolite clustering methods. We also thank Camila Carlos, Lily Khadempour, Heidi Horn, and Lindsay Kalan for critical feedback on the manuscript.

**References**

1. Scallan, E.; Hoekstra, R. M.; Angulo, F. J.; Tauxe, R. V.; Widdowson, M. A.; Roy, S. L.; Jones, J. L.; Griffin, P. M., Foodborne illness acquired in the United States--major pathogens. *Emerg Infect Dis* **2011,** *17* (1), 7-15.

2. Center; Prevention, f. D. C. a., Antibiotic Resistance Threats in the United States, 2013. 2013.

3. Hacquard, S.; Garrido-Oter, R.; Gonzalez, A.; Spaepen, S.; Ackermann, G.; Lebeis, S.; McHardy, A. C.; Dangl, J. L.; Knight, R.; Ley, R.; Schulze-Lefert, P., Microbiota and Host Nutrition across Plant and Animal Kingdoms. *Cell Host Microbe* **2015,** *17* (5), 603-16.

4. Sharon, G.; Garg, N.; Debelius, J.; Knight, R.; Dorrestein, P. C.; Mazmanian, S. K., Specialized metabolites from the microbiome in health and disease. *Cell Metab* **2014,** *20* (5), 719-730.

5. Lynch, S. V.; Pedersen, O., The Human Intestinal Microbiome in Health and Disease. *N Engl J Med* **2016,** *375* (24), 2369-2379.

6. Sender, R.; Fuchs, S.; Milo, R., Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol* **2016,** *14* (8), e1002533.

7. Walter, J.; Ley, R., The human gut microbiome: ecology and recent evolutionary changes. *Annu Rev Microbiol* **2011,** *65*, 411-29.

8. Scanlan, P. D.; Marchesi, J. R., Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *Isme j* **2008,** *2* (12), 1183-93.

9. Hoffmann, C.; Dollive, S.; Grunberg, S.; Chen, J.; Li, H.; Wu, G. D.; Lewis, J. D.; Bushman, F. D., Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One* **2013,** *8* (6), e66019.

10. Minot, S.; Sinha, R.; Chen, J.; Li, H.; Keilbaugh, S. A.; Wu, G. D.; Lewis, J. D.; Bushman, F. D., The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res* **2011,** *21* (10), 1616-25.

11. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K. S.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; Mende, D. R.; Li, J.; Xu, J.; Li, S.; Li, D.; Cao, J.; Wang, B.; Liang, H.; Zheng, H.; Xie, Y.; Tap, J.; Lepage, P.; Bertalan, M.; Batto, J. M.; Hansen, T.; Le Paslier, D.; Linneberg, A.; Nielsen, H. B.; Pelletier, E.; Renault, P.; Sicheritz-Ponten, T.; Turner, K.; Zhu, H.; Yu, C.; Jian, M.; Zhou, Y.; Li, Y.; Zhang, X.; Qin, N.; Yang, H.; Wang, J.; Brunak, S.; Dore, J.; Guarner, F.; Kristiansen, K.; Pedersen, O.; Parkhill, J.; Weissenbach, J.; Bork, P.; Ehrlich, S. D., A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **2010,** *464* (7285), 59-65.

12. Lloyd-Price, J.; Mahurkar, A.; Rahnavard, G.; Crabtree, J.; Orvis, J.; Hall, A. B.; Brady, A.; Creasy, H. H.; McCracken, C.; Giglio, M. G.; McDonald, D.; Franzosa, E. A.; Knight, R.; White, O.; Huttenhower, C., Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* **2017,** *550* (7674), 61-66.

13. Faith, J. J.; Guruge, J. L.; Charbonneau, M.; Subramanian, S.; Seedorf, H.; Goodman, A. L.; Clemente, J. C.; Knight, R.; Heath, A. C.; Leibel, R. L.; Rosenbaum, M.; Gordon, J. I., The long-term stability of the human gut microbiota. *Science* **2013,** *341* (6141), 1237439.

14. Buffie, C. G.; Jarchum, I.; Equinda, M.; Lipuma, L.; Gobourne, A.; Viale, A.; Ubeda, C.; Xavier, J.; Pamer, E. G., Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to Clostridium difficile-induced colitis. *Infect Immun* **2012,** *80* (1), 62-73.

15. Ubeda, C.; Taur, Y.; Jenq, R. R.; Equinda, M. J.; Son, T.; Samstein, M.; Viale, A.; Socci, N. D.; van den Brink, M. R.; Kamboj, M.; Pamer, E. G., Vancomycin-resistant Enterococcus domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. *J Clin Invest* **2010,** *120* (12), 4332-41.

16. Lawley, T. D.; Bouley, D. M.; Hoy, Y. E.; Gerke, C.; Relman, D. A.; Monack, D. M., Host transmission of Salmonella enterica serovar Typhimurium is controlled by virulence factors and indigenous intestinal microbiota. *Infect Immun* **2008,** *76* (1), 403-16.

17. Bohnhoff, M.; Drake, B. L.; Miller, C. P., Effect of streptomycin on susceptibility of intestinal tract to experimental Salmonella infection. *Proc Soc Exp Biol Med* **1954,** *86* (1), 132-7.

18. Buffie, C. G.; Pamer, E. G., Microbiota-mediated colonization resistance against intestinal pathogens. *Nat Rev Immunol* **2013,** *13* (11), 790-801.

19. Spees, A. M.; Lopez, C. A.; Kingsbury, D. D.; Winter, S. E.; Baumler, A. J., Colonization resistance: battle of the bugs or Menage a Trois with the host? *PLoS Pathog* **2013,** *9* (11), e1003730.

20. Steed, A. L.; Christophi, G. P.; Kaiko, G. E.; Sun, L.; Goodwin, V. M.; Jain, U.; Esaulova, E.; Artyomov, M. N.; Morales, D. J.; Holtzman, M. J.; Boon, A. C. M.; Lenschow, D. J.; Stappenbeck, T. S., The microbial metabolite desaminotyrosine protects from influenza through type I interferon. *Science* **2017,** *357* (6350), 498-502.

21. Buffie, C. G.; Bucci, V.; Stein, R. R.; McKenney, P. T.; Ling, L.; Gobourne, A.; No, D.; Liu, H.; Kinnebrew, M.; Viale, A.; Littmann, E.; van den Brink, M. R.; Jenq, R. R.; Taur, Y.; Sander, C.; Cross, J. R.; Toussaint, N. C.; Xavier, J. B.; Pamer, E. G., Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. *Nature* **2015,** *517* (7533), 205-8.

22. Hsiao, A.; Ahmed, A. M.; Subramanian, S.; Griffin, N. W.; Drewry, L. L.; Petri, W. A., Jr.; Haque, R.; Ahmed, T.; Gordon, J. I., Members of the human gut microbiota involved in recovery from Vibrio cholerae infection. *Nature* **2014,** *515* (7527), 423-6.

23. Donia, M. S.; Cimermancic, P.; Schulze, C. J.; Wieland Brown, L. C.; Martin, J.; Mitreva, M.; Clardy, J.; Linington, R. G.; Fischbach, M. A., A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell* **2014,** *158* (6), 1402-1414.

24. Claus, S. P.; Tsang, T. M.; Wang, Y.; Cloarec, O.; Skordi, E.; Martin, F. P.; Rezzi, S.; Ross, A.; Kochhar, S.; Holmes, E.; Nicholson, J. K., Systemic multicompartmental effects of the gut microbiome on mouse metabolic phenotypes. *Mol Syst Biol* **2008,** *4*, 219.

25. Martin, F. P.; Dumas, M. E.; Wang, Y.; Legido-Quigley, C.; Yap, I. K.; Tang, H.; Zirah, S.; Murphy, G. M.; Cloarec, O.; Lindon, J. C.; Sprenger, N.; Fay, L. B.; Kochhar, S.; van Bladeren, P.; Holmes, E.; Nicholson, J. K., A top-down systems biology view of microbiome-mammalian metabolic interactions in a mouse model. *Mol Syst Biol* **2007,** *3*, 112.

26. Zheng, X. J.; Xie, G. X.; Zhao, A. H.; Zhao, L. J.; Yao, C.; Chiu, N. H. L.; Zhou, Z. X.; Bao, Y. Q.; Jia, W. P.; Nicholson, J. K.; Jia, W., The Footprints of Gut Microbial-Mammalian Co-Metabolism. *Journal of Proteome Research* **2011,** *10* (12), 5512-5522.

27. Lu, K.; Abo, R. P.; Schlieper, K. A.; Graffam, M. E.; Levine, S.; Wishnok, J. S.; Swenberg, J. A.; Tannenbaum, S. R.; Fox, J. G., Arsenic exposure perturbs the gut microbiome and its metabolic profile in mice: an integrated metagenomics and metabolomics analysis. *Environ Health Perspect* **2014,** *122* (3), 284-91.

28. Hou, Y.; Braun, D. R.; Michel, C. R.; Klassen, J. L.; Adnani, N.; Wyche, T. P.; Bugni, T. S., Microbial Strain Prioritization Using Metabolomics Tools for the Discovery of Natural Products. *Analytical Chemistry* **2012,** *84* (10), 4277-4283.

29. Krug, D.; Zurek, G.; Revermann, O.; Vos, M.; Velicer, G. J.; Muller, R., Discovering the hidden secondary metabolome of Myxococcus xanthus: a study of intraspecific diversity. *Appl Environ Microbiol* **2008,** *74* (10), 3058-68.

30. Barman, M.; Unold, D.; Shifley, K.; Amir, E.; Hung, K.; Bos, N.; Salzman, N., Enteric salmonellosis disrupts the microbial ecology of the murine gastrointestinal tract. *Infect Immun* **2008,** *76* (3), 907-15.

31. Stecher, B.; Robbiani, R.; Walker, A. W.; Westendorf, A. M.; Barthel, M.; Kremer, M.; Chaffron, S.; Macpherson, A. J.; Buer, J.; Parkhill, J.; Dougan, G.; von Mering, C.; Hardt, W. D., Salmonella enterica serovar typhimurium exploits inflammation to compete with the intestinal microbiota. *PLoS Biol* **2007,** *5* (10), 2177-89.

32. Rivera-Chavez, F.; Baumler, A. J., The Pyromaniac Inside You: Salmonella Metabolism in the Host Gut. *Annu Rev Microbiol* **2015,** *69*, 31-48.

33. Kumamoto, C. A., Inflammation and gastrointestinal Candida colonization. *Curr Opin Microbiol* **2011,** *14* (4), 386-91.

34. Mason, K. L.; Erb Downward, J. R.; Falkowski, N. R.; Young, V. B.; Kao, J. Y.; Huffnagle, G. B., Interplay between the gastric bacterial microbiota and Candida albicans during postantibiotic recolonization and gastritis. *Infect Immun* **2012,** *80* (1), 150-8.

35. Koh, A. Y., Murine models of Candida gastrointestinal colonization and dissemination. *Eukaryot Cell* **2013,** *12* (11), 1416-22.

36. Auchtung, T. A.; Fofanova, T. Y.; Stewart, C. J.; Nash, A. K.; Wong, M. C.; Gesell, J. R.; Auchtung, J. M.; Ajami, N. J.; Petrosino, J. F., Investigating Colonization of the Healthy Adult Gastrointestinal Tract by Fungi. *mSphere* **2018,** *3* (2).

37. Schofield, D. A.; Westwater, C.; Balish, E., Divergent chemokine, cytokine and beta-defensin responses to gastric candidiasis in immunocompetent C57BL/6 and BALB/c mice. *J Med Microbiol* **2005,** *54* (Pt 1), 87-92.

38. Kennedy, M. J.; Volz, P. A., Effect of various antibiotics on gastrointestinal colonization and dissemination by Candida albicans. *Sabouraudia* **1985,** *23* (4), 265-73.

39. Romano, K. A.; Vivas, E. I.; Amador-Noguez, D.; Rey, F. E., Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. *MBio* **2015,** *6* (2), e02481.

40. Faith, J. J.; Ahern, P. P.; Ridaura, V. K.; Cheng, J.; Gordon, J. I., Identifying gut microbe-host phenotype relationships using combinatorial communities in gnotobiotic mice. *Sci Transl Med* **2014,** *6* (220), 220ra11.

41. Dunne, C., Adaptation of bacteria to the intestinal niche: probiotics and gut disorder. *Inflamm Bowel Dis* **2001,** *7* (2), 136-45.

42. Turnbaugh, P. J.; Ridaura, V. K.; Faith, J. J.; Rey, F. E.; Knight, R.; Gordon, J. I., The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med* **2009,** *1* (6), 6ra14.

43. McNulty, N. P.; Yatsunenko, T.; Hsiao, A.; Faith, J. J.; Muegge, B. D.; Goodman, A. L.; Henrissat, B.; Oozeer, R.; Cools-Portier, S.; Gobert, G.; Chervaux, C.; Knights, D.; Lozupone, C. A.; Knight, R.; Duncan, A. E.; Bain, J. R.; Muehlbauer, M. J.; Newgard, C. B.; Heath, A. C.; Gordon, J. I., The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. *Sci Transl Med* **2011,** *3* (106), 106ra106.

44. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S. L., Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **2009,** *10* (3), R25.

45. Li, H.; Durbin, R., Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009,** *25* (14), 1754-60.

46. Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de Los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H., antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* **2017,** *45* (W1), W36-w41.

47. Xia, J.; Sinelnikov, I. V.; Han, B.; Wishart, D. S., MetaboAnalyst 3.0--making metabolomics more meaningful. *Nucleic Acids Res* **2015,** *43* (W1), W251-7.

48. Xia, J.; Wishart, D. S., Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. *Curr Protoc Bioinformatics* **2016,** *55*, 14.10.1-14.10.91.

49. Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S., MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform* **2016,** *8*, 3.

50. Costello, E. K.; Stagaman, K.; Dethlefsen, L.; Bohannan, B. J.; Relman, D. A., The application of ecological theory toward an understanding of the human microbiome. *Science* **2012,** *336* (6086), 1255-62.

51. Antunes, L. C.; McDonald, J. A.; Schroeter, K.; Carlucci, C.; Ferreira, R. B.; Wang, M.; Yurist-Doutsch, S.; Hira, G.; Jacobson, K.; Davies, J.; Allen-Vercoe, E.; Finlay, B. B., Antivirulence activity of the human gut metabolome. *MBio* **2014,** *5* (4), e01183-14.

52. Byrd, A. L.; Segre, J. A., Adapting Koch's postulates. *Science* **2016,** *351* (6270), 224.

53. Wesolowska-Andersen, A.; Bahl, M. I.; Carvalho, V.; Kristiansen, K.; Sicheritz-Ponten, T.; Gupta, R.; Licht, T. R., Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. *Microbiome* **2014,** *2*, 19.

54. Deatherage Kaiser, B. L.; Li, J.; Sanford, J. A.; Kim, Y. M.; Kronewitter, S. R.; Jones, M. B.; Peterson, C. T.; Peterson, S. N.; Frank, B. C.; Purvine, S. O.; Brown, J. N.; Metz, T. O.; Smith, R. D.; Heffron, F.; Adkins, J. N., A Multi-Omic View of Host-Pathogen-Commensal Interplay in Salmonella-Mediated Intestinal Infection. *PLoS One* **2013,** *8* (6), e67155.

55. Stecher, B.; Chaffron, S.; Kappeli, R.; Hapfelmeier, S.; Freedrich, S.; Weber, T. C.; Kirundi, J.; Suar, M.; McCoy, K. D.; von Mering, C.; Macpherson, A. J.; Hardt, W. D., Like will to like: abundances of closely related species can predict susceptibility to intestinal colonization by pathogenic and commensal bacteria. *PLoS Pathog* **2010,** *6* (1), e1000711.

56. Rivera-Chavez, F.; Lopez, C. A.; Baumler, A. J., Oxygen as a driver of gut dysbiosis. *Free Radic Biol Med* **2017,** *105*, 93-101.

57. Thiennimitr, P.; Winter, S. E.; Winter, M. G.; Xavier, M. N.; Tolstikov, V.; Huseby, D. L.; Sterzenbach, T.; Tsolis, R. M.; Roth, J. R.; Baumler, A. J., Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota. *Proc Natl Acad Sci U S A* **2011,** *108* (42), 17480-5.

58. Deriu, E.; Liu, J. Z.; Pezeshki, M.; Edwards, R. A.; Ochoa, R. J.; Contreras, H.; Libby, S. J.; Fang, F. C.; Raffatellu, M., Probiotic bacteria reduce salmonella typhimurium intestinal colonization by competing for iron. *Cell Host Microbe* **2013,** *14* (1), 26-37.

59. Letzel, A. C.; Pidot, S. J.; Hertweck, C., A genomic approach to the cryptic secondary metabolome of the anaerobic world. *Nat Prod Rep* **2013,** *30* (3), 392-428.

60. Just-Baringo, X.; Albericio, F.; Alvarez, M., Thiopeptide antibiotics: retrospective and recent advances. *Mar Drugs* **2014,** *12* (1), 317-51.

61. Marcobal, A.; Kashyap, P. C.; Nelson, T. A.; Aronov, P. A.; Donia, M. S.; Spormann, A.; Fischbach, M. A.; Sonnenburg, J. L., A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. *Isme j* **2013,** *7* (10), 1933-43.

62. Antunes, L. C.; Arena, E. T.; Menendez, A.; Han, J.; Ferreira, R. B.; Buckner, M. M.; Lolic, P.; Madilao, L. L.; Bohlmann, J.; Borchers, C. H.; Finlay, B. B., Impact of salmonella infection on host hormone metabolism revealed by metabolomics. *Infect Immun* **2011,** *79* (4), 1759-69.

63. Zhu, X.; Lei, H.; Wu, J.; Li, J. V.; Tang, H.; Wang, Y., Systemic responses of BALB/c mice to Salmonella typhimurium infection. *J Proteome Res* **2014,** *13* (10), 4436-45.

64. Arrieta, M. C.; Walter, J.; Finlay, B. B., Human Microbiota-Associated Mice: A Model with Challenges. *Cell Host Microbe* **2016,** *19* (5), 575-8.

65. Stecher, B.; Macpherson, A. J.; Hapfelmeier, S.; Kremer, M.; Stallmach, T.; Hardt, W. D., Comparison of Salmonella enterica serovar Typhimurium colitis in germfree mice and mice pretreated with streptomycin. *Infect Immun* **2005,** *73* (6), 3228-41.

66. Li, J. W.; Vederas, J. C., Drug discovery and natural products: end of an era or an endless frontier? *Science* **2009,** *325* (5937), 161-5.

67. Wang, M. X.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W. T.; Criisemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderon, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C. C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C. C.; Yang, Y. L.; Humpf, H. U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya, C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J. Q.; Neupane, R.; Gurr, J.; Rodriguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P. M.; Phapale, P.; Nothias, L. F.; Alexandrovr, T.; Litaudon, M.; Wolfender, J. L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D. T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Muller, R.; Waters, K. M.; Shi, W. Y.; Liu, X. T.; Zhang, L. X.; Knight, R.; Jensen, P. R.; Palsson, B. O.; Pogliano, K.; Linington, R. G.; Gutierrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N., Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology* **2016,** *34* (8), 828-837.

68. Farr, S. B.; Kogoma, T., Oxidative stress responses in Escherichia coli and Salmonella typhimurium. *Microbiol Rev* **1991,** *55* (4), 561-85.

69. Hayes, J. D.; McLellan, L. I., Glutathione and glutathione-dependent enzymes represent a co-ordinately regulated defence against oxidative stress. *Free Radic Res* **1999,** *31* (4), 273-300.

70. Zhang, Y.; Brady, A.; Jones, C.; Song, Y.; Darton, T. C.; Blohmke, C. J.; Pollard, A. J.; Magder, L. S.; Fasano, A.; Sztein, M. B.; Fraser, C. M., Compositional and Functional Differences in the Human Gut Microbiome Correlate with Clinical Outcome following Infection with Wild-Type Salmonella enterica Serovar Typhi. *MBio* **2018,** *9* (3).

71. Oz, H. S.; Chen, T. S.; Nagasawa, H., Comparative efficacies of 2 cysteine prodrugs and a glutathione delivery agent in a colitis model. *Transl Res* **2007,** *150* (2), 122-9.

72. Mardinoglu, A.; Shoaie, S.; Bergentall, M.; Ghaffari, P.; Zhang, C.; Larsson, E.; Backhed, F.; Nielsen, J., The gut microbiota modulates host amino acid and glutathione metabolism in mice. *Mol Syst Biol* **2015,** *11* (10), 834.

73. Duvallet, C.; Gibbons, S. M.; Gurry, T.; Irizarry, R. A.; Alm, E. J., Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat Commun* **2017,** *8* (1), 1784.

74. Kassam, Z.; Lee, C. H.; Yuan, Y.; Hunt, R. H., Fecal microbiota transplantation for Clostridium difficile infection: systematic review and meta-analysis. *Am J Gastroenterol* **2013,** *108* (4), 500-8.

75. Suez, J.; Elinav, E., The path towards microbiome-based metabolite treatment. *Nat Microbiol* **2017,** *2,* 17075.

**Figures**

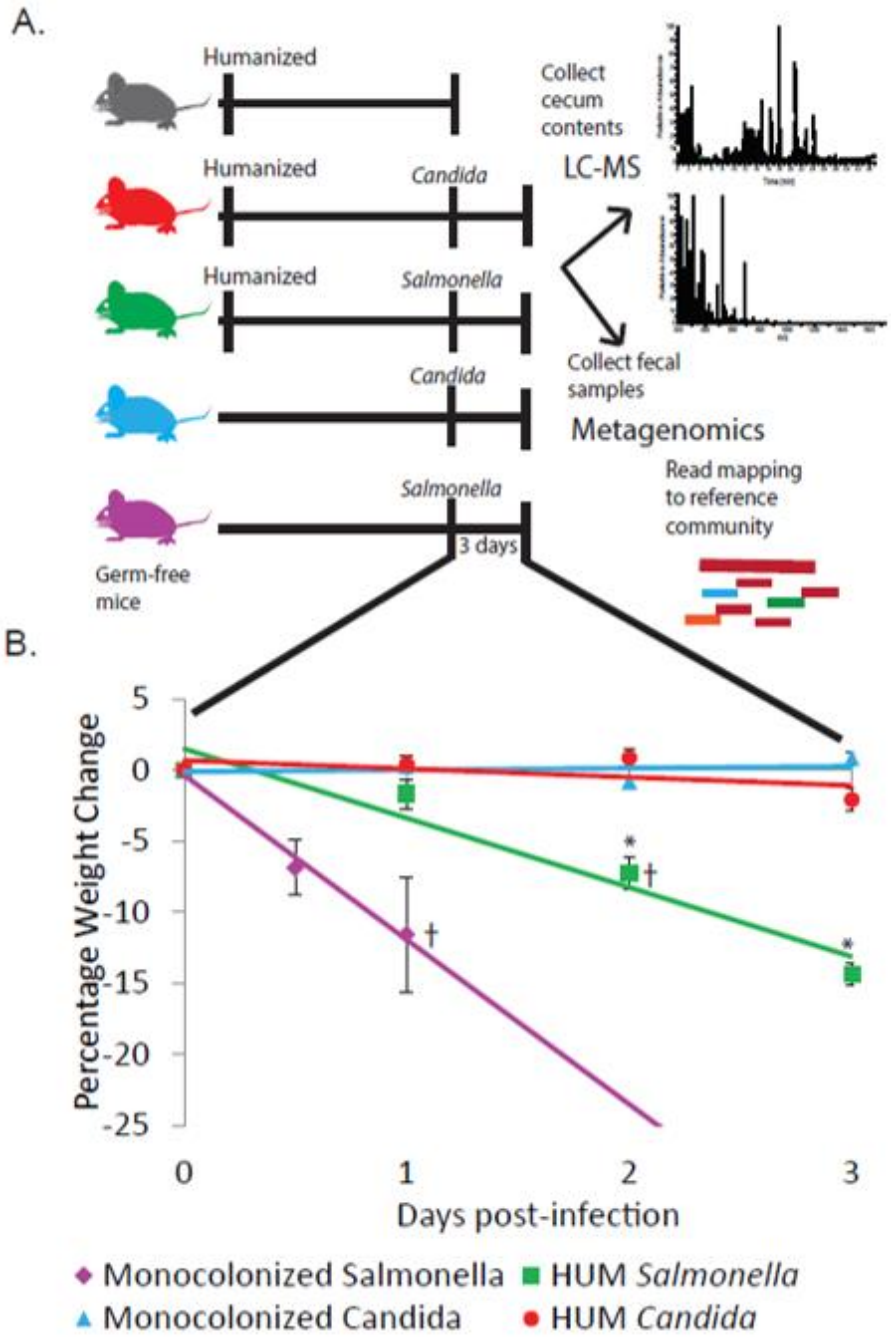



**Figure 1.** (A) Overview of experimental design (B) Percent body weight loss during three days of infection. Errors bars indicate standard error. Significant difference from HUM *Candida* ($p < 0.05$) using Wilcoxon test denoted by * next to relevant group. Mice sacrificed early indicated with † (Monocolonized *Salmonella*, 3 at 12hr and 3 at 24hr, HUM *Salmonella* 1 at 24 hr).

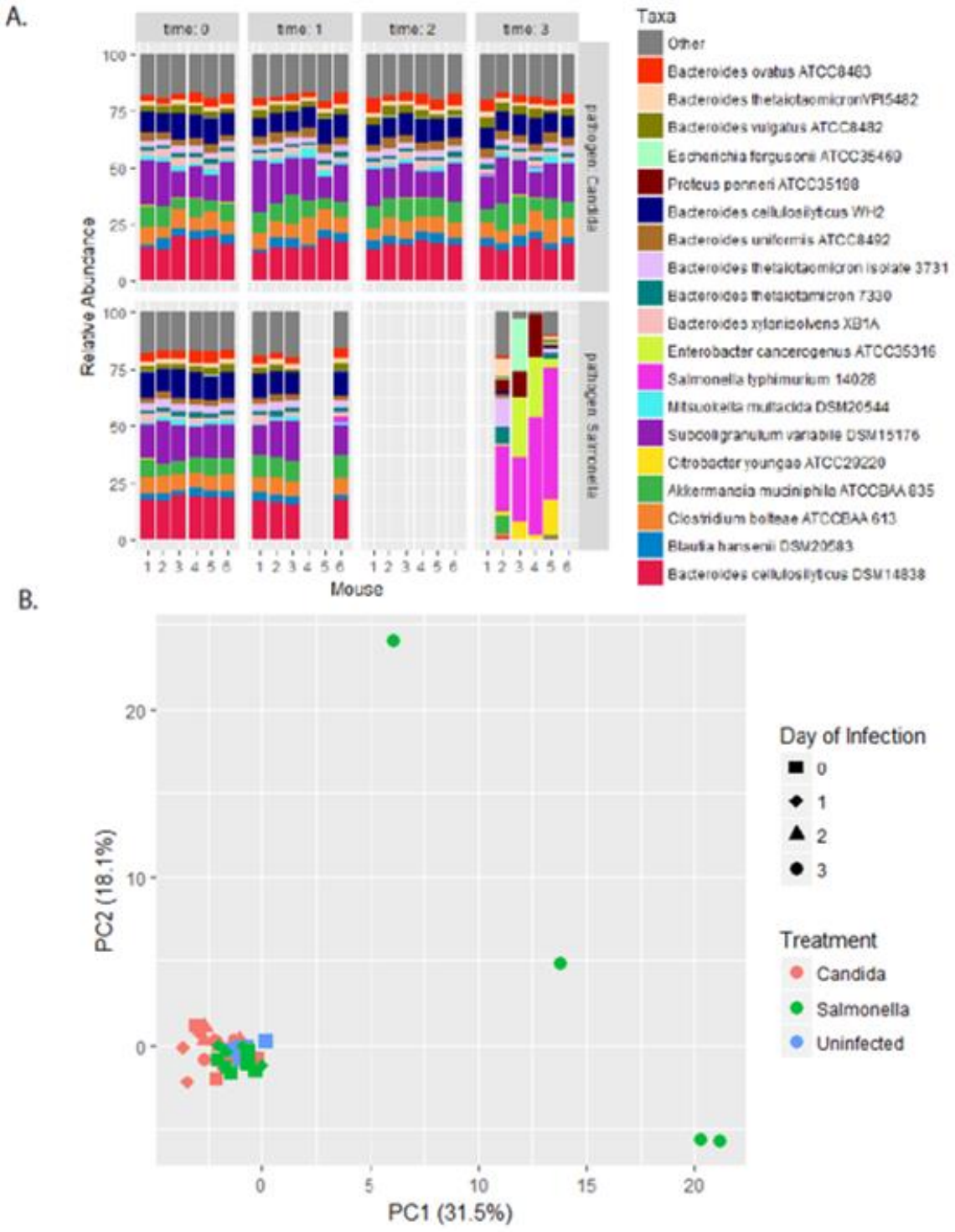



**Figure 2.** Variation in fecal microbiota metagenomes during infection. (A) Relative abundance of top 19 strains in HUM *Candida albicans* and *Salmonella enterica* Typhimurium infection group. (B) PCA of strain relative abundance.

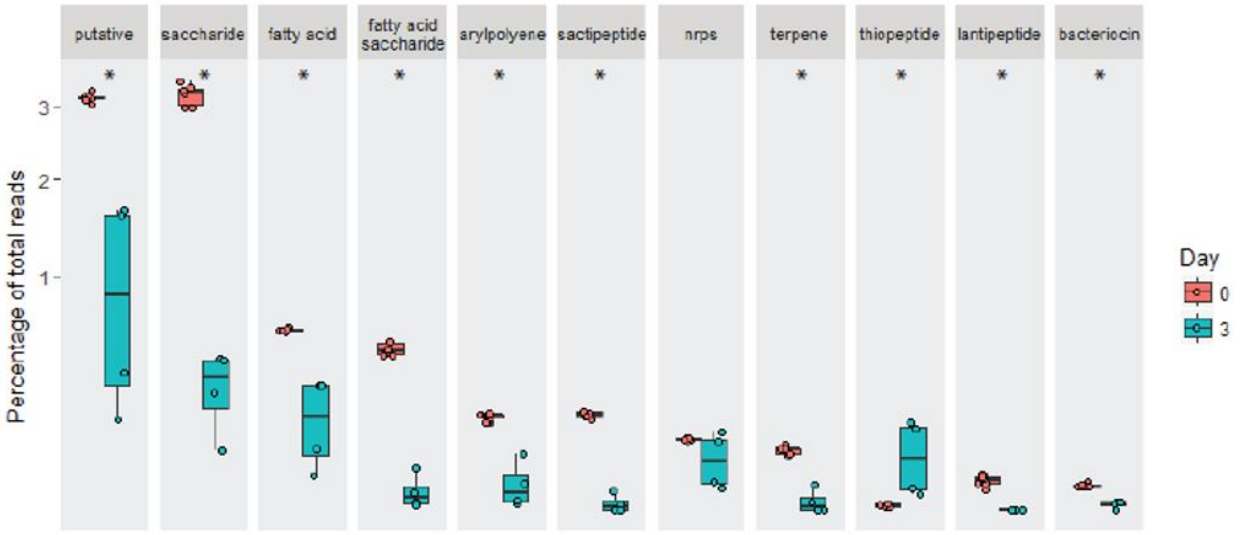



**Figure 3.** Percentage abundance of reads mapping to biosynthetic gene clusters out of total reads that were mapped from the metagenome from HUM *Salmonella* infected mice prior to infection (n=6) and three days into infection (n=4), on a square root adjusted axis. Significance ($p<0.05$ with Benjamini Hochberg correction) is indicated with *.

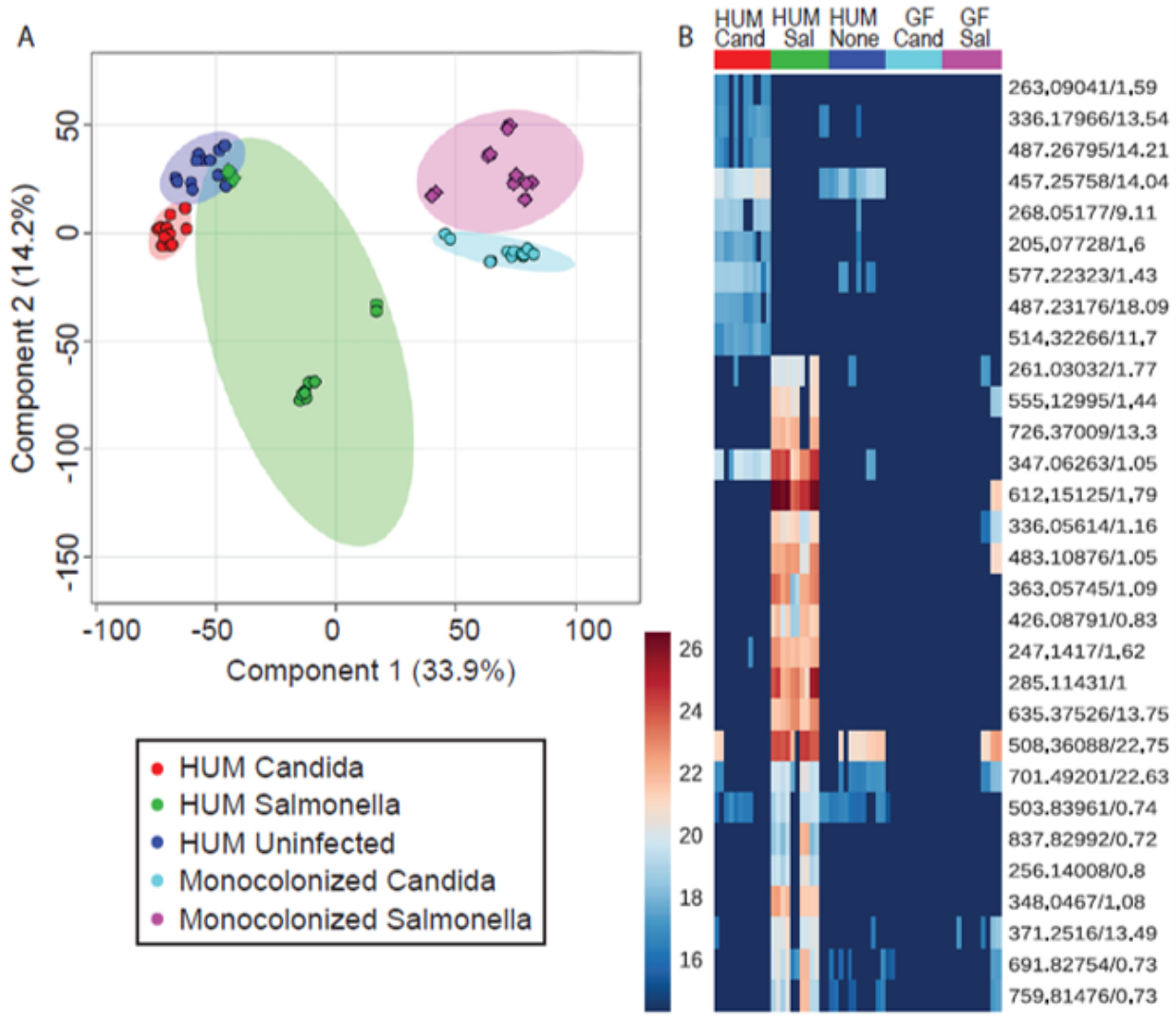



**Figure 4.** (A) PLS-DA of metabolites from all groups, with 95% confidence intervals. (B) Metabolites of interest 1.5x higher in HUM infected groups than uninfected mice, absent in 8/10 technical replicates for monocolonized mice. Circles are samples collected three post-infection, diamonds are from animals sacrificed 1 day post-infection.

## Supplemental Information

**Supplemental Table 1.** Strains used in humanized community

| Genus | Species | ATCC | DSMZ | Also known as |
|---|---|---|---|---|
| *Akkermansia* | *muciniphila* | BAA-835 | 22959 | |
| *Alistipes* | *indistinctus* | NA | 22520 | |
| *Anaerococcus* | *hydrogenalis* | 49630 | 7454 | |
| *Anaerotruncus* | *colihominis* | na | 17241 | |
| *Bacteroides* | *caccae* | 43185 | 19024 | |
| *Bacteroides* | *cellulosilyticus* | na | 14838 | |
| *Bacteroides* | *coprophilus* | na | 18228 | |
| *Bacteroides* | *dorei* | na | 17855 | |
| *Bacteroides* | *eggerthii* | 27754 | 20697 | |
| *Bacteroides* | *finegoldii* | na | 17565 | |
| *Bacteroides* | *intestinalis* | na | 17393 | |
| *Bacteroides* | *ovatus* | 8483 | na | |
| *Bacteroides* | *plebeius* | na | 17135 | |
| *Bacteroides* | *stercoris* | 43183 | na | |
| *Bacteroides* | *thetaiotaomicron3731* | na | na | |
| *Bacteroides* | *thetaiotaomicron7330* | na | na | |
| *Bacteroides* | *thetaiotaomicronVPI-5482* | 29148 | na | |
| *Bacteroides* | *uniformis* | 8492 | na | |
| *Bacteroides* | *vulgatus* | 8482 | na | |
| *Bacteroides* | *WH2* | na | na | Bacteroides thetaiotamicron; Bacteroides cellulolyticus |
| *Bacteroides* | *xylanisolvens* | na | 18836 | |
| *Bifidobacterium* | *adolescentis* | 15703 | na | |
| *Bifidobacterium* | *angulatum* | 27535 | 20098 | |
| *Bifidobacterium* | *bifidum* | 29521 | 20456 | |
| *Bifidobacterium* | *dentium* | 27678 | na | |
| *Bifidobacterium* | *pseudocatenulatum* | 27919 | 20438 | |
| *Blautia* | *hansenii* | 27752 | 20583 | |
| *Blautia* | *luti* | na | 14534 | |
| *Catenibacterium* | *mitsuokai* | na | 15897 | |
| *Citrobacter* | *youngae* | 29220 | na | |
| *Clostridium* | *asparagiforme* | na | 15981 | |
| *Clostridium* | *bartlettii* | na | 16795 | Intestinibacter bartlettii |
| *Clostridium* | *bolteae* | BAA-613 | 15670 | |
| *Clostridium* | *hathewayi* | na | 13479 | |

| | | | | |
|---|---|---|---|---|
| *Clostridium* | *hiranonis* | na | 13275 | |
| *Clostridium* | *hylemonae* | na | 15053 | |
| *Clostridium* | *leptum* | 29065 | 753 | |
| Clostridium | M62_1 | na | na | |
| *Clostridium* | *nexile* | 27757 | 1787 | |
| *Clostridium* | *nexile-related* | na | na | Tyzzerella nexilis |
| *Clostridium* | *ramosum* | 25582 | 1402 | |
| *Clostridium* | *scindens* | 35704 | 5676 | |
| *Clostridium* | *spiroforme* | 29900 | 1552 | |
| *Clostridium* | *sporogenes* | 15579 | na | |
| *Clostridium* | *symbiosum* | 14940 | 934 | |
| *Collinsella* | *aerofaciens* | 25986 | 3979 | |
| Collinsella | aerofaciens | 25986 | 3979 | |
| *Collinsella* | *intestinalis* | na | 13280 | |
| *Collinsella* | *stercoris* | na | 13279 | |
| *Coprococcus* | *comes* | 27758 | na | |
| *Coprococcus* | *eutactus* | 27759 | na | |
| *Desulfovibrio* | *piger* | na | na | GOR1 |
| *Dorea* | *formicigenerans* | 27755 | 3992 | |
| *Dorea* | *longicatena* | na | 13814 | |
| *Edwardsiella* | *tarda* | 23685 | na | |
| Edwardsiella | tarda | 23685 | na | |
| *Enterobacter* | *cancerogenus* | 35316 | na | |
| *Escherichia* | *coliK12* | na | | |
| *Escherichia* | *fergusonii* | 35469 | 13698 | |
| *Eubacterium* | *biforme* | 27806 | 3989 | |
| *Eubacterium* | *cylindroides* | na | na | |
| *Eubacterium* | *dolichum* | 29143 | 3991 | |
| *Eubacterium* | *eligens* | 27750 | 3376 | |
| *Eubacterium* | *hallii* | 27751 | 3353 | |
| *Eubacterium* | *plautii* | 29863 | na | Clostridium orbscindens; Flavonifractor plautii |
| Eubacterium | rectale | 33656 | | |
| *Eubacterium* | *ventriosum* | 27560 | na | |
| *Faecalibacterium* | *prausnitzii M21/2* | na | na | |
| *Fusobacterium* | *varium* | na | na | JCM6320 |
| *Holdemania* | *filiformis* | 51649 | 12042 | |
| *Lactobacillus* | *reuteri* | na | 20016 | |
| *Lactobacillus* | *ruminis* | 27780 | 20403 | |
| *Marvinbryantia* | *formatexigens* | na | 14469 | |
| *Megamonas* | *funiformis* | na | 19343 | |
| *Mitsuokella* | *multacida* | 27723 | 20544 | |
| *Parabacteroides* | *distasonis* | 8503 | 20701 | |

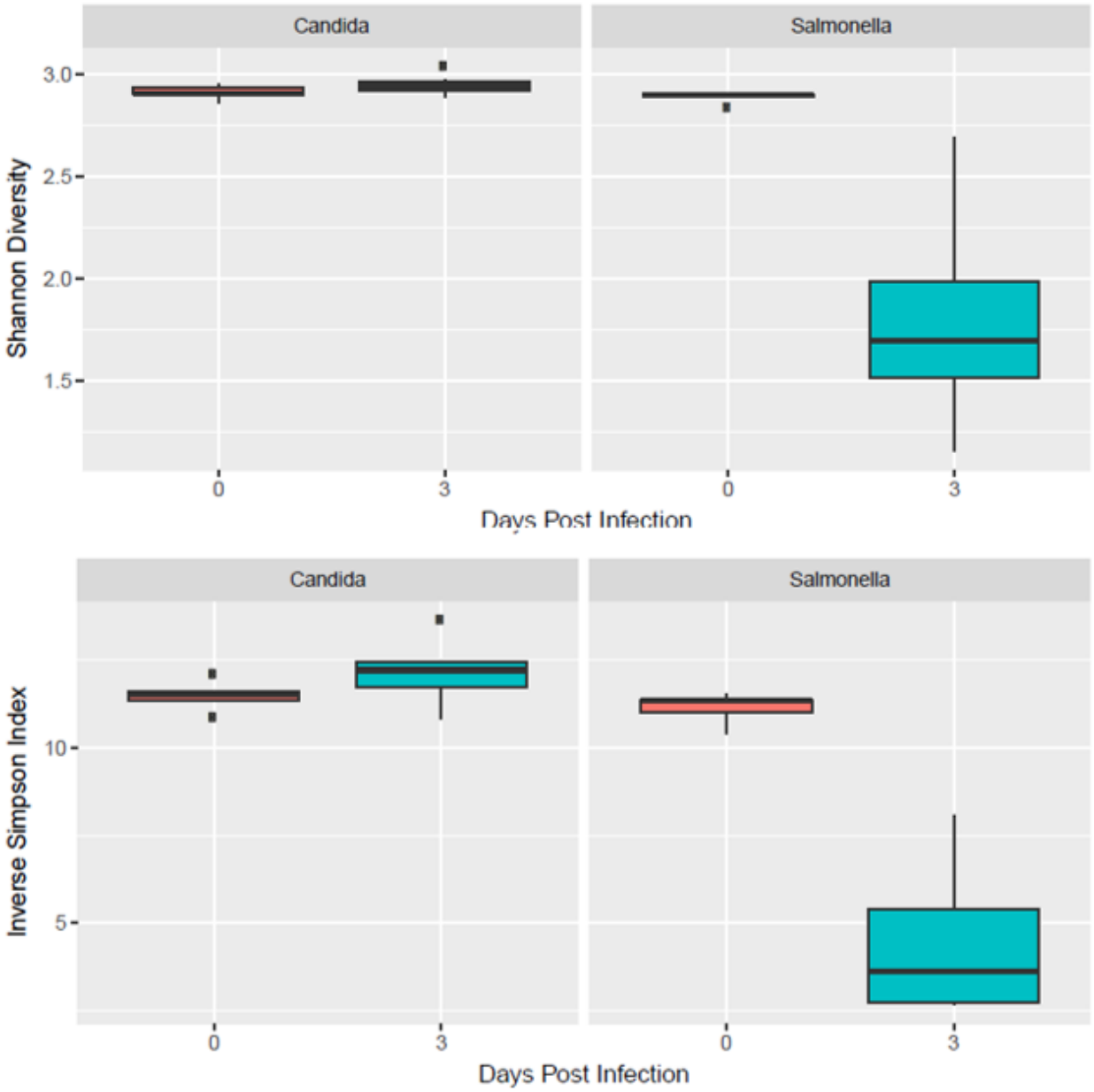
| | | | | |
|---|---|---|---|---|
| *Parabacteroides* | *johnsonii* | na | 18315 | |
| *Parabacteroides* | *merdae* | 43184 | 19495 | |
| *Proteus* | *penneri* | 35198 | na | |
| *Providencia* | *alcalifaciens* | na | na | |
| *Providencia* | *rettgeri* | na | 1131 | |
| *Providencia* | *rustigianii* | 33673 | 4541 | |
| *Providencia* | *stuartii* | 25827 | na | Clostridium sp. GM2/1 |
| *Roseburia* | *intestinalis* | na | 14610 | |
| *Ruminococcus* | *gnavus* | 29149 | na | |
| *Ruminococcus* | *hydrogenotrophicus* | na | 10507 | Blautia hydrogenotrophicus |
| *Ruminococcus* | *lactaris* | 29176 | na | |
| *Ruminococcus* | *obeum* | na | na | |
| *Ruminococcus* | *torques* | 27756 | na | |
| *Streptococcus* | *infantarius* | BAA-102 | na | |
| *Subdoligranulum* | *variabile* | na | 15176 | |
| *Victivallis* | *vadensis* | BAA-548 | 14823 | |

**Supplemental Table 2.** Biosynthetic gene clusters predicted in humanized microbiota

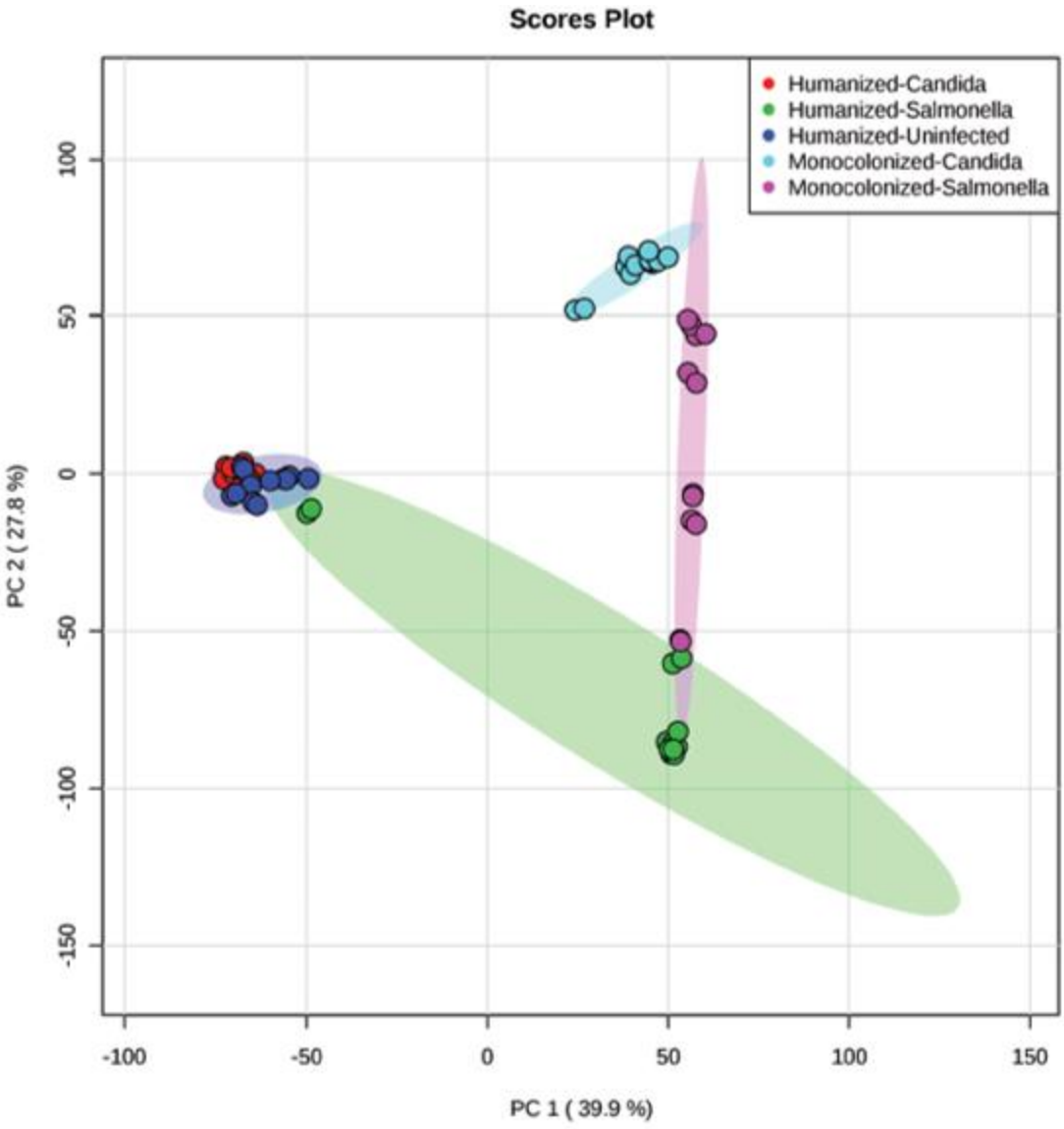
| **Cluster Name** | **Count** |
|---|---|
| putative | 486 |
| saccharide | 345 |
| fatty_acid | 117 |
| sactipeptide | 34 |
| fatty acid -saccharide | 22 |
| nrps | 19 |
| arylpolyene | 14 |
| thiopeptide | 12 |
| bacteriocin | 10 |
| siderophore | 4 |
| lantipeptide | 3 |
| other | 3 |
| resorcinol | 3 |
| terpene | 2 |
| hserlactone | 2 |
| sactipeptide-cf_saccharide | 1 |
| sactipeptide-nrps | 1 |
| bacteriocin-proteusin | 1 |
| sactipeptide-lantipeptide | 1 |
| t1pks-nrps | 1 |

**Supplemental Table 3.** Features of interested in humanized infected mice

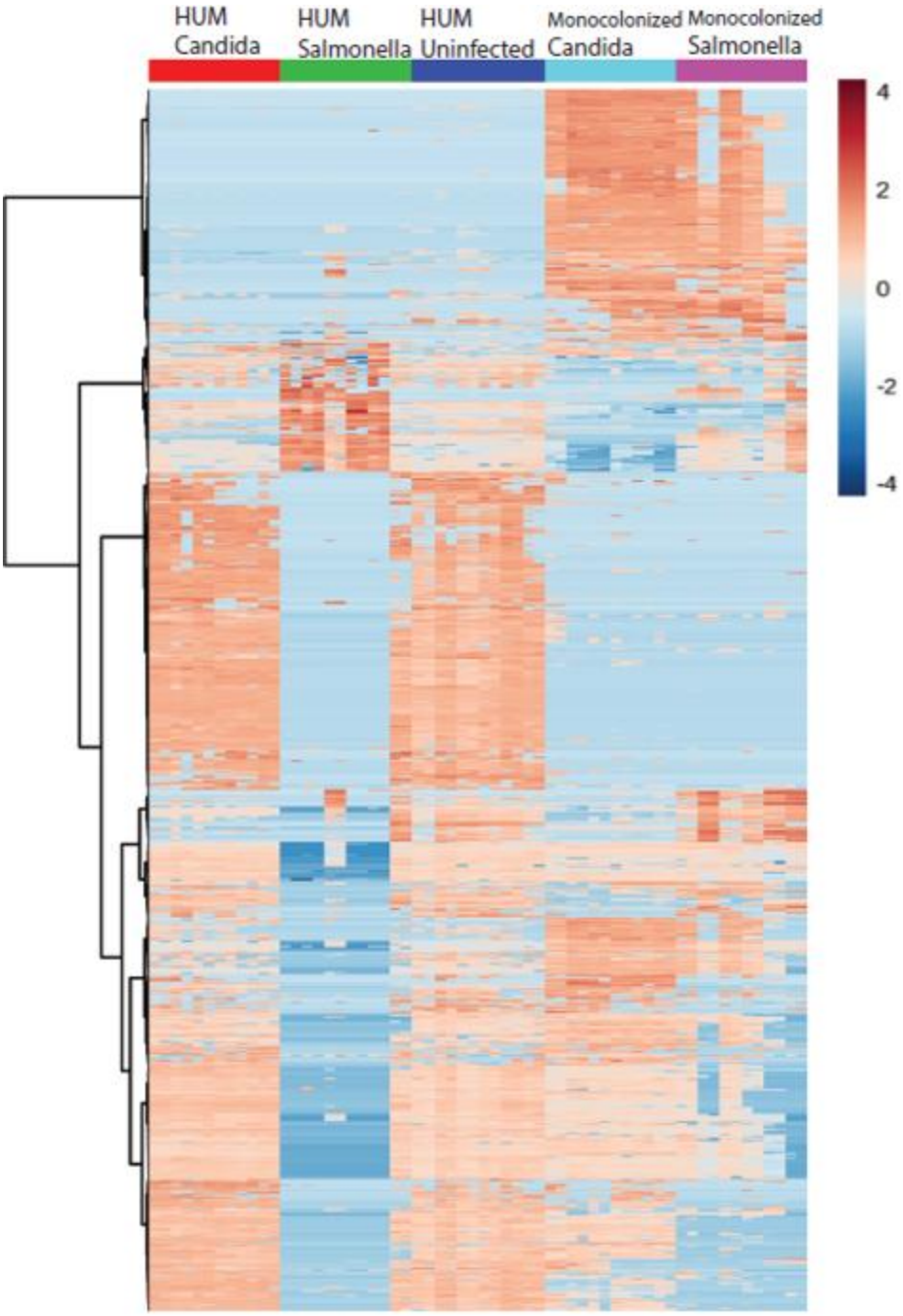
| Molecular weight | Retention time (min) | charge | MS/MS Identification | Overabundant in Salmonella or Candida |
|---|---|---|---|---|
| 247.1417 | 1.391 | 1 | Hydroxybutyrylcarnitine | Salmonella |
| 247.1417 | 1.625 | 1 | Hydroxybutyrylcarnitine | Salmonella |
| 256.1401 | 0.799 | 1 | | Salmonella |
| 261.0303 | 1.772 | 1 | | Salmonella |
| 285.1143 | 1.002 | 1 | | Salmonella |
| 336.0561 | 1.16 | 1 | | Salmonella |
| 347.0626 | 1.052 | 1 | | Both |
| 348.0467 | 1.08 | 1 | Inosine monophosphate | Salmonella |
| 363.0575 | 1.093 | 1 | | Salmonella |
| 371.2516 | 13.487 | 1 | | Salmonella |
| 426.0879 | 0.832 | 1 | Glutathione-cysteine disulfide | Salmonella |
| 483.1088 | 1.048 | 2 | | Salmonella |
| 503.8396 | 0.741 | 1 | | Salmonella |
| 508.3609 | 22.75 | 1 | | Salmonella |
| 555.13 | 1.437 | 2 | | Salmonella |
| 612.1513 | 1.789 | 1 | Glutathione disulfide | Salmonella |
| 635.3753 | 13.747 | 2 | | Salmonella |
| 691.8275 | 0.733 | 1 | | Salmonella |
| 701.492 | 22.63 | 1 | | Salmonella |
| 726.3701 | 13.302 | 2 | | Salmonella |
| 759.8148 | 0.731 | 1 | | Salmonella |
| 837.8299 | 0.72 | 1 | | Salmonella |
| 205.0773 | 1.597 | 1 | | Candida |
| 263.0904 | 1.592 | 1 | | Candida |
| 268.0518 | 9.114 | 1 | | Candida |
| 336.1797 | 13.536 | 1 | | Candida |
| 457.2576 | 14.042 | 1 | | Candida |
| 487.2318 | 18.094 | 1 | | Candida |
| 487.268 | 14.205 | 1 | | Candida |
| 514.3227 | 11.698 | 2 | | Candida |
| 577.2232 | 1.429 | 1 | | Candida |



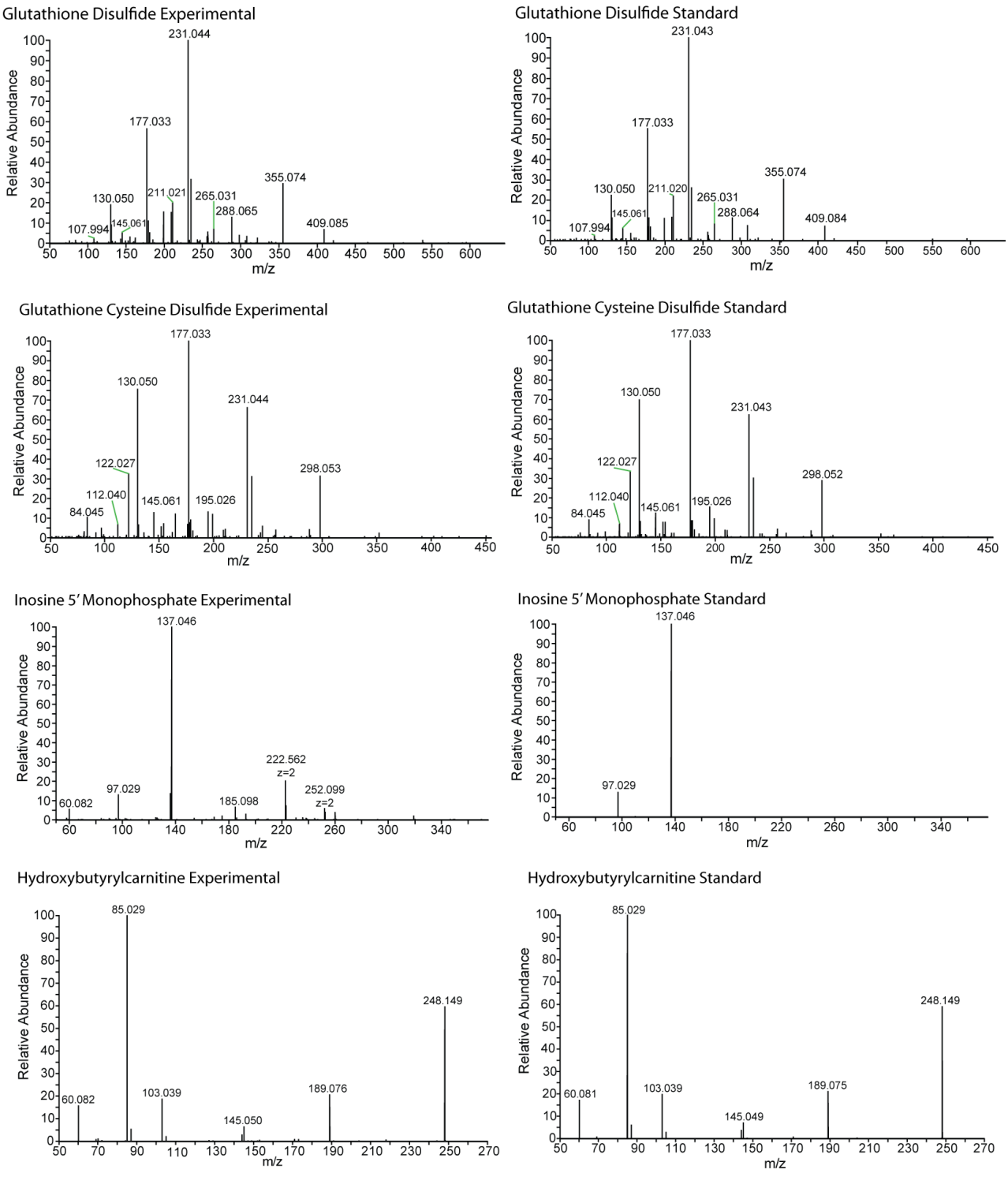

**Supplemental Figure 1**. Community diversity before and 3 days post-infection.




**Supplemental Figure 2**. Principal component analysis of all metabolites detected.




**Supplemental Figure 3.** Heatmap of all *m/z* detected.




**Supplemental Figure 4**. MS/MS spectra of experimental compounds and matching standards.

# Chapter 8

# Combined Metabolomics, Peptidomics, and Proteomics Profiling of Gut Microbial Responses to Infection by *Salmonella enterica* Typhimurium and *Candida albicans*

Keywords: LC-MS; Metabolomics; Peptidomics; Proteomics; Microbiome; Infection; Multiomics

**Abstract**

The gut microbiome plays many critical roles in human health, including immune system modulation and defense against disruptive pathogens. Mass spectrometry is a sensitive and robust technique to profile not only small molecule changes in the microbiome due to infection, but also peptide and protein changes due to infection. Thus, multiomics studies, such as metabolomics and proteomics studies, are commonly employed to investigate the role of the microbiome in health and disease. Here, a combination of metabolomics, peptidomics, and proteomics was used to profile the effect of the microbiome during infection by both *Salmonella enterica* Typhimurium and *Candida albicans*. Peptidomics was integrated with metabolomics and proteomics to study host–microbiota-pathogen interaction. The identified bioactive peptides, endogenous peptides Fibrinopeptide B and Thymosin alpha present in the gut could provide insight into microbiome mediated communication between the gut and the brain. Metabolomics and proteomics studies also reveal molecular changes that highlight the role of the microbiome in the immune and inflammatory response to infection.

**Introduction**

In the gastrointestinal tract, a large and dynamic community of microbes, the gut microbiome, reside, where three domains of life, namely archaea, bacteria, and eukaryote, are represented.[1, 2] These microbes have an important role in human health. They affect the nutrition of the host as they carry out various digestive processes, and they modulate immune function.[3, 4] Additionally, disruptions to the normal composition of the microbiome are associated with various diseases, including obesity,[5, 6] inflammatory bowel disease,[7] and cardiovascular disease.[8, 9] Consequently, a variety of "omic" techniques have been employed to study the microbiome. For

example, there are metagenomic and metaproteomic studies on the gut microbiome as well as studies on how the microbiome influences metabolites in other systems of the body, such as the blood and cerebrospinal fluid.[10-13]

Microbial metabolism produces both primary and secondary metabolites through the gut microbiome metagenome that influence health and disease.[4] The role of metabolites produced by the microbiome is not limited to nutrition and metabolism, but can also provide defense against pathogens, and immune regulation among many other functions.[14, 15] Furthermore, the ability of secondary metabolites synthesized by the microbiome to provide protection to the host against pathogens could potentially be used for discovery of new natural products, which is of interest due to the rise of antibiotic resistance.[16] Mass spectrometry (MS) is a powerful tool to study metabolites in a sample, and in conjunction with nuclear magnetic resonance (NMR) can aid in natural product discovery.[17] For example, metabolomic changes of the microbiome in response to disturbances to the microbiome has been profiled with MS techniques.[18, 19]

Metaproteomics studies on the microbiome provide valuable information on expressed proteins involved in various biological processes. For example, metaproteomics has been used to study the metabolism of the gut microbiome in premature infants,[20] and the proteome changes in the microbiome due to disease, including Crohn's disease and inflammatory bowel disease.[21, 22] Additionally, metaproteomics is often coupled with metabolomics, as well as other "omics" for multiomics analysis of the biological processes occurring in the gut.[23] For example, metagenomics and metabolomics are often combined to study how different disruptions to the gut microbiome affect microbiome composition and metabolism.[19, 24] Metaproteomics studies can also be used with metagenomic and metabolomics techniques to provide a comprehensive analysis of active processes in the gut microbiome or of the effect of stressors on the microbiome.[25, 26]

Here a multiomics approach will be applied to study the response of a model human gut microbiome to infection by *Salmonella enterica* Typhimurium and *Candida albicans* in mice. Previously a metagenomics and metabolomics approach was utilized to investigate microbiome community and small molecule changes in the gut microbiome in response to infection.[27] To further understand the microbiome's response to infection, another multiomics approach combining metabolomics, peptidomics, and proteomics will be investigated here. The combination of peptidomics with metabolomics and proteomics is less commonly employed but will provide a great depth of information. Peptidomics analysis resulted in identification of multiple endogenous peptides in the gut microbiome. The combined metabolomics, peptidomics, and proteomics analysis of the gut microbiome resulted in discovery of immune system changes in response to infection.

**Experimental Methods**

Sample Collection and Extraction

Mouse experiments, including inoculation with approximately 90 strains representing a model human microbiome and infection with either *Salmonella enterica* Typhimurium (*Salmonella*) and *Candida albicans* (*Candida*), were performed as previously described.[27] A methanol chloroform water extraction was performed and a 3 kDa molecular weight cut-off performed on the aqueous and organic fractions.[27] The 3 kDa molecular weight cut-off device was equilibrated with 0.4 mL of 50% methanol and then flipped over and centrifuged (14,000 x g, 2 mins) to collect content above three kDa. A 30 kDa MWCO was performed on the contents above three kDa to separate the extract into peptidomics (three kDa to 30kDa) and proteomics (above 30kDa) fractions.

Metabolomics

Aqueous fractions (below 3 kDa) were resuspended at 10 mg/mL in optima grade water with 0.1% formic acid. Liquid chromatography-mass spectrometry (LC-MS) for label-free quantitation was performed as previously described.[27] LC-MS/MS (tandem MS) was performed on samples with the same gradient as the MS1 data acquisition using a top 5 DDA method with inclusion lists for targets upregulated in each sample. Full MS parameters were 70 k resolution, 1E6 AGC target, 100 ms maximum injection time, and 200 to 1700 *m/z*. MS/MS parameters were 70 k resolution, 5E5 AGC target, 100 ms maximum injection time, and 1.0 isolation width. A 10 s dynamic exclusion was employed. Three technical replicates were collected for MS/MS runs, and each technical replicate had a different collision energy (25, 30, and 40 were used).

Metabolomics Data Analysis

MS1 and MS/MS data was analyzed in Compound Discoverer 2.1 software. Initially, LC-MS data was analyzed. Retention time alignment was performed using adaptive curve with 5 ppm mass accuracy and 2.0 min max retention time shift. All following steps utilized 5 ppm mass accuracy and 0.2 retention time error unless otherwise stated. The workflow nodes were as follows: detect unknown compounds (30% intensity tolerance, 3 S/N ratio, and 1000000 min peak intensity), group unknown compounds, fill gasps (1.5 S/N threshold), normalize areas (constant sum), mark background compounds, predict compositions, differential analysis, and assign compound annotations. In addition to group ratios, the 6 biological replicates (each with two technical replicates) were also used to calculate biological replicate ratios. *Salmonella* targets were selected based upon 2-fold down regulation of the HumUn/HumSal and MonoSal/HumSal ratio with p-value less than 0.05. Additionally, the both the HumUn/HumSal and MonoSal/HumSal ratios needed be less than 0.5 in 4 out of the 6 biological replicate ratios. *Candida* targets were

selected based upon the same workflow but using the HumUn/HumCand and the MonoCand/HumCand ratios. LC-MS/MS data was analyzed in Compound Discoverer 2.1 software with the same parameters as the LC-MS data, except search mzCloud and search Chemspider nodes were added. The mzCloud search was conducted against all compound classes and activation types with a 20 CE activation energy tolerance and 10 ppm parent and fragment mass tolerance. Identifications in mzCloud were manually verified by checking fragmentation patterns of the top hits. Additional identifications were made by searching the Metlin spectral database with 5 ppm mass accuracy against target molecular weights. MS/MS database spectra were manually compared against experimental spectra. LC-MS normalized peak areas data from Compound Discoverer was imported into Metaboanalyst[28, 29] with log transformation to generate principal component analysis results.

Peptidomics Acquisition and Analysis

The sample preparation, data acquisition and data analysis of peptide fraction is similar as previously described (Chapter 6). Briefly, Sep-Pak C18 Cartridges was used for peptide desalting, and then peptide samples were dried down in a speed vac and resuspended in optima grade water with 3% acetonitrile and 0.1% formic acid. UPLC-MS/MS analysis was conducted using a Thermo Dionex UltiMateTM 3000 nanoLC system coupled to a Thermo Fusion Lumos Orbitrap MS. Specially, a decision tree-driven MS scheme was used here to allow improved peptide sequencing by alternating between Higher-energy collisional dissociation(HCD), Electron transfer dissociation (ETD) and Electron Transfer Higher Energy Dissociation (EThcD) for each specific charge state (Charge 2, combined HCD and EThcD, Charge 3-6 combined HCD and EThcD). The .raw data files from the Orbitrap MS analysis were searched against SwePep neuropeptide preprohormone database using PEAKS 8.5 software and a precursor tolerance of 10 ppm and a

fragment mass tolerance of 0.02 Da were allowed. Acetylation (N-term), amidation, oxidation (M), pyro-glu from E, pyro-glu from Q, sulfation (STY), were set as rare dynamic modifications and three maximum variable PTM were allowed per peptide. Confident peptide identification was threshold with Ascore (PTM site confidence) higher than 20, FDR lower than 1%, and at least one unique peptide was found.

Proteomics Acquisition and Analysis

The protein fractions from the supernatant contents above 30 kDa (aqueous and organic fractions) and from the pellets were combined. Ice-cold PBS was used to dissolved protein mixture samples to remove debris from the pellet at a low centrifuge speed (300 x g, 4$^{o}$C for 5 mins, 3 times). The supernatant was then centrifuged at 20,000 x g for 10 min to pellet the bacterial cells and host cells and then lysed with 4% SDS and 6 M urea Lysis buffer with sonication (On 15 sec, Off 30 sec, 3 cycles). Acetone protein precipitation was used to remove SDS from protein sample and the total protein concentration of each pellet was determined by BCA assay. 100 µg proteins were aliquoted and digested with Trypsin/Lys-C mixture overnight. Then, the digested proteins were desalted, dried down in a speed vac, and saved in the -80$^{o}$C until labeling. Twelve-plex isobaric DiLeu reagent were synthesized and used for labeling reaction as previously described.[30] The labeled samples were combined in equal ratios to form pooled 12-plex samples (Labeling map in **Supplemental Table S1**). Excess DiLeu reagents were removed from pooled 12-plex samples via SCX Ziptips (OMIX-SCX, Agilent, Santa Clara, CA) and high pH RPLC fractionation approach was used to improve protein identification coverage.[31] 4 fractions were collected from each pooled sample and dried down in a speed vac and saved in the -80$^{o}$C until LC-MS/MS analysis.

Fusion Lumos Orbitrap MS was also adopted here for protein fraction analysis. Full MS scans were acquired from *m/z* 300 to 1500 at a resolution of 60 K, AGC at $2 \times 10^5$, and maximum injection time (IT) of 100 ms. The top 20 precursors were then selected for higher-energy C-trap dissociation tandem mass spectrometry (HCD MS2) analysis with an isolation window of 1 *m/z*, a normalized collision energy (NCE) of 30, a resolving power of 60 K, an AGC target of $5 \times 10^4$, a maximum injection time of 100 ms, and a lower mass limit of 110 *m/z*. PEAKS software was used for protein identification and quantification. Trypsin with D&P enzyme was selected for this bottom-up proteomics study and three missed cleavages allowed. Searches were performed with a precursor mass tolerance of 25 ppm and a fragment mass tolerance of 0.02 Da. Static modifications consisted of DiLeu labels on peptide N-termini (+145.12801 Da) and carbamidomethylation of cysteine residues (+57.02146 Da). Dynamic modifications consisted of DiLeu labels on lysine residues, oxidation of methionine residues (+15.99492 Da). Peptide spectral matches (PSMs) were validated based on 1% FDR and quantitation was performed with a reporter ion integration tolerance of 20 ppm for the most confident centroid.

## Results

### Metabolomics

Label-free metabolomics was performed to investigate small molecule changes by the microbiome in response to both *Salmonella* and *Candida* infections. Humanized, *Salmonella* infected mice (HumSal) were compared to humanized, uninfected mice (HumUn) and germ-free mice monocolonized with *Salmonella* (MonoSal). Similarly, humanized, *Candida* infected mice (HumCand) were compared to humanized, uninfected mice (HumUn) and germ-free mice monocolonized with *Candida* (MonoCand). Volcano plots were used to analyze the differences in

the data by fold change and p-value. Additionally, potential targets were filtered based upon 4/6 biological replicates having ratios below the 2-fold threshold. The HumSal vs. HumUn and HumSal vs. MonoSal comparisons were analyzed separately with volcano plots as shown in **Figure 1(A)** and **Figure 1(B)**. Then, the overlap in the 2-fold downregulated compounds between two infections was compared with a Venn diagram, resulting in 384 small molecule targets for compounds upregulated in the humanized, *Salmonella* infected group (**Figure 1(C)**). For the *Candida* infection, the same analysis procedure resulted in 83 targets upregulated in the humanized, *Candida* infected mice (**Figure 1(D-F)**). The *Salmonella* targets are listed in **Supplemental Table S2** and the *Candida* targets are listed in **Supplemental Table S3**. Of the targets, 13 were in both the *Salmonella* list and the *Candida* list.

To investigate trends in the data in an unsupervised manner, principal component analysis (PCA) was utilized. PCA linear transforms the data to display the properties and variance of the sample data along the coordinate axis. **Figure 2** shows the PCA results for the data. In **Figure 2(A)**, the separation of the samples along the principal component axis is shown. The HumSal, MonoSal, and MonoCand groups separate well from the HumUn and HumCand groups, which correlates to the fact that the *Candida* infection was not as severe as the *Salmonella* infection.[27] In **Figure 2(B)**, the PCA loadings plots is shown, where *m/z* retention time points are plotted. Points on the outside contribute most to the variation in the data. The points in the bottom right corner are where the compounds up-regulated in the HumSal group are located. **Figure 2(C-E)** provides examples of selected *Salmonella* targets that are localized to this area of the PCA plot. The targets that were selected for both infections are slightly to the left of the *Salmonella* only targets, an example of which is in **Figure 2(F)**. Thus, the target list is composed of compounds that correlate well to the compounds causing the variance in the data as determined by the PCA plot.

LC-MS/MS data in combination with mzCloud and Metlin spectra libraries was used for compound identification. **Table 1** provides identifications for compounds detected in this study, with the associated error. The compounds listed in italics were previously identified as being up-regulated in the presence of the microbiome after infection.[27] Interestingly, multiple additional putative annotations were for carnitine molecules. Additional adenosine related compounds were also observed.

Peptidomics

The aqueous fraction contents between 3 and 30 kDa were analyzed with nanoLC-MS/MS to determine endogenous peptides in the gut produced in response to *Salmonella* infection. For peptidomic analysis, the data was searched against the SwePep neuropeptide preprohormone database using PEAKS 8.5 software. Overall, for *Salmonella* infection, 1166 peptides were identified in the SwePep neuropeptide database belonging to 23 preprohormones (**Supplemental Table S4**). For *Candida* identifications, 17 peptides were identified belonging to 6 preprohormones (**Supplemental Table S5**).

Changes in the peptidome between different sample groups were identified with label-free quantification. Peptide targets were selected based upon 1.5-fold upregulated in the HumSal group and an overall FDR less than 0.01 for both the HumSal vs. HumUn and the HumSal vs. MonoSal comparisons. Similarly, for the *Candida* infection, peptides were selected as targets if they were 1.5 fold upregulated in the HumCand vs. HumUn and HumCand vs. MonoCand comparisons with an overall FDR less than 0.01. In total 877 peptides were found to be upregulated in the HumSal group and 1 peptide was found to be upregulated in the HumCand group. In the HumSal up-regulated peptides, the intact peptide of fibrinopeptide B and a truncated peptide of Thymosin alpha were identified.

Proteomics

For proteomics analysis, nanoLC-MS/MS with 12-plex DiLeu labeling for relative quantitation was employed.[30] PEAKS Studio software package was used for data analysis of the 12-plex DiLeu labeled samples. From the PEAKS analysis, 1318 proteins were identified from 6821 peptide spectral matches (PSMs). Up-regulated proteins were selected based upon 1.5-fold changes in the HumSal or HumCand groups compared to their respective controls (HumUn and either MonoSal or MonoCand). This results in 14 up-regulated proteins for HumSal. The gene ontology (GO) analysis of the up-regulated proteins in the HumSal group (**Figure 3**), display the biological process and molecular functions of the up-regulated proteins. Three proteins, Alpha-1-acid glycoprotein, Alpha-2-HS-glycoprotein, and Haptoglobin, were detected, which modulate the activity of the immune system during the acute-phase reaction. No proteins were found to be significantly up-regulated in HumCand group, which also correlates to the fact that the *Candida* infection was not as severe as the *Salmonella* infection.

**Discussion**

Metabolites identified that are up-regulated in response to infection and in the presence of the model microbiome provide evidence of metabolism changes induced by the microbiome in response to infection. Here, we focused on *m/z* greater than 200 to limit detection of common primary metabolites and instead focus on larger secondary metabolites and potential natural products. Previously, importance of glutathione metabolism against potential oxidative stress was discussed due to the presence of glutathione disulfide and glutathione cysteine disulfide after *Salmonella* infection.[27] Multiple carnitine molecules were identified in this study. Acetyl-L-carnitine is involved in energy metabolism; it increases acetyl-CoA uptake in the mitochondria in

fatty acid oxidation, increases acetylcholine production, and is a free radical scavenger.[32-34] Acetyl-L carnitine can cross the blood brain barrier easier compared to carnitine, and consequently, its positive effects for various neurological disorders, including depression and Alzheimer's disease, has been observed.[35, 36] Acylcarnitines are involved in β-oxidation through transportation of long chain fatty acids across the mitochondrial membrane.[37] Studies have focused on various biological and metabolic roles of acylcarnities,[38] including potential proinflammatory roles.[38, 39]

From the peptidomics analysis, intact endogenous peptide fibrinopeptide B and truncated endogenous peptide Thymosin alpha were identified. Both augment host immune defense. Fibrinopeptide B is formed by cleavage of fibrinogen by thrombin during the coagulation process.[40] Previous studies have shown fibrinopeptide B involvement in the inflammatory processes.[40] Investigations with sepsis, rheumatoid arthritis, and pilocytic astrocytoma, which are pathogenic inflammatory conditions, found increases in fibrinogen degradation fragments and fibrinopeptide B.[41-43] Additionally, Thymosin alpha is involved in restoration of immune functions as it can modulate both the innate and adaptive immune system.[44] The stimulation of the adaptive immune function is critical for fighting infections, both viral and bacterial, and other diseases.[45] Thus, fibrinopeptide B and Thymosin alpha stimulation by the human microbiome during infection could enhance cell-mediated immunity. These signaling peptides induced by model human microbiome during infection furthers understanding of the bidirectional communication between gut microbiota, and immune system.

The proteomics results also display up-regulated proteins after infection that are involved in the immune response and regulation of the inflammatory response (**Figure 3**). In particular, alpha-1-acid glycoprotein is an acute phase protein that has been shown to be involved in immune modulation and in the inflammatory response.[46] The alpha-1-acid glycoprotein has been shown to

have both pro-inflammatory and anti-inflammatory effects.[47] Overall, metabolomics, peptidomics, and proteomics analysis reveals microbiome-mediated immune and inflammatory responses in response to *Salmonella* infection. Further analysis of the detected up-regulated molecules, including peptidomic analysis of endogenous peptides from the organic fraction of the extraction, could reveal additional systems and endogenous peptides involved in the response to infection.

**Acknowledgements**

Jenny Bratburd performed mouse experimental work. Pingli Wei performed peptidomics and proteomics experiments and data analysis. Caitlin Keller performed metabolomics experiments and data analysis. This work was supported by T32AI55397 and U19AI109673 from the National Institute of Health, and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin–Madison with funding from the Wisconsin Alumni Research Foundation. This work was supported in part by grants NIH DK108259 (to F.E.R.). L. L. acknowledges funding from the NIH (R01DK071801, RF1AG052324), a Vilas Distinguished Achievement Professorship, and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy. The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531 to LL) and Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. We would like to thank Eugenio Vivas for assistance in mouse experimental procedures.

**References**

1. Hoffmann, C.; Dollive, S.; Grunberg, S.; Chen, J.; Li, H.; Wu, G. D.; Lewis, J. D.; Bushman, F. D., Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One* **2013,** *8* (6), e66019.

2. Scanlan, P. D.; Marchesi, J. R., Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *Isme j* **2008,** *2* (12), 1183-93.

3. Kau, A. L.; Ahern, P. P.; Griffin, N. W.; Goodman, A. L.; Gordon, J. I., Human nutrition, the gut microbiome and the immune system. *Nature* **2011,** *474* (7351), 327-336.

4. Sharon, G.; Garg, N.; Debelius, J.; Knight, R.; Dorrestein, P. C.; Mazmanian, S. K., Specialized metabolites from the microbiome in health and disease. *Cell Metab* **2014,** *20* (5), 719-730.

5. Turnbaugh, P. J.; Backhed, F.; Fulton, L.; Gordon, J. I., Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* **2008,** *3* (4), 213-23.

6. Rosenbaum, M.; Knight, R.; Leibel, R. L., The gut microbiota in human energy homeostasis and obesity. *Trends Endocrinol Metab* **2015,** *26* (9), 493-501.

7. Kostic, A. D.; Xavier, R. J.; Gevers, D., The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* **2014,** *146* (6), 1489-1499.

8. Tang, W. H.; Kitai, T.; Hazen, S. L., Gut Microbiota in Cardiovascular Health and Disease. *Circ Res* **2017,** *120* (7), 1183-1196.

9. Jie, Z.; Xia, H.; Zhong, S. L.; Feng, Q.; Li, S.; Liang, S.; Zhong, H.; Liu, Z.; Gao, Y.; Zhao, H.; Zhang, D.; Su, Z.; Fang, Z.; Lan, Z.; Li, J.; Xiao, L.; Li, R.; Li, X.; Li, F.; Ren, H.; Huang, Y.; Peng, Y.; Li, G.; Wen, B.; Dong, B.; Chen, J. Y.; Geng, Q. S.; Zhang, Z. W.; Yang, H.; Wang, J.; Zhang, X.; Madsen, L.; Brix, S.; Ning, G.; Xu, X.; Liu, X.; Hou, Y.; Jia, H.; He, K.; Kristiansen, K., The gut microbiome in atherosclerotic cardiovascular disease. *Nat Commun* **2017,** *8* (1), 845.

10. Turnbaugh, P. J.; Hamady, M.; Yatsunenko, T.; Cantarel, B. L.; Duncan, A.; Ley, R. E.; Sogin, M. L.; Jones, W. J.; Roe, B. A.; Affourtit, J. P.; Egholm, M.; Henrissat, B.; Heath, A. C.; Knight, R.; Gordon, J. I., A core gut microbiome in obese and lean twins. *Nature* **2009,** *457* (7228), 480-U7.

11. Verberkmoes, N. C.; Russell, A. L.; Shah, M.; Godzik, A.; Rosenquist, M.; Halfvarson, J.; Lefsrud, M. G.; Apajalahti, J.; Tysk, C.; Hettich, R. L.; Jansson, J. K., Shotgun metaproteomics of the human distal gut microbiota. *Isme Journal* **2009,** *3* (2), 179-189.

12. Wikoff, W. R.; Anfora, A. T.; Liu, J.; Schultz, P. G.; Lesley, S. A.; Peters, E. C.; Siuzdak, G., Metabolomics analysis reveals large effects of gut microflora on mammalian blood

metabolites. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (10), 3698-3703.

13. Vogt, N. M.; Romano, K. A.; Darst, B. F.; Engelman, C. D.; Johnson, S. C.; Carlsson, C. M.; Asthana, S.; Blennow, K.; Zetterberg, H.; Bendlin, B. B.; Rey, F. E., The gut microbiota-derived metabolite trimethylamine N-oxide is elevated in Alzheimer's disease. *Alzheimer's Research & Therapy* **2018,** *10* (1), 124.

14. Brestoff, J. R.; Artis, D., Commensal bacteria at the interface of host metabolism and the immune system. *Nat Immunol* **2013,** *14* (7), 676-84.

15. Clarke, T. B.; Davis, K. M.; Lysenko, E. S.; Zhou, A. Y.; Yu, Y.; Weiser, J. N., Recognition of peptidoglycan from the microbiota by Nod1 enhances systemic innate immunity. *Nat Med* **2010,** *16* (2), 228-31.

16. Wilson, M. R.; Zha, L.; Balskus, E. P., Natural product discovery from the human microbiome. *J Biol Chem* **2017,** *292* (21), 8546-8552.

17. Wolfender, J. L.; Marti, G.; Thomas, A.; Bertrand, S., Current approaches and challenges for the metabolite profiling of complex natural extracts. *J Chromatogr A* **2015,** *1382*, 136-64.

18. Zheng, X. J.; Xie, G. X.; Zhao, A. H.; Zhao, L. J.; Yao, C.; Chiu, N. H. L.; Zhou, Z. X.; Bao, Y. Q.; Jia, W. P.; Nicholson, J. K.; Jia, W., The Footprints of Gut Microbial-Mammalian Co-Metabolism. *Journal of Proteome Research* **2011,** *10* (12), 5512-5522.

19. Lu, K.; Abo, R. P.; Schlieper, K. A.; Graffam, M. E.; Levine, S.; Wishnok, J. S.; Swenberg, J. A.; Tannenbaum, S. R.; Fox, J. G., Arsenic exposure perturbs the gut microbiome and its metabolic profile in mice: an integrated metagenomics and metabolomics analysis. *Environ Health Perspect* **2014,** *122* (3), 284-91.

20. Xiong, W.; Brown, C. T.; Morowitz, M. J.; Banfield, J. F.; Hettich, R. L., Genome-resolved metaproteomic characterization of preterm infant gut microbiota development reveals species-specific metabolic shifts and variabilities during early life. *Microbiome* **2017,** *5* (1), 72.

21. Erickson, A. R.; Cantarel, B. L.; Lamendella, R.; Darzi, Y.; Mongodin, E. F.; Pan, C.; Shah, M.; Halfvarson, J.; Tysk, C.; Henrissat, B.; Raes, J.; Verberkmoes, N. C.; Fraser, C. M.; Hettich, R. L.; Jansson, J. K., Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. *PLoS One* **2012,** *7* (11), e49138.

22. Presley, L. L.; Ye, J.; Li, X.; Leblanc, J.; Zhang, Z.; Ruegger, P. M.; Allard, J.; McGovern, D.; Ippoliti, A.; Roth, B.; Cui, X.; Jeske, D. R.; Elashoff, D.; Goodglick, L.; Braun, J.; Borneman, J., Host-microbe relationships in inflammatory bowel disease detected by bacterial and metaproteomic analysis of the mucosal-luminal interface. *Inflamm Bowel Dis* **2012,** *18* (3), 409-17.

23. Mondot, S.; Lepage, P., The human gut microbiome and its dysfunctions through the meta-omics prism. *Ann N Y Acad Sci* **2016,** *1372* (1), 9-19.

24. Zhang, Y.; Zhao, F.; Deng, Y.; Zhao, Y.; Ren, H., Metagenomic and metabolomic analysis of the toxic effects of trichloroacetamide-induced gut microbiome and urine metabolome perturbations in mice. *J Proteome Res* **2015,** *14* (4), 1752-61.

25. Maier, T. V.; Lucio, M.; Lee, L. H.; VerBerkmoes, N. C.; Brislawn, C. J.; Bernhardt, J.; Lamendella, R.; McDermott, J. E.; Bergeron, N.; Heinzmann, S. S.; Morton, J. T.; Gonzalez, A.; Ackermann, G.; Knight, R.; Riedel, K.; Krauss, R. M.; Schmitt-Kopplin, P.; Jansson, J. K., Impact of Dietary Resistant Starch on the Human Gut Microbiome, Metaproteome, and Metabolome. *MBio* **2017,** *8* (5).

26. Deatherage Kaiser, B. L.; Li, J.; Sanford, J. A.; Kim, Y. M.; Kronewitter, S. R.; Jones, M. B.; Peterson, C. T.; Peterson, S. N.; Frank, B. C.; Purvine, S. O.; Brown, J. N.; Metz, T. O.; Smith, R. D.; Heffron, F.; Adkins, J. N., A Multi-Omic View of Host-Pathogen-Commensal Interplay in Salmonella-Mediated Intestinal Infection. *PLoS One* **2013,** *8* (6), e67155.

27. Bratburd, J. R.; Keller, C.; Vivas, E.; Gemperline, E.; Li, L.; Rey, F. E.; Currie, C. R., Gut Microbial and Metabolic Responses to Salmonella enterica Serovar Typhimurium and Candida albicans. *MBio* **2018,** *9* (6).

28. Xia, J.; Sinelnikov, I. V.; Han, B.; Wishart, D. S., MetaboAnalyst 3.0--making metabolomics more meaningful. *Nucleic Acids Res* **2015,** *43* (W1), W251-7.

29. Xia, J.; Wishart, D. S., Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. *Curr Protoc Bioinformatics* **2016,** *55*, 14.10.1-14.10.91.

30. Frost, D. C.; Greer, T.; Li, L., High-resolution enabled 12-plex DiLeu isobaric tags for quantitative proteomics. *Anal Chem* **2015,** *87* (3), 1646-54.

31. Yang, F.; Shen, Y.; Camp, D. G., 2nd; Smith, R. D., High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis. *Expert Rev Proteomics* **2012,** *9* (2), 129-34.

32. Mansour, H. H., Protective role of carnitine ester against radiation-induced oxidative stress in rats. *Pharmacol Res* **2006,** *54* (3), 165-71.

33. Di Cesare Mannelli, L.; Ghelardini, C.; Toscano, A.; Pacini, A.; Bartolini, A., The neuropathy-protective agent acetyl-L-carnitine activates protein kinase C-gamma and MAPKs in a rat model of neuropathic pain. *Neuroscience* **2010,** *165* (4), 1345-52.

34. Imperato, A.; Ramacci, M. T.; Angelucci, L., Acetyl-L-carnitine enhances acetylcholine release in the striatum and hippocampus of awake freely moving rats. *Neurosci Lett* **1989,** *107* (1-3), 251-5.

35. Veronese, N.; Stubbs, B.; Solmi, M.; Ajnakina, O.; Carvalho, A. F.; Maggi, S., Acetyl-L-Carnitine Supplementation and the Treatment of Depressive Symptoms: A Systematic Review and Meta-Analysis. *Psychosom Med* **2018,** *80* (2), 154-159.

36. Pettegrew, J. W.; Levine, J.; McClure, R. J., Acetyl-L-carnitine physical-chemical, metabolic, and therapeutic properties: relevance for its mode of action in Alzheimer's disease and geriatric depression. *Mol Psychiatry* **2000,** *5* (6), 616-32.

37. Rinaldo, P.; Matern, D.; Bennett, M. J., Fatty acid oxidation disorders. *Annu Rev Physiol* **2002,** *64*, 477-502.

38. McCoin, C. S.; Knotts, T. A.; Adams, S. H., Acylcarnitines--old actors auditioning for new roles in metabolic physiology. *Nat Rev Endocrinol* **2015,** *11* (10), 617-25.

39. Adams, S. H.; Hoppel, C. L.; Lok, K. H.; Zhao, L.; Wong, S. W.; Minkler, P. E.; Hwang, D. H.; Newman, J. W.; Garvey, W. T., Plasma acylcarnitine profiles suggest incomplete long-chain fatty acid beta-oxidation and altered tricarboxylic acid cycle activity in type 2 diabetic African-American women. *J Nutr* **2009,** *139* (6), 1073-81.

40. Jennewein, C.; Tran, N.; Paulus, P.; Ellinghaus, P.; Eble, J. A.; Zacharowski, K., Novel aspects of fibrin(ogen) fragments during inflammation. *Mol Med* **2011,** *17* (5-6), 568-73.

41. Ichikawa, Y.; Yamada, C.; Horiki, T.; Hoshina, Y.; Uchiyama, M., Serum matrix metalloproteinase-3 and fibrin degradation product levels correlate with clinical disease activity in rheumatoid arthritis. *Clin Exp Rheumatol* **1998,** *16* (5), 533-40.

42. Iba, T.; Kidokoro, A.; Fukunaga, M.; Sugiyama, K.; Sawada, T.; Kato, H., Association between the severity of sepsis and the changes in hemostatic molecular markers and vascular endothelial damage markers. *Shock* **2005,** *23* (1), 25-9.

43. Inserra, I.; Iavarone, F.; Martelli, C.; D'Angelo, L.; Delfino, D.; Rossetti, D. V.; Tamburrini, G.; Massimi, L.; Caldarelli, M.; Di Rocco, C.; Messana, I.; Castagnola, M.; Desiderio, C., Proteomic study of pilocytic astrocytoma pediatric brain tumor intracystic fluid. *J Proteome Res* **2014,** *13* (11), 4594-606.

44. Romani, L.; Moretti, S.; Fallarino, F.; Bozza, S.; Ruggeri, L.; Casagrande, A.; Aversa, F.; Bistoni, F.; Velardi, A.; Garaci, E., Jack of all trades: thymosin alpha1 and its pleiotropy. *Ann N Y Acad Sci* **2012,** *1269*, 1-6.

45. King, R.; Tuthill, C., Immune Modulation with Thymosin Alpha 1 Treatment. *Vitam Horm* **2016,** *102*, 151-78.

46. Hochepied, T.; Berger, F. G.; Baumann, H.; Libert, C., Alpha(1)-acid glycoprotein: an acute phase protein with inflammatory and immunomodulating properties. *Cytokine Growth Factor Rev* **2003,** *14* (1), 25-34.

47. Pukhal'skii, A. L.; Shmarina, G. V.; Lyutov, A. G.; Novikova, L. I.; Shemyakin, I. G.; Aleshkin, V. A., alpha1-acid glycoprotein possesses in vitro pro- and antiinflammatory activities. *Bull Exp Biol Med* **2001,** *131* (5), 479-81.

**Tables**

**Table 1.** Putative identifications from metabolomics analysis. Compounds from previous studies are shown in italics.

| Molecular Weight | RT (min) | Compound Name | Infection | Literature Molecular Weight | ppm |
|---|---|---|---|---|---|
| 203.11562 | 1.219 | Acetyl-L-Carnitine | Salmonella | 203.11575 | -0.65 |
| 231.14701 | 6.577 | Butyryl-L-carnitine, Isobutyryl carnitine | Salmonella | 231.14706 | -0.23 |
| 245.16276 | 12.672 | 2-Methylbutyroylcarnitine, pivaloylcarnitine | Salmonella | 245.16270 | 0.23 |
| 247.14170 | 1.497/1.724 | *hydroxybutyrylcarnitine* | Salmonella | 247.14197 | -1.08 |
| 259.17830 | 14.895 | Hexanoylcarnitine | Salmonella | 259.17835 | -0.17 |
| 267.09656 | 2.413 | Adenosine | Salmonella &Candida | 267.09674 | -0.68 |
| 347.06261 | 1.452 | Adenosine 3’ monophosphate | Salmonella &Candida | 347.06308 | -1.35 |
| 347.06263 | 1.160 | *similar to Adenosine monophosphate* | Salmonella | 347.06308 | -1.30 |
| 348.04670 | 1.188 | *Inosine monophosphate* | Salmonella | 348.04709 | -1.12 |
| 426.08792 | 0.935 | *cysteine glutathione disulfide* | Salmonella | 426.08789 | 0.07 |
| 612.15125 | 1.881 | *L-Glutathione oxidized* | Salmonella | 612.15198 | -1.19 |

## Figures

### Figure 1.

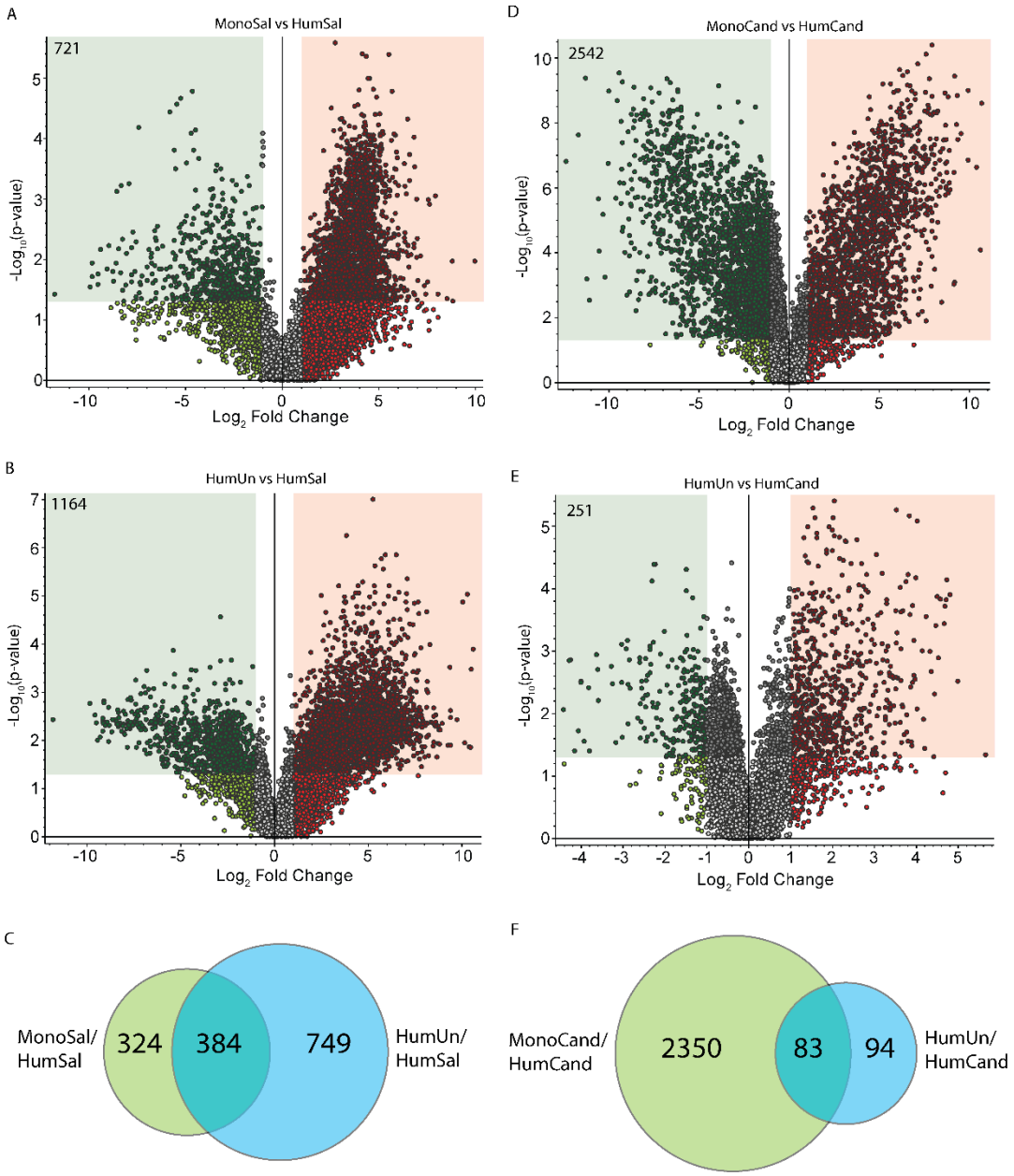



**Figure 1.** Target selection for metabolomics results. Volcano plots for MonoSal/HumSal (A), HumUn/HumSal (B), MonoCand (D), and HumUn/HumCand (E) selected statistically significant targets in each ratio. The overlap of the targets for the HumSal group (C) and HumCand group (F) were used to select final targets.

**Figure 2**

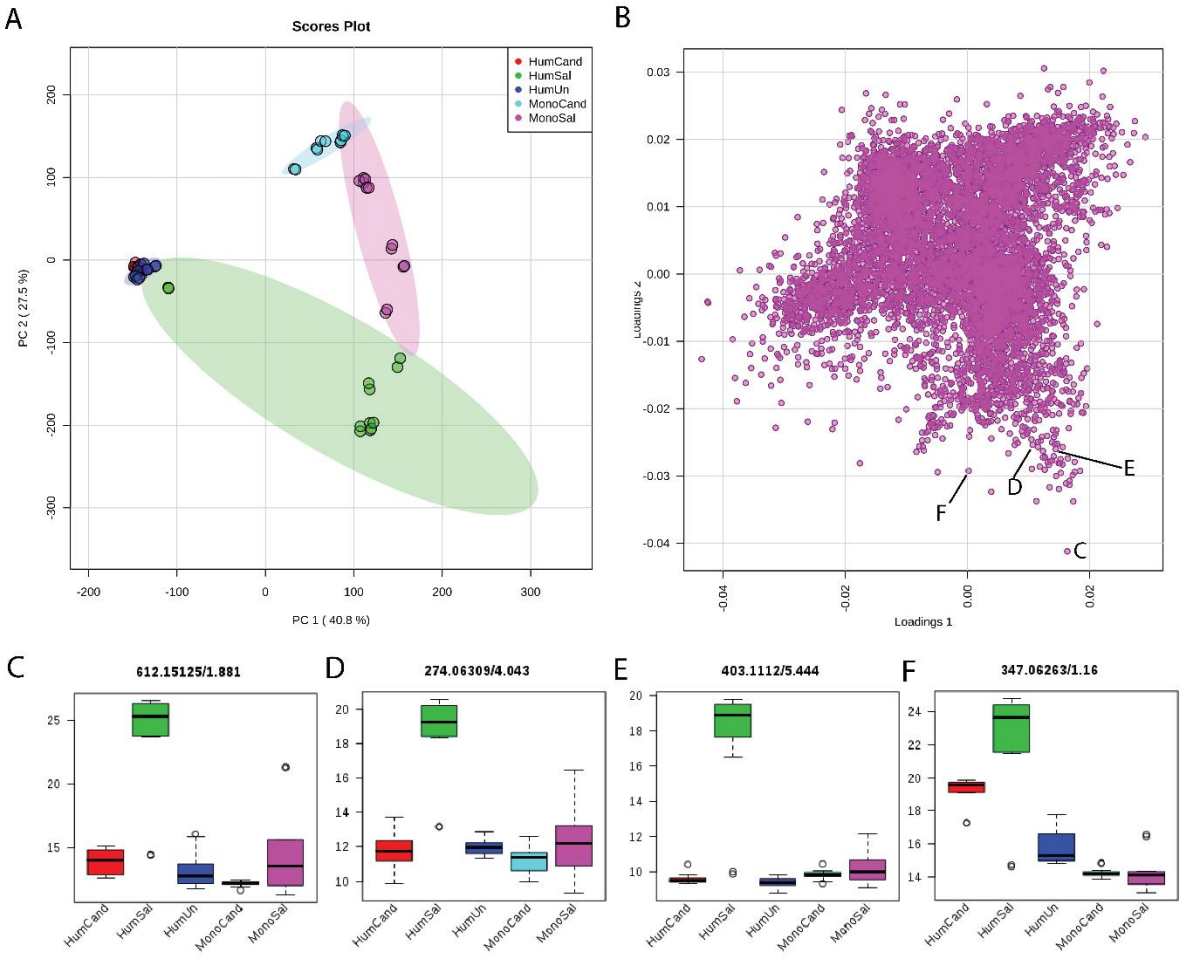
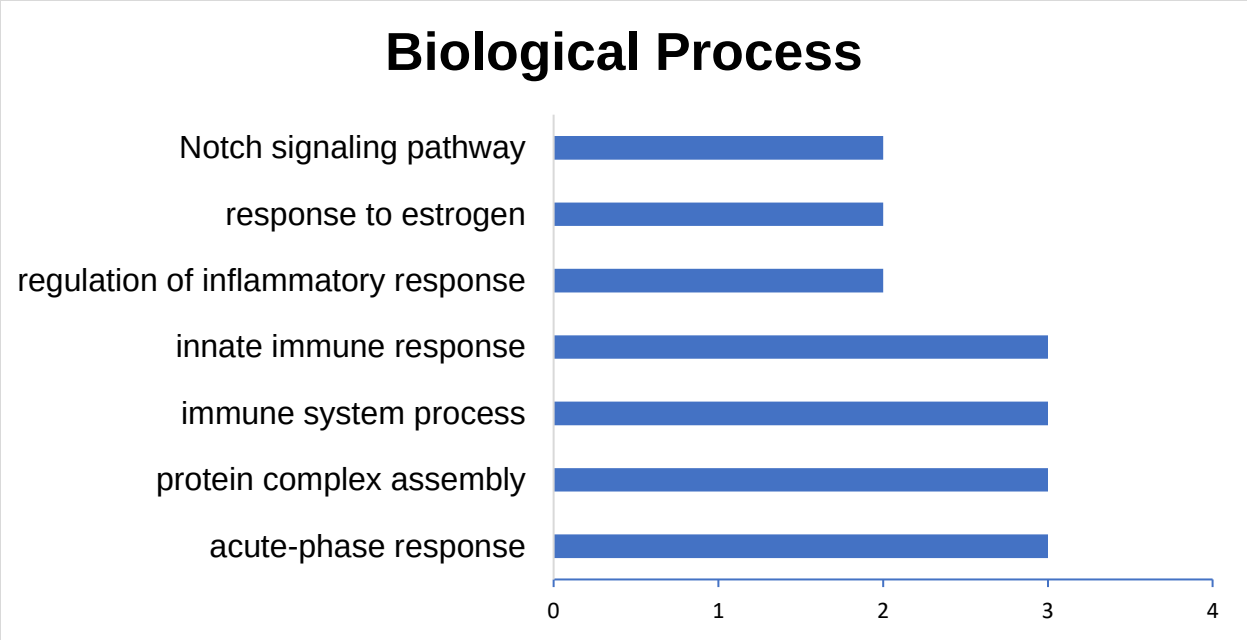



**Figure 2.** PCA results for the metabolomics data. **(A)** shows the PCA plot for separation of sample groups, while **(B)** shows the loadings plot for the various detected compounds. **(C-F)** show various compounds on the edge of the loadings plot as being up-regulated in the HumSal (and HumCand for **(F)**).



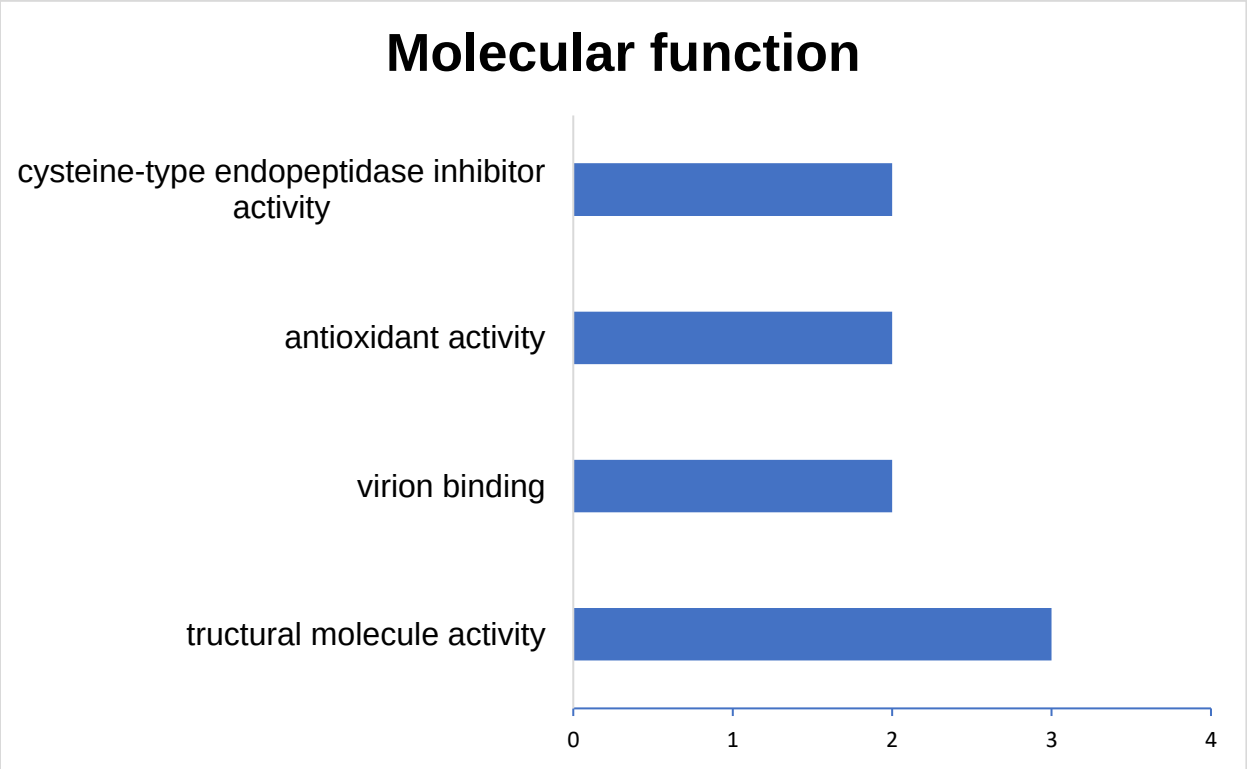




**Figure 3.** Gene ontology results for proteins up-regulated in the HumSal group. The biological process is shown on top and the molecular function is shown on the bottom.

## Supplemental Tables

**Supplemental Table S1.** 12-Plex DiLeu channels Labeling Map.

| **Sample** | HS1 | HS2 | HS3 | HS4 | HS5 | HS6 | GS1 | GS2 | GS3 | GS4 | GS5 | GS6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Label** | 117b | 117c | 118a | 118b | 118c | 118d | 115a | 115b | 116a | 116b | 116c | 117a |
| | | | | | | | | | | | | |
| **Sample** | | | | | | | NO1 | NO2 | NO3 | NO4 | NO5 | NO6 |
| **Label** | | | | | | | 115a | 115b | 116a | 116b | 116c | 117a |
| | | | | | | | | | | | | |
| **Sample** | HC1 | HC2 | HC3 | HC4 | HC5 | HC6 | GC1 | GC2 | GC3 | GC4 | GC5 | GC6 |
| **Label** | 117b | 117c | 118a | 118b | 118c | 118d | 115a | 115b | 116a | 116b | 116c | 117a |
| | | | | | | | | | | | | |
| **4 mixtures** | HS& GS | | HS& NO | | HC& GC | | HC& GC | | | | | |

**Supplemental Table S2.** Target Compounds produced in response to *Salmonella* infection.

| Molecular Weight | RT [min] | Log2 Fold Change: (HumUn) / (HumSal) | Log2 Fold Change: (MonoSal) / (HumSal) | P-value: (HumUn) / (HumSal) | P-value: (MonoSal) / (HumSal) |
|---|---|---|---|---|---|
| 203.11562 | 1.219 | -4.41 | -4.45 | 3.51E-04 | 7.27E-05 |
| 205.04819 | 0.854 | -5.65 | -6.98 | 3.02E-02 | 1.33E-02 |
| 206.01046 | 28.177 | -2.49 | -2.36 | 7.96E-03 | 4.71E-02 |
| 206.09148 | 18.084 | -2.64 | -3.45 | 1.92E-02 | 1.10E-02 |
| 211.13601 | 22.812 | -2.54 | -2.45 | 6.89E-03 | 4.28E-02 |
| 212.04077 | 0.883 | -6.12 | -6.97 | 9.05E-03 | 2.23E-02 |
| 216.09972 | 14.883 | -4.39 | -4.01 | 2.77E-03 | 2.89E-02 |
| 217.03840 | 0.961 | -4.80 | -3.23 | 3.46E-03 | 1.16E-02 |
| 222.08928 | 19.206 | -1.90 | -3.54 | 2.85E-02 | 1.29E-02 |
| 224.10196 | 15.413 | -2.38 | -4.03 | 3.75E-02 | 1.32E-02 |
| 225.00491 | 28.173 | -2.63 | -2.37 | 6.54E-03 | 4.78E-02 |
| 226.19290 | 22.922 | -2.63 | -2.97 | 1.73E-02 | 1.88E-02 |
| 227.00028 | 28.172 | -2.55 | -2.57 | 7.82E-03 | 4.46E-02 |
| 231.14701 | 6.577 | -1.19 | -4.85 | 4.38E-02 | 4.14E-03 |
| 231.88960 | 0.845 | -3.41 | -2.73 | 1.77E-03 | 3.75E-02 |
| 234.16187 | 22.151 | -2.43 | -2.32 | 5.84E-03 | 4.28E-02 |
| 236.96618 | 0.858 | -2.20 | -4.23 | 4.14E-02 | 3.75E-02 |
| 239.02031 | 0.954 | -4.42 | -3.86 | 4.24E-03 | 1.01E-02 |
| 242.11247 | 12.709 | -2.27 | -2.73 | 1.38E-02 | 3.74E-02 |
| 243.21976 | 21.301 | -2.75 | -2.31 | 6.89E-03 | 4.96E-02 |
| 244.12815 | 10.160 | -2.99 | -3.29 | 1.27E-02 | 2.76E-02 |
| 245.16276 | 12.672 | -2.19 | -1.99 | 2.61E-03 | 2.17E-02 |
| 247.14170 | 1.724 | -7.40 | -8.32 | 1.79E-03 | 6.14E-04 |
| 247.14175 | 1.497 | -3.66 | -5.57 | 4.41E-03 | 1.58E-04 |
| 247.86382 | 0.849 | -2.07 | -2.59 | 7.98E-03 | 4.21E-03 |
| 248.97206 | 8.038 | -3.98 | -3.86 | 1.86E-02 | 4.99E-02 |
| 254.15157 | 19.808 | -1.33 | -2.40 | 4.39E-02 | 2.84E-02 |
| 259.17830 | 14.895 | -1.99 | -3.95 | 1.83E-02 | 5.21E-03 |
| 261.03032 | 1.867 | -3.47 | -3.68 | 1.32E-02 | 4.39E-02 |
| 261.15741 | 14.883 | -5.75 | -4.15 | 3.15E-03 | 4.74E-02 |
| 262.01367 | 1.723 | -6.25 | -6.74 | 8.08E-03 | 8.13E-03 |
| 265.06713 | 1.155 | -2.82 | -3.84 | 8.24E-03 | 3.71E-02 |
| 266.16454 | 22.052 | -2.39 | -2.15 | 3.11E-03 | 3.68E-02 |
| 267.09656 | 2.413 | -3.65 | -3.55 | 5.18E-03 | 3.47E-02 |
| 268.12810 | 16.669 | -1.95 | -3.18 | 1.98E-02 | 2.10E-02 |
| 270.13037 | 0.921 | -2.65 | -3.29 | 8.40E-03 | 1.90E-02 |
| 270.14366 | 15.454 | -2.95 | -4.42 | 1.67E-02 | 1.09E-02 |
| 270.14386 | 14.877 | -2.37 | -4.31 | 3.27E-02 | 1.05E-02 |
| 270.90932 | 28.178 | -2.46 | -2.55 | 6.51E-03 | 2.47E-02 |
| 271.21468 | 22.026 | -2.72 | -2.61 | 1.44E-02 | 4.11E-02 |
| 271.25093 | 22.512 | -2.59 | -2.32 | 6.89E-03 | 3.95E-02 |
| 272.90979 | 28.176 | -2.32 | -2.36 | 6.68E-03 | 4.41E-02 |
| 274.06309 | 4.043 | -7.12 | -8.07 | 2.37E-03 | 2.65E-02 |
| 274.12737 | 1.373 | -2.65 | -3.23 | 1.30E-02 | 1.33E-03 |
| 274.13860 | 12.705 | -2.38 | -2.93 | 1.33E-02 | 3.17E-02 |
| 274.25072 | 23.413 | -2.17 | -2.27 | 9.99E-03 | 3.61E-02 |
| 278.15163 | 22.775 | -2.53 | -2.40 | 3.59E-03 | 1.50E-02 |
| 280.13075 | 21.019 | -2.00 | -2.56 | 1.18E-02 | 3.31E-02 |
| 280.24010 | 22.494 | -3.17 | -2.94 | 6.56E-04 | 1.66E-03 |
| 282.14367 | 18.510 | -2.32 | -3.95 | 1.73E-02 | 1.38E-02 |
| 282.14370 | 18.087 | -2.38 | -4.11 | 2.08E-02 | 1.24E-02 |
| 282.14378 | 16.668 | -1.82 | -3.34 | 1.99E-02 | 2.13E-02 |
| 283.32368 | 22.788 | -3.12 | -2.27 | 3.82E-02 | 3.13E-02 |
| 283.32368 | 22.598 | -2.07 | -2.16 | 7.91E-03 | 4.80E-02 |
| 285.11431 | 1.108 | -6.84 | -6.50 | 3.73E-03 | 4.88E-03 |
| 286.13869 | 15.420 | -2.25 | -3.41 | 3.00E-02 | 6.54E-03 |
| 286.88142 | 28.291 | -2.58 | -2.31 | 1.59E-02 | 4.56E-02 |
| 287.24576 | 21.124 | -2.90 | -3.21 | 1.24E-02 | 4.96E-02 |
| 288.88233 | 28.178 | -2.56 | -2.43 | 4.43E-03 | 4.08E-02 |
| 289.12702 | 1.420 | -2.05 | -2.46 | 3.53E-03 | 2.36E-02 |
| 289.99846 | 8.040 | -2.94 | -2.95 | 2.75E-02 | 3.89E-02 |
| 290.88296 | 28.183 | -2.49 | -2.44 | 1.72E-03 | 2.95E-02 |
| 292.08625 | 21.483 | -1.88 | -2.62 | 2.26E-02 | 2.55E-02 |
| 294.16173 | 22.781 | -2.42 | -2.23 | 5.66E-03 | 4.75E-02 |
| 294.16175 | 22.642 | -2.36 | -2.33 | 6.84E-03 | 4.40E-02 |
| 298.15676 | 21.409 | -2.34 | -2.59 | 1.45E-02 | 4.99E-02 |
| 299.28219 | 23.468 | -2.32 | -2.25 | 9.30E-03 | 4.02E-02 |
| 299.87702 | 0.846 | -3.18 | -3.10 | 2.68E-03 | 2.81E-02 |
| 299.89790 | 0.826 | -2.90 | -5.96 | 4.33E-02 | 1.46E-02 |
| 300.15435 | 15.915 | -2.01 | -3.59 | 2.69E-02 | 1.15E-02 |
| 300.15454 | 15.411 | -2.28 | -3.89 | 3.70E-02 | 1.17E-02 |
| 302.17006 | 14.878 | -2.86 | -4.78 | 2.50E-02 | 9.04E-03 |
| 303.88420 | 28.332 | -2.66 | -1.89 | 7.42E-03 | 3.45E-02 |
| 303.91471 | 28.183 | -2.22 | -2.35 | 1.12E-02 | 3.97E-02 |
| 306.09237 | 0.933 | -5.08 | -4.87 | 4.70E-02 | 1.86E-02 |
| 306.16166 | 22.932 | -2.30 | -2.06 | 9.73E-03 | 3.22E-02 |
| 307.09615 | 1.114 | -9.49 | -8.75 | 4.40E-03 | 1.59E-02 |
| 309.09181 | 1.108 | -5.18 | -4.09 | 5.51E-03 | 5.43E-03 |

| | | | | | |
|---|---|---|---|---|---|
| 311.22231 | 22.052 | -2.46 | -1.77 | 2.48E-03 | 3.90E-02 |
| 312.13588 | 22.614 | -2.40 | -2.35 | 9.73E-03 | 3.74E-02 |
| 314.17009 | 16.669 | -1.86 | -3.28 | 2.09E-02 | 2.21E-02 |
| 314.92232 | 28.180 | -2.46 | -2.34 | 8.17E-03 | 4.76E-02 |
| 316.26106 | 20.551 | -3.03 | -3.17 | 1.79E-02 | 4.85E-02 |
| 317.25637 | 22.950 | -2.80 | -2.57 | 6.23E-03 | 4.17E-02 |
| 317.27177 | 20.480 | -2.38 | -2.64 | 2.35E-02 | 3.95E-02 |
| 317.87950 | 28.218 | -2.64 | -2.58 | 6.62E-03 | 4.67E-02 |
| 318.16479 | 15.417 | -2.31 | -3.86 | 3.63E-02 | 1.37E-02 |
| 318.16494 | 14.442 | -2.48 | -3.47 | 1.74E-02 | 2.30E-02 |
| 318.27670 | 23.408 | -2.29 | -2.32 | 9.47E-03 | 3.39E-02 |
| 321.28769 | 22.952 | -2.21 | -1.90 | 8.66E-03 | 4.10E-02 |
| 322.10709 | 1.191 | -5.65 | -5.09 | 1.32E-02 | 1.31E-02 |
| 323.05147 | 1.189 | -5.55 | -6.15 | 5.32E-03 | 2.98E-02 |
| 323.05156 | 0.963 | -4.88 | -5.04 | 1.32E-02 | 4.04E-02 |
| 323.07018 | 1.109 | -7.32 | -6.43 | 4.81E-03 | 6.84E-03 |
| 323.20939 | 22.718 | -2.95 | -2.89 | 1.02E-02 | 3.35E-02 |
| 326.29308 | 23.314 | -2.57 | -2.27 | 7.07E-03 | 3.01E-02 |
| 326.91170 | 28.182 | -2.40 | -2.37 | 1.21E-02 | 4.39E-02 |
| 327.90807 | 28.314 | -2.46 | -2.31 | 2.29E-02 | 4.56E-02 |
| 329.90863 | 28.181 | -2.59 | -2.39 | 5.54E-03 | 4.22E-02 |
| 330.16487 | 16.668 | -1.78 | -2.92 | 2.66E-02 | 2.08E-02 |
| 331.06787 | 1.423 | -5.34 | -4.10 | 5.98E-03 | 2.11E-02 |
| 331.14169 | 21.402 | -1.80 | -2.70 | 3.17E-02 | 3.55E-02 |
| 331.90955 | 28.179 | -2.59 | -2.48 | 9.13E-04 | 2.37E-02 |
| 332.18065 | 15.413 | -2.70 | -4.19 | 2.25E-02 | 1.31E-02 |
| 333.90459 | 28.181 | -2.51 | -2.43 | 7.33E-03 | 4.35E-02 |
| 334.11785 | 22.607 | -2.44 | -2.41 | 1.09E-02 | 3.70E-02 |
| 334.15423 | 21.345 | -3.24 | -2.46 | 4.04E-03 | 3.76E-02 |
| 335.30321 | 23.408 | -2.27 | -2.30 | 9.71E-03 | 3.56E-02 |
| 337.14409 | 21.483 | -2.08 | -2.32 | 1.30E-02 | 3.32E-02 |
| 337.24626 | 20.458 | -1.79 | -2.89 | 1.79E-02 | 2.54E-02 |
| 337.90054 | 28.192 | -2.76 | -2.56 | 7.69E-03 | 4.50E-02 |
| 337.90312 | 28.179 | -2.52 | -2.41 | 7.20E-03 | 4.96E-02 |
| 340.30875 | 23.344 | -2.39 | -2.18 | 8.17E-03 | 4.30E-02 |
| 340.91667 | 28.186 | -2.57 | -2.39 | 6.13E-03 | 4.94E-02 |
| 343.30851 | 21.889 | -2.19 | -2.88 | 3.38E-02 | 4.54E-02 |
| 344.07956 | 18.793 | -2.24 | -2.48 | 9.78E-03 | 3.71E-02 |
| 344.18057 | 18.087 | -2.14 | -4.01 | 2.12E-02 | 1.17E-02 |
| 344.18069 | 16.669 | -1.72 | -3.26 | 2.66E-02 | 1.99E-02 |
| 344.18072 | 17.176 | -1.74 | -2.80 | 2.69E-02 | 2.65E-02 |
| 345.03346 | 0.968 | -4.93 | -4.41 | 4.41E-03 | 3.93E-02 |
| 346.11426 | 0.860 | -2.34 | -2.89 | 5.43E-03 | 3.44E-02 |
| 347.06261 | 1.452 | -8.02 | -9.37 | 7.12E-03 | 6.86E-03 |
| 347.06263 | 1.160 | -7.34 | -9.80 | 1.39E-02 | 1.16E-02 |
| 348.04670 | 1.188 | -7.69 | -8.04 | 5.67E-03 | 8.01E-03 |
| 348.16980 | 22.705 | -4.32 | -3.16 | 1.12E-03 | 4.66E-02 |
| 348.25100 | 21.770 | -2.24 | -2.87 | 2.59E-02 | 4.83E-02 |
| 349.28249 | 22.941 | -1.94 | -2.41 | 1.24E-02 | 6.70E-03 |
| 349.31893 | 22.955 | -3.34 | -2.90 | 8.58E-03 | 4.30E-02 |
| 349.31894 | 23.403 | -2.69 | -2.20 | 4.48E-03 | 2.96E-02 |
| 349.31894 | 23.493 | -2.36 | -2.03 | 7.29E-03 | 3.67E-02 |
| 350.08862 | 1.720 | -4.98 | -7.52 | 3.95E-02 | 5.19E-03 |
| 352.11324 | 23.513 | -2.33 | -2.18 | 1.53E-02 | 4.62E-02 |
| 354.23803 | 22.941 | -1.79 | -2.21 | 2.22E-02 | 7.77E-03 |
| 355.21445 | 22.932 | -2.28 | -2.31 | 1.62E-02 | 2.95E-02 |
| 357.23025 | 21.351 | -2.61 | -2.17 | 4.67E-03 | 3.29E-02 |
| 358.19620 | 18.514 | -2.19 | -3.70 | 1.65E-02 | 1.45E-02 |
| 358.19631 | 18.086 | -2.29 | -3.97 | 2.21E-02 | 1.16E-02 |
| 359.17285 | 21.402 | -2.44 | -1.81 | 3.57E-03 | 3.63E-02 |
| 359.33971 | 22.625 | -2.76 | -2.77 | 7.67E-03 | 2.52E-02 |
| 360.96415 | 8.035 | -4.28 | -4.36 | 2.26E-02 | 4.50E-02 |
| 361.31883 | 20.537 | -2.99 | -3.53 | 1.63E-02 | 3.31E-02 |
| 362.19111 | 15.916 | -2.54 | -3.97 | 2.00E-02 | 1.82E-02 |
| 362.19120 | 16.668 | -1.74 | -3.24 | 2.66E-02 | 2.06E-02 |
| 362.21251 | 22.586 | -2.22 | -2.07 | 9.82E-03 | 4.15E-02 |
| 362.30296 | 23.400 | -2.31 | -2.25 | 8.41E-03 | 3.31E-02 |
| 363.05745 | 1.201 | -8.82 | -9.04 | 4.89E-03 | 1.43E-02 |
| 363.05749 | 1.539 | -6.91 | -8.56 | 1.62E-02 | 1.18E-02 |
| 363.33444 | 23.281 | -2.58 | -2.22 | 3.82E-03 | 3.09E-02 |
| 363.33446 | 22.962 | -3.72 | -3.61 | 4.92E-03 | 3.75E-02 |
| 364.04170 | 1.258 | -5.77 | -6.66 | 2.49E-02 | 2.48E-02 |
| 365.05323 | 1.715 | -8.13 | -8.45 | 6.58E-03 | 6.80E-03 |
| 365.27753 | 21.772 | -2.25 | -2.88 | 2.68E-02 | 4.93E-02 |
| 365.31391 | 22.948 | -2.24 | -2.06 | 8.29E-03 | 3.86E-02 |
| 367.86476 | 0.845 | -3.03 | -3.41 | 3.65E-03 | 2.66E-02 |
| 369.28763 | 23.087 | -2.79 | -2.95 | 5.31E-03 | 3.99E-02 |
| 371.24574 | 22.735 | -3.85 | -4.07 | 1.10E-02 | 3.25E-02 |
| 371.24591 | 22.643 | -2.52 | -2.23 | 7.98E-03 | 2.74E-02 |
| 371.35092 | 23.315 | -2.47 | -2.30 | 8.72E-03 | 4.67E-02 |
| 373.31893 | 20.534 | -3.00 | -3.50 | 2.11E-02 | 4.77E-02 |
| 373.31913 | 21.315 | -2.33 | -3.12 | 3.56E-02 | 3.96E-02 |
| 376.20663 | 18.509 | -2.21 | -3.83 | 1.38E-02 | 1.41E-02 |
| 376.20672 | 18.087 | -2.29 | -3.98 | 2.19E-02 | 1.16E-02 |

| | | | | | |
|---|---|---|---|---|---|
| 376.20683 | 16.668 | -1.97 | -3.38 | 2.02E-02 | 2.35E-02 |
| 376.22160 | 23.854 | -2.40 | -2.14 | 7.25E-03 | 4.70E-02 |
| 376.28219 | 22.927 | -2.41 | -2.93 | 1.23E-02 | 1.43E-02 |
| 377.31387 | 22.941 | -2.75 | -2.87 | 8.67E-03 | 1.61E-02 |
| 377.35013 | 23.484 | -2.24 | -2.63 | 3.81E-03 | 2.75E-02 |
| 379.32943 | 23.400 | -2.32 | -2.26 | 8.57E-03 | 3.33E-02 |
| 379.32945 | 22.924 | -2.29 | -2.19 | 8.06E-03 | 4.91E-02 |
| 383.24585 | 23.125 | -2.47 | -2.31 | 1.01E-02 | 4.58E-02 |
| 383.24585 | 23.035 | -2.58 | -2.16 | 1.01E-02 | 4.70E-02 |
| 383.24595 | 22.985 | -4.08 | -3.64 | 9.26E-03 | 7.40E-03 |
| 383.30340 | 22.464 | -2.75 | -2.92 | 1.21E-02 | 3.84E-02 |
| 383.83925 | 0.849 | -1.50 | -3.04 | 1.17E-02 | 2.99E-03 |
| 387.20441 | 22.489 | -2.80 | -2.94 | 3.68E-02 | 2.16E-02 |
| 387.20452 | 22.421 | -2.53 | -2.65 | 1.94E-02 | 3.75E-02 |
| 388.20699 | 18.510 | -2.20 | -4.10 | 2.05E-02 | 1.22E-02 |
| 390.22231 | 18.088 | -2.35 | -4.02 | 2.04E-02 | 1.19E-02 |
| 392.07644 | 0.831 | -2.03 | -2.75 | 4.15E-02 | 3.98E-02 |
| 393.30875 | 22.921 | -2.39 | -2.94 | 5.75E-03 | 8.85E-03 |
| 393.30880 | 21.773 | -2.22 | -2.37 | 3.72E-03 | 1.68E-02 |
| 393.34507 | 23.352 | -3.07 | -2.61 | 2.31E-02 | 4.48E-02 |
| 393.34510 | 23.616 | -2.14 | -2.68 | 3.64E-03 | 2.17E-02 |
| 393.34517 | 23.481 | -2.01 | -2.45 | 9.84E-03 | 2.39E-02 |
| 394.17787 | 14.702 | -4.10 | -3.58 | 8.35E-03 | 4.33E-02 |
| 401.25645 | 21.799 | -2.39 | -1.94 | 7.05E-03 | 4.10E-02 |
| 401.25650 | 21.893 | -2.59 | -2.60 | 8.42E-03 | 4.91E-02 |
| 401.25658 | 21.980 | -2.29 | -2.17 | 7.96E-03 | 3.64E-02 |
| 402.29158 | 22.251 | -2.98 | -2.89 | 1.48E-02 | 1.69E-02 |
| 403.11120 | 5.444 | -9.55 | -8.93 | 2.98E-03 | 5.91E-03 |
| 406.04158 | 1.260 | -5.05 | -5.19 | 4.44E-03 | 4.20E-02 |
| 406.21738 | 18.514 | -2.27 | -4.05 | 1.88E-02 | 1.36E-02 |
| 406.21761 | 17.175 | -2.05 | -2.95 | 2.10E-02 | 2.82E-02 |
| 408.17576 | 24.492 | -2.50 | -2.40 | 1.74E-02 | 4.32E-02 |
| 409.30355 | 20.512 | -3.38 | -3.73 | 1.80E-02 | 4.60E-02 |
| 409.33998 | 22.930 | -2.24 | -1.99 | 9.27E-03 | 4.41E-02 |
| 412.20433 | 18.093 | -3.00 | -4.60 | 8.00E-03 | 1.51E-02 |
| 413.27749 | 20.628 | -2.96 | -3.34 | 2.67E-02 | 4.11E-02 |
| 414.20382 | 21.166 | -2.88 | -2.78 | 2.96E-03 | 1.63E-02 |
| 414.20384 | 20.657 | -4.04 | -4.66 | 5.45E-03 | 1.82E-02 |
| 414.20384 | 21.374 | -2.59 | -3.99 | 1.10E-02 | 2.67E-02 |
| 414.20389 | 21.020 | -2.10 | -2.53 | 1.36E-02 | 4.02E-02 |
| 416.30700 | 22.710 | -5.22 | -4.39 | 1.88E-02 | 2.87E-02 |
| 416.30713 | 22.752 | -3.52 | -3.53 | 2.12E-02 | 2.72E-02 |
| 418.23265 | 20.628 | -3.31 | -3.10 | 1.82E-02 | 4.79E-02 |
| 420.23298 | 18.514 | -2.18 | -3.85 | 1.83E-02 | 1.28E-02 |
| 420.30831 | 22.916 | -2.72 | -3.04 | 1.72E-02 | 2.50E-02 |
| 421.27916 | 23.859 | -2.38 | -2.29 | 4.28E-03 | 4.62E-02 |
| 421.33996 | 22.927 | -2.81 | -2.91 | 1.39E-02 | 3.29E-02 |
| 421.37627 | 23.486 | -1.76 | -2.78 | 3.10E-02 | 3.36E-02 |
| 421.37631 | 23.304 | -1.87 | -2.25 | 2.92E-03 | 4.96E-02 |
| 421.37651 | 23.244 | -3.04 | -3.72 | 1.88E-02 | 1.65E-03 |
| 423.35558 | 23.388 | -2.33 | -2.32 | 9.08E-03 | 3.14E-02 |
| 425.84380 | 0.836 | -4.60 | -6.55 | 2.10E-02 | 2.24E-02 |
| 426.08792 | 0.935 | -6.39 | -5.64 | 4.47E-03 | 4.40E-02 |
| 427.32967 | 22.461 | -3.02 | -3.44 | 1.42E-02 | 3.47E-02 |
| 431.23034 | 21.015 | -1.99 | -2.44 | 1.47E-02 | 3.98E-02 |
| 431.23046 | 21.377 | -2.54 | -3.98 | 1.98E-02 | 3.28E-02 |
| 431.23047 | 20.657 | -3.95 | -4.66 | 6.35E-03 | 1.74E-02 |
| 431.24354 | 20.891 | -3.56 | -3.87 | 7.93E-03 | 3.30E-02 |
| 433.16213 | 22.712 | -2.67 | -2.68 | 1.36E-02 | 3.70E-02 |
| 434.13253 | 21.400 | -3.57 | -2.84 | 1.53E-02 | 3.57E-02 |
| 434.13255 | 21.597 | -3.46 | -3.15 | 1.54E-02 | 2.49E-02 |
| 435.85203 | 0.845 | -2.81 | -3.60 | 9.06E-03 | 2.28E-02 |
| 436.18550 | 21.132 | -3.40 | -4.17 | 5.05E-03 | 3.16E-02 |
| 436.18557 | 21.025 | -2.48 | -2.61 | 1.02E-02 | 4.03E-02 |
| 436.18560 | 21.347 | -2.80 | -3.83 | 2.87E-03 | 1.66E-02 |
| 437.33500 | 22.916 | -2.72 | -3.04 | 1.72E-02 | 2.46E-02 |
| 437.33502 | 21.780 | -2.65 | -2.82 | 2.66E-03 | 7.87E-03 |
| 437.37119 | 23.728 | -2.67 | -2.49 | 9.10E-03 | 4.92E-02 |
| 437.37120 | 23.344 | -3.31 | -3.24 | 1.89E-02 | 2.72E-02 |
| 437.37126 | 23.468 | -2.34 | -2.77 | 1.28E-02 | 3.59E-02 |
| 438.37474 | 23.481 | -2.61 | -3.08 | 9.36E-03 | 2.17E-02 |
| 439.40232 | 22.669 | -2.36 | -2.28 | 7.42E-03 | 4.05E-02 |
| 440.10311 | 1.229 | -6.36 | -6.50 | 3.60E-03 | 3.57E-02 |
| 442.29020 | 22.916 | -2.67 | -3.01 | 1.82E-02 | 2.70E-02 |
| 443.07082 | 1.096 | -5.29 | -4.72 | 5.46E-03 | 9.76E-03 |
| 443.30108 | 22.491 | -2.51 | -2.30 | 1.22E-02 | 4.30E-02 |
| 443.32451 | 22.454 | -2.32 | -1.98 | 4.17E-03 | 4.36E-02 |
| 444.28037 | 16.102 | -3.05 | -3.06 | 1.05E-02 | 3.22E-02 |
| 447.28302 | 23.282 | -2.32 | -2.41 | 7.40E-03 | 4.29E-02 |
| 448.14852 | 22.267 | -2.88 | -2.59 | 1.65E-02 | 3.26E-02 |
| 450.24367 | 18.774 | -2.11 | -3.87 | 1.78E-02 | 1.46E-02 |
| 450.24379 | 18.911 | -2.25 | -3.81 | 1.48E-02 | 1.72E-02 |
| 452.15985 | 21.037 | -2.21 | -2.26 | 9.11E-03 | 4.28E-02 |
| 453.36608 | 22.907 | -2.33 | -2.32 | 6.07E-03 | 2.96E-02 |

| | | | | | |
|---|---|---|---|---|---|
| 454.18120 | 14.876 | -4.61 | -4.91 | 7.53E-03 | 3.05E-02 |
| 457.34002 | 23.072 | -2.96 | -2.38 | 8.82E-03 | 4.51E-02 |
| 459.26158 | 20.656 | -4.17 | -3.74 | 3.51E-03 | 2.30E-02 |
| 459.26162 | 21.759 | -2.44 | -3.67 | 4.11E-03 | 3.58E-03 |
| 459.26162 | 21.373 | -3.54 | -3.36 | 2.19E-03 | 2.40E-02 |
| 460.26482 | 21.375 | -3.72 | -3.25 | 2.48E-03 | 2.52E-02 |
| 460.26500 | 20.656 | -4.29 | -3.81 | 3.59E-03 | 2.32E-02 |
| 461.24148 | 21.020 | -2.11 | -2.82 | 1.03E-02 | 3.64E-02 |
| 462.12303 | 23.440 | -1.96 | -1.97 | 1.08E-02 | 3.85E-02 |
| 462.16372 | 22.707 | -2.71 | -2.85 | 1.08E-02 | 2.66E-02 |
| 464.33454 | 22.895 | -2.83 | -3.13 | 2.14E-02 | 3.44E-02 |
| 465.36619 | 22.916 | -2.73 | -3.27 | 1.59E-02 | 4.26E-02 |
| 465.40245 | 23.227 | -3.92 | -3.71 | 1.87E-02 | 3.39E-02 |
| 465.40263 | 23.413 | -2.52 | -3.43 | 3.35E-02 | 4.90E-02 |
| 466.30869 | 28.086 | -2.58 | -3.21 | 9.06E-03 | 1.30E-02 |
| 466.30882 | 28.052 | -2.27 | -2.35 | 7.99E-03 | 3.14E-02 |
| 467.38195 | 23.371 | -2.51 | -2.31 | 5.87E-03 | 4.03E-02 |
| 467.43341 | 23.300 | -2.35 | -2.30 | 7.39E-03 | 4.60E-02 |
| 472.09474 | 28.164 | -2.38 | -2.38 | 5.90E-03 | 4.32E-02 |
| 472.28957 | 16.032 | -3.27 | -2.37 | 1.64E-03 | 6.22E-03 |
| 475.25646 | 21.017 | -2.32 | -3.03 | 5.65E-03 | 3.94E-02 |
| 479.14957 | 23.441 | -1.98 | -1.96 | 1.14E-02 | 3.72E-02 |
| 479.15146 | 24.004 | -2.28 | -2.68 | 4.28E-02 | 3.80E-02 |
| 481.36114 | 22.896 | -2.84 | -3.13 | 2.15E-02 | 3.45E-02 |
| 481.36120 | 21.778 | -2.64 | -2.80 | 5.40E-03 | 1.52E-02 |
| 481.39736 | 23.339 | -2.82 | -2.63 | 1.15E-02 | 4.80E-02 |
| 481.39739 | 23.455 | -2.56 | -3.02 | 9.61E-03 | 2.08E-02 |
| 481.39740 | 23.586 | -3.01 | -3.02 | 9.37E-03 | 2.08E-02 |
| 483.10872 | 1.872 | -7.21 | -7.76 | 5.33E-03 | 2.40E-02 |
| 486.09259 | 24.167 | -2.20 | -2.68 | 2.88E-02 | 3.81E-02 |
| 486.31600 | 22.897 | -2.78 | -3.15 | 2.12E-02 | 3.34E-02 |
| 489.38476 | 28.052 | -2.56 | -2.42 | 6.37E-03 | 4.69E-02 |
| 493.83133 | 0.835 | -4.29 | -6.69 | 2.13E-02 | 1.75E-02 |
| 494.28378 | 1.494 | -3.78 | -5.51 | 4.51E-03 | 3.19E-04 |
| 501.36622 | 23.053 | -2.62 | -2.67 | 6.91E-03 | 4.70E-02 |
| 503.83960 | 0.844 | -3.03 | -3.81 | 9.42E-03 | 1.88E-02 |
| 508.36088 | 22.865 | -2.84 | -3.06 | 2.12E-02 | 3.76E-02 |
| 509.39232 | 22.896 | -2.44 | -2.60 | 3.58E-03 | 2.33E-02 |
| 509.42876 | 23.269 | -2.70 | -2.88 | 8.39E-03 | 3.79E-02 |
| 511.40776 | 22.875 | -2.52 | -2.37 | 8.33E-03 | 4.41E-02 |
| 511.40814 | 23.352 | -2.58 | -2.55 | 1.06E-02 | 4.95E-02 |
| 513.30871 | 21.020 | -1.94 | -2.45 | 1.50E-02 | 4.89E-02 |
| 513.35097 | 20.592 | -2.50 | -3.65 | 3.37E-02 | 4.69E-02 |
| 515.30667 | 14.510 | -6.33 | -6.89 | 4.37E-03 | 7.01E-03 |
| 520.31206 | 13.855 | -6.36 | -6.55 | 3.87E-03 | 3.42E-03 |
| 525.38738 | 21.770 | -2.38 | -2.52 | 2.57E-03 | 2.07E-02 |
| 525.38744 | 22.866 | -2.84 | -3.06 | 2.12E-02 | 3.75E-02 |
| 525.42358 | 23.437 | -2.57 | -2.62 | 1.41E-02 | 2.05E-02 |
| 525.42359 | 23.322 | -2.00 | -2.51 | 3.65E-02 | 3.61E-02 |
| 525.42375 | 23.567 | -2.90 | -4.83 | 1.47E-02 | 2.07E-02 |
| 529.32274 | 14.525 | -3.44 | -5.81 | 1.54E-03 | 3.65E-05 |
| 531.43187 | 27.250 | -2.19 | -1.81 | 3.29E-03 | 4.59E-03 |
| 531.43199 | 26.734 | -1.93 | -2.22 | 8.05E-03 | 4.05E-02 |
| 531.43204 | 27.142 | -1.72 | -1.80 | 1.83E-02 | 4.67E-02 |
| 532.37565 | 22.983 | -4.11 | -3.18 | 9.09E-03 | 1.91E-02 |
| 534.19309 | 2.412 | -3.57 | -3.38 | 5.87E-03 | 3.74E-02 |
| 537.32722 | 24.371 | -2.07 | -1.86 | 1.91E-02 | 4.87E-02 |
| 541.27495 | 13.989 | -6.53 | -5.29 | 4.34E-03 | 3.02E-03 |
| 549.11071 | 5.693 | -9.08 | -9.51 | 1.66E-03 | 2.61E-02 |
| 552.38139 | 28.161 | -2.46 | -2.72 | 3.97E-03 | 1.86E-02 |
| 552.38696 | 22.835 | -2.85 | -3.00 | 2.02E-02 | 3.79E-02 |
| 553.16846 | 23.884 | -2.22 | -2.09 | 8.89E-03 | 4.20E-02 |
| 553.45469 | 23.256 | -2.33 | -2.70 | 1.29E-02 | 3.97E-02 |
| 557.37699 | 20.607 | -2.52 | -3.71 | 3.38E-02 | 4.49E-02 |
| 559.46277 | 27.070 | -2.44 | -2.47 | 8.49E-03 | 4.24E-02 |
| 561.79806 | 0.846 | -4.84 | -3.89 | 1.64E-03 | 3.62E-02 |
| 563.12627 | 12.165 | -6.55 | -7.71 | 4.14E-03 | 3.55E-03 |
| 569.41353 | 22.835 | -2.86 | -3.00 | 2.03E-02 | 3.78E-02 |
| 569.44962 | 23.286 | -2.07 | -2.61 | 7.91E-03 | 1.87E-02 |
| 569.44976 | 23.421 | -2.44 | -3.00 | 1.06E-02 | 1.82E-02 |
| 571.82691 | 0.844 | -3.01 | -4.62 | 2.33E-02 | 1.93E-02 |
| 573.28653 | 1.993 | -2.86 | -7.42 | 2.25E-04 | 6.71E-05 |
| 586.15316 | 12.074 | -7.55 | -6.54 | 4.95E-03 | 4.48E-02 |
| 587.10269 | 2.257 | -4.38 | -8.55 | 7.37E-03 | 7.59E-04 |
| 587.13751 | 12.300 | -7.68 | -7.54 | 5.28E-03 | 4.13E-02 |
| 589.11674 | 10.136 | -9.82 | -8.17 | 1.74E-03 | 4.96E-02 |
| 589.41852 | 22.991 | -2.63 | -2.53 | 7.93E-03 | 4.81E-02 |
| 590.26534 | 1.221 | -2.82 | -3.88 | 3.42E-03 | 1.89E-03 |
| 596.26293 | 22.856 | -2.90 | -2.95 | 1.03E-02 | 4.21E-02 |
| 596.26341 | 22.693 | -2.45 | -2.53 | 1.13E-02 | 4.68E-02 |
| 597.23977 | 12.158 | -2.40 | -3.04 | 5.49E-04 | 1.05E-03 |
| 599.46011 | 23.308 | -2.56 | -2.27 | 7.68E-03 | 4.52E-02 |
| 600.30906 | 15.411 | -2.26 | -3.86 | 3.65E-02 | 1.12E-02 |
| 601.40315 | 20.619 | -2.52 | -3.69 | 3.23E-02 | 4.24E-02 |

| | | | | | |
|---|---|---|---|---|---|
| 603.13231 | 12.160 | -7.42 | -7.60 | 4.09E-03 | 1.61E-02 |
| 612.15104 | 2.017 | -8.64 | -9.90 | 3.69E-03 | 2.86E-02 |
| 612.15125 | 1.881 | -11.80 | -11.75 | 3.69E-03 | 3.79E-02 |
| 613.43969 | 22.805 | -2.87 | -2.97 | 2.00E-02 | 4.05E-02 |
| 613.47567 | 23.290 | -2.49 | -2.48 | 3.08E-03 | 2.17E-02 |
| 626.15955 | 12.291 | -8.09 | -6.76 | 5.07E-03 | 4.98E-02 |
| 627.18689 | 24.470 | -2.16 | -1.98 | 8.78E-03 | 4.56E-02 |
| 628.09287 | 3.234 | -6.79 | -7.34 | 1.95E-02 | 2.44E-02 |
| 629.78451 | 0.844 | -4.73 | -4.23 | 2.86E-03 | 3.52E-02 |
| 633.44475 | 22.956 | -2.86 | -2.62 | 8.81E-03 | 4.28E-02 |
| 635.37526 | 13.849 | -9.14 | -8.72 | 3.94E-03 | 2.33E-02 |
| 637.29439 | 20.125 | -1.63 | -2.21 | 5.85E-03 | 3.93E-02 |
| 639.81335 | 0.842 | -2.88 | -5.52 | 4.87E-02 | 2.16E-02 |
| 642.15453 | 12.122 | -8.37 | -7.37 | 4.66E-03 | 3.64E-02 |
| 645.42931 | 20.631 | -2.68 | -3.67 | 2.93E-02 | 4.10E-02 |
| 652.10374 | 5.751 | -9.29 | -9.82 | 8.47E-03 | 1.97E-02 |
| 657.46570 | 22.774 | -2.83 | -2.95 | 2.00E-02 | 3.92E-02 |
| 657.50174 | 23.285 | -1.98 | -2.40 | 6.35E-03 | 1.62E-02 |
| 660.31836 | 2.761 | -7.37 | -7.92 | 5.62E-04 | 5.61E-04 |
| 664.09245 | 1.275 | -5.36 | -5.22 | 4.31E-03 | 7.32E-03 |
| 665.32553 | 20.123 | -2.64 | -1.93 | 3.58E-02 | 4.74E-02 |
| 666.06998 | 1.888 | -5.78 | -6.52 | 4.82E-03 | 4.50E-03 |
| 668.09826 | 5.748 | -7.83 | -8.32 | 1.40E-02 | 1.70E-02 |
| 677.47048 | 22.928 | -2.89 | -2.82 | 1.09E-02 | 4.40E-02 |
| 691.82752 | 0.835 | -4.07 | -5.03 | 3.85E-02 | 3.96E-02 |
| 697.77088 | 0.844 | -4.35 | -4.54 | 5.36E-03 | 2.93E-02 |
| 701.49201 | 22.744 | -2.75 | -2.90 | 2.07E-02 | 4.27E-02 |
| 701.52782 | 23.226 | -1.96 | -1.92 | 1.02E-02 | 1.23E-02 |
| 707.50340 | 27.201 | -2.78 | -3.45 | 1.12E-02 | 1.62E-02 |
| 719.36329 | 13.108 | -4.81 | -5.59 | 1.21E-02 | 2.07E-02 |
| 721.49632 | 22.894 | -2.89 | -2.82 | 8.95E-03 | 3.73E-02 |
| 726.37009 | 13.401 | -7.69 | -7.23 | 4.18E-03 | 1.90E-02 |
| 727.39695 | 12.381 | -3.99 | -3.34 | 8.06E-03 | 2.35E-02 |
| 734.40276 | 20.539 | -2.58 | -4.57 | 2.40E-02 | 1.16E-02 |
| 759.81475 | 0.833 | -4.12 | -5.60 | 4.66E-02 | 4.73E-02 |
| 794.42888 | 13.703 | -6.96 | -6.56 | 5.51E-03 | 1.28E-02 |
| 803.39223 | 12.162 | -3.52 | -4.23 | 3.43E-02 | 2.56E-02 |
| 811.40724 | 11.820 | -2.97 | -3.90 | 6.80E-03 | 2.66E-02 |
| 850.38994 | 21.025 | -1.61 | -2.93 | 4.16E-02 | 2.75E-02 |
| 870.23685 | 1.890 | -5.25 | -5.58 | 8.44E-03 | 4.87E-03 |
| 873.18136 | 1.872 | -7.78 | -7.78 | 4.96E-03 | 1.79E-02 |
| 877.17912 | 11.247 | -7.59 | -7.55 | 2.56E-03 | 4.63E-02 |
| 878.27300 | 24.374 | -2.60 | -2.36 | 1.54E-02 | 4.63E-02 |
| 879.14672 | 11.764 | -7.86 | -7.28 | 5.72E-03 | 4.54E-02 |
| 888.56171 | 16.100 | -3.08 | -3.00 | 1.13E-02 | 4.68E-02 |
| 894.15731 | 11.400 | -8.35 | -8.23 | 4.67E-03 | 3.23E-02 |
| 1039.59072 | 14.691 | -6.49 | -5.86 | 5.77E-03 | 1.97E-02 |
| 1143.65445 | 0.843 | -3.04 | -4.62 | 2.42E-02 | 1.93E-02 |
| 1192.65652 | 14.069 | -2.45 | -3.16 | 1.38E-02 | 3.33E-02 |
| 1259.56968 | 0.843 | -4.76 | -4.61 | 2.76E-03 | 3.33E-02 |
| 1279.62761 | 0.841 | -2.86 | -5.42 | 4.26E-02 | 1.93E-02 |
| 1690.86495 | 15.055 | -7.79 | -7.77 | 3.97E-03 | 6.40E-03 |
| 1795.82207 | 14.031 | -8.26 | -8.24 | 4.36E-03 | 4.68E-03 |
| 1908.91261 | 14.475 | -6.62 | -7.30 | 1.03E-02 | 2.64E-02 |

**Supplemental Table S3.** Target Compounds produced in response to *Candida* infection.

| Molecular Weight | RT [min] | Log2 Fold Change: (MonoCand) / (HumCand) | Log2 Fold Change: (HumUn) / (HumCand) | P-value: (MonoCand) / (HumCand) | P-value: (HumUn) / (HumCand) |
|---|---|---|---|---|---|
| 201.07486 | 0.975 | -1.09 | -1.04 | 3.71E-05 | 1.39E-03 |
| 205.07728 | 1.701 | -2.71 | -1.24 | 1.58E-04 | 7.52E-03 |
| 211.09564 | 1.649 | -1.17 | -1.51 | 9.99E-03 | 3.31E-03 |
| 211.09570 | 2.585 | -1.19 | -1.29 | 1.31E-02 | 3.70E-02 |
| 243.11071 | 13.181 | -1.31 | -1.33 | 4.33E-03 | 1.57E-02 |
| 248.10457 | 16.250 | -5.65 | -1.26 | 2.32E-06 | 1.77E-03 |
| 251.10170 | 2.650 | -2.75 | -1.82 | 3.73E-05 | 1.38E-03 |
| 258.09629 | 1.041 | -1.46 | -1.20 | 2.71E-05 | 1.02E-03 |
| 259.07598 | 0.908 | -1.95 | -2.41 | 6.63E-03 | 2.47E-03 |
| 265.64015 | 11.918 | -1.67 | -1.56 | 2.25E-02 | 2.78E-02 |
| 267.09656 | 2.413 | -4.55 | -2.23 | 4.09E-04 | 4.14E-05 |
| 268.05177 | 9.113 | -8.56 | -1.71 | 3.27E-07 | 1.13E-02 |
| 271.11657 | 1.529 | -2.65 | -1.41 | 3.54E-04 | 2.84E-02 |
| 272.11193 | 1.546 | -2.11 | -1.42 | 7.92E-06 | 5.94E-03 |
| 272.13700 | 1.577 | -2.26 | -1.62 | 9.87E-04 | 6.64E-03 |
| 275.08271 | 13.147 | -1.92 | -1.62 | 3.32E-03 | 1.66E-02 |
| 275.08282 | 12.026 | -1.96 | -1.49 | 1.96E-03 | 1.53E-02 |
| 285.23016 | 16.114 | -2.81 | -1.85 | 9.23E-05 | 4.78E-02 |
| 297.19393 | 18.697 | -2.31 | -2.30 | 6.15E-03 | 2.15E-02 |
| 301.12715 | 1.318 | -1.30 | -1.43 | 1.23E-02 | 1.57E-02 |
| 301.12718 | 1.771 | -3.33 | -1.13 | 2.44E-03 | 2.94E-03 |
| 324.17987 | 11.899 | -7.21 | -2.59 | 6.40E-06 | 1.30E-03 |
| 325.18873 | 16.306 | -4.06 | -1.15 | 5.19E-07 | 4.46E-02 |
| 330.11748 | 1.154 | -1.30 | -1.15 | 8.03E-04 | 3.48E-03 |
| 331.06787 | 1.423 | -3.83 | -4.41 | 3.97E-03 | 8.79E-03 |
| 343.17429 | 13.260 | -2.91 | -1.29 | 2.61E-03 | 7.21E-03 |
| 344.07956 | 18.793 | -1.21 | -1.51 | 3.99E-03 | 1.05E-03 |
| 347.06263 | 1.160 | -5.43 | -3.59 | 1.98E-04 | 1.99E-03 |
| 361.31891 | 20.303 | -7.44 | -1.09 | 1.65E-05 | 4.79E-02 |
| 362.15090 | 12.765 | -2.31 | -1.15 | 3.45E-03 | 9.42E-03 |
| 363.05745 | 1.201 | -2.10 | -2.45 | 1.45E-02 | 5.23E-03 |
| 363.05749 | 1.539 | -2.47 | -2.55 | 2.43E-03 | 3.93E-03 |
| 374.17233 | 15.957 | -5.01 | -1.54 | 3.02E-05 | 9.40E-03 |
| 377.35012 | 23.394 | -1.17 | -1.43 | 3.07E-02 | 9.74E-03 |
| 383.26684 | 18.042 | -2.67 | -1.56 | 3.86E-03 | 4.92E-02 |
| 384.18942 | 1.861 | -5.29 | -1.48 | 1.83E-05 | 1.10E-04 |
| 400.10655 | 12.780 | -5.28 | -2.46 | 1.02E-05 | 1.32E-03 |
| 402.29149 | 22.011 | -3.07 | -1.83 | 3.55E-02 | 8.62E-03 |
| 409.20930 | 15.959 | -5.78 | -1.55 | 8.63E-06 | 2.75E-03 |
| 409.84062 | 0.846 | -2.90 | -2.29 | 9.37E-05 | 6.72E-03 |
| 414.12236 | 13.887 | -5.12 | -3.02 | 3.17E-05 | 7.86E-04 |
| 422.19370 | 15.731 | -8.38 | -1.26 | 1.90E-06 | 7.66E-03 |
| 434.13253 | 21.400 | -1.26 | -1.81 | 1.19E-02 | 4.19E-02 |
| 436.18077 | 1.044 | -2.60 | -1.91 | 1.27E-04 | 1.77E-03 |
| 439.22054 | 15.722 | -12.37 | -1.21 | 1.53E-07 | 1.12E-02 |
| 444.22551 | 15.979 | -7.87 | -1.34 | 3.32E-06 | 2.04E-03 |
| 457.25758 | 14.152 | -9.28 | -1.60 | 2.77E-07 | 2.74E-03 |
| 458.24107 | 16.480 | -3.02 | -1.66 | 8.24E-03 | 1.80E-02 |
| 459.27311 | 15.826 | -6.91 | -1.26 | 2.29E-06 | 1.73E-02 |
| 466.20760 | 15.980 | -8.53 | -1.32 | 2.03E-06 | 9.55E-04 |
| 477.82814 | 0.847 | -3.06 | -2.52 | 4.74E-05 | 5.79E-03 |
| 487.23176 | 18.208 | -6.26 | -1.48 | 4.29E-05 | 6.32E-03 |
| 487.26795 | 14.316 | -5.60 | -1.50 | 4.11E-06 | 5.49E-03 |
| 489.28332 | 15.982 | -8.74 | -1.53 | 3.67E-06 | 1.98E-03 |
| 491.26225 | 19.414 | -3.55 | -1.25 | 5.09E-03 | 3.86E-02 |
| 492.13306 | 14.741 | -4.65 | -1.88 | 6.35E-04 | 4.41E-02 |
| 514.32266 | 11.735 | -1.53 | -2.29 | 7.79E-03 | 7.65E-05 |
| 531.27977 | 11.920 | -1.99 | -1.94 | 2.01E-02 | 4.72E-02 |
| 534.19309 | 2.412 | -4.31 | -2.25 | 2.32E-04 | 4.07E-05 |
| 544.27427 | 1.565 | -2.25 | -1.62 | 8.25E-04 | 6.79E-03 |
| 568.28570 | 12.585 | -6.25 | -2.54 | 3.07E-06 | 3.09E-03 |
| 569.44962 | 23.286 | -2.16 | -1.86 | 4.00E-02 | 4.74E-03 |
| 576.28608 | 2.273 | -4.19 | -4.05 | 1.95E-02 | 1.93E-02 |
| 599.46011 | 23.308 | -1.36 | -1.61 | 2.33E-02 | 1.31E-03 |
| 613.47567 | 23.290 | -1.65 | -1.53 | 4.50E-03 | 8.56E-03 |
| 657.50174 | 23.285 | -2.06 | -1.95 | 1.54E-02 | 3.50E-03 |
| 671.39622 | 13.210 | -1.12 | -1.24 | 1.60E-02 | 2.98E-03 |
| 701.52782 | 23.226 | -2.23 | -3.03 | 2.77E-02 | 3.24E-03 |
| 729.15936 | 12.708 | -7.37 | -1.79 | 6.83E-07 | 4.80E-02 |
| 743.17508 | 13.423 | -5.52 | -2.43 | 3.36E-05 | 2.83E-02 |
| 787.41922 | 13.610 | -1.37 | -1.13 | 1.44E-04 | 5.67E-03 |
| 789.33925 | 8.794 | -6.24 | -1.53 | 9.69E-04 | 1.76E-02 |
| 852.42355 | 13.088 | -3.38 | -1.65 | 4.87E-04 | 1.89E-02 |
| 887.48635 | 14.818 | -3.99 | -2.87 | 6.11E-05 | 3.43E-03 |
| 900.49216 | 13.681 | -2.09 | -2.19 | 1.05E-02 | 1.18E-02 |
| 906.44739 | 12.381 | -1.74 | -1.13 | 1.04E-03 | 8.45E-03 |
| 910.43236 | 15.981 | -4.70 | -2.88 | 1.32E-03 | 6.83E-04 |
| 911.47117 | 12.548 | -5.18 | -3.93 | 3.81E-05 | 2.81E-02 |
| 939.50251 | 14.865 | -1.94 | -1.37 | 9.25E-04 | 1.63E-02 |
| 1028.50089 | 15.029 | -2.79 | -2.87 | 1.79E-02 | 1.49E-02 |
| 1040.59812 | 14.478 | -3.96 | -3.63 | 1.02E-03 | 1.15E-03 |
| 1173.64252 | 15.720 | -3.15 | -4.24 | 4.19E-02 | 1.40E-03 |
| 1193.34153 | 9.749 | -10.04 | -1.29 | 1.81E-07 | 1.99E-02 |

**Supplemental Table S4.** Peptidomics results from *Salmonella* infection.

| Protein Accession | Peptide | GS:HS:NO (Ratio) | Start | End | PTM |
|---|---|---|---|---|---|
| O08677\|KNG1_MOUSE | I.AQLNAENDHPF.Y | 1.00:34.00:29.61 | 286 | 296 | |
| O08677\|KNG1_MOUSE | K.SGNQYM(+15.99)LH.R | 1.00:0.44:0.03 | 50 | 57 | Oxidation (M) |
| O08677\|KNG1_MOUSE | R.ENEFFIV.T | 1.00:0.08:0 | 115 | 121 | |
| O08677\|KNG1_MOUSE | R.RPPGFSPF.R | 1.00:2.02:0 | 380 | 387 | |
| O08677\|KNG1_MOUSE | R.SVTVQETK.E | 1.00:96.35:0 | 389 | 396 | |
| O08677\|KNG1_MOUSE | I.STDSPDLEPV.L | 1.00:2.02:0 | 149 | 158 | |
| O08677\|KNG1_MOUSE | A.QLNAENDHPFYYK.I | 1.00:27.07:0 | 287 | 299 | |
| O08677\|KNG1_MOUSE | K.SGNQYM(+15.99)LHR.V | 1.00:143.54:0 | 50 | 58 | Oxidation (M) |
| O08677\|KNG1_MOUSE | R.VIEGTK.T | 0.000694444 | 59 | 64 | |
| O08677\|KNG1_MOUSE | R.DIPVDSPELK.E | 1.00:22.68:0 | 269 | 278 | |
| O08677\|KNG1_MOUSE | A.QLNAENDHPFYY.K | 1.00:0:0 | 287 | 298 | |
| O08677\|KNG1_MOUSE | I.SDFPEATSPK.C | 1.00:1.68:0 | 621 | 630 | |
| O08677\|KNG1_MOUSE | R.DAETEQGPTHGHGWLHEK.Q | 1.00:4.66:0 | 414 | 431 | |
| O08677\|KNG1_MOUSE | R.DIPVDSPELKEVLGHSIA.Q | 1.00:1.29:0 | 269 | 286 | |
| O08677\|KNG1_MOUSE | K.EVLGHSIA.Q | 1.00:0.74:0 | 279 | 286 | |
| O08677\|KNG1_MOUSE | R.VIEGTKTDGSPTFYSFK.Y | 1.00:22.37:0 | 59 | 75 | |
| O08677\|KNG1_MOUSE | K.ATSQVVAGTK.Y | 1.00:2.90:0 | 306 | 315 | |
| O08677\|KNG1_MOUSE | A.QLNAENDHPF.Y | 1.00:0.07:0 | 287 | 296 | |
| O08677\|KNG1_MOUSE | L.ISDFPEATSPK.C | 1.00:1.00:0 | 620 | 630 | |
| O08677\|KNG1_MOUSE | R.SVTVQETKEGR.T | 0.000694444 | 389 | 399 | |
| O08677\|KNG1_MOUSE | R.RPPGFSPFR.S | 1.00:22.17:0 | 380 | 388 | |
| O08677\|KNG1_MOUSE | I.AQLNAENDHPFYYK.I | 1.00:3.29:0 | 286 | 299 | |
| O08677\|KNG1_MOUSE | K.YVIEFIAR.E | 0.000694444 | 316 | 323 | |
| O08677\|KNG1_MOUSE | T.FYSFK.Y | 0.000694444 | 71 | 75 | |
| O08677\|KNG1_MOUSE | A.QLNAENDHPFYYKIDTVK.K | 0.000694444 | 287 | 304 | |
| O08677\|KNG1_MOUSE | R.VIEGTKTDGSPT.F | 1.00:1.85:0 | 59 | 70 | |
| O08677\|KNG1_MOUSE | R.DIPVDSPELKEV.L | 1.00:0.25:0 | 269 | 280 | |
| O08677\|KNG1_MOUSE | K.VVPTVK.C | 1.00:1.14:0 | 363 | 368 | |
| O08677\|KNG1_MOUSE | R.EDGDDHTHTV.G | 0.000694444 | 484 | 493 | |
| O08677\|KNG1_MOUSE | K.KATSQVVAGTK.Y | 1.00:12.54:0 | 305 | 315 | |
| O08677\|KNG1_MOUSE | A.QLNAENDHPFY.Y | 1.00:0:0 | 287 | 297 | |
| O08677\|KNG1_MOUSE | H.DDLIPDIHVQPD.S | 0.000694444 | 602 | 613 | |
| O08677\|KNG1_MOUSE | Y.SGDDLVEALPK.P | 1.00:0.47:0 | 251 | 261 | |
| O08677\|KNG1_MOUSE | R.ENEFFIVT.Q | 1.00:2.95:0.32 | 115 | 122 | |
| O08677\|KNG1_MOUSE | H.DDLIPDIHVQPDSLS.F | 1.00:1.59:0 | 602 | 616 | |
| O08677\|KNG1_MOUSE | L.IPDIHVQPDSLS.F | 0.000694444 | 605 | 616 | |
| O08677\|KNG1_MOUSE | K.TDGSPTFYSFK.Y | 1.00:0:0 | 65 | 75 | |
| O08677\|KNG1_MOUSE | K.IDTVKK.A | 1.00:1.89:0 | 300 | 305 | |
| O08677\|KNG1_MOUSE | T.EFSDFDLLDALS | 1.00:35.47:0 | 650 | 661 | |
| O08677\|KNG1_MOUSE | I.AQLNAENDHPFYY.K | 1.00:0.04:0 | 286 | 298 | |
| O08677\|KNG1_MOUSE | R.HQREDGDDHTHTV.G | 0:1.00:0.14 | 481 | 493 | |
| O08677\|KNG1_MOUSE | R.DIPVDSPEL.K | 1.00:0:0 | 269 | 277 | |
| O08677\|KNG1_MOUSE | H.DDLIPDIHVQPDSLSFK.L | 0.000694444 | 602 | 618 | |
| O08677\|KNG1_MOUSE | D.DLVEALPK.P | 1.00:0:0 | 254 | 261 | |
| O08677\|KNG1_MOUSE | R.GNLFM(+15.99)DI.N | 1.00:0:0 | 230 | 236 | Oxidation (M) |
| O08677\|KNG1_MOUSE | H.DDLIPDIHVQPDSLSF.K | 0.000694444 | 602 | 617 | |
| O08677\|KNG1_MOUSE | V.DFSLKQFNPGVK.S | 0.000694444 | 38 | 49 | |
| O08677\|KNG1_MOUSE | V.TVQETKEGR.T | 0.000694444 | 391 | 399 | |
| O08677\|KNG1_MOUSE | N.DHPFYY.K | 1.00:0:0 | 293 | 298 | |
| O08677\|KNG1_MOUSE | H.DDLIPDIHVQPDSL.S | 0:1.00:0.10 | 602 | 615 | |
| O08677\|KNG1_MOUSE | K.LISDFPEATSPK.C | 0.000694444 | 619 | 630 | |
| P01027\|CO3_MOUSE | I.HTLDPEK.L | 1.00:1.02:0 | 945 | 951 | |
| P01027\|CO3_MOUSE | I.HTLDPEKLGQGGVQK.V | 0.000694444 | 945 | 959 | |
| P01027\|CO3_MOUSE | S.DQVPDTDSETRIIL.L | 0.000694444 | 969 | 981 | |
| P01027\|CO3_MOUSE | R.SEETKQNEAFSLT.A | 1.00:25.23:0 | 1321 | 1333 | |
| P01027\|CO3_MOUSE | V.LEAHDAQGDIPV.T | 1.00:1.24:0 | 46 | 57 | |
| P01027\|CO3_MOUSE | R.DILSSNNQHGILPL.S | 1.00:0.27:0 | 179 | 192 | |
| P01027\|CO3_MOUSE | I.LQGSPVVQM(+15.99)AEDAVD.G | 0.000694444 | 982 | 996 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.IGQQEVEVK.A | 1.00:1.84:0 | 905 | 913 | |
| P01027\|CO3_MOUSE | V.LEAHDAQGDIPVT.V | 1.00:0.01:0 | 46 | 58 | |
| P01027\|CO3_MOUSE | I.VLEAHDAQGDIPV.T | 1.00:1.44:0 | 45 | 57 | |
| P01027\|CO3_MOUSE | D.QVPDTDSETR.I | 1.00:22.11:0 | 970 | 979 | |
| P01027\|CO3_MOUSE | R.SEETKQNEAFS.L | 1.00:2.59:0 | 1321 | 1331 | |
| P01027\|CO3_MOUSE | I.HTLDPEKLGQGGVQKV.D | 1.00:7.61:0 | 945 | 960 | |
| P01027\|CO3_MOUSE | I.HTLDPEKLGQGGVQ.K | 1.00:19.35:0 | 945 | 958 | |
| P01027\|CO3_MOUSE | K.SSVAVPYVIVPLK.I | 1.00:11.47:0 | 892 | 904 | |
| P01027\|CO3_MOUSE | K.ALTQDDGVAK.L | 1.00:2.08:0 | 399 | 408 | |
| P01027\|CO3_MOUSE | K.VLVVTQGSNAK.A | 1.00:2.70:0 | 388 | 398 | |
| P01027\|CO3_MOUSE | L.DPEKLGQGGVQ.K | 0.000694444 | 948 | 958 | |
| P01027\|CO3_MOUSE | H.TLDPEKLGQGGVQ.K | 1.00:6.40:0 | 946 | 958 | |
| P01027\|CO3_MOUSE | R.DNHLAPGQQT.T | 1.00:0.27:0.02 | 571 | 580 | |
| P01027\|CO3_MOUSE | R.YFQTIK.I | 1.00:0.32:0 | 882 | 887 | |
| P01027\|CO3_MOUSE | S.EETKQNEAFSLT.A | 0.000694444 | 1322 | 1333 | |
| P01027\|CO3_MOUSE | K.LSINTPNSR.Q | 1.00:1.24:0 | 409 | 417 | |
| P01027\|CO3_MOUSE | A.QYQTDVPDHKDLNM(+15.99)DV.S | 1.00:2.52:0 | 1281 | 1296 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.VELLHNPA.F | 1.00:0.08:0 | 864 | 871 | |
| P01027\|CO3_MOUSE | K.TVLTGASGHL.R | 1.00:0.07:0 | 75 | 84 | |
| P01027\|CO3_MOUSE | R.VSIRPAPET.A | 1.00:0:0 | 1365 | 1373 | |
| P01027\|CO3_MOUSE | I.EELKEPEK.N | 1.00:0.29:0 | 776 | 783 | |
| P01027\|CO3_MOUSE | I.AVHYLDQTEQWEK.F | 1.00:1.75:0 | 1024 | 1036 | |
| P01027\|CO3_MOUSE | E.DAVDGER.L | 1.00:1.28:1.63 | 993 | 999 | |

| | | | | | |
|---|---|---|---|---|---|
| P01027\|CO3_MOUSE | Q.EALELIK.K | 1.00:33.89:0 | 1044 | 1050 | |
| P01027\|CO3_MOUSE | K.TVAIHTLDPEKLGQGGVQ.K | 0.000694444 | 941 | 958 | |
| P01027\|CO3_MOUSE | K.DFDSVPPVVR.W | 0.000694444 | 1245 | 1254 | |
| P01027\|CO3_MOUSE | K.TSQGLQTEQR.A | 0.000694444 | 647 | 656 | |
| P01027\|CO3_MOUSE | A.SYM(+15.99)NLQRPYTV.A | 1.00:10.92:0 | 1179 | 1189 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.M(+15.99)ELKPGDNLNVNFHLR.T | 0.000694444 | 463 | 478 | Oxidation (M) |
| P01027\|CO3_MOUSE | T.SPYQI.H | 1.00:0:0 | 352 | 356 | |
| P01027\|CO3_MOUSE | K.LEEPYLGK.F | 1.00:20.21:0 | 1202 | 1209 | |
| P01027\|CO3_MOUSE | K.LGQGGVQK.V | 1.00:10.12:0 | 952 | 959 | |
| P01027\|CO3_MOUSE | D.Q(-17.03)VPDTDSETRII.L | 0.000694444 | 970 | 981 | Pyro-glu from Q |
| P01027\|CO3_MOUSE | D.TDSETRII.L | 1.00:126.65:0 | 974 | 981 | |
| P01027\|CO3_MOUSE | K.VDVPAADLSDQVPDTDSETRII.L | 0.000694444 | 960 | 981 | |
| P01027\|CO3_MOUSE | K.VDVPAADLSDQVPDTDSET.R | 1.00:0.95:0 | 960 | 978 | |
| P01027\|CO3_MOUSE | K.AGQYTDK.G | 0.000694444 | 682 | 688 | |
| P01027\|CO3_MOUSE | A.ADLSDQVPDTDSETR.I | 1.00:62.91:0 | 965 | 979 | |
| P01027\|CO3_MOUSE | K.VDVPAADLSDQVPDTDSETR.I | 0.000694444 | 960 | 979 | |
| P01027\|CO3_MOUSE | S.NADALVGK.S | 1.00:0.64:0 | 316 | 323 | |
| P01027\|CO3_MOUSE | I.VLEAHDAQGDIPVT.V | 1.00:0:0 | 45 | 58 | |
| P01027\|CO3_MOUSE | G.IPM(+15.99)YSIITPNVL.R | 1.00:0.41:0 | 25 | 36 | Oxidation (M) |
| P01027\|CO3_MOUSE | V.HYLDQTEQWEK.F | 1.00:0.82:0 | 1026 | 1036 | |
| P01027\|CO3_MOUSE | I.FGVQDGDK.K | 1.00:0.69:0 | 274 | 281 | |
| P01027\|CO3_MOUSE | R.DNHLAPGQQTTLR.I | 0.000694444 | 571 | 583 | |
| P01027\|CO3_MOUSE | R.AFYEHAPK.Q | 1.00:0.17:0 | 209 | 216 | |
| P01027\|CO3_MOUSE | R.WEEPDQQLYNVEAT.S | 1.00:3.28:0 | 1220 | 1233 | |
| P01027\|CO3_MOUSE | K.AAVFNHFI.S | 1.00:0.12:0.04 | 914 | 921 | |
| P01027\|CO3_MOUSE | Q.VPDTDSETR.I | 1.00:46.96:1.48 | 971 | 979 | |
| P01027\|CO3_MOUSE | K.AAVFNHFISDGVK.K | 1.00:1.80:0 | 914 | 926 | |
| P01027\|CO3_MOUSE | K.TM(+15.99)EAHPYST.M | 1.00:0.48:0 | 440 | 448 | Oxidation (M) |
| P01027\|CO3_MOUSE | I.LQGSPVVQM(+15.99)AEDAVDGERLK.H | 0.000694444 | 982 | 1001 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.TVAIHTLDPEKLGQGGVQK.V | 0.000694444 | 941 | 959 | |
| P01027\|CO3_MOUSE | V.TVQDFLK.R | 1.00:73.72:0 | 60 | 66 | |
| P01027\|CO3_MOUSE | V.DVPAADLSDQVPDTDSETR.I | 0.000694444 | 961 | 979 | |
| P01027\|CO3_MOUSE | R.DNHLAPGQQTT.L | 1.00:0.18:0 | 571 | 581 | |
| P01027\|CO3_MOUSE | R.NRWEEPDQQLYNV.E | 1.00:0.62:0 | 1218 | 1230 | |
| P01027\|CO3_MOUSE | A.ADLSDQVPDTDSET.R | 1.00:2.60:6.10 | 965 | 978 | |
| P01027\|CO3_MOUSE | V.LEAHDAQGDIPVTV.T | 0.000694444 | 46 | 59 | |
| P01027\|CO3_MOUSE | M.ELKPGDNLNVN.F | 1.00:6.60:0 | 464 | 474 | |
| P01027\|CO3_MOUSE | V.ANFGETV.V | 1.00:0.59:0.22 | 111 | 117 | |
| P01027\|CO3_MOUSE | R.SGIPIVTSPYQI.H | 1.00:0.75:0 | 345 | 356 | |
| P01027\|CO3_MOUSE | K.FFKPAM(+15.99)PFDLM(+15.99)V.F | 1.00:0.33:0.08 | 364 | 375 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.SELEEDIIPEEDIIS.R | 1.00:0.03:0 | 749 | 763 | |
| P01027\|CO3_MOUSE | R.SVQLM(+15.99)ER.R | 0.000694444 | 671 | 677 | Oxidation (M) |
| P01027\|CO3_MOUSE | S.IITPNVLR.L | 0.000694444 | 30 | 37 | |
| P01027\|CO3_MOUSE | I.LIETPDGIPVKR.D | 0.000694444 | 167 | 178 | |
| P01027\|CO3_MOUSE | A.VFNHFISDGVK.K | 1.00:4.39:0 | 916 | 926 | |
| P01027\|CO3_MOUSE | K.IWDVVEK.A | 1.00:0:0 | 615 | 621 | |
| P01027\|CO3_MOUSE | K.ALTQDDGVAKL.S | 1.00:2.19:0 | 399 | 409 | |
| P01027\|CO3_MOUSE | R.SEETKQNEAFSLTAK.G | 0.000694444 | 1321 | 1335 | |
| P01027\|CO3_MOUSE | I.LIETPDGIPV.K | 1.00:0.07:0 | 167 | 176 | |
| P01027\|CO3_MOUSE | K.VLM(+15.99)EGVRPS.N | 1.00:0:0 | 307 | 315 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.SELEEDIIPEEDII.S | 1.00:19.73:0 | 749 | 762 | |
| P01027\|CO3_MOUSE | R.EQEELKV.R | 1.00:0.23:0 | 856 | 862 | |
| P01027\|CO3_MOUSE | R.IFTVDNNLLPV.G | 1.00:0.31:0 | 150 | 160 | |
| P01027\|CO3_MOUSE | A.DLSDQVPDTDSETR.I | 1.00:1.31:0.04 | 966 | 979 | |
| P01027\|CO3_MOUSE | R.VVIEDGVGDAVLTR.K | 0.000694444 | 292 | 305 | |
| P01027\|CO3_MOUSE | R.MELKPGDNLNV.N | 1.00:0.01:0 | 463 | 473 | |
| P01027\|CO3_MOUSE | V.DVPAADLSDQVPDTDSET.R | 1.00:1.09:0 | 961 | 978 | |
| P01027\|CO3_MOUSE | K.NYAGVFM(+15.99)DAGLAFK.T | 0.000694444 | 633 | 646 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.QKPDGVFQEDGPVIHQEM(+15.99)I.G | 1.00:1.68:0 | 1112 | 1130 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.SEETKQNEA.F | 1.00:4.89:0 | 1321 | 1329 | |
| P01027\|CO3_MOUSE | A.LLLLKDFDSVPPVVR.W | 0.000694444 | 1240 | 1254 | |
| P01027\|CO3_MOUSE | I.EELKEPEKNGI.S | 1.00:0.67:0 | 776 | 786 | |
| P01027\|CO3_MOUSE | S.EETKQNEAFSL.T | 0:1.00:17.94 | 1322 | 1332 | |
| P01027\|CO3_MOUSE | R.QEALELIK.K | 1.00:57.21:0 | 1043 | 1050 | |
| P01027\|CO3_MOUSE | V.IFGVQDGDKK.I | 1.00:1.41:0 | 273 | 282 | |
| P01027\|CO3_MOUSE | K.TVAIHTLDPEK.L | 0.000694444 | 941 | 951 | |
| P01027\|CO3_MOUSE | K.RQEALELIK.K | 0.000694444 | 1042 | 1050 | |
| P01027\|CO3_MOUSE | A.LLLLKDFDSVPPVV.R | 1.00:0.60:0 | 1240 | 1253 | |
| P01027\|CO3_MOUSE | K.QKPDGVFQEDGPVIHQEM(+15.99)IGGFR.N | 0.000694444 | 1112 | 1134 | Oxidation (M) |
| P01027\|CO3_MOUSE | D.Q(-17.03)VPDTDSETR.I | 1.00:384.00:0 | 970 | 979 | Pyro-glu from Q |
| P01027\|CO3_MOUSE | A.ADLSDQVPDTDSETRII.L | 0.000694444 | 965 | 981 | |
| P01027\|CO3_MOUSE | V.IEDGVGDAVLTR.K | 0.000694444 | 294 | 305 | |
| P01027\|CO3_MOUSE | R.DNHLAPGQQTTL.R | 1.00:0.42:0 | 571 | 582 | |
| P01027\|CO3_MOUSE | L.SLPITPEFIPSF.R | 1.00:0:0 | 518 | 529 | |
| P01027\|CO3_MOUSE | L.SLPITPEFIPSFR.L | 0.000694444 | 518 | 530 | |
| P01027\|CO3_MOUSE | T.FYYIDDPNGLEV.S | 1.00:0.85:0.87 | 243 | 254 | |
| P01027\|CO3_MOUSE | S.DQVPDTDSETR.I | 0.000694444 | 969 | 979 | |
| P01027\|CO3_MOUSE | K.AGEYIEASYM(+15.99)NLQRPYTV.A | 0.000694444 | 1172 | 1189 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.TVAIHTLDPEKLGQGGVQKV.D | 0.000694444 | 941 | 960 | |
| P01027\|CO3_MOUSE | V.ANFGETVVEK.A | 1.00:0:0 | 111 | 120 | |
| P01027\|CO3_MOUSE | R.IEGNQGAR.V | 0.000694444 | 584 | 591 | |
| P01027\|CO3_MOUSE | R.DILSSNNQHGILPLSWNIPELV.N | 1.00:4.77:0 | 179 | 200 | |
| P01027\|CO3_MOUSE | S.SNNQHGILPL.S | 1.00:0:12.93 | 183 | 192 | |
| P01027\|CO3_MOUSE | R.KVLM(+15.99)EGVRPSNADALVGK.S | 1.00:9.41:0 | 306 | 323 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.TM(+15.99)EAHPYSTM(+15.99)HNSNNYLHL.S | 0.000694444 | 440 | 458 | Oxidation (M) |
| P01027\|CO3_MOUSE | E.EPDQQLYNVEAT.S | 1.00:1.86:0 | 1222 | 1233 | |
| P01027\|CO3_MOUSE | T.M(+15.99)HNSNNYLHLS.V | 1.00:0.48:0 | 449 | 459 | Oxidation (M) |

| | | | | | |
|---|---|---|---|---|---|
| P01027\|CO3_MOUSE | K.FFKPAM(+15.99)PFDLM(+15.99)VFVTNPDGSPA.S | 1.00:7.17:0 | 364 | 385 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.SGIPIVTSPYQIH.F | 1.00:0:0 | 345 | 357 | |
| P01027\|CO3_MOUSE | R.VVIEDGVGDAV.L | 1.00:0.69:0 | 292 | 302 | |
| P01027\|CO3_MOUSE | R.SEETKQNEAF.S | 1.00:15.70:0 | 1321 | 1330 | |
| P01027\|CO3_MOUSE | R.VGLVAVDK.G | 1.00:0:0 | 592 | 599 | |
| P01027\|CO3_MOUSE | R.SEETKQNEAFSL.T | 1.00:346.23:0 | 1321 | 1332 | |
| P01027\|CO3_MOUSE | I.LQGSPVVQM(+15.99).A | 1.00:40.64:0 | 982 | 990 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.VLM(+15.99)EGVRPSNADALVGK.S | 0.000694444 | 307 | 323 | Oxidation (M) |
| P01027\|CO3_MOUSE | S.EETKQNEAF.S | 0:1.00:0.38 | 1322 | 1330 | |
| P01027\|CO3_MOUSE | R.VSIRPAPETAK.K | 0.000694444 | 1365 | 1375 | |
| P01027\|CO3_MOUSE | V.DVPAADLSDQVPDTDSETRII.L | 0.000694444 | 961 | 981 | |
| P01027\|CO3_MOUSE | R.YYTYLVM(+15.99)NK.G | 1.00:1.57:0 | 489 | 497 | Oxidation (M) |
| P01027\|CO3_MOUSE | L.SDQVPDTDSETRII.L | 1.00:1.07:0 | 968 | 981 | |
| P01027\|CO3_MOUSE | S.NNQHGILPL.S | 1.00:0:0 | 184 | 192 | |
| P01027\|CO3_MOUSE | V.DNNLLPV.G | 1.00:0.09:0 | 154 | 160 | |
| P01027\|CO3_MOUSE | S.SDLWGEKPNTS.Y | 1.00:0.07:0 | 1609 | 1619 | |
| P01027\|CO3_MOUSE | W.EEPDQQLYNVEAT.S | 0.000694444 | 1221 | 1233 | |
| P01027\|CO3_MOUSE | T.DVPDHKDL.N | 0.000694444 | 1285 | 1292 | |
| P01027\|CO3_MOUSE | L.DPEKLGQGGVQK.V | 0.000694444 | 948 | 959 | |
| P01027\|CO3_MOUSE | A.IHTLDPEK.L | 0.000694444 | 944 | 951 | |
| P01027\|CO3_MOUSE | R.DNHLAPGQQTT(+79.96)LRIE(-.98).G | 1.00:2.81:0 | 571 | 585 | Sulfation; Amidation |
| P01027\|CO3_MOUSE | S.DLWGEKPNTS.Y | 1.00:0.95:0 | 1610 | 1619 | |
| P01027\|CO3_MOUSE | K.SSVAVPYVIVPL.K | 1.00:0.18:0 | 892 | 903 | |
| P01027\|CO3_MOUSE | R.SELEEDIIPEEDIISR.S | 0.000694444 | 749 | 764 | |
| P01027\|CO3_MOUSE | V.FQEDGPVIHQEM(+15.99)I.G | 1.00:7.95:0 | 1118 | 1130 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.M(+15.99)ELKPGDNLNVNFHL.R | 1.00:0.12:0 | 463 | 477 | Oxidation (M) |
| P01027\|CO3_MOUSE | A.IHTLDPEKLGQGGVQK.V | 0.000694444 | 944 | 959 | |
| P01027\|CO3_MOUSE | K.QKPDGVFQEDGPVIH.Q | 1.00:0.02:0 | 1112 | 1126 | |
| P01027\|CO3_MOUSE | K.RQEALELI.K | 1.00:0.67:0 | 1042 | 1049 | |
| P01027\|CO3_MOUSE | K.TVLTGASGHLR.S | 0.000694444 | 75 | 85 | |
| P01027\|CO3_MOUSE | K.FLNTAK.D | 1.00:10.75:0 | 1210 | 1215 | |
| P01027\|CO3_MOUSE | I.HTLDPEKLGQGGVQKVD.V | 0.000694444 | 945 | 961 | |
| P01027\|CO3_MOUSE | V.LEAHDAQG.D | 1.00:2.68:0 | 46 | 53 | |
| P01027\|CO3_MOUSE | V.VLSLPITPEFIPSFR.L | 0.000694444 | 516 | 530 | |
| P01027\|CO3_MOUSE | D.QVPDTDSETRIILQGSPVVQM(+15.99).A | 0.000694444 | 970 | 990 | Oxidation (M) |
| P01027\|CO3_MOUSE | V.TVQDFLKR.Q | 0.000694444 | 60 | 67 | |
| P01027\|CO3_MOUSE | A.EDAVDGER.L | 1.00:2.06:0 | 992 | 999 | |
| P01027\|CO3_MOUSE | V.VLSLPITPEFIPSF.R | 1.00:0.18:0 | 516 | 529 | |
| P01027\|CO3_MOUSE | E.DGPVIH.Q | 1.00:0.09:0 | 1121 | 1126 | |
| P01027\|CO3_MOUSE | K.WLILEK.Q | 1.00:0.25:0 | 1106 | 1111 | |
| P01027\|CO3_MOUSE | I.LIETPDGIPVK.R | 1.00:2.13:0 | 167 | 177 | |
| P01027\|CO3_MOUSE | M.ELKPGDNLNV.N | 1.00:0:0 | 464 | 473 | |
| P01027\|CO3_MOUSE | R.DHVLGLAR.S | 0.000694444 | 741 | 748 | |
| P01027\|CO3_MOUSE | A.QYQTDVPDHK.D | 1.00:0.63:0 | 1281 | 1290 | |
| P01027\|CO3_MOUSE | E.DGVGDAVLTR.K | 1.00:1.55:0 | 296 | 305 | |
| P01027\|CO3_MOUSE | I.LHSGSDM(+15.99)VEAE.R | 1.00:0:0 | 333 | 343 | Oxidation (M) |
| P01027\|CO3_MOUSE | S.DQVPDTDSET.R | 1.00:0.41:0 | 969 | 978 | |
| P01027\|CO3_MOUSE | R.IFTVDNNLLPVGK.T | 1.00:1.91:0 | 150 | 162 | |
| P01027\|CO3_MOUSE | I.E(-18.01)ELKEPEK.N | 1.00:0.16:0 | 776 | 783 | Pyro-glu from E |
| P01027\|CO3_MOUSE | I.VLEAHDAQGDIPVTV.T | 0:1.00:0.97 | 45 | 59 | |
| P01027\|CO3_MOUSE | K.HLIVTPA.G | 1.00:1.20:0 | 1002 | 1008 | |
| P01027\|CO3_MOUSE | Q.ATFM(+15.99)V.F | 1.00:1.08:0 | 1271 | 1275 | Oxidation (M) |
| P01027\|CO3_MOUSE | T.SQGLQTEQR.A | 0.000694444 | 648 | 656 | |
| P01027\|CO3_MOUSE | R.NRWEEPDQQLYNVEAT.S | 0.000694444 | 1218 | 1233 | |
| P01027\|CO3_MOUSE | K.PAM(+15.99)PFDLM(+15.99)VFVTNPDGSPA.S | 1.00:2.68:0 | 367 | 385 | Oxidation (M) |
| P01027\|CO3_MOUSE | L.SINTPNSR.Q | 0.000694444 | 410 | 417 | |
| P01027\|CO3_MOUSE | I.EELKEPEKNGISTK.V | 0.000694444 | 776 | 789 | |
| P01027\|CO3_MOUSE | R.VGLVAVDKGVFV.L | 1.00:1.94:0 | 592 | 603 | |
| P01027\|CO3_MOUSE | Q.VPDTDSETRII.L | 1.00:1.27:0 | 971 | 981 | |
| P01027\|CO3_MOUSE | K.QKPDGVFQEDGPVIHQEM(+15.99).I | 1.00:0.11:0 | 1112 | 1129 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.LVAYYTLIGASGQR.E | 0.000694444 | 531 | 544 | |
| P01027\|CO3_MOUSE | S.EETKQNEAFSLTA.K | 0.000694444 | 1322 | 1334 | |
| P01027\|CO3_MOUSE | R.SEETKQNEAFSLTAKGK.G | 0.000694444 | 1321 | 1337 | |
| P01027\|CO3_MOUSE | G.IPM(+15.99)YSIITPNVLR.L | 0.000694444 | 25 | 37 | Oxidation (M) |
| P01027\|CO3_MOUSE | I.LQGSPVVQ.M | 0.000694444 | 982 | 989 | |
| P01027\|CO3_MOUSE | M.AEDAVDGER.L | 1.00:7.61:0 | 991 | 999 | |
| P01027\|CO3_MOUSE | R.QVLTSEK.T | 0.000694444 | 68 | 74 | |
| P01027\|CO3_MOUSE | M.NLQRPYTV.A | 1.00:1.37:0 | 1182 | 1189 | |
| P01027\|CO3_MOUSE | K.SGSDEVQAGQQR.K | 0.000694444 | 1571 | 1582 | |
| P01027\|CO3_MOUSE | A.DLSDQVPDTDSETRII.L | 0.000694444 | 966 | 981 | |
| P01027\|CO3_MOUSE | K.AAVFNHFISDGVKK.T | 0.000694444 | 914 | 927 | |
| P01027\|CO3_MOUSE | K.FFKPAM(+15.99)PFDLM(+15.99).V | 1.00:0:0 | 364 | 374 | Oxidation (M) |
| P01027\|CO3_MOUSE | E.DGPVIHQEM(+15.99).I | 1.00:1.88:0.74 | 1121 | 1129 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.M(+15.99)ELKPGDNLNVNFH.L | 1.00:0.35:0 | 463 | 476 | Oxidation (M) |
| P01027\|CO3_MOUSE | S.DQVPDTDSETRIILQGSPVVQM(+15.99).A | 0.000694444 | 969 | 990 | Oxidation (M) |
| P01027\|CO3_MOUSE | L.DPEKLGQGGVQKVD.V | 0.000694444 | 948 | 961 | |
| P01027\|CO3_MOUSE | R.SEETKQNEAFSLTAKGKGR.G | 0.000694444 | 1321 | 1339 | |
| P01027\|CO3_MOUSE | R.WEEPDQQLYNV.E | 1.00:0.46:0 | 1220 | 1230 | |
| P01027\|CO3_MOUSE | R.SHFPQSWLWTI.E | 1.00:15.95:0 | 765 | 775 | |
| P01027\|CO3_MOUSE | K.EFNSDKEGHK.Y | 1.00:0:0 | 96 | 105 | |
| P01027\|CO3_MOUSE | A.IHTLDPEKLGQGGVQ.K | 0.000694444 | 944 | 958 | |
| P01027\|CO3_MOUSE | T.NIELLDDFDEYTM(+15.99)T.I | 1.00:0:0 | 1551 | 1564 | Oxidation (M) |
| P01027\|CO3_MOUSE | V.DVPAADLSD.Q | 0.000694444 | 961 | 969 | |
| P01027\|CO3_MOUSE | H.TLDPEKLGQGGVQKV.D | 1.00:0.74:0 | 946 | 960 | |
| P01027\|CO3_MOUSE | A.EDAVDGERL.K | 1.00:1.19:0 | 992 | 1000 | |
| P01027\|CO3_MOUSE | V.VQM(+15.99)AEDAVD.G | 1.00:3.05:0 | 988 | 996 | Oxidation (M) |
| P01027\|CO3_MOUSE | A.SYM(+15.99)NLQRPYTVA.I | 0.000694444 | 1179 | 1190 | Oxidation (M) |

| | | | | | |
|---|---|---|---|---|---|
| P01027\|CO3_MOUSE | G.IPM(+15.99)YSIITPNVLRLESEETIV.L | 0.000694444 | 25 | 45 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.SHFPQSWLWT.I | 1.00:0:0 | 765 | 774 | |
| P01027\|CO3_MOUSE | E.DGPVIHQEM(+15.99)IGGFR.N | 0.000694444 | 1121 | 1134 | Oxidation (M) |
| P01027\|CO3_MOUSE | A.LM(+15.99)NKLEEPYLG.K | 1.00:0:0 | 1198 | 1208 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.M(+15.99)DKAGQYTDK.G | 0.000694444 | 679 | 688 | Oxidation (M) |
| P01027\|CO3_MOUSE | I.LQGSPVVQM(+15.99)AEDA.V | 1.00:2.65:0 | 982 | 994 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.TIYTPGSTVLYR.I | 0.000694444 | 138 | 149 | |
| P01027\|CO3_MOUSE | K.VDVPAADLSD.Q | 0.000694444 | 960 | 969 | |
| P01027\|CO3_MOUSE | E.DGPVIHQEM(+15.99)IGGF.R | 0.000694444 | 1121 | 1133 | Oxidation (M) |
| P01027\|CO3_MOUSE | D.TDSETRIILQGSPVVQM(+15.99).A | 0.000694444 | 974 | 990 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.GVFVLNK.K | 0.000694444 | 600 | 606 | |
| P01027\|CO3_MOUSE | L.SWNIPELV.N | 1.00:0.59:0 | 193 | 200 | |
| P01027\|CO3_MOUSE | R.EPGQDLVVLSLPITPEFIPSFR.L | 0.000694444 | 509 | 530 | |
| P01027\|CO3_MOUSE | E.DGPVIHQEM(+15.99)I.G | 0.000694444 | 1121 | 1130 | Oxidation (M) |
| P01027\|CO3_MOUSE | E.TKQNEAFSLT.A | 0.000694444 | 1324 | 1333 | |
| P01027\|CO3_MOUSE | E.DIIPEEDIIS.R | 1.00:0:0 | 754 | 763 | |
| P01027\|CO3_MOUSE | K.DTWVEHWPEA.E | 1.00:0:0 | 1625 | 1634 | |
| P01027\|CO3_MOUSE | L.EKQKPDGVF.Q | 0.000694444 | 1110 | 1118 | |
| P01027\|CO3_MOUSE | R.NRWEEPDQQL.Y | 1.00:0.93:0 | 1218 | 1227 | |
| P01027\|CO3_MOUSE | V.SFQSGYLFI.Q | 1.00:0:0 | 125 | 133 | |
| P01027\|CO3_MOUSE | S.TM(+15.99)HNSNNYLHLS.V | 1.00:0:0 | 448 | 459 | Oxidation (M) |
| P01027\|CO3_MOUSE | V.PDTDSETRIILQGSPVVQM(+15.99)AEDAVDGER.L | 0.000694444 | 972 | 999 | Oxidation (M) |
| P01027\|CO3_MOUSE | V.FQEDGPVIH.Q | 1.00:0:0 | 1118 | 1126 | |
| P01027\|CO3_MOUSE | Y.QTDVPDHKDLNM(+15.99)DV.S | 1.00:0.63:0 | 1283 | 1296 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.KVLM(+15.99)EGVRPS.N | 1.00:0.23:0 | 306 | 315 | Oxidation (M) |
| P01027\|CO3_MOUSE | M.DKAGQYTDK.G | 0.000694444 | 680 | 688 | |
| P01027\|CO3_MOUSE | R.YYGGGYGSTQA.T | 0.000694444 | 1261 | 1271 | |
| P01027\|CO3_MOUSE | R.M(+15.99)ELKPGDNLNVN.F | 1.00:0:0 | 463 | 474 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.LGQGGVQKVD.V | 0.000694444 | 952 | 961 | |
| P01027\|CO3_MOUSE | K.AAVFNHFISDGV.K | 1.00:0:0 | 914 | 925 | |
| P01027\|CO3_MOUSE | R.MELKPGDNLN.V | 1.00:0:0 | 463 | 472 | |
| P01027\|CO3_MOUSE | T.GFAPDTKDLEL.L | 1.00:0:0 | 1410 | 1420 | |
| P01027\|CO3_MOUSE | K.AAVFNHFIS.D | 1.00:0:0 | 914 | 922 | |
| P01027\|CO3_MOUSE | V.RNEQVEIR.A | 0.000694444 | 841 | 848 | |
| P01027\|CO3_MOUSE | S.DQVPDTDSETRI.I | 0.000694444 | 969 | 980 | |
| P01027\|CO3_MOUSE | K.VLMEGVRPS.N | 1.00:0:0 | 307 | 315 | |
| P01027\|CO3_MOUSE | L.EKQKPDGV.F | 0.000694444 | 1110 | 1117 | |
| P01027\|CO3_MOUSE | V.VQM(+15.99)AEDAVDGER.L | 0.000694444 | 988 | 999 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.LSINTPNS.R | 1.00:0:0 | 409 | 416 | |
| P01027\|CO3_MOUSE | S.DQVPDTDSETRIILQGSPVVQM(+15.99)AE.D | 0.000694444 | 969 | 992 | Oxidation (M) |
| P01027\|CO3_MOUSE | V.VLSLPITPEFIPS.F | 1.00:0:0 | 516 | 528 | |
| P01027\|CO3_MOUSE | A.VFNHFISDGV.K | 1.00:0:0 | 916 | 925 | |
| P01027\|CO3_MOUSE | K.FFKPAMPFDLM(+15.99)V.F | 1.00:0:0 | 364 | 375 | Oxidation (M) |
| P01027\|CO3_MOUSE | L.KIGQQEVEV.K | 1.00:0:0 | 904 | 912 | |
| P01027\|CO3_MOUSE | A.IHTLDPEKLGQGGVQKVD.V | 0.000694444 | 944 | 961 | |
| P01027\|CO3_MOUSE | K.IGQQEVEV.K | 1.00:0:0 | 905 | 912 | |
| P01027\|CO3_MOUSE | D.QVPDTDSETRIILQGSPV.V | 0.000694444 | 970 | 987 | |
| P01027\|CO3_MOUSE | T.GFAPDTKDLELL.A | 1.00:0.34:0 | 1410 | 1421 | |
| P01027\|CO3_MOUSE | G.IPM(+15.99)YSIITPNVLRLESEETI.V | 0.000694444 | 25 | 44 | Oxidation (M) |
| P01027\|CO3_MOUSE | T.LDPEKLGQGGVQKV.D | 0.000694444 | 947 | 960 | |
| P01027\|CO3_MOUSE | E.ETKQNEAFSLT.A | 0.000694444 | 1323 | 1333 | |
| P01027\|CO3_MOUSE | W.ENGNLLR.S | 0.000694444 | 1314 | 1320 | |
| P01027\|CO3_MOUSE | K.VDVPAADLS.D | 0.000694444 | 960 | 968 | |
| P01027\|CO3_MOUSE | T.DSETRII.L | 0.000694444 | 975 | 981 | |
| P01027\|CO3_MOUSE | L.DPEKLGQGGVQKV.D | 0.000694444 | 948 | 960 | |
| P01027\|CO3_MOUSE | I.HTLDPEKLGQG.G | 0.000694444 | 945 | 955 | |
| P01027\|CO3_MOUSE | A.NFGETVVEK.A | 1.00:0:0 | 112 | 120 | |
| P01027\|CO3_MOUSE | R.IILQGSPVVQM(+15.99)AE.D | 0.000694444 | 980 | 992 | Oxidation (M) |
| P01027\|CO3_MOUSE | E.DGPVIHQEM(+15.99)IG.G | 0:1.00:0.04 | 1121 | 1131 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.VDVPAADLSDQVPDTDSETRI.I | 0.000694444 | 960 | 980 | |
| P01027\|CO3_MOUSE | R.VVIEDGVGDAVLT.R | 1.00:0:0 | 292 | 304 | |
| P01027\|CO3_MOUSE | L.DQTEQWEK.F | 1.00:2.13:0 | 1029 | 1036 | |
| P01027\|CO3_MOUSE | D.TDSETRIILQGSPV.V | 0.000694444 | 974 | 987 | |
| P01027\|CO3_MOUSE | S.DQVPDTDSETRIILQGSPV.V | 0.000694444 | 969 | 987 | |
| P01027\|CO3_MOUSE | A.TFM(+15.99)VFQALA.Q | 0.000694444 | 1272 | 1280 | Oxidation (M) |
| P01027\|CO3_MOUSE | Q.M(+15.99)AEDAVDGERL.K | 1.00:0:0 | 990 | 1000 | Oxidation (M) |
| P01027\|CO3_MOUSE | K.TVAIHTLDPEKLGQGGVQKVDVPAA.D | 0.000694444 | 941 | 965 | |
| P01027\|CO3_MOUSE | W.GLSSDLWGEKPNTS.Y | 1.00:0:0 | 1606 | 1619 | |
| P01027\|CO3_MOUSE | K.YLM(+15.99)WGLSSDLWGEKPNTS.Y | 1.00:0:0 | 1602 | 1619 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.YISKYEM(+15.99)NK.A | 0.000694444 | 1428 | 1436 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.DILSSNNQHGILP.L | 1.00:0:0 | 179 | 191 | |
| P01027\|CO3_MOUSE | V.PYVIVPLK.I | 0.000694444 | 897 | 904 | |
| P01027\|CO3_MOUSE | R.MELKPGDNLNVNFH.L | 1.00:0:0 | 463 | 476 | |
| P01027\|CO3_MOUSE | R.M(+15.99)ELKPGDNLNV.N | 1.00:0.10:0 | 463 | 473 | Oxidation (M) |
| P01027\|CO3_MOUSE | A.IHTLDPEKLGQGGVQKV.D | 0.000694444 | 944 | 960 | |
| P01027\|CO3_MOUSE | S.LPITPEFIPS.F | 1.00:0:0 | 519 | 528 | |
| P01027\|CO3_MOUSE | Q.M(+15.99)AEDAVDGER.L | 1.00:0:0 | 990 | 999 | Oxidation (M) |
| P01027\|CO3_MOUSE | H.TLDPEKLGQGGVQK.V | 0.000694444 | 946 | 959 | |
| P01027\|CO3_MOUSE | R.MELKPGDNLNVN.F | 1.00:0:0 | 463 | 474 | |
| P01027\|CO3_MOUSE | D.QVPDTDSETRIILQGSPVVQM(+15.99)AE.D | 0.000694444 | 970 | 992 | Oxidation (M) |
| P01027\|CO3_MOUSE | D.TDSETRIILQGSPVVQM(+15.99)AE.D | 0.000694444 | 974 | 992 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.IFTVDNNLLPVG.K | 1.00:0:0 | 150 | 161 | |
| P01027\|CO3_MOUSE | T.LDPEKLGQGGVQK.V | 0.000694444 | 947 | 959 | |
| P01027\|CO3_MOUSE | Y.LFIQTDK.T | 1.00:0:0 | 131 | 137 | |
| P01027\|CO3_MOUSE | R.SEETKQNEAFSLTAKG.K | 0.000694444 | 1321 | 1336 | |
| P01027\|CO3_MOUSE | M.QDFFIDL.R | 1.00:0:0 | 827 | 833 | |
| P01027\|CO3_MOUSE | P.DTDSETRII.L | 1.00:13.46:0 | 973 | 981 | |

| | | | | | |
|---|---|---|---|---|---|
| P01027\|CO3_MOUSE | H.NSNNYLHL.S | 1.00:0:0 | 451 | 458 | |
| P01027\|CO3_MOUSE | D.QVPDTDSETRIILQGSPVVQ.M | 0.000694444 | 970 | 989 | |
| P01027\|CO3_MOUSE | R.SELEEDII.P | 1.00:0:0 | 749 | 756 | |
| P01027\|CO3_MOUSE | M.HNSNNYLHLS.V | 1.00:0:0 | 450 | 459 | |
| P01027\|CO3_MOUSE | R.VM(+15.99)QDFFIDL.R | 1.00:0:0 | 825 | 833 | Oxidation (M) |
| P01027\|CO3_MOUSE | P.DTDSETRIILQGSPVVQM(+15.99).A | 0.000694444 | 973 | 990 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.MELKPGDNLNVNFHL.R | 1.00:0:0 | 463 | 477 | |
| P01027\|CO3_MOUSE | D.TDSETRIILQGSPVVQ.M | 0.000694444 | 974 | 989 | |
| P01027\|CO3_MOUSE | I.HTLDPEKLGQGGVQ(-.98).K | 0.000694444 | 945 | 958 | Amidation |
| P01027\|CO3_MOUSE | D.TDSETRIILQG.S | 0.000694444 | 974 | 984 | |
| P01027\|CO3_MOUSE | P.DTDSETRIILQGSPV.V | 0.000694444 | 973 | 987 | |
| P01027\|CO3_MOUSE | Q.VPDTDSETRIILQGSPVVQM(+15.99).A | 0.000694444 | 971 | 990 | Oxidation (M) |
| P01027\|CO3_MOUSE | S.EETKQNEAFSLTAK.G | 0.000694444 | 1322 | 1335 | |
| P01027\|CO3_MOUSE | R.VM(+15.99)QDFFIDLRLPYSVV.R | 1.00:0:0 | 825 | 840 | Oxidation (M) |
| P01027\|CO3_MOUSE | R.EPGQDLVVLSLPITPEFIPSF.R | 1.00:0:0 | 509 | 529 | |
| P11859\|ANGT_MOUSE | A.IYEQDSGTLHFLG.R | 1.00:0.07:0 | 456 | 468 | |
| P11859\|ANGT_MOUSE | R.GSYNLQDLLAEDKLPT.L | 1.00:0.81:0 | 379 | 394 | |
| P11859\|ANGT_MOUSE | R.GFSQLPGVHEFWV.D | 1.00:0:0 | 281 | 293 | |
| P16014\|SCG1_MOUSE | R.NHPDSELEST.A | 1.00:4.53:0 | 386 | 395 | |
| P16014\|SCG1_MOUSE | R.ADQRVLTAEEKKELENLA.A | 0.000694444 | 643 | 660 | |
| P16014\|SCG1_MOUSE | R.SEESQEREY.K | 0:1.00:0.03 | 375 | 383 | |
| P16014\|SCG1_MOUSE | R.GREPGAHSALDTREE.K | 0.000694444 | 421 | 435 | |
| P16014\|SCG1_MOUSE | R.RPFSEDVNWGYE.K | 1.00:1.88:0 | 574 | 585 | |
| P16014\|SCG1_MOUSE | K.IYHSEER.V | 1.00:0.68:0 | 152 | 158 | |
| P16014\|SCG1_MOUSE | R.GREPGAHSALDTR.E | 0.000694444 | 421 | 433 | |
| P16014\|SCG1_MOUSE | R.GNPDDSFLEDEGEDRNGVTLT.E | 0.000694444 | 538 | 558 | |
| P16014\|SCG1_MOUSE | A.EEKKELENLA.A | 0.000694444 | 651 | 660 | |
| P16014\|SCG1_MOUSE | R.NHPDSELESTANRHGEETEEERSYEGANGR.Q | 0.000694444 | 386 | 415 | |
| P16014\|SCG1_MOUSE | R.LGALFNPYFDPLQWK.N | 1.00:17.90:0 | 516 | 530 | |
| P16014\|SCG1_MOUSE | R.LLDEGHYPVRESPIDTAK.R | 0.000694444 | 438 | 455 | |
| P16014\|SCG1_MOUSE | R.HGEETEEER.S | 0.000694444 | 399 | 407 | |
| P16014\|SCG1_MOUSE | K.SSYEGHPLSEERRPSPK.E | 0.000694444 | 300 | 316 | |
| P16014\|SCG1_MOUSE | R.VLTAEEKKELENLA.A | 0.000694444 | 647 | 660 | |
| P16014\|SCG1_MOUSE | R.QYDGVAELDQLLHY.R | 0.000694444 | 600 | 613 | |
| P16014\|SCG1_MOUSE | R.GNPDDSFLEDEGEDRN.G | 0.000694444 | 538 | 553 | |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 27 | 50 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSK.L | 1.00:383.50:0 | 8 | 23 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPSK.E | 1.00:30.97:0 | 27 | 38 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEK.F | 1.00:1.51:0.03 | 8 | 18 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPSKETIEQEKQA.G | 1.00:383.68:0 | 27 | 47 | |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 26 | 50 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 8 | 50 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | K.NPLPSKETIEQEKQAGES | 0.000694444 | 33 | 50 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKS.K | 1.00:2.41:0 | 8 | 22 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPS.K | 1.00:0.66:0.00 | 27 | 37 | |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSKETIEQEKQA.G | 0.000694444 | 26 | 47 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSKL.K | 1.00:22.29:0 | 8 | 24 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDMAEIEKFDK.S | 1.00:3.03:0.07 | 8 | 21 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | K.NPLPSK.E | 1.00:25.80:0 | 33 | 38 | |
| P20065\|TYB4_MOUSE | K.NPLPSKETIEQEKQA.G | 1.00:342.35:0 | 33 | 47 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99).A | 1.00:8.83:0 | 8 | 13 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPS.K | 1.00:0.52:0 | 26 | 37 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDMAEIEK.F | 1.00:0.76:0.39 | 8 | 18 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | E.TQEKNPLPSKETIEQEKQA.G | 0.000694444 | 29 | 47 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFD.K | 1.00:2.04:0 | 8 | 20 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | E.TQEKNPLPSKETIEQEKQAGES | 0.000694444 | 29 | 50 | |
| P20065\|TYB4_MOUSE | S.DKPDM(+15.99)AEIEKFDKSKLK.K | 0.000694444 | 9 | 25 | Oxidation (M) |
| P20065\|TYB4_MOUSE | M.AEIEKFDK.S | 1.00:3.67:0 | 14 | 21 | |
| P20065\|TYB4_MOUSE | K.LKKTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 24 | 50 | |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPSKETIEQEK.Q | 1.00:151.22:0 | 27 | 45 | |
| P20065\|TYB4_MOUSE | S.KETIEQEK.Q | 1.00:0.96:0 | 38 | 45 | |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPSKETIE.Q | 0.000694444 | 27 | 42 | |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSK.E | 1.00:41.48:0 | 26 | 38 | |
| P20065\|TYB4_MOUSE | K.E(-18.01)TIEQEK.Q | 1.00:2.01:0.02 | 39 | 45 | Pyro-glu from E |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQA.G | 0.000694444 | 8 | 47 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | A.EIEKFDKSK.L | 1.00:72.19:0 | 15 | 23 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSKLKK.T | 0.000694444 | 8 | 26 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | S.DKPDM(+15.99)AEIEKFDKSK.L | 0.000694444 | 9 | 23 | Oxidation (M) |
| P20065\|TYB4_MOUSE | L.PSKETIEQEKQAGES | 1.00:307.02:0 | 36 | 50 | |
| P20065\|TYB4_MOUSE | E.TQEKNPLPSKETIE.Q | 0.000694444 | 29 | 42 | |
| P20065\|TYB4_MOUSE | S.KETIEQEKQAGES | 1.00:383.53:0 | 38 | 50 | |
| P20065\|TYB4_MOUSE | M.AEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 14 | 50 | |
| P20065\|TYB4_MOUSE | K.LKKTETQEKNPLPSKETIEQEKQA.G | 0.000694444 | 24 | 47 | |
| P20065\|TYB4_MOUSE | E.TQEKNPLPSK.E | 1.00:5.95:0 | 29 | 38 | |
| P20065\|TYB4_MOUSE | K.ETIEQEKQ.A | 1.00:0.26:0.20 | 39 | 46 | |
| P20065\|TYB4_MOUSE | E.KNPLPSKETIEQEKQAGES | 0.000694444 | 32 | 50 | |
| P20065\|TYB4_MOUSE | D.M(+15.99)AEIEKFDK.S | 1.00:2.22:0 | 13 | 21 | Oxidation (M) |
| P20065\|TYB4_MOUSE | Q.EKNPLPS.K | 1.00:0.00:0 | 31 | 37 | |
| P20065\|TYB4_MOUSE | K.LKKTETQEK.N | 0.000694444 | 24 | 32 | |
| P20065\|TYB4_MOUSE | M.AEIEKFDKSK.L | 0.000694444 | 14 | 23 | |
| P20065\|TYB4_MOUSE | K.SKLKKTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 22 | 50 | |
| P20065\|TYB4_MOUSE | T.QEKNPLPSKETIEQEKQAGES | 0.000694444 | 30 | 50 | |
| P20065\|TYB4_MOUSE | D.M(+15.99)AEIEKFDKSK.L | 0.000694444 | 13 | 23 | Oxidation (M) |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSKETIEQEK.Q | 0.000694444 | 26 | 45 | |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPSKETI.E | 0.000694444 | 27 | 41 | |

| | | | | | |
|---|---|---|---|---|---|
| P20065\|TYB4_MOUSE | T.QEKNPLPSKETIEQEKQA.G | 0.000694444 | 30 | 47 | |
| P20065\|TYB4_MOUSE | E.IEKFDK.S | 1.00:1.90:0 | 16 | 21 | |
| P20065\|TYB4_MOUSE | A.EIEKFDKSKL.K | 1.00:11.52:0 | 15 | 24 | |
| P20065\|TYB4_MOUSE | K.NPLPSKETIE.Q | 1.00:13.51:0 | 33 | 42 | |
| P20065\|TYB4_MOUSE | P.A(+42.01)TMSDKPDMAEIEKFDKSKLK.K | 0.000694444 | 5 | 25 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | D.M(+15.99)AEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 13 | 50 | Oxidation (M) |
| P20065\|TYB4_MOUSE | S.KETIEQEKQA.G | 1.00:6.62:0 | 38 | 47 | |
| P20065\|TYB4_MOUSE | L.PATMSDKPDM(+15.99)AEIEKFDKSKLK.K | 0.000694444 | 4 | 25 | Oxidation (M) |
| P20065\|TYB4_MOUSE | L.PSKETIEQEKQA.G | 0.000694444 | 36 | 47 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDMAEIEKFDKS.K | 1.00:0.17:0 | 8 | 22 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | P.ATMSDKPDM(+15.99)AEIEKFDKSK.L | 0.000694444 | 5 | 23 | Oxidation (M) |
| P20065\|TYB4_MOUSE | K.LKKTETQEKNPLPS.K | 0.000694444 | 24 | 37 | |
| P20065\|TYB4_MOUSE | P.ATMSDKPDM(+15.99)AEIEKFDKSKLKK.T | 0.000694444 | 5 | 26 | Oxidation (M) |
| P20065\|TYB4_MOUSE | E.TQEKNPLPS.K | 1.00:2.12:0 | 29 | 37 | |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSKETIEQEKQAG.E | 0.000694444 | 26 | 48 | |
| P20065\|TYB4_MOUSE | T.Q(-17.03)EKNPLPSKETIEQEKQA.G | 0.000694444 | 30 | 47 | Pyro-glu from Q |
| P20065\|TYB4_MOUSE | K.LKKTETQEKNPLPSK.E | 0.000694444 | 24 | 38 | |
| P20065\|TYB4_MOUSE | M.AEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQA.G | 0.000694444 | 14 | 47 | |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSKETIE.Q | 0.000694444 | 26 | 42 | |
| P20065\|TYB4_MOUSE | T.QEKNPLPSK.E | 1.00:126.03:0 | 30 | 38 | |
| P20065\|TYB4_MOUSE | L.PATMSDKPDM(+15.99)AEIEKFDKSK.L | 0.000694444 | 4 | 23 | Oxidation (M) |
| P20065\|TYB4_MOUSE | K.ETIEQEKQAG.E | 1.00:1.21:0 | 39 | 48 | |
| P20065\|TYB4_MOUSE | T.Q(-17.03)EKNPLPSKETIE.Q | 0.000694444 | 30 | 42 | Pyro-glu from Q |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPS(+79.96)KETIEQEKQAGES | 0.000694444 | 26 | 50 | Sulfation |
| P20065\|TYB4_MOUSE | K.KT(+79.96)ETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 26 | 50 | Sulfation |
| P20065\|TYB4_MOUSE | K.NPLPSKET.I | 1.00:0.38:0 | 33 | 40 | |
| P20065\|TYB4_MOUSE | A.E(-18.01)IEKFDKSK.L | 0.000694444 | 15 | 23 | Pyro-glu from E |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKF.D | 0.000694444 | 8 | 19 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | E.TQEKNPLPSKETIEQEK.Q | 0.000694444 | 29 | 45 | |
| P20065\|TYB4_MOUSE | K.LKKTETQEKNPLPSKETIE.Q | 0.000694444 | 24 | 42 | |
| P20065\|TYB4_MOUSE | K.FDKSKL.K | 1.00:0.78:0 | 19 | 24 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDMAEIEKFDKSK.L | 0.000694444 | 8 | 23 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | K.ETIEQEKQAGES | 1.00:383.22:0 | 39 | 50 | |
| P20065\|TYB4_MOUSE | A.EIEKFD.K | 1.00:0:0 | 15 | 20 | |
| P20065\|TYB4_MOUSE | K.SKLKKTETQEKNPLPSKETIEQEKQA.G | 0.000694444 | 22 | 47 | |
| P20065\|TYB4_MOUSE | Q.EKNPLPSK.E | 1.00:0.34:0 | 31 | 38 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)A.E | 1.00:0.44:0.02 | 8 | 14 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSKLKKTETQEKNPLPSKETIE.Q | 0.000694444 | 8 | 42 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | D.M(+15.99)AEIEKFDKSKLK.K | 0.000694444 | 13 | 25 | Oxidation (M) |
| P20065\|TYB4_MOUSE | M.AEIEKFDKSKLK.K | 0.000694444 | 14 | 25 | |
| P20065\|TYB4_MOUSE | L.KKTETQEKNPLPSKETIEQEKQA.G | 0.000694444 | 25 | 47 | |
| P20065\|TYB4_MOUSE | T.ETQEKNPLPSK.E | 0.000694444 | 28 | 38 | |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSKETIEQEKQ.A | 0.000694444 | 26 | 46 | |
| P20065\|TYB4_MOUSE | E.KNPLPSK.E | 0.000694444 | 32 | 38 | |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSKETI.E | 0.000694444 | 26 | 41 | |
| P20065\|TYB4_MOUSE | M.SDKPDM(+15.99)AEIEKFDKSKLK.K | 0.000694444 | 8 | 25 | Oxidation (M) |
| P20065\|TYB4_MOUSE | P.SKETIEQEK.Q | 1.00:0:0 | 37 | 45 | |
| P20065\|TYB4_MOUSE | I.EKFDK.S | 1.00:0.95:0.09 | 17 | 21 | |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSKET.I | 0.000694444 | 26 | 40 | |
| P20065\|TYB4_MOUSE | T.Q(-17.03)EKNPLPSKETIEQEKQAGES | 0.000694444 | 30 | 50 | Pyro-glu from Q |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPSKETIEQE.K | 0.000694444 | 27 | 44 | |
| P20065\|TYB4_MOUSE | L.KKTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 25 | 50 | |
| P20065\|TYB4_MOUSE | K.KT(+79.96)ETQEKNPLPSKETIEQEKQA.G | 0.000694444 | 26 | 47 | Sulfation |
| P20065\|TYB4_MOUSE | A.EIEKFDKS.K | 1.00:1.53:0 | 15 | 22 | |
| P20065\|TYB4_MOUSE | S.DKPDM(+15.99)AEIEKFDKSKL.K | 0.000694444 | 9 | 24 | Oxidation (M) |
| P20065\|TYB4_MOUSE | L.KKTETQEKNPLPSKETIE.Q | 0.000694444 | 25 | 42 | |
| P20065\|TYB4_MOUSE | E.KNPLPSKETIEQEKQA.G | 0.000694444 | 32 | 47 | |
| P20065\|TYB4_MOUSE | S.DKPDM(+15.99)AEIEKFDKSKLKK.T | 0.000694444 | 9 | 26 | Oxidation (M) |
| P20065\|TYB4_MOUSE | M.AEIEKFDKSKL.K | 0.000694444 | 14 | 24 | |
| P20065\|TYB4_MOUSE | Q.EKNPLPSKETIEQEKQAGES | 0.000694444 | 31 | 50 | |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSKETIEQEKQAGES(-.98) | 0.000694444 | 26 | 50 | Amidation |
| P20065\|TYB4_MOUSE | T.ETQEKNPLPSKETIEQEKQA.G | 0.000694444 | 28 | 47 | |
| P20065\|TYB4_MOUSE | D.M(+15.99)AEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQA.G | 0.000694444 | 13 | 47 | Oxidation (M) |
| P20065\|TYB4_MOUSE | P.ATM(+15.99)SDKPDM(+15.99)AEIEKFDKSK.L | 0.000694444 | 5 | 23 | Oxidation (M) |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPSKET.I | 0.000694444 | 27 | 40 | |
| P20065\|TYB4_MOUSE | E.TQEKNPLPSKETIEQE.K | 1.00:50.01:0 | 29 | 44 | |
| P20065\|TYB4_MOUSE | M.AEIEKFDKSKLKKTETQEKNPLPSKETIE.Q | 0.000694444 | 14 | 42 | |
| P20065\|TYB4_MOUSE | M.A(+42.01)EIEKFDK.S | 1.00:1.68:0 | 14 | 21 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | A.EIEKFDKSKLK.K | 0.000694444 | 15 | 25 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AE.I | 1.00:0.62:0 | 8 | 15 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPSKETIEQEKQAG.E | 0.000694444 | 27 | 48 | |
| P20065\|TYB4_MOUSE | Q.EKNPLPSKETIEQEKQA.G | 0.000694444 | 31 | 47 | |
| P20065\|TYB4_MOUSE | L.PSKETIEQEK.Q | 0.000694444 | 36 | 45 | |
| P20065\|TYB4_MOUSE | A.E(-18.01)IEKFDKSKLK.K | 0.000694444 | 15 | 25 | Pyro-glu from E |
| P20065\|TYB4_MOUSE | L.P(+42.01)ATM(+15.99)SDKPDM(+15.99)AEIEKFDKSKLK.K | 0.000694444 | 4 | 25 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | T.QEKNPLPS.K | 1.00:0.12:0 | 30 | 37 | |
| P20065\|TYB4_MOUSE | K.TETQEKNPLP.S | 1.00:0:0 | 27 | 36 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSKLKKTETQEK.N | 0.000694444 | 8 | 32 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | A.E(-18.01)IEKFDK.S | 1.00:0.72:0 | 15 | 21 | Pyro-glu from E |
| P20065\|TYB4_MOUSE | K.SKLKKTETQEKNPLPSKETIE.Q | 0.000694444 | 22 | 42 | |
| P20065\|TYB4_MOUSE | A.T(+42.01)M(+15.99)SDKPDM(+15.99)AEIEKFDKSKLK.K | 0.000694444 | 6 | 25 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDMAEIEKFD.K | 1.00:0:0 | 8 | 20 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | M.A(+42.01)EIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 14 | 50 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDMA.E | 1.00:0.34:2.20 | 8 | 14 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | N.PLPSKETIEQEKQAGES | 0.000694444 | 34 | 50 | |
| P20065\|TYB4_MOUSE | P.SKETIEQEKQA.G | 1.00:0.37:0 | 37 | 47 | |

| | | | | | |
|---|---|---|---|---|---|
| P20065\|TYB4_MOUSE | A.EIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 15 | 50 | |
| P20065\|TYB4_MOUSE | T.ETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 28 | 50 | |
| P20065\|TYB4_MOUSE | E.KNPLPSKETIEQEK.Q | 0.000694444 | 32 | 45 | |
| P20065\|TYB4_MOUSE | D.M(+15.99)AEIEKFDKSKLKKTETQEKNPLPSKETIE.Q | 0.000694444 | 13 | 42 | Oxidation (M) |
| P20065\|TYB4_MOUSE | K.LKKTETQEKNPLPSKETIEQEK.Q | 0.000694444 | 24 | 45 | |
| P20065\|TYB4_MOUSE | D.M(+15.99)AEIEKFDKS.K | 1.00:4.37:0 | 13 | 22 | Oxidation (M) |
| P20065\|TYB4_MOUSE | L.KKTETQEKNPLPS.K | 0.000694444 | 25 | 37 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDMAEIEKFDKSKL.K | 1.00:1.02:0 | 8 | 24 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | Q.E(-18.01)KNPLPS.K | 1.00:0:0 | 31 | 37 | Pyro-glu from E |
| P20065\|TYB4_MOUSE | K.PDMAEIEKFDKSKLKK(-.98).T | 0.000694444 | 11 | 26 | Amidation |
| P20065\|TYB4_MOUSE | I.E(-18.01)KFDK.S | 1.00:0.23:0 | 17 | 21 | Pyro-glu from E |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDMAEIEKFDKSKLK.K | 0.000694444 | 8 | 25 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | L.PSKETIE.Q | 0.000694444 | 36 | 42 | |
| P20065\|TYB4_MOUSE | D.M(+15.99)AEIEKFDKSKLKK.T | 0.000694444 | 13 | 26 | Oxidation (M) |
| P20065\|TYB4_MOUSE | K.FDKSKLK.K | 0.000694444 | 19 | 25 | |
| P20065\|TYB4_MOUSE | L.KKTETQEKNPLPSK.E | 0.000694444 | 25 | 38 | |
| P20065\|TYB4_MOUSE | E.IEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 16 | 50 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSKLKKTETQEKNPLPS.K | 0.000694444 | 8 | 37 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | M.A(+42.01)EIEKFDKSK.L | 0.000694444 | 14 | 23 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKS(+79.96)KLK.K | 0.000694444 | 8 | 25 | Acetylation (N-term); Oxidation (M); Sulfation |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLPSKETIEQEKQAGE.S | 0.000694444 | 26 | 49 | |
| P20065\|TYB4_MOUSE | A.EIEKFDKSKLKK.T | 0.000694444 | 15 | 26 | |
| P20065\|TYB4_MOUSE | K.TETQEKNPLPSKETIEQEKQ.A | 0.000694444 | 27 | 46 | |
| P20065\|TYB4_MOUSE | Q.EKNPLPSKETIE.Q | 0.000694444 | 31 | 42 | |
| P20065\|TYB4_MOUSE | D.MAEIEKFDK.S | 1.00:0:0 | 13 | 21 | |
| P20065\|TYB4_MOUSE | E.TQEKNPLPSKETIEQEKQAG.E | 0.000694444 | 29 | 48 | |
| P20065\|TYB4_MOUSE | A.EIEKFDK.S | 1.00:0.95:0.01 | 15 | 21 | |
| P20065\|TYB4_MOUSE | M.SDKPDM(+15.99)AEIEKFDKSK.L | 0.000694444 | 8 | 23 | Oxidation (M) |
| P20065\|TYB4_MOUSE | K.KTETQEKNPLP.S | 1.00:0:0 | 26 | 36 | |
| P20065\|TYB4_MOUSE | E.IEKFDKSKL.K | 0.000694444 | 16 | 24 | |
| P20065\|TYB4_MOUSE | A.E(-18.01)IEKFDKSKL.K | 0.000694444 | 15 | 24 | Pyro-glu from E |
| P20065\|TYB4_MOUSE | L.PSKETIEQEKQAG(-.98).E | 0.000694444 | 36 | 48 | Amidation |
| P20065\|TYB4_MOUSE | E.KNPLPSKETIE.Q | 0.000694444 | 32 | 42 | |
| P20065\|TYB4_MOUSE | I.EKFDKSKLK.K | 0.000694444 | 17 | 25 | |
| P20065\|TYB4_MOUSE | I.EKFDKSK.L | 1.00:4.43:0 | 17 | 23 | |
| P20065\|TYB4_MOUSE | D.M(+15.99)(+42.01)AEIEKFDKSKLKKTETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 13 | 50 | Oxidation (M); Acetylation (N-term) |
| P20065\|TYB4_MOUSE | K.ETIEQEKQA.G | 1.00:1.27:0 | 39 | 47 | |
| P20065\|TYB4_MOUSE | K.KTETQEK.N | 1.00:10.46:0 | 26 | 32 | |
| P20065\|TYB4_MOUSE | L.PATM(+15.99)SDKPDM(+15.99)AEIEKFDKSKLK.K | 0.000694444 | 4 | 25 | Oxidation (M) |
| P20065\|TYB4_MOUSE | L.P(+42.01)ATM(+15.99)SDKPDM(+15.99)AEIEKFDKSKL.K | 0.000694444 | 4 | 24 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | E.TIEQEKQAGES | 0.000694444 | 40 | 50 | |
| P20065\|TYB4_MOUSE | N.PLPSKETIEQEKQA.G | 0.000694444 | 34 | 47 | |
| P20065\|TYB4_MOUSE | L.PATMSDKPDM(+15.99)AEIEKFDKSKL.K | 0.000694444 | 4 | 24 | Oxidation (M) |
| P20065\|TYB4_MOUSE | M.AEIEKFDKSKLKK.T | 0.000694444 | 14 | 26 | |
| P20065\|TYB4_MOUSE | K.K(+42.01)TETQEKNPLPSKETIEQEKQAGES | 0.000694444 | 26 | 50 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | M.AEIEKFDKS.K | 0.000694444 | 14 | 22 | |
| P20065\|TYB4_MOUSE | E.IEKFDKSKLK.K | 0.000694444 | 16 | 25 | |
| P20065\|TYB4_MOUSE | L.PATM(+15.99)SDKPDM(+15.99)AEIEKFDKSK.L | 0.000694444 | 4 | 23 | Oxidation (M) |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSKLKKT.E | 0.000694444 | 8 | 27 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDK.S | 1.00:3.32:0.01 | 8 | 21 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSKLK.K | 0.000694444 | 8 | 25 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | S.DKPDM(+15.99)AEIEKFDKSKLKKTETQEK.N | 0.000694444 | 9 | 32 | Oxidation (M) |
| P20065\|TYB4_MOUSE | E.IEKFDKSK.L | 0.000694444 | 16 | 23 | |
| P20065\|TYB4_MOUSE | K.ETIEQEK.Q | 1.00:2.83:0 | 39 | 45 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKSKLKKTET.Q | 0.000694444 | 8 | 29 | Acetylation (N-term); Oxidation (M) |
| P20065\|TYB4_MOUSE | K.PDM(+15.99)AEIEKFDKSKL.K | 0.000694444 | 11 | 24 | Oxidation (M) |
| P20065\|TYB4_MOUSE | E.IEKFDKSKLKK.T | 0.000694444 | 16 | 26 | |
| P20065\|TYB4_MOUSE | M.S(+42.01)DKPDM(+15.99)AEIEKFDKS(+79.96)K.L | 0.000694444 | 8 | 23 | Acetylation (N-term); Oxidation (M); Sulfation |
| P20065\|TYB4_MOUSE | D.KPDM(+15.99)AEIEKFDKSKLK.K | 0.000694444 | 10 | 25 | Oxidation (M) |
| P20065\|TYB4_MOUSE | M.SDKPDM(+15.99)AEIEKFDKSKLK(-.98).K | 0.000694444 | 8 | 25 | Oxidation (M); Amidation |
| P20065\|TYB4_MOUSE | M.A(+42.01)EIEKFDKSKLK.K | 0.000694444 | 14 | 25 | Acetylation (N-term) |
| P20065\|TYB4_MOUSE | M.SDKPDM(+15.99)AEIEKFDKSKLKK.T | 0.000694444 | 8 | 26 | Oxidation (M) |
| P20065\|TYB4_MOUSE | K.SKLKKTETQEK.N | 0.000694444 | 22 | 32 | |
| P20918\|PLMN_MOUSE | R.HSIFTPQTNPR.A | 1.00:0.38:0 | 513 | 523 | |
| P20918\|PLMN_MOUSE | G.DSGGPL.V | 1.00:0:0.29 | 761 | 766 | |
| P20918\|PLMN_MOUSE | R.DVILFEK.R | 1.00:0.79:0 | 90 | 96 | |
| P20918\|PLMN_MOUSE | A.SLFSLTK.K | 1.00:63.80:0 | 32 | 38 | |
| P20918\|PLMN_MOUSE | K.TSSIIR.M | 1.00:2.34:0 | 82 | 87 | |
| P20918\|PLMN_MOUSE | R.M(+15.99)RDVILFEK.R | 1.00:1.59:0 | 88 | 96 | Oxidation (M) |
| P20918\|PLMN_MOUSE | R.FVDWIER.E | 1.00:0.73:0 | 801 | 807 | |
| P20918\|PLMN_MOUSE | R.PNKPGVYV.R | 1.00:4.87:0 | 789 | 796 | |
| P20918\|PLMN_MOUSE | K.TPENFPDAGLEM(+15.99).N | 1.00:0:0 | 412 | 423 | Oxidation (M) |
| P20918\|PLMN_MOUSE | R.GTSSTTITGK.K | 1.00:0:0 | 387 | 396 | |
| P20918\|PLMN_MOUSE | K.VILGAHEEYIR.G | 0.000694444 | 637 | 647 | |
| P20918\|PLMN_MOUSE | K.WGATFPHVPNY.S | 1.00:0:0.37 | 127 | 137 | |
| P20918\|PLMN_MOUSE | R.SFQYHS.K | 1.00:0.50:0.53 | 63 | 68 | |
| P26339\|CMGA_MOUSE | V.FENQSPDAK.H | 1.00:0.13:0.03 | 128 | 136 | |
| P26339\|CMGA_MOUSE | R.AQQPLK.Q | 0.000694444 | 91 | 96 | |
| P26339\|CMGA_MOUSE | V.FENQSPDAKHR.D | 0.000694444 | 128 | 138 | |
| P26339\|CMGA_MOUSE | S.AIEAELEKV.A | 0.000694444 | 445 | 453 | |
| P26339\|CMGA_MOUSE | K.AIQKDDGQSDSQ.A | 0.000694444 | 272 | 283 | |
| P26350\|PTMA_MOUSE | V.AEDDEDDDVDTKKQKTEEDD | 1.00:383.74:0 | 92 | 111 | |
| P26350\|PTMA_MOUSE | K.EVVEEAENGR.D | 1.00:8.22:0 | 22 | 31 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLK.E | 0.000694444 | 2 | 18 | Acetylation (N-term) |

| | | | | | |
|---|---|---|---|---|---|
| P26350\|PTMA_MOUSE | K.RVAEDDEDDDVDTKKQKTEEDD | 0.000694444 | 90 | 111 | |
| P26350\|PTMA_MOUSE | E.EEDGDEDEEAEAPTGKRV.A | 1.00:25.05:0 | 74 | 91 | |
| P26350\|PTMA_MOUSE | K.EVVEEAENGRDAPANGNAQNEENGEQEA.D | 1.00:6.56:0 | 22 | 49 | |
| P26350\|PTMA_MOUSE | R.DAPANGNAQNEENGEQEA.D | 1.00:2.30:0.02 | 32 | 49 | |
| P26350\|PTMA_MOUSE | S.SEITTK.D | 0.000694444 | 10 | 15 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEK.K | 0.000694444 | 2 | 20 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | R.VAEDDEDDDVDTKKQKTEEDD | 0.000694444 | 91 | 111 | |
| P26350\|PTMA_MOUSE | K.RVAEDDEDDDVD.T | 1.00:8.58:0 | 90 | 101 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKKEVVE.E | 0.000694444 | 2 | 25 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | A.VDTSSEITTKDLKEKKEVVEEAENGRDAPANGNAQNEENGEQEA<br>.D | 0.000694444 | 6 | 49 | |
| P26350\|PTMA_MOUSE | V.AEDDEDDDVD.T | 0.000694444 | 92 | 101 | |
| P26350\|PTMA_MOUSE | K.KEVVEEAENG.R | 1.00:0.04:0 | 21 | 30 | |
| P26350\|PTMA_MOUSE | V.AEDDEDDDVDTK.K | 1.00:5.83:0 | 92 | 103 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKKEVVEEAENGRDAPANGNA<br>QNEENGEQEA.D | 0.000694444 | 2 | 49 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | D.AAVDTSSEITTK.D | 1.00:116.40:0 | 4 | 15 | |
| P26350\|PTMA_MOUSE | V.AEDDEDDDVDTKK.Q | 1.00:11.23:0 | 92 | 104 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKK.E | 1.00:26.83:0 | 2 | 21 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | E.DEEAEAPTGK.R | 1.00:3.30:0.07 | 80 | 89 | |
| P26350\|PTMA_MOUSE | A.VDTSSEITTK.D | 0.000694444 | 6 | 15 | |
| P26350\|PTMA_MOUSE | K.EVVEEAENGRDAPA.N | 1.00:2.81:0 | 22 | 35 | |
| P26350\|PTMA_MOUSE | K.KEVVEEAENGRDAPANGNAQNEENGEQEA.D | 0.000694444 | 21 | 49 | |
| P26350\|PTMA_MOUSE | K.EKKEVVEEAENGRDAPANGNAQNEENGEQEA.D | 0.000694444 | 19 | 49 | |
| P26350\|PTMA_MOUSE | T.SSEITTK.D | 1.00:5.46:0 | 9 | 15 | |
| P26350\|PTMA_MOUSE | K.EVVEEAENG.R | 1.00:0:0.10 | 22 | 30 | |
| P26350\|PTMA_MOUSE | E.EEDGDEDEEAEAPTGK.R | 0.000694444 | 74 | 89 | |
| P26350\|PTMA_MOUSE | E.DEEAEAPTGKRV.A | 1.00:19.33:0 | 80 | 91 | |
| P26350\|PTMA_MOUSE | V.DTSSEITTK.D | 0.000694444 | 7 | 15 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKKEV.V | 0.000694444 | 2 | 23 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | K.EKKEVVE.E | 0.000694444 | 19 | 25 | |
| P26350\|PTMA_MOUSE | S.EITTKDLKEKKEVVE.E | 0.000694444 | 11 | 25 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKKEVVEEAENGR.D | 0.000694444 | 2 | 31 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | E.DDDVDTKKQKTEEDD | 0.000694444 | 97 | 111 | |
| P26350\|PTMA_MOUSE | K.KEVVEEAENGR.D | 1.00:25.41:0 | 21 | 31 | |
| P26350\|PTMA_MOUSE | V.AEDDEDDDVDTKKQK.T | 1.00:45.08:0 | 92 | 106 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKKEVVEEAENGRDAPA.N | 0.000694444 | 2 | 35 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | E.DEEAEAPT.G | 1.00:0.26:0 | 80 | 87 | |
| P26350\|PTMA_MOUSE | E.DDDVDTKKQKTEE.D | 0.000694444 | 97 | 109 | |
| P26350\|PTMA_MOUSE | T.TKDLKEKKEVVE.E | 0.000694444 | 14 | 25 | |
| P26350\|PTMA_MOUSE | A.VDTSSEITTKDLK.E | 0.000694444 | 6 | 18 | |
| P26350\|PTMA_MOUSE | S.DAAVDTSSEITTKDLKEKKEVVE.E | 0.000694444 | 3 | 25 | |
| P26350\|PTMA_MOUSE | K.DLKEKKEVVE.E | 0.000694444 | 16 | 25 | |
| P26350\|PTMA_MOUSE | D.DDVDTKKQKTEEDD | 0.000694444 | 98 | 111 | |
| P26350\|PTMA_MOUSE | S.DAAVDTSSEITTK.D | 1.00:2.05:0 | 3 | 15 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKKEVV.E | 0.000694444 | 2 | 24 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | R.VAEDDEDDDVDTK.K | 1.00:0.07:0 | 91 | 103 | |
| P26350\|PTMA_MOUSE | A.EDDEDDDVDTKKQKTEEDD | 0.000694444 | 93 | 111 | |
| P26350\|PTMA_MOUSE | E.DDEDDDVDTKK.Q | 1.00:1.47:0 | 94 | 104 | |
| P26350\|PTMA_MOUSE | K.DLKEKK.E | 0.000694444 | 16 | 21 | |
| P26350\|PTMA_MOUSE | K.DLKEKKEVVEEAENGRDAPANGNAQNEENGEQEA.D | 0.000694444 | 16 | 49 | |
| P26350\|PTMA_MOUSE | V.AEDDEDDDVDTKKQKT.E | 0.000694444 | 92 | 107 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEI.T | 0:1.00:0.07 | 2 | 12 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | D.DDVDTKKQKTEE.D | 0.000694444 | 98 | 109 | |
| P26350\|PTMA_MOUSE | D.AAVDTSSEITTKDLKEKKEVVE.E | 0.000694444 | 4 | 25 | |
| P26350\|PTMA_MOUSE | K.RVAEDDEDDDVDT.K | 1.00:0:0 | 90 | 102 | |
| P26350\|PTMA_MOUSE | E.DDEDDDVDTKKQK.T | 0.000694444 | 94 | 106 | |
| P26350\|PTMA_MOUSE | E.DEEAEAPTGKRVAE.D | 0.000694444 | 80 | 93 | |
| P26350\|PTMA_MOUSE | E.EAEAPTGKRV.A | 0.000694444 | 82 | 91 | |
| P26350\|PTMA_MOUSE | E.EAENGRDAPANGNAQNEENGEQEA.D | 0.000694444 | 26 | 49 | |
| P26350\|PTMA_MOUSE | D.DEDDDVDTK.K | 0.000694444 | 95 | 103 | |
| P26350\|PTMA_MOUSE | E.KKEVVEE.A | 0.000694444 | 20 | 26 | |
| P26350\|PTMA_MOUSE | D.TSSEITTK.D | 1.00:3.72:0 | 8 | 15 | |
| P26350\|PTMA_MOUSE | S.SEITTKDLK.E | 0.000694444 | 10 | 18 | |
| P26350\|PTMA_MOUSE | V.DTSSEITTKDLK.E | 0.000694444 | 7 | 18 | |
| P26350\|PTMA_MOUSE | E.ITTKDLKEKKEVVE.E | 0.000694444 | 12 | 25 | |
| P26350\|PTMA_MOUSE | R.VAEDDEDDDVDTKK.Q | 1.00:0:0.02 | 91 | 104 | |
| P26350\|PTMA_MOUSE | A.NGNAQNEENGEQEA.D | 0.000694444 | 36 | 49 | |
| P26350\|PTMA_MOUSE | A.VDTSSEITTKDLKEKK.E | 0.000694444 | 6 | 21 | |
| P26350\|PTMA_MOUSE | D.DEDDDVDTKK.Q | 0.000694444 | 95 | 104 | |
| P26350\|PTMA_MOUSE | E.AEAPTGKRVAED.D | 0.000694444 | 83 | 94 | |
| P26350\|PTMA_MOUSE | V.AEDDEDDDVDTKKQ.K | 1.00:25.03:0 | 92 | 105 | |
| P26350\|PTMA_MOUSE | E.EEDGDEDEEAEAPTGKR.V | 0.000694444 | 74 | 90 | |
| P26350\|PTMA_MOUSE | E.DEEAEAPTGKRVAED.D | 0.000694444 | 80 | 94 | |
| P26350\|PTMA_MOUSE | T.SSEITTKDLKEKKEVVE.E | 0.000694444 | 9 | 25 | |
| P26350\|PTMA_MOUSE | T.SSEITTKDLK.E | 0.000694444 | 9 | 18 | |
| P26350\|PTMA_MOUSE | E.EEDGDEDEEAEAPTG.K | 0.000694444 | 74 | 88 | |
| P26350\|PTMA_MOUSE | E.DDEDDDVDTKKQKTEEDD | 0.000694444 | 94 | 111 | |
| P26350\|PTMA_MOUSE | D.AAVDTSSEITTKDLK.E | 0.000694444 | 4 | 18 | |
| P26350\|PTMA_MOUSE | V.VEEAENGRDAPA.N | 1.00:0.93:0.22 | 24 | 35 | |
| P26350\|PTMA_MOUSE | E.AEAPTGKRVAE.D | 1.00:1.69:0 | 83 | 93 | |
| P26350\|PTMA_MOUSE | E.DGDEDEEAEAPTGKRV.A | 0.000694444 | 76 | 91 | |
| P26350\|PTMA_MOUSE | R.VAEDDEDDDVDT.K | 1.00:0.14:1.08 | 91 | 102 | |
| P26350\|PTMA_MOUSE | I.TTKDLKEKKEVVE.E | 0.000694444 | 13 | 25 | |
| P26350\|PTMA_MOUSE | A.VDTSSEITTKDLKEKKEV.V | 0.000694444 | 6 | 23 | |
| P26350\|PTMA_MOUSE | D.DEDDDVDTKKQK.T | 0.000694444 | 95 | 106 | |
| P26350\|PTMA_MOUSE | E.DEEAEAPTGKR.V | 1.00:1.80:0 | 80 | 90 | |

| | | | | | |
|---|---|---|---|---|---|
| P26350\|PTMA_MOUSE | R.VAEDDEDDDVDTKKQK.T | 0.000694444 | 91 | 106 | |
| P26350\|PTMA_MOUSE | A.VDTSSEITTKDLKEKKEVVE.E | 0.000694444 | 6 | 25 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKKEVVEEAENGRDAPANGNAQN.E | 0.000694444 | 2 | 41 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | A.VDTSSEITTKDLKEKKEVVEEAENGR.D | 0.000694444 | 6 | 31 | |
| P26350\|PTMA_MOUSE | V.EEAENGRDAPA.N | 0:1.00:0.43 | 25 | 35 | |
| P26350\|PTMA_MOUSE | S.DAAVDTSSEITTKDLK.E | 0.000694444 | 3 | 18 | |
| P26350\|PTMA_MOUSE | A.AVDTSSEITTKDLKEKKEVVE.E | 0.000694444 | 5 | 25 | |
| P26350\|PTMA_MOUSE | D.DVDTKKQKTEE.D | 0.000694444 | 99 | 109 | |
| P26350\|PTMA_MOUSE | A.DNEVDEEEEEGGEEEEEEEEGDGEEEDGDEDEEAEAPTGKRV.A | 0.000694444 | 50 | 91 | |
| P26350\|PTMA_MOUSE | E.DEEAEAPTGKRVAEDDEDDDVD.T | 1.00:0.60:0 | 80 | 101 | |
| P26350\|PTMA_MOUSE | D.DVDTKKQKTEEDD | 0.000694444 | 99 | 111 | |
| P26350\|PTMA_MOUSE | R.VAEDDEDDDVDTKKQ.K | 1.00:0:0 | 91 | 105 | |
| P26350\|PTMA_MOUSE | A.VDTSSEITTKDLKEKKEVVEEAENGRDAPA.N | 0.000694444 | 6 | 35 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKKEVVEEAENGRD.A | 0.000694444 | 2 | 32 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | D.DEDDDVDTKKQKTEEDD | 1.00:7.11:0 | 95 | 111 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKKEVVEEA.E | 0.000694444 | 2 | 27 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDLKEKKEVVEE.A | 0.000694444 | 2 | 26 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | V.AEDDEDDDVDTKKQKTEED.D | 0.000694444 | 92 | 110 | |
| P26350\|PTMA_MOUSE | D.TSSEITTKDLK.E | 0.000694444 | 8 | 18 | |
| P26350\|PTMA_MOUSE | K.RVAEDDEDDDVDTK.K | 1.00:0:0 | 90 | 103 | |
| P26350\|PTMA_MOUSE | E.DDEDDDVDTK.K | 1.00:1.00:0 | 94 | 103 | |
| P26350\|PTMA_MOUSE | D.EDDDVDTKKQKTEEDD | 0.000694444 | 96 | 111 | |
| P26350\|PTMA_MOUSE | E.EAEAPTGKRVAE.D | 0.000694444 | 82 | 93 | |
| P26350\|PTMA_MOUSE | D.VDTKKQKTEE.D | 0.000694444 | 100 | 109 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTK.D | 0.000694444 | 2 | 15 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | A.EAPTGKRVAED.D | 0.000694444 | 84 | 94 | |
| P26350\|PTMA_MOUSE | S.DAAVDTSSEITTKDLKEKK.E | 0.000694444 | 3 | 21 | |
| P26350\|PTMA_MOUSE | T.KDLKEKKEVVE.E | 0.000694444 | 15 | 25 | |
| P26350\|PTMA_MOUSE | S.SEITTKDLKEKKEVVE.E | 0.000694444 | 10 | 25 | |
| P26350\|PTMA_MOUSE | N.GRDAPANGNAQN.E | 1.00:0.44:0 | 30 | 41 | |
| P26350\|PTMA_MOUSE | A.ENGRDAPANGNAQN.E | 0.000694444 | 28 | 41 | |
| P26350\|PTMA_MOUSE | V.DTSSEITTKDLKEKKEVVE.E | 0.000694444 | 7 | 25 | |
| P26350\|PTMA_MOUSE | E.NGRDAPANGNAQN.E | 0.000694444 | 29 | 41 | |
| P26350\|PTMA_MOUSE | M.S(+42.01)DAAVDTSSEITTKDL.K | 1.00:383.98:0 | 2 | 17 | Acetylation (N-term) |
| P26350\|PTMA_MOUSE | D.AAVDTSSEITTKDLKEKK.E | 0.000694444 | 4 | 21 | |
| P26350\|PTMA_MOUSE | D.TSSEITTKDLKEKKEVVE.E | 0.000694444 | 8 | 25 | |
| P26350\|PTMA_MOUSE | E.EAEAPTGKRVAED.D | 0.000694444 | 82 | 94 | |
| P26350\|PTMA_MOUSE | V.VEEAENGR.D | 0.000694444 | 24 | 31 | |
| P26350\|PTMA_MOUSE | D.VDTKKQKTEEDD | 0.000694444 | 100 | 111 | |
| P26350\|PTMA_MOUSE | E.AENGRDAPANGNAQN.E | 0.000694444 | 27 | 41 | |
| P26350\|PTMA_MOUSE | E.EAENGRDAPANGNAQN.E | 0.000694444 | 26 | 41 | |
| P26350\|PTMA_MOUSE | L.KEKKEVVE.E | 0.000694444 | 18 | 25 | |
| P28798\|GRN_MOUSE | R.LPDPQILK.S | 0.000694444 | 352 | 359 | |
| P28798\|GRN_MOUSE | R.WDM(+15.99)FLRDPVPRPLL | 0.000694444 | 576 | 589 | Oxidation (M) |
| P28798\|GRN_MOUSE | R.DPVPRPLL | 0.000694444 | 582 | 589 | |
| P28798\|GRN_MOUSE | R.WDM(+15.99)FLRDPVPRPL.L | 0.000694444 | 576 | 588 | Oxidation (M) |
| P31786\|ACBP_MOUSE | V.GDVNTDRPGL.L | 1.00:1.94:0.25 | 38 | 47 | |
| P31786\|ACBP_MOUSE | K.QATVGDVNTDRPGLLDLK.G | 1.00:8.45:0 | 34 | 51 | |
| P31786\|ACBP_MOUSE | V.GDVNTDRPGLL.D | 1.00:1.54:1.34 | 38 | 48 | |
| P31786\|ACBP_MOUSE | K.TYVEKVDELK.K | 1.00:1.94:0 | 73 | 82 | |
| P31786\|ACBP_MOUSE | M.S(+42.01)QAEFDKAAEEVKR.L | 0.000694444 | 2 | 15 | Acetylation (N-term) |
| P31786\|ACBP_MOUSE | M.S(+42.01)QAEFDKAAEEVK.R | 1.00:12.74:0 | 2 | 14 | Acetylation (N-term) |
| P31786\|ACBP_MOUSE | G.DVNTDRPGLL.D | 1.00:0.52:0 | 39 | 48 | |
| P31786\|ACBP_MOUSE | V.GDVNTDRPGLLDLK.G | 1.00:45.87:0 | 38 | 51 | |
| P31786\|ACBP_MOUSE | G.DVNTDRPGLLDLK.G | 0.000694444 | 39 | 51 | |
| P31786\|ACBP_MOUSE | K.TYVEKVDELKK.K | 0.000694444 | 73 | 83 | |
| P31786\|ACBP_MOUSE | K.QATVGDVNTDRPGLLDLKGK.A | 0.000694444 | 34 | 53 | |
| P31786\|ACBP_MOUSE | K.TYVEKVDELKKK.Y | 0.000694444 | 73 | 84 | |
| P31786\|ACBP_MOUSE | G.DVNTDRPGLLDL.K | 1.00:0:0 | 39 | 50 | |
| P31786\|ACBP_MOUSE | K.QATVGDVNTDRPGL.L | 1.00:0:0 | 34 | 47 | |
| P31786\|ACBP_MOUSE | V.GDVNTDRPGLLDL.K | 1.00:0:0.26 | 38 | 50 | |
| P31786\|ACBP_MOUSE | K.TQPTDEEM(+15.99)LFIYSHFK.Q | 1.00:383.98:0 | 18 | 33 | Oxidation (M) |
| P31786\|ACBP_MOUSE | K.QATVGDVNTDRPGLLDL.K | 1.00:0:0 | 34 | 50 | |
| P31786\|ACBP_MOUSE | M.S(+42.01)QAEFDKAAEEVKRLK.T | 0.000694444 | 2 | 17 | Acetylation (N-term) |
| P31786\|ACBP_MOUSE | A.TVGDVNTDRPGLL.D | 0:1.00:1.86 | 36 | 48 | |
| P31786\|ACBP_MOUSE | K.TQPTDEEML.F | 1.15741E-05 | 18 | 26 | |
| P31786\|ACBP_MOUSE | N.TDRPGLLDLKGK.A | 0.000694444 | 42 | 53 | |
| P31786\|ACBP_MOUSE | G.DVNTDRPGLLDLKGK.A | 0.000694444 | 39 | 53 | |
| P31786\|ACBP_MOUSE | V.NTDRPGLLDLKGK.A | 0.000694444 | 41 | 53 | |
| P33680\|GUC2A_MOUSE | H.SALPEALRPV.C | 1.00:8.54:0.88 | 73 | 82 | |
| P33680\|GUC2A_MOUSE | G.VTVQDGDLSFPL.E | 0.000694444 | 22 | 33 | |
| P33680\|GUC2A_MOUSE | G.VTVQDGDLSFPLESVKK.L | 0.000694444 | 22 | 38 | |
| P33680\|GUC2A_MOUSE | S.HSALPEALRPV.C | 0.000694444 | 72 | 82 | |
| P33680\|GUC2A_MOUSE | G.VTVQDGDLSFPLESV.K | 1.00:1.95:1.61 | 22 | 36 | |
| P33680\|GUC2A_MOUSE | S.SHSALPEALRPV.C | 1.00:4.11:0 | 71 | 82 | |
| P33680\|GUC2A_MOUSE | Q.DGDLSFPLESV.K | 1.00:0:0 | 26 | 36 | |
| P33680\|GUC2A_MOUSE | A.EEILQRLE.A | 1.00:1.58:0 | 89 | 96 | |
| P33680\|GUC2A_MOUSE | A.EEILQRLEAIAQ.D | 0.000694444 | 89 | 100 | |
| P33680\|GUC2A_MOUSE | A.EEILQRLEAI.A | 0.000694444 | 89 | 98 | |
| P33680\|GUC2A_MOUSE | Q.DGDLSFPLESVKK.L | 0.000694444 | 26 | 38 | |
| P33680\|GUC2A_MOUSE | G.VTVQDGDLSFPLESVK.K | 0.000694444 | 22 | 37 | |
| P51437\|CRAMP_MOUSE | L.DLDPEPQGDEDPDTPK.S | 1.00:2.04:0.01 | 57 | 72 | |
| P51437\|CRAMP_MOUSE | K.EQGVVK.Q | 1.00:6.18:0 | 98 | 103 | |
| P51437\|CRAMP_MOUSE | L.DPEPQGDEDPDTPK.S | 1.00:60.08:0.35 | 59 | 72 | |
| P51437\|CRAMP_MOUSE | L.DLDPEPQGDEDPDTPKSV.R | 1.00:7.86:0 | 57 | 74 | |
| P51437\|CRAMP_MOUSE | E.DPDTPK.S | 1.00:13.16:0 | 67 | 72 | |

| | | | | | |
|---|---|---|---|---|---|
| P51437\|CRAMP_MOUSE | G.DEDPDTPK.S | 1.00:6.43:0 | 65 | 72 | |
| P51437\|CRAMP_MOUSE | V.DDFNQQSLDTNLYR.L | 1.00:276.51:1.07 | 41 | 54 | |
| P51437\|CRAMP_MOUSE | G.DEDPDTPKSV.R | 1.00:8.23:0 | 65 | 74 | |
| P51437\|CRAMP_MOUSE | L.LDLDPEPQGDEDPDTPKSV.R | 0.000694444 | 56 | 74 | |
| P51437\|CRAMP_MOUSE | R.AVDDFNQQSLDTNL.Y | 1.00:0.16:0 | 39 | 52 | |
| P51437\|CRAMP_MOUSE | R.AVDDFNQQSLDTNLYR.L | 1.00:384.00:0 | 39 | 54 | |
| P51437\|CRAMP_MOUSE | L.DLDPEPQGDEDPDTPKS.V | 1.00:0.77:0 | 57 | 73 | |
| P51437\|CRAMP_MOUSE | L.LDLDPEPQGDEDPDTPK.S | 1.00:5.78:0 | 56 | 72 | |
| P51437\|CRAMP_MOUSE | D.EDPDTPKSV.R | 1.00:16.50:0 | 66 | 74 | |
| P51437\|CRAMP_MOUSE | Q.GDEDPDTPKSV.R | 1.00:3.28:0.06 | 64 | 74 | |
| P51437\|CRAMP_MOUSE | R.LLDLDPEPQG.D | 1.00:3.62:0 | 55 | 64 | |
| P51437\|CRAMP_MOUSE | V.DDFNQQSLDTNL.Y | 1.00:1.55:0 | 41 | 52 | |
| P51437\|CRAMP_MOUSE | L.DPEPQG.D | 1.00:9.90:0.30 | 59 | 64 | |
| P51437\|CRAMP_MOUSE | R.AVDDFNQQSLDTNLY.R | 1.00:14.84:0 | 39 | 53 | |
| P51437\|CRAMP_MOUSE | L.DLDPEPQGDEDPDTPKSVR.F | 0.000694444 | 57 | 75 | |
| P51437\|CRAMP_MOUSE | Q.GDEDPDTPK.S | 1.00:1.39:0.23 | 64 | 72 | |
| P51437\|CRAMP_MOUSE | D.DFNQQSLDTNLYR.L | 1.00:5.39:0 | 42 | 54 | |
| P51437\|CRAMP_MOUSE | V.DDFNQQSLDTN.L | 1.00:0.25:0.05 | 41 | 51 | |
| P51437\|CRAMP_MOUSE | R.LLDLDPEPQGDEDPDTPKS.V | 1.00:0:0 | 55 | 73 | |
| P51437\|CRAMP_MOUSE | D.LDPEPQGDEDPDTPKSV.R | 1.00:0.24:0 | 58 | 74 | |
| P51437\|CRAMP_MOUSE | L.LDLDPEPQGDEDPDTPKS.V | 0.000694444 | 56 | 73 | |
| P51437\|CRAMP_MOUSE | Y.RLLDLDPEPQGDEDPDTPKSV.R | 0.000694444 | 54 | 74 | |
| P51437\|CRAMP_MOUSE | L.DPEPQGDEDPDTP.K | 0.000694444 | 59 | 71 | |
| P51437\|CRAMP_MOUSE | R.LLDLDPEPQGDEDPDTPKSV.R | 0.000694444 | 55 | 74 | |
| P51437\|CRAMP_MOUSE | V.DDFNQQSLDTNLY.R | 0.000694444 | 41 | 53 | |
| P51437\|CRAMP_MOUSE | L.DLDPEPQ.G | 0.000694444 | 57 | 63 | |
| P51437\|CRAMP_MOUSE | A.VDDFNQQSLDT.N | 0.000694444 | 40 | 50 | |
| P51437\|CRAMP_MOUSE | D.LDPEPQGDEDPDTPK.S | 1.00:0.56:0 | 58 | 72 | |
| P51437\|CRAMP_MOUSE | R.AVDDFNQQSLDT.N | 1.00:0:0 | 39 | 50 | |
| P51437\|CRAMP_MOUSE | G.DEDPDTPKSVR.F | 1.00:1.26:0.03 | 65 | 75 | |
| P51437\|CRAMP_MOUSE | F.FQKLVPQPE.Q | 0.000694444 | 164 | 172 | |
| P51437\|CRAMP_MOUSE | V.DDFNQQSLDT.N | 0.000694444 | 41 | 50 | |
| P51437\|CRAMP_MOUSE | V.DDFNQQSLDTNLYRLL.D | 0:1.00:0.02 | 41 | 56 | |
| P51437\|CRAMP_MOUSE | R.LLDLDPEPQGDEDPDTPK.S | 1.00:1.18:0 | 55 | 72 | |
| P51437\|CRAMP_MOUSE | R.AVDDFNQQSLDTNLYRLL.D | 0.000694444 | 39 | 56 | |
| P55095\|GLUC_MOUSE | Q.HALQDTEENPRSFPA.S | 1.00:0.02:0 | 21 | 35 | |
| P55095\|GLUC_MOUSE | K.RHDEF.E | 1.00:8.52:0 | 91 | 95 | |
| P55095\|GLUC_MOUSE | R.HADGSFSDEM(+15.99)ST.I | 1.00:0.12:0 | 146 | 157 | Oxidation (M) |
| P55095\|GLUC_MOUSE | Q.HALQDTEENPR.S | 1.00:0.31:0 | 21 | 31 | |
| P55095\|GLUC_MOUSE | R.HADGSFSDEM(+15.99)STI.L | 0.000694444 | 146 | 158 | Oxidation (M) |
| P55095\|GLUC_MOUSE | T.SDYSKYLDS.R | 0.000694444 | 60 | 68 | |
| P55095\|GLUC_MOUSE | R.DFINWLIQTK.I | 0.000694444 | 166 | 175 | |
| P55095\|GLUC_MOUSE | R.DFINWLIQT.K | 1.00:0:0 | 166 | 174 | |
| P70296\|PEBP1_MOUSE | I.SWDGLDPGKLYT.L | 1.00:0.49:0 | 54 | 65 | |
| P70296\|PEBP1_MOUSE | R.EWHHFLVV.N | 1.00:0.46:0 | 83 | 90 | |
| P70296\|PEBP1_MOUSE | S.DYVGSGPPSGTGLH.R | 1.00:0:0 | 105 | 118 | |
| P70296\|PEBP1_MOUSE | K.GNDISSGTVLSDYVGSGPPSGTGLHR.Y | 0.000694444 | 94 | 119 | |
| P70296\|PEBP1_MOUSE | K.LYEQLSGK | 1.00:1.38:0 | 180 | 187 | |
| P70296\|PEBP1_MOUSE | K.VLTPTQVM(+15.99)NRPSSI.S | 1.00:1.27:0 | 40 | 53 | Oxidation (M) |
| P70296\|PEBP1_MOUSE | W.DDYVPK.L | 1.00:3.30:0 | 174 | 179 | |
| P70296\|PEBP1_MOUSE | I.SWDGLDPGKLYTLV.L | 1.00:1.79:0 | 54 | 67 | |
| P70296\|PEBP1_MOUSE | K.GNDISSGTVLSDYVGSGPPSGTGLH.R | 1.00:0:0 | 94 | 118 | |
| P70296\|PEBP1_MOUSE | S.DYVGSGPPSGTGLHR.Y | 0.000694444 | 105 | 119 | |
| P70296\|PEBP1_MOUSE | S.GPPSGTGLH.R | 1.00:0:0 | 110 | 118 | |
| P70296\|PEBP1_MOUSE | V.DELGKVLTPTQVM(+15.99)NRPSSI.S | 0.000694444 | 35 | 53 | Oxidation (M) |
| P81117\|NUCB2_MOUSE | R.TRLDELKR.Q | 0.000694444 | 101 | 108 | |
| P81117\|NUCB2_MOUSE | T.EDELK.E | 1.00:1.26:0 | 344 | 348 | |
| P81117\|NUCB2_MOUSE | K.LQQGIAPSGPAGELKFEPHT | 0.000694444 | 401 | 420 | |
| P81117\|NUCB2_MOUSE | K.VHNTEPVENA.R | 1.00:1.93:0 | 34 | 43 | |
| P81117\|NUCB2_MOUSE | Q.GIAPSGPAGELKFEPHT | 1.00:4.73:0 | 404 | 420 | |
| P81117\|NUCB2_MOUSE | R.LVTLEEFLR.A | 1.00:78.88:0 | 312 | 320 | |
| P81117\|NUCB2_MOUSE | R.EHVM(+15.99)SEIDNNKDR.L | 0.000694444 | 299 | 311 | Oxidation (M) |
| P81117\|NUCB2_MOUSE | K.LDALQDTGM(+15.99)NHHLLLK.Q | 1.00:6.76:0 | 122 | 137 | Oxidation (M) |
| P81117\|NUCB2_MOUSE | V.IEVLETDPHFR.E | 0.000694444 | 62 | 72 | |
| P81117\|NUCB2_MOUSE | K.LQQGIAPSGPAGELK.F | 0.000694444 | 401 | 415 | |
| P81117\|NUCB2_MOUSE | E.PPDTGLYYDEYLKQVIE.V | 1.00:24.23:0 | 47 | 63 | |
| P81117\|NUCB2_MOUSE | A.VPIDVDKTK.V | 0.000694444 | 25 | 33 | |
| P81117\|NUCB2_MOUSE | Q.GIAPSGPAGELK.F | 0.000694444 | 404 | 415 | |
| P81117\|NUCB2_MOUSE | K.KLQQGIAPSGPAGELKFEPHT | 0.000694444 | 400 | 420 | |
| P81117\|NUCB2_MOUSE | A.RIEPPDTGLYYDEYLKQV.I | 0.000694444 | 44 | 61 | |
| P81117\|NUCB2_MOUSE | V.IEVLETDPHFREKLQK.A | 0.000694444 | 62 | 77 | |
| P81117\|NUCB2_MOUSE | V.IEVLETDPHFREK.L | 0.000694444 | 62 | 74 | |
| P81117\|NUCB2_MOUSE | R.LSQELDLVSHKVR.T | 0.000694444 | 88 | 100 | |
| P81117\|NUCB2_MOUSE | K.ADIEEIRSGR.L | 0.000694444 | 78 | 87 | |
| P81117\|NUCB2_MOUSE | K.LHDVNNDGFLDEQELEALFTR.E | 0.000694444 | 252 | 272 | |
| P81117\|NUCB2_MOUSE | K.QFEHLNHQNPNTFESRDLDM(+15.99)LIK.A | 0.000694444 | 138 | 160 | Oxidation (M) |
| P81117\|NUCB2_MOUSE | V.IEVLETDPHFREKLQKA.D | 0.000694444 | 62 | 78 | |
| P81117\|NUCB2_MOUSE | R.LSQELDLV.S | 0.000694444 | 88 | 95 | |
| P81117\|NUCB2_MOUSE | R.TRLDELK.R | 0.000694444 | 101 | 107 | |
| P81117\|NUCB2_MOUSE | R.LSQELDLVSHKV.R | 0.000694444 | 88 | 99 | |
| Q02819\|NUCB1_MOUSE | K.QEPNLQVDHM(+15.99)NLLK.Q | 0.000694444 | 122 | 135 | Oxidation (M) |
| Q02819\|NUCB1_MOUSE | K.QFEHLDPQNQHTFEA.R | 1.00:0.59:0 | 136 | 150 | |
| Q02819\|NUCB1_MOUSE | K.VNVPGSQAQLK.E | 1.00:7.66:0 | 218 | 228 | |
| Q02819\|NUCB1_MOUSE | R.LSQETEALGR.S | 1.00:25.09:0 | 367 | 376 | |
| Q02819\|NUCB1_MOUSE | K.KVPEQPPELPQLDSQH.L | 1.00:0:0 | 443 | 458 | |
| Q02819\|NUCB1_MOUSE | K.KVPEQPPELPQLDSQ.H | 1.00:45.28:0 | 443 | 457 | |
| Q02819\|NUCB1_MOUSE | L.EEFLA.S | 1.00:1.88:0 | 314 | 318 | |

| | | | | | |
|---|---|---|---|---|---|
| Q02819\|NUCB1_MOUSE | K.QFEHLDPQNQHTFEAR.D | 1.00:23.58:0 | 136 | 151 | |
| Q02819\|NUCB1_MOUSE | I.NVLETDGHF.R | 0:1.00:2.96 | 59 | 67 | |
| Q02819\|NUCB1_MOUSE | K.EVWEELDGLDPNRFNPK.T | 1.00:3.59:0 | 229 | 245 | |
| Q02819\|NUCB1_MOUSE | R.SQDRLEAQ.K | 1.00:0:0.01 | 377 | 384 | |
| Q02819\|NUCB1_MOUSE | R.LVTLEEFLA.S | 1.00:0.62:0.65 | 310 | 318 | |
| Q02819\|NUCB1_MOUSE | K.QFEHLDPQNQHT.F | 1.00:0.64:0 | 136 | 147 | |
| Q02819\|NUCB1_MOUSE | I.NVLETDGHFREK.L | 1.00:7.48:0 | 59 | 70 | |
| Q02819\|NUCB1_MOUSE | I.NVLETDGHFREKLQAA.N | 0.000694444 | 59 | 74 | |
| Q02819\|NUCB1_MOUSE | K.QEPNLQVDHM(+15.99)NLL.K | 1.00:0.76:0 | 122 | 134 | Oxidation (M) |
| Q02819\|NUCB1_MOUSE | R.ADTDDAPVPAPA.G | 1.00:1.85:0 | 420 | 431 | |
| Q02819\|NUCB1_MOUSE | H.HEEFKRY(+79.96)EM(+15.99)LKEH.E | 1.00:0:0.41 | 172 | 184 | Sulfation; Oxidation (M) |
| Q02819\|NUCB1_MOUSE | K.KVPEQPPELPQLDSQHL | 1.00:2.51:0 | 443 | 459 | |
| Q02819\|NUCB1_MOUSE | K.NVDTNQDR.L | 0:1.00:0.08 | 302 | 309 | |
| Q02819\|NUCB1_MOUSE | K.EVWEELDGLDPNR.F | 1.00:6.47:1.06 | 229 | 241 | |
| Q02819\|NUCB1_MOUSE | R.YLESLGEEQR.K | 1.00:5.51:0 | 189 | 198 | |
| Q02819\|NUCB1_MOUSE | K.KVPEQPPELPQL.D | 1.00:0.67:0 | 443 | 454 | |
| Q02819\|NUCB1_MOUSE | R.YLESLGEEQRK.E | 1.00:5.37:0 | 189 | 199 | |
| Q02819\|NUCB1_MOUSE | R.DLAQYDAAHHEEFKR.Y | 0.000694444 | 163 | 177 | |
| Q02819\|NUCB1_MOUSE | K.EFGDTGEGWK.T | 1.00:7.28:0 | 324 | 333 | |
| Q02819\|NUCB1_MOUSE | A.NAEDIKSGKLSQELDFVSHNVR.T | 0.000694444 | 75 | 96 | |
| Q02819\|NUCB1_MOUSE | A.HHEEF.K | 1.00:0:0 | 171 | 175 | |
| Q02819\|NUCB1_MOUSE | E.QPPELPQLDSQHL | 1.00:0.72:0 | 447 | 459 | |
| Q02819\|NUCB1_MOUSE | K.EVWEELDGLDPN.R | 1.00:6.40:0.72 | 229 | 240 | |
| Q02819\|NUCB1_MOUSE | R.YLESLGEEQRKEAERKLQEQQRR.H | 0.000694444 | 189 | 211 | |
| Q02819\|NUCB1_MOUSE | I.LHDINSDGVLDEQELEA.L | 1.00:1.22:0.43 | 250 | 266 | |
| Q02819\|NUCB1_MOUSE | K.TFFILHDINSDGVLDEQELEA.L | 1.00:14.31:0 | 246 | 266 | |
| Q02819\|NUCB1_MOUSE | R.TKLDELKR.Q | 0.000694444 | 97 | 104 | |
| Q02819\|NUCB1_MOUSE | R.ADTDDAPVPAPAGDQKDVPA.S | 0.000694444 | 420 | 439 | |
| Q02819\|NUCB1_MOUSE | K.LSQELDFVSHNVR.T | 1.00:7.24:0 | 84 | 96 | |
| Q02819\|NUCB1_MOUSE | R.YLESLGEEQRKEAERKLQEQQR.R | 0.000694444 | 189 | 210 | |
| Q02819\|NUCB1_MOUSE | Y.DAAHHEEFKR.Y | 0.000694444 | 168 | 177 | |
| Q02819\|NUCB1_MOUSE | R.LVTLEEFLASTQR.K | 0.000694444 | 310 | 322 | |
| Q02819\|NUCB1_MOUSE | R.DLAQYDAAHHEEFK.R | 1.00:0.83:0 | 163 | 176 | |
| Q02819\|NUCB1_MOUSE | I.NVLETDGHFR.E | 0.000694444 | 59 | 68 | |
| Q02819\|NUCB1_MOUSE | K.VYDPKNEE.D | 0.000694444 | 275 | 282 | |
| Q02819\|NUCB1_MOUSE | R.TKLDELKRQEVSR.L | 0.000694444 | 97 | 109 | |
| Q02819\|NUCB1_MOUSE | R.LVTLEEFLASTQRK.E | 0.000694444 | 310 | 323 | |
| Q02819\|NUCB1_MOUSE | K.M(+15.99)DAKQEPNLQVDHM(+15.99)NLLK.Q | 0.000694444 | 118 | 135 | Oxidation (M) |
| Q02819\|NUCB1_MOUSE | R.AAPPQEDSQATETPDTGLYYHR.Y | 0.000694444 | 31 | 52 | |
| Q02819\|NUCB1_MOUSE | A.GDQKDVPA.S | 1.00:6.12:0 | 432 | 439 | |
| Q02819\|NUCB1_MOUSE | K.LSQELDFVSHNV.R | 1.00:1.71:0 | 84 | 95 | |
| Q02819\|NUCB1_MOUSE | K.VPEQPPELPQLDSQHL | 1.00:2.11:0 | 444 | 459 | |
| Q02819\|NUCB1_MOUSE | K.VYDPKNEEDDM(+15.99)REM(+15.99)EEERLR.M | 0.000694444 | 275 | 294 | Oxidation (M) |
| Q02819\|NUCB1_MOUSE | K.LSQELDFV.S | 1.00:4.27:0 | 84 | 91 | |
| Q02819\|NUCB1_MOUSE | R.SQDRLEAQK.R | 1.00:2.27:0 | 377 | 385 | |
| Q02819\|NUCB1_MOUSE | A.TRDLAQYDAAHHEEFKR.Y | 0.000694444 | 161 | 177 | |
| Q02819\|NUCB1_MOUSE | A.NAEDIKSGKLSQELDFVSHNV.R | 0.000694444 | 75 | 95 | |
| Q02819\|NUCB1_MOUSE | V.LETDGHFR.E | 1.00:0.82:0 | 61 | 68 | |
| Q02819\|NUCB1_MOUSE | A.YTEEELKR.F | 0.000694444 | 341 | 348 | |
| Q02819\|NUCB1_MOUSE | K.LQEQQR.R | 0.000694444 | 205 | 210 | |
| Q02819\|NUCB1_MOUSE | A.GDQKDVPASEKKVPEQPPELPQLDSQHL | 0.000694444 | 432 | 459 | |
| Q02819\|NUCB1_MOUSE | R.ADTDDAPVPAPAGDQKDVPASEKKVPEQPPELPQLDSQHL | 0.000694444 | 420 | 459 | |
| Q02819\|NUCB1_MOUSE | V.LETDGHF.R | 1.00:0:0 | 61 | 67 | |
| Q02819\|NUCB1_MOUSE | R.DLAQYDAAHHEEF.K | 1.00:0:0 | 163 | 175 | |
| Q02819\|NUCB1_MOUSE | R.SQDRLEAQKR.E | 0.000694444 | 377 | 386 | |
| Q02819\|NUCB1_MOUSE | K.LDELKRQEVS.R | 0.000694444 | 99 | 108 | |
| Q02819\|NUCB1_MOUSE | R.YLESLGEEQRKEAERKLQEQQ.R | 0.000694444 | 189 | 209 | |
| Q02819\|NUCB1_MOUSE | A.YTEEELKRFEEELA.A | 1.00:0:0 | 341 | 354 | |
| Q02819\|NUCB1_MOUSE | K.LQAANAEDIKSGK.L | 0.000694444 | 71 | 83 | |
| Q02819\|NUCB1_MOUSE | Q.PPELPQLDSQHL | 1.00:0.66:0 | 448 | 459 | |
| Q02819\|NUCB1_MOUSE | A.NAEDIKSGKLSQELDFV.S | 0.000694444 | 75 | 91 | |
| Q02819\|NUCB1_MOUSE | A.NAEDIKSGK.L | 0.000694444 | 75 | 83 | |
| Q02819\|NUCB1_MOUSE | R.AAPPQEDSQATETPDTGLYYHRYLQEVI.N | 0.000694444 | 31 | 58 | |
| Q02819\|NUCB1_MOUSE | K.LDELKRQEVSR.L | 0.000694444 | 99 | 109 | |
| Q02819\|NUCB1_MOUSE | R.KLQEQQR.R | 0.000694444 | 204 | 210 | |
| Q02819\|NUCB1_MOUSE | K.SGKLSQELDFVSHNVR.T | 0.000694444 | 81 | 96 | |
| Q02819\|NUCB1_MOUSE | A.GDQKDVPASEK.K | 0.000694444 | 432 | 442 | |
| Q02819\|NUCB1_MOUSE | R.YLESLGEEQ.R | 1.00:0:0 | 189 | 197 | |
| Q02819\|NUCB1_MOUSE | V.EM(+15.99)SPAYTEEELKRFEEELAAREAELNAR.A | 0.000694444 | 336 | 363 | Oxidation (M) |
| Q02819\|NUCB1_MOUSE | R.RRYLESLGEEQR.K | 0.000694444 | 187 | 198 | |
| Q02819\|NUCB1_MOUSE | A.SEKKVPEQPPELPQLDSQHL | 0.000694444 | 440 | 459 | |
| Q02819\|NUCB1_MOUSE | K.LQAANAEDIKSGKLSQELDFVSHNVR.T | 0.000694444 | 71 | 96 | |
| Q08535\|SECR_MOUSE | R.SEQDTENIPENSLAR.S | 0:1.00:0.41 | 62 | 76 | |
| Q60994\|ADIPO_MOUSE | T.GFLLYHDTN | 1.00:0:0 | 239 | 247 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDKAKLKKTETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 2 | 44 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | K.NTLPTKETIEQEKRSEIS | 0.000694444 | 27 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | K.ETIEQEKRSEIS | 0.000694444 | 33 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | K.TETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 21 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | K.NTLPTK.E | 1.00:112.67:0 | 27 | 32 | |
| Q6ZWY8\|TYB10_MOUSE | T.KETIEQEK.R | 1.00:0.96:0 | 32 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | K.E(-18.01)TIEQEK.R | 1.00:2.01:0.02 | 33 | 39 | Pyro-glu from E |
| Q6ZWY8\|TYB10_MOUSE | K.ETIEQEKR.S | 0:1.00:0.09 | 33 | 40 | |
| Q6ZWY8\|TYB10_MOUSE | K.NTLPTKETIEQEK.R | 0.000694444 | 27 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | K.TETQEKNTLPT.K | 1.00:2.34:0 | 21 | 31 | |
| Q6ZWY8\|TYB10_MOUSE | K.LKKTETQEK.N | 0.000694444 | 18 | 26 | |
| Q6ZWY8\|TYB10_MOUSE | A.SFDKAKLKKTETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 12 | 44 | |

| | | | | | |
|---|---|---|---|---|---|
| Q6ZWY8\|TYB10_MOUSE | A.SFDKAK.L | 0.000694444 | 12 | 17 | |
| Q6ZWY8\|TYB10_MOUSE | T.KETIEQEKRSEIS | 0.000694444 | 32 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99).G | 0:1.00:0.06 | 2 | 7 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | K.TETQEKNTLPTKETIEQEK.R | 0.000694444 | 21 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | T.QEKNTLPTKETIEQEKRSEIS | 0.000694444 | 24 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | E.TQEKNTLPTKETIEQEK.R | 1.00:92.94:0 | 23 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | K.E(-18.01)TIEQEKR.S | 1.00:106.06:0 | 33 | 40 | Pyro-glu from E |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDMGEIA.S | 1.00:2.71:7.98 | 2 | 11 | Acetylation (N-term) |
| Q6ZWY8\|TYB10_MOUSE | K.NTLPTKETIEQEKR.S | 0.000694444 | 27 | 40 | |
| Q6ZWY8\|TYB10_MOUSE | M.GEIASFDKAKLKKTETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 8 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | E.TQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 23 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | L.KKTETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 19 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | E.KNTLPTKETIEQEKRSEIS | 0.000694444 | 26 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | K.LKKTETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 18 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | E.TQEKNTLPTKETIEQE.K | 0.000694444 | 23 | 38 | |
| Q6ZWY8\|TYB10_MOUSE | K.TETQEKNTLPTKETIEQEKR.S | 0.000694444 | 21 | 40 | |
| Q6ZWY8\|TYB10_MOUSE | K.LKKTETQEKNTLPTKETIEQEK.R | 0.000694444 | 18 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | K.KTETQEKNTLPTKETIEQEKR.S | 0.000694444 | 20 | 40 | |
| Q6ZWY8\|TYB10_MOUSE | E.KNTLPTKETIEQEK.R | 0.000694444 | 26 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | N.TLPTKETIEQE.K | 0.000694444 | 28 | 38 | |
| Q6ZWY8\|TYB10_MOUSE | A.SFDKAKL.K | 0.000694444 | 12 | 18 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDKAKLK.K | 0.000694444 | 2 | 19 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | E.TQEKNTLPT.K | 1.00:2.23:0 | 23 | 31 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDKAKLKKTETQEKNTLPTKETIEQEK.R | 0.000694444 | 2 | 39 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | A.SFDKAKLK.K | 0.000694444 | 12 | 19 | |
| Q6ZWY8\|TYB10_MOUSE | K.TETQEKNTLPTK.E | 0.000694444 | 21 | 32 | |
| Q6ZWY8\|TYB10_MOUSE | T.ETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 22 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | N.TLPTKETIEQEK.R | 0.000694444 | 28 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | T.IEQEKR.S | 0.000694444 | 35 | 40 | |
| Q6ZWY8\|TYB10_MOUSE | N.TLPTKETIEQEKRSEIS | 0.000694444 | 28 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | E.TQEKNTLPTKETIE.Q | 0.000694444 | 23 | 36 | |
| Q6ZWY8\|TYB10_MOUSE | D.M(+15.99)GEIASFDKAKLKKTETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 7 | 44 | Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | K.KTETQEKNTLPT.K | 0.000694444 | 20 | 31 | |
| Q6ZWY8\|TYB10_MOUSE | K.NTLPTKETIE.Q | 0.000694444 | 27 | 36 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDKAKL.K | 0:1.00:0.03 | 2 | 18 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDKAKLKKTETQEKNTLPT.K | 0.000694444 | 2 | 31 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | A.SFDKAKLKK.T | 0.000694444 | 12 | 20 | |
| Q6ZWY8\|TYB10_MOUSE | K.KTETQEKNTLPTK.E | 0.000694444 | 20 | 32 | |
| Q6ZWY8\|TYB10_MOUSE | K.NTLPTKETIEQEKRSE.I | 0.000694444 | 27 | 42 | |
| Q6ZWY8\|TYB10_MOUSE | K.NTLPTKETIEQEKRS.E | 0.000694444 | 27 | 41 | |
| Q6ZWY8\|TYB10_MOUSE | M.GEIASFDKA.K | 1.00:0.36:0 | 8 | 16 | |
| Q6ZWY8\|TYB10_MOUSE | M.GEIASFDKAKLKKTETQEKNTLPTKETIEQEK.R | 0.000694444 | 8 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | T.LPTKETIEQEKRSEIS | 0.000694444 | 29 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | E.TQEKNTLPTK.E | 0.000694444 | 23 | 32 | |
| Q6ZWY8\|TYB10_MOUSE | K.KTETQEKNTLPTKETIEQEK.R | 0.000694444 | 20 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | T.QEKNTLPTKETIEQEK.R | 0.000694444 | 24 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | I.EQEKRSEIS | 1.00:1.07:0 | 36 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | A.DKPDM(+15.99)GEIA.S | 0.000694444 | 3 | 11 | Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | K.KTETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 20 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | Q.EKNTLPTKETIEQEKRSEIS | 0.000694444 | 25 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | M.GEIASFDKAKLK.K | 0.000694444 | 8 | 19 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDK.A | 0.000694444 | 2 | 15 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | K.LKKTETQEKNTLPTKETIEQE.K | 0.000694444 | 18 | 38 | |
| Q6ZWY8\|TYB10_MOUSE | P.DM(+15.99)GEIASFDKAKLKKTETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 6 | 44 | Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDKAKLKKTETQEK.N | 1.00:137.07:0 | 2 | 26 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | M.GEIASFDKAKLKK.T | 0.000694444 | 8 | 20 | |
| Q6ZWY8\|TYB10_MOUSE | E.TIEQEKR.S | 0.000694444 | 34 | 40 | |
| Q6ZWY8\|TYB10_MOUSE | K.KTETQEKNTLPTKETIE.Q | 0.000694444 | 20 | 36 | |
| Q6ZWY8\|TYB10_MOUSE | P.TKETIEQEKRSEIS | 0.000694444 | 31 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | T.IEQEKRSEIS | 1.00:2.77:0 | 35 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | K.NTLPTKETI.E | 0.000694444 | 27 | 35 | |
| Q6ZWY8\|TYB10_MOUSE | E.TQEKNTLPTKETIEQEKRSE.I | 0.000694444 | 23 | 42 | |
| Q6ZWY8\|TYB10_MOUSE | M.GEIASFDKAKLKKTETQEK.N | 0.000694444 | 8 | 26 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDKAKLKK.T | 0.000694444 | 2 | 20 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | L.KKTETQEKNTLPTKETIEQEK.R | 0.000694444 | 19 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | T.Q(-17.03)EKNTLPTKETIEQEK.R | 0.000694444 | 24 | 39 | Pyro-glu from Q |
| Q6ZWY8\|TYB10_MOUSE | L.PTKETIEQEK.R | 0.000694444 | 30 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEI.A | 1.00:1.86:0.52 | 2 | 10 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | K.TETQEKNTLPTKETIE.Q | 0.000694444 | 21 | 36 | |
| Q6ZWY8\|TYB10_MOUSE | K.KTETQEK.N | 1.00:10.46:0 | 20 | 26 | |
| Q6ZWY8\|TYB10_MOUSE | P.TKETIEQEK.R | 0.000694444 | 31 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | K.LKKTETQEKNTLPT.K | 0.000694444 | 18 | 31 | |
| Q6ZWY8\|TYB10_MOUSE | K.AKLKKTETQEKNTLPTKETIEQEKRSEIS | 0.000694444 | 16 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | K.LKKTETQEKNTLPTKETIEQEKR.S | 0.000694444 | 18 | 40 | |
| Q6ZWY8\|TYB10_MOUSE | L.PTKETIEQEKRSEIS | 0.000694444 | 30 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDKAKLKKTE.T | 0.000694444 | 2 | 22 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDKAKLKKTETQEKN.T | 0.000694444 | 2 | 27 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | M.GEIASFDKAK.L | 0.000694444 | 8 | 17 | |
| Q6ZWY8\|TYB10_MOUSE | Q.EKNTLPTKETIEQEK.R | 0.000694444 | 25 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | E.TIEQEKRSEIS | 1.00:0.95:0 | 34 | 44 | |
| Q6ZWY8\|TYB10_MOUSE | A.SFDKAKLKKTETQEK.N | 0.000694444 | 12 | 26 | |
| Q6ZWY8\|TYB10_MOUSE | E.KNTLPTK.E | 1.00:3.97:0 | 26 | 32 | |
| Q6ZWY8\|TYB10_MOUSE | M.GEIASFDKAKLKKTETQEKNTLPT.K | 0.000694444 | 8 | 31 | |
| Q6ZWY8\|TYB10_MOUSE | D.M(+15.99)GEIASFDKAKLKKTETQEKNTLPTKETIEQEK.R | 0.000694444 | 7 | 39 | Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | A.SFDKAKLKKTETQEKNTLPTKETIEQEK.R | 0.000694444 | 12 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | K.TETQEKNTLPTKETIEQE.K | 0.000694444 | 21 | 38 | |

| | | | | | |
|---|---|---|---|---|---|
| Q6ZWY8\|TYB10_MOUSE | M.GEIASFDKAKL.K | 0.000694444 | 8 | 18 | |
| Q6ZWY8\|TYB10_MOUSE | K.ETIEQEK.R | 1.00:2.83:0 | 33 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDM(+15.99)GEIASFDKAKLKKTETQEKNTLPTK.E | 0.000694444 | 2 | 32 | Acetylation (N-term); Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | A.DKPDM(+15.99)GEIASFDKAKLK.K | 0.000694444 | 3 | 19 | Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | E.TQEKNTLPTKETIEQEKR.S | 0.000694444 | 23 | 40 | |
| Q6ZWY8\|TYB10_MOUSE | D.M(+15.99)GEIASFDKAKLK.K | 0.000694444 | 7 | 19 | Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | D.M(+15.99)GEIASFDKAKLKK.T | 0.000694444 | 7 | 20 | Oxidation (M) |
| Q6ZWY8\|TYB10_MOUSE | S.FDKAKLK.K | 0.000694444 | 13 | 19 | |
| Q6ZWY8\|TYB10_MOUSE | K.AKLKKTETQEK.N | 0.000694444 | 16 | 26 | |
| Q8K0E8\|FIBB_MOUSE | Y.DEPTDSLDAR.G | 1.00:315.93:0.83 | 25 | 34 | |
| Q8K0E8\|FIBB_MOUSE | K.EEPPSLRPAPPPISGGG.Y | 1.00:0.58:0 | 44 | 60 | |
| Q8K0E8\|FIBB_MOUSE | E.EPPSLRPAPPPISGGGYR.A | 0.000694444 | 45 | 62 | |
| Q8K0E8\|FIBB_MOUSE | K.EEPPSLRPAPPPISGGGYR.A | 0.000694444 | 44 | 62 | |
| Q8K0E8\|FIBB_MOUSE | E.EPPSLRPAPPPISGGG.Y | 1.00:4.01:0 | 45 | 60 | |
| Q8K0E8\|FIBB_MOUSE | E.EPPSLRPAPPPISGGGY.R | 1.00:52.75:0 | 45 | 61 | |
| Q8K0E8\|FIBB_MOUSE | K.EEPPSLRPAPPPISGGGY.R | 1.00:3.16:0 | 44 | 61 | |
| Q8K0E8\|FIBB_MOUSE | I.LEDLR.S | 0.000694444 | 192 | 196 | |
| Q8K0E8\|FIBB_MOUSE | R.PAPPPISGGGY.R | 1.00:0.44:0 | 51 | 61 | |
| Q8K0E8\|FIBB_MOUSE | L.RPAPPPISGGGY.R | 1.00:1.11:0 | 50 | 61 | |
| Q8K0E8\|FIBB_MOUSE | Y.DEPTDSLDA.R | 1.00:54.10:0 | 25 | 33 | |
| Q8K0E8\|FIBB_MOUSE | R.KEEPPSLRPAPPPISGGG.Y | 1.00:0.41:0 | 43 | 60 | |
| Q8K0E8\|FIBB_MOUSE | Y.DEPTDSLDARGHRPVD.R | 0.000694444 | 25 | 40 | |
| Q8K0E8\|FIBB_MOUSE | M.YLIQPDTSIKPY.R | 1.00:0.11:0 | 245 | 256 | |
| Q8K0E8\|FIBB_MOUSE | R.KEEPPSLRPAPPPISGGGYR.A | 0.000694444 | 43 | 62 | |
| Q8K0E8\|FIBB_MOUSE | P.TDSLDAR.G | 0.000694444 | 28 | 34 | |
| Q8K0E8\|FIBB_MOUSE | A.ADDDYDEPT(+79.96)DSLDARGHRPVDR.R | 0.000694444 | 20 | 41 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | K.SSIAELNNNIQS.V | 1.00:0.13:0 | 115 | 126 | |
| Q8K0E8\|FIBB_MOUSE | D.EPTDSLDARGHRPVDR.R | 0.000694444 | 26 | 41 | |
| Q8K0E8\|FIBB_MOUSE | L.RPAPPPISGGG.Y | 1.00:0:0 | 50 | 60 | |
| Q8K0E8\|FIBB_MOUSE | R.PAPPPISGGG.Y | 1.00:0.08:0 | 51 | 60 | |
| Q8K0E8\|FIBB_MOUSE | V.TFQYLTLLKDM(+15.99).W | 0.000694444 | 134 | 144 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | R.RKEEPPSLRPAPPPISGGGY.R | 1.00:116.56:0 | 42 | 61 | |
| Q8K0E8\|FIBB_MOUSE | E.PTDSLDARGHRPVDR.R | 0.000694444 | 27 | 41 | |
| Q8K0E8\|FIBB_MOUSE | R.GHRPVD.R | 0.000694444 | 35 | 40 | |
| Q8K0E8\|FIBB_MOUSE | A.ADDDY(+79.96)DEPTDSLDARGHRPVDR.R | 0.000694444 | 20 | 41 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | K.IRPFFPQQ | 1.00:0.79:0 | 474 | 481 | |
| Q8K0E8\|FIBB_MOUSE | R.RKEEPPSL.R | 1.00:2.24:0 | 42 | 49 | |
| Q8K0E8\|FIBB_MOUSE | E.PTDSLDAR.G | 1.00:6.20:0 | 27 | 34 | |
| Q8K0E8\|FIBB_MOUSE | P.TDSLDARGHRPVDR.R | 0.000694444 | 28 | 41 | |
| Q8K0E8\|FIBB_MOUSE | A.ADDDY(+79.96)DEPTDSLDAR.G | 0.000694444 | 20 | 34 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | Y.DEPTDSLDARGHRPVDR.R | 0.000694444 | 25 | 41 | |
| Q8K0E8\|FIBB_MOUSE | T.LLNQERPI.K | 1.00:0.85:0 | 106 | 113 | |
| Q8K0E8\|FIBB_MOUSE | P.APPPISGGGY.R | 1.00:260.09:0 | 52 | 61 | |
| Q8K0E8\|FIBB_MOUSE | R.GHRPVDR.R | 0.000694444 | 35 | 41 | |
| Q8K0E8\|FIBB_MOUSE | K.GGETSEM(+15.99)YLIQPDTSIKPYR.V | 0.000694444 | 238 | 257 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | V.TFQYLT.L | 1.00:1.67:0 | 134 | 139 | |
| Q8K0E8\|FIBB_MOUSE | K.GTAGNALM(+15.99)DGASQLV.G | 1.00:2.72:0 | 367 | 381 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | R.DNDGWVTTDPR.K | 1.00:0.29:0 | 401 | 411 | |
| Q8K0E8\|FIBB_MOUSE | V.TFQYLTLLK.D | 0.000694444 | 134 | 142 | |
| Q8K0E8\|FIBB_MOUSE | E.EPPSLRPAPPP.I | 1.00:56.28:0 | 45 | 55 | |
| Q8K0E8\|FIBB_MOUSE | E.E(-18.01)PPSLRP.A | 0:1.00:0.07 | 45 | 51 | Pyro-glu from E |
| Q8K0E8\|FIBB_MOUSE | D.DDY(+79.96)DEPTDSLDARGHRPVDR.R | 0.000694444 | 22 | 41 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | E.EPPSLRPAPPPIS.G | 0.000694444 | 45 | 57 | |
| Q8K0E8\|FIBB_MOUSE | V.KENENVI.N | 1.00:2.83:0.52 | 153 | 159 | |
| Q8K0E8\|FIBB_MOUSE | R.PAPPPISGGGYR.A | 0.000694444 | 51 | 62 | |
| Q8K0E8\|FIBB_MOUSE | K.GTAGNALM(+15.99)DGA.S | 1.00:0.03:0.02 | 367 | 377 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | Y.DEPTDSLDARGH.R | 0.000694444 | 25 | 36 | |
| Q8K0E8\|FIBB_MOUSE | Y.DEPTDSLDARGHRPVDRRKE.E | 0.000694444 | 25 | 44 | |
| Q8K0E8\|FIBB_MOUSE | K.IRPFFPQ.Q | 1.00:0.21:0 | 474 | 480 | |
| Q8K0E8\|FIBB_MOUSE | A.ADDDY(+79.96)DEPTDSLDARGHRPV.D | 0.000694444 | 20 | 39 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | I.AELNNNIQSV.S | 1.00:0.94:0 | 118 | 127 | |
| Q8K0E8\|FIBB_MOUSE | K.AHYGGFTV.Q | 1.00:0.66:0 | 344 | 351 | |
| Q8K0E8\|FIBB_MOUSE | Y.DEPTDSLDARGHRPVDRRKEEPPSLRPAPPPISGGGY.R | 0.000694444 | 25 | 61 | |
| Q8K0E8\|FIBB_MOUSE | A.ADDDY(+79.96)DEPTDSLDARGHRPVD.R | 0.000694444 | 20 | 40 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | R.RKEEPPSLR.P | 0.000694444 | 42 | 50 | |
| Q8K0E8\|FIBB_MOUSE | A.ADDDY(+79.96)DEPTDSLDARGHRPVDRR.K | 0.000694444 | 20 | 42 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | D.NDGWVTTDPR.K | 1.00:10.67:0 | 402 | 411 | |
| Q8K0E8\|FIBB_MOUSE | I.EM(+15.99)EDWKGDKVK.A | 0.000694444 | 333 | 343 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | K.SSIAELNNNI.Q | 1.00:0:0 | 115 | 124 | |
| Q8K0E8\|FIBB_MOUSE | A.ADDDYDEPTDSLDAR.G | 1.00:5.92:0 | 20 | 34 | |
| Q8K0E8\|FIBB_MOUSE | S.SIAELNNNIQSV.S | 1.00:2.14:0 | 116 | 127 | |
| Q8K0E8\|FIBB_MOUSE | M.YLIQPDTSIKP.Y | 1.00:1.18:0 | 245 | 255 | |
| Q8K0E8\|FIBB_MOUSE | D.YDEPT(+79.96)DSLDARGHRPV.D | 0.000694444 | 24 | 39 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | K.YQVSVNK.Y | 1.00:0.84:0 | 358 | 364 | |
| Q8K0E8\|FIBB_MOUSE | D.EPTDSLDAR.G | 1.00:79.74:0 | 26 | 34 | |
| Q8K0E8\|FIBB_MOUSE | Y.DEPTDSLDARGHRPV.D | 0.000694444 | 25 | 39 | |
| Q8K0E8\|FIBB_MOUSE | I.KSSIAELNNNIQS.V | 1.00:0:0 | 114 | 126 | |
| Q8K0E8\|FIBB_MOUSE | K.EEPPSLRP.A | 1.00:6.84:0 | 44 | 51 | |
| Q8K0E8\|FIBB_MOUSE | A.ADDDYDEPTDSLDARGHRPVDR.R | 0.000694444 | 20 | 41 | |
| Q8K0E8\|FIBB_MOUSE | P.TDSLDARGHRPVD.R | 0.000694444 | 28 | 40 | |
| Q8K0E8\|FIBB_MOUSE | T.DSLDAR.G | 1.00:320.43:63.56 | 29 | 34 | |
| Q8K0E8\|FIBB_MOUSE | E.E(-18.01)PPSLRPAPPPISGGGY.R | 0.000694444 | 45 | 61 | Pyro-glu from E |
| Q8K0E8\|FIBB_MOUSE | I.LEDQR.L | 0:1.00:0.36 | 166 | 170 | |
| Q8K0E8\|FIBB_MOUSE | L.DARGHRPVDR.R | 0.000694444 | 32 | 41 | |
| Q8K0E8\|FIBB_MOUSE | L.NQERPI.K | 1.00:0.03:0 | 108 | 113 | |
| Q8K0E8\|FIBB_MOUSE | D.DDYDEPT(+79.96)DSLDARGHRPVDR.R | 0.000694444 | 22 | 41 | Sulfation |

| | | | | | |
|---|---|---|---|---|---|
| Q8K0E8\|FIBB_MOUSE | K.EDGGGWWYNR.C | 1.00:5.19:0 | 417 | 426 | |
| Q8K0E8\|FIBB_MOUSE | A.ADDDY(+79.96)DEPTDSLDARGHRPVDRRK.E | 0.000694444 | 20 | 43 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | D.E(-18.01)PTDSLDARGHRPVDR.R | 0.000694444 | 26 | 41 | Pyro-glu from E |
| Q8K0E8\|FIBB_MOUSE | P.TDSLDARGHRPV.D | 0.000694444 | 28 | 39 | |
| Q8K0E8\|FIBB_MOUSE | Y.DEPTDSLDARGHRPVDRR.K | 0.000694444 | 25 | 42 | |
| Q8K0E8\|FIBB_MOUSE | R.M(+15.99)GPTELLI.E | 0:1.00:0.12 | 325 | 332 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | E.E(-18.01)PPSLRPAPPPISGGGYR.A | 0.000694444 | 45 | 62 | Pyro-glu from E |
| Q8K0E8\|FIBB_MOUSE | K.ENENVI.N | 0.000694444 | 154 | 159 | |
| Q8K0E8\|FIBB_MOUSE | I.DETVNDNIPL.N | 1.00:0:0 | 174 | 183 | |
| Q8K0E8\|FIBB_MOUSE | D.EPTDSLDARGHRPVD.R | 0.000694444 | 26 | 40 | |
| Q8K0E8\|FIBB_MOUSE | K.GGETSEM(+15.99)YLIQPDTSIKPY.R | 1.00:0.31:0 | 238 | 256 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | S.EM(+15.99)YLIQPDTSIKPYR.V | 0.000694444 | 243 | 257 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | I.EM(+15.99)EDWKGDKV.K | 1.00:0:0 | 333 | 342 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | A.ADDDY(+79.96)DEPTDSLDARGHRPVDRRKEEPPSLRPAPPPISGGGYR.A | 0.000694444 | 20 | 62 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | E.EPPSLRPAPPPISG.G | 0.000694444 | 45 | 58 | |
| Q8K0E8\|FIBB_MOUSE | E.PTDSLDARGHRPV.D | 0.000694444 | 27 | 39 | |
| Q8K0E8\|FIBB_MOUSE | C.GLPGEYWLGNDK.I | 1.00:0:0 | 307 | 318 | |
| Q8K0E8\|FIBB_MOUSE | D.DDY(+79.96)DEPTDSLDARGHRPVDRR.K | 0.000694444 | 22 | 42 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | Y.DEPTDSLDARGHRPVDRRKEEPPSL.R | 0.000694444 | 25 | 49 | |
| Q8K0E8\|FIBB_MOUSE | R.YYWGGLYSWDM(+15.99)SK.H | 1.00:3.78:0 | 436 | 448 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | V.TFQYLTLL.K | 0.000694444 | 134 | 141 | |
| Q8K0E8\|FIBB_MOUSE | V.TFQYLTLLKDM(+15.99)W.K | 0.000694444 | 134 | 145 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | A.ADDDY(+79.96)DEPTDSLDARGHRPVDRRKEEPPSLRPAPPPISGGGY.R | 0.000694444 | 20 | 61 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | K.SSIAELNNNIQSV.S | 1.00:0:0 | 115 | 127 | |
| Q8K0E8\|FIBB_MOUSE | R.SILEDLR.S | 0.000694444 | 190 | 196 | |
| Q8K0E8\|FIBB_MOUSE | T.DSLDARGHRPVDR.R | 0.000694444 | 29 | 41 | |
| Q8K0E8\|FIBB_MOUSE | T.FQYLTLLKDM(+15.99).W | 0.000694444 | 135 | 144 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | E.PTDSLDARGHRPVDRRKE.E | 0.000694444 | 27 | 44 | |
| Q8K0E8\|FIBB_MOUSE | D.DDY(+79.96)DEPTDSLDARGHRPV.D | 0.000694444 | 22 | 39 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | D.EPTDSLDARGHRPV.D | 0.000694444 | 26 | 39 | |
| Q8K0E8\|FIBB_MOUSE | P.TDSLDARGHRPVDRRKE.E | 0.000694444 | 28 | 44 | |
| Q8K0E8\|FIBB_MOUSE | V.SDTSSV.T | 0.000694444 | 128 | 133 | |
| Q8K0E8\|FIBB_MOUSE | K.EEPPSLRPAPPP.I | 1.00:13.43:0 | 44 | 55 | |
| Q8K0E8\|FIBB_MOUSE | Y.LIQPDTSIKP.Y | 1.00:0:0 | 246 | 255 | |
| Q8K0E8\|FIBB_MOUSE | D.EPTDSLDARGHRPVDRRKE.E | 0.000694444 | 26 | 44 | |
| Q8K0E8\|FIBB_MOUSE | S.LRPAPPPISGGGY.R | 0.000694444 | 49 | 61 | |
| Q8K0E8\|FIBB_MOUSE | D.SLDARGHRPVDR.R | 0.000694444 | 30 | 41 | |
| Q8K0E8\|FIBB_MOUSE | D.DDYDEPTDS(+79.96)LDARGHRPVD.R | 0.000694444 | 22 | 40 | Sulfation |
| Q8K0E8\|FIBB_MOUSE | E.EPPSLRPAPPPISGGGYRA.R | 0.000694444 | 45 | 63 | |
| Q8K0E8\|FIBB_MOUSE | H.RPVDR.R | 0.000694444 | 37 | 41 | |
| Q8K0E8\|FIBB_MOUSE | V.SDTSSVTFQYLTLLKDM(+15.99).W | 0.000694444 | 128 | 144 | Oxidation (M) |
| Q8K0E8\|FIBB_MOUSE | R.KEEPPSLRPAPPPISGGGY.R | 1.00:4.76:0 | 43 | 61 | |
| Q99KQ4\|NAMPT_MOUSE | V.TLEEGKGDLEEYGHDL.L | 1.00:93.58:76.76 | 442 | 457 | |
| Q99KQ4\|NAMPT_MOUSE | K.NAQLNIEQDVAPH | 1.00:1.80:0 | 479 | 491 | |
| Q9D3P9\|NEUT_MOUSE | V.NDKNEKEEV.I | 1.00:0:0 | 130 | 138 | |
| Q9D3P9\|NEUT_MOUSE | C.SDSEEDVRALEADLLTNM(+15.99)HTSKI.S | 0.000694444 | 23 | 45 | Oxidation (M) |
| Q9D3P9\|NEUT_MOUSE | R.KLPLVLDGFSLEA.M | 1.00:0:0 | 87 | 99 | |
| Q9EPS2\|PYY_MOUSE | A.SPEEL.S | 1.00:0.25:0.30 | 41 | 45 | |
| Q9EPS2\|PYY_MOUSE | P.AKPEAPGEDASPEELS.R | 1.00:0.42:0 | 31 | 46 | |
| Q9EPS2\|PYY_MOUSE | E.NLPFRPEGLDQW | 1.00:0.43:2.68 | 87 | 98 | |
| Q9EPS2\|PYY_MOUSE | K.LLFTDDSDSENLPFRPEGLDQ.W | 0.000694444 | 77 | 97 | |
| Q9EPS2\|PYY_MOUSE | K.LLFTDDSDSENLPFR.P | 0:1.00:0.11 | 77 | 91 | |
| Q9EPS2\|PYY_MOUSE | K.LLFTDDSDSENLPFRPEGLDQW | 0.000694444 | 77 | 98 | |

**Supplemental Table S5.** Peptidomics results from *Candida* infection.

| Protein Accession | Peptide | Group Profile (Ratio) | Start | End | PTM |
|---|---|---|---|---|---|
| P26350\|PTMA_MOUSE | K.EVVEEAENG.R | 1.15741E-05 | 22 | 30 | |
| P26350\|PTMA_MOUSE | R.VAEDDEDDDVDT.K | 1.15741E-05 | 91 | 102 | |
| P26350\|PTMA_MOUSE | R.DAPANGNAQNEENGEQEA.D | 1.15741E-05 | 32 | 49 | |
| P31786\|ACBP_MOUSE | K.TQPTDEEM(+15.99)LFIYSHF.K | 1.00:0.27:7.32 | 18 | 32 | Oxidation (M) |
| P31786\|ACBP_MOUSE | K.QATVGDVNTDRPGLL.D | 0:1.00:28.31 | 34 | 48 | |
| P31786\|ACBP_MOUSE | V.GDVNTDRPGLL.D | 1.00:0:40.48 | 38 | 48 | |
| P31786\|ACBP_MOUSE | K.QATVGDVNTDRPGLLDL.K | 0.000694444 | 34 | 50 | |
| P31786\|ACBP_MOUSE | K.TQPTDEEML.F | 1.15741E-05 | 18 | 26 | |
| P31786\|ACBP_MOUSE | A.TVGDVNTDRPGLL.D | 1.15741E-05 | 36 | 48 | |
| P33680\|GUC2A_MOUSE | G.VTVQDGDLSFPLESV.K | 1.15741E-05 | 22 | 36 | |
| P33680\|GUC2A_MOUSE | H.SALPEALRPV.C | 1.15741E-05 | 73 | 82 | |
| Q02819\|NUCB1_MOUSE | I.NVLETDGHF.R | 1.00:4.96:50.76 | 59 | 67 | |
| Q02819\|NUCB1_MOUSE | R.LVTLEEFLA.S | 1.15741E-05 | 310 | 318 | |
| Q6ZWY8\|TYB10_MOUSE | K.ETIEQEK.R | 1.00:6.39:23.11 | 33 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | E.TIEQEK.R | 1.00:0.70:12.47 | 34 | 39 | |
| Q6ZWY8\|TYB10_MOUSE | M.A(+42.01)DKPDMGEIA.S | 1.15741E-05 | 2 | 11 | Acetylation (N-term) |
| Q9EPS2\|PYY_MOUSE | E.NLPFRPEGLDQW | 1.15741E-05 | 87 | 98 | |

# Chapter 9

# Conclusions and Future Directions

**Conclusions**

This thesis describes the development and application of mass spectrometry techniques, including mass spectrometry imaging (MSI), to study symbiotic relationships. In Chapter 3, an atmospheric pressure (AP)- matrix assisted laser desorption/ionization (MALDI)- MSI method was developed for metabolite imaging and then applied to study the effects of salt stress in *Medicago truncatula* (Medicago) root nodules. The AP-MALDI Orbitrap platform used here enables high resolution imaging for both mass and space, as it has a smaller minimum spatial resolution than the traditional vacuum MALDI Orbitrap instrument. Although the AP-MALDI platform resulted in fewer detected features compared to the traditional vacuum MALDI platform, the detected features were highly complementary, indicating that the combined use of both MALDI platforms would enable identification of the highest number of small molecules. For the salt stress analysis, 44 *m/z* ions were present at relatively higher abundances in the control samples, and 77 *m/z* were enriched in the salt samples. LC-MS/MS was used for identification of enriched ions. In particular, arginine was identified in the salt nodules.[1] Arginine is often observed as upregulated in response to stress and may help the plant tolerate the salt stress.[2, 3]

To investigate the important roles of endogenous peptides in plants, a MALDI-MSI method was developed in Chapter 4 to image peptides in the root nodules of Medicago. By using both DHB and CHCA matrices, 277 peptides and protein fragments in root nodules and 496 peptides and protein fragments in seedling plants were imaged. Wild-type plants were compared to mutant lines that overexpressed a specific class of endogenous peptides for method development. They were also compared to mutants that lack a class of endogenous peptides, which resulted in non-functional nodules. Differences in peptide content were observed between wild-type root nodules and the non-functional mutant root nodules and between seedling plants and mature root nodules.

Through LC-MS/MS 22 peptides were *de novo* sequenced and 10 peptides and protein fragments were identified.[4]

As the identified peptides in Chapter 4 appeared to be protein fragments rather than endogenous peptides, the MALDI-MSI workflow was optimized for detection of nodule-specific cysteine-rich (NCR) peptides in Chapter 5. NCR peptides are involved in bacteria differentiation in the root nodule and are important for formation of functional nodules capable of biological nitrogen fixation. As these peptides are fairly large (3-5 kDa), multi-step washing procedures to remove lipid species, which interfere with ionization of larger molecules, were found to be beneficial. The optimal wash was determined to be 70% ethanol, 30s; 100% ethanol, 30s; Carnoy's solution, 2 min; 100% ethanol, 30s. After washing, 71 ions were observed above *m/z* 2500, which is the approximate start of the NCR peptide mass range. Sixteen of these peaks were matched to NCR peptides via accurate mass matching, with three of the peptides also being confirmed via LC-MS/MS. These peptides were found to have differing distributions in the root nodules, including peptides localized to just the lower half of the root nodule, and peptides localized to the entire root nodule.

The other half of this thesis focused on studying the microbiome with multiomic strategies. In Chapter 6, three different extraction methods were compared, a methanol/ chloroform/ water extraction, an acidified methanol extraction, and a 40/40/20 acetonitrile/ methanol/ water extraction, for analysis of metabolites, peptides, and proteins in cecum content. The methanol/ chloroform/ water extraction performed the best for metabolomics and peptidomics analyses, while the acidified methanol extraction performed best for the proteomics analysis. The acidified methanol extraction was especially proficient in identifying microbiome proteins. Overall, the acidified methanol and methanol/ chloroform/ water extractions detected a similar number of total

molecules, but as the methanol/ chloroform/ water extraction's total compounds came from all three molecule types, rather than being predominantly proteins, that extraction was selected as the more desirable option. The gene ontology results for the peptidomics and proteomics results agreed well with each other, providing a deeper understanding of the biological processes in the gut.

The chloroform/ methanol/ water extraction was then used to study the small molecules involved in the microbiome's response to infection in Chapter 7, especially as it relates to the potential of the microbiome to produce natural products.[5] An approximately 90 strain model of the human gut microbiome was used in this study to create "humanized" mice from germ-free mice. Humanized, pathogen-infected mice were compared to humanized, uninfected mice and mice noncolonized with the pathogen. Two different infections, *Salmonella enterica* Typhimurium and *Candida albicans*, were investigated. Metagenomics analysis revealed changes in the gut microbiome composition after *Salmonella* infection, but few changed in the gut microbiome after *Candida* infection. Metabolomics analysis revealed 22 compounds up-regulated in the humanized, *Salmonella*-infected mice, and 10 compound up-regulated in the humanized, *Candida* infected mice, with one compound being shared in both infections. Of these, only 10 compounds had putative identifications via accurate mass matching (4 were confirmed with MS/MS), leaving the rest to be potentially novel compounds.[6]

To continue to study the microbiome's response to infection, Chapter 8 explored another multiomics study on the mice model system described in Chapter 7. In this chapter, proteomics and peptidomics were added to the metabolomics analysis. A new analysis strategy for metabolomics analysis resulted in additional metabolomic targets for analysis; there were 384 targets from the *Salmonella* infection and 83 targets from the *Candida* infection. Identification of these targets revealed multiple carnitine and adenosine compounds that were changing due to

infection. The peptidomics analysis found Fibrinopeptide B and Thymosin alpha that are present in the gut that could provide insight into microbiome mediated communication between the gut and the brain. Proteomics analysis determined that upregulated proteins were involved in immune function and inflammatory response. Thus, metabolomics, peptidomics, and proteomics analysis revealed a multi system microbiome-mediated response to infection.

**Future Directions**

In Chapter 3, one major challenge of the analysis of salt stress via MALDI-MSI was the abundance of sodium ions in the high salt sample. While the higher tolerance of MALDI to salts enabled data collection without a major reduction in signal, the presence of sodium in the high salt sample did make it challenging to identify changing metabolites. The increase in sodium levels altered the relative ratios of the different adducts between the low and high salt sample so that molecules with the same relative abundance between the low and high salt samples appeared higher in the low salt samples for the $[M+H]^+$ and $[M+K]^+$ adducts and higher in the high salt samples for the $[M+Na]^+$ adduct. By ensuring that the $[M+H]^+$, $[M+Na]^+$ and $[M+K]^+$ adducts all follow the same trend, this problem can be circumvented, but this analysis is time-consuming. In the future, alternative set-ups for studying salt stress could be very advantageous and lead to more identifications that are also more confident. For example, a similar stress can also be achieved through non-ionic osmotic stress, which is caused by mannitol or polyethylene glycol. Studying non-ionic osmotic stress may achieve more accurate results.

For the MALDI-MSI analysis of endogenous peptides in Medicago, the developed MALDI-MSI procedure for NCR peptide detection could be applied to study Medicago mutants. For example the Medicago mutant *dnf*1-1, forms non-functional nodules lacking certain

endogenous peptides and are therefore deficient in biological nitrogen fixation.[7] The developed MALDI-MSI procedure could be applied to compare wild-type and Medicago mutant *dnf*1-1 root nodules, which may provide insights into the role of certain NCR peptides. Additionally, for the best results, the LC-MS/MS sample preparation procedure should be optimized. In this work, many NCR degradation products were observed, resulting in five potential full-length NCR peptides detected via LC-MS/MS. Potentially, degradation could be minimized during the sample preparation steps through use of a device called Denator, to deactivate proteases, on the root nodules prior to extraction, or through use of acid or protease inhibitors in the extraction solvent to minimize degradation of these NCR peptides.

Finally, MALDI-MSI analysis of Medicago root nodules is not limited to small molecules and peptides. MALDI-MSI methods could also be developed to study the spatial distributions of proteins and glycans on the tissue. Protein analysis by MALDI-MSI is very common.[8] Smaller proteins (less than approximately 20-30 kDa) can be analyzed intact with MALDI-ToF mass spectrometers. Additionally, on-tissue trypsin application will digest proteins on the tissue section, allowing for protein detection in a bottom-up strategy.[9] Bottom-up LC-MS/MS studies are usually performed in parallel to aid in protein identification in MALDI-MSI. For glycans, PNGaseF enzyme can also be applied on-tissue to cleave N-glycans.[10]

The multiomics analysis described in the microbiome work could be integrated further with pathway analysis, especially the metabolomic, peptidomic and proteomic studies. Multiomic integration is challenging, but various software platforms exist to aid in the process for integration of data from metabolomics and proteomics analyses.[11-13] Increased integration of the data could provide a more detailed and complete picture of the biological processes occurring the gut microbiome. Furthermore, increased metabolomics identifications would also improve the

multiomics analysis for the multiomics study in Chapter 8 describing the microbiome's response to infection. Finally, for future microbiome studies, it would be best to lower the starting *m/z* from 200 to 100. Initially, the higher starting *m/z* was selected to focus on secondary metabolites and natural products, but as the study shifted more toward general metabolomics, the higher starting *m/z* became disadvantageous. Other study designs for the microbiome study could also be employed. For example, time course studies looking at metabolite changes as the infection progresses could yield insightful information about the small molecules produced by the microbiome during disease progression. Additionally, other pathogens could also be investigated.

Finally, a MALDI-MSI method to analyze the spatial distribution of small molecules in the gut could help to discover and confirm small molecules produced by the microbiome. As described in Appendix VI, imaging of the mucosal layer, where the microbiome resides, is proving to be challenging. However, by combining a staining method or fluorescence in situ hybridization, which will help to localize the microbiome, with MALDI-MSI, small molecules could be directly localized to the microbiome.[14] Although this is potentially a challenging project, it has the potential to be very impactful due to the localization capabilities.

In summary, this dissertation research focuses on developing and applying MS techniques to investigate two symbiotic relationships. MALDI-MSI strategies were developed to study the plant *Medicago truncatula*. Also, LC-MS strategies were applied to investigate metabolites, peptides, and proteins in the microbiome. I developed methods to study challenging biological questions, including low concentration signaling peptides and complex relationships between the microbiome, host, and pathogenic infection. The methods developed here could also be widely applied to a variety of other biological systems.

**References**

1. Keller, C.; Maeda, J.; Jayaraman, D.; Chakraborty, S.; Sussman, M. R.; Harris, J. M.; Ané, J.-M.; Li, L., Comparison of Vacuum MALDI and AP-MALDI Platforms for the Mass Spectrometry Imaging of Metabolites Involved in Salt Stress in Medicago truncatula. *Frontiers in Plant Science* **2018,** *9*, 1238.

2. Rare, E., Stress physiology: The functional significance of the accumulation of nitrogen-containing compounds. *Journal of Horticultural Science* **1990,** *65* (3), 231-243.

3. Kakkar, R. K.; Bhaduri, S.; Rai, V. K.; Kumar, S., Amelioration of NaCl stress by arginine in rice seedlings: changes in endogenous polyamines. *Biologia Plantarum* **2000,** *43* (3), 419-422.

4. Gemperline, E.; Keller, C.; Jayaraman, D.; Maeda, J.; Sussman, M. R.; Ane, J. M.; Li, L., Examination of Endogenous Peptides in Medicago truncatula Using Mass Spectrometry Imaging. *J Proteome Res* **2016,** *15* (12), 4403-4411.

5. Sharon, G.; Garg, N.; Debelius, J.; Knight, R.; Dorrestein, P. C.; Mazmanian, S. K., Specialized metabolites from the microbiome in health and disease. *Cell Metab* **2014,** *20* (5), 719-730.

6. Bratburd, J. R.; Keller, C.; Vivas, E.; Gemperline, E.; Li, L.; Rey, F. E.; Currie, C. R., Gut Microbial and Metabolic Responses to Salmonella enterica Serovar Typhimurium and Candida albicans. *MBio* **2018,** *9* (6).

7. Starker, C. G.; Parra-Colmenares, A. L.; Smith, L.; Mitra, R. M.; Long, S. R., Nitrogen fixation mutants of Medicago truncatula fail to support plant and bacterial symbiotic gene expression. *Plant Physiol* **2006,** *140* (2), 671-80.

8. Chaurand, P.; Norris, J. L.; Cornett, D. S.; Mobley, J. A.; Caprioli, R. M., New developments in profiling and imaging of proteins from tissue sections by MALDI mass spectrometry. *Journal of Proteome Research* **2006,** *5* (11), 2889-2900.

9. Taverna, D.; Norris, J. L.; Caprioli, R. M., Histology-Directed Microwave Assisted Enzymatic Protein Digestion for MALDI MS Analysis of Mammalian Tissue. *Analytical Chemistry* **2015,** *87* (1), 670-676.

10. Powers, T. W.; Jones, E. E.; Betesh, L. R.; Romano, P. R.; Gao, P.; Copland, J. A.; Mehta, A. S.; Drake, R. R., Matrix Assisted Laser Desorption Ionization Imaging Mass Spectrometry Workflow for Spatial Profiling Analysis of N-Linked Glycan Expression in Tissues. *Analytical Chemistry* **2013,** *85* (20), 9799-9806.

11. Caspi, R.; Billington, R.; Fulcher, C. A.; Keseler, I. M.; Kothari, A.; Krummenacker, M.; Latendresse, M.; Midford, P. E.; Ong, Q.; Ong, W. K.; Paley, S.; Subhraveti, P.; Karp, P. D., The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res* **2018,** *46* (D1), D633-d639.

12. Zhou, G.; Xia, J., OmicsNet: a web-based tool for creation and visual analysis of biological networks in 3D space. *Nucleic Acids Research* **2018,** *46* (W1), W514-W522.

13. Zhou, G.; Xia, J., Using OmicsNet for Network Integration and 3D Visualization. *Current Protocols in Bioinformatics* **2019,** *65* (1), e69.

14. Geier, B. K.; Sogin, E.; Michellod, D.; Janda, M.; Kompauer, M.; Spengler, B.; Dubilier, N.; Liebeke, M., Spatial metabolomics of in situ, host-microbe interactions. *bioRxiv* **2019**, 555045.

# Chapter 10

# Investigation of Symbiotic Relationships through Mass Spectrometry: Application to Plant/Bacteria Symbiosis

Written for the Wisconsin Initiative for Scientific Literacy in order to describe this thesis for a broader audience.

Keywords: Symbiosis, *Medicago truncatula*, microbiome, metabolites, peptides

**Symbiotic relationships**

The diversity of life forms on Earth is truly astounding. Organisms can range from tiny single cells not visible to the naked eye to massive multi-organ animals. Across all levels of life, from the tiniest microorganisms (microbes) at 1/1000$^{th}$ of a millimeter or less in size to complex organisms weighing tons, relationships exist. These relationships can involve food, such as predator-prey relationships, but do not have to be harmful to one member of the relationship. Multitudes of organisms exhibit mutually beneficial, or symbiotic relationships.

My research explores two very different types of relationships. In one case, Rhizobia bacteria invade the plant *Medicago truncatula* (Medicago) to facilitate the exchange of nutrients between the two organisms. In the other, trillions of microbial cells inhabit the human intestines; these cells are commonly referred to as the gut microbiota and greatly impact human health. While the microbiota residing in our bodies are typically beneficial, disturbances in the microbiota community are associated with a number of harmful diseases. In this chapter, I will focus solely on my work on the Medicago, Rhizobia symbiotic relationships.

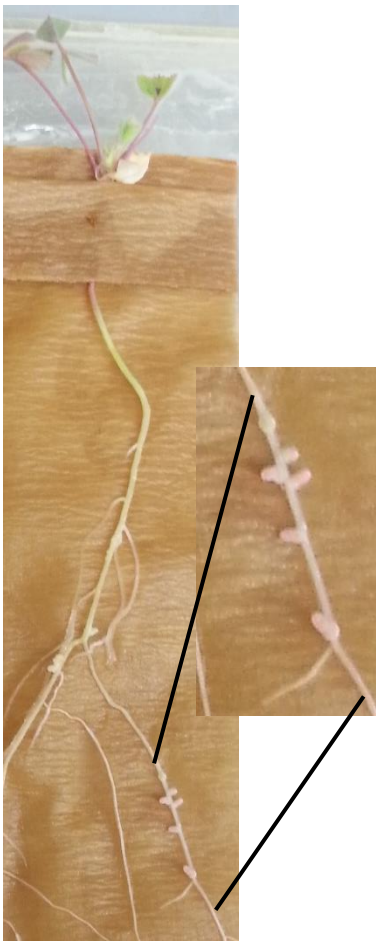


**Figure 1.** Picture of a Medicago Plant used for this study. An enlarged section of roots containing root nodules is also shown.

*Medicago truncatula*-rhizobia Symbiosis

While atmospheric nitrogen ($N_2$) is the most abundant gas in the atmosphere, only a handful of organisms, namely certain bacteria and archaea, which are single-celled organisms without a defined nucleus, can use the $N_2$ in the atmosphere for their nitrogen needs by converting $N_2$ into ammonia ($NH_3$), which is the form of

nitrogen most organism require. This process is called biological nitrogen fixation. To meet crops' nitrogen demand, farmers use nitrogen fertilizers, which increase available nitrogen.[1] Fertilizers require considerable resources to manufacture,[2] and fertilizer leaching contributes to environmental pollution through release of greenhouse gases.[3] Also, fertilizer usage can lead to eutrophication of waterways, where excess nutrients results in an overgrowth of algae, depleting oxygen and resulting in the death of aquatic life. Thus, it is beneficial to use biological nitrogen fixation in agriculture to reduce fertilizer usage.[4] This is accomplished through plants that form a symbiotic relationship with nitrogen fixing bacteria. For example, legume plants (i.e. peas, beans, soybeans) can form a symbiotic relationship with rhizobia bacteria, which can perform biological nitrogen fixation. The rhizobia bacteria provide fixed nitrogen for the plant, and in exchange, the plant provides a protected environment and a carbon source for the rhizobia bacteria.[5] The use of legume plants in agriculture reduces the requirements for nitrogen fertilizers.[6, 7]

*Medicago truncatula* (plant)-*Sinorhizobium meliloti* (bacteria) is a model legume-rhizobia symbiotic relationship capable of performing biological nitrogen fixation.[8] When rhizobia bacteria are near Medicago plants, they detect chemical signals produced by the plant and then secrete their own chemical signals, which starts a series of signaling events in the plant to from specialized organs on the plant roots, called root nodules (**Figure 1**), that provide an environment for rhizobia to live in.[9, 10] My research develops and applies analytical methods involving an instrument called a mass spectrometer to further understand the molecules present in the root nodules during biological nitrogen fixation.

**Mass Spectrometry**

When studying complex systems, mass spectrometry (MS) is a powerful tool, as it can detect a wide variety of biomolecules present in a system. These biomolecules include proteins, peptides, and metabolites, among others. Proteins are encoded for in our genes and play a wide variety of roles in the cells, such as structural support and transport and storage. Peptides are often signaling molecules throughout our body (insulin, for example), but can have other roles, such as defense. Metabolites are produced by proteins and also have a wide variety of biological roles, including energy, defense, structure, and signaling. My research primarily focuses on studying the metabolite and peptide content in plant tissue via mass spectrometry, including developing methods to more effectively analyze these molecules.

MS measures the mass to charge ratio (*m/z*) of a charged molecule (either a positively or negatively charged ion). Ions are created when uncharged molecules either lose electrons, which are negatively charged particles, creating a net positive electrical charge, or gain electrons, creating a net negative electrical charge. The general set-up of a mass spectrometer is (1) ion source (2) mass analyzer and (3) detector (**Figure 2**). The ion source creates charged molecules in the gaseous phase, allowing ions to enter the mass spectrometer. The ion source is important because uncharged ions are not able to travel through various parts of the mass spectrometer. In the mass analyzer, the ions are separated based upon their mass to charge ratio (*m/z*). Thus, a molecule with +1 charge will have an *m/z* twice as large as a molecule of the same mass but with a +2 charge. The detector monitors the ions as they are separated by the mass analyzer and sends signals produced by the ions hitting the detector to a computer to create a mass spectrum. The mass spectrum plots intensity (y axis) for all the *m/z* (x axis). An example mass spectrum is shown in **Figure 2**.

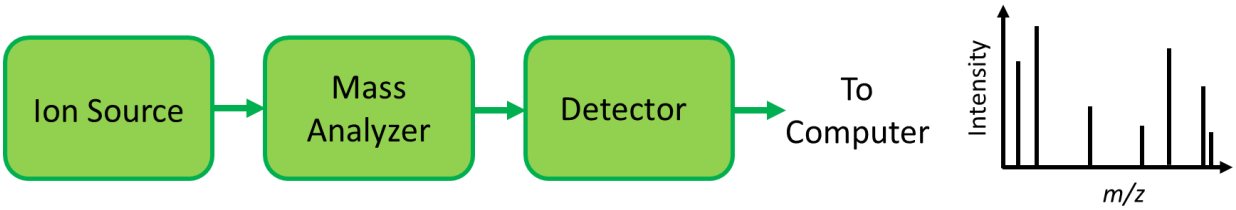



**Figure 2.** Basic diagram depicting the set-up of a mass spectrometer.

The basic workflow for analyzing the molecular content of a tissue involves homogenizing the tissue, *i.e.* making the sample uniform in nature, and extracting molecules from the tissue into an appropriate solvent. However, during the homogenization procedure, the spatial information of where the molecule was located in the tissue is lost. Mass spectrometry imaging (MSI) analyzes tissue sections directly in order to preserve the spatial locations of the molecules within the tissue. Although there are multiple different ionization techniques (i.e. different ion sources for creating ions from the tissue sections), my research utilizes a specific ionization technique utilizing a laser to selectively ionize specific locations of the sample at a time. The general workflow for MSI is shown in **Figure 3**. The sample is sliced into thin sections (16 µm, about ¼ the width of a human hair) and a small molecule, called the matrix, is applied uniformly across the tissue section to aid in ionization. After inserting the sample into the instrument, a box is drawn around the area to be imaged and the box divided into pixels. To acquire the image, a laser is fired at each pixel across the matrix covered tissue section. This generates a mass spectrum at each pixel. After acquisition, software is used to extract the ion intensity of specific *m/z* from each mass spectrum to create a heatmap for the *m/z* across the tissue section. Different distributions can be observed for different *m/z* (**Figure 3**). In this way, hundreds of different molecules can be imaged in a single run. First, I will discuss my efforts to improve MSI methods to study symbiosis in the plant Medicago, then I will discuss the application of developed methods to study the plant's response to stress.

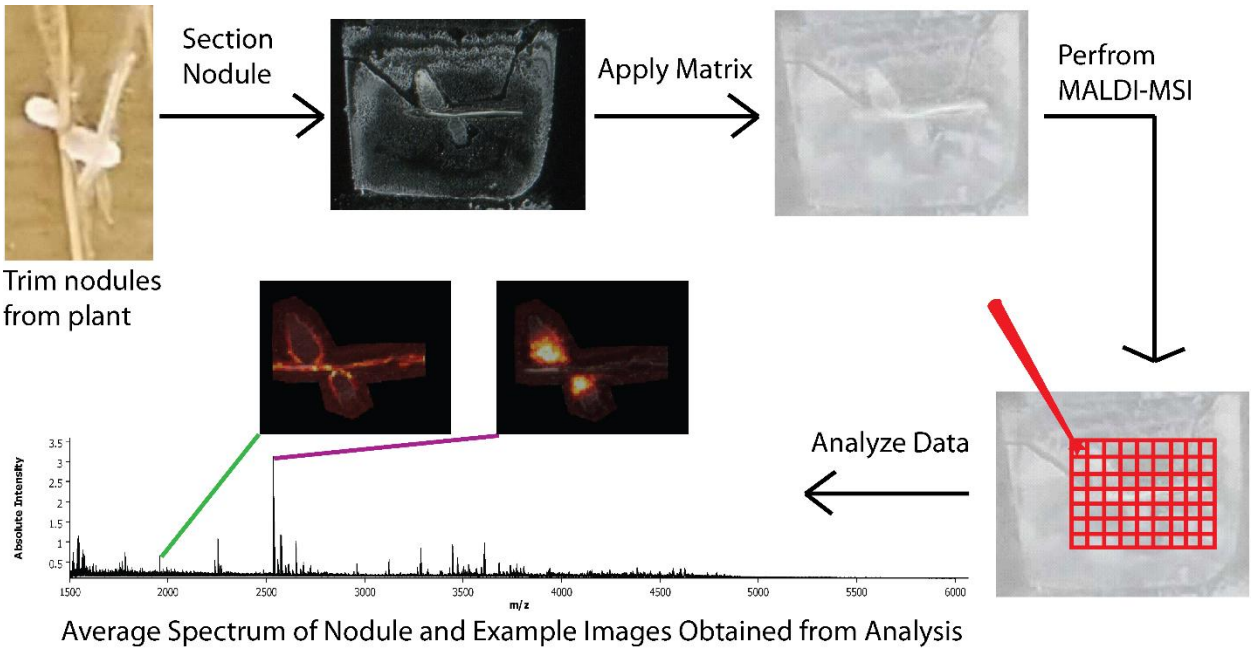



**Figure 3. Typical workflow for MSI on Medicago root nodules.**

## Experimental Results- Developing a MSI Method to Study Peptides in *Medicago truncatula* Rhizobia Symbiosis

My research uses mass spectrometry to not only study symbiotic relationships, but also to develop and improve methods that increase our ability to study symbiotic relationships. Previously, our lab developed a MSI method to image metabolites in the root nodules of Medicago.[11] Since then, MSI of plant metabolites has become increasingly more prevalent. Peptides, however, are much less frequently studied by MSI despite their importance in plants. Signaling peptides, in particular have important roles in communication within the plant, but they are difficult to investigate with MSI due to the fact that they are present at very low concentrations. The various roles of signalling peptides in plants include growth, development, and symbiosis.[12] Thus, I set out to develop a MSI method to image signalling peptides in Medicago root nodules, specifically focusing on a class of peptides that are critical for proper formation of root nodules

capable of performing biological nitrogen fixation. To accomplish this, I altered the sample preparation workflow in **Figure 3**. Traditionally, metabolite MSI sample preparation involves sectioning the tissue section and then matrix application. I added a washing step in between the sectioning step and the matrix application step in the MSI workflow (**Figure 4**). By washing tissue sections with organic solvents (ethanol, chloroform, for example), interfering molecules are removed. Removing these species increases our ability to detect peptides. **Figure 4** shows the increased signal of larger peptides after washing.

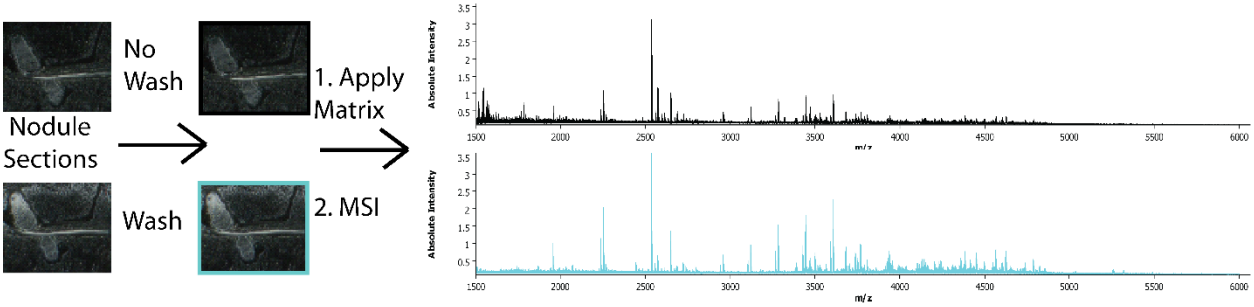



**Figure 4.** Optimization of MALDI-MSI for detection of larger endogenous peptides in plants.

**Experimental Results: Studying *Medicago truncatula* Rhizobia Symbiosis in Salt Stress**

Although basic understanding of symbiotic relationships is important, additional knowledge of what occurs when these relationships are disturbed is highly relevant as well. Stressors can cause harm to either or both organisms present in symbiotic relationships. One example of an environmental stressor is soil salinity. For example, in Medicago, high levels of salt in the soil has major effects on the root nodules, despite the relatively minor effects of salt stress on other aspects of plant growth. Due to the importance of biological nitrogen fixation in the root nodules, it is important to understand why salt has such a relatively large impact on root nodules.

To study salt stress, I utilized a MSI method based upon a new ionization source to study the metabolites in root nodules. The source used here to study salt stress allows for easier and

faster insertion of the sample into the instrument. The employed source also allows for higher spatial resolution (decreased pixel size) due to the smaller laser diameter. Higher spatial resolution allows for better resolution of smaller features in the tissue, in a similar way that higher resolution cameras better resolve images.

The developed method was applied to study the metabolites present at high salt and low salt conditions. **Figure 5** shows example MSI results for small molecule changes due to salt stress in Medicago. Each image in **Figure 5** shows control (low salt) nodules on the left side and high salt (salt) nodules on the right side. **Figure 5 (A,E)** provide the picture of the root nodule sections (optical image) prior to MSI analysis. The MSI images are shown on a hot intensity scale, which goes from black to red to yellow. Thus, the area on the tissue with the highest amount of the molecule will be shown in yellow. Example *m/z* whose distribution shows a higher amount in control root nodules **Figure 5 (B-D)** and salt root nodules **Figure 5 (F-H)** are shown. This analysis reveals which molecules are being altered in salt stress, providing insights into how the plant adapts to the stress.

**Conclusions and Future Directions**

I have discussed my research focusing on determining molecular distribution changes in Medicago by using MSI to image *m/z* in the plant. On the one hand, I have developed a MALDI-MSI method to investigate higher molecular weight peptides in Medicago root nodules and developed an MSI method to study metabolites in the Medicago root nodules. However, I have also applied developed methods to study small molecule changes due salt stress in the root nodules. In the future, my method to study peptides in the root nodule, could be applied to better understand the role of peptides in root nodules by comparing wild-type root nodules to root nodules from a

non-functional mutant version of Medicago. Comparing peptides observed in functional and non-functional nodules could determine peptides that are critical for forming nodules capable of biological nitrogen fixation. This would improve understanding of the needed molecule components for formation of legume-rhizobia symbiosis.

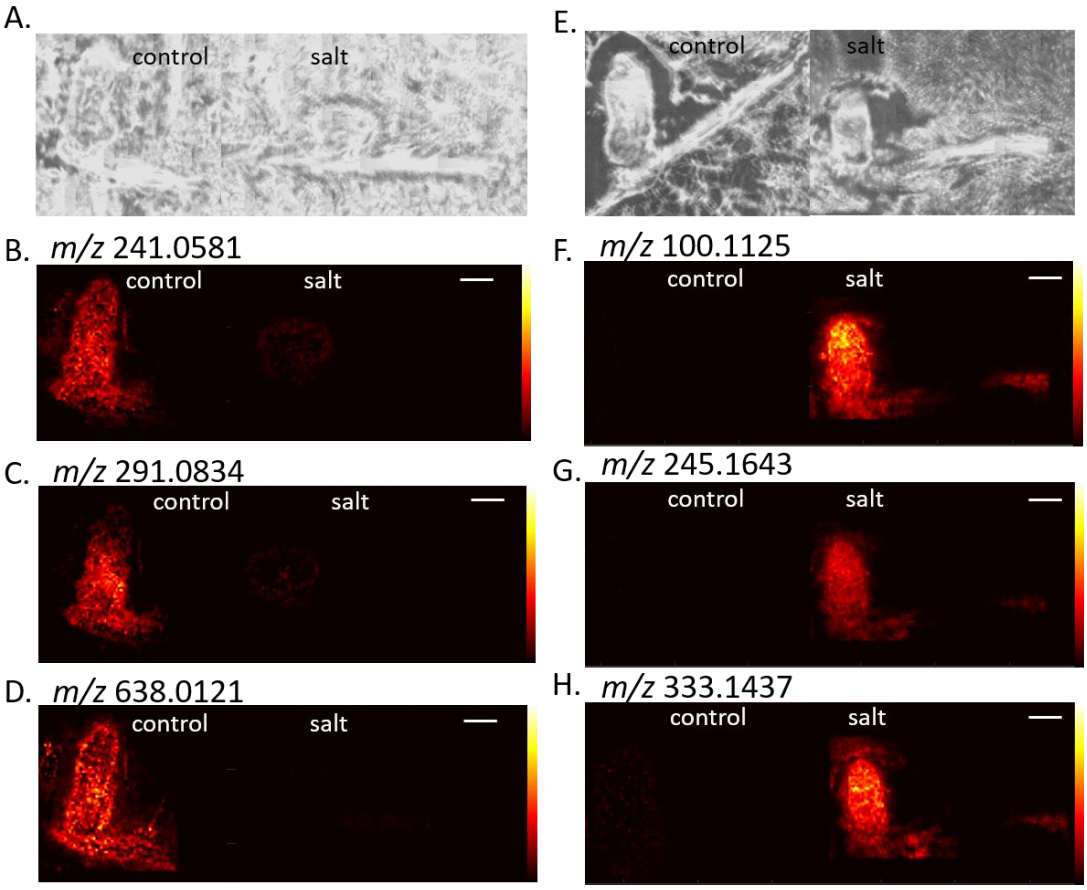



**Figure 5.** Example MALDI-MSI images for *m/z* changing during salt stress. **(A,E)** are the optical images and **(B-D)** are *m/z* higher in control nodules, while **(F-H)** are *m/z* higher in salt nodules. The white scale bar indicated 1 mm.

## References

1. Mueller, N. D.; Gerber, J. S.; Johnston, M.; Ray, D. K.; Ramankutty, N.; Foley, J. A., Closing yield gaps through nutrient and water management. *Nature* **2012,** *490* (7419), 254-7.

2. Jensen, E. S.; Peoples, M. B.; Boddey, R. M.; Gresshoff, P. M.; Hauggaard-Nielsen, H.; Alves, B. J. R.; Morrison, M. J., Legumes for mitigation of climate change and the provision of

feedstock for biofuels and biorefineries. A review. *Agronomy for Sustainable Development* **2012,** *32* (2), 329-364.

3. Gresshoff, P. M.; Hayashi, S.; Biswas, B.; Mirzaei, S.; Indrasumunar, A.; Reid, D.; Samuel, S.; Tollenaere, A.; van Hameren, B.; Hastwell, A.; Scott, P.; Ferguson, B. J., The value of biodiversity in legume symbiotic nitrogen fixation and nodulation for biofuel and food production. *J Plant Physiol* **2015,** *172*, 128-36.

4. Graham, P. H.; Vance, C. P., Legumes: importance and constraints to greater use. *Plant Physiol* **2003,** *131* (3), 872-7.

5. Haag, A. F.; Arnold, M. F.; Myka, K. K.; Kerscher, B.; Dall'Angelo, S.; Zanda, M.; Mergaert, P.; Ferguson, G. P., Molecular insights into bacteroid development during Rhizobium-legume symbiosis. *FEMS Microbiol Rev* **2013,** *37* (3), 364-83.

6. Kopke, U.; Nemecek, T., Ecological services of faba bean. *Field Crops Research* **2010,** *115* (3), 217-233.

7. Zentner, R. P.; Campbell, C. A.; Biederbeck, V. O.; Miller, P. R.; Selles, F.; Fernandez, M. R., In search of a sustainable cropping system for the semiarid Canadian prairies. *Journal of Sustainable Agriculture* **2001,** *18* (2-3), 117-136.

8. Mus, F.; Crook, M. B.; Garcia, K.; Costas, A. G.; Geddes, B. A.; Kouri, E. D.; Paramasivan, P.; Ryu, M. H.; Oldroyd, G. E. D.; Poole, P. S.; Udvardi, M. K.; Voigt, C. A.; Ane, J. M.; Peters, J. W., Symbiotic Nitrogen Fixation and the Challenges to Its Extension to Nonlegumes. *Applied and Environmental Microbiology* **2016,** *82* (13), 3698-3710.

9. Via, V. D.; Zanetti, M. E.; Blanco, F., How legumes recognize rhizobia. *Plant Signal Behav* **2016,** *11* (2), e1120396.

10. Ferguson, B. J.; Indrasumunar, A.; Hayashi, S.; Lin, M. H.; Lin, Y. H.; Reid, D. E.; Gresshoff, P. M., Molecular analysis of legume nodule development and autoregulation. *J Integr Plant Biol* **2010,** *52* (1), 61-76.

11. Gemperline, E.; Li, L., MALDI-mass spectrometric imaging for the investigation of metabolites in Medicago truncatula root nodules. *J Vis Exp* **2014,** (85).

12. Tavormina, P.; De Coninck, B.; Nikonorova, N.; De Smet, I.; Cammue, B. P., The Plant Peptidome: An Expanding Repertoire of Structural Features and Biological Functions. *Plant Cell* **2015,** *27* (8), 2095-118.

# Appendix 1

# MALDI Mass Spectrometry Imaging of Peptides in *Medicago truncatula* Root Nodules

Adapted from: Keller C, Gemperline E, Li L. MALDI Mass Spectrometry Imaging of Peptides in *Medicago Truncatula* Root Nodules. ("Plant Proteomics Methods and Protocols," 3rd edition Invited book chapter, submitted)

Keywords: MALDI, MSI, peptides, *Medicago truncatula*, imaging

**Abstract**

Mass spectrometry imaging (MSI) is routinely used to visualize the distributions of biomolecules in tissue sections. In plants, MSI of metabolites is commonly observed, but the imaging of larger molecules is less frequently performed despite the importance of proteins and endogenous peptides to the plant. Here, we describe a matrix-assisted laser desorption/ionization (MALDI) MSI method for the imaging of peptides in *Medicago truncatula* root nodules. Sample preparation steps, including embedding in gelatin, sectioning, and matrix application are described. The method described is employed to determine the spatial distribution of hundreds of peptide peaks.

**Introduction**

Matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) is a powerful tool to visualize the spatial distribution of molecules in a tissue.[1] In MALDI-MSI, a laser is fired at discrete positions, or pixels, across a matrix-covered tissue. At each pixel, a mass spectrum is collected. Once the instrument collects mass spectra at all of the pixels, software programs extract the ion intensity for a particular *m/z* across all pixels to create an image, or heatmap, for that *m/z*. In this way, hundreds of images can be generated from a single instrument run. To prepare a sample for analysis, the general sample preparation steps are flash freezing and embedding, sectioning, and applying a suitable matrix. Sample preparation is a critical step to preserve the sample and to achieve good signal of the chosen analytes.[2, 3] For example, the matrix coating, which assists in ionizing analyte molecules in the tissue section, can influence the type of analytes in a sample that will ionize and the spatial resolution of the imaging experiment. MALDI-MSI has been applied to many different analyte types, including metabolites,[4, 5] neuropeptides,[6]

and proteins[7] in many different organisms. However, applications of the technique to plants have focused on small molecules,[8] with only a few focusing on larger molecules.[9-12]

Here, we provide a detailed protocol focusing on applying MALDI-MSI to investigate peptides present in the root nodules of *Medicago truncatula* (Medicago. [9] Medicago forms specialized organs, called root nodules, on its roots as a result of a symbiotic relationship with rhizobia bacteria for biological nitrogen fixation. Plant peptides are involved in the formation of the nodule on the roots of the plant, as well as in plant growth and development.[13, 14] For example, nodule-specific cysteine-rich peptides are involved in the differentiation of bacteria into bacteroids in the root nodules,[15] and CLAVATA3/embryo-surrounding region (CLE) peptides are involved in autoregulation of nodulation.[16, 17] Thus, the protocol here aims to provide a method that can be used to determine the spatial distribution of plant peptides via MALDI-MSI to further our understanding about these important biomolecules.

## Materials

Embedding Nodules

1. Plant material: *Medicago truncatula* plants inoculated with *Sinorhizobium meliloti* (Rm1021)
2. Embedding Media: 100 mg/mL gelatin
3. Plastic embedding containers suitable for storage in the -80ºC
4. Dry ice

MALDI-MSI Sample Preparation

1. Optimal cutting temperature (OCT) compound
2. 25 x 75 mm glass slides
3. 50% Methanol: HPLC-grade methanol, MilliQ water (50:50 v:v)

4. 50% Methanol 0.1% FA: HPLC-grade methanol, MilliQ water (50:50 v:v), 0.1% formic acid (FA)
5. DHB matrix solution: 40 mg/mL 2,5-dihydroxybenzoic acid (DHB) in 50% methanol 0.1% FA. Sonicate the matrix until completely dissolved.
6. 50% Acetonitrile: HPLC-grade acetonitrile, MilliQ water (50:50 v:v)
7. 50% Acetonitrile 0.1% FA: HPLC-grade acetonitrile, MilliQ water (50:50 v:v), 0.1% formic acid
8. CHCA matrix solution: 5 mg/mL α-cyano-4-hydroxycinnamic acid (CHCA) in 50% acetonitrile 0.1% FA. Sonicate the matrix until completely dissolved.
9. SA matrix solution: 5 mg/mL sinapic acid (SA) in 50% acetonitrile 0.1% FA. Sonicate the matrix to completely dissolve it.

**Methods**

Sample preparation for MALDI-MSI is a critical step to obtain the best results during the MALDI-MSI analysis. Sample preparation steps, such as sample preservation, washing, matrix choice, and matrix application method will all influence the sample analysis. Here, we describe sample preparation steps for flash freezing the nodules and embedding in gelatin, followed by matrix application with a TM Sprayer automatic sprayer system (HTX Technologies). The sample is analyzed on the MALDI LTQ Orbitrap XL (Thermo Scientific) equipped with a nitrogen laser, and data analysis is performed in ImageQuest (Thermo Scientific) and MSiReader.[18] **Figure 1** demonstrates the sample workflow for MALDI-MSI of Medicago root nodules.

<u>Embedding Nodules</u>

1. Trim nodules from the plant with about 2-4 mm of surrounding roots (see note 1).

2. Place nodule in a plastic cup or similar holding container of appropriate size for your sample (for example a 5 mm x 5 mm x 5 mm square plastic cup for very small samples) with a drop of 100 mg/mL gelatin (see note 2).
3. Place on dry ice and wait for nodule and gelatin to freeze. The gelatin will turn white when frozen.
4. Once the nodule is frozen, fill the embedding container with 100 mg/mL gelatin. Wait for the entire embedding container with gelatin to freeze. Once the gelatin is completely white, the nodule can be stored at -80ºC (see note 3) prior to MSI analysis.

MALDI-MSI Sample Preparation

1. Take the embedded nodule and trim sample to rectangle with a couple mm of gelatin surrounding the sample on all sides. Do this quickly to minimize the time the sample is at room temperature.
2. Attach sample to a cryostat chuck with a drop of OCT compound (see note 4).
3. Allow sample on the chuck to equilibrate in the cryostat at -20ºC for 15 minutes.
4. Align the sample so that the cryostat is cutting sections evenly across the root and root nodule. This can be done by taking about five sections and adjusting the chuck if part of the sample is being missed. For plant root nodules, our lab typically uses 16 µm, but other section thicknesses between 8-35 µm,[19] can be used.
5. Once the center of the nodule (or other desired depth) is reached, thaw mount sections onto a glass slide by warming the back of the slide against your hand and then placing the front of the slide gently onto the tissue section.
6. Continue until desired number of sections across the z stack of the root nodule are obtained.

7. Keep the sections in a dry environment (i.e. dry box) while preparing the TM Sprayer for matrix application (see note 5).
8. Turn nitrogen gas on TM Sprayer to 10 psi, and the solvent pump to 0.25 mL/min. The solvent for the pump should match what the matrix is dissolved in (without the FA), so for DHB this would be 50% methanol and for CHCA this would be 50% acetonitrile. Turn on the TM Sprayer and laptop (see note 6).
9. Set the temperature on the software to the appropriate temperature for the desired solvent and TM Sprayer system (see note 7). As a starting point, 80$^{o}$C is the appropriate temperature for 50% methanol.
10. Load the dissolved matrix (i.e. DHB, CHCA, SA see note 8) into the sample loop with the knob in the load position.
11. Load the TM Sprayer method and manually change gas pressure and flow rate if method differs from the initial parameters of 10 psi and 0.25 mL/min. The TM Sprayer has recommended methods for specific matrices and analyte types, although method parameters may need to be optimized for a specific application. For DHB imaging of peptides, method parameters typically used are 1250 velocity, 0.1 mL/min, 12 passes, 30s dry time, rotate and offset (cc pattern), 10 psi, 80$^{o}$C. For imaging of peptides using CHCA and SA as matrices, method parameters to start from are 1100 velocity, 0.2 mL/min, 8 passes, 30 s dry time, rotate and offset (cc pattern), 10 psi, 85$^{o}$C (see note 9).
12. Once the TM Sprayer has reached the appropriate temperature, add slides containing sample to the sample holder. Secure slides in place as necessary to prevent movement during matrix application.

13. Switch the sample loop knob to the spray position. Once matrix is coming out of the nozzle, start the TM Sprayer program.
14. After the matrix application is finished, cool down the system while flushing with the solvent the matrix is dissolved in (for DHB, this would be 50% methanol) at 0.25 mL/min. Rinse the sample loop 3 times with solvent and toggle the knob. Once the system is below 50°C, the system can be turned off.
15. Store the sample in a dry box at -20°C if running on the instrument the following day.

MSI data acquisition on the MALDI LTQ Orbitrap XL

1. Place glass slide(s) with sample into the slide adapter. If importing the image of the glass slide, scan the slide in the adapter with a scanner. Then add the backing plate and insert the plate into the instrument. Alternatively, the slide can be scanned after inserting the plate into the instrument with the camera in the instrument (see note 10).
2. Open the plate image in the MALDI source dialog box in the Tune software. Zoom in as necessary to see sample, depending on sample size. Draw boxes around the areas to be imaged (see note 11). Save this as a MALDI position file. For MS1 imaging, using a rectangle box and raster motion works best. Also set the desired spatial step size (75 μm is the smallest raster size without oversampling).
3. In Xcalibur, set up the sequence by adding the file name, path location, instrument method, and MALDI position file. The instrument method contains parameters controlling the mass resolution, mass range, and centroid/profile data. The instrument method also requires a tune file, which controls the laser energy and the microscans (microscans/step is controlled in the instrument file). The microscans and microscans/step should match to ensure that one pixel is one mass spectrum in the data file.

4. Check the laser energy by shooting the laser on a matrix only area that is not being imaged and checking the signal level. You can adjust the laser energy in your tune file as necessary to get the optimal signal.
5. Start the sequence.

Data Processing

1. Once the data is collected, the data can be viewed in ImageQuest, or exported to another software program. To visualize the data in ImageQuest, use the average spectra within a selected area tool to view an average spectrum of a certain area of the sample. In the bottom window of ImageQuest, there should be a spectrum from the sample. **Figure 2** shows example spectra averaged over the nodules for peptide imaging results with DHB, CHCA, and SA matrices.
2. Look through the peaks in the collected spectrum, zooming in as appropriate, and when one wants to visualize the distribution of a certain peak in the tissue, select add new data set. Select the single dataset option with plot type Mass Range/TIC. Use the *m/z* for the mass range and select the desired tolerance window (i.e. 5 ppm). Repeat as necessary to visualize the *m/z* in the sample. Under the 2D tab, there are other color bar options as well as smoothing options.
3. To view in MSiReader [18], export the data in ImageQuest into an imzML format, keeping the data in profile.
4. Load the imzML file into MSiReader and select the desired mass tolerance, image smoothing, and color bar parameters. Insert a *m/z* that is localized across the sample to visualize the sample (one can find a good *m/z* for this in ImageQuest). Normalize to the total ion count (TIC). To pull out *m/z* unique to the sample, use the polygon tool to create interrogated and reference

zones. Outline around the sample to create an interrogated zone, then create a matrix only region for the reference zone.

5. Use the extract peaks unique to the interrogated zone tool to create a list of *m/z* present in the image. One will need to set percentage numbers for the threshold a *m/z* needs to be above in the interrogated zone and the threshold a *m/z* needs to be below in the reference zone to be added to the list. Also set the algorithm for peak centroid calculation (typically parabolic centroid works well).
6. Once the list has been created, use the generate an image for each peak in a list tool to create images for all the *m/z*. Manually go through the images and remove any bad images (i.e. images that have signal in the matrix as well as the sample or do not appear to have any signal anywhere). **Figure 3** shows example MALDI-MSI images generated from peptide imaging of root nodules with either DHB or CHCA as the matrix. Different distributions across the root and root nodules are observed.

**Notes**

1. For best results, select nodules that are red in color and elongated rod in shape rather than round. These are the nodules in which the symbiosis is well developed.
2. To make the sectioning process easier, ensure that the nodule is as flat as possible with the root in line with the nodule. This will help to get both the root and the nodule in the same plane when sectioning.
3. If the nodule is not completely frozen when covered in gelatin, it will not stick to the bottom of the cup and instead will float up to the middle or top of the cup. This makes the nodule harder to find and may result in the positioning of the nodule being lost. After adding the

gelatin, the cup should be kept level while waiting for the rest of the gelatin to freeze. If the gelatin freezes at an angle, it will be harder to level the nodule while sectioning to get both the root and root nodule in a single section. Avoid air bubbles close to the nodule when adding the gelatin, as this also will make the nodules harder to section.

4. OCT compound is a polymeric species and will suppress analyte signal if it comes into contact with the sample. Thus, care should be taken to ensure that the compound does not come into contact with the sample or with the blade or stage of the cryostat.
5. After sectioning and before matrix application, washing steps to remove highly abundant lipid species can increase signal intensity and observed protein peaks.[20] For protein imaging, ethanol washes and potentially a Carnoy wash are typically used to remove the lipid species that can suppress protein signal. For endogenous peptide imaging, washes may (or may not) remove the target peptides, depending on the chemical properties of the peptides. Thus, care should be taken when using washing techniques with peptides to ensure that they are not being removed in the washing steps.
6. Here the TM Sprayer is used to apply the matrix to evenly across the sample. It is important that the matrix is applied in a homogenous manner at all points on the tissue so that matrix inhomogeneity does not skew the results. A matrix application method should be reproducible run-to-run to ensure that results remain consistent. Other automatic sprayers can be used (i.e. home-build or the Bruker ImagePrep). Other matrix application techniques include the airbrush and sublimation.[21] Airbrush application can be achieved easily with minimal expense, however, user-to-user variation can be high and reproducibility can be a challenge. Sublimation provides very small crystal size and good imaging results for metabolomics

studies, but due to the dry application, the method requires further re-crystallization steps for analysis of larger molecules (i.e. peptides and proteins).[22]

7. The temperature of the TM Sprayer should be about 5°C below the temperature at which the "puffing" sound starts. This sound indicates that the matrix is not being sprayed in a consistent manner. If run at a temperature when the solvent is "puffing" the matrix will not cover the sample homogeneously, which will negatively affect results.
8. There are many different matrices to choose from. DHB and CHCA are both common matrices and can be used for a variety of analytes. Other matrices may be used primarily for larger peptides and proteins (i.e. SA) or primarily for negative mode (i.e. 9-aminoacrilamide). Matrices other than DHB and CHCA may work well depending on your desired analyte.
9. If this method is too wet, you can cut the flow rate in half and double the number of passes to achieve the same matrix density, but with a drier spray.
10. The preferred scanning method depends on the sample and time considerations. For the nodules, scanning in with the camera on the instrument provides good alignment and image quality, but this takes 25 minutes per slide. For larger tissues, the scanner separate from the instrument works well and saves time.
11. To check the alignment of the image to the slide in the instrument you can click a point on the image and check the cursor position on the camera box on the tune page to see where the actual position is. It can also be helpful to check the outside of the boxes to ensure the sample is not being cut-off.

**Acknowledgements**

This work was supported in part by funding from the National Science Foundation (NSF) Division of Integrative Organismal Systems (IOS) RESEARCH PGR award #1546742, the University of Wisconsin-Madison Graduate School and the Wisconsin Alumni Research Foundation (WARF). The MALDI-Orbitrap and Q-Exactive instruments were purchased through an NIH shared instrument grant (NCRR S10RR029531). LL acknowledges a Vilas Distinguished Achievement Professorship and Charles Melbourne Johnson Professorship with funding provided by the WARF and University of Wisconsin-Madison School of Pharmacy.

**References**

1. Caprioli, R. M.; Farmer, T. B.; Gile, J., Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* **1997,** *69* (23), 4751-60.

2. Goodwin, R. J.; Pennington, S. R.; Pitt, A. R., Protein and peptides in pictures: imaging with MALDI mass spectrometry. *Proteomics* **2008,** *8* (18), 3785-800.

3. Buchberger, A. R.; DeLaney, K.; Johnson, J.; Li, L., Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal Chem* **2018,** *90* (1), 240-265.

4. Ye, H.; Gemperline, E.; Venkateshwaran, M.; Chen, R.; Delaux, P. M.; Howes-Podoll, M.; Ane, J. M.; Li, L., MALDI mass spectrometry-assisted molecular imaging of metabolites during nitrogen fixation in the Medicago truncatula-Sinorhizobium meliloti symbiosis. *Plant J* **2013,** *75* (1), 130-45.

5. Gemperline, E.; Jayaraman, D.; Maeda, J.; Ane, J. M.; Li, L., Multifaceted investigation of metabolites during nitrogen fixation in Medicago via high resolution MALDI-MS imaging and ESI-MS. *J Am Soc Mass Spectrom* **2015,** *26* (1), 149-58.

6. Chen, R. B.; Li, L. J., Mass spectral imaging and profiling of neuropeptides at the organ and cellular domains. *Analytical and Bioanalytical Chemistry* **2010,** *397* (8), 3185-3193.

7. Chaurand, P.; Norris, J. L.; Cornett, D. S.; Mobley, J. A.; Caprioli, R. M., New developments in profiling and imaging of proteins from tissue sections by MALDI mass spectrometry. *Journal of Proteome Research* **2006,** *5* (11), 2889-2900.

8. Lee, Y. J.; Perdian, D. C.; Song, Z.; Yeung, E. S.; Nikolau, B. J., Use of mass spectrometry for imaging metabolites in plants. *Plant J* **2012,** *70* (1), 81-95.

9. Gemperline, E.; Keller, C.; Jayaraman, D.; Maeda, J.; Sussman, M. R.; Ane, J. M.; Li, L., Examination of Endogenous Peptides in Medicago truncatula Using Mass Spectrometry Imaging. *J Proteome Res* **2016,** *15* (12), 4403-4411.

10. Poth, A. G.; Mylne, J. S.; Grassl, J.; Lyons, R. E.; Millar, A. H.; Colgrave, M. L.; Craik, D. J., Cyclotides associate with leaf vasculature and are the products of a novel precursor in petunia (Solanaceae). *J Biol Chem* **2012,** *287* (32), 27033-46.

11. Cavatorta, V.; Sforza, S.; Mastrobuoni, G.; Pieraccini, G.; Francese, S.; Moneti, G.; Dossena, A.; Pastorello, E. A.; Marchelli, R., Unambiguous characterization and tissue localization of Pru P 3 peach allergen by electrospray mass spectrometry and MALDI imaging. *J Mass Spectrom* **2009,** *44* (6), 891-7.

12. Grassl, J.; Taylor, N. L.; Millar, A. H., Matrix-assisted laser desorption/ionisation mass spectrometry imaging and its development for plant protein imaging. *Plant Methods* **2011,** *7*, 11.

13. Tavormina, P.; De Coninck, B.; Nikonorova, N.; De Smet, I.; Cammue, B. P., The Plant Peptidome: An Expanding Repertoire of Structural Features and Biological Functions. *Plant Cell* **2015,** *27* (8), 2095-118.

14. Batut, J.; Mergaert, P.; Masson-Boivin, C., Peptide signalling in the rhizobium-legume symbiosis. *Curr Opin Microbiol* **2011,** *14* (2), 181-7.

15. Van de Velde, W.; Zehirov, G.; Szatmari, A.; Debreczeny, M.; Ishihara, H.; Kevei, Z.; Farkas, A.; Mikulass, K.; Nagy, A.; Tiricz, H.; Satiat-Jeunemaitre, B.; Alunni, B.; Bourge, M.; Kucho, K.; Abe, M.; Kereszt, A.; Maroti, G.; Uchiumi, T.; Kondorosi, E.; Mergaert, P., Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* **2010,** *327* (5969), 1122-6.

16. Mortier, V.; Den Herder, G.; Whitford, R.; Van de Velde, W.; Rombauts, S.; D'Haeseleer, K.; Holsters, M.; Goormachtig, S., CLE Peptides Control Medicago truncatula Nodulation Locally and Systemically. *Plant Physiology* **2010,** *153* (1), 222-237.

17. Mortier, V.; De Wever, E.; Vuylsteke, M.; Holsters, M.; Goormachtig, S., Nodule numbers are governed by interaction between CLE peptides and cytokinin signaling. *Plant Journal* **2012,** *70* (3), 367-376.

18. Robichaud, G.; Garrard, K. P.; Barry, J. A.; Muddiman, D. C., MSiReader: An Open-Source Interface to View and Analyze High Resolving Power MS Imaging Files on Matlab Platform. *Journal of the American Society for Mass Spectrometry* **2013,** *24* (5), 718-721.

19. Qin, L.; Zhang, Y.; Liu, Y.; He, H.; Han, M.; Li, Y.; Zeng, M.; Wang, X., Recent advances in matrix-assisted laser desorption/ionisation mass spectrometry imaging (MALDI-MSI) for in situ analysis of endogenous molecules in plants. *Phytochem Anal* **2018,** *29* (4), 351-364.

20. Seeley, E. H.; Oppenheimer, S. R.; Mi, D.; Chaurand, P.; Caprioli, R. M., Enhancement of protein sensitivity for MALDI imaging mass spectrometry after chemical treatment of tissue sections. *Journal of the American Society for Mass Spectrometry* **2008,** *19* (8), 1069-1077.

21. Gemperline, E.; Rawson, S.; Li, L. J., Optimization and Comparison of Multiple MALDI Matrix Application Methods for Small Molecule Mass Spectrometric Imaging. *Analytical Chemistry* **2014,** *86* (20), 10030-10035.

22. Yang, J.; Caprioli, R. M., Matrix sublimation/recrystallization for imaging proteins by mass spectrometry at high spatial resolution. *Anal Chem* **2011,** *83* (14), 5728-34.

**Figures**

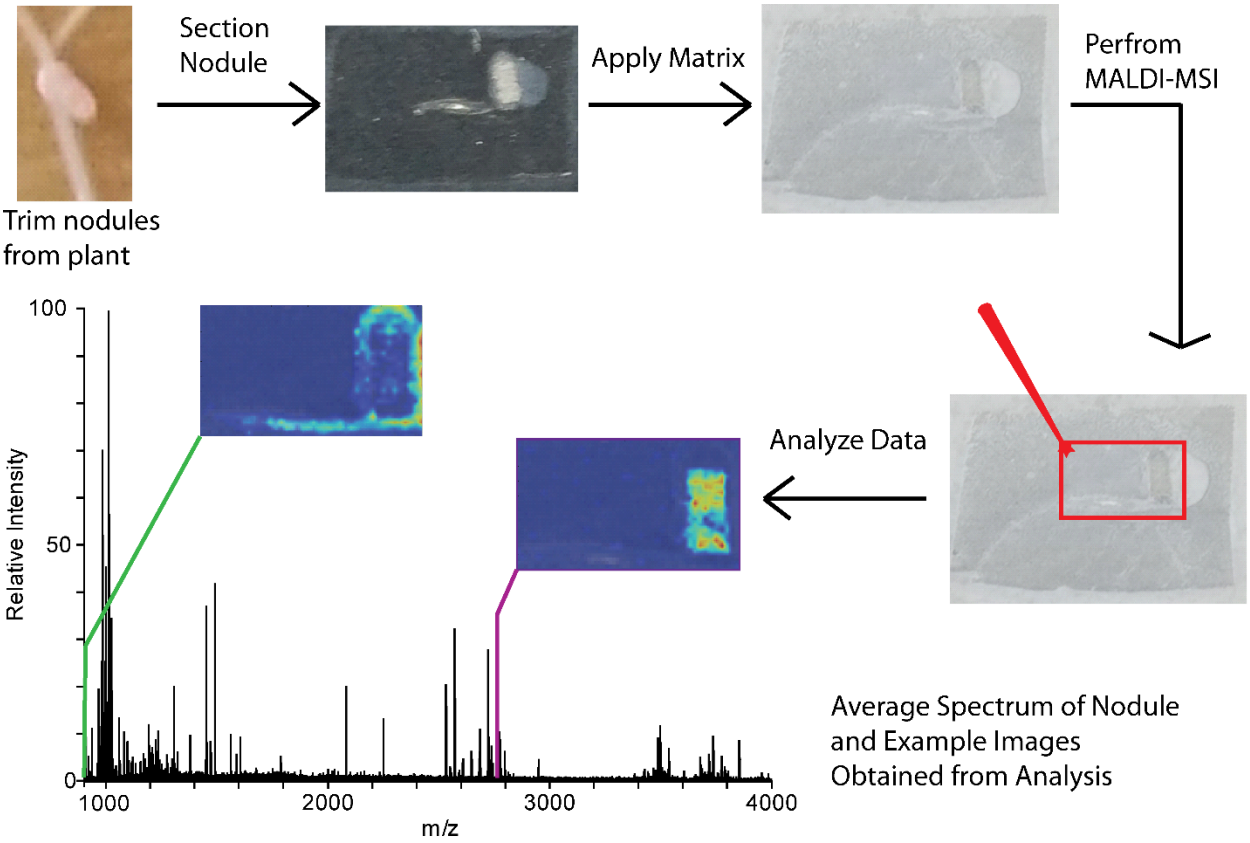



**Figure 1**. MALDI-MSI scheme showing the sample preparation, instrument analysis, and data analysis steps for a typical experiment.

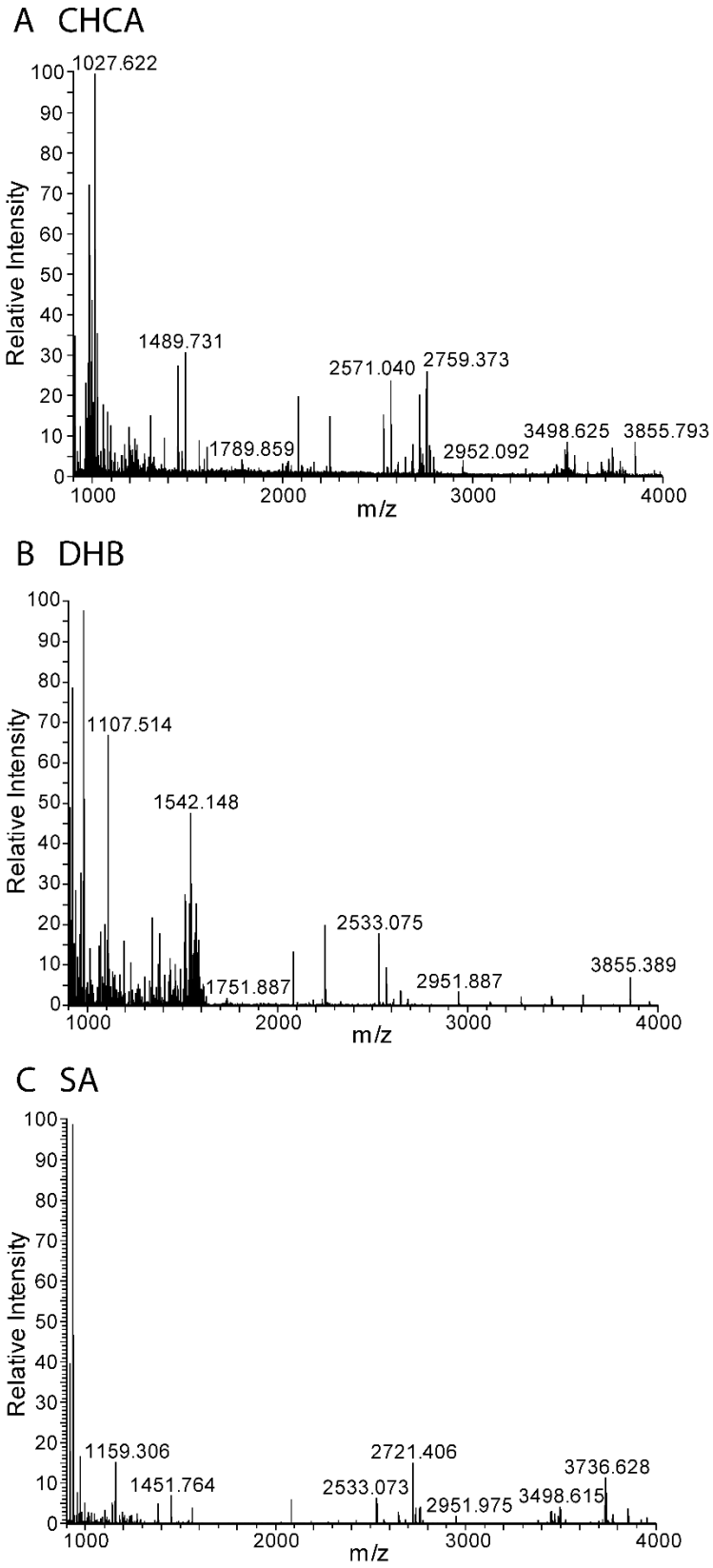



**Figure 2**. Example spectra average over the entire root nodules for MALDI-MSI on the root nodules with different matrices. The matrices are CHCA **(A)**, DHB **(B)**, SA **(C)**.

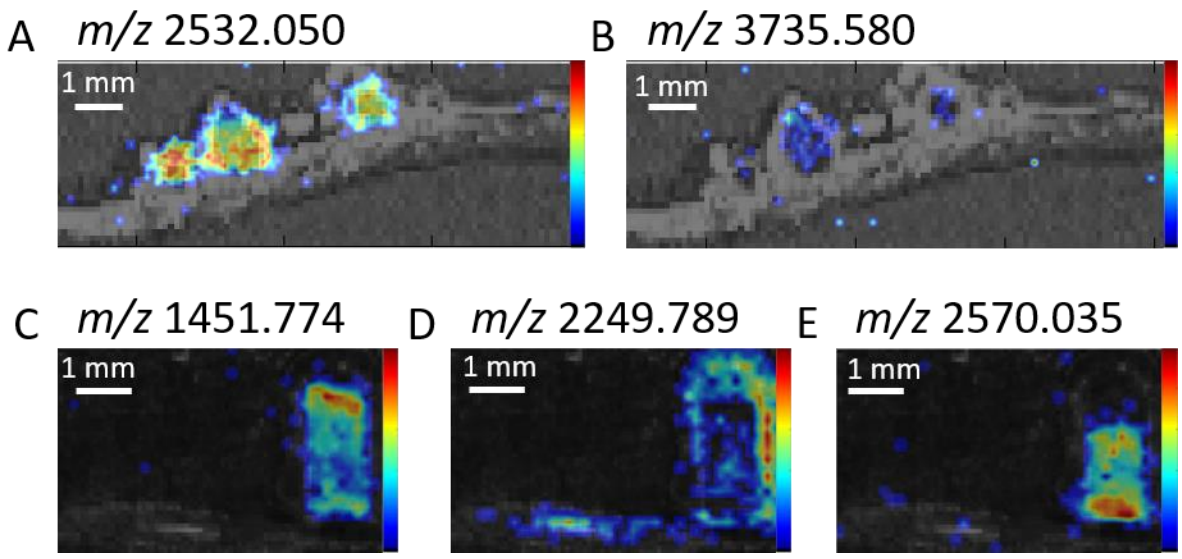



**Figure 3**. MALDI-MSI images of peptides with either DHB **(A, B)** or CHCA **(C, D, E)** as the matrix. The images are generated at +/- 5 ppm.

# Appendix II

# Strategies for Metabolite Identification in Liquid Chromatograph-Mass Spectrometry

Keyword: LC-MS; LC-MS/MS; metabolites; small molecules; Identification

**Abstract**

Metabolite identification is a challenging and time-consuming process in the metabolomics field. However, new software is constantly being developed to aid in the identification process. *In silico* fragmentation tools can be applied to a wide variety of molecules. Additionally, mass spectral databases are continually being expanded to include data on a greater variety of molecules, increasing the chances of identifying one's unknown molecule within the database. Here, I will discuss the small molecule identification strategies that are commonly used to identify unknowns in my metabolomics projects.

**Introduction**

Small molecules have a variety of functions, including metabolites with roles in signaling, defense, and other biological processes and also natural products produced by bacteria and plants with potential medicinal purposes.[1-4] The important roles of small molecules make them highly studied. Mass spectrometry (MS) is a useful technique to profile small molecules as it has lower sample requirements and higher sensitivity and specificity compared to other techniques, such as nuclear magnetic resonance (NMR) spectroscopy.[5] Liquid chromatography-mass spectrometry (LC-MS) allows for the profiling of hundreds to thousands of small molecule features in a sample extract, due to the assistance of chromatographic separation to help simplify complex systems. Consequently, LC-MS and LC-MS/MS have been used to profile small molecules in a wide variety of samples, including different types of bacteria, plants, and animals.[6, 7] LC-MS is also used for relative quantification of metabolites between different conditions and for absolute quantification of certain metabolite(s) in a sample.[8] A major limitation of LC-MS studies, however, is the identification process. Fragmentation patterns from tandem MS (MS/MS) data is necessary to

confidently identify metabolites from LC-MS data. As the small molecule ultimately needs to be confirmed by retention time and fragmentation matching of the experimental data to a pure standard for best identification confidence, small molecule identification is time-consuming in any metabolomics study. Here, various identification strategies for metabolomics studies with LC-MS are demanding and will be discussed.

**Figure 1** provides an overview of metabolite identification and an example workflow that I typically use for my research. From the raw data, the *m/z* and charge, or molecular weight, from data analysis software, and the MS/MS spectra can be obtained. Initially, searches against various databases to obtain a list of a few to potentially hundreds of compounds that match to the measured molecular weight. The MS/MS spectrum is used to narrow down the list of potential candidates to one, or a few, best matches. Typically, spectral databases and *in silico* fragmentation are used to analyze the MS/MS data. Once the MS/MS spectrum has been putatively annotated, the standard is obtained for the putative annotation, and the MS/MS spectrum and the retention time of the standard are compared to the experimental data to confirm identification.

The Metabolomics Standard Initiative has proposed a series of four levels for identification of metabolites via LC-MS/MS.[9] Level four is for unknown compounds, meaning the *m/z* is unidentified and unclassified. Level 3 describes putatively characterized to compound classes, meaning that a certain class of compound is identified by the data. Level 2 is for putatively annotated compounds, where publicly available or commercial databases are used to match experimental MS/MS spectra to database MS/MS spectra of metabolites. Leve l is given to identified compounds. To be confidently identified, the criteria dictates that 2 orthogonal methods are used to compare data from the authentic standards analyzed under the exact experimental conditions as how the experimental data were generated for unknown compounds. Thus, while

spectra matching or *in silico* fragmentation can putatively identify (or annotate) experimental *m/z*, confident identification should be made with an obtained standard with, for example, retention time matching and MS/MS spectral matching as the two orthogonal methods.

**Exact Mass Searches**

Initially, experimental *m/z* of interest can be searched against a variety of databases to acquire a list of potential compounds whose mass matches to the experimental *m/z* within the set tolerance. The tolerance window for this MS1 searching is dictated by the instrument accuracy. For example, Orbitrap instruments can typically have mass accuracy below 5 ppm, so searches can be done +/- 5 ppm. The wider the search window, the longer the list of compounds that match to the experimental *m/z* will be. Consequently, a high-resolution accurate mass (HRAM) instrument is very beneficial for identification purposes. The available databases for accurate mass matching vary in scope and size. Databases such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) compound database[10, 11] and Chemical Entities of Biological Interest (ChEBI) database[12] focus on biological small molecules across many species, while the human metabolome database (HMDB)[13] focuses solely on human small molecule metabolites. The ChemSpider database is a chemical structure database of 67 million structures, and is not limited to biological small molecules, meaning it will also include synthetic compounds and inorganic compounds.[14] The PubChem chemical database is an open chemistry database, where scientists can submit a variety of chemical substances and any corresponding data for others to use.[15] While the comprehensive databases include significantly more compound options for identification, many of these may not make sense to identify unknowns in a biological system, leading to false matches. The smaller biologically focused databases have more relevant compound to match to, but one may miss unique

compounds (*i.e* previously unidentified in a biological system). If a specific class of molecules are of interest, databases exist that focus only on that small subset. For example, there are many databases focusing on lipids, such as Lipid Maps and LipidBlast, among others. There are also natural product databases, such as Global Natural Products Social Molecular Networking (GNPS)[16] and Antibase. Databases can also be custom-built by a user to fit their special purposes. Selecting the database or databases to search is an important step of metabolite identification as it influences the molecules that can potentially be identified. Selecting a very large database can result in hundreds of molecules that match up to one *m/z*, which is then very hard to narrow down. However, too small of a database can result in missed identification or even incorrect identification. For example, the HMDB database would not be the best option to search against for identifying bacteria metabolites.

**Tandem MS**

Fragmentation information from experimentally obtained MS/MS spectra is used to refine the list of tens to hundreds of potential matches obtained from an accurate mass search of the *m/z*. To acquire MS/MS data with CID and HCD fragmentation, the *m/z* of interest are isolated, fragmented by collision with an inert gas, and the fragments detected in a mass analyzer. The MS/MS spectra can then be compared to database MSMS spectra of standard compounds, to *in silico* fragmentation, where fragmentation patterns for molecular are predicted, or to MS/MS spectra for pure standards of the molecules. In the next sections, method to narrow down the list of potential compounds to one (or a couple) potential identifications is described. Afterwards, pure standards can be obtained to confirm the putative annotation.

**Tandem MS Databases**

Mass spectra databases contain thousands to hundreds of thousands of MS/MS for standards compounds. Many databases exist for both LC-MS/MS and GC-MS analysis; however, the number of compounds with experimentally obtained MS/MS spectra can vary greatly among databases. Some of the previously mentioned databases for accurate mass searching, such as the HMDB database, also contain MS/MS information as well. Other mass spectral databases are only available commercially (NIST). More information on various LC-MS databases with spectral matching capabilities can be found in the literature.[17] Here, I will highlight three spectral databases that I have used in my research.

Metlin

The Metlin MS/MS metabolite database[18] started in 2003 and has grown to include over 300,000 molecular standards with MS/MS spectra and approximately 1 million total small molecules. *In silico* spectra are provided for many of the molecules without experimental MS/MS spectra. Metlin allows for simple searching, batch search of numerous *m/z* values, MS/MS spectra match searching, and fragment similarity searching, among other functionalites.[18] Multiple adducts in both positive and negative ion mode can be searched against. In batch searching mode, compounds matching to the *m/z* (or neutral weight if molecular weights from Compound Discoverer are used instead of *m/z*) are returned with whether or not MS/MS is present for that molecule. If MS/MS is present, the spectra can be viewed and compared to experimental spectra manually. One pitfall is that not all spectra provided in the batch searching are experimentally obtained, many are *in silico* predicted spectra, which are less reliable then experimentally obtained spectra. The experimental spectra in the Metlin database are acquired on various time of flight (ToF) mass spectrometers, and therefore may be slightly different to spectra obtained from other

mass analyzers. As the online formatting doesn't make it convenient for inserting into excel or another program for storage of one's results, manual inspection is slightly time-consuming. However, automatic searching of the Metlin database is available through XCMS software, along with other features.[19]

mzCloud

MzCloud (HighChem LLC, Slovakia) is a high-resolution spectral database of small molecule MS/MS obtained from orbitrap mass analyzers. The database includes 4,444,777 spectra from 16,750 compounds as of 3/21/19 and is constantly growing. Additionally, spectra trees are present for many molecules, showing $MS^n$ spectra (*ie*. $MS^2$, $MS^3$ or $MS^4$ spectra). Spectra are usually present for both CID and HCD ionization types under positive and/or negative modes at a wide variety of collision energies. For example, Metlin usually provides spectra at collision energies of 10, 20, and 40, whereas mzCloud may have spectra from collision energies 10-100 at 10 CE intervals. MzCloud has a wide range of molecular classes, ranging from endogenous metabolites to natural products/medicines to extractables and leachables. Spectra are contributed from various laboratories and are curated to ensure optimal quality. Raw data was filtered and recalibrated, and the filtered, recalibrated spectra are searchable. The use of HCD fragmentation is very useful for comparison to data obtained in the Li lab, which is usually obtained with HCD fragmentation. However, the mzCloud database is smaller than other databases, especially for endogenous metabolites.

In Compound Discoverer, batch searching of MS/MS information in one's LC-MS/MS data can be performed. Their scoring system provides the top scoring compound with its score. Usually compounds with scores above 85-90 (out of 100) are good matches. Care should be taken to look in the secondary mzCloud table to examine the second-best match. If two or more

compounds have similar MS/MS spectra, they will be very close in score, and may only be separated in score by 0.1. Even though mzCloud may call the compound that is 0.1 higher than the second identified, the identification of the first ranking compound is not confident in this case. In this case, it would be best to obtain the standards of the top two compounds to differentiate the two and decide which is the best match based on retention time or another orthogonal parameter.

MassBank of North America

The MassBank of North America (MoNA) mass spectral database is a public repository database for MS/MS of standards.[20] It auto-curates submissions and is designed to be a collaborative, centralized, platform for metabolites mass spectra. There are 77,296 unique compounds across the MoNA database as of (3/21/19). The databases (MoNA, LipidBlast, among others) are available to download from the Fiehn lab and can be used in conjunction with the mzDial software (also created from the Fiehn lab) for automatic search and score of one's experimental data to the database. There are 87,758 LC-MS/MS spectra, 60,555 of which are in positive mode and 27,185 of which are negative mode as of (3/21/19). Submitted spectra can be collected with a variety of mass analyzers and collision energies, which may affect comparison results when comparing one's own experimental spectra from a different instrument platform. The Fiehn lab has also recently added a HILIC library for LC-MS containing over 1000 standards with retention time and MS/MS information.

**In Silico Fragmentation**

As mass spectral databases do not currently cover all potential metabolites (let alone all potential small molecules), other strategies to identify small molecules are necessary. One strategy is to use fragment similarity matching to gain insight into potential functional classes or potential

structures present in the unknown molecule by matching experimental fragmentation patterns to similar fragmentation patterns observed in MS/MS spectra of known standards.[21] This spectra similarity matching is available with both the Metlin database (available online) and mzCloud database (through Compound Discoverer software) and is useful in situations where the experimental data fail to provide any direct matches to compounds in the database. Another strategy is to use *in silico* fragmentation, where fragmentation is predicted for molecules of interest, and then experimental spectra is compared to *in silico* predicted fragments. Multiple software platforms have been created for this task, including MetFrag,[22] CFM-ID,[23] and MSFinder.[24, 25] Different software will use different algorithms and may provide slightly different results. It should also be noted that the success rate of *in silico* prediction tools varies. As shown in **Figure 1**, not all fragments observed in the spectra will be explained by the *in silico* fragmentation, even in cases where the identification is correct. Most of the time, the correct hit can be observed in the top hits, but not always as the first, or best match to the data. This is demonstrated in the Critical Assessment of Small Molecule Identification (CASMI) contest, in which groups compete to identify multiple different MS/MS spectra using automatic methods, such as *in silico* fragmentation. In the 2016 results, the correct compound was in the top 10 rankings for the identification system applied about 50% of the time.[26] Consequently, user verification and manual inspection of all the top hits is necessary for *in silico* annotation. Good putative annotations describe almost all of the experimental fragments. The good putative annotations can then be verified by comparison to an authentic standard.

## Conclusions

Although metabolite identification and annotation have become bottleneck steps in metabolomics studies, current and novel software tools are constantly being developed to meet the needs of the metabolomics field. By applying a wide range of MS/MS identification tools, the most identifications can be made, assuming one has the time to put into some of the more user demanding techniques or to analyze one's data through multiple different LC-MS software platforms. The mzCloud, Metlin, and MoNA MS/MS spectral databases can be automatically searched using Compound Discoverer, XCMS, and msDial software, respectively. However, the list of putative annotations from the automatic search should still be manually inspected and if confident identification is required, an authentic standard is necessary. *In silico* fragmentation tools offer the advantage of being applicable on a wider range of molecules, including molecules that are not present in spectral databases. These tools, however, have varying success rates of identifying the correct molecule. In the end, the most confident identification strategy is to compare retention times and MS/MS to standards, but the above strategies can help to narrow down the list of possible candidates to one, or a couple, which saves time and money when obtaining authentic standards for comparison.

**Acknowledgements**

C.K would like to thank Erin Gemperline, whose mentorship in small molecule identification in the lab introduced C.K. to some of these techniques. Also, Qinjingwen Cao provided edits and her own insight into metabolite identification for this appendix.

**References**

1. Bero, J.; Quetin-Leclercq, J., Natural Products Published in 2009 from Plants Traditionally Used to Treat Malaria. *Planta Medica* **2011,** *77* (6), 631-640.

2. Perez-Victoria, I.; Martin, J.; Reyes, F., Combined LC/UV/MS and NMR Strategies for the Dereplication of Marine Natural Products. *Planta Med* **2016,** *82* (9-10), 857-71.

3. Luan, H. H.; Medzhitov, R., Food Fight: Role of Itaconate and Other Metabolites in Antimicrobial Defense. *Cell Metab* **2016,** *24* (3), 379-387.

4. Gemperline, E.; Chen, B. M.; Li, L. J., Challenges and recent advances in mass spectrometric imaging of neurotransmitters. *Bioanalysis* **2014,** *6* (4), 525-540.

5. Dunn, W. B.; Bailey, N. J. C.; Johnson, H. E., Measuring the metabolome: current analytical technologies. *Analyst* **2005,** *130* (5), 606-625.

6. Gemperline, E.; Jayaraman, D.; Maeda, J.; Ane, J. M.; Li, L., Multifaceted investigation of metabolites during nitrogen fixation in Medicago via high resolution MALDI-MS imaging and ESI-MS. *J Am Soc Mass Spectrom* **2015,** *26* (1), 149-58.

7. Cao, Q. J. W.; Ouyang, C. Z.; Zhong, X. F.; Li, L. J., Profiling of small molecule metabolites and neurotransmitters in crustacean hemolymph and neuronal tissues using reversed-phase LC-MS/MS. *Electrophoresis* **2018,** *39* (9-10), 1241-1248.

8. Hao, L.; Greer, T.; Page, D.; Shi, Y.; Vezina, C. M.; Macoska, J. A.; Marker, P. C.; Bjorling, D. E.; Bushman, W.; Ricke, W. A.; Li, L., In-Depth Characterization and Validation of Human Urine Metabolomes Reveal Novel Metabolic Signatures of Lower Urinary Tract Symptoms. *Sci Rep* **2016,** *6*, 30869.

9. Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R., Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007,** *3* (3), 211-221.

10. Kanehisa, M.; Goto, S., KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **2000,** *28* (1), 27-30.

11. Tanabe, M.; Kanehisa, M., Using the KEGG database resource. *Curr Protoc Bioinformatics* **2012,** *Chapter 1*, Unit1.12.

12. Hastings, J.; Owen, G.; Dekker, A.; Ennis, M.; Kale, N.; Muthukrishnan, V.; Turner, S.; Swainston, N.; Mendes, P.; Steinbeck, C., ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res* **2016,** *44* (D1), D1214-9.

13. Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vazquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.;

Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A., HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res* **2018,** *46* (D1), D608-d617.

14. Little, J. L.; Williams, A. J.; Pshenichnov, A.; Tkachenko, V., Identification of "known unknowns" utilizing accurate mass data and ChemSpider. *J Am Soc Mass Spectrom* **2012,** *23* (1), 179-85.

15. Gindulyte, A.; Shoemaker, B. A.; Yu, B.; He, J.; Zhang, J.; Chen, J.; Zaslavsky, L.; Thiessen, P. A.; Li, Q.; He, S.; Kim, S.; Cheng, T.; Bolton, E. E., PubChem 2019 update: improved access to chemical data. *Nucleic Acids Research* **2018,** *47* (D1), D1102-D1109.

16. Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W. T.; Crusemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderon, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C. C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C. C.; Yang, Y. L.; Humpf, H. U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; P, C. A. B.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodriguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P. M.; Phapale, P.; Nothias, L. F.; Alexandrov, T.; Litaudon, M.; Wolfender, J. L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D. T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Muller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. O.; Pogliano, K.; Linington, R. G.; Gutierrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N., Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* **2016,** *34* (8), 828-837.

17. Vinaixa, M.; Schymanski, E. L.; Neumann, S.; Navarro, M.; Salek, R. M.; Yanes, O., Mass spectral databases for LC/MS- and GC/MS-based metabolomics: State of the field and future prospects. *TrAC Trends in Analytical Chemistry* **2016,** *78*, 23-35.

18. Guijas, C.; Montenegro-Burke, J. R.; Domingo-Almenara, X.; Palermo, A.; Warth, B.; Hermann, G.; Koellensperger, G.; Huan, T.; Uritboonthai, W.; Aisporna, A. E.; Wolan, D. W.; Spilker, M. E.; Benton, H. P.; Siuzdak, G., METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Analytical Chemistry* **2018,** *90* (5), 3156-3164.

19. Huan, T.; Forsberg, E. M.; Rinehart, D.; Johnson, C. H.; Ivanisevic, J.; Benton, H. P.; Fang, M.; Aisporna, A.; Hilmers, B.; Poole, F. L.; Thorgersen, M. P.; Adams, M. W. W.; Krantz, G.; Fields, M. W.; Robbins, P. D.; Niedernhofer, L. J.; Ideker, T.; Majumder, E. L.;

Wall, J. D.; Rattray, N. J. W.; Goodacre, R.; Lairson, L. L.; Siuzdak, G., Systems biology guided by XCMS Online metabolomics. *Nature Methods* **2017,** *14*, 461.

20. Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T., MassBank: a public repository for sharing mass spectral data for life sciences. *J Mass Spectrom* **2010,** *45* (7), 703-14.

21. Benton, H. P.; Wong, D. M.; Trauger, S. A.; Siuzdak, G., XCMS2: processing tandem mass spectrometry data for metabolite identification and structural characterization. *Anal Chem* **2008,** *80* (16), 6382-9.

22. Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S., MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform* **2016,** *8*, 3.

23. Allen, F.; Pon, A.; Wilson, M.; Greiner, R.; Wishart, D., CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. *Nucleic Acids Res* **2014,** *42* (Web Server issue), W94-9.

24. Lai, Z.; Tsugawa, H.; Wohlgemuth, G.; Mehta, S.; Mueller, M.; Zheng, Y.; Ogiwara, A.; Meissen, J.; Showalter, M.; Takeuchi, K.; Kind, T.; Beal, P.; Arita, M.; Fiehn, O., Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics. *Nat Methods* **2018,** *15* (1), 53-56.

25. Tsugawa, H.; Kind, T.; Nakabayashi, R.; Yukihira, D.; Tanaka, W.; Cajka, T.; Saito, K.; Fiehn, O.; Arita, M., Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. *Anal Chem* **2016,** *88* (16), 7946-58.

26. Schymanski, E. L.; Ruttkies, C.; Krauss, M.; Brouard, C.; Kind, T.; Duhrkop, K.; Allen, F.; Vaniya, A.; Verdegem, D.; Bocker, S.; Rousu, J.; Shen, H.; Tsugawa, H.; Sajed, T.; Fiehn, O.; Ghesquiere, B.; Neumann, S., Critical Assessment of Small Molecule Identification 2016: automated methods. *J Cheminform* **2017,** *9* (1), 22.

## Figures

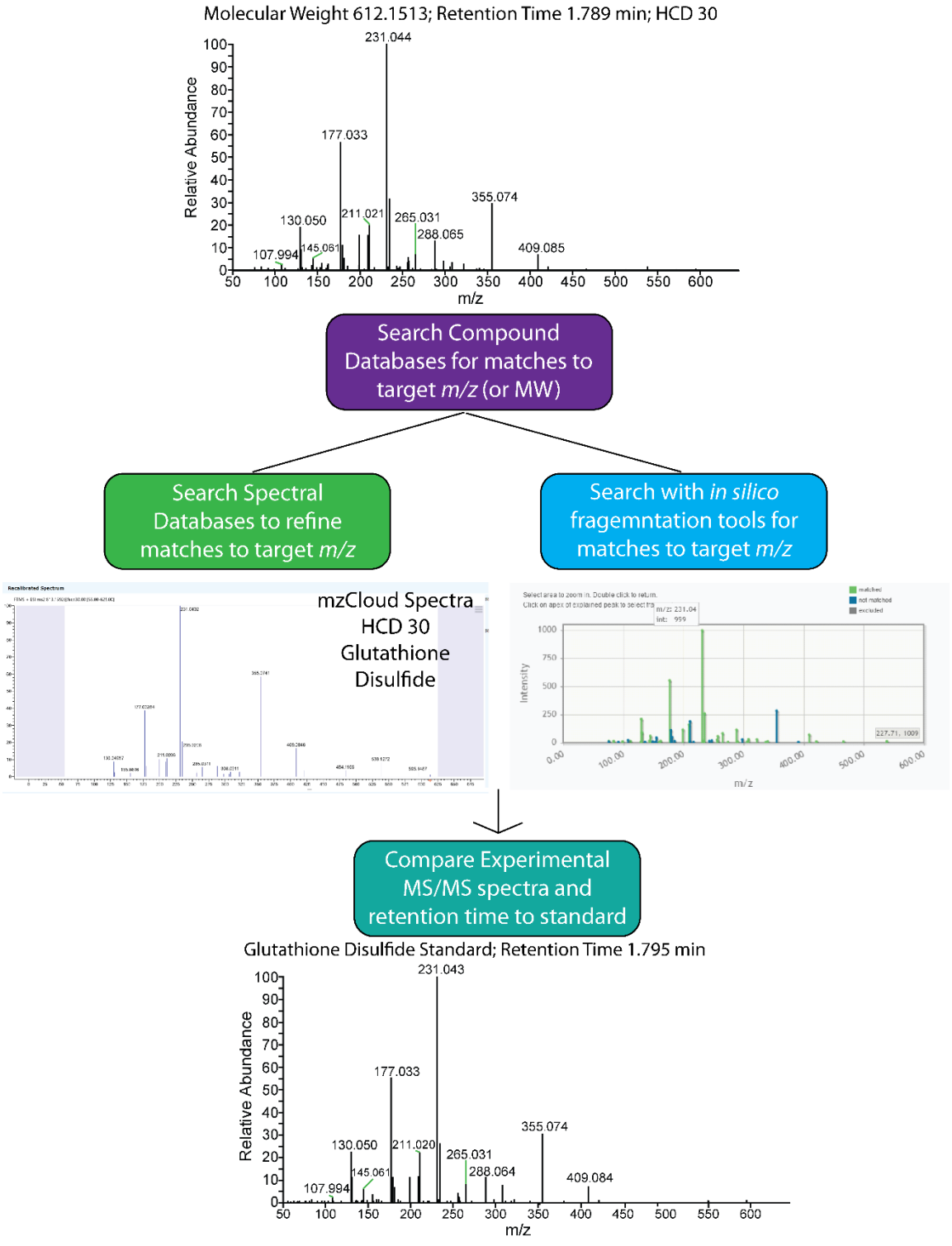



**Figure 1.** Overview of Metabolite Identification and Example Workflow for Identifying Glutathione Disulfide.

# Appendix III

# Multifaceted Investigation of Metabolites During Nitrogen Fixation in *Medicago truncatula* via High Resolution MALDI-MS Imaging and ESI-MS

Adapted from: Gemperline E, Keller C, Li L. Multifaceted Investigation of Metabolites During Nitrogen Fixation in *Medicago truncatula* via High Resolution MALDI-MS Imaging and ESI-MS. (“The Model Legume *Medicago truncatula*” Invited book chapter, Submitted).

Keywords: Nitrogen Fixation, *Medicago truncatula*, Metabolites, MALDI, Orbitrap, Mass Spectrometry, Imaging, Q-Exactive

**Abstract**

Legumes have developed the unique symbiotic relationship with soil bacteria known as rhizobia. This interaction results in the formation of specialized organs on the plants roots, called root nodules, in which rhizobia thrive and perform nitrogen fixation, a process in which the bacteria reduce atmospheric dinitrogen into a form of nitrogen (ammonium) that the plant can actually use for its own biological processes. Although metabolites are important in this symbiotic association, few studies have investigated the metabolites that are involved in the nitrogen fixation process. Here, we used a multifaceted mass spectrometric approach to detect and identify metabolites that are specifically present during nitrogen fixation using the *Medicago truncatula–Sinorhizobium meliloti* association as the model system. We used well-characterized plant and bacterial mutants to elucidate differences between the metabolites that are present in functional, nitrogen fixing nodules vs. non-functional nodules. Our study highlights the benefits of using a combination of mass spectrometric techniques for a more comprehensive look at the differences in metabolite composition and the distributions of these metabolites for biological studies.

**Introduction**

Nitrogen fixation, in which atmospheric dinitrogen ($N_2$) is converted to fixed forms such as ammonia ($NH_3$) or nitrate ($NO_3^-$), is critical to maintaining life. As an essential nutrient, nitrogen availability plays a key role in agriculture productivity and crop growth.[1, 2] However, most organisms are unable to utilize $N_2$, which is readily available in the atmosphere, and instead rely on fixed forms of nitrogen.[3] To provide the necessary amounts of nitrogen to crops, agriculture currently depends on the application of nitrogen fertilizers, which is possible through heavy fossil fuel usage.[4] Additionally, fertilizers contribute to environmental pollution through release of

carbon dioxide during fossil fuel combustion and through release of nitrous oxide, a significant greenhouse gas.[5] Furthermore, fertilizer leeching also contributes to eutrophication of waterways.[6] Although most organisms are not able to convert $N_2$ to $NH_3$ and $NO_3^-$, certain prokaryotes, including *Rhizobium* bacteria, use the enzyme nitrogenase to catalyze the transformation of $N_2$ to fixed forms, which is known as biological nitrogen fixation (BNF).[2] BNF is a key component of creating accessible nitrogen to support life, and can occur in both free-living and plant associated prokaryotes.[7]

Certain plants are capable of creating a highly efficient BNF process by forming a symbiotic relationship with nitrogen-fixing bacteria. Many legumes have the ability to form a symbiotic relationship with *Rhizobium* bacteria for the exchange of nutrients, and the *Medicago truncatula-Sinorhizobium meliloti* symbiosis is an example of a legume, rhizobia symbiosis.[8] The symbiosis process begins with the secretion of flavonoids and isoflavonoids by the plant, and when rhizobia recognize these molecules, the bacteria synthesize nod factors, or lipochitooligosaccharides.[9] The nod factors start a signaling cascade that results in rhizobia entry into the plant and the formation of specialized organs, called nodules, on the plant roots.[10] Rhizobia enter through an infection thread that forms from a root hair and breach the plant cell wall while simultaneously changes in the root cortical cells begin to form nodules.[11] In the nodule, rhizobia bacteria differentiate into bacteroids capable of performing BNF in the low oxygen environment provided by the nodule, which is important for nitrogenase activity.[11] Currently, many studies focus on understanding the legume/rhizobia symbiosis and on applying plant/bacteria symbiosis to enable a wider range of plants to meet some of their nitrogen requirement on their own.[8, 12]

Mass Spectrometry (MS) coupled to either gas or liquid chromatography (LC-MS/GC-MS) is a powerful tool for the analysis of complex mixtures of small molecules. GC-MS uses either

electron ionization or chemical ionization to ionize compounds coming off of the column for MS analysis,[13] whereas LC-MS uses electrospray ionization (ESI) to desolvate and ionize compounds prior to entry into the MS. MS allows for increased sensitivity in the detection of small molecules compared to other techniques, such as NMR.[14] Multiple GC-MS and LC-MS experiments have been conducted to study metabolites in *Medicago truncatula* (Medicago),[15-17] including metabolite changes in response to stress.[18, 19] Although these experiments allow the characterization of complex metabolite mixtures, valuable spatial information is lost during the extraction process.

Matrix-assisted laser desorption/ionization- mass spectrometry imaging (MALDI-MSI) allows for localization of hundreds of molecules in a tissue. In MALDI-MSI, a laser is fired on discrete points, or pixels, on a matrix-covered tissue section. Analyte crystallization with the matrix molecules on the tissue is critical for ionization in in MALDI-MSI as the matrix absorbs the laser energy and assists in producing analyte ions in the gas phase. Images are created by extracting the intensity for a given *m/z* value at each pixel on the sample. Thus, many images can be generated in a single imaging experiment.[20] This is an advantage over other imaging techniques that can only look at a few molecules at a time and require knowledge of the target molecule.[21] MSI is becoming increasingly used to study metabolites in a wide range of plants.[21] Time-of-flight (TOF), mass analyzers have historically been used in MALDI-MSI experiments, and has been used to map metabolites in Medicago.[22] However, FTICR and Orbitrap mass analyzers offer increased mass resolution compared to TOF mass analyzers, allowing for the differentiation of small *m/z* changes.[23] MALDI-MSI with an Orbitrap mass analyzer has been used to study soybean-aphid and rice-bacteria interactions,[24] phospholipids in *A. thaliana* seeds,[25] maize leaf sections,[26, 27] maize root sections,[28] and lipids in cotton embryos.[29]

Previous work from our lab utilized MALDI-MSI to study how the spatial distributions of endogenous peptides change during growth and maturation of *Medicago truncatula*.[30] Although MALDI-MSI analysis of plant metabolites has become increasingly prevalent, the use of MALDI-MSI to study plant peptides and proteins is still limited, with only a handful of examples.[21, 31-34] As endogenous peptides play key roles in the growth and development of Medicago, including autoregulation of nodules and bacteria differentiation in the nodules,[35] investigating the location of these peptides with MALDI-MSI could reveal insights about the roles of these molecules in the plant. In our study, we demonstrated the applicability of MALDI-MSI to study peptides in Medicago as we observed differences in the peptides detected and in the spatial distributions of detected peptides between seedling roots and mature roots and nodules.

Here, we discuss our use of both LC-MS and MALDI-MSI to gain a more thorough picture of the metabolites involved in biological nitrogen fixation.[36] As ESI and MALDI have different ionization mechanisms, they provide complementary coverage of molecules, allowing for increased coverage of the metabolome. MALDI-MSI studies were carried out with a high resolution MALDI-Orbitrap instrument with both wild-type and mutant plants and bacteria to better understand the metabolites involved in biological nitrogen fixation, which provides insight into the mechanism of biological nitrogen fixation. The *Medicago truncatula– Sinorhizobium meliloti* association is an excellent model for examining the nitrogen fixation process due to the availability of genetic information for both of the symbiotic partners.

**Nitrogen Fixation Metabolites Detected with MALDI-MS Imaging**

We previously used MALDI-MSI to study the localizations of different metabolites in wild-type Medicago root nodules. Clear ion images of metabolites localized to the Medicago root or

different zones in the nodule were observed and we were able to detect several amino acids and many key players in the citric acid cycle.[22] One exciting aspect of our study compared wild-type Medicago root nodules to the root nodules of several Medicago mutants that are not capable of nitrogen fixation. We identified one metabolite, Heme (part of the leghemoglobin protein which is responsible for regulation the oxygen environment in the plant during nitrogen fixation), that was present in the nitrogen fixing, wild-type nodules, and absent in the mutants. This result led us to delve more deeply into the identification of metabolites involved in the nitrogen fixation process in Medicago that is presented here.

One key advantage utilized in the presented study was the use of a high-resolution, accurate-mass (HRAM) MALDI-Orbitrap mass spectrometer. This instrument provides significant advantages for discovery metabolomics studies compared to the MALDI-ToF-ToF instrument used in the previous study that has modest mass spectral resolution, because accurate-mass measurements (defined as a mass error $\leq$ 5ppm) are essential for identifying the detected compounds. Many studies using MALDI-ToF-ToF instruments require accurate masses to be acquired using other instrumentation, typically ESI instruments, but there is no guarantee that the molecular species of interest that are detected by MALDI will also ionize with ESI. Additionally the high resolving power of the MALDI-Orbitrap also greatly increases the number of molecular species that can be detected because it can differentiate between compounds that separated in mass by as little as 0.003 Da. The capabilities of the HRAM MALDI-Orbitrap allows for increased detection of the Medicago metabolome, which can lead to greater insight into biological processes such as nitrogen fixation.

In addition to taking advantage of high-end mass spectrometer technology, this study utilized well-characterized mutants of both *Medicago truncatula* and its nitrogen-fixing symbiotic partner,

*Sinorhizobium meliloti*. The plant mutant (*dnf1*)[37, 38] and bacterial mutant (*fixJ*)[22] are deficient in nitrogen fixation capabilities. In order to identify metabolites that are relevant to the nitrogen fixation process, we compared wild-type Medicago with wild-type rhizobia (wt-wt) samples, that are capable of performing nitrogen fixation, to combinations of mutant samples, wild-type Medicago with fixJ rhizobia (wt-fixJ), dnf1 Medicago with wild-type rhizobia (dnf1-wt), and dnf1 Medicago with fixJ rhizobia (dnf1-fixJ), which are all incapable of nitrogen fixation. MALDI-MSI was performed on Medicago sections from all four sample types and ion images were extracted for all masses that were detected in the wt-wt samples but were absent in the mutant samples. Each image was manually confirmed as a metabolite, and not a MALDI matrix ion, and a list of "metabolites of interest" was generated from the *m/z* values that produced ion distributions in the wt-wt (functional) nodules but not in any of the three mutant (non-functional) nodules. Representative metabolites specifically detected in the wt-wt samples and not in the mutant samples are shown in **Figure 1**.

**Figure 1(A)** shows a photograph of each of the four sample types that were used in this study and thin sections or slices of the root nodules, coated with MALDI matrix and fixed onto a glass slide immediately prior to MALDI-MSI acquisition is shown in **Figure 1(B)**. The functional, wt-wt nodules have a larger, more-elongated shape in comparison to the non-functional, mutant nodules. Ten representative ion images of metabolites that were found only in the functional wt-wt samples can be seen in **Figure 1(C)**. These representative metabolites show different spatial distributions throughout the root and nodule portions of the plant. Understanding the spatial localization of these compounds provides further information about these metabolites and their role in nitrogen fixation.

**Nitrogen Fixation Metabolites Detected with ESI-MS**

In order to identify as many metabolites involved in the nitrogen fixation process as possible, we used second ionization type, ESI to complement the MALDI-MSI data. Since MALDI and ESI have different modes of ionization, they ionize some of the same classes of molecules and some different classes of molecules and thus are considered complementary techniques. The combination of the HRAM MALDI and ESI Orbitrap instrumentation provides novel advantages over previous studies of biological processes in plant models that utilize a single type of mass spectrometry. For this study, we grew additional Medicago plants and generated root nodule extracts for each of the four sample types listed above; these sample extracts were analyzed by LC-ESI-MS using a HRAM Q-Exactive mass spectrometer. Similarly to the MALDI-MSI samples, we wanted to focus on the metabolites that were detected via LC-MS in only the functional wt-wt nodules and were absent in the non-functional mutant samples. SIEVE software was used to compare three biological replicates of the four sample types, and a list of "metabolites of interest" was compiled. Representative data from SIEVE are shown in **Figure 2**.

The panels on the left show XIC (extracted ion chromatogram) peaks for metabolites that show at least 2-fold greater intensity in the wt-wt samples (blue) compared to any of the mutant samples (red, green, and yellow). The panels on the right display the information on the left in bar graph form, in which each bar represents a different biological replicate.

Overall, 90 "metabolites of interest" (metabolites detected in the functional, wt-wt nodules and absent in the non-functional, mutant nodules) were detected via MALDI-MSI, and 21 "metabolites of interest" were detected via LC-ESI-MS. Of the combined 111 "metabolites of interest" that were detected using both methods, only 7 were detected using both MALDI and ESI, thus demonstrating the complementary nature of these two different ionization techniques. See (Gemperline et al. 2015) for full list of "metabolites of interest".

**Identification of Nitrogen Fixation- relevant Metabolites**

Assigning confident molecular identifications for *m/z* values is typically quite challenging in discovery metabolomics studies. Searching the accurate mass information alone using the many metabolomics databases that are available online often results in tens to hundreds of possible identifications; therefore, information is also required in order to identify metabolites Accurate mass and the comparison of experimental MS/MS fragmentation patterns to theoretical (or *in silico*) fragmentation allows for putative identifications of the "metabolites of interest". For unambiguous metabolite identification, experimental data (accurate mass data, retention time, and MS/MS fragmentation patterns) should be compared with metabolite standards.

MS/MS experiments were conducted on the "metabolites of interest" from both the MALDI and ESI studies. MALDI is notoriously poor for collecting high quality MS/MS data and acceptable MS/MS data could not be collected for all of the "metabolites of interest" that were detected by MALDI-MSI. The *in silico* fragmentation database, MetFrag,[39] was used to search the accurate mass and MS/MS data. **Figure 3** presents representative results from MetFrag.

The peaks correspond to experimental MS/MS data, where the precursor ion (or the mass of the "metabolite of interest") is shown in blue and green indicates experimental peaks that matched to *in silico* fragmentation peaks and red indicates experimental peaks that did not match to *in silico* fragmentation peaks for the selected metabolite. The experimental mass spectra show good correlation with the *in silico* fragmentation data. Red peaks could be a result of noise/ background ions or other molecular species/isomers with the same or very similar *m/z* values that were also selected in the fragmentation window. A list of 34 annotated "metabolites of interest" is shown in **Table 1**.

We were able to identify some of the "metabolites of interest" as various organic acids and amino acids. The possible role of some of these metabolites in legume nodulation and nitrogen fixation has either been established or hypothesized. Excitingly, for most other identified "metabolites of interest", the roles of these molecules in regards to nodule development or nitrogen fixation have yet to be ascertained. The roles of some of the amino acids in nitrogen fixation have been previously established. For example, we identified asparagine, arginine, glutamic acid, and proline. Asparagine is involved in nitrogen cycling between various plant organs, and the fixed nitrogen in indeterminate nodules is exported as asparagine for use by the plant. In addition glutamine, glutamate, and arginine are thought to act as signals indicating the plant's nitrogen status, which in turn is hypothesized to regulate nodule growth and activity.[40-42]

Proline betaine (the N-methylated form of proline) has a significant known function in the symbiosis between *Medicago truncatula* and its nitrogen-fixing partner *Sinorhizobium meliloti*. Betaines induce nodulation (*nod*) genes in *Sinorhizobium meliloti* which allows the root nodules to form; furthermore, proline betaine acts as an energy source, in particular as a carbon and nitrogen source for *Sinorhizobium meliloti* in low-osmolarity environments, which enables the rhizobia to colonize legume roots efficiently and offers a competitive advantage against other soil bacteria that compete for plant carbon sources for colonization. Without root nodule development and inhabitation by rhizobia, nitrogen fixation could not take place. Additionally, proline and/or its N-methylated form may act as an important energy source for bacteria during early stages of symbiotic interaction and also prior to bacteroid differentiation in the root infection process.[43]

In addition to various amino acids and organic acids, we also detected Heme. Leghemoglobins are essential for maintaining microaerobic conditions in the nodule environment, which is required for nitrogen fixation to occur, and a heme moiety is a critical component of these leghemoglobins.

The enzyme nitrogenase, which is responsible for nitrogen fixation activity, contains Fe and Mo protein components, which are incompatible with free oxygen, indicating that nitrogenase can function efficiently only under microaerobic conditions.[44] Leghemoglobin monitors the oxygen level in the cytosol of infected cells and maintains optimum oxygen levels for the functioning of nitrogenase and also for bacterial respiration.

Furthermore, S-adenosylmethionine (SAM) was detected and identified in only nitrogen-fixation-capable root nodules. SAM is a known precursor of polyamine and ethylene biosynthesis and has implications in nitrogen fixation signaling by controlling ethylene production.[45, 46]

The roles of the majority of the metabolites we detected and identified from only the functional, wt-wt nodules are still unknown. The metabolite identifications along with some of the spatial information provided by MALDI-MSI could provide key insight to researchers who are studying the nitrogen fixation process in *Medicago truncatula* and other legume species.

**Acknowledgements**

This work was supported in part by funding from the National Science Foundation (NSF) Division of Integrative Organismal Systems (IOS) RESEARCH PGR award #1546742, the University of Wisconsin-Madison Graduate School and the Wisconsin Alumni Research Foundation (WARF), a Romnes Faculty Research Fellowship program to L.L. and a NSF grant to JMA (NSF#0701846). E.G. acknowledges an NSF Graduate Research Fellowship (DGE-1256259). The MALDI-Orbitrap and Q-Exactive instruments were purchased through an NIH shared instrument grant (NCRR S10RR029531).

**References**

1. Mueller, N. D.; Gerber, J. S.; Johnston, M.; Ray, D. K.; Ramankutty, N.; Foley, J. A., Closing yield gaps through nutrient and water management. *Nature* **2012,** *490* (7419), 254-7.

2. Oldroyd, G. E.; Dixon, R., Biotechnological solutions to the nitrogen problem. *Curr Opin Biotechnol* **2014,** *26*, 19-24.

3. Hoffman, B. M.; Lukoyanov, D.; Yang, Z. Y.; Dean, D. R.; Seefeldt, L. C., Mechanism of nitrogen fixation by nitrogenase: the next stage. *Chem Rev* **2014,** *114* (8), 4041-62.

4. Jensen, E. S.; Peoples, M. B.; Boddey, R. M.; Gresshoff, P. M.; Hauggaard-Nielsen, H.; Alves, B. J. R.; Morrison, M. J., Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. A review. *Agronomy for Sustainable Development* **2012,** *32* (2), 329-364.

5. Gresshoff, P. M.; Hayashi, S.; Biswas, B.; Mirzaei, S.; Indrasumunar, A.; Reid, D.; Samuel, S.; Tollenaere, A.; van Hameren, B.; Hastwell, A.; Scott, P.; Ferguson, B. J., The value of biodiversity in legume symbiotic nitrogen fixation and nodulation for biofuel and food production. *J Plant Physiol* **2015,** *172*, 128-36.

6. Graham, P. H.; Vance, C. P., Legumes: importance and constraints to greater use. *Plant Physiol* **2003,** *131* (3), 872-7.

7. Pii, Y.; Mimmo, T.; Tomasi, N.; Terzano, R.; Cesco, S.; Crecchio, C., Microbial interactions in the rhizosphere: beneficial influences of plant growth-promoting rhizobacteria on nutrient acquisition process. A review. *Biology and Fertility of Soils* **2015,** *51* (4), 403-415.

8. Mus, F.; Crook, M. B.; Garcia, K.; Costas, A. G.; Geddes, B. A.; Kouri, E. D.; Paramasivan, P.; Ryu, M. H.; Oldroyd, G. E. D.; Poole, P. S.; Udvardi, M. K.; Voigt, C. A.; Ane, J. M.; Peters, J. W., Symbiotic Nitrogen Fixation and the Challenges to Its Extension to Nonlegumes. *Applied and Environmental Microbiology* **2016,** *82* (13), 3698-3710.

9. Via, V. D.; Zanetti, M. E.; Blanco, F., How legumes recognize rhizobia. *Plant Signal Behav* **2016,** *11* (2), e1120396.

10. Ferguson, B. J.; Indrasumunar, A.; Hayashi, S.; Lin, M. H.; Lin, Y. H.; Reid, D. E.; Gresshoff, P. M., Molecular analysis of legume nodule development and autoregulation. *J Integr Plant Biol* **2010,** *52* (1), 61-76.

11. Haag, A. F.; Arnold, M. F.; Myka, K. K.; Kerscher, B.; Dall'Angelo, S.; Zanda, M.; Mergaert, P.; Ferguson, G. P., Molecular insights into bacteroid development during Rhizobium-legume symbiosis. *FEMS Microbiol Rev* **2013,** *37* (3), 364-83.

12. Lau, W.; Fischbach, M. A.; Osbourn, A.; Sattely, E. S., Key applications of plant metabolic engineering. *PLoS Biol* **2014,** *12* (6), e1001879.

13. Bhardwaj, C.; Hanley, L., Ion sources for mass spectrometric identification and imaging of molecular species. *Natural Product Reports* **2014,** *31* (6), 756-767.

14. Dunn, W. B.; Bailey, N. J. C.; Johnson, H. E., Measuring the metabolome: current analytical technologies. *Analyst* **2005,** *130* (5), 606-625.

15. Barsch, A.; Carvalho, H. G.; Cullimore, J. V.; Niehaus, K., GC-MS based metabolite profiling implies three interdependent ways of ammonium assimilation in Medicago truncatula root nodules. *J Biotechnol* **2006,** *127* (1), 79-83.

16. Staszkow, A.; Swarcewicz, B.; Banasiak, J.; Muth, D.; Jasinski, M.; Stobiecki, M., LC/MS profiling of flavonoid glycoconjugates isolated from hairy roots, suspension root cell cultures and seedling roots of Medicago truncatula. *Metabolomics* **2011,** *7* (4), 604-613.

17. Zhang, N.; Venkateshwaran, M.; Boersma, M.; Harms, A.; Howes-Podoll, M.; den Os, D.; Ane, J. M.; Sussman, M. R., Metabolomic profiling reveals suppression of oxylipin biosynthesis during the early stages of legume-rhizobia symbiosis. *Febs Letters* **2012,** *586* (19), 3150-3158.

18. Zhang, J. Y.; de Carvalho, M. H. C.; Torres-Jerez, I.; Kang, Y.; Allen, S. N.; Huhman, D. V.; Tang, Y. H.; Murray, J.; Sumner, L. W.; Udvardi, M. K., Global reprogramming of transcription and metabolism in Medicago truncatula during progressive drought and after rewatering. *Plant Cell and Environment* **2014,** *37* (11), 2553-2576.

19. Del-Saz, N. F.; Florez-Sarasa, I.; Clemente-Moreno, M. J.; Mhadhbi, H.; Flexas, J.; Fernie, A. R.; Ribas-Carbo, M., Salinity tolerance is related to cyanide-resistant alternative respiration in Medicago truncatula under sudden severe stress. *Plant Cell Environ* **2016,** *39* (11), 2361-2369.

20. Bjarnholt, N.; Li, B.; D'Alvise, J.; Janfelt, C., Mass spectrometry imaging of plant metabolites--principles and possibilities. *Nat Prod Rep* **2014,** *31* (6), 818-37.

21. Kaspar, S.; Peukert, M.; Svatos, A.; Matros, A.; Mock, H. P., MALDI-imaging mass spectrometry - An emerging technique in plant biology. *Proteomics* **2011,** *11* (9), 1840-50.

22. Ye, H.; Gemperline, E.; Venkateshwaran, M.; Chen, R.; Delaux, P. M.; Howes-Podoll, M.; Ane, J. M.; Li, L., MALDI mass spectrometry-assisted molecular imaging of metabolites during nitrogen fixation in the Medicago truncatula-Sinorhizobium meliloti symbiosis. *Plant J* **2013,** *75* (1), 130-45.

23. Lee, Y. J.; Perdian, D. C.; Song, Z. H.; Yeung, E. S.; Nikolau, B. J., Use of mass spectrometry for imaging metabolites in plants. *Plant Journal* **2012,** *70* (1), 81–95.

24. Klein, A. T.; Yagnik, G. B.; Hohenstein, J. D.; Ji, Z.; Zi, J.; Reichert, M. D.; MacIntosh, G. C.; Yang, B.; Peters, R. J.; Vela, J.; Lee, Y. J., Investigation of the Chemical Interface in the

Soybean-Aphid and Rice-Bacteria Interactions Using MALDI-Mass Spectrometry Imaging. *Anal Chem* **2015,** *87* (10), 5294-301.

25. Sturtevant, D.; Duenas, M. E.; Lee, Y. J.; Chapman, K. D., Three-dimensional visualization of membrane phospholipid distributions in Arabidopsis thaliana seeds: A spatial perspective of molecular heterogeneity. *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids* **2017,** *1862* (2), 268-281.

26. Korte, A. R.; Lee, Y. J., MALDI-MS analysis and imaging of small molecule metabolites with 1,5-diaminonaphthalene (DAN). *Journal of Mass Spectrometry* **2014,** *49* (8), 737-741.

27. Duenas, M. E.; Klein, A. T.; Alexander, L. E.; Yandeau-Nelson, M. D.; Nikolau, B. J.; Lee, Y. J., High spatial resolution mass spectrometry imaging reveals the genetically programmed, developmental modification of the distribution of thylakoid membrane lipids among individual cells of maize leaf. *Plant Journal* **2017,** *89* (4), 825-838.

28. Feenstra, A. D.; Duenas, M. E.; Lee, Y. J., Five Micron High Resolution MALDI Mass Spectrometry Imaging with Simple, Interchangeable, Multi-Resolution Optical System. *Journal of the American Society for Mass Spectrometry* **2017,** *28* (3), 434-442.

29. Horn, P. J.; Korte, A. R.; Neogi, P. B.; Love, E.; Fuchs, J.; Strupat, K.; Borisjuk, L.; Shulaev, V.; Lee, Y. J.; Chapman, K. D., Spatial Mapping of Lipids at Cellular Resolution in Embryos of Cotton. *Plant Cell* **2012,** *24* (2), 622-636.

30. Gemperline, E.; Keller, C.; Jayaraman, D.; Maeda, J.; Sussman, M. R.; Ane, J. M.; Li, L., Examination of Endogenous Peptides in Medicago truncatula Using Mass Spectrometry Imaging. *J Proteome Res* **2016,** *15* (12), 4403-4411.

31. Cavatorta, V.; Sforza, S.; Mastrobuoni, G.; Pieraccini, G.; Francese, S.; Moneti, G.; Dossena, A.; Pastorello, E. A.; Marchelli, R., Unambiguous characterization and tissue localization of Pru P 3 peach allergen by electrospray mass spectrometry and MALDI imaging. *J Mass Spectrom* **2009,** *44* (6), 891-7.

32. Grassl, J.; Taylor, N. L.; Millar, A. H., Matrix-assisted laser desorption/ionisation mass spectrometry imaging and its development for plant protein imaging. *Plant Methods* **2011,** *7*, 11.

33. Poth, A. G.; Mylne, J. S.; Grassl, J.; Lyons, R. E.; Millar, A. H.; Colgrave, M. L.; Craik, D. J., Cyclotides associate with leaf vasculature and are the products of a novel precursor in petunia (Solanaceae). *J Biol Chem* **2012,** *287* (32), 27033-46.

34. Bencivenni, M.; Faccini, A.; Zecchi, R.; Boscaro, F.; Moneti, G.; Dossena, A.; Sforza, S., Electrospray MS and MALDI imaging show that non-specific lipid-transfer proteins (LTPs) in tomato are present as several isoforms and are concentrated in seeds. *J Mass Spectrom* **2014,** *49* (12), 1264-71.

35. Batut, J.; Mergaert, P.; Masson-Boivin, C., Peptide signalling in the rhizobium-legume symbiosis. *Curr Opin Microbiol* **2011,** *14* (2), 181-7.

36. Gemperline, E.; Jayaraman, D.; Maeda, J.; Ane, J. M.; Li, L., Multifaceted investigation of metabolites during nitrogen fixation in Medicago via high resolution MALDI-MS imaging and ESI-MS. *J Am Soc Mass Spectrom* **2015,** *26* (1), 149-58.

37. Mitra, R. M.; Long, S. R., Plant and bacterial symbiotic mutants define three transcriptionally distinct stages in the development of the *Medicago truncatula*/*Sinorhizobium meliloti* symbiosis. *Plant Physiology* **2004,** *134* (2), 595–604.

38. Wang, D.; Griffitts, J.; Starker, C.; Fedorova, E.; Limpens, E.; Ivanov, S.; Bisseling, T.; Long, S. R., A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* **2010,** *327* (5969), 1126–1129.

39. Wolf, S.; Schmidt, S.; Muller-Hannemann, M.; Neumann, S., In silico fragmentation for computer assisted identification of metabolite mass spectra. *BMC Bioinformatics* **2010,** *11*, 148.

40. Schulze, J., Source-sink manipulations suggest an N-feedback mechanism for the drop in N2 fixation during pod-filling in pea and broad bean. *J Plant Physiol* **2003,** *160* (5), 531-7.

41. Fischinger, S. A.; Drevon, J.-J.; Claassen, N.; Schulze, J., Nitrogen from senescing lower leaves of common bean is re-translocated to nodules and might be involved in a N-feedback regulation of nitrogen fixation. *Journal of Plant Physiology* **2006,** *163* (10), 987-995.

42. Sulieman, S.; Fischinger, S. A.; Gresshoff, P. M.; Schulze, J., Asparagine as a major factor in the N-feedback regulation of N2 fixation in Medicago truncatula. *Physiol Plant* **2010,** *140* (1), 21-31.

43. Luyten, E.; Vanderleyden, J., Survey of genes identified in Sinorhizobium meliloti spp., necessaryfor the development of an efficient symbiosis. *European Journal of Soil Biology* **2000,** *36* (1), 1-26.

44. Appleby, C. A., Leghemoglobin and Rhizobium Respiration. *Annual Review of Plant Physiology* **1984,** *35* (1), 443-478.

45. Larrainzar, E.; Molenaar, J. A.; Wienkoop, S.; Gil-Quintana, E.; Alibert, B.; Limami, A. M.; Arrese-Igor, C.; Gonzalez, E. M., Drought stress provokes the down-regulation of methionine and ethylene biosynthesis pathways in Medicago truncatula roots and nodules. *Plant Cell Environ* **2014,** *37* (9), 2051-63.

46. Gil-Quintana, E.; Lyon, D.; Staudinger, C.; Wienkoop, S.; Gonzalez, E. M., Medicago truncatula and Glycine max: Different Drought Tolerance and Similar Local Response of the Root Nodule Proteome. *J Proteome Res* **2015,** *14* (12), 5240-51.

47. Gemperline, E.; Jayaraman, D.; Maeda, J.; Ane, J. M.; Li, L. J., Multifaceted Investigation of Metabolites During Nitrogen Fixation in Medicago via High Resolution MALDI-MS Imaging and ESI-MS. *Journal of the American Society for Mass Spectrometry* **2015,** *26* (1), 149-158.

## Tables

**Table 1. Annotations of "Metabolites of Interest" Detected in Medicago Root Nodules**

| Metabolite | Measured *m/z* | Theoretical *m/z* | Δ ppm |
|---|---|---|---|
| Aminobutryic acid | 104.0708 | 104.0706 | 1.9 |
| Methyl-piperidin-iumone [M+] | 114.0915 | 114.0913 | 1.4 |
| Proline | 116.0706 | 116.0706 | 0.0 |
| Aminopentene-diol* | 118.0863 | 118.0863 | 0.0 |
| Isoleucinol | 118.1228 | 118.1226 | 1.3 |
| Asparagine* | 133.0607 | 133.0608 | 0.8 |
| Glutamic Acid | 148.0603 | 148.0604 | 0.9 |
| Asparagine [M+Na]* | 155.0425 | 155.0427 | 1.3 |
| 3-thiophen-1-yl propanoic acid | 158.0397 | 158.0396 | 0.7 |
| Ethyl-aminocyclopentane carboxylic acid | 158.1173 | 158.1176 | 1.6 |
| Phenylalanine | 166.0859 | 166.0863 | 2.1 |
| Asparagine [M+K]* | 171.0165 | 171.0167 | 1.2 |
| Arginine | 175.1187 | 175.1190 | 1.4 |
| Tyrosine | 182.0810 | 182.0812 | 1.0 |
| Methyl-alpha-galactopyranoside | 195.0862 | 195.0863 | 0.6 |
| Ethyl-aminocyclopentane carboxylic acid [M+K] | 196.0731 | 196.0734 | 1.7 |
| Proclavaminic acid* | 203.1024 | 203.1026 | 1.0 |
| Methyl-alpha-galactopyranoside [M+Na]* | 217.0678 | 217.0682 | 1.8 |
| Proclavaminic acid [M+Na] | 225.0840 | 225.0846 | 2.4 |
| Proclavaminic acid [M+K] | 241.0577 | 241.0585 | 3.3 |
| Dihydroxybenzoic acid succinimido ester | 252.0500 | 252.0503 | 1.2 |
| 5-amino-2-(aminomethyl)-6-butoxyoxane-3,4-diol [M+Na] | 257.1462 | 257.1472 | 3.9 |
| 9H-fluoren-9-yl-di(propan-2-yl)phosphane | 283.1618 | 283.1610 | 2.9 |
| N-(4-guanidinobutyl)-3-methyldecanamide | 299.2802 | 299.2805 | 1.1 |
| SAM (S-adenosyl-L-methionine) | 399.1441 | 399.1445 | 1.1 |
| 2-hydroxy-5-[[2-phenyl-2-[4-(phenylcarbamoylamino)phenyl]sulfanylacetyl]amino]benzoic acid | 514.1423 | 514.1431 | 1.5 |
| 2-[hydroxy-[(2R)-3-hydroxy-2-[(9E,12E)-octadeca-9,12-dienoyl]oxypropoxy]phosphoryl]oxyethyl-trimethylazanium [M+] | 520.3395 | 520.3398 | 0.5 |
| Oleoyl lysophosphatidylcholine | 522.3552 | 522.3554 | 0.5 |
| 18-[(4Z)-4-[(2-hydroxy-5-nitrophenyl)hydrazinylidene]-3-oxocyclohexa-1,5-dien-1-yl]octadecanoic acid [M+] | 541.3125 | 541.3146 | 3.9 |
| 3-[[[[(2R,3S,4R,5R)-5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methoxy-hydroxyphosphoryl]oxy-hydroxyphosphoryl]oxy-hydroxyphosphoryl]oxypropanoic acid | 580.0223 | 580.0242 | 3.1 |
| [[(2R,3S,5R)-5-[4-amino-5-(4-aminobutyldisulfanyl)-2-oxopyrimidin-1-yl]-3-hydroxyoxolan-2-yl]methoxy-hydroxyphosphoryl] phosphono hydrogen phosphate [M+] | 602.0052 | 602.0067 | 2.4 |
| (Z)-4-oxo-2-[(Z)-1-oxooctadec-9-enyl]-12-henicosenoic acid | 603.5340 | 603.5347 | 1.1 |
| 1,3-dilinolenin | 613.4814 | 613.4827 | 2.0 |
| Heme | 617.1844 | 617.1846 | 0.4 |

Ions are [M+H] unless otherwise specified.
* Found in both MALDI and ESI

**Figures**

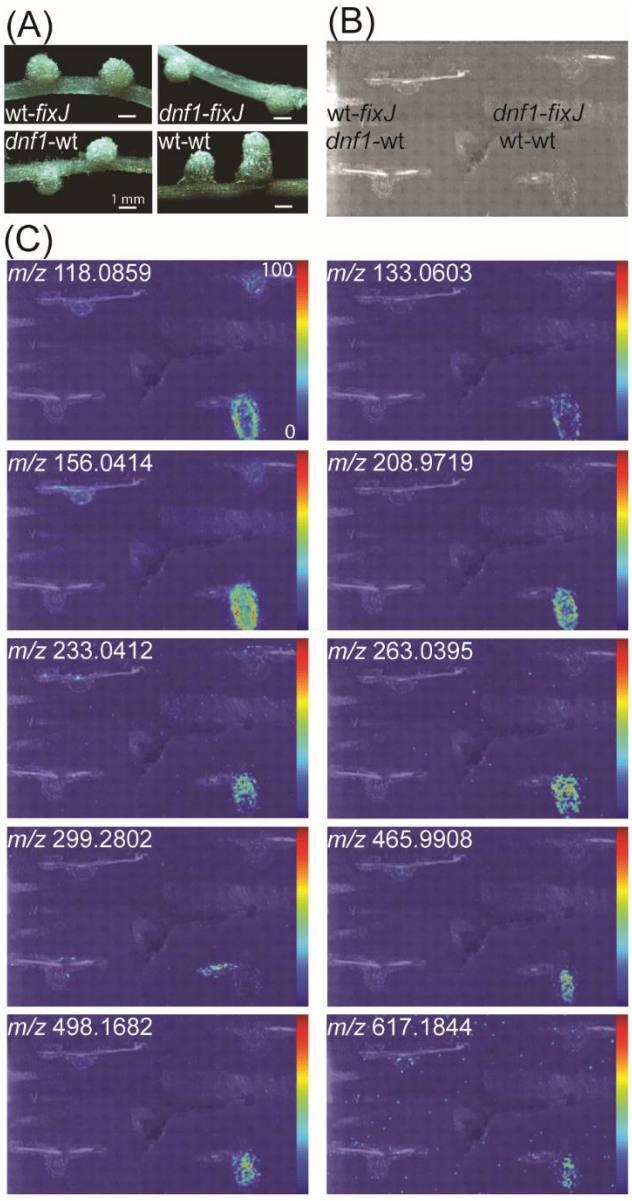



**Figure 1. (**A) A photograph of all four of the sample types that were used in this study: wt-*fixJ*, *dnf1-fixJ*, *dnf1*-wt, and wt-wt. (B) Thaw-mounted Medicago sections on a glass slide that was covered with DHB matrix prior to MALDI-MSI. (C) Representative ion images of metabolites with distinct spatial distributions that were found only in the wt-wt samples. Scale bar = 1 mm**.** Reprinted with permission from (Gemperline et al. 2015). Copyright 2015 Springer.

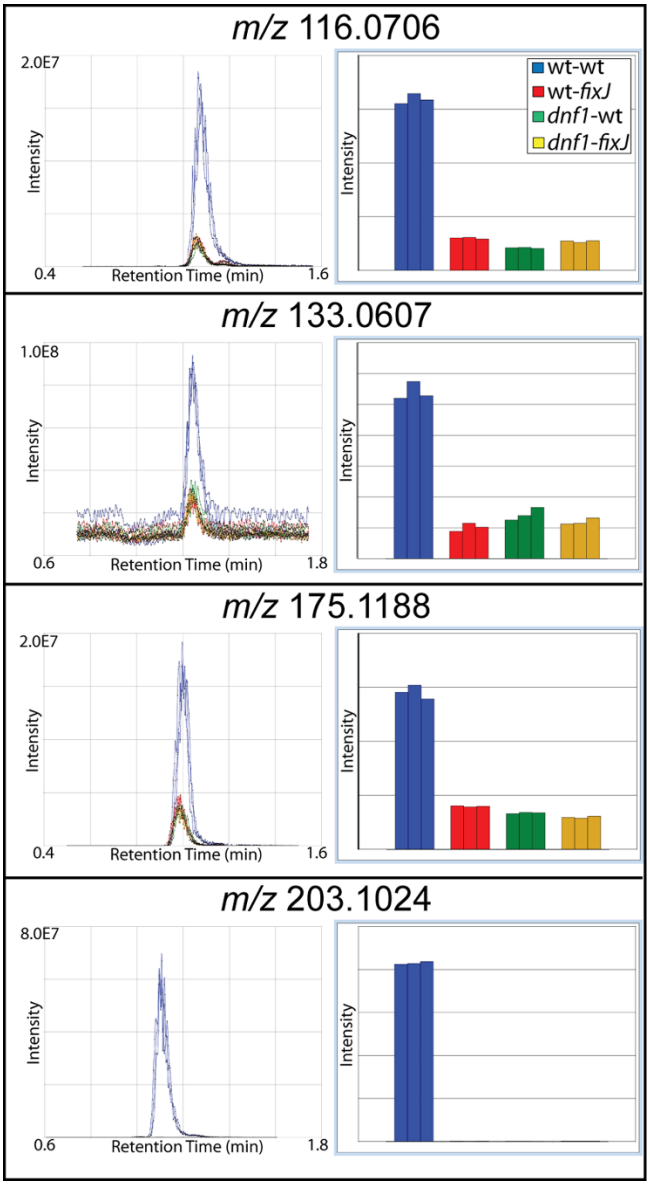



**Figure 2.** Representative data from SIEVE, which was used to select the "metabolites of interest". Blue indicates wt-wt samples; red indicates wt-*fixJ* samples; green indicates *dnf1*-wt samples; and yellow indicates *dnf1-fixJ* samples. The panels on the left show XIC peaks, demonstrating that the metabolite has at least 2-fold greater intensity in the wt-wt samples compared to any of the mutant samples. The panels on the right visualize the information on the left in bar graph form. Reprinted with permission from (Gemperline et al. 2015). Copyright 2015 Springer.

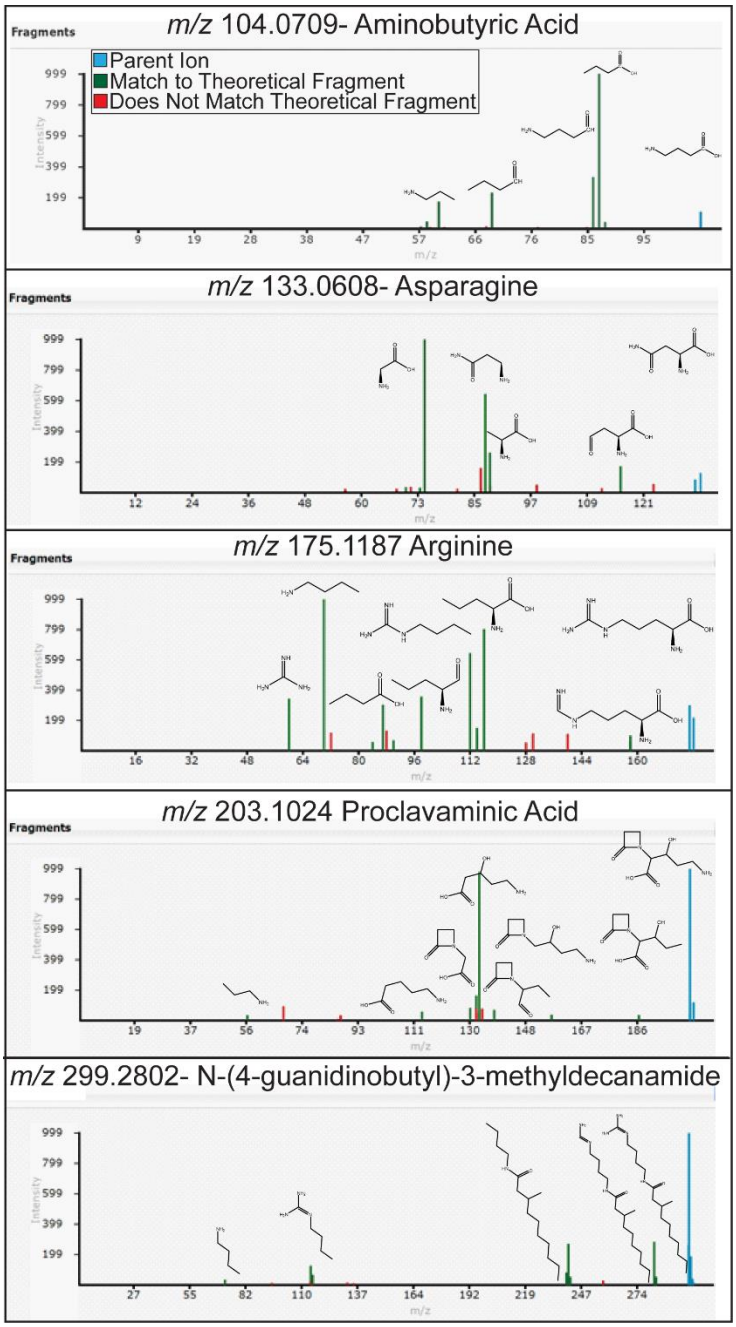



**Figure 3.** Representative MS/MS spectra of the identified "metabolites of interest" and the molecular structures of the matched fragments. Blue indicates the parent ion; green indicates MS/MS peaks that matched the *in silico* fragmentation in MetFrag; and red indicates the experimental peaks that did not match the *in silico* fragmentation. Reprinted with permission from (Gemperline et al. 2015). Copyright 2015 Springer.

# Appendix IV

# Regulation of Nodule Differentiation in the Indeterminate Nodulator *Medicago truncatula* by Moderate Salt Stress

Adapted from: Chapter three of Sanhita Chakraborty's thesis. This material is in collaboration with Dr. Jean-Michel Ané and Dr. Jean Harris's (University of Vermont) and experiments performed by undergraduate student Emily Millar and volunteer Ashley Lantigua in Harris lab. Junko Maeda and Dr. Dhileepkumar Jayaraman from the Ané lab grew *M. truncatula* seedlings for nodules, and I performed MALDI-MSI. Emily Millar grew the seedlings for nodules and performed ABA ELISA with Sanhita. Ashley Lantigua grew and inoculated plants and performed histochemical assays for *bacA* and *nifH*.

Keywords: Salt stress, *Medicago truncatula*, Metabolites, MALDI, Orbitrap, Mass Spectrometry Imaging

**Abstract**

Nodules are symbiotic root organs that legumes develop in association with the soil bacteria rhizobia. These symbiotic organs are mini-factories where atmospheric dinitrogen is reduced to ammonia by the rhizobial enzyme nitrogenase. Nodule development is strongly inhibited by salt stress. We have previously shown that salt interferes with the initial stages of symbiosis, where it hyperinduces several Nod factor (NF) signaling and infection-associated genes and inhibits infection in the epidermis. Here, we studied the development of the rare nodules that develop on *Medicago truncatula* roots under salinity. Indeterminate nodulators, such as *M. truncatula,* develop nodules that show a gradient of differentiation due to the presence of a persistent meristem supplying cells to form the bulk of the nodule. We exploited this developmental pattern to address whether salinity impacts nodule differentiation. Our data revealed the upregulation of several genes involved in NF-signaling and infection under moderate salt treatment, coincident with very little expression of host and symbiont genes involved in the differentiation of rhizobia into nitrogen-fixing bacteroids. In order to capture metabolic alterations in the nodules *in situ,* we exploited MALDI-MSI to find distinct spatial accumulation of metabolites, such as proline betaine and homostachydrine, in salt-stressed nodules. We further observed differential accumulation of abscisic acid between the control and salt-stressed nodules. Our combined findings reveal a later effect of salinity on development of the nodule organ, where it interferes with differentiation of rhizobia into bacteroids.

**Introduction**

Legumes are prized agriculturally for their high protein content. The biological process that forms the basis of their internal source of nitrogenous fertilizer is a symbiotic association

between legume roots and a group of soil-residing bacteria, collectively known as rhizobia. The legume-rhizobium symbiosis culminates in the formation of a new symbiotic root organ, a nodule, that houses differentiated rhizobia, capable of reducing atmospheric dinitrogen into ammonia. Nodule development is influenced by environmental factors, such as light, temperature, pH, drought and salinity.[1] Hence changing environmental conditions can have a profound impact on development of the symbiosis.

Nodule development involves the entry of rhizobia into the host (infection), their differentiation into a symbiotic form called a bacteroid, and the concomitant synthesis of an organ (organogenesis) in the host root that houses these rhizobia.[2] Nodule organogenesis closely follows perception of rhizobial signal molecules by the host epidermis.[3] In a signaling cascade that is mediated by the phytohormone cytokinin, the inner cell layers of the host root become mitotically activated and undergo multiple rounds of anticlinal and periclinal divisions, forming a nodule primordium.[4-6] Rhizobia populate these nodule primordia and together, they mature into a compound organ, the nodule.[7] In indeterminate nodulators, such as *Medicago* sp., a single nodule exhibits a longitudinal gradient of differentiation due to the presence of a persistent meristem at the tip of the organ.[2, 8] This meristem encompasses a region of division that produces differentiating cells contributing to the nodule organ. Newly produced cells become infected with rhizobia, delivered within transcellular apoplastic compartments called infection threads (IT). These are formed by the inward growth of host cell wall and plasma membrane, such that the rhizobia are topologically outside of the host cell.[8-10] Individual rhizobial cells become enclosed within the host-derived membrane and are released into the host cell. The membrane at this stage becomes the peribacteroid membrane, or symbiosome membrane, and together with the enclosed rhizobium, forms an organelle-like structure, the symbiosome. In species of *Medicago*, rhizobia

within a symbiosome undergo cell enlargement and endoreduplication and terminally differentiate into bacteroids. Once a bacterium metamorphoses into a bacteroid, it is able to reduce atmospheric dinitrogen into ammonia.

Nodule initiation requires the activation of cytokinin signaling, which induces downstream genes, including *CLAVATA3/embryo-surrounding region 13* (*CLE13*) and the Cytokinin type A *Response Regulator 4* (*RR4*).[3-5, 11-13] *Lateral root-organ defective/ Numerous infections and polyphenolics (LATD/NIP)* encodes a nitrate transporter and is required for the induction of *Nodulation Signaling Pathway 2* (*NSP2*) by cytokinin during nodule development.[14] *CLE13* and *LATD/NIP* expression are induced in the meristematic regions during nodule development.[11, 15, 16] The transcription factor *WUSCHEL RELATED HOMEOBOX5* (*WOX5*), thought to be required for meristematic activity in the root, is also highly expressed in nodule meristem[17] and might have a subtle interaction with cytokinin signaling.[18] Cytokinin also plays a role slightly later in nodule development, during the transition of a meristematic cell into a differentiated cell, as observed from high expression of *RR4* at this stage.[4, 17] As a nodule matures, the dividing cells differentiate and become infected with rhizobia. Several host genes, such as, *Early Nodulin 11* (*ENOD11*), *ERF required for nodulation 1* and *2*, (*ERN1* and *ERN2*), *Nuclear Factor YA-1* (*NF*-YA1) and *Nodulation pectate lyase* (*NPL*) that are involved during early rhizobium-legume signaling and rhizobial infection in the root, also participate during and after this transition.[17, 19-21]

Differentiation of infecting rhizobia relies on the activities of both host and symbiont gene products. Host *Nodules with activated defense 1 (NAD1),* encoding an endoplasmic reticulum-localized product, is required to suppress plant defense responses, following rhizobial release from the infection threads but prior to differentiation into bacteroids.[22] Host *Doesn't Fix Nitrogen* (*DNF*) genes are part of a large gene family encoding nodule-specific cysteine-rich (NCR)

peptides, required for the successful transition from non-symbiotic rhizobia to nitrogen-fixing bacteroids.[23] These are secreted peptides targeted to the symbiosomes, which interact with rhizobia, often entering their cytoplasm.[24, 25] *DNF7* (*NCR169*) and *DNF4* (*NCR211*) are crucial for bacteroid differentiation.[25, 26] *DNF4* (*NCR211*) participates during the final stages of differentiation by helping bacteroids adapt to the changing cellular environment.[27] Deletion in *DNF4* or *DNF7* leads to impaired differentiation of the symbiont resulting in inefficient small, white nodules  Both genes are strongly expressed in infected cells.[26, 27] Unlike the products of *DNF7* and *DNF4* NCR peptides, many others exert an antimicrobial effect on the symbiont, as the host controls its relationship with the symbiont.[25] BacA, a plasma membrane-localized rhizobial ABC transporter-like protein, functions to import NCR peptides. This import function is thought to redirect NCR peptides away from the plasma membrane, thus protecting it from the antimicrobial properties of these peptides.[28, 29] Consequently, *bacA* mutants form an ineffective symbiosis as they are unable to differentiate into bacteroids and senesce prematurely.[30] Once differentiated, bacteroids begin to express the *nitrogen fixation* (*nif*) genes and produce rhizobial enzyme nitrogenase, which fixes atmospheric nitrogen.[31] *nifH* encodes the dinitrogenase reductase subunit, which reduces the dinitrogenase component of the nitrogenase enzyme complex. Nodule maturation into a nitrogen-fixing organ is accompanied by the accumulation of the hemoprotein, leghemoglobin, which may be required to produce the low oxygen environment required for nitrogenase activity.[32] The nodule eventually senesces; senescing nodules are green in color due to the degradation of heme.[33] Heme oxygenases function to break down heme.[34] In *Medicago truncatula* the *GIRAFFE* gene encodes a  heme oxygenase,[35, 36] and is strongly expressed in senescing nodules  (https://mtgea.noble.org/v3/probeset.php?id=Mtr.40298.1.S1_at&print=true).

Salinity strongly impacts nodule development by inhibiting rhizobial colonization, their release from the IT and differentiation into bacteroids.[37-41] Plant response to salinity is mediated by phytohormones, primarily abscisic acid (ABA), which accumulates in salt-stressed tissues.[42, 43] In order to survive the inhibitory effects of salt, nodules often accumulate osmoprotectants in the form of organic acids, such as lactate; free amino acids, such as asparagine and proline; amino acid derivatives, such as proline betaine; and certain carbohydrates, such as pinnitol.[44-46]

Reactive oxygen species (ROS) function as a balance between plant development and environmental stress response; while they can act as signal molecules in response to various abiotic stresses, excess ROS within a cell can cause oxidative stress.[47] *Respiratory Burst Oxidase Homolog* (*Rboh*) genes play an important role in generating ROS, where their products convert molecular oxygen ($O_2$) to superoxide radical ($O_2^{\cdot}$).[48, 49] In *M. truncatula, RbohC* regulates root growth in an ABA-dependent manner, is induced in rhizobium-inoculated roots, and is highly expressed in nitrate-induced senescing nodules.[35, 50-52] *MtRbohA* on the other hand, is required for nitrogen fixation and is expressed in mature nodules.[53] In *Arabidopsis thaliana*, *AtRbohD* and *AtRbohF* are upregulated under salt stress and *AtrbohF* mutants are hypersensitive to $Na^+$ accumulation.[54]

Despite the strong negative effect of salinity on nodule development, the molecular mechanism is poorly understood. We have previously shown that salt interferes with early rhizobium-legume signaling and rhizobial infection, and that there is a significant reduction in the number of nodules formed under salt stress (Chakraborty et al., 2018). The few nodules that do form in the presence of salt are infected with rhizobia, but appear morphologically abnormal, suggesting that salinity also interferes with nodule maturation. Here, we examine the effect of moderate salt stress on the development of the nodule organ. Our study reveals that salinity delays

nodule development and modulates the expression of host and symbiont genes required for nodule maturation. We show that while salinity induces the expression of several host genes involved during the early stages of nodule maturation, it reduces the expression of both host and symbiont genes associated with differentiation into bacteroids, nitrogen fixation and heme degradation. We further find that these changes are reflected in differential accumulation of distinct metabolites in nodules in the presence or absence of salt stress.

**Experimental Methods**

Plant Growth Conditions and inoculation

Seedlings were grown and inoculated as described by Chakraborty et al (Chakraborty et al., in prep.). *M. truncatula* A17 seeds were scarified in concentrated sulfuric acid for 15 minutes, surface sterilized with 30% bleach for 10 minutes, and imbibed for 4-6 hours at room temperature. Imbibed seeds were stratified at 4°C for 1-4 days and germinated overnight. Seven germinated seedlings were placed in cyg seed germination pouches (https://mega-international.com), in Modified Nodulation Medium (MNM) and wrapped with aluminum foil to shade the roots. Pouches flanked by two water-saturated empty pouches were placed in a styrofoam box, covered with a transparent lid, sealed with tape and placed in MTR30 Conviron growth chambers, set at 20°C, 50% humidity, and a 16-h-light/8-h-dark cycle, with an intensity of 100 μE $m^{-2}$ $s^{-1}$. Media were replenished twice a week. Rhizobial strains used were in the *S. meliloti* (*Rm*1021) background. Pelleted cells from overnight cultures grown in LB medium supplemented with appropriate antibiotic(s) were re-suspended in 10 mM $MgSO_4$, and the $OD_{600}$ adjusted to ~0.07. Each pouch was then inoculated with 1 mL of the rhizobial suspension, added near the tip of the roots or mock-inoculated with an equivalent volume of 10 mM $MgSO_4$.

*bacA* and *nifH* expression

Seedlings were inoculated with *Rm1021*[55] strains harboring either the transcriptional fusion *bacA:uidA* (VO2196) (V. Oke, personal communication) or *nifH:uidA* (CSB357).[23] Whole seedlings with nodules were harvested and stained for β-glucuronidase activity with 1 mM 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-gluc) + 0.02% SDS in 50 mM sodium phosphate buffer, pH 7.0,[56] for five hours at 37°C. Nodules attached to roots were imaged under a Leica dissecting microscope (LeicaMZ8).

Quantitative RT-PCR

Total RNA was extracted either from the primary roots (1, 2, 3 or 5 dpi) or the oldest nodule and any nodule within 3 cm of it on the primary root (14 and 21 dpi). For the root samples, 21 seedlings per treatment per experiment were used; for the nodule samples, 21 seedlings (control) or 63 seedlings (salt) per treatment per experiment were used. Tissue was ground, frozen and RNA was extracted using the RNeasy Plant Mini kit (Qiagen), genomic DNA was removed using the TURBO DNA-free kit (Ambion), and the resulting RNA further cleaned-up using the RNeasy MinElute Clean-up kit (Qiagen). RNA quality was assessed with an Agilent 2100 Bioanalyzer and RNA samples with RNA Integrity Number (RIN) >8 were used for analysis. cDNA was synthesized from 1 μg total RNA using the iScript cDNA synthesis kit (BIO RAD). qPCR was performed with SYBR green (Quanta PerfecTa) in the StepOnePlus Real time PCR system (Applied biosystems) using the ddCT method, run in duplicate with UBC9 used as endogenous control (Kakar et al., 2008).[57]

MALDI-MSI

Root nodules from seedlings grown in low salt (control) and high salt plants were prepared for MALDI-MSI as previously described.[58] Briefly, nodules removed from plants were embedded

in gelatin and gently frozen, sectioned at 16 μm thickness, and sprayed with 2,5-dihydroxybenzoic acid (DHB; Acros Organics) matrix with a TM Sprayer (HTX Technologies, LLC, Carrboro, NC, USA). MALDI-MSI was performed on a MALDI LTQ Orbitrap XL instrument from Thermo Scientific equipped with a $N_2$ laser in positive ion mode. LTQ Tune software (Thermo Scientific, Waltham, MA, USA) and Xcaliber (Thermo Scientific, Waltham, MA, USA) were used to select the imaging region and step size and the instrument parameters respectively. Imaging was performed on three biological replicates (each with three technical replicates) at 75 μm raster step size. The mass range was set to 100-1000 *m/z* and the resolution to 60,000. Two microscans were averaged at each pixel. MSiReader software[59] was used to generate images from the data. Images of *m/z* consistently present in only the control root nodules and roots were grouped into a control target list, and images of *m/z* consistently in the high salt root nodules and roots were grouped into a salt target list.

MALDI-Orbitrap MS/MS

Tandem MS (MS/MS) was performed on the targets using on-tissue MALDI-MS/MS. On-tissue MS/MS was manually optimized for collision-induced dissociation (CID) or higher energy collisional dissociation (HCD) activation with detection in the orbitrap or ion trap (CID only) by isolated the parent *m/z* and adjusting the collision energy. Further on-tissue MS/MS in the ion trap (CID activation) was performed with a 4-step spiral automatic sequence. The raster step size was 100 μm and the spiral step size was 50 μm. A full MS scan in the orbitrap was followed by MS/MS scans at collision energies of 25, 30, and 35 respectively for scans 2-4. An inclusion list with the target *m/z* was used. The dynamic exclusion parameters used were a repeat count of 2, repeat duration of 120 s, and exclusion duration of 600 s.

Tissue Extractions

Approximately 100 control nodules and 100 salt treated nodules (with 2-4 mm of surrounding root) were trimmed from the plants and flash frozen. Nodules were ground with a mortar and pestle under liquid nitrogen, and a methanol/chloroform/water extraction was performed by adding in order 3 parts methanol (600 µL), 1 part chloroform (200 µL), and 4 parts water (800 µL). Samples were vortexed and then centrifuged for 10 min at 5000 x g and 4ºC. After removing the upper aqueous layer, 4 parts methanol were added, the extraction was vortexed briefly, and centrifuged again at 1500 x g for 5 min and 4ºC. The supernatant (organic layer) was removed. The aqueous and organic fractions were dried down in a speedvac and saved in a -80 ºC freezer until analysis.

LC-MS/MS on Tissue Extractions

Aqueous samples were resuspended in optima grade water with 0.1 % FA at 10 mg/mL. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was performed with a Dionex Ultimate 3000 UHPLC system (Thermo Scientific, Waltham, MA, USA) equipped with a Kintex C18 column (2.1 mm internal diameter x 150 mm length, 1.7 µm particle size; Phenomenex, Torrance, CA, USA) with a corresponding guard column, and a Q Exactive MS (Thermo Scientific, Waltham, MA, USA). The column temperature was 35ºC, and the mobile phases were optima grade water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). A 35-minute multi-step gradient at a flow rate of 0.3 mL/min was used with the following parameters: 0-5 min, held at 1% B; 5-10 min, linear gradient from 1-3% B; 10-18 min, linear gradient from 3-40% B; 18-22 min, linear gradient from 40-80% B; 22-27 min, column cleaning at 95% B; and 27-35 min, re-equilibration at 1% B. The injection volume was 4 µL, and the samples were kept at 10ºC during analysis. The MS parameters were as follows: positive ion mode, scan range of *m/z* 100-1500, 70,000 resolution, 1 E6 AGC, and 100 ms maximum injection. MS/MS was acquired

on targets using a top 5 method with an inclusion list with the following parameters: 35,000 resolution, 1 E5 AGC, 100 ms max injection time, 15 s dynamic exclusion, and collision energies of 30, 35, and 40.

Metabolite Identification

Metabolite identifications were made using *in silico* fragmentation with MetFrag software.[60] Target *m/z* were searched against the KEGG and ChEBI databases with 5 ppm mass error. Multiple positive ion adducts were searched. The top fragments from the MS/MS spectra and their intensities were compared to the *in silico* fragmentation. Fragment ion tolerances were set at 5 ppm relative error and 0.01 absolute error for MS/MS detected in the orbitrap and 1500 ppm, 0.15 absolute for MS/MS detected in the ion trap. Putative identifications were made when almost all major fragments were explained by the *in silico* fragmentation. MS/MS from the mzCloud database (HighChem LLC, Slovakia) and obtained standards were used where possible to verify putative identifications.

**Results**

Salt stress upregulates early-nodulation and infection-associated genes in the nodules.

We wanted to determine the way in which salinity interferes with nodule maturation. M. truncatula forms indeterminate nodules, in which a meristem at the distal end of the nodule continuously provides a new population of cells that differentiate and contribute to the nodule organ. As a result, all the stages of nodule differentiation, cell division, infection, and symbiosome formation, occur simultaneously in a graded fashion with the youngest tissue at the tip (distal end) and the oldest tissue at the base of the nodule, adjacent to the root. Previously, we had shown that a moderate salt stress affects the expression of many genes involved in signaling and infection at

an early stage in the interaction: at 1 day post-inoculation (dpi), prior to the formation of any nodule primordia. We wondered whether salt continued to affect the expression of these genes as the early signaling pathway is recapitulated at the nodule apex. To address this question, we measured the expression of these early signaling and infection-related genes in nodules at 14 and 21 dpi. At this point, under control conditions, both plant and bacteria have differentiated and produced a mature, nitrogen-fixing nodule, with an active distal meristem and a gradient of development. ERN1 and ERN2 are required for early signaling during nodule development,[20] and we had observed upregulation of ERN1 at 1 and 2 dpi by salt (Chakraborty et al., in prep.). Here, we found that salt could still upregulate ERN1 at 14 dpi (**Figure 1**).

We also previously observed upregulation of several infection-associated genes by salt (Chakraborty et al., in prep.). NF-YA1 is essential for the progression of infection thread (IT) growth - nf-ya1 mutants develop aberrant ITs[21] - and is also required for nodule meristem development, with an expression pattern confined to the youngest part of the nodule, the apical meristem.[19] MtNPL encodes a Nodule Pectate Lyase; its Lotus japonicus ortholog, LjNPL has been shown to be involved in localized degradation of host cell wall (CW) to facilitate rhizobial entry.[61] ENOD11 encodes a CW-localized glycoprotein that accumulates in the infection chamber prior to the beginning of an IT formation.[62] We found upregulation of these infection-associated genes in 14 dpi nodules by salt (**Figure 1**) just as we did at 1 dpi (Chakraborty et al., in prep.). Interestingly, expression of most of these genes had dropped by 21 dpi and expression in salt and control were not significantly different at that time point (**Figures 1, 2. Table 1**).

Given the dual role of Rboh genes in abiotic stress responses and legume-rhizobium symbiosis, we quantified the expression of MtRbohA and MtRbohC in nodules on plants grown with or without salinity. Similar to other infection-associated genes, MtRbohC was induced by salt in 14

dpi nodules (**Figure 1**), whereas MtRbohA did not show any difference in expression (**Supplemental Figure 1**). Next, to determine if the cell division genes, WOX5 and LATD/NIP, might be responsible for the callus-like appearance of salt-stressed nodules, we conducted qPCR on 14 dpi nodules, and found that salt upregulated WOX5 (**Figure 1**) but not LATD/NIP. Similar to our findings in the nodule, salt also failed to alter LATD/NIP expression in 1, 2 ,3 and 5 dpi roots (**Supplemental Figure 1**).

Transcriptome profiling at 1 dpi previously revealed that the activation of Nod factor (NF)-triggered cytokinin signaling is broadly robust under salinity, and that NIN and NSP2, two genes acting downstream of cytokinin signaling have unaltered expression (1, 2, 3 and 5 dpi) (Chakraborty et al., in prep.). Further to these results, we found that the cytokinin-induced nodulation gene RR4[5] is upregulated in 14 dpi nodules in the presence of salinity, whereas salt had no effect on the expression of CLE13.[12] To check whether CLE13 expression is affected at an earlier stage of nodule development, we also measured its transcript abundance in the roots, when cytokinin signaling is initially activated, and found that CLE13 induction by rhizobia was unaltered in the presence of salinity (**Supplemental Figure 1**).

Salt stress downregulates host and symbiont genes involved in bacteroid differentiation and nitrogen fixation.

Since several host genes associated with early stages of nodule differentiation are upregulated in the presence of salt in 14 dpi nodules, we wondered whether that meant that nodule development had slowed or ceased. If that were true, the genes involved in the later stages of differentiation, such as bacteroid formation, nitrogen fixation and heme metabolism must be expressed at a lower level compared to the control condition at this time point. To test this possibility, we studied the expression of host and rhizobial genes required for bacteroid

differentiation, nitrogen fixation and heme degradation. NAD1 is required for host immune suppression prior to bacteroid differentiation but after the release of rhizobia into the host cell from the IT.[22] Although an induction in its expression need not necessarily coincide with its function, it can be used to get an idea about rhizobial release from IT in the presence of salinity. We found that NAD1 is normally induced under salinity, indicating that the symbiosis has proceeded to a stage, where the host is able to prepare itself for the next stages of differentiation. This requires successful host-symbiont signaling, in turn, indicating that the bacteria are ready to differentiate (**Figure 2**). We then asked whether the subsequent step of differentiation into bacteroids can take place normally under salinity. Rhizobial differentiation into bacteroids requires sequential induction of DNF genes in the host.[23] We measured the expression of DNF4 (NCR211) and DNF7 (NCR169)[26, 27] (**Figure 2**). As expected, DNF7 expression was lower under salinity (**Figure 2**) but we did not observe any effect of salt on DNF4 expression (**Supplemental Figure 1**).

The Rhizobial BacA protein is required to combat the antimicrobial properties of the NCR peptides,[28, 29] and plays an important role in membrane alterations during bacteroid differentiation.[63] If salinity slows down bacteroid differentiation, that is likely to be reflected in delayed bacA expression. To test this hypothesis, we studied bacA promoter activity using reporter gene expression, and found that salt significantly inhibited bacA expression (**Figure 3A**). We further tested whether the exposure time of the host to salinity made any difference in bacA expression. To test whether length of salt exposure affected differentiation of infecting rhizobia, we added salt either one or four days prior to inoculation, and compared to differentiation under control conditions, using bacA expression as a marker for rhizobium differentiation. We found that bacA expression was reduced only with a prolonged prior salt treatment and not the shorter

treatment. Additionally, there was no significant difference in bacA expression for a given treatment between the two time-points tested (9 and 14 dpi) (**Figure 3A**).

Successful differentiation into bacteroids is a prerequisite for nitrogen fixation and several dnf mutants as well as bacA mutants fail to fix nitrogen.[26, 27, 30] We wanted to ascertain whether salinity inhibits rhizobial infection such that nitrogen fixation cannot take place. One marker for nitrogen fixation is the expression of nifH, the rhizobial gene encoding one of nitrogenase subunits. Supporting our previous inference, we found that nifH expression was decreased under salinity. We also found that the length of exposure to salinity prior to inoculation had an impact on nifH expression (**Figure 3B**). While both salt treatments reduced nifH expression, this downregulation was more severe when the seedlings experienced salt for 4 dpi.

Accumulation of leghemoglobin can be used as a proxy for nodule maturation. Salt-stressed nodules are rarely pink (Chakraborty et al., 2018), indicative of little to no leghemoglobin. This phenotype can be caused either by reduced production of the protein or increased degradation. GIRAFFE is involved in heme degradation in nodules by encoding the only heme oxygenase found in M. truncatula.[35] We found significantly less expression of GIRAFFE at 14 dpi under salinity, further supporting that nodule differentiation is hindered under salinity (**Figure 2**). Combining the results from DNF7, bacA, nifH and GIRAFFE expression, we conclude that salt interferes with the expression of host and symbiont genes involved in bacteroid differentiation and nitrogen fixation.

Host genes show altered temporal expression profiles under salinity.

In the previous sections, we studied the expression of fifteen host genes that are involved during various stages of nodule development (**Table 1**). Based on differences between treatments (control and salt) and stages of development (14 and 21 dpi), we grouped the expression pattern

into six categories (**Table 1**). We have already discussed the differences we saw in expression patterns between treatments at 14 dpi (**Table 1: groups A, D, E; Figure 2, 3**). In addition to that, we found that while control nodules had genes in groups C and E downregulated at 21 dpi, the salt-stressed nodules had genes in group A downregulated as well (**Table 1**). Interestingly, these were the genes that were upregulated in the presence of salt at 14 dpi (**Figure 1**). These trends reveal that the temporal expression dynamics of host genes are changed under salinity.

Distinct metabolites accumulate in the control and salt-stressed nodules.

Since salt shifted the expression dynamics of several nodulation genes, we wondered if this intersection between salt signaling and nodule development would be reflected in the accumulation of metabolites in the nodules. To this end, we used matrix-assisted laser desorption/ionization-mass spectrometry imaging (MALDI-MSI), a powerful tool for determining the spatial distribution of metabolites within nodules.[58] To confidently identify small metabolites, MS/MS was performed on *m/z* consistently increased in either the control or salt-stressed nodules. Salt-stressed nodules and the subtending roots showed increased accumulation of proline betaine and homostachydrine, but reduced heme accumulation. Control roots adjacent to the nodule accumulated the flavonoid, formononetin, which was not observed in the salt-stressed roots (**Figure 4**). Several other metabolites showed distinct spatial distribution patterns unique to a treatment, indicating that salinity causes a major metabolic shift in the nodules (**Figure 4**). One technical issue encountered was that the increased availability of sodium in the salt nodules could create changes in the ion images that are due to the increased sodium ions rather than an actual change in concentration of the molecule. However, these artificial changes were found and ruled out.[64] **Supplemental Table S2** provides accurate mass and MS/MS details for the identified metabolites.

ABA is the primary mediator of salt stress and accumulates in high amounts in salt-stressed tissues.[42] We were curious to see whether an increased accumulation of ABA could be seen in salt-stressed nodules as well. To this end, we quantified ABA in young (white) and mature (pink) nodules that developed under control or salt-stressed conditions. We observed a significant decrease in ABA levels between young and mature nodules in the control condition that was absent from the salt-stressed nodules. ABA accumulation in the salt-stressed nodules were comparable between young and mature stages (**Figure 5**), indicating that it accumulates differently in the presence of salt.

**Discussion**

We have previously shown that salt interferes with early rhizobium-legume signaling by transcriptional regulation and that it inhibits rhizobial infection in the epidermis (Chakraborty et al., in prep.). Here, we wanted to inquire whether salinity also interferes with nodulation at later developmental stages. We find that salt stress transcriptionally regulates several host genes involved in nodule differentiation and inhibits the expression of the rhizobial bacA and nifH genes. These data indicate that nodule development is delayed or blocked under salinity. Our study further reveals major metabolic shift during nodule maturation under salinity.

Transcriptional alterations under salinity suggest either a delay or a block in nodule development.

We observed upregulation of several genes involved in the early stages of nodule development, such as early rhizobium-legume signaling, meristem development and infection, and a simultaneous reduction in the expression of genes involved in the later stages, such as bacteroid differentiation, heme degradation and nitrogen fixation (**Figures 1, 2, 3**). These results suggest a delay or block in the development of nodule under salt stress. Salinity has earlier been shown to

delay developmental processes, such as flowering time in rice (Oryza sativa)[65] and primary xylem differentiation in soybean (Glycine max) roots.[66] Delaying developmental processes is a common strategy employed by plants in the face of imminent stress, and often involves microRNAs (miRNA).[67] miRNAs post-transcriptionally silence target RNAs, often transcription factors, and are differentially expressed under stressful conditions.[68] We observed an upregulation of several genes encoding the transcription factors NF-YA1, WOX5, ERN1 and ERN2. WOX5 overexpression causes prolonged meristematic activity in the nodule compared to the wild-type,[18] further corroborating the idea that salt delays nodule maturation. NF-YA1 is post-transcriptionally regulated by miR169 in nodules.[19] In rice, several members of the miR169 family are induced, and NF-YA expression repressed, by salinity[69] In A. thaliana, miR169 targets members of the NF-YA family and MIR169a overexpression renders the plants hypersensitive to nitrogen starvation.[70] These findings clearly reveal miR169 in an interesting position in salt regulation of nodulation.

We observed decreased expression of the heme oxygenase, GIRAFFE, in 14 dpi salt-stressed nodules (**Figure 2**). The A. thaliana heme oxygenase, HY1, participates in salt acclimation by K+ retention via H+-ATPase and transcriptional regulation of the Salt Overly Sensitive 1 (SOS1) gene encoding a plasma membrane-localized Na+/H+ antiporter,[71] and is required for salt acclimation.[72, 73] A decrease in GIRAFFE expression could indicate an inability of the nodules to adjust to the highly saline environment. Alternatively, it could be a part of a feedback mechanism that prevents further salt-induced damage (i.e. involved in salt acclimation). Contrary to our findings, in soybean nodules, 100 mM NaCl induces heme oxygenase 1 (HO-1), a gene suggested to be closely related to mammalian HO-1 based on induction patterns.[74] The apparent disparity between these two results could be attributed to different nodule developmental patterns in M. truncatula and soybean, where the former develops indeterminate nodules harboring a persistent

meristem. Meristem-derived cells gradually become infected with rhizobia and differentiate into nitrogen-fixing cells. If salinity delayed nodule maturation as we suggest, it would cause a delayed accumulation of leghemoglobin. GIRAFFE is also involved in heme degradation.[36] Hence, the reduction in GIRAFFE expression could be an outcome of delayed nodule development under salt stress, or an acclimation to salinity.

MALDI-MSI reveals distinct distribution of key metabolites in salt-stressed nodules.

The distinct phenotype of salt-stressed nodules suggests major physiological alterations that underlie this change. We used MALDI-MSI-based metabolomics to get an insight into the biochemical pathways that might be differentially regulated in the presence of salt (**Figure 4**). Our results demonstrate distinct spatial accumulation of the metabolites, indicating their relative contribution during a particular stage of development (**Figure 4**). We detected very little heme accumulation in salt-stressed nodules (**Figure 4C**), which might result in reduced expression of GIRAFFE, further corroborating the idea that development of nodules is delayed under salt stress. Alternatively, it is possible that very little accumulation of heme causes reduced GIRAFFE expression via a feedback mechanism, and in that case, GIRAFFE expression would not be a reflection of any delay in nodule development.

We also observed increased accumulation of proline betaine (stachydrine) and pipecolatebetaine (homostachydrine) in salt-stressed nodules (**Figure 4B, D**). These quaternary ammonium compounds are two amongst some key compatible osmolytes that accumulate in high concentration under salt stress.[75] Proline betaine is an osmoprotectant that accumulates in alfalfa (Medicago sativa) nodules under salt stress, particularly in the cytosol and bacteroids, and is thought to be associated with an increase in turgor pressure in the symbiosomes.[45] In free-living R. meliloti 102F34, high osmolarity in the environment enhances the uptake of proline betaine,

which in turn, helps in the revival of growth rate from salinity-induced inhibition.[76] Overexpression of betS, a bacterial gene encoding a high affinity glycine and proline betaine transporter, relieves bacteroids isolated from alfalfa from some of the negative effects of salt stress.[77] It remains to be tested whether an increased uptake of certain host metabolite influences differentiation into bacteroids.

We found a correlation between a decrease in ABA content and maturation of nodules (**Figure 5**). Interestingly, this difference is lost in nodules that develop under salt-stress. At the same time, the latter accumulated high amounts of proline betaine (**Figure 4D**). In sorghum (Sorghum bicolor) leaves, ABA induces betaine aldehyde dehydrogenase, an enzyme involved in the synthesis of glycine betaine.[78] It would be interesting to test what effect an inhibition of ABA synthesis could have on the synthesis of proline betaine.

In conclusion, our work glimpses into the interaction between salinity and nodule maturation in the indeterminate nodulator M. truncatula. Given the gradient of differentiation the nodules exhibit, it would be interesting to examine host cell division patterns in the developing nodule and bacteroids ultrastructure. In addition, it will be exciting to test whether ABA accumulation plays a role in bacteroid differentiation.

**Acknowledgments**

J.M. and D.J. from the Ané lab grew *M. truncatula* seedlings for nodules, and C.K. performed MALDI-MSI. E.M. grew the seedlings for nodules and performed ABA ELISA with S.C.. A.L. grew and inoculated plants and performed histochemical assays for *bacA* and *nifH*. This work was supported in part by funding from the National Science Foundation (NSF) Division of Integrative Organismal Systems (IOS) RESEARCH PGR award #1546742, University of

Wisconsin-Madison Graduate School and the Wisconsin Alumni Research Foundation (WARF), a Vilas Distinguished Achievement Professorship to L.L., and a NSF grant to JMA (NSF#0701846). LL acknowledges funding support from NIH through grants R56MH110215 and R01 DK071801. The MALDI LTQ Orbitrap XL was purchased through an NIH shared instrument grant (NCRR S10RR029531 to LL).

**References**

1. Zahran, H. H., Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol Mol Biol Rev* **1999,** *63* (4), 968-89, table of contents.

2. Ferguson, B. J.; Indrasumunar, A.; Hayashi, S.; Lin, M. H.; Lin, Y. H.; Reid, D. E.; Gresshoff, P. M., Molecular analysis of legume nodule development and autoregulation. *J Integr Plant Biol* **2010,** *52* (1), 61-76.

3. Vernie, T.; Kim, J.; Frances, L.; Ding, Y.; Sun, J.; Guan, D.; Niebel, A.; Gifford, M. L.; de Carvalho-Niebel, F.; Oldroyd, G. E., The NIN Transcription Factor Coordinates Diverse Nodulation Programs in Different Tissues of the Medicago truncatula Root. *Plant Cell* **2015,** *27* (12), 3410-24.

4. Plet, J.; Wasson, A.; Ariel, F.; Le Signor, C.; Baker, D.; Mathesius, U.; Crespi, M.; Frugier, F., MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in Medicago truncatula. *Plant J* **2011,** *65* (4), 622-33.

5. Ariel, F.; Brault-Hernandez, M.; Laffont, C.; Huault, E.; Brault, M.; Plet, J.; Moison, M.; Blanchet, S.; Ichante, J. L.; Chabaud, M.; Carrere, S.; Crespi, M.; Chan, R. L.; Frugier, F., Two direct targets of cytokinin signaling regulate symbiotic nodulation in Medicago truncatula. *Plant Cell* **2012,** *24* (9), 3838-52.

6. Xiao, T. T.; Schilderink, S.; Moling, S.; Deinum, E. E.; Kondorosi, E.; Franssen, H.; Kulikova, O.; Niebel, A.; Bisseling, T., Fate map of Medicago truncatula root nodules. *Development* **2014,** *141* (18), 3517-28.

7. Monahan-Giovanelli, H.; Pinedo, C. A.; Gage, D. J., Architecture of infection thread networks in developing root nodules induced by the symbiotic bacterium Sinorhizobium meliloti on Medicago truncatula. *Plant Physiol* **2006,** *140* (2), 661-70.

8. Vasse, J.; de Billy, F.; Camut, S.; Truchet, G., Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J Bacteriol* **1990,** *172* (8), 4295-306.

9. Gage, D. J., Analysis of infection thread development using Gfp- and DsRed-expressing Sinorhizobium meliloti. *J Bacteriol* **2002,** *184* (24), 7042-6.

10. Gage, D. J., Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol Mol Biol Rev* **2004,** *68* (2), 280-300.

11. Mortier, V.; Den Herder, G.; Whitford, R.; Van de Velde, W.; Rombauts, S.; D'Haeseleer, K.; Holsters, M.; Goormachtig, S., CLE Peptides Control Medicago truncatula Nodulation Locally and Systemically. *Plant Physiology* **2010,** *153* (1), 222-237.

12. Mortier, V.; De Wever, E.; Vuylsteke, M.; Holsters, M.; Goormachtig, S., Nodule numbers are governed by interaction between CLE peptides and cytokinin signaling. *Plant Journal* **2012,** *70* (3), 367-376.

13. Azarakhsh, M.; Lebedeva, M. A.; Lutova, L. A., Identification and Expression Analysis of Medicago truncatula Isopentenyl Transferase Genes (IPTs) Involved in Local and Systemic Control of Nodulation. *Front Plant Sci* **2018,** *9*, 304.

14. C, Z., Abscisic Acid and Nitrate Transporter Mtlatd/nip signalling in Root and Nodule Development in *Medicago Truncatula*. University of Vermont: 2015.

15. Bright, L. J.; Liang, Y.; Mitchell, D. M.; Harris, J. M., The LATD gene of Medicago truncatula is required for both nodule and root development. *Mol Plant Microbe Interact* **2005,** *18* (6), 521-32.

16. Yendrek, C. R.; Lee, Y. C.; Morris, V.; Liang, Y.; Pislariu, C. I.; Burkart, G.; Meckfessel, M. H.; Salehin, M.; Kessler, H.; Wessler, H.; Lloyd, M.; Lutton, H.; Teillet, A.; Sherrier, D. J.; Journet, E. P.; Harris, J. M.; Dickstein, R., A putative transporter is essential for integrating nutrient and hormone signaling with lateral root growth and nodule development in Medicago truncatula. *Plant J* **2010,** *62* (1), 100-12.

17. Limpens, E.; Moling, S.; Hooiveld, G.; Pereira, P. A.; Bisseling, T.; Becker, J. D.; Kuster, H., cell- and tissue-specific transcriptome analyses of Medicago truncatula root nodules. *PLoS One* **2013,** *8* (5), e64377.

18. Osipova, M. A.; Mortier, V.; Demchenko, K. N.; Tsyganov, V. E.; Tikhonovich, I. A.; Lutova, L. A.; Dolgikh, E. A.; Goormachtig, S., Wuschel-related homeobox5 gene expression and interaction of CLE peptides with components of the systemic control add two pieces to the puzzle of autoregulation of nodulation. *Plant Physiol* **2012,** *158* (3), 1329-41.

19. Combier, J. P.; Frugier, F.; de Billy, F.; Boualem, A.; El-Yahyaoui, F.; Moreau, S.; Vernie, T.; Ott, T.; Gamas, P.; Crespi, M.; Niebel, A., MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula. *Genes Dev* **2006,** *20* (22), 3084-8.

20. Andriankaja, A.; Boisson-Dernier, A.; Frances, L.; Sauviac, L.; Jauneau, A.; Barker, D. G.; de Carvalho-Niebel, F., AP2-ERF transcription factors mediate Nod factor dependent Mt ENOD11 activation in root hairs via a novel cis-regulatory motif. *Plant Cell* **2007,** *19* (9), 2866-85.

21. Laporte, P.; Lepage, A.; Fournier, J.; Catrice, O.; Moreau, S.; Jardinaud, M. F.; Mun, J. H.; Larrainzar, E.; Cook, D. R.; Gamas, P.; Niebel, A., The CCAAT box-binding transcription factor NF-YA1 controls rhizobial infection. *J Exp Bot* **2014,** *65* (2), 481-94.

22. Domonkos, A.; Kovacs, S.; Gombar, A.; Kiss, E.; Horvath, B.; Kovats, G. Z.; Farkas, A.; Toth, M. T.; Ayaydin, F.; Boka, K.; Fodor, L.; Ratet, P.; Kereszt, A.; Endre, G.; Kalo, P., NAD1 Controls Defense-Like Responses in Medicago truncatula Symbiotic Nitrogen Fixing Nodules Following Rhizobial Colonization in a BacA-Independent Manner. *Genes (Basel)* **2017,** *8* (12).

23. Starker, C. G.; Parra-Colmenares, A. L.; Smith, L.; Mitra, R. M.; Long, S. R., Nitrogen fixation mutants of Medicago truncatula fail to support plant and bacterial symbiotic gene expression. *Plant Physiol* **2006,** *140* (2), 671-80.

24. Mergaert, P.; Nikovics, K.; Kelemen, Z.; Maunoury, N.; Vaubert, D.; Kondorosi, A.; Kondorosi, E., A novel family in Medicago truncatula consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiol* **2003,** *132* (1), 161-73.

25. Pan, H.; Wang, D., Nodule cysteine-rich peptides maintain a working balance during nitrogen-fixing symbiosis. *Nat Plants* **2017,** *3* (5), 17048.

26. Horvath, B.; Domonkos, A.; Kereszt, A.; Szucs, A.; Abraham, E.; Ayaydin, F.; Boka, K.; Chen, Y.; Chen, R.; Murray, J. D.; Udvardi, M. K.; Kondorosi, E.; Kalo, P., Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the Medicago truncatula dnf7 mutant. *Proc Natl Acad Sci U S A* **2015,** *112* (49), 15232-7.

27. Kim, M.; Chen, Y.; Xi, J.; Waters, C.; Chen, R.; Wang, D., An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proc Natl Acad Sci U S A* **2015,** *112* (49), 15238-43.

28. Haag, A. F.; Baloban, M.; Sani, M.; Kerscher, B.; Pierre, O.; Farkas, A.; Longhi, R.; Boncompagni, E.; Herouart, D.; Dall'angelo, S.; Kondorosi, E.; Zanda, M.; Mergaert, P.; Ferguson, G. P., Protection of Sinorhizobium against host cysteine-rich antimicrobial peptides is critical for symbiosis. *PLoS Biol* **2011,** *9* (10), e1001169.

29. Mergaert, P., Role of antimicrobial peptides in controlling symbiotic bacterial populations. *Nat Prod Rep* **2018,** *35* (4), 336-356.

30. Glazebrook, J.; Ichige, A.; Walker, G. C., A Rhizobium meliloti homolog of the Escherichia coli peptide-antibiotic transport protein SbmA is essential for bacteroid development. *Genes Dev* **1993,** *7* (8), 1485-97.

31. Kereszt, A.; Mergaert, P.; Kondorosi, E., Bacteroid development in legume nodules: evolution of mutual benefit or of sacrificial victims? *Mol Plant Microbe Interact* **2011,** *24* (11), 1300-9.

32. Dakora, F. D., A functional relationship between leghaemoglobin and nitrogenase based on novel measurements of the two proteins in legume root nodules. *Ann Bot* **1995,** *75* (1), 49-54.

33. Roponen, I., The effect of darkness on the leghemoglobin content and amino acid levels in the root nodules of pea plants. *Physiologia Plan-tarum.* **1970,** *23* (3), 452-460.

34. Shekhawat, G. S.; Verma, K., Haem oxygenase (HO): an overlooked enzyme of plant metabolism and defence. *J Exp Bot* **2010,** *61* (9), 2255-70.

35. Ghosh, P., Role of Heme Oxygenase in modulating expression of ROS-regulatory enzymes in Medicago truncatula. 2014.

36. Zhang, Y., Light Regulation of legume nodulation through a plant hormone signaling network. University of Vermont: 2014.

37. Tu, J., Effect of salinity on Rhizobium-root hair interaction, nodulation and growth of soybean. Canadian Journal of Plant Science: 1981; Vol. 61, pp 231-239.

38. Singleton, P. W.; Bohlool, B. B., Effect of salinity on nodule formation by soybean. *Plant Physiol* **1984,** *74* (1), 72-6.

39. Zahran, H. H.; Sprent, J. I., Effects of sodium chloride and polyethylene glycol on root-hair infection and nodulation of Vicia faba L. plants by Rhizobium leguminosarum. *Planta* **1986,** *167* (3), 303-9.

40. Lluch, C.; Ligero, F.; Cordovilla, M. P., The effect of salinity on N fixation and assimilation in Vicia faba. *Journal of Experimental Botany* **1994,** *45* (10), 1483-1488.

41. Borucki, W.; Sujkowska, M., The effects of sodium chloride-salinity upon growth, nodulation, and root nodule structure of pea (Pisum sativum L.) plants. *Acta Physiologiae Plantarum* **2008,** *30* (3), 293-301.

42. Jia, W.; Wang, Y.; Zhang, S.; Zhang, J., Salt-stress-induced ABA accumulation is more sensitively triggered in roots than in shoots. *J Exp Bot* **2002,** *53* (378), 2201-6.

43. Geng, Y.; Wu, R.; Wee, C. W.; Xie, F.; Wei, X.; Chan, P. M.; Tham, C.; Duan, L.; Dinneny, J. R., A spatio-temporal understanding of growth regulation during the salt stress response in Arabidopsis. *Plant Cell* **2013,** *25* (6), 2132-54.

44. Fougere, F.; Le Rudulier, D.; Streeter, J. G., Effects of Salt Stress on Amino Acid, Organic Acid, and Carbohydrate Composition of Roots, Bacteroids, and Cytosol of Alfalfa (Medicago sativa L.). *Plant Physiol* **1991,** *96* (4), 1228-36.

45. Trinchant, J. C.; Boscari, A.; Spennato, G.; Van de Sype, G.; Le Rudulier, D., Proline betaine accumulation and metabolism in alfalfa plants under sodium chloride stress. Exploring its compartmentalization in nodules. *Plant Physiol* **2004,** *135* (3), 1583-94.

46. Verdoy, D.; Coba De La Pena, T.; Redondo, F. J.; Lucas, M. M.; Pueyo, J. J., Transgenic Medicago truncatula plants that accumulate proline display nitrogen-fixing activity with enhanced tolerance to osmotic stress. *Plant Cell Environ* **2006,** *29* (10), 1913-23.

47. Choudhury, F. K.; Rivero, R. M.; Blumwald, E.; Mittler, R., Reactive oxygen species, abiotic stress and stress combination. *Plant J* **2017,** *90* (5), 856-867.

48. Lambeth, J. D., NOX enzymes and the biology of reactive oxygen. *Nat Rev Immunol* **2004,** *4* (3), 181-9.

49. Marino, D.; Dunand, C.; Puppo, A.; Pauly, N., A burst of plant NADPH oxidases. *Trends Plant Sci* **2012,** *17* (1), 9-15.

50. Cabeza, R.; Koester, B.; Liese, R.; Lingner, A.; Baumgarten, V.; Dirks, J.; Salinas-Riester, G.; Pommerenke, C.; Dittert, K.; Schulze, J., An RNA sequencing transcriptome analysis reveals novel insights into molecular aspects of the nitrate impact on the nodule activity of Medicago truncatula. *Plant Physiol* **2014,** *164* (1), 400-11.

51. Zhang, C.; Bousquet, A.; Harris, J. M., Abscisic acid and lateral root organ defective/NUMEROUS INFECTIONS AND POLYPHENOLICS modulate root elongation via reactive oxygen species in Medicago truncatula. *Plant Physiol* **2014,** *166* (2), 644-58.

52. Chen, T.; Duan, L.; Zhou, B.; Yu, H.; Zhu, H.; Cao, Y.; Zhang, Z., Interplay of Pathogen-Induced Defense Responses and Symbiotic Establishment in Medicago truncatula. *Front Microbiol* **2017,** *8*, 973.

53. Marino, D.; Andrio, E.; Danchin, E. G.; Oger, E.; Gucciardo, S.; Lambert, A.; Puppo, A.; Pauly, N., A Medicago truncatula NADPH oxidase is involved in symbiotic nodule functioning. *New Phytol* **2011,** *189* (2), 580-92.

54. Kurusu, T.; Kuchitsu, K.; Tada, Y., Plant signaling networks involving Ca(2+) and Rboh/Nox-mediated ROS production under salinity stress. *Front Plant Sci* **2015,** *6*, 427.

55. Meade, H. M.; Long, S. R.; Ruvkun, G. B.; Brown, S. E.; Ausubel, F. M., Physical and genetic characterization of symbiotic and auxotrophic mutants of Rhizobium meliloti induced by transposon Tn5 mutagenesis. *J Bacteriol* **1982,** *149* (1), 114-22.

56. Swanson, J. A.; Mulligan, J. T.; Long, S. R., Regulation of syrM and nodD3 in Rhizobium meliloti. *Genetics* **1993,** *134* (2), 435-44.

57. Kakar, K.; Wandrey, M.; Czechowski, T.; Gaertner, T.; Scheible, W. R.; Stitt, M.; Torres-Jerez, I.; Xiao, Y.; Redman, J. C.; Wu, H. C.; Cheung, F.; Town, C. D.; Udvardi, M. K., A community resource for high-throughput quantitative RT-PCR analysis of transcription factor gene expression in Medicago truncatula. *Plant Methods* **2008,** *4*, 18.

58. Gemperline, E.; Li, L., MALDI-mass spectrometric imaging for the investigation of metabolites in Medicago truncatula root nodules. *J Vis Exp* **2014,** (85).

59. Robichaud, G.; Garrard, K. P.; Barry, J. A.; Muddiman, D. C., MSiReader: An Open-Source Interface to View and Analyze High Resolving Power MS Imaging Files on Matlab Platform. *Journal of the American Society for Mass Spectrometry* **2013,** *24* (5), 718-721.

60. Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S., MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform* **2016,** *8*, 3.

61. Xie, F.; Murray, J. D.; Kim, J.; Heckmann, A. B.; Edwards, A.; Oldroyd, G. E.; Downie, J. A., Legume pectate lyase required for root infection by rhizobia. *Proc Natl Acad Sci U S A* **2012,** *109* (2), 633-8.

62. Fournier, J.; Teillet, A.; Chabaud, M.; Ivanov, S.; Genre, A.; Limpens, E.; de Carvalho-Niebel, F.; Barker, D. G., Remodeling of the infection chamber before infection thread formation reveals a two-step mechanism for rhizobial entry into the host legume root hair. *Plant Physiol* **2015,** *167* (4), 1233-42.

63. Karunakaran, R.; Haag, A. F.; East, A. K.; Ramachandran, V. K.; Prell, J.; James, E. K.; Scocchi, M.; Ferguson, G. P.; Poole, P. S., BacA is essential for bacteroid development in nodules of galegoid, but not phaseoloid, legumes. *J Bacteriol* **2010,** *192* (11), 2920-8.

64. Keller, C.; Maeda, J.; Jayaraman, D.; Chakraborty, S.; Sussman, M. R.; Harris, J. M.; Ané, J.-M.; Li, L., Comparison of Vacuum MALDI and AP-MALDI Platforms for the Mass Spectrometry Imaging of Metabolites Involved in Salt Stress in Medicago truncatula. *Frontiers in Plant Science* **2018,** *9*, 1238.

65. Bouharmont, J.; Kinet, J. M.; Lutts, S., Changes in plant response to NaCl during development of rice (Oryza sativa L.) varieties differing in salinity resistance. *Journal of Experimental Botany* **1995,** *46* (12), 1843-1852.

66. Hilal, M.; Zenoff, A. M.; Ponessa, G.; Moreno, H.; Massa, E. M., Saline stress alters the temporal patterns of xylem differentiation and alternative oxidase expression in developing soybean roots. *Plant Physiol* **1998,** *117* (2), 695-701.

67. Cui, L. G.; Shan, J. X.; Shi, M.; Gao, J. P.; Lin, H. X., The miR156-SPL9-DFR pathway coordinates the relationship between development and abiotic stress tolerance in plants. *Plant J* **2014,** *80* (6), 1108-17.

68. Shriram, V.; Kumar, V.; Devarumath, R. M.; Khare, T. S.; Wani, S. H., MicroRNAs As Potential Targets for Abiotic Stress Tolerance in Plants. *Front Plant Sci* **2016,** *7*, 817.

69. Zhao, B.; Ge, L.; Liang, R.; Li, W.; Ruan, K.; Lin, H.; Jin, Y., Members of miR-169 family are induced by high salinity and transiently inhibit the NF-YA transcription factor. *BMC Mol Biol* **2009,** *10*, 29.

70. Zhao, M.; Ding, H.; Zhu, J. K.; Zhang, F.; Li, W. X., Involvement of miR169 in the nitrogen-starvation responses in Arabidopsis. *New Phytol* **2011,** *190* (4), 906-15.

71. Bose, J.; Xie, Y.; Shen, W.; Shabala, S., Haem oxygenase modifies salinity tolerance in Arabidopsis by controlling K(+) retention via regulation of the plasma membrane H(+)-ATPase and by altering SOS1 transcript levels in roots. *J Exp Bot* **2013,** *64* (2), 471-81.

72. Xie, Y. J.; Xu, S.; Han, B.; Wu, M. Z.; Yuan, X. X.; Han, Y.; Gu, Q.; Xu, D. K.; Yang, Q.; Shen, W. B., Evidence of Arabidopsis salt acclimation induced by up-regulation of HY1 and the regulatory role of RbohD-derived reactive oxygen species synthesis. *Plant J* **2011,** *66* (2), 280-92.

73. Xie, Y.; Cui, W.; Yuan, X.; Shen, W.; Yang, Q., Heme oxygenase-1 is associated with wheat salinity acclimation by modulating reactive oxygen species homeostasis. *J Integr Plant Biol* **2011,** *53* (8), 653-70.

74. Zilli, C. G.; Balestrasse, K. B.; Yannarelli, G. G.; Polizio, A. H.; Santa-Cruz, D. M.; Tomaro, M. L., Heme oxygenase up-regulation under salt stress protects nitrogen metabolism in nodules of soybean plants. *Environmental and Experimental Botany* **2008,** *64* (1), 83-89.

75. Ashraf, M.; Harris, P. J. C., Potential biochemical indicators of salinity tolerance in plants. *Plant Science* **2004,** *166* (1), 3-16.

76. Gloux, K.; Le Rudulier, D., Transport and catabolism of proline betaine in salt-stressed Rhizobium meliloti. *Archives of Microbiology* **1989,** *151* (2), 143-148.

77. Boscari, A.; Van de Sype, G.; Le Rudulier, D.; Mandon, K., Overexpression of BetS, a Sinorhizobium meliloti high-affinity betaine transporter, in bacteroids from Medicago sativa nodules sustains nitrogen fixation during early salt stress adaptation. *Mol Plant Microbe Interact* **2006,** *19* (8), 896-903.

78. Saneoka, H.; Ishiguro, S.; Moghaieb, R. E. A., Effect of salinity and abscisic acid on accumulation of glycinebetaine and betaine aldehyde dehydrogenase mRNA in Sorghum leaves (Sorghum bicolor). *Journal of Plant Physiology* **2001,** *158* (7), 853-859.

**Tables**

**Table 1: Patterns of nodule gene expression in response to salt stress.**

| GROUP | 14dpi: salt vs control | 21 vs 14dpi (control) | 21 vs 14dpi (salt) | Genes |
|---|---|---|---|---|
| A | **up** | none | **down** | *ENOD11, NF-YA1, WOX5, ERN1, ERN2, RR4, NPL, MtRbohC* |
| B | none | none | none | *LATD, MtRbohA, CLE13* |
| C | none | **down** | **down** | *NAD1, DNF4* |
| D | **down** | none | none | *GIRAFFE* |
| E | **down** | **down** | **down** | *DNF7* |

## Figures

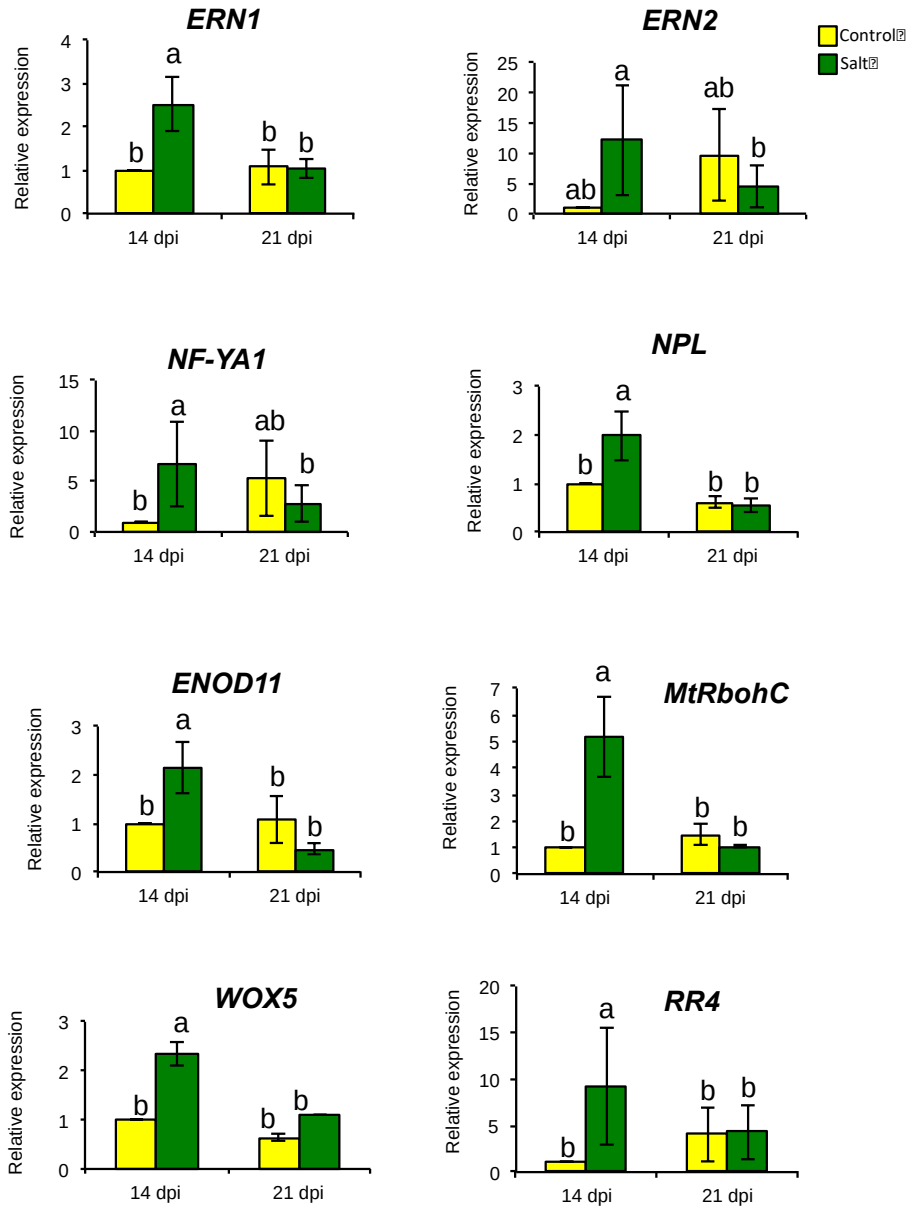



**Figure 1.** Salinity upregulates *RR4* and several meristem, early signaling, and infection-associated genes in nodules. Transcript levels measured by quantitative RT-PCR. Bars represent mean of three independent experiments +/- SEM. Statistical analysis done using a 2-way ANOVA followed by a t-test on the LS-Means (α=0.05). Columns not connected by the same letter are significantly different.

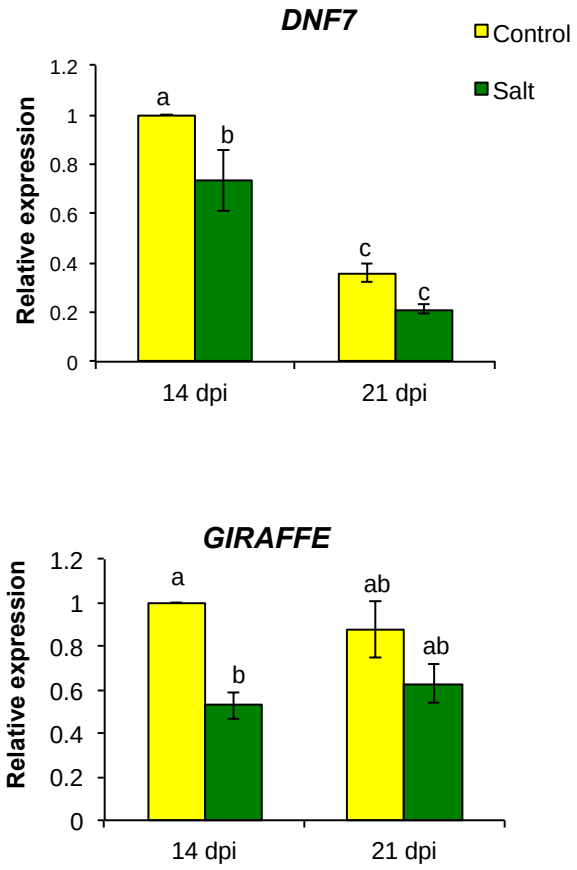



**Figure 2.** Salinity downregulates *DNF7* and *GIRAFFE* at 14 dpi. Transcript levels measured by quantitative RT-PCR. Bars represent mean of three independent experiments +/- SEM. Statistical analysis done using a 2-way ANOVA followed by a t-test on the LS-Means ($\alpha$=0.05). Columns not connected by the same letter are significantly different.

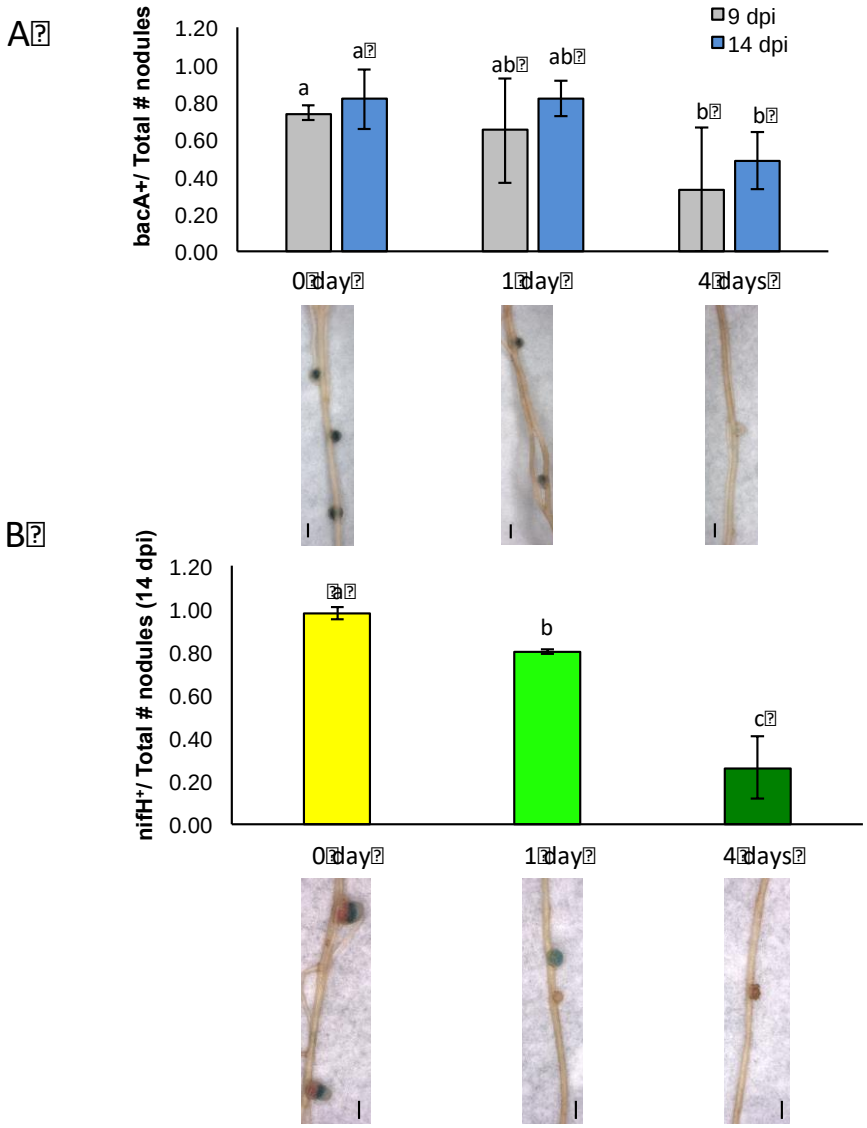



**Figure 3.** Decreased expression of rhizobial genes required for bacteroid differentiation and nitrogen fixation with increasing exposure to salt stress. *bacA* (A) and *nifH* (B) promoter activity as measured by *uidA* (*GUS*) reporter gene expression. Graph representing the mean of three independent experiments with *n*= 14 seedlings per treatment per experiment. Representative images are shown underneath each column. Days underneath the graphs represent duration of salt treatment at the time of inoculation. Scale bar=1 mm. Statistical analysis done using a 2-way ANOVA followed by Tukey's HSD test on the LS-Means (α=0.05). Columns not connected by the same letter are significantly different.

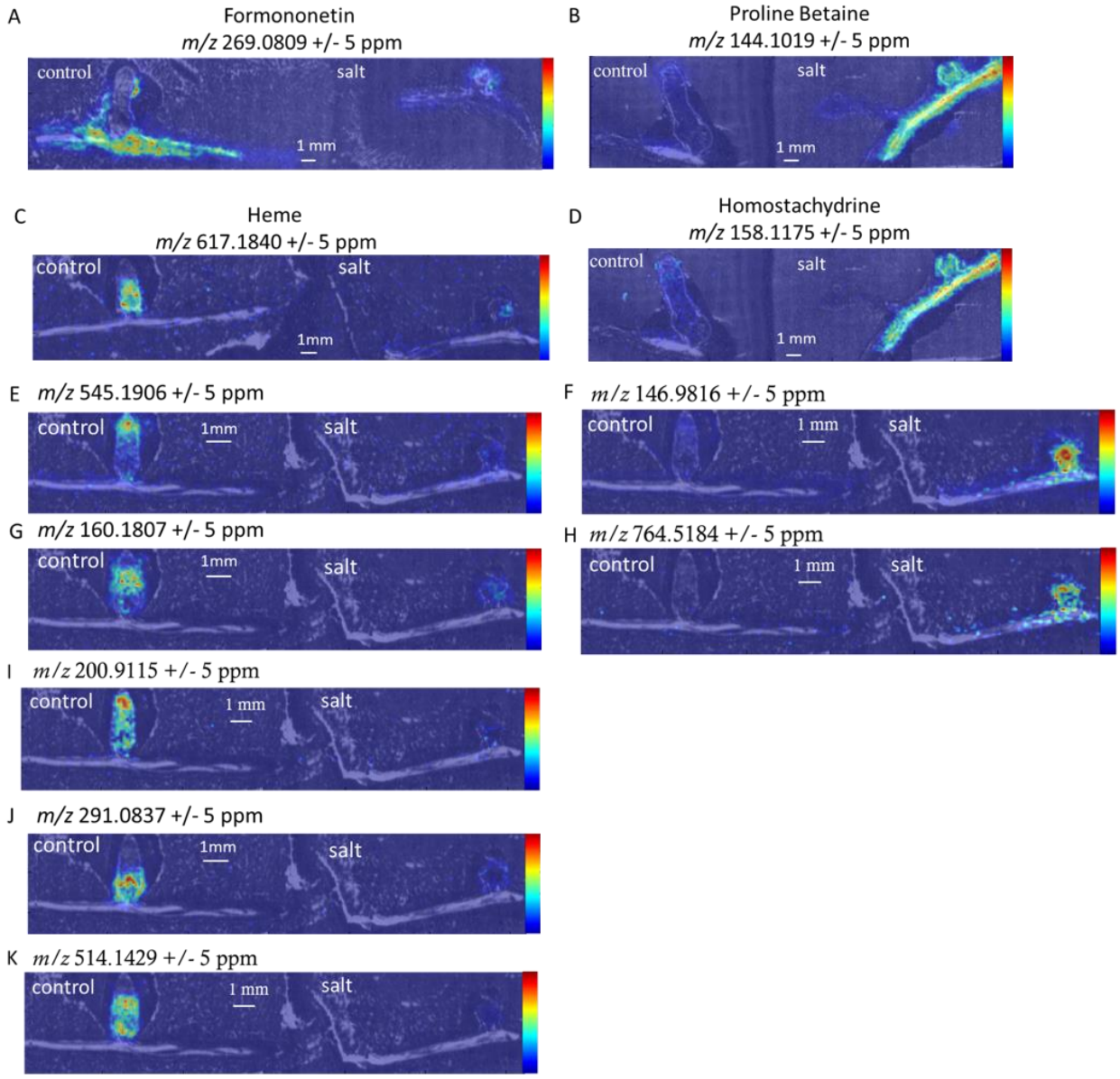



**Figure 4.** Salinity alters accumulation of small metabolites in nodules. MALDI-MSI images for *m/z* present in either control or salt-stressed nodules or roots **(A-K)**. Ion images are overlaid with the optical images. Left and right panels show accumulation of metabolites exclusively in control or salt-stressed nodules or roots, respectively. Accumulation of formononetin **(A)** and heme **(C)** are observed only in the control condition; accumulation of proline betaine **(B)** and homostachydrine **(D)** are observed only under salt stress. Accumulation of unidentified small metabolites associated with definite stages of nodule development that are not observed under salt stress **(E, G, I, J and K)** or only observed under salinity **(F, H)**. Metabolites studied from 21 dpi nodules.

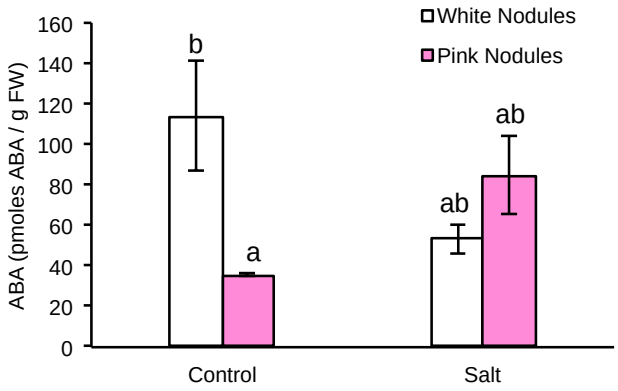



**Figure 5.** Differential ABA accumulation in the control and salt-stressed nodules. ABA quantification from 21 dpi nodules as determined by ELISA. Data represent the average of three independent experiments. All reactions were done in duplicate for a technical replicate. Statistical analysis done using a 2-way ANOVA followed by Student's t- test on the LS-Means ($\alpha$=0.05). Columns not connected by the same letter are significantly different. Error bars represent SEM.

## Supplemental

### Supplemental Table S1: Primers for qRT-PCR.

| Gene | Forward (5'-3') | Reverse (5'-3') | Reference |
|---|---|---|---|
| *CLE13* | TCAACTTTGCAGGCT CGTAG | TGTAGAAGGCTTCGG CTGTC | (Zhang, 2014) |
| *LATD* | TCGGCCAGAAAGAAAG ACGAT | GTTCTTCTCTCGATGAC CTAAACA | This work. |
| *NF-YA1* | TCGGATCTACTGTCCAC TCTTTGG | TTGGCATGACGATACCG TGTC | (Fonouni-Farde et al., 2016) |
| *WOX5* | GGAATCCAACAACAGA ACAAGTTAAA | TGAATTTGATCAGTGCT TGGAGTT | (Imin et al., 2007) |
| *ENOD11* | TAGGGCTTGCTGATA AATCTC | TAATTGGAGGCTTGT AAGTAG | (Ding et al., 2008) |
| *ERN1* | TGTCTCATTGGATTC TCCTCTTGC | TTGGAGCAGAAGCA ACAGCA | (Cerri et al., 2012) |
| *ERN2* | CTAGCCATGCTGACA CGACTAAT | ACTGGCTGTGCCAAT ACAGTTAG | (Cerri et al., 2012) |
| *NPL* | CGTGCATTCCCATAAAG GGC | TGTTAAATGCCACCGTT CCG | (Chakraborty et al., 2018, in prep) |
| *RR4* | ATGCTTTTGTTCCGG GTTTA | CTGCACCTTCCTCCA AACAT | (Zhang, 2014) |
| *MtRbohA* | GCGCAACTCC TTTGATTAGC | GAAATAGGCT CGCTTGGTTG | (Marino et al., 2011) |
| *MtRbohC* | GGGAGACCTGATGCTAT TCAAG | TGTCTTCAACAATAATG TCCATCTG | (Marino et al., 2011) |
| *NAD1* | TCAAGAAGTGTTGTGGG CAG | GCAATCTTGGTGCTACT TTGG | (Wang et al., 2016). |
| *DNF4* | GCGTTAATGCACATCAT TTGTGG | TTATTCTCGGACACAAA CACCTTG | (Horváth et al., 2015) |
| *DNF7* | GGAAATGCGTTGAAAAT GTTTGTG | AACATTTCTCCACTTTA TTCTCGGG | (Horváth et al., 2015) |
| *GIRAFFE* | TGGGAAGAAG ATTGCTGGTC | TGTCCCTCACA TTCTGCAAC | (Zhang, 2014) |
| *Helicase* | GTACGAGGTCGGTGCTC TTGAA | GCAACCGAAAATTGCA CCATAC | (Kakar et al., 2008) |
| *UBC9* | GGTTGATTGCTC TTCTCTCCCC | AAGTGATTGCTC GTCCAACCC | (Kakar et al., 2008) |

**Supplemental Table S2.** Identifications in Control and Salt nodules (*m/z* are for $[M+H]^+$ adduct). Asterisk* indicate identification confirmed with obtained standard.

| ***m/z*** | **Adduct** | **Identification** | **Theoretical *m/z*** | **Delta ppm** | **Source of MS/MS** |
|---|---|---|---|---|---|
| 144.1018 | $[M+H]^+$<br>$[M+Na]^+$<br>$[M+K]^+$ | Proline Betaine (Stachydrine)* | 144.1019 | -1.03 | 1. On-tissue (CID, ion trap)<br>2.LC-MS/MS (HCD) |
| 158.1175 | $[M+H]^+$<br>$[M+Na]^+$<br>$[M+K]^+$ | Homostachydrine | 158.1176 | -0.45 | 1. LC-MS/MS (HCD) |
| 269.0809 | $[M+H]^+$<br>$[M+K]^+$ | Formononetin* | 269.0808 | 0.24 | 1. On-tissue (CID, ion trap and orbitrap)<br>2. On-tissue (HCD) |
| 617.1840 | $[M+H]^+$<br>$[M+K]^+$ | Heme | 617.1846 | -0.99 | 1. On-tissue (CID, ion trap and orbitrap)<br>2. LC-MS/MS (HCD) |

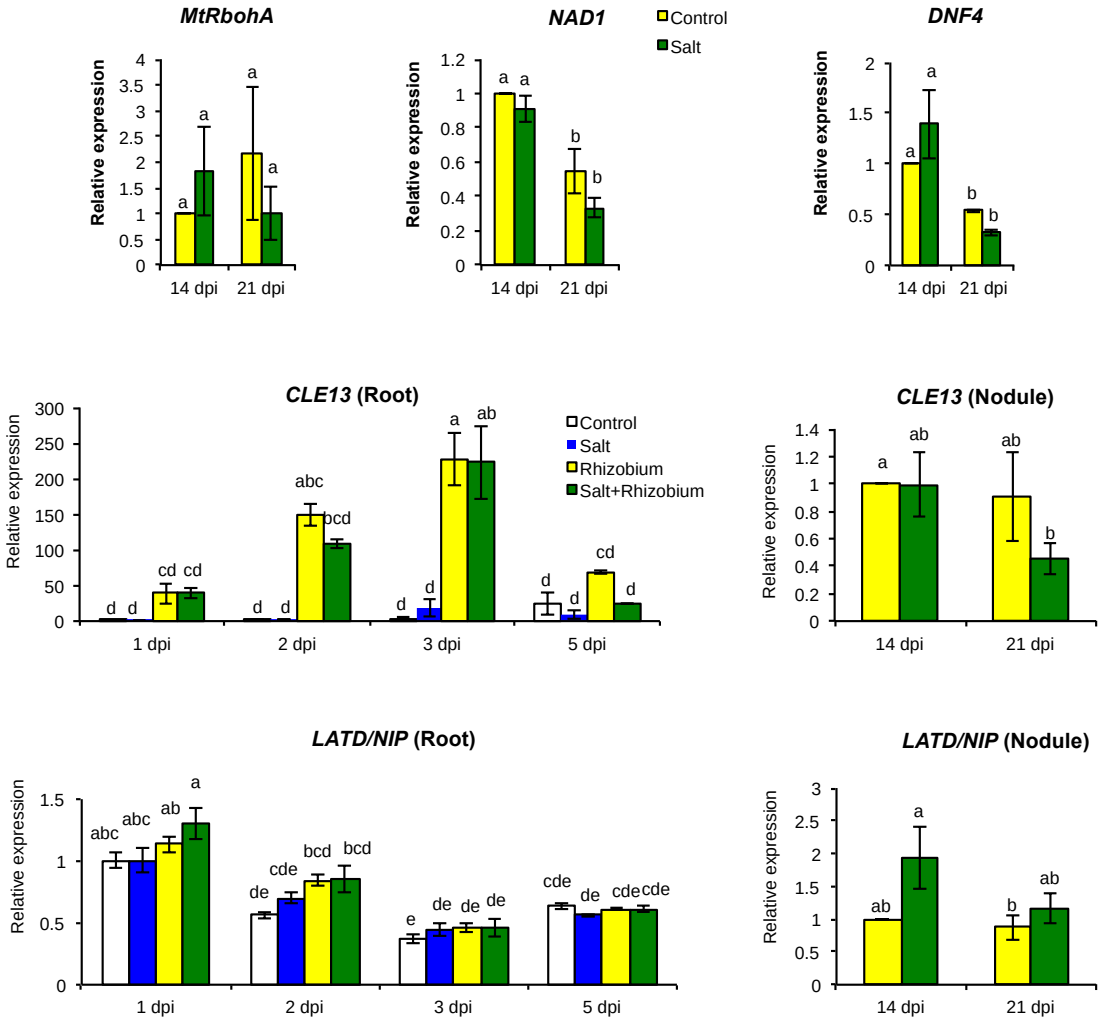



**Figure S1.** Host genes whose expression is not altered in the presence of salinity. Transcript levels measured by quantitative RT-PCR. Bars represent mean of three independent experiments +/- SEM. Statistical analysis done using a 2-way ANOVA followed by a t-test (nodules) or Tukey's HSD test (roots) on the LS-Means ($\alpha$=0.05). Columns not connected by the same letter are significantly different.

# Appendix V

# Optimization of the Extraction Procedure for Identification of Endogenous Peptides in *Medicago truncatula* seedling plants by LC-MS/MS

This project was in collaboration with Junko Maeda and Dhileep Jayaraman from the Ané lab (UW-Madison Department of Bacteriology) and Micheal R. Sussman (UW-Department of Biochemistry). Alyssa Eisold was an undergrad researcher assisting with this project.

Keyword: LC-MS/MS; *Medicago truncatula*; extractions; endogenous peptides

**Abstract**

Plant peptide hormones have numerous roles in signaling and in the growth and development of plants. Despite a large number of predicted plant peptide hormones, only a small number of peptides have been fully characterized. Due to its ability to detect small concentrations of endogenous peptides and their many post translational modifications, mass spectrometry is a valuable tool to study peptide hormones. Here, a variety of plant extraction techniques were evaluated for optimal detection of plant peptides. Aqueous extraction solvents provided the highest number of detected peptides. Although the extractions did not detect many endogenous peptides, many peptide sequences were detected. Most of these detected peptides are proteolytic cleavage fragments of larger proteins. Further analysis is necessary to determine whether these detected protein fragments have functional roles in the plant.

**Introduction**

Peptide hormones play important roles in both short and long range signaling in plants, covering a wide range of functions, from plant growth and development, to formation of symbiosis interactions and stress response.[1-3] In *Medicago truncatula* (Medicago), nodules form on the roots of the plant as a result of the symbiotic relationship between the plant and rhizobia bacteria. Once the symbiotic relationship is formed in the root nodules, biological nitrogen fixation, where the rhizobia bacteria reduce atmospheric $N_2$ into ammonia, occurs. Various peptides are involved in the development and regulation of the symbiotic relationship between Medicago and rhizobium bacteria.[3] Two members of the CLAVATA3/embryo-surrounding region (CLE) peptide family (*Mt*CLE*12* and *Mt*CLE*13*) are involved in the regulation of nodules.[4, 5] A C-terminally encoded peptide (*Mt*CEP1) was shown to regulate nodulation and lateral root formation.[6, 7] Furthermore,

nodule specific cysteine rich (NCR) peptides are involved in the differentiation of bacteria into bacteriods in the root nodules.[8-10]

Currently, the *in vivo* discovery and characterization of peptide hormones in plants is challenging due to the low concentrations of these peptides and has only been performed by a few laboratories.[7, 11-18] Despite the low number of characterized peptide hormones, a study in Arabidopsis thaliana found numerous unannotated potential secreted peptides, indicating that only a small amount of the plant peptidome has been fully characterized.[19] While numerous endogenous signaling peptides have been predicted based upon genetic sequencing data,[19] the genomic data does not provide the active form of the peptide in the plant or potential post translational modifications (PTMs). Methods to isolate active peptide hormones include using plant cell cultures, peptide overexpressor lines, xylem sap, and submerged seedlings in order to achieve higher peptide concentration or lower amounts of background interfering peptides to isolate enough of the peptide to detect and characterize.[7, 11, 16, 18, 20] As cell cultures and overexpressor lines are altered version of the plant, they may not fully represent the in vivo, non-mutant version of the plant. Also, mutant lines of plants take considerable time and effort to create. The other methods focus on secreted peptides rather than all peptides present in the plant. Thus, a method to investigate endogenous peptide distribution in wild-type plants would be beneficial.

Mass spectrometry is a powerful tool to detect peptides hormones as it allows for sensitive detection of many peptide modifications. PTMs are often found on peptide hormones and can play crucial roles in the function of the peptide.[21, 22] CLE and CEP peptides are hydroxylated and arabinosylated, and NCR peptides have multiple disulfide bonds.[1, 23] Currently, detection of peptide hormones is challenging due to the low concentrations of the peptide *in vivo*, which is why studies usually use overexpressor lines, which create additional copies of one specific gene, in this

case an endogenous peptide coding gene, or focus on peptides secreted form the plant, which results in a simpler background.[23] Here the possibility of endogenous peptide extraction from wild-type plant cells, which provides the information from wild-type plant was investigated. Various extraction methods were attempted to find an optimal extraction method for a wide variety of peptide hormones directly from Medicago seedling plants without the need for overexpressor lines.

**Experimental Methods**

Tissue Extractions

Seedling plants (50-100) were frozen on dry ice and stored at -80$^{o}$C until the extraction was performed. Multiple extractions were attempted based upon extractions found in literature for endogenous peptides, including:

1. 40% Ethanol[24]
2. 40% Ethanol[24] with 1% Acetic Acid
3. 10% AcN with 1 M acetic acid and a protease inhibitor tablet[25]
4. $H_2O$ with 1% TFA[26]
5. dichloromethane:methanol (1:1,v/v)[27]
6. Intercellular fluid extraction[28]
7. 8 M Urea[29]
8. 8 M Urea[29] denator
9. 40% Ethanol[24] denator

Plant material was ground up with a mortar and pestle under liquid nitrogen. The general procedure for the extractions was to add the extraction liquid, vortex, probe sonicate (select extractions), centrifuge, collect the liquid, and dry down the supernatant in a speed vac unless otherwise noted.

The extraction liquid composition and amount and centrifugation parameters were as described in the original paper unless otherwise noted below.

1. 40% Ethanol:[24] After addition of 40% ethanol, the mixture was probe sonicated for 6 cycles (15 s on and 30 s off). After centrifugation, the supernatant was dried down in a speed vac (further clean-up steps from the original paper were not followed).
2. 40% Ethanol[24] with 1% Acetic Acid: 1% acetic acid was added to the 40% ethanol extraction liquid. After addition of 40% ethanol 1% acetic acid, the mixture was probe sonicated for 6 cycles (15 s on and 30 s off). After centrifugation, the supernatant was dried down in a speed vac (further clean-up steps from the original paper were not followed).
3. 10% AcN with 1 M acetic acid and a protease inhibitor tablet:[25] A mini EDTA-free protease inhibitor cocktail tablet was used for the protease inhibitor tablet. After addition of the 10% AcN extraction mixture, the mixture was probe sonicated for 6 cycles (15 s on and 30 s off). Centrifugation occurred at 4000 x g for 30 min at 4$^{o}$C. The supernatant was removed and dried in a rotovap. After the extraction was dried down, the pellet was rinsed with 1 mL ethanol and dried in the rotovap.
4. $H_2O$ with 1% TFA:[26] After addition of $H_2O$ with 1% TFA at a ratio of 4 mL/g plant material, the mixture was probe sonicated for 6 cycles (15 s on and 30 s off). Plant material was pelleted by centrifuging at 4500 x g for 2 hr at 4$^{o}$C. An aliquot of the extraction underwent an acetone precipitation procedure.[30] To 750 μL of the extraction resuspended in water, 7.5 μL of N-ethylmorpholine (NEM) and 750 μL of o-chlorophenol with 1% NEM were added. The mixture was vortexed and shook for 1 min at room temperature. The extract was then centrifuged at 10,000 x g for 10 min at 4$^{o}$C. The aqueous phase was removed, dried down, and stored in the -80$^{o}$C freezer. The phenolic phase was transferred to a falcon tube and 20 volumes

of acetone were added, mixed and placed at -20°C overnight. The acetone precipitation mixture was centrifuged at 4500 x g for 2 hour. The acetone was decanted off. The pellet was rinsed with acetone, switched to centrifuge tubes, and dried in a speed vac.

5. dichloromethane:methanol (1:1,v/v):[27] For the fifth extraction, three temperatures for the overnight incubation in 1:1 dichloromethane:methanol were attempted, room temperature, 4°C, and -20°C, and the methanol/water layer was collected the next day.
6. Intercellular fluid extraction:[28] For the intercellular fluid extraction, seedling plants were placed in the citric acid buffer described in the paper except the thiourea was replaced with urea. The submerged plants were placed in the vacuum desicator for 10 minutes, and the centrifuge set-up to release the ICF fluid was repeated twice. Both the ICF sample and 14 mL of infiltration buffer were collected.
7. 8 M Urea:[29] The extraction followed as described in the paper except the centrifuge time during the molecular weight cut-off (MWCO) step was increased until everything had passed through the filter, approximately 60 minutes. As the MWCO took longer than expected, half the extract was desalted (SepPak)without going through the MWCO. This half was kept separate from the half that went through the MWCO. After the MWCO, samples were desalted (SepPak). All desalted samples were dried down in a speed vac.
8. 8 M Urea[29] denator: Seedling plants (5 at a time) were heat stabilized using a denator for 15 s at 5.0 mm and 95°C. The extraction followed as described in the paper except the MWCO device was rinsed with 0.2 mL 0.1 M NaOH and 0.5 mL $H_2O$ prior to loading the sample and 0.1 mL $H_2O$ after loading the sample and centrifuge time during the molecular weight cut-off (MWCO) step were increased until everything had passed through the filter, approximately 60 minutes.

9. 40% Ethanol[24] denator: Seedling plants (5 at a time) were heat stabilized using a denator 15 s at 5.0 mm and 95ºC. After addition of 40% ethanol, the mixture was probe sonicated for 6 cycles (15 s on and 30 s off). After centrifugation, the supernatant was dried down in a speed vac (further clean-up steps from the original paper were not followed).

Sample Preparation

After the extractions, peptide concentrations were calculated with a peptide assay (Pierce Quantitative Colorimetric Peptide Assay, Thermo Scientific). Samples were reduced and alkylated in a 50 mM tris buffer, pH=8 at a concentration of 3.5 mg/mL. One aliquot of extraction 9 was also reduced and alkylated in an 8 M Urea 50 mM tris buffer, pH=8 at a concentration of 3.5 mg/mL, to test whether urea was necessary for reduction and alkylation. Reduction occurred with 5 mM dithiothreitol (DTT) for 1 hour at room temperature followed by alkylation with 15 mM iodoacetamide for 30 min in the dark at room temperature. The reaction was quenched with 5 mM DTT for 5 min at room temperature followed by acidification with addition of 0.3% TFA. Samples were desalted with SepPak $C_{18}$ solid phase extraction cartridge (Waters). SCX clean-up was performed with SCX SpinTips (Protea) according to kit protocol 1, except SpinTip rinse was performed with 100 µL of the wash solution, the sample was reconstituted in 200 µL of the reconstitution solution and loaded three times through the SpinTip. The sample was eluted with two 150 µL aliquots, and all centrifuges occurred for 1.5 min. Samples were desalted with three additions of 1 mL $H_2O$ with subsequent dry down in a speed vac. The peptide assay was used to calculate the peptide concentration prior to analysis. Certain samples, which were not reduced and alkylated in the literature, were run with and without reduction and alkylation to test whether reduction and alkylation was necessary. Sample that were not reduced are labeled NR and were desalted after extraction prior to analysis.

LC-MS/MS

Peptide samples were resuspended at 1 ug/uL. Samples were run on a nanoAcquity LC (Waters) coupled to a Q Exactive MS (Thermo Scientific). Separation occurred with a 15 cm long C18 nano LC column at a flow rate of 0.3 uL/min. Mobile phase A was optima $H_2O$ with 0.1% FA and mobile phase B was optima AcN with 0.1% FA. The LC parameters used 16 min trapping initially, followed by a gradient from 0-4% B from 0 to 2 min, 4-35% B from 0-70 min, 35-75% B from 70-71 min, 71-81 min hold at 75% B, 75-95% from 81-82 min, hold at 95% B from 82-92 min, 95-0% B from 92-93 min, and re-equilibration at 0% B for 15 min. Samples were kept at 4$^{o}$C during analysis. A top 15 DDA method was used for MS analysis. Full MS spectra were taken from 300-2000 *m/z* at 35,500 resolution, 1E6 AGC target, and 100 ms max injection time. The default charge state was selected as 2. MS/MS spectra were acquired with HCD activation at a collision energy of 30 at 17,500 resolution, 1E5 AGC target, 200 ms max inject time, and a 2.0 *m/z* isolation window. A dynamic exclusion of 30 s and a charge exclusion of unassigned, 1, and greater than or equal to 8 was used.

Peptide Identification

PEAKS *de novo* sequencing software was used to analyze the LC-MS/MS data. *De novo* sequencing occurred with a parent mass error of 20 ppm and a fragment mass error of 0.02 Da. Database searching was performed against a signaling peptide database and a full protein database from uniprot. Carbamidomethylation (C) was used as a fixed modification, and oxidation (M), hydroxylation (P), acetylation (N-term), amidation (C-term), and arabinosylation (P) were used as variable modifications. Three variable modifications were allowed per peptide and up to 5 peptides were reported. Matches to the signaling databases were manually reviewed for data quality while a 1% FDR cut-off was used for the protein database.

## Results

De Novo Sequencing

Initially, *de novo* sequencing was performed to determine peptide sequences without the input of a protein database. **Table 1** shows the number of *de novo* sequences with an average local confidence (ALC) score over 80. The ALC score provides overall strength of the *de novo* identification by averaging the likelihood of each amino acid at each position of the sequence. Although the PEAKS default ALC score is 50, 80 was chosen to show only the most confident sequences where each amino acid is likely correct. Overall, higher numbers of *de novo* sequenced peptides were achieved with aqueous buffers, especially with the 8 M Urea extraction. The denatured plants with the ethanol extraction also resulted in a high number of sequences. The denatured sample with 8 M Urea extraction had the highest number of *de novo* sequences. Further analysis would be needed to determine whether the higher number of peptides with the denatured samples in both extractions is due decreased degradation of peptides or due to increased levels of proteolytic cleavage of proteins during the process of denaturing the sample. Reduction and alkylation of the samples followed by SCX purification generally increased the number of detected peptides. The molecular weight cut-off step in the 8 M urea extraction potentially resulted in loss of peptides, as the no MWCO, SCX sample had a higher number of peptide sequences that the MWCO, SCX sample.

To determine whether the different extractions were detecting similar peptides, the detected peptide signaling were compared. The overlap of peptide sequences with *de novo* sequencing is shown in **Figure 1**. Venn Diagrams illustrate the shared peptide sequences between different extractions. Overall, most cases show a low number of shared sequences and a high number of unique sequences. Each extraction was performed on 50-100 seedling plants, meaning that a good

variety of biological variability should be averaged in each extraction. Also, technical replicates were not assessed; thus, some variability could be due to technical variability. In **Figure 1 (A)** the presence of acetic acid in the 40% ethanol extraction solvent to limit peptide degradation is compared with the 40% ethanol solvent extraction. The very low overlap between ethanol with and without acetic acid indicates that the addition of acetic acid changes the extraction properties with the lowing of the pH. In **Figure 1 (B)**, the three acid extractions (extraction #s 2-4) were compared. Only 40 sequences were in all three, leaving hundreds of sequences unique to one of the extractions. A similar pattern was observed for the three different temperatures of extraction 5.

Different sample preparations for the extractions were also compared for the intercellular fluid extraction **Figure 1 (D)** and Urea extraction **Figure 1 (E)**. For the intracellular fluid extraction, the buffer was compared to the intercellular fluid extraction (both non-reduced and SCX purified, reduced, and alkylated). There appears to be some leakage from the intercellular fluid into the buffer, but the most sequences came from the SCX purified, reduced, and alkylated sample. For the urea sample, the molecular weight cut-off device type was switched after the initial extraction preparation due to excessive time to run the sample through the device, the second MWCO device is labeled new. There is potential loss during the MWCO step as there SCX, reduced and alkylated sample preparation with no MWCO yielded a high number of unique sequences.

Finally, a comparison of the denator and non-denatored samples with the 8 M urea extraction is shown for **Figure 1 (F-G)** for the nonreduced and reduced, alkylated, and SCX purified samples respectively. The shared sequences are still a small portion of the total sequences. The denatored samples with the ethanol extraction were reduced and alkylated with and without

urea in the reduction buffer. The shared sequences was higher for this comparison, which is expected since aliquots of the same extraction were used. Finally, the use of acid to limit degradation, the use of the denator to limit degradation, and no addition to limit degradation were all compared for the ethanol extraction in **Figure 1 (I)**. There were only 13 shared sequences between the three extractions, while 1602 sequences were unique to the denatored sample with the ethanol extraction. As only one technical replicate was injected, more technical replicates would be necessary to tell whether the low number of shared sequences is due to biological variability, run to run variability, or actual difference in the extractions.

Identification of Signaling Peptides

An endogenous peptide database was created by searching for peptide classes found in a plant peptide review in the uniprot Medicago protein database and compiling all the hits. It should be noted that this method included some protein sequences that are not endogenous peptides through the search. For example, when searching for nodule specific cysteine rich peptides, cysteine proteinases were also included in the search hits. Despite this, the peptide database is much shorter than the full protein database and requires substantially less time to search with in PEAKS software. **Table 1** shows the number of PSMs to the peptide database for all the attempted extractions. Overall, the number of PSMs https://vrs-us.com/job/analytical-chemist-lc-ms/was very low and usually only matched up to a couple proteins. Many of the protein groups identified were miscellaneous proteins that are not expected to be endogenous peptides (i.e. cysteine proteinases) or lipid transfer proteins. **Table 2** shows the protein groups identified for extraction 1 with the peptide database. Most of these are larger proteins, but a couple potentially interesting protein groups are seen, such as the NCR, thionin, and defensin groups. However, NCR peptides are not expected to be in the seedling plants and may be due to contamination on the LC column.

As **Figure 2** shows for the thionin **(A)** and defensin **(B)** peptide matches to the ethanol extraction (extraction 1), the peptides matching to the thionin and defensin proteins are small pieces and may not be the full version of the peptides. After manually looking through the data, short peptides matching to either thionin, defensin, or occasionally RALF or KDEL-tailed cysteine endopeptidase CEP1 were found with varying MS/MS spectral qualities. However, no extractions provided multiple strong (i.e. good MS/MS spectral quality) full length endogenous peptides.

Identification of Protein Fragments

As there were few matches to endogenous peptides in the extractions, select extractions were analyzed in PEAKS with the full Medicago protein database. Due to long analysis times, only the first 5 extractions were analyzed with the full protein database. As the goal of the project was targeted towards endogenous peptides select extractions were analyzed against the protein database to test whether the observed peptides in the LC-MS/MS runs could be due to protein fragments. **Table 3** lists the PSMs for extractions 1-5 matching to the full protein database with an FDR of 1% and the number of protein groups for each extraction. There were significantly more PSMs for the protein database compared to the peptide database, indicating that there are a lot of protein fragments in the LC-MS/MS runs. As the number of PSMs in the peptide databases was typically very low, the peptide PSMs should not contribute significantly contribute to the protein PSMs. **Figure 3** compares the overlap in proteins detected between the different extractions analyzed with the full protein database. Overall, there was more overlap between the proteins detected than with the *de novo* sequenced peptides (**Figure 1**), indicating that the proteins detected might be more consistent than the individual peptide sequences detected. In **(A)**, the ethanol extractions with and without acetic acid were compared. Slightly less than half of the detected proteins in each of the two extractions were shared. In **Figure 3 (B)**, the three acid extractions

were compared. Again, there is much better overlap with the proteins detected than with the *de novo* peptides detected. The 1% TFA extraction detected the most protein groups, although in **Table 3**, the 10% AcN 1 M acetic acid extraction had the most PSMs matching to proteins. The three different temperatures for extraction 5 (methanol) had high overlap **Figure 3 (C)**, which is expected as the extraction was split three ways for the overnight incubation at the three temperatures. Further technical and biological replicates would be necessary to determine how much of the variation observed is due to biological variation, different efficiencies of the extraction, or due to technical variation in the peptides selected for MS/MS. One possibility would be to check the peptides detected in the MS1 level to asses if technical variation of *m/z* selected for MS/MS contributed to variation.

**Conclusions**

Multiple extraction solvents were investigated to find an optimal method for the detection of endogenous peptides. *De novo* sequencing results indicate that the aqueous extraction solvents provide high numbers of detected peptide sequences. However, not many of these matched to known peptide hormones. Instead, most peptide sequences were fragments of bigger proteins. Although there is increasing evidence for protein fragments playing roles within the plant, further studies would be required to determine if the protein fragments detected here have biological roles. Future studies should perform more biological and technical replicates of the extractions to determine whether the observed variation in the extractions performed here is due to technical variation, biological variation, or due to actual differences in the extraction efficiency.

**References**

1. Tavormina, P.; De Coninck, B.; Nikonorova, N.; De Smet, I.; Cammue, B. P., The Plant Peptidome: An Expanding Repertoire of Structural Features and Biological Functions. *Plant Cell* **2015,** *27* (8), 2095-118.

2. Marmiroli, N.; Maestri, E., Plant peptides in defense and signaling. *Peptides* **2014,** *56*, 30-44.

3. Batut, J.; Mergaert, P.; Masson-Boivin, C., Peptide signalling in the rhizobium-legume symbiosis. *Curr Opin Microbiol* **2011,** *14* (2), 181-7.

4. Mortier, V.; Den Herder, G.; Whitford, R.; Van de Velde, W.; Rombauts, S.; D'Haeseleer, K.; Holsters, M.; Goormachtig, S., CLE Peptides Control Medicago truncatula Nodulation Locally and Systemically. *Plant Physiology* **2010,** *153* (1), 222-237.

5. Mortier, V.; De Wever, E.; Vuylsteke, M.; Holsters, M.; Goormachtig, S., Nodule numbers are governed by interaction between CLE peptides and cytokinin signaling. *Plant Journal* **2012,** *70* (3), 367-376.

6. Imin, N.; Mohd-Radzman, N. A.; Ogilvie, H. A.; Djordjevic, M. A., The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in Medicago truncatula. *Journal of Experimental Botany* **2013,** *64* (17), 5395-5409.

7. Mohd-Radzman, N. A.; Binos, S.; Truong, T. T.; Imin, N.; Mariani, M.; Djordjevic, M. A., Novel MtCEP1 peptides produced in vivo differentially regulate root development in Medicago truncatula. *Journal of Experimental Botany* **2015,** *66* (17), 5289-5300.

8. Farkas, A.; Maroti, G.; Durgo, H.; Gyorgypal, Z.; Lima, R. M.; Medzihradszky, K. F.; Kereszt, A.; Mergaert, P.; Kondorosi, E., Medicago truncatula symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proc Natl Acad Sci U S A* **2014,** *111* (14), 5183-8.

9. Van de Velde, W.; Zehirov, G.; Szatmari, A.; Debreczeny, M.; Ishihara, H.; Kevei, Z.; Farkas, A.; Mikulass, K.; Nagy, A.; Tiricz, H.; Satiat-Jeunemaitre, B.; Alunni, B.; Bourge, M.; Kucho, K.; Abe, M.; Kereszt, A.; Maroti, G.; Uchiumi, T.; Kondorosi, E.; Mergaert, P., Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* **2010,** *327* (5969), 1122-6.

10. Guefrachi, I.; Nagymihaly, M.; Pislariu, C. I.; de Velde, W. V.; Ratet, P.; Mars, M.; Udvardi, M. K.; Kondorosi, E.; Mergaert, P.; Alunni, B., Extreme specificity of NCR gene expression in Medicago truncatula. *Bmc Genomics* **2014,** *15*, 16.

11. Tabata, R.; Sumida, K.; Yoshii, T.; Ohyama, K.; Shinohara, H.; Matsubayashi, Y., Perception of root-derived peptides by shoot LRR-RKs mediates systemic N-demand signaling. *Science* **2014,** *346* (6207), 343-6.

12. Fukuda, H.; Higashiyama, T., Diverse functions of plant peptides: entering a new phase. In *Plant Cell Physiol*, Japan, 2011; Vol. 52, pp 1-4.

13. Matsubayashi, Y.; Hanai, H.; Hara, O.; Sakagami, Y., Active fragments and analogs of the plant growth factor, phytosulfokine: structure-activity relationships. *Biochem Biophys Res Commun* **1996,** *225* (1), 209-14.

14. Roberts, I.; Smith, S.; Stes, E.; De Rybel, B.; Staes, A.; van de Cotte, B.; Njo, M. F.; Dedeyne, L.; Demol, H.; Lavenus, J.; Audenaert, D.; Gevaert, K.; Beeckman, T.; De Smet, I., CEP5 and XIP1/CEPR1 regulate lateral root initiation in Arabidopsis. *J Exp Bot* **2016,** *67* (16), 4889-99.

15. Whitford, R.; Fernandez, A.; Tejos, R.; Perez, A. C.; Kleine-Vehn, J.; Vanneste, S.; Drozdzecki, A.; Leitner, J.; Abas, L.; Aerts, M.; Hoogewijs, K.; Baster, P.; De Groodt, R.; Lin, Y. C.; Storme, V.; Van de Peer, Y.; Beeckman, T.; Madder, A.; Devreese, B.; Luschnig, C.; Friml, J.; Hilson, P., GOLVEN secretory peptides regulate auxin carrier turnover during plant gravitropic responses. *Dev Cell* **2012,** *22* (3), 678-85.

16. Ohyama, K.; Ogawa, M.; Matsubayashi, Y., Identification of a biologically active, small, secreted peptide in Arabidopsis by in silico gene screening, followed by LC-MS-based structure analysis. *Plant J* **2008,** *55* (1), 152-60.

17. Okamoto, S.; Shinohara, H.; Mori, T.; Matsubayashi, Y.; Kawaguchi, M., Root-derived CLE glycopeptides control nodulation by direct binding to HAR1 receptor kinase. *Nat Commun* **2013,** *4*, 2191.

18. Ito, Y.; Nakanomyo, I.; Motose, H.; Iwamoto, K.; Sawa, S.; Dohmae, N.; Fukuda, H., Dodeca-CLE peptides as suppressors of plant stem cell differentiation. *Science* **2006,** *313* (5788), 842-5.

19. Lease, K. A.; Walker, J. C., The Arabidopsis unannotated secreted peptide database, a resource for plant peptidomics. *Plant Physiol* **2006,** *142* (3), 831-8.

20. Okamoto, S.; Suzuki, T.; Kawaguchi, M.; Higashiyama, T.; Matsubayashi, Y., A comprehensive strategy for identifying long-distance mobile peptides in xylem sap. *Plant J* **2015,** *84* (3), 611-20.

21. Imin, N.; Patel, N.; Corcilius, L.; Payne, R. J.; Djordjevic, M. A., CLE peptide tri-arabinosylation and peptide domain sequence composition are essential for SUNN-dependent autoregulation of nodulation in Medicago truncatula. *New Phytol* **2018,** *218* (1), 73-80.

22. Haag, A. F.; Kerscher, B.; Dall'Angelo, S.; Sani, M.; Longhi, R.; Baloban, M.; Wilson, H. M.; Mergaert, P.; Zanda, M.; Ferguson, G. P., Role of cysteine residues and disulfide bonds in the activity of a legume root nodule-specific, cysteine-rich peptide. *J Biol Chem* **2012,** *287* (14), 10791-8.

23. Patel, N.; Mohd-Radzman, N. A.; Corcilius, L.; Crossett, B.; Connolly, A.; Cordwell, S. J.; Ivanovici, A.; Taylor, K.; Williams, J.; Binos, S.; Mariani, M.; Payne, R. J.; Djordjevic, M. A., Diverse Peptide Hormones Affecting Root Growth Identified in the Medicago truncatula Secreted Peptidome. *Mol Cell Proteomics* **2018,** *17* (1), 160-174.

24. Serra, A.; Hemu, X.; Nguyen, G. K.; Nguyen, N. T.; Sze, S. K.; Tam, J. P., A high-throughput peptidomic strategy to decipher the molecular diversity of cyclic cysteine-rich peptides. *Sci Rep* **2016,** *6*, 23005.

25. Fesenko, I. A.; Arapidi, G. P.; Skripnikov, A. Y.; Alexeev, D. G.; Kostryukova, E. S.; Manolov, A. I.; Altukhov, I. A.; Khazigaleeva, R. A.; Seredina, A. V.; Kovalchuk, S. I.; Ziganshin, R. H.; Zgoda, V. G.; Novikova, S. E.; Semashko, T. A.; Slizhikova, D. K.; Ptushenko, V. V.; Gorbachev, A. Y.; Govorun, V. M.; Ivanov, V. T., Specific pools of endogenous peptides are present in gametophore, protonema, and protoplast cells of the moss Physcomitrella patens. *BMC Plant Biol* **2015,** *15*, 87.

26. Zhang, K.; McKinlay, C.; Hocart, C. H.; Djordjevic, M. A., The Medicago truncatula small protein proteome and peptidome. *J Proteome Res* **2006,** *5* (12), 3355-67.

27. Hellinger, R.; Koehbach, J.; Soltis, D. E.; Carpenter, E. J.; Wong, G. K.; Gruber, C. W., Peptidomics of Circular Cysteine-Rich Plant Peptides: Analysis of the Diversity of Cyclotides from Viola tricolor by Transcriptome and Proteome Mining. *J Proteome Res* **2015,** *14* (11), 4851-62.

28. Weinhold, A.; Wielsch, N.; Svatos, A.; Baldwin, T., Label-free nanoUPLC-MSE based quantification of antimicrobial peptides from the leaf apoplast of Nicotiana attenuata. *Bmc Plant Biology* **2015,** *15*, 14.

29. Secher, A.; Kelstrup, C. D.; Conde-Frieboes, K. W.; Pyke, C.; Raun, K.; Wulff, B. S.; Olsen, J. V., Analytic framework for peptidomics applied to large-scale neuropeptide identification. *Nat Commun* **2016,** *7*, 11436.

30. Ohyama, K.; Ogawa, M.; Matsubayashi, Y., Identification of a biologically active, small, secreted peptide in Arabidopsis by in silico gene screening, followed by LC-MS-based structure analysis. *Plant Journal* **2008,** *55* (1), 152-160.

**Tables**

**Table 1.** Overview of *de novo* Sequencing and Peptide Database Search Results for all Attempted Extractions. NR=not reduced SCX=strong cation exchange purification

| **Extraction** | ***De novo* (80 ALC)** | **Peptide Database PSMs (1% FDR)** |
|---|---|---|
| 1: EtOH SCX | 588 | 43 |
| 2: EtOH Acetic Acid SCX | 633 | 10 |
| 3: 10% Acetonitrile SCX | 823 | 43 |
| 4: 1% TFA SCX | 1003 | 2 |
| 5: MeOH room temperature SCX | 558 | 16 |
| 5: MeOH 4C SCX | 716 | 30 |
| 5: MeOH -20C SCX | 625 | 23 |
| 6: ICF NR | 604 | 1 |
| 6: Infiltration Buffer NR | 387 | 6 |
| 6: ICF SCX | 2053 | 5 |
| 7: Urea NR | 1523 | 1 |
| 7: Urea no MWCU SCX | 2434 | 70 |
| 7: Urea new MWCO NR | 630 | 1 |
| 7: Urea new MWCO SCX | 1802 | 19 |
| 8: Den Urea NR | 1507 | 1 |
| 8: Den Urea SCX | 3227 | 35 |
| 9: Den EtOH SCX | 1812 | 7 |
| 9: Den EtOH Urea SCX | 1753 | 6 |

**Table 2.** Identified Protein Groups for Extraction 1 (40% Ethanol) with the Peptide Database.

| Accession | Score (-10lgP) | Coverage (%) | Unique Peptides | PTM | Description |
|---|---|---|---|---|---|
| A0A072UTH7\|A0A072UTH7_MEDTR | 111.79 | 53 | 5 | Carbamidomethylation | Non-specific lipid-transfer protein (Fragment) OS=Medicago truncatula GN=MTR_4g428370 PE=3 SV=1 |
| G7JI87\|G7JI87_MEDTR | 87.98 | 37 | 2 | Carbamidomethylation | Non-specific lipid-transfer protein OS=Medicago truncatula GN=MTR_4g027800 PE=3 SV=1 |
| B7FFE5\|B7FFE5_MEDTR | 82.93 | 9 | 2 | | Lipid transfer protein OS=Medicago truncatula GN=MTR_3g079210 PE=2 SV=1 |
| G7IL76\|G7IL76_MEDTR | 80.64 | 29 | 1 | Carbamidomethylation | Thionin-like protein OS=Medicago truncatula GN=MTR_2g007950 PE=2 SV=1 |
| A0A072TXA9\|A0A072TXA9_MEDTR | 69.14 | 28 | 1 | Carbamidomethylation; Oxidation (M); Amidation; Hydroxylation | Nodule Cysteine-Rich (NCR) secreted peptide OS=Medicago truncatula GN=MTR_7g029540 PE=4 SV=1 |
| I3T793\|I3T793_MEDTR | 68.49 | 12 | 1 | Carbamidomethylation | Leginsulin related MtN11/16/17 family OS=Medicago truncatula GN=MTR_3g436100 PE=2 SV=1 |
| G7JBG6\|G7JBG6_MEDTR | 67.65 | 17 | 1 | Carbamidomethylation | Lipid transfer protein OS=Medicago truncatula GN=MTR_3g085210 PE=3 SV=1 |
| G7JJK0\|G7JJK0_MEDTR | 67.39 | 9 | 2 | Carbamidomethylation | Non-specific lipid-transfer protein OS=Medicago truncatula GN=MTR_4g029390 PE=2 SV=1 |
| G7LA76\|G7LA76_MEDTR | 66.79 | 5 | 1 | Carbamidomethylation | Chitinase (Class Ib) / Hevein OS=Medicago truncatula GN=MTR_8g074330 PE=2 SV=2 |
| G7IPC6\|G7IPC6_MEDTR | 62.14 | 11 | 1 | Carbamidomethylation; Amidation | Lipid transfer protein OS=Medicago truncatula GN=MTR_2g026820 PE=4 SV=1 |
| G7KX64\|G7KX64_MEDTR | 59.72 | 13 | 1 | Carbamidomethylation | Non-specific lipid-transfer protein OS=Medicago truncatula GN=MTR_7g073060 PE=3 SV=2 |
| G7KX37\|G7KX37_MEDTR | 57.17 | 15 | 1 | Carbamidomethylation | Non-specific lipid-transfer protein OS=Medicago truncatula GN=MTR_7g072730 PE=3 SV=2 |
| G7K519\|G7K519_MEDTR | 52.08 | 7 | 1 | Carbamidomethylation | Lipid transfer protein OS=Medicago truncatula GN=MTR_5g006940 PE=3 SV=1 |
| G7IPD4\|G7IPD4_MEDTR | 45.79 | 32 | 1 | Carbamidomethylation; Amidation | Lipid transfer protein OS=Medicago truncatula GN=MTR_2g026920 PE=4 SV=1 |
| A0A072V9M7\|A0A072V9M7_MEDTR | 42.49 | 15 | 1 | Carbamidomethylation | Defensin MtDef1.1/MtDef1.2 OS=Medicago truncatula GN=MTR_2g079440 PE=4 SV=1 |
| G7IX57\|G7IX57_MEDTR | 29.24 | 9 | 1 | | Nodule Cysteine-Rich (NCR) secreted peptide OS=Medicago truncatula GN=MTR_3g033925 PE=4 SV=1 |

**Table 3.** Results for Selected Extractions Searched Against the Protein Database

| **Extraction** | **Protein Database PSMs (1% FDR)** | **Protein Groups** |
|---|---|---|
| 1: EtOH | 1990 | 314 |
| 2: EtOH AA | 1590 | 312 |
| 3: 10% AcN | 3721 | 431 |
| 4: 1% TFA | 2911 | 569 |
| 5: MeOH RT | 1341 | 259 |
| 5: MeOH 4C | 1292 | 252 |
| 5: MeOH -20C | 1344 | 253 |

**Figures**

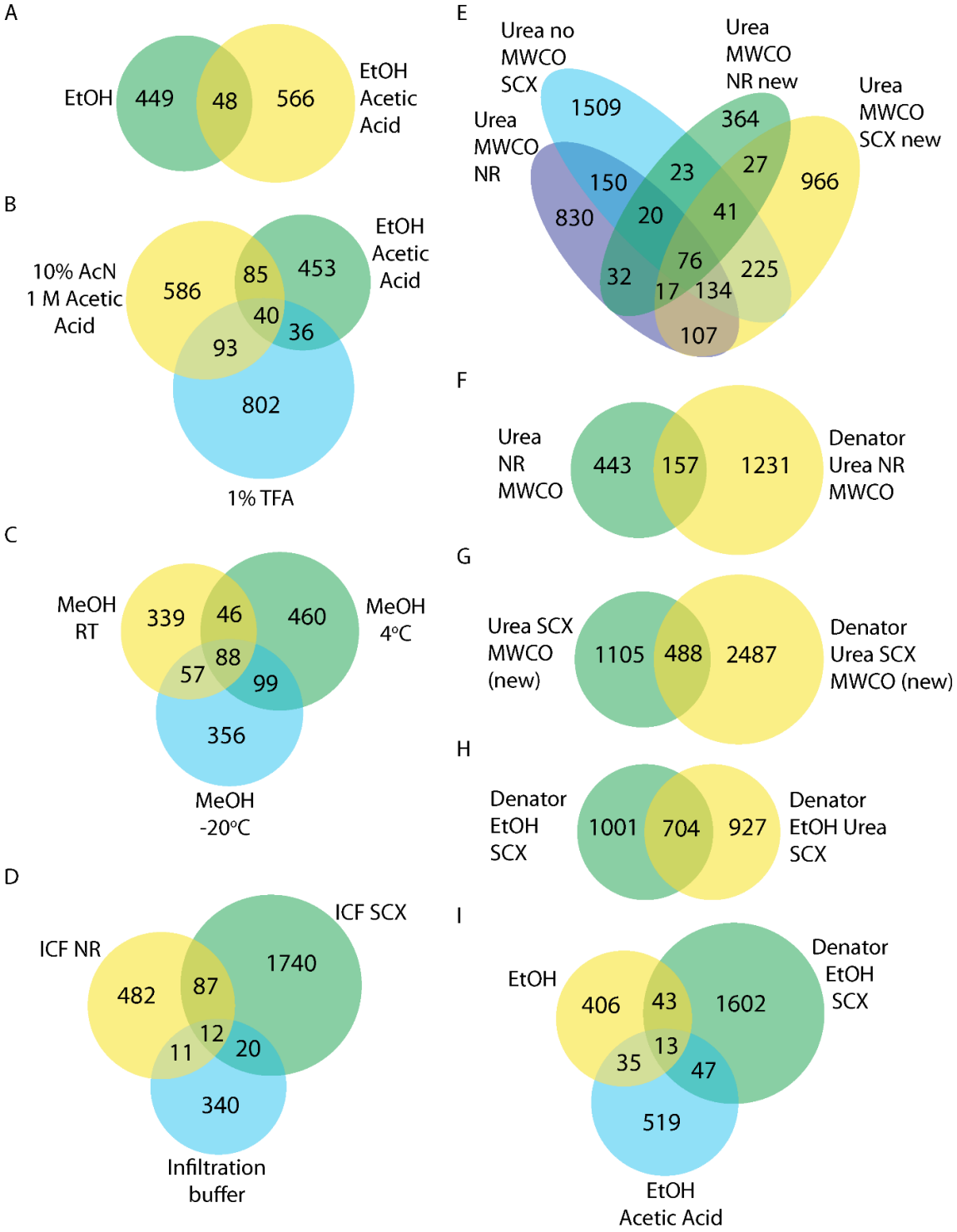
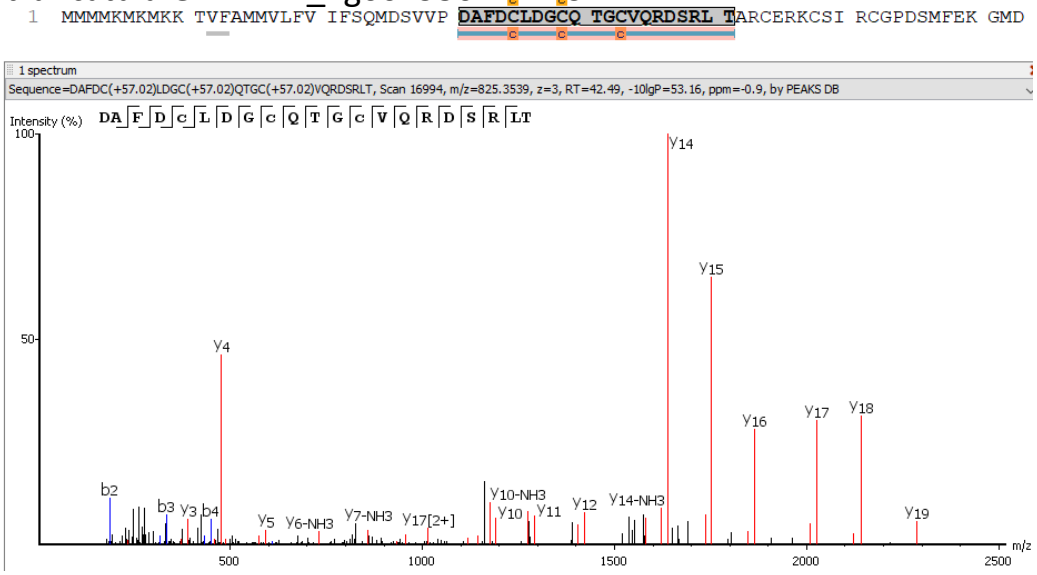



**Figure 1.** Comparison of *de novo* sequenced peptides detected with different extractions and different sample preparation.

>tr|G7IL76|G7IL76_MEDTR Thionin-like protein OS=Medicago truncatula GN=MTR_2g007950 PE=2 SV=1




>tr|A0A072V9M7|A0A072V9M7_MEDTR Defensin MtDef1.1/MtDef1.2 OS=Medicago truncatula GN=MTR_2g079440 PE=4 SV=1

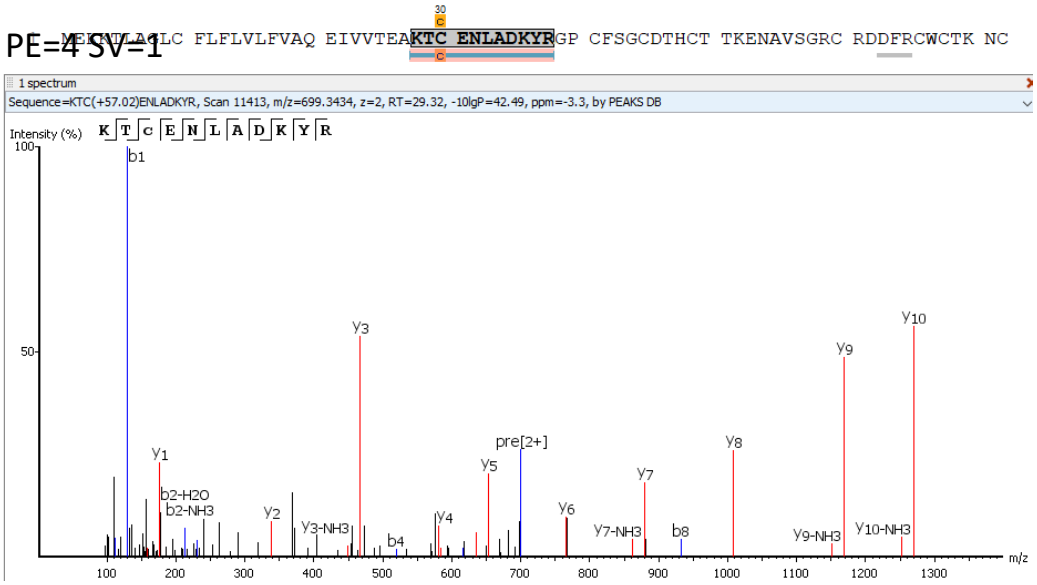



**Figure 2.** Example PEAKS MS/MS sequencing information for the thionin **(A)** and defensin **(B)** peptide matches to the ethanol extraction (extraction 1).

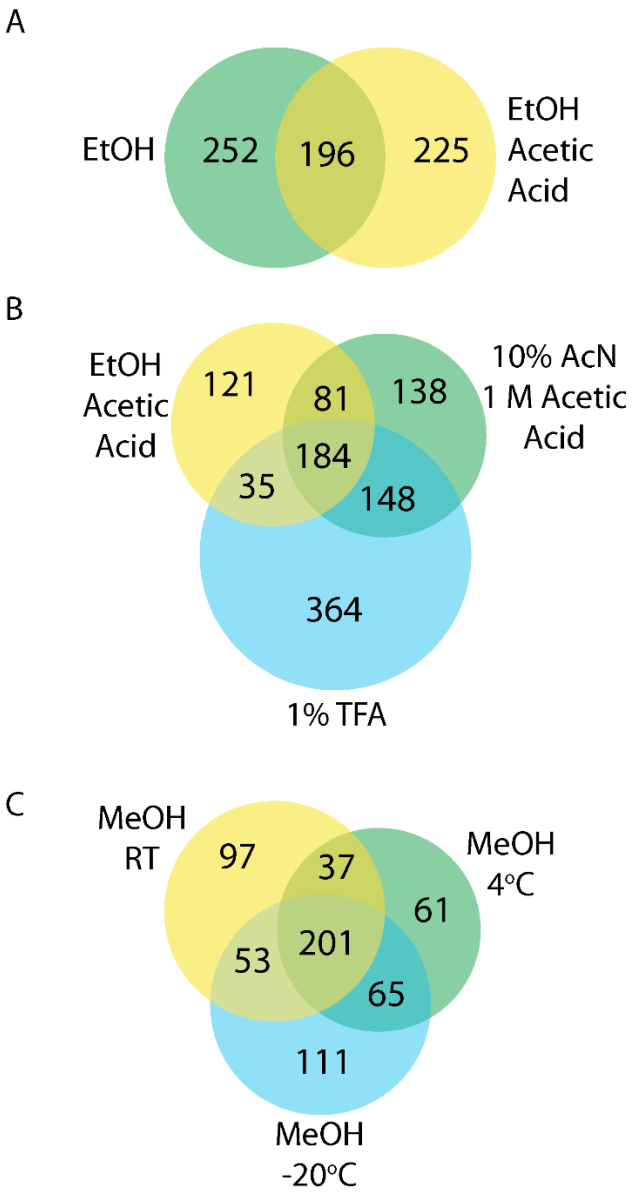



**Figure 3.** Comparison of the overlap in proteins detected with the full protein database between the different extractions.

# Appendix VI

# Imaging the Mucosal Layer in the Mouse Gastrointestinal Tract by MALDI-MSI

This project was in collaboration with Erin Gemperline (Li lab) and Jenny Bratburd from the Currie lab (UW-Madison Department of bacteriology). The Frederico Rey lab assisted with mice experiments.

Keyword: MALDI; MSI; microbiome; small molecules; cecum

**Abstract**

The importance of the microbiome in human health has been well documented. Small molecules produced by the microbiome not only are involved in nutrition, but also can potentially inhibit pathogen colonization. As matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) can localize molecules to specific areas of a tissue section, a MALDI-MSI method to determine the small molecules localized to the mucus membrane, where the microbiome predominantly live, would aid in determination of microbiome produced metabolites. Here various MALDI-MSI sample preparation steps were investigated to optimize a method to study microbiome metabolites. Different matrices, embedding options, and MALDI-MSI instruments were all explored. While detection of small molecules at high spatial resolution was possible with two time-of-flight (ToF)/ToF MALDI mass spectrometers, in the future, staining procedures are necessary to confirm that small molecule detected are indeed localized to the microbiome in the mucus layer.

**Introduction**

The large and dynamic community of microbes in the gastrointestinal tract, known as the gut microbiome, encompasses a variety of microorganisms, including bacteria, archaea, and eukaryotes.[1, 2] The microbiome-host relationship is critical for maintaining health. For example, the microbiome in humans is involved in nutrition and immune function modulation.[3] Furthermore, disruptions to the microbiome are associated with various diseases, including inflammatory bowel disease and cardiovascular disease.[4, 5] The microbiome produces small molecules, including both primary and secondary metabolites, that influence host health, not only in nutrition, but also in other various roles, such as protection against pathogens.[6, 7] Thus,

investigation of the small molecules produced by the microbiome would enable further information about host-microbiome metabolism, and additionally could aid in discovery of natural products produced by the microbiome for defense against pathogenic threats.

In the gastrointestinal tract, the microbiome is typically localized to the mucosal layer. In the colon and stomach, there are two mucosal layers, an outer, loose layer and an inner layer that is firmly attached to the epithelium.[8] The small intestine, however, has a less defined mucus membrane. In the colon, the inner layer is estimated to be 50 μm thick in mouse, while the outer layer is more variable, but is generally twice as thick.[9] Bacteria are found in the outer mucus layer, and were absent from the inner mucus layer.[10] The mucus membrane is challenging to study due to the fact that if collapses when it desiccates and when water based fixatives that crosslink proteins are used. Currently, the mucus membrane is studied by staining for the Muc2 protein and by 16S RNA fluorescence in situ hybridization (FISH), which stains for bacteria. To accomplish this, studies preserve the intestine samples, with fecal content inside to increase sample stability, utilizing Carnoy's fixative as this does not contain any water and does not crosslink proteins.[9]

Matrix-assisted laser-desorption/ionization mass spectrometry imaging (MALDI-MSI) has been used to study a variety of biological molecules, including small molecules, in numerous different types of tissues.[11] For example, small molecules produced by bacteria on the exoskeletons of ants in response to pathogen exposure has been investigated with MALDI-MSI.[12] Here, the use of MALDI-MSI to study small molecules produced by the microbiome in the mucosal layer was investigated. Additionally, an approximately 90 strain model of the human microbiome[13] was introduced to germ-free mice as a model system (termed humanized mice). Humanized, *Salmonella*-infected mice (HumSal) were compared to humanized, uninfected mice (HumNo) and germ-free, *Salmonella*-infected mice (GFSal) to determine if metabolites produced by the

microbiome in response to infection could be localized to the mucus membrane. This research focused on optimizing MALDI-MSI sample preparation and MALDI-MSI instrumentation for detection of the mucus membrane in the intestines.

**Experimental Procedures**

MALDI-MSI Sample Preparation

Mouse experiments were prepared as previously described.[13] Intestine samples were either embedded in gelatin (100 mg/mL) prior to sectioning (for MALDI LTQ Orbitrap XL and UltrafleXtreme experiments) or sectioned without embedding material (for RapifleX experiments). Cecum samples were sectioned at 12 µm (for MALDI LTQ Orbitrap XL and UltrafleXtreme experiments) or 20 µm (for RapifleX experiments) across the circumference of the intestines with a cryostat at -20$^{o}$C, creating circular sections. Sections were thaw-mounted onto glass slides for Orbitrap analysis or Indium Tin Oxide (ITO) glass slides for ToF/ToF analysis. Matrix was applied with the TM Sprayer. Both 2,5-dihydroxybenzoic acid (DHB) matrix (40 mg/mL in 50% MeOH, 0.1% formic acid) and α-cyano-4-hydroxycinnamic acid (CHCA) matrix (5 mg/mL in 50% acetonitrile, 0.1% formic acid) were tested. Samples were saved at -20$^{o}$C until analysis.

MALDI-MSI

MALDI-MSI was performed on three different instruments to test different instrument platforms and to achieve the proper spatial resolution. The three instruments were (1) MALDI LTQ Orbitrap XL (2) UltrafleXtreme ToF/ToF and (3) RapifleX ToF/ToF mass spectrometers. MALDI-Orbitrap MSI was acquired from 100-1700 *m/z* with 2 µscans per pixel at 75 µm spatial resolution. Centroid data was imported into SCiLs software for analysis (linear interpolation,

0.0005 Da mass accuracy). UltrafleXtreme ToF/ToF data acquisition was collected at 20 µm spatial resolution with 500 laser shots at each pixel from *m/z* 100-1700. RapifleX ToF/ToF data acquisition was collected at 15 µm spatial resolution from *m/z* 100-1200 averaging 500 laser shots per pixel at 10,000 frequency. RapifleX and UltrafleX mass spectra were calibrated with a standard mixture. For ToF/ToF analysis, raw profile data was imported into SCiLS software (auto detect axis settings) and no baseline removal for analysis. Segmentation was performed with weak denoising, and normalization to the total ion count. MALDI LTQ Orbitrap XL interval width was set at 5 ppm and ToF/ToF interval width was set to 0.020%.

**Results**

MALDI-MSI sample preparation procedures for metabolite imaging in mouse intestine samples were explored to optimize a procedure for imaging of small molecules produced by the microbiome in the gut. Initially, CHCA and DHB matrices were tested on the MALDI LTQ Orbitrap XL mass spectrometer due to the high-resolution accurate mass (HRAM) abilities of the instrument, which aids in confident metabolite identification. In order to visualize any layers that may appear in the sample, segmentation was performed on the data (**Figure 1**). For the segmentation analysis, the initial two groups provided by the software were split into new groups (segmentation diagram provided for each analysis) until further segmentation no longer split apart unique areas. The CHCA sample was unable to be imaged in its entirety due to time constraints. With CHCA as the matrix, the spectra from the sample were not consistently separated from the matrix (**Figure 1C**). To work around this problem, segmentation was performed on only the humanized, no infection (HumNo) sample (**Figure 2B**). For segmentation on just the HumNo sample, the data does segment out into circular layers; however, further analysis would be

necessary to ensure that this is reproducible and not an artifact. The DHB segmentation results (Figure 1E), demonstrate better separation between the sample and matrix, with the exception being the center of the HumNo sample. Certain samples show some potential segmentation out into layers, but further verification would be necessary as HumSal sample in particular did not segment out in a similar manner to the other samples. Although the segmentation results showed some promising patterns in the data, further attempts on the MALDI LTQ Orbitrap XL were not attempted due to the relatively large step size of the instrument (75 µm).

As the mucus layer in the cecum is estimated to be around 50 µm in mice, analysis of intestine samples was performed on two ToF/ToF instruments capable of lower spatial resolution. The UltrafleXtreme mass spectrometer minimal laser setting allows for 15-20 µm imaging, while the RapifleX mass spectrometer has a 5 µm laser capable of 10-20 µm imaging. For the RapifleX, 15 µm was selected for MALDI-MSI experiments as this would provide approximately 3 pixels across the inner mucus layer. **Figure 2** shows the UltrafleXtreme results for segmentation. The CHCA data did not segment well; it appears as if the data separated based upon time frame of when the data was collected as the data segments into vertical areas (**Figure 2B**). The entire sample again was not imaged. The DHB sample shows some segmentation of the data into layers **(Figure 2D)**. The data did separate into intestine and intestine content, but further segmentation of the data was not useful. For testing MALDI-imaging of colon samples on the Rapiflex, samples that were not embedded in gelatin were attempted due to concerns about the stability of the mucus layer in water. The HumNo sample tested was attached to the chuck by applying optimal cutting temperature media to excess intestine tissue after the fecal pellet ended (the fecal pellet was the target area to be imaged). Although the sample did section without any embedding material, it was nearly impossible to thaw mount the section onto a glass slide. Consequently, sections were not

flat to the glass slide for MALDI-MSI, which resulted in very poor results (**Figure 3**). As the presence of water was a concern, a 50% methanol matrix application method and a 100% methanol matrix application method were both attempted. Both methods were for DHB matrix (40 mg/mL) but pass number and flow rate were adjusted on the TM sprayer method due to the much drier 100% methanol solvent. The poor sample adherence to the slide prevented comparison of these two matrix application methods.

**Conclusions**

Although optimal MALD-MSI procedures for cecum imaging were not found, the spatial resolution provided by the UltrafleXtreme and RapifleX ToF/ToF instruments is necessary to be able to differentiate the inner mucus membrane on the sample. While the RapifleX data was affected by sample preparation difficulties, the RapifleX would be the best instrument to pursue this project in the future due to the high spatial resolution achievable and the increased speed of acquisition compared to the other two instruments. For sample preparation, embedding is necessary for proper adherence of the sample to the glass slide. If one embeds the sample in gelatin with excess intestinal tissue on either side of the desired area to be imaged, there will be minimal contact between the water in the gelatin embedding substance and the mucus layer inside the tissue. Other aspects of the sample procedure, such as the matrix application parameters and solvent, may still need to be adjusted for intestinal imaging. However, despite instrumental difficulties with the performed CHCA imaging, DHB matrix does appear to provide better results as it is able to more reliably separate the sample from the matrix areas.

As there are various concerns about the stability of the mucus membrane, which would greatly affect the ability to investigate small molecules produced by the microbiome in the mucus

membrane, having a histological method to stain for the mucus membrane would be beneficial. Concerns about mucosal layer stability when in contact with water or upon desiccation, which also collapses the mucus membrane, would then be able to be investigated. Studies have also utilized Carnoy fixation on the tissue, which is water free and does not crosslink proteins, to enable imaging of the mucosal layer. It is also recommended to use an area of the colon with fecal pellet still intact as this protects the mucus membrane from washing.[9] One method to perform staining would be to stain with antibodies for the muc2 protein. Alternatively, staining could be performed for bacteria with fluorescence in situ hybridization (FISH). This has recently been performed on deep sea mussel to investigate small molecules involved in host-microbe symbiosis by combining FISH with atmospheric pressure MALDI-MSI.[14] The combination of FISH with MALDI-MSI could be very beneficial for this study to correlate the areas of bacteria detection via FISH to the small molecules detected in MALDI-MSI.

**Acknowledgements**

C.K. would like to thank Erin Gemperline who initiated this project and aided in MALDI-MSI experiments. Also, Jenny Bratburd performed all mouse experiments and provided samples with aid form the Frederico Rey lab. The RapifleX MALDI TOF/TOF mass spectrometer was purchased through the support of an NIH shared instrument grant S10OD025084. The MALDI LTQ Orbitrap XL was purchased through an NIH shared instrument grant (NCRR S10RR029531 to LL).

**References**

1. Scanlan, P. D.; Marchesi, J. R., Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *Isme j* **2008,** *2* (12), 1183-93.

2. Hoffmann, C.; Dollive, S.; Grunberg, S.; Chen, J.; Li, H.; Wu, G. D.; Lewis, J. D.; Bushman, F. D., Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One* **2013,** *8* (6), e66019.

3. Kau, A. L.; Ahern, P. P.; Griffin, N. W.; Goodman, A. L.; Gordon, J. I., Human nutrition, the gut microbiome and the immune system. *Nature* **2011,** *474* (7351), 327-336.

4. Jie, Z.; Xia, H.; Zhong, S. L.; Feng, Q.; Li, S.; Liang, S.; Zhong, H.; Liu, Z.; Gao, Y.; Zhao, H.; Zhang, D.; Su, Z.; Fang, Z.; Lan, Z.; Li, J.; Xiao, L.; Li, R.; Li, X.; Li, F.; Ren, H.; Huang, Y.; Peng, Y.; Li, G.; Wen, B.; Dong, B.; Chen, J. Y.; Geng, Q. S.; Zhang, Z. W.; Yang, H.; Wang, J.; Zhang, X.; Madsen, L.; Brix, S.; Ning, G.; Xu, X.; Liu, X.; Hou, Y.; Jia, H.; He, K.; Kristiansen, K., The gut microbiome in atherosclerotic cardiovascular disease. *Nat Commun* **2017,** *8* (1), 845.

5. Kostic, A. D.; Xavier, R. J.; Gevers, D., The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* **2014,** *146* (6), 1489-1499.

6. Clarke, T. B.; Davis, K. M.; Lysenko, E. S.; Zhou, A. Y.; Yu, Y.; Weiser, J. N., Recognition of peptidoglycan from the microbiota by Nod1 enhances systemic innate immunity. *Nat Med* **2010,** *16* (2), 228-31.

7. Sharon, G.; Garg, N.; Debelius, J.; Knight, R.; Dorrestein, P. C.; Mazmanian, S. K., Specialized metabolites from the microbiome in health and disease. *Cell Metab* **2014,** *20* (5), 719-730.

8. Atuma, C.; Strugala, V.; Allen, A.; Holm, L., The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo. *Am J Physiol Gastrointest Liver Physiol* **2001,** *280* (5), G922-9.

9. Johansson, M. E.; Larsson, J. M.; Hansson, G. C., The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc Natl Acad Sci U S A* **2011,** *108 Suppl 1*, 4659-65.

10. Johansson, M. E.; Phillipson, M.; Petersson, J.; Velcich, A.; Holm, L.; Hansson, G. C., The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci U S A* **2008,** *105* (39), 15064-9.

11. Buchberger, A. R.; DeLaney, K.; Johnson, J.; Li, L., Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal Chem* **2018,** *90* (1), 240-265.

12. Gemperline, E.; Horn, H. A.; DeLaney, K.; Currie, C. R.; Li, L., Imaging with Mass Spectrometry of Bacteria on the Exoskeleton of Fungus-Growing Ants. *ACS Chem Biol* **2017,** *12* (8), 1980-1985.

13. Bratburd, J. R.; Keller, C.; Vivas, E.; Gemperline, E.; Li, L.; Rey, F. E.; Currie, C. R., Gut Microbial and Metabolic Responses to Salmonella enterica Serovar Typhimurium and Candida albicans. *MBio* **2018,** *9* (6).

14. Geier, B. K.; Sogin, E.; Michellod, D.; Janda, M.; Kompauer, M.; Spengler, B.; Dubilier, N.; Liebeke, M., Spatial metabolomics of in situ, host-microbe interactions. *bioRxiv* **2019**, 555045.

**Figures**

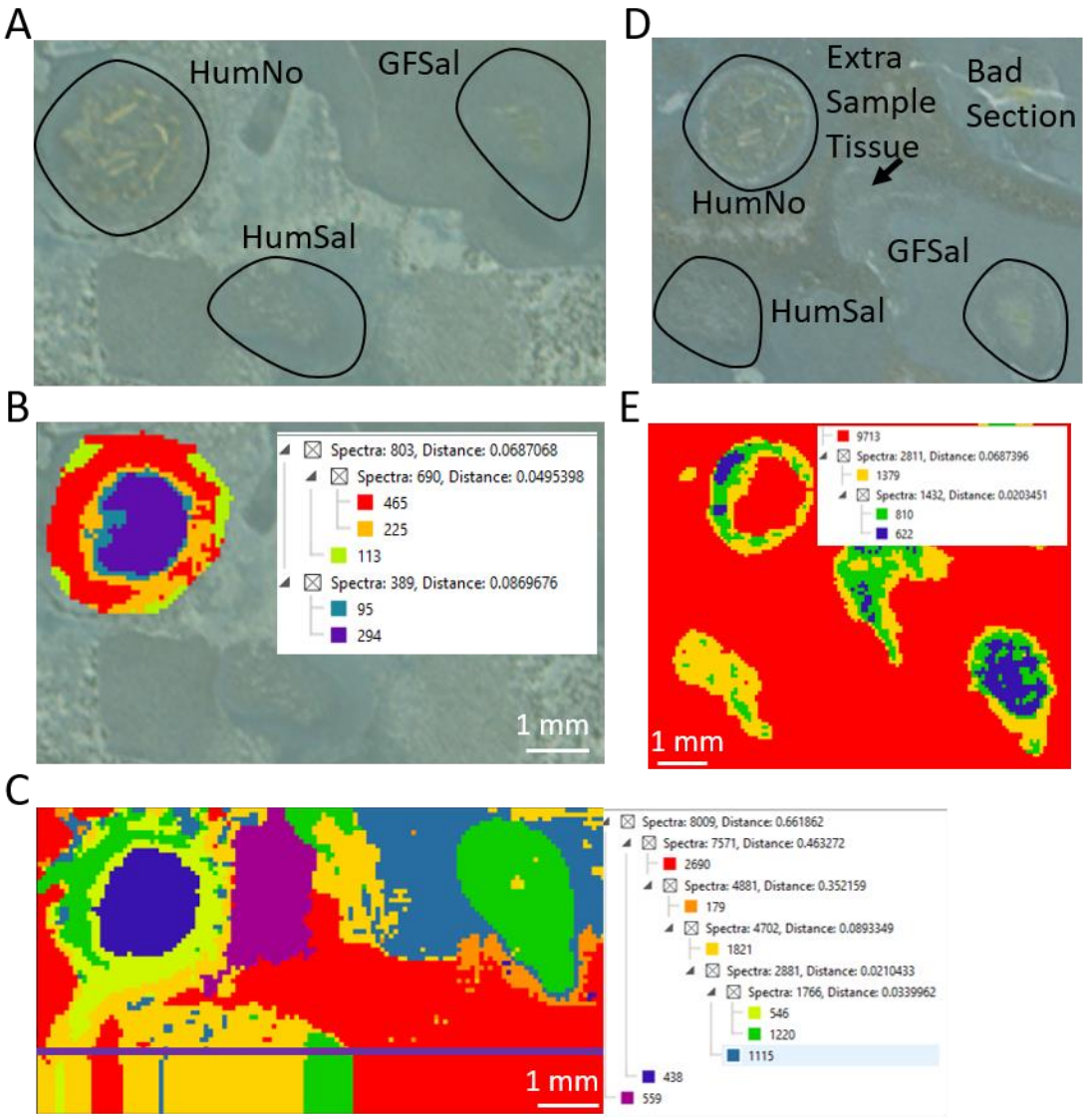



**Figure 1.** Results for cecum imaging on MALDI LTQ Orbitrap XL with CHCA **(A,B,C)** and DHB **(D,E)**. The optical image is shown in **(A)** for CHCA and **(D)** for DHB. Segmentation of the full imaged area is shown in **(C)** for CHCA and **(E)** for DHB. Segmentation of only the HumNo sample for CHCA is shown in **(B)**. The CHCA sample area was not completely imaged due to long instrument time requirements. The purple line indicates where the instrument acquisition stops.

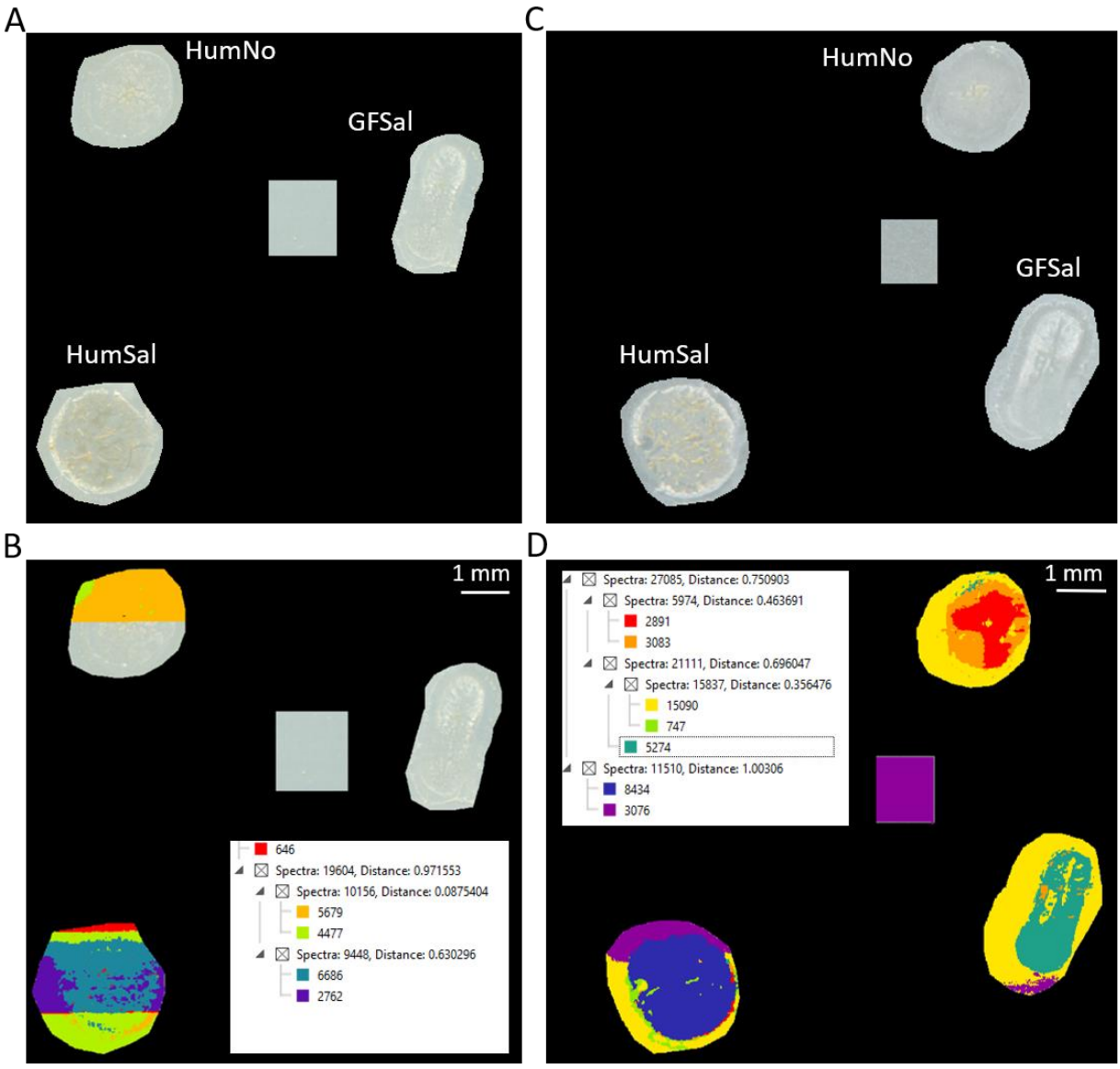



**Figure 2.** Results for UltrafleXtreme imaging on cecum samples. Optical images are shown in **(A)** for CHCA and **(C)** for DHB. Segmentation results are shown in **(B)** for CHCA and **(D)** for DHB.

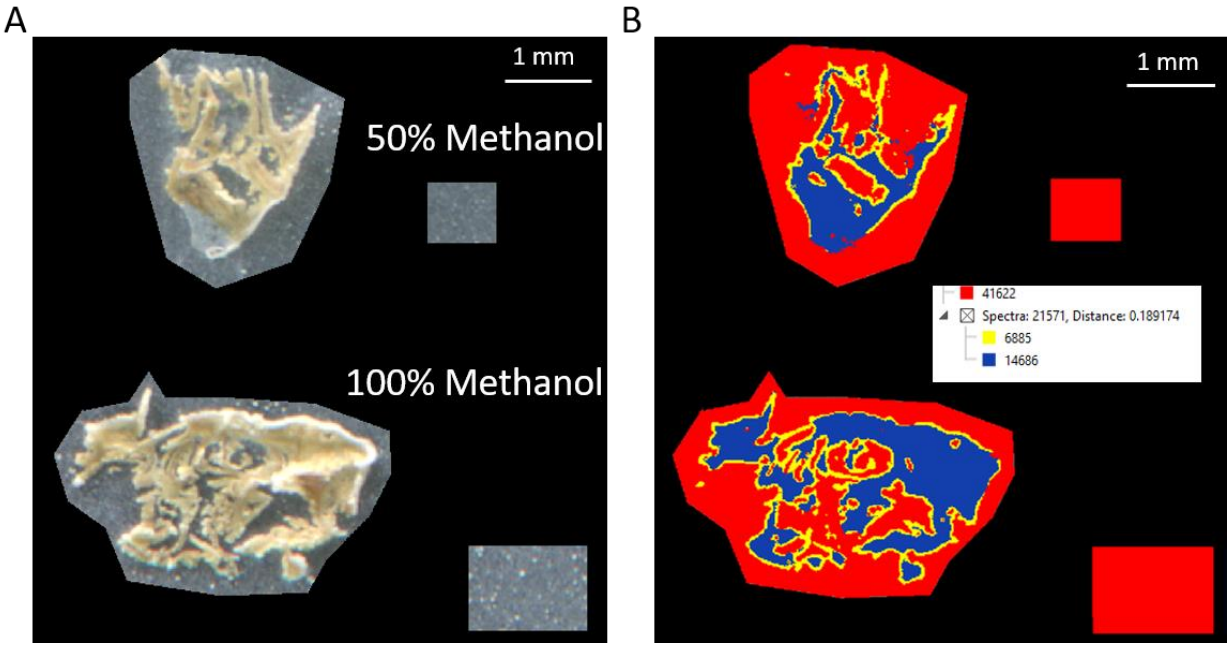



**Figure 3.** Results for RapifleX imaging on cecum samples. Optical image for DHB is in **(A)** and segmentation results are shown in **(B)**.

# Appendix VII

# Studying Bacterial Interactions via MALDI-MSI of small molecules on a PVDF Membrane

This project was in collaboration with Duane S Juang from the Beebe lab (UW-Madison Department of Biomedical Engineering)

Keyword: MALDI-MSI; PVDF; bacteria; small molecules; microbiome

**Abstract**

The microbiome plays an important role in human health. However, microbial interactions within the microbiome are difficult to study. Matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) has previously been used to study microbial interactions on agar to investigate secreted small molecules and potential natural products. In this appendix, a MALDI-MSI method for studying bacterial interactions on PVDF membrane is explored. PVDF membranes are typically used for binding of proteins in western blot analysis, but as demonstrated in this study, also allow small molecules to travel through the membrane. It was demonstrated in this study that MALDI-MSI of small molecules on PVDF membranes is a promising approach to study bacteria interactions as small molecule detection on the PVDF membrane is achievable.

**Introduction**

The microorganisms in the gastrointestinal tract, known as the gut microbiome, is composed of trillions of microorganisms, which encompass the 3 domains of life: bacterial, archaea and eukaryotes.[1, 2] The gut microbiome is critical to human health. For example, roles of the microbiome include nutrition and immune function modulation.[3, 4] Furthermore, changes or disruptions to the gut microbiome community has been associated with various inflammatory chronic conditions such as inflammatory bowel disease,[5] obesity,[6, 7] and cardiovascular disease.[8, 9] Due to the complex and dynamic nature of the gut microbiome, studying the interactions between the various members of the microbiome is a challenge.

Matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) is a useful technique to study the distribution of various biomolecules in a thin sample or tissue section. MALDI-MSI has been applied to a variety of biomolecules, including metabolites,[10]

neuropeptides,[11] and proteins.[12] In MALDI-MSI, a laser is fired at discrete positions, or pixels, across a sample. A mass spectrum is collected at each pixel, and software can then extract the ion intensity for a specific *m/z* across all the pixels to create an ion image for that *m/z*. In this way, hundreds of images can be generated for a specific tissue section. MALDI-MSI has been used to study ant-bacteria interactions on the ant exoskeleton,[13] as well as other host-microorganism interactions.[14] Hundreds of metabolites can be detected by MALDI analysis of microbe interactions without the need to perform an extraction from the agar growth media.[15, 16] The MALDI-MSI direct analysis of agar method has been used to study fungal-microbe interactions,[17] as well as various other microbial interactions to determine small molecule interactions and natural products.[18, 19]

Various polymer films have been used for MALDI-MS and MALDI-MSI analysis. Polyvinylidene fluoride (PVDF) is typically used in western blots, where its uniform pore structure and non-specific affinity for binding proteins enables detection and quantification of various proteins. PVDF has also been used as a substrate for protein detection via MALDI-MS.[20, 21] Here we investigated using a PVDF membrane as a substrate for MALDI-MSI imaging of bacteria interactions. Bacteria were grown in wells on top of a PVDF membrane and small molecules were allowed to travel across the PVDF membrane between the wells. MALDI-MSI is a quick and efficient method to investigate the small molecules traveling between the wells directly from the PVDF membrane.

## Experimental Procedures

### MALDI-MSI Sample Preparation

PVDF samples were prepared by the Beebe lab. Prepared PVDF membrane samples were trimmed from the PCR tape, with excess PCR tape surrounding the PVDF membrane. Enough excess PCR tape was left around the PVDF membrane so that if the sample came off in the ion source in the instrument (unlikely) then the sample would still be too big to enter the mass analyzer. The PVDF sample was taped onto the glass slide with double sided tape to the back of the PCR tape and with regular masking tape around the top and bottom of the PCR tape. 2,5-Dihydroxybenzoic acid (DHB; 40 mg/mL in 50% methanol, 0.1% formic acid) matrix was applied with the TM sprayer automatic sprayer system (1250 v, 24 passes, 30 s dry, 3 mm spacing, 0.05 mL/min, 10 psi, CC pattern). Samples were stored at -20$^{o}$C until analysis.

MALDI-MSI

MALDI-MSI was performed on a MALDI LTQ Orbitrap XL instrument. A rectangular area across the entire width of the PVDF membrane, but with a small height (investigating 2D movement only) was selected as the imaging area. A raster step size of 75 µm was used. Full MS parameters were as follows: 12-15 µJ laser energy, 60k resolution, 100-1000 *m/z*, 2 µscans per pixel. ImageQuest (Thermo Scientific) and MSiReader[22] software was used to analyze the results.

**Results**

**Figure 1** shows the workflow for imaging small molecules on PVDF membranes. PVDF membrane sections were attached to PCR tape, which is sticky when put under pressure, but not as sticky at normal pressures. Wells (cut into plastic) were placed over top of the PVDF membrane (**Figure 2**) and sample is placed in the wells. Molecules then can diffuse across the membrane between the wells. The PCR tape, or sections of the PCR tape, were attached to glass slides, matrix sprayed on the sample, and the sample was analyzed on the MALDI LTQ Orbitrap XL.

Small Molecule Standard

Adenosine (1 mg/mL in water) was used as a test sample to assess the ability of the PVDF membrane workflow to detect small molecules. Two TM sprayer methods were compared to test whether the flow rate needs to be higher to extract molecules off of the PVDF membrane. The 24 pass, 0.05 mL/min TM sprayer method, was compared to an 8 pass, 0.20 mL/min method, with all other parameters being equal. **Figure 3** shows the MALDI-MSI image for adenosine with the 24 pass method **(A)** and the 8 pass method **(B)**. The 8 pass method shows diffusion outside of the PVDF membrane, indicating that delocalization was an issue for the higher flow rate method. The 24 pass method showed high signal without diffusion, so the 24 pass method appears to be able to extract small molecule off of the PVDF membranes.

Bacteria Test System

To test whether bacteria interactions could be detected on the PVDF membrane, a quorum sensing bacteria system was utilized. Two autoinducer bacteria samples were cultured on the wells on top of the PVDF membrane. The autoinducer bacteria produce the quorum sensing (QS) molecule, N-Hexanoyl-L-homoserine lactone, which when sensed by the autoinducer bacteria, produces a purple molecule, violacein. **Figure 4** shows the MALDI-MSI results for the detection of these two molecules on the PVDF system. The QS molecule was not detected on the PVDF **(A)**, however, violacein was detected **(B)**. The lack of detection of the QS molecule could be due to quick degradation of the molecule when it is sensed by the bacteria, or perhaps be a concentration issue, where the QS molecule is present at too low of concentration to detect.

## Conclusions

The PVDF system is a promising method to detect small molecule interactions between two bacteria systems. A small molecule standard was observed in between the two wells in a test system. Also, in a test bacteria system, violacein, which is produced after a signal molecule is sensed by the bacteria, was detected. Further optimization is necessary to determine the lack of detection of the QS molecule. Analysis of a standard of the QS molecule would be useful to determine if the molecule ionizes well in MALDI and what the limit of detection is for the molecule. Different PVDF membranes could also be analyzed to find the best properties, such as the optimal pore size.[23]

## Acknowledgements

D.S.J. from the Beebe lab prepared samples by growing bacteria on top of the PVDF membrane. C.K. performed MALDI-MSI sample preparation, data collection, and analysis. C.K. would like to thank Yuan Liu, who is taking over future work on this project. The MALDI LTQ Orbitrap XL was purchased through an NIH shared instrument grant (NCRR S10RR029531 to LL).

## References

1. Scanlan, P. D.; Marchesi, J. R., Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *Isme j* **2008,** *2* (12), 1183-93.

2. Hoffmann, C.; Dollive, S.; Grunberg, S.; Chen, J.; Li, H.; Wu, G. D.; Lewis, J. D.; Bushman, F. D., Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One* **2013,** *8* (6), e66019.

3. Danneskiold-Samsøe, N. B.; Dias de Freitas Queiroz Barros, H.; Santos, R.; Bicas, J. L.; Cazarin, C. B. B.; Madsen, L.; Kristiansen, K.; Pastore, G. M.; Brix, S.; Maróstica Júnior, M.

R., Interplay between food and gut microbiota in health and disease. *Food Research International* **2019,** *115*, 23-31.

4. Kau, A. L.; Ahern, P. P.; Griffin, N. W.; Goodman, A. L.; Gordon, J. I., Human nutrition, the gut microbiome and the immune system. *Nature* **2011,** *474* (7351), 327-336.

5. Kostic, A. D.; Xavier, R. J.; Gevers, D., The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* **2014,** *146* (6), 1489-1499.

6. Turnbaugh, P. J.; Backhed, F.; Fulton, L.; Gordon, J. I., Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* **2008,** *3* (4), 213-23.

7. Rosenbaum, M.; Knight, R.; Leibel, R. L., The gut microbiota in human energy homeostasis and obesity. *Trends Endocrinol Metab* **2015,** *26* (9), 493-501.

8. Jie, Z.; Xia, H.; Zhong, S. L.; Feng, Q.; Li, S.; Liang, S.; Zhong, H.; Liu, Z.; Gao, Y.; Zhao, H.; Zhang, D.; Su, Z.; Fang, Z.; Lan, Z.; Li, J.; Xiao, L.; Li, R.; Li, X.; Li, F.; Ren, H.; Huang, Y.; Peng, Y.; Li, G.; Wen, B.; Dong, B.; Chen, J. Y.; Geng, Q. S.; Zhang, Z. W.; Yang, H.; Wang, J.; Zhang, X.; Madsen, L.; Brix, S.; Ning, G.; Xu, X.; Liu, X.; Hou, Y.; Jia, H.; He, K.; Kristiansen, K., The gut microbiome in atherosclerotic cardiovascular disease. *Nat Commun* **2017,** *8* (1), 845.

9. Tang, W. H.; Kitai, T.; Hazen, S. L., Gut Microbiota in Cardiovascular Health and Disease. *Circ Res* **2017,** *120* (7), 1183-1196.

10. Gemperline, E.; Jayaraman, D.; Maeda, J.; Ane, J. M.; Li, L., Multifaceted investigation of metabolites during nitrogen fixation in Medicago via high resolution MALDI-MS imaging and ESI-MS. *J Am Soc Mass Spectrom* **2015,** *26* (1), 149-58.

11. OuYang, C. Z.; Chen, B. M.; Li, L. J., High Throughput In Situ DDA Analysis of Neuropeptides by Coupling Novel Multiplex Mass Spectrometric Imaging (MSI) with Gas-Phase Fractionation. *Journal of the American Society for Mass Spectrometry* **2015,** *26* (12), 1992-2001.

12. Caprioli, R. M.; Farmer, T. B.; Gile, J., Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* **1997,** *69* (23), 4751-60.

13. Gemperline, E.; Horn, H. A.; DeLaney, K.; Currie, C. R.; Li, L., Imaging with Mass Spectrometry of Bacteria on the Exoskeleton of Fungus-Growing Ants. *ACS Chem Biol* **2017,** *12* (8), 1980-1985.

14. Klein, A. T.; Yagnik, G. B.; Hohenstein, J. D.; Ji, Z.; Zi, J.; Reichert, M. D.; MacIntosh, G. C.; Yang, B.; Peters, R. J.; Vela, J.; Lee, Y. J., Investigation of the Chemical Interface in the Soybean-Aphid and Rice-Bacteria Interactions Using MALDI-Mass Spectrometry Imaging. *Anal Chem* **2015,** *87* (10), 5294-301.

15. Watrous, J. D.; Dorrestein, P. C., Imaging mass spectrometry in microbiology. In *Nat Rev Microbiol*, England, 2011; Vol. 9, pp 683-94.

16. Yang, Y. L.; Xu, Y.; Straight, P.; Dorrestein, P. C., Translating metabolic exchange with imaging mass spectrometry. *Nat Chem Biol* **2009,** *5* (12), 885-7.

17. Moree, W. J.; Yang, J. Y.; Zhao, X.; Liu, W. T.; Aparicio, M.; Atencio, L.; Ballesteros, J.; Sanchez, J.; Gavilan, R. G.; Gutierrez, M.; Dorrestein, P. C., Imaging mass spectrometry of a coral microbe interaction with fungi. *J Chem Ecol* **2013,** *39* (7), 1045-54.

18. Gonzalez, D. J.; Xu, Y.; Yang, Y. L.; Esquenazi, E.; Liu, W. T.; Edlund, A.; Duong, T.; Du, L.; Molnar, I.; Gerwick, W. H.; Jensen, P. R.; Fischbach, M.; Liaw, C. C.; Straight, P.; Nizet, V.; Dorrestein, P. C., Observing the invisible through imaging mass spectrometry, a window into the metabolic exchange patterns of microbes. *J Proteomics* **2012,** *75* (16), 5069-76.

19. Hoffmann, T.; Dorrestein, P. C., Homogeneous matrix deposition on dried agar for MALDI imaging mass spectrometry of microbial cultures. *J Am Soc Mass Spectrom* **2015,** *26* (11), 1959-62.

20. Chen, S. H.; Liao, H. K.; Chang, C. Y.; Juo, C. G.; Chen, J. H.; Chan, S. I.; Chen, Y. J., Targeted protein quantitation and profiling using PVDF affinity probe and MALDI-TOF MS. *Proteomics* **2007,** *7* (17), 3038-50.

21. Chang, C.-Y.; Liao, H.-K.; Juo, C.-G.; Chen, S.-H.; Chen, Y.-J., Improved analysis of membrane protein by PVDF-aided, matrix-assisted laser desorption/ionization mass spectrometry. *Analytica Chimica Acta* **2006,** *556* (1), 237-246.

22. Robichaud, G.; Garrard, K. P.; Barry, J. A.; Muddiman, D. C., MSiReader: An Open-Source Interface to View and Analyze High Resolving Power MS Imaging Files on Matlab Platform. *Journal of the American Society for Mass Spectrometry* **2013,** *24* (5), 718-721.

23. Ino, Y.; Hirano, H., Mass spectrometric characterization of proteins transferred from polyacrylamide gels to membrane filters. *Febs j* **2011,** *278* (20), 3807-14.

**Figures**

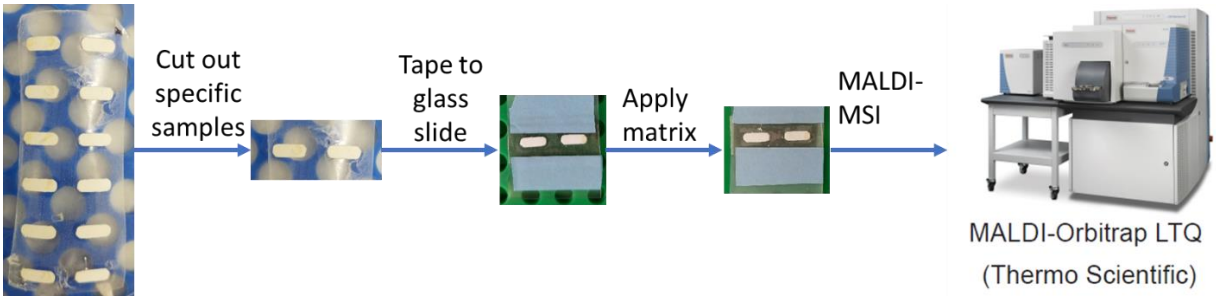



**Figure 1.** Workflow for MALDI-MSI of PVDF membranes.

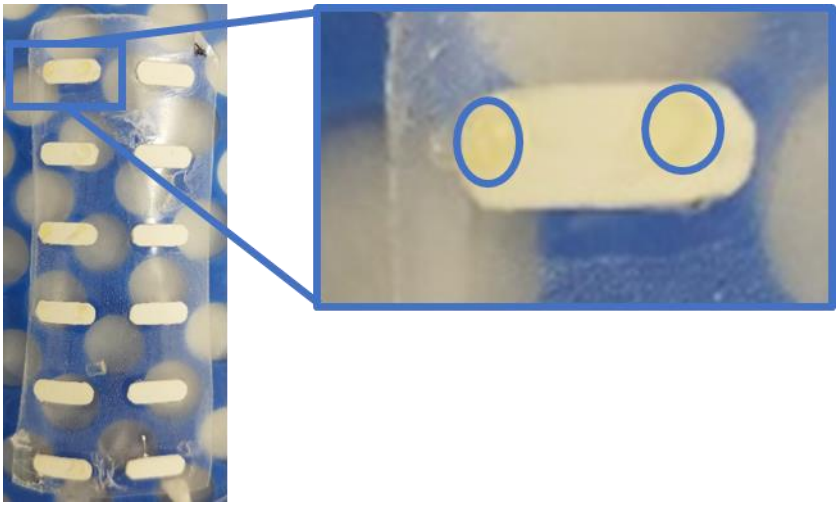


**Figure 2.** Picture of PVDF membrane set-up. Blue circles on PVDF membrane indicate where wells containing the sample would go over top of the membrane.

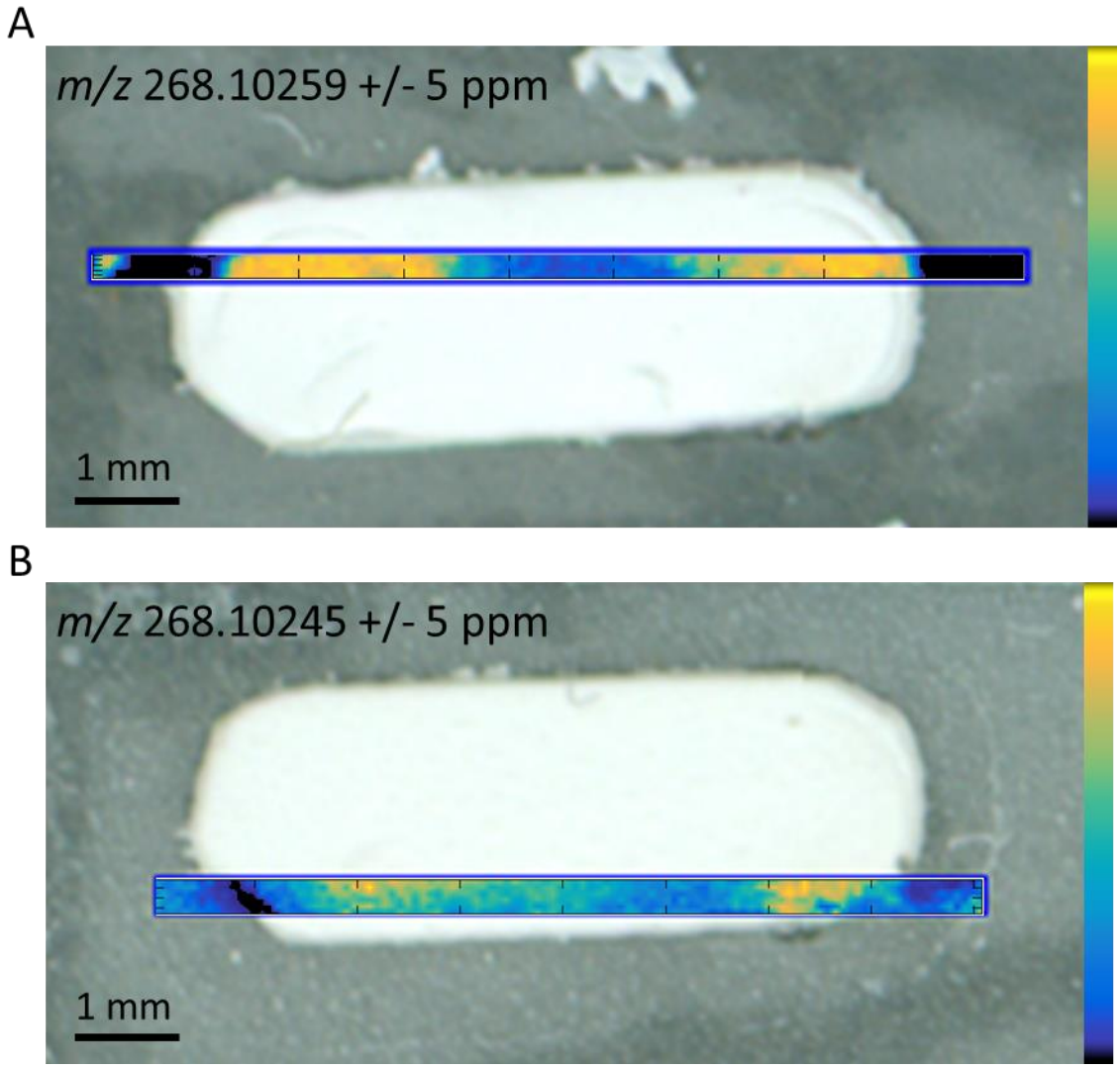



**Figure 3.** MADLI-MSI results for adenosine standard on PVDF membrane with low flow rate TM sprayer method **(A)** and high flow rate MALDI-MSI method **(B)**. The rectangular area imaged is show overtop of the whole PVDF membrane.

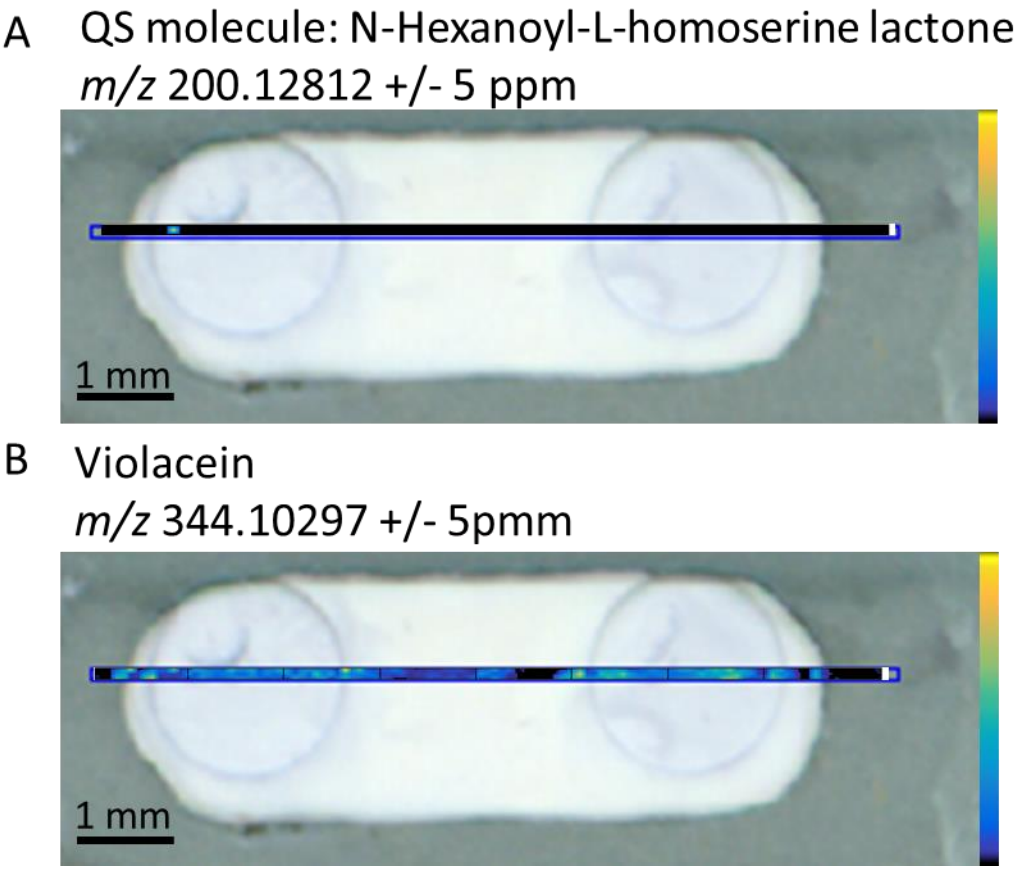



**Figure 4.** MALDI-MSI results for autoinducer quorum sensing bacteria system. **(A)** shows the quorum sensing molecules and **(B)** shows the violacein molecule (purple) produced after detection of quorum sensing molecule.

# Appendix VIII

# MALDI-MSI of Pancreatic Hormones in the Islets of Langerhans

This project was in collaboration with Donnie Stapleton and Dr. Mark P. Keller from the Attie lab (UW-Madison Department of Biochemistry)

Keyword: MALDI-MSI; pancreas islets; hormones; diabetes

**Abstract**

Matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) allows for label-free localization of many biomolecules in a single experiment. Although endogenous peptide imaging is challenging, determining the spatial distribution of neuropeptides and peptide hormones in tissues is important for gaining a better understanding of the roles of these molecules in tissue. Peptide hormones such as insulin and glucagon are critical for healthy blood sugar levels, and dysregulation of these peptides can lead to diseases such as diabetes. Here, a method was developed for optimal detection of peptide hormones by investigating different matrices, different MALDI-MSI instrumentation, and wash methods. Although the washes generally increased insulin detection, DHB without any washes resulted in the most peptides detected. A MALDI time of flight mass spectrometer was used for the studies over a MALDI Orbitrap instrument due to the larger mass range of the time of flight instrument.

**Introduction**

Matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) is a powerful tool to determine the localization of molecules in tissue.[1, 2] Although MALDI-MSI of peptides is routinely performed in neuropeptide-rich tissues, such as the brain, imaging of endogenous peptides remains challenging due to the low concentration of these molecules.[3] However, endogenous peptides play important roles throughout the body, and so developing MALDI-MSI methods to determine the localization of endogenous peptides in various tissues is important.

The pancreas, a member of the endocrine system, contains Islet of Langerhans that secrete peptide hormones that regulate blood sugar. The islets are groups of cells located throughout the

pancreas and make up only 1-2% of the overall pancreas. There are multiple cell types that make up the islets of Langerhans, including alpha cells, which make glucagon, beta cells, which make insulin, and delta cells, which make somatostatin.[4, 5] These cell types have specific distributions within the islets of Langerhans. Consequently, the peptides produced by each cell type should have a specific distribution in the islets. In mouse, islets are predominantly made up of the beta cells with other cell types localized to the outside of the islet.[6]

Numerous diseases are known to be a result of defects in the production of the pancreatic hormones. For example, in type 2 diabetes the body cannot properly use the hormone insulin and as a result the body over-produces it. The goal of this project was to use MALDI-MSI to map the distribution changes of pancreatic peptide hormones in the pancreatic islets due to type 2 diabetes using a mouse model for type 2 diabetes. Initially, sample preparation methods, including matrix application and wash methods were optimized. Multiple MSI instruments were also investigated.

**Experimental Procedures**

MALDI-MSI Sample Preparation

Pancreas organs were dissected from lean mice and rats and placed in double distilled water prior to freezing in liquid nitrogen or embedded in gelatin (100 mg/mL) prior to freezing on dry ice. While, the sequences and exact masses of the peptides present in mice and rat pancreas have slight differences, overall, mice and rat pancreas have the same complement of peptides present. Pancreas tissue was sectioned at 12 μm thickness and thaw-mounted onto either a glass slide (MALDI-Orbitrap) or an indium tin oxide (ITO)-coated glass slide (UltrafleXtreme). Washes were tested on some sections prior to matrix application, including 70% EtOH for 30 s followed by 100% EtOH for 30s, 70% IPA for 30 s followed by 100% IPA for 30s, and Carnoy's fluid (6:3:1

ethanol: chloroform: acetic acid) for 30 s followed by 100% ethanol for 30 s. A TM sprayer system (HTX Technologies) was used to apply matrix. DHB matrix (40 mg/mL in 50% MeOH 0.1% FA) was applied with a 24 pass method (1250 velocity, 0.1 mL/min, 80 C, 3 mm spacing, no dry time, rotate and offset). CHCA and SA matrix (5 mg/mL in 50% AcN 0.1% FA) was applied with an 8 pass method (1100 velocity, 0.2 mL/min, 90 C, 2 mm spacing, 30 s dry time, rotate and offset).

MALDI-MSI

MALDI-MSI was performed on both the MALDI-Orbitrap LTQ-XL instrument (Thermo Scientific) and on the UltrafleXtreme instrument (Bruker Daltronics). Runs on the MALDI-Orbitrap LTQ-XL were performed at 75 μm raster step size, and 60,000 resolution from 1000-4000 *m/z*. Xcalibur and LTQ tune software (Thermo Scientific) were used to create the imaging methods. MALDI-Orbitrap MSI data was exported into an imzML format in ImageQuest (Thermo Scientific) and images present in the tissue were selected using MSiReader [7]. MALDI-MSI on the UltrafleXtreme was performed at 50 μm raster step size in reflectron mode from 1000-6000 *m/z*. Prior to acquiring data, the UltrafleXtreme was calibrated with the proteomass High mass calibration mix (Thermo Scientific) with Insulin (Sigma Aldrich) spiked in. Data acquired on the UltrafleXtreme was analyzed in Flex Imaging (Bruker Daltonics).

## Results

**Table 1** provides some expected pancreatic peptide hormones and their *m/z* with and without post translational modifications (PTMs). Insulin is composed of two chains, the A and B chains, which are held together by disulfide bonds. During processing of the insulin preprohormone to form insulin, the A and B chains are released, along with the insulin C peptide, which is composed of amino acids between the A and B chains in the insulin preprohormone.

Furthermore, mice and rats have two function forms of Insulin, Insulin I and Insulin II, which are produced in a 1:2 ratio, respectively. Initial experiments performed on the MALDI-Orbitrap revealed many peptide signals, but not many of the peptide signals matched up to known pancreatic hormones. Additionally, initial experiments of peptides embedded in gelatin did not reveal any endogenous peptides. Due to concerns about peptide degradation when the tissue is place in gelatin, which is at 37$^{o}$C, all further studies were conducted on pancreas sections embedded in ice. **Figure 1** shows some example MALDI-MSI images. There were numerous *m/z* peaks located throughout the pancreas as shown in **Figure 1 (A)**. A small number of these were localized to discrete parts of the pancreas section. In **Figure 1 (B)**, the image of the distribution of *m/z* 3120.5624, which matches to a known peptide hormone in the pancreas (insulin 1 C peptide, 2.3 ppm). While the *m/z* in **Figure 1 (C)** is unknown, it shows a similar distribution to the insulin 1 C peptide. One major problem was that the hormone insulin is not detectable on the MALDI-Orbitrap as it is too large for the instrument's 4000 *m/z* upper limit. Although it has been shown that Insulin can be reduced on-tissue into its A and B chains,[8] attempts to reduce insulin into its A and B chain were unsuccessful in this experiment. As a result, MALDI-MSI experiments were carried out on the UltrafleXtreme, which is capable of imaging intact Insulin.

Of the three different matrices compared on the UltrafleXtreme for the best detection of pancreas hormones (DHB, CHCA, and SA), DHB was found to give the highest signal for pancreatic peptides. Washes were also investigated as possible sample preparation steps. Washes are a key step in protein imaging as they wash away lipids, which ionize very well and can suppress signal of other compounds. However, washes can sometimes also wash away peptides, making it problematic for peptide imaging. Here, washes were investigated due to larger molecule weight of intact insulin. The larger molecular weight means that insulin is less likely to be removed upon

washes. Additionally, as insulin is larger than typical peptides, the ionization efficiency for it in reflection mode of the instrument is less than what it would be for other peptides. By washing away interfering species, the detection of insulin could be increased. The washes investigated were: (1) 70% EtOH for 30 s followed by 100% EtOH for 30s, (2) 70% IPA for 30 s followed by 100% IPA for 30s, and (3) Carnoy's fluid (6:3:1 ethanol: chloroform: acetic acid) for 30 s followed by 100% ethanol for 30 s. **Figure 2** compares the ethanol wash (1) to an unwashed section with DHB as the matrix. Although the insulin signal is higher after the wash (**Figure 2 (C)**), many *m/z* values below 4,000 are only in the unwashed section (**Figure 2 (A-B)**). **Figure 3** shows the results of the other two washes for an example peptide in the 3,000-4,000 *m/z* range (**Figure 3(A)**) and for insulin II (**Figure 3 (B)**). The results of the isopropanol wash (2) were similar to the ethanol wash results as insulin was present after the wash, but many peptides signals under *m/z* 4,000 were not. The Carnoy wash (3), however, did not wash away all signals below 4,000 *m/z* (**Figure 3 (A)**). Thus, while washing the tissue section is potentially beneficial for insulin, the other peptide hormones are not benefited by washes. As the purpose of the study is to detect as many peptide hormones as possible, overall, not washing the tissue wash found to be the most advantageous strategy in order to maximize the number of small peptide hormones detected.

**Conclusions**

The optimal sample preparation procedure was determined to be DHB matrix with no washes as this yielded the most peptides detected on the pancreas section. Pancreas sections were embedded in ice rather than gelatin to minimize degradation of peptides during the embedding process. Further experiments to compare lean and obese B6 (normal) and BTBR (genetically

modified for type 2 diabetes) mice were not carried out due to similarity to and inability to expand upon previously published papers.[9, 10]

**Acknowledgements**

C.K. would like to thank Donnie Stapleton and Dr. Mark P. Keller from the Attie lab for collaborating on this project. D.S handled the mice for pancreas sample collection. C.K. performed MALDI-MSI sample preparation, data collection, and data analysis. The MALDI LTQ Orbitrap XL was purchased through an NIH shared instrument grant (NCRR S10RR029531 to LL).

**References**

1. Caldwell, R. L.; Caprioli, R. M., Tissue profiling by mass spectrometry: a review of methodology and applications. *Mol Cell Proteomics* **2005,** *4* (4), 394-401.

2. Caprioli, R. M.; Farmer, T. B.; Gile, J., Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* **1997,** *69* (23), 4751-60.

3. Chen, R.; Li, L., Mass spectral imaging and profiling of neuropeptides at the organ and cellular domains. *Anal Bioanal Chem* **2010,** *397* (8), 3185-93.

4. Barreto, S. G.; Carati, C. J.; Toouli, J.; Saccone, G. T., The islet-acinar axis of the pancreas: more than just insulin. *American journal of physiology. Gastrointestinal and liver physiology* **2010,** *299* (1), G10-22.

5. Chandra, R.; Liddle, R. A., Modulation of pancreatic exocrine and endocrine secretion. *Current opinion in gastroenterology* **2013,** *29* (5), 517-22.

6. Caicedo, A., Paracrine and autocrine interactions in the human islet: more than meets the eye. *Seminars in cell & developmental biology* **2013,** *24* (1), 11-21.

7. Robichaud, G.; Garrard, K. P.; Barry, J. A.; Muddiman, D. C., MSiReader: An Open-Source Interface to View and Analyze High Resolving Power MS Imaging Files on Matlab Platform. *Journal of the American Society for Mass Spectrometry* **2013,** *24* (5), 718-721.

8. Green-Mitchell, S. M.; Cazares, L. H.; Semmes, O. J.; Nadler, J. L.; Nyalwidhe, J. O., On-tissue identification of insulin: in situ reduction coupled with mass spectrometry imaging. *Proteomics. Clinical applications* **2011,** *5* (7-8), 448-53.

9. Minerva, L.; Boonen, K.; Menschaert, G.; Landuyt, B.; Baggerman, G.; Arckens, L., Linking Mass Spectrometric Imaging and Traditional Peptidomics: A Validation in the Obese Mouse Model. *Analytical chemistry* **2011,** *83* (20), 7682-7691.

10. Minerva, L.; Clerens, S.; Baggerman, G.; Arckens, L., Direct profiling and identification of peptide expression differences in the pancreas of control and ob/ob mice by imaging mass spectrometry. *Proteomics* **2008,** *8* (18), 3763-3774.

**Tables**

**Table 1: Expected Pancreatic Hormones and their Theoretical *m/z* in Mice**

| **Peptide** | **$[M+H]^+$** | **PTMs** | **$[M+H]^+$ w/ PTM** |
|---|---|---|---|
| Somatostatin-14 (103-116) | 1639.734 | disulfide bond 105-116 | 1637.724 |
| Insulin II C peptide (57-87) | 3132.570 | | |
| Insulin I C peptide (57-85) | 3120.570 | | |
| Glucagon (53-81) | 3481.623 | phosphoserine 54 | 3561.589 |
| Glicentin-related polypeptide (GRFP) (21-50) | 3439.425 | | |
| Amylin (38-74) | 3921.970 | disulfide bond (39-44) and tryosine amide (74) | 3918.970 |
| Insulin 1 | 5804.718 | Interchain disulfide bonds (31-94 and 43-107) | 5800.686 |
| Insulin 2 | 5797.642 | Interchain disulfide bonds (31-96 and 43-109) | 5793.611 |

**Figures**

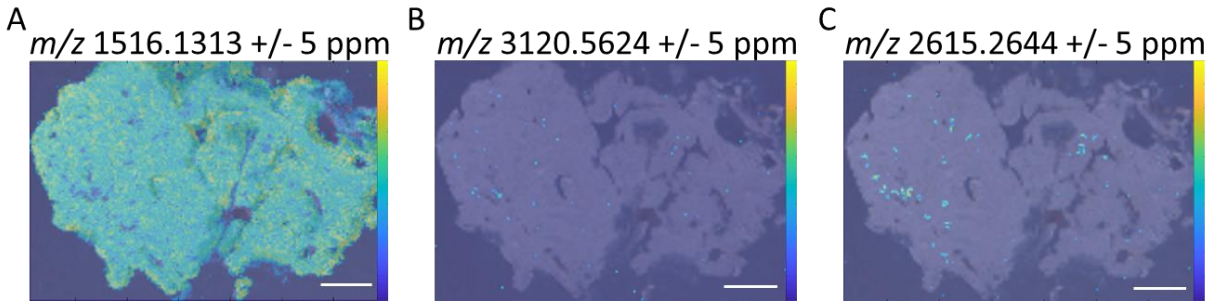



**Figure 1**. Three example MALDI-MSI images acquired on the MALDI-Orbitrap with DHB as the matrix. The white scale bar indicates 2 mm.

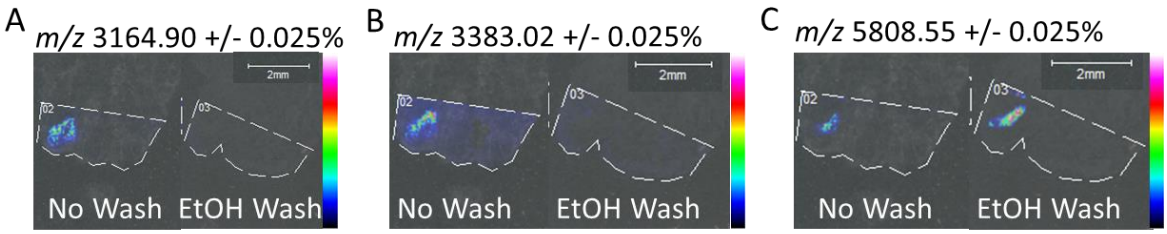



**Figure 2**. Example images from MALDI-MSI on the UltrafleXtreme for a rat Pancreas with DHB matrix. The images show both the unwashed section and the ethanol washed section.

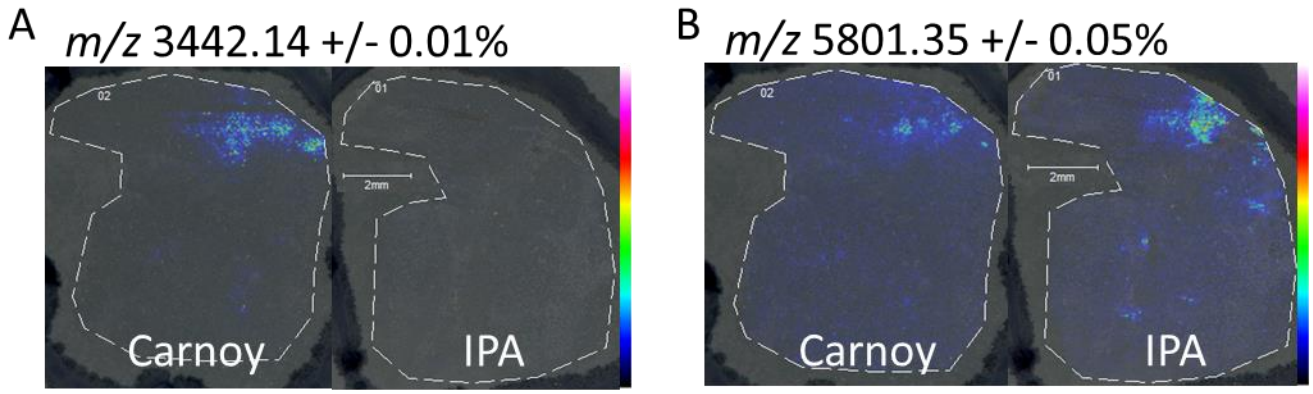



**Figure 3.** MALDI-MSI images obtained from the UltrafleXtreme with CHCA as a matrix. Both the Carnoy/EtOH wash and the 70%/100% isopropanol washes are shown.

# Appendix IX

# Next Generation Cancer Magnetic Resonance Imaging with Tumor-Targeted Alkylphosphocholine Metal Analogs

Adapted from: Zhang RR, Brunnquell CL, Hernandez R, Pinchuk AN, Keller C, Grudzinski JG, McMillan AB, Clark PA, Jeffrey J, Li L, Audhya A, Kuo JS, Weichert JP. Next Generation Cancer Magnetic Resonance Imaging with Tumor-Targeted Alkylphosphocholine Metal Analogs. *Radiology.* Submitted

Keywords: MRI; PET Imaging; MALDI-MSI; Gd chelates; alkylphosphocholines

**Abstract**

Purpose: Tumor-targeted alkylphosphocholines (APC) were developed as broad-spectrum cancer imaging and therapy agents to exploit the abundant selective retention of phospholipids by cancer cells relative to normal cells. Radioactive APC analogs have exhibited selective uptake and prolonged tumor retention in over 50 cancer types in preclinical models as well as over 15 human cancer types in over a dozen clinical trials. To advance the structural limits of this platform, we recently added a chelating moiety capable of binding gadolinium and many other metals for cancer-targeted magnetic resonance imaging (MRI), PET imaging, and targeted radionuclide therapy. A major aim was to synthesize, characterize, and validate the tumor selectivity of a new broad-spectrum, tumor-targeted, macrocyclic MRI chelate, Gd-NM600 in xenograft and orthotopic tumor models. A secondary aim was to identify and track the chemical speciation and spatial localization of this new chelate Gd-NM600 in order to assess Gd deposition.

Materials and Methods: T1-weighted relaxivities of Gd-NM600 were characterized in water and plasma at 1.5T and 3.0T. Tumor uptake and subcellular localization studies were performed using transmission electron microscopy. We imaged 8 different rodent models of human cancer over time and compared the T1-weighted imaging using Gd-NM600 with T1-weighted imaging using Gd-DOTA, a commercial macrocyclic Gd chelate. Finally, MALDI-MSI was used to characterize and map the distribution of the chemical species of Gd-NM600.

Results: Gd-NM600 exhibits high T1 relaxivities (approximately 16.4 s-1/mM at 1.5T), excellent tumor uptake (3.95 %ID/g at 48 hours) and MRI conspicuity. In rodent models of cancer, we observed significant uptake, prolonged tumor retention, and minimal uptake saturability of Gd-NM600. Broad-spectrum tumor-specific uptake was demonstrated in eight different human cancer models. Cancer cell uptake of Gd-NM600 via endosomal internalization and processing was

revealed with transmission electron microscopy. Importantly, tissue mass spectrometry imaging successfully interrogated the spatial localization and chemical speciation of Gd compounds and also identified breakdown products of Gd species.

Conclusion: We present a new macrocyclic cancer-targeted Gd chelate that achieves broad-spectrum tumor uptake and prolonged retention. Furthermore, we demonstrate its apparent *in vivo* stability by ultrahigh resolution MS tissue imaging This new tumor targeted contrast agent coupled with the enhanced imaging resolution of MRI relative to PET may transform oncologic imaging.

**Introduction**

Most clinically approved contrast agents for MRI are paramagnetic gadolinium chelates that are specific for cancer and exhibit relatively low T1-weighted relaxivities.[1-3] The low sensitivity of MRI (detectable at μM-nM concentrations) presents a barrier to targeted MRI of molecular processes. Although receptor-targeting strategies for PET and SPECT imaging are useful for tumor imaging, receptor-based MRI agents have failed to achieve sufficient contrast signal due to inadequate levels of even the most highly-expressed receptors.[3, 4] Cancer uptake of these non-targeted contrast agents is passive and transient, via blood-brain barrier disruption and/or the enhanced permeability and retention (EPR) effect.[2, 5] Therefore, the observed uptake signal more accurately reflects aberrations in vascularity, permeability, and tissue structure rather than molecular differences between cancer and normal tissue, thus contributing to diagnostic inaccuracies.[4, 6, 7]

In addition, most current Gd-based MR imaging agents exhibit relatively low relaxivities, whereas high relaxivities will likely enhance cancer detectability at the same contrast agent dose.[8-11] This is compounded by the controversy surrounding off-target gadolinium deposition detected

in patients that has led to Food and Drug Administration warnings on all commercial Gd chelates, as well as the suspended use of some linear chelates by the European Medicine Agency.[12] Although more evidence suggests that the linear chelates do deposit more in brain structures and other organs compared to macrocyclic chelates, some of the linear chelate agents banned in Europe have the highest relaxivities (Magnevist; Bayer Healthcare, Leverskusen, Germany) and superior diagnostic accuracy of cancer compared to macrocyclic Gd chelates.[12]

We report here for the first time a tumor-targeted macrocyclic Gd chelate that demonstrates superior relaxivity and selective uptake by cancer cells *in vitro* and *in vivo* compared to commercial macrocyclic Gd chelates. This novel agent circumvents many limitations of conventional MRI contrast agents through an observed high-capacity tumor cell uptake mechanism.[13-15] Moreover, this agent demonstrates significantly higher tumor and persistent adjacent tissue contrast relative to commercial macrocyclic MRI agents.

This new MRI contrast agent is derived from previous radioactive and fluorescent APC analogs that have demonstrated extensive broad tumor uptake and retention in numerous nuclear medicine imaging clinical trials, targeted radiotherapy clinical trials, and in preclinical fluorescence-guided surgical studies.[15-19] We now report the synthesis and biodistribution properties of Gd-NM600, a DOTA chelate of the tumor targeting APC molecule conjugated with gadolinium, in order to assess its potential for targeted MRI of cancer. We demonstrate the preservation of tumor cell avidity with the substitution of a chelate containing Gd.

To our knowledge, this represents the first small molecule Gd chelate that exhibits broad-spectrum tumor uptake and retention and demonstrably impressive T1-weighted relaxivities compared to other commercial macrocylic agents, enabling the use of lower Gd doses in imaging due to its high tumor uptake and prolonged retention, high T1 relaxivities, and potentially less

deposition into normal tissues. Furthermore, this agent may enhance diagnostic accuracy in several body regions including head and neck, spinal cord, breast, liver, and musculoskeletal imaging where MRI outperforms other imaging modalities.[20-23]

Lastly, because the chemical speciation of Gd chelates remains an important area of clinical research, we also developed a mass spectrometry methodology to interrogate the chemical speciation of this agent with high specificity and spatial localization. We demonstrate that this method can be used to characterize the chemical speciation of commercial linear and macrocyclic Gd chelates. This methodology may enhance our understanding of the behavior of Gd chelates used extensively in MRI and may finally shed light on the mechanisms of gadolinium deposition.

## Materials and Methods

### Animals

All animal studies were performed under University of Wisconsin-Madison Institutional Animal Care and Use Committee approved protocols. Studies were performed on mice bearing orthotopic glioblastoma (U87, n=7 and GSC 12.1, n=6), αβ-overexpressing triple negative breast cancer (n=4), flank prostate cancer (PC3, n=3), flank non-small cell lung cancer (A549, n=9), flank colorectal cancer (HT29, n=3), flank pancreatic cancer (MiaPaca, n=3), and flank glioblastoma (n=12). For flank xenograft models, $1\text{-}2\times10^6$ cells were injected into the right flank of nude athymic mice, and tumors were grown to a diameter of 5-15mm. For orthotopic xenograft glioblastoma models, $1\text{-}2\times10^6$ U87 or glioblastoma stem cells (GSC 12.1) were stereotactically implanted in the right striatum of immunodeficient NOD/SCID mice [24-26]. For breast cancer models, $10^6$ αβ-overexpressing triple negative breast cancer cells were prepared as previously

described and injected into the mammary ducts [27]. APC analogs for injection were formulated as previously reported.[13]

Relaxivity Measurements

Relaxivity of Gd-NM600 was measured at 1.5T and 3.0T (GE Signa HDxt and Signa PET/MR, GE Healthcare, Waukesha, WI) at 37°C. Gd-NM600 samples were prepared at concentrations of 0.125-1 mM in water, human plasma, and 0.125 mM human serum albumin. Sample temperature was maintained at 37°C with an MRI-compatible sample holder and warm water circulation system. For $T_1$ measurement, an inversion recovery pulse sequence with inversion time TI=50-750ms, repetition time TR=4000ms/5000ms (1.5T/3.0T), and echo time TE=8-9ms was utilized. For $T_2$ measurement, a spin-echo sequence with TR=5000ms and TE=50-600ms was utilized. $T_1$ and $T_2$ times were estimated using nonlinear least squares fitting of the signal magnitude vs. inversion time and echo time, respectively. Longitudinal ($r_1$) and transverse ($r_2$) relaxivities of Gd-NM600 were estimated from the slope of the linear relationship between relaxation rate and agent concentration.

Tumor Imaging

Gd-NM600 was administered intravenously (typically 0.12g/kg=0.11mmol/kg, up to 0.18g/kg=0.165mmol/kg) in eight rodent models of human cancer: mouse models include an orthotopic glioblastoma, a triple negative breast cancer, and flank xenografts of prostate cancer, non-small cell lung cancer, colorectal cancer, pancreatic cancer, and glioblastoma, and a rat flank glioblastoma xenograft model. *In vivo* MR imaging of mice was performed on a 4.7T preclinical scanner (Agilent Technologies, Santa Clara, CA) pre-contrast and at multiple time points for up to seven days following i.v. contrast administration. $T_1$-weighted tumor imaging was performed with a 2D fast spin echo pulse sequence, and tumor $R_1$ maps were estimated using 3D SPGR

acquisitions with variable flip angles and $B_1$ field correction.[28] A $T_1$-weighted 3D SPGR scan was used to image the abdomen and visualize biodistribution of Gd-NM600 in the heart, liver, and kidneys. To compare uptake of Gd-NM600 to a clinical agent, Gd-DOTA, a flank xenograft glioma model (U87) was used (n=3). To assess tumor-specific uptake, the tumor to muscle $T_1$-weighted signal ratio was computed across multiple time points in two models.

Tumor Saturation Imaging Study

To investigate saturation of uptake and contrast loading in tumor cells, an *in vivo* study involving serial daily imaging and Gd-NM600 injections was performed. $T_1$-weighted tumor imaging on 6 nude athymic female mice bearing A549 flank xenografts was performed prior to administration of Gd-NM600. Mice were delivered Gd-NM600 via tail vein (0.12g/kg) and underwent $T_1$-weighted imaging 24 hr after administration. Immediately following imaging, mice were delivered another equal dose of Gd-NM600. Three total iterations of Gd-NM600 injection followed 24 hours later by contrast-enhanced $T_1$-weighted imaging were performed.

MALDI-Mass Spectrometry Imaging Sample Preparation

Six animals bearing A549 flank xenografts were injected with 0.12g/kg of Gd-NM600 (n=3) or vehicle alone (n=3). Flash frozen organs (tumor, liver, and kidney) were embedded in 100 mg/mL gelatin (Becton, Dickinson, and Company) and stored at -80˚C. Organs were sectioned at 12 µm thickness on a cryostat (Microtom HM 525, Thermo Scientific) at -20˚C. Sections were thaw mounted onto indium tin oxide coated glass slides (Delta Technologies, Loveland, CO). 2,5-dihydroxybenzoic acid (DHB, Acros Organics) matrix (40 mg/mL in 50:50 methanol:$H_2O$) was applied using the TM Sprayer (HTX Technologies, LLC, Carrboro, NC, USA) automatic sprayer system. The TM sprayer method used 12 passes with 30 s dry time (rotation and offset), 3 mm spacing, 1250 velocity, 80˚C temperature, and 0.1 mL/min flow rate. Methanol was HPLC grade

(Fisher Scientific) and Milli-Q ultrapure water (Millipure Sigma) was used. Matrix covered samples were stored in a dry box at -20˚C until analysis.

MALDI-Mass Spectrometry Imaging Instrumentation

Matrix-assisted laser desorption/ionization-mass spectrometry imaging (MALDI-MSI) was performed in positive ion mode on a MALDI LTQ Orbitrap XL mass spectrometer (Thermo Scientific, Waltham, MA) equipped with a N2 laser. LTQ Tune software (Thermo Scientific, Waltham, MA) was used to select the imaging region and step size, and Xcalibur (Thermo Scientific, Waltham, MA) was used to select the instrument parameters. Imaging was performed on three control mice and three mice dosed with the Gd-NM600 compound at 75 μm raster step size, from 130-2000 mass to charge ratio (*m/z*) at 60,000 resolution. Two microscans were averaged at each pixel. ImageQuest software (Thermo Scientific, Waltham, MA) was used to view raw data and export the raw data to an imzML format. MSiReader software[29] was used to generate images unique to mice dosed with the compound (experimental mice). Briefly, *m/z* elements that were present in at least 10% of the interrogated zone (experimental mice) and in less than 5% of the reference zone (control mice) or in over 5% of the reference zone with a ratio greater than 2 were selected. Images for these *m/z* were then identified and manually inspected.

For MALDI mass spectra of Gd-DOTA, and Gd-BOPTA (MultiHance), a mixture of these compounds formulated at 1 μmol in distilled water and 2,5-dihydroxybenzoic acid (DHB, Acros Organics) matrix (40 mg/mL) was used to obtain the MALDI mass spectra.

Statistical Analysis

Statistical analyses were performed with Microsoft Excel or GraphPad Prism 5.0. Paired t-tests were used to determine differences in Gd-NM600 uptake based on imaging and biodistribution data. All $p<0.05$ were considered statistically significant.

**Results**

Synthesis and chelation of APC-chelates

Tri-benzyl ester **2** was prepared from commercially available cyclen **1** by treatment with 3.3 equivalents of benzyl bromoacetate and sodium acetate in *N*,*N*-dimethylacetamide using a modification of a published procedure (Supplemental Figure S1).[30] *N*-Alkylation of **2** with *tert*-butyl bromoacetate in acetonitrile using $K_2CO_3$ as a base afforded DOTA *t*-butyl tri-benzyl ester **3**. Deprotection of the *tert*-butyl ester in **3** with HCl in dioxane resulted in DOTA tri-benzyl ester **4**. This compound was coupled with 18-(*p*-aminophenyl)octadecyl phosphocholine **5** using COMU as a coupling reagent[31] to yield conjugate **6**. The simultaneous deprotection of the three benzyl ester groups in **6** by the catalytic hydrogenation resulted in DO3A derivative **7** which upon complexation with gadolinium (III) chloride produced the Gd chelate **8** (Supplementary Figure 1). Alternatively, $^{64}$Cu chelate of DOTA-APC **7** was also prepared as described above.

Gd-NM600 exhibits high relaxivity

The longitudinal and transverse relaxivities of Gd-NM600 in water, human plasma, and 0.125 mM human serum albumin were measured ($r_1$ and $r_2$, respectively). For relaxivity measurements, conical tubes containing 0.125-1 mM Gd-NM600 were warmed to 37°C in an MRI-compatible circulating water bath. With 1.5T field strength, the measured longitudinal relaxivity was between 14.9-16.5 $s^{-1}$/mM; at 3.0T, it ranged from 8.7-10.0 $s^{-1}$/mM (**Figure 1**). The transverse relaxivity ranged from 32.8-41.2 $s^{-1}$/mM at 1.5T and 23.7-33.2 $s^{-1}$/mM at 1.5T and 3.0T, respectively (**Figure 1**). Transverse relaxivity was always lowest in water compared to other preparations, while longitudinal relaxivity was more consistent across all preparations. The significantly higher longitudinal relaxivity of this agent compares favorably to current clinical agents with $r_1$ values ranging from 3.5-7.0 $s^{-1}$/mM at 1.5T and 2.5-6.5 $s^{-1}$/mM at 3.0T.[9]

Transmission electron microscopy confirms selective uptake in cancer cells

In order to investigate the mechanism of selective uptake, U87 glioblastoma cells and SK fibroblasts were incubated with 1 μM Gd-NM600 and observed under TEM at various time points up to 24 hours. Starting after 15 minutes of incubation and at all later time points, Gd-NM600 uptake is observed in U87 cells (**Figure 2(a)**). At early time points, the agent accumulates in early endosomes. As endosomal compartments mature and acidify, they become more electron dense and ultimately fuse with lysosomes. These electron-dense clusters corresponding to accumulation in endosomes and lysosomes are observed markedly less in skin fibroblast control cells. Notably, no dense clusters corresponding to gadolinium are observed in the nucleus (**Figure 2(c)**), consistent with previous studies that show no nuclear uptake of APCs and APC-like molecules.[15, 32, 33] Surface transmission electron microscopy confirmed high metal contrast inside the U87 cancer cells (**Figure 2(d)**).

$T_1$-weighted imaging and biodistribution demonstrate tumor-specific uptake and retention in multiple cancer models

*In vivo* uptake of Gd-NM600 was characterized with pilot studies using two flank xenograft models of human cancer, A549 (non-small cell lung cancer) and U87 (glioma). $T_1$-weighted MR images and $T_1$ maps were acquired in all subjects prior to intravenous contrast administration and imaging was repeated at multiple time points up to seven days (**Figure 3**). Flank tumors of both models showed tumor-specific enhancement over the 7-day period. Tumor enhancement was maintained for up to 4 and 7 days in A549 and U87, respectively ($p<0.05$). Specifically, the tumor to control (muscle) signal ratio at 24 hours post Gd-NM600 administration was 1.7 and 1.9 times that in pre-contrast imaging for A549 and U87 tumors, respectively. $T_1$ maps acquired at pre-contrast and 48 hour time points confirmed that a significant increase in the whole-tumor median

$R_1$ relaxation rate was observed at 48 hours post-contrast in both tumor models (increased to approximately 2.2 times pre-contrast $R_1$ in both tumor models, $p<0.05$). In comparison, gadoterate meglumine (Dotarem®), Gd chelated by DOTA, was investigated in mouse flank tumor xenografts. Three nude mice with U87 flank xenografts were scanned with a $T_1$-weighted sequence prior to administration of Gd-DOTA, immediately following contrast administration, and at multiple time points up to 24 hours. After Gd-DOTA administration, the $T_1$-weighted tumor to muscle signal ratio rapidly increased over the course of five minutes, reduced over the course of one hour, and returned to baseline signal levels by 24 hours (**Figure 3(d-e)**). Gd-DOTA administration increased ratio of tumor to muscle $T_1$-weighted signal from 1.15 pre-contrast to 1.67 five minutes post-contrast. With Gd-NM600 administration, tumor to muscle signal ratio increased from 1.24 pre-contrast to a maximum of 2.12 at 24 hours post-contrast. These results indicate increased uptake and prolonged retention of Gd-NM600 in cancer cells that are not observed with the extracellular macrocyclic agent Gd-DOTA.

Several additional tumor models were investigated to further demonstrate the broad-spectrum tumor uptake of Gd-NM600. These models include colorectal flank xenograft models (HT29), a flank xenograft prostate cancer model (PC3), an orthotopic αβ-overexpressing triple negative breast cancer model[34], pancreatic cancer (MiaPaca), and two orthotopic brain cancer models (histologically confirmed U87 and a cancer stem cell model, GSC 12.1). Prolonged $T_1$-weighted signal enhancement was observed with *in vivo* $T_1$-weighted imaging in all of these models between 24 and 96 hours following Gd-NM600 administration (Fig 4). The uptake observed in a wide variety of tested cancer types suggests a broad-spectrum uptake mechanism.

Biodistribution in U87 flank tumor-bearing mice was assessed both *ex vivo* and *in vivo*. Seventy-two hours following administration of Gd-NM600, organs were harvested and Gd uptake

was quantified using ICP-MS. Gd content in the tumor (3.95% injected dose per gram, (%ID/g)) was higher than in all other tissues except the organs of clearance, the liver and kidney (Supplemental figure S2). *Ex vivo* biodistribution of bulk administration of Gd-NM600 was quite similar to that of trace levels of $^{64}$Cu-DOTA-APC, including in tumors, despite three orders of magnitude difference in amount of agent delivered ($10^{-6}$ mol vs. $10^{-9}$ mol, Supplemental Figure S3a). *In* vivo biodistribution of Gd-NM600 and $^{64}$Cu-DOTA-APC in a flank-bearing U87 rat model demonstrated excellent localization of both PET signal and T1-weighted enhancement on simultaneous PET/MR imaging with the co-administration of these two agents at different mass doses (**Supplemental Figure S3(b-d)**). Ex *vivo* biodistribution was compared to *in vivo* observations of Gd-NM600 in the blood pool, liver, and kidneys. $T_1$-weighted imaging of the abdomen demonstrated Gd-NM600 clearance from the liver, kidneys, and blood over the 7-day period (**Supplemental Figure S2**). In contrast, 24 hours following Gd-DOTA administration, little observable agent remained in circulation (**Supplemental Figure S2(b)**). High uptake in the liver and kidneys was observed, with maximum liver signal 24 hours and maximum kidney uptake one hour following contrast administration.

Chemical Speciation of Gadolinium Chelates by MALDI-MSI of Gd-DOTA APC

Matrix-assisted laser desorption/ionization (MALDI) is a soft ionization technique that harnesses a deposited matrix, which absorbs laser energy and ionizes neighboring molecules such that the chemical form of compounds of interest is preserved.[35, 36] Coupled with a mass analyzer, MALDI can serve as a powerful tool to identify compounds in tissues with high mass accuracy and also map their spatially distribution.[35] Notably, this method can identify Gd deposits with high specificity due to the unique isotopic fingerprint of Gd.

Mice bearing A549-flank xenografts were injected with either Gd-NM600 (n=3), or the formulation alone (n=3). Based on prior biodistribution data, we chose to collect the tissues with the highest Gd content (liver, kidney, and tumor; **Supplemental Figure S2(a)**). After 7 days, the livers, kidneys, and tumor tissues were harvested, and prepared for MALDI-MSI. MALDI-MSI of the tissues demonstrated the intact chemical species of Gd-NM600 in all three tissues from treated animals, but not in tissues of control animals (**Figure 5(a-c)**). Overlay of the histology and the Gd-NM600 distribution revealed homogenous signal in the kidneys and liver, and heterogeneous signal in the tumor (**Figure 5(c)**). The unique isotopic distribution of the Gd species is identifiable in the mass spectra (**Figure 5(d)**). Notably, a rim of high signal around the tumor margin was observed in the tumor tissues (**Figure 5(a)**), suggesting the highest abundance of the compound at the edge of the tumor margin, most likely representing dividing tumor cells.

To assess if MALDI-MSI is useful to characterize deposition of commercial chelates, we also employed this method for two commercial chelates. Gd-DOTA, and Gd-BOPTA (MultiHance) at the same concentration were incorporated into a mixture with the same matrix and analyzed using MALDI-MSI (**Supplemental Figure S4(a-b)**). Identification of these compounds was achieved with high mass accuracy. Again, Gd's unique isotopic distribution is observed in the mass spectra. These results suggest that this methodology is able to identify the integrity and chemical species of Gd being deposited in normal tissues. This methodology can be used to identify Gd species with high mass accuracy and specificity, and also assess chemical integrity of the chelated species of gadolinium.

In order to demonstrate that this methodology can be applied to characterize the breakdown products of Gd-NM600, we searched for other breakdown products including ones that contain Gd under positive ion mode. Three mice containing A549 flank xenografts (n=3) were injected with

Gd-NM600. After 7 days, the liver, kidneys and tumors were harvested and analyzed using MALDI-MSI. Although the highest signals corresponded to the parent compound, cleavage of the amide linker was observed in the liver and tumor sections, but not in the kidneys (**Supplemental Figure S5(a-c)**). Loss of carboxylic acid was also observed in the tumors (**Supplemental Figure S5(d-e)**), but not in the liver and kidney. Further work is underway to assess the chemical speciation of our Gd-APC chelates at longer time points using MALDI-MSI.

Gd-NM600 tumor uptake mechanism demonstrates low saturability

We hypothesized that the tumor uptake mechanism of Gd-NM600 would exhibit minimal saturability. This potentially allows increasing contrast delivery to tumors to address MRI's low sensitivity for molecular targeted contrast agents. To test this hypothesis, six animals bearing A549 flank xenografts were administered three times the standard mass dose (3×0.12g/kg) of Gd-NM600 over the course of three days (one administration per day). Twenty-four hours after each contrast injection, $T_1$-weighted imaging was performed. All six animals in this study tolerated the increased total dose of 0.36g/kg (0.34mmol/kg) Gd-NM600 over 72 hours. Tumor enhancement increased over the three days of post-contrast imaging and did not reach saturation even after three consecutive days of contrast delivery (**Figure 6**). The average tumor to muscle ratio increased from 1.89× to 2.43× the baseline value over three days. In comparison, this ratio slowly decreased from 1.90× to 1.41× the baseline value in the three animals administered a single mass dose of Gd-NM600. In this study, complete saturation of Gd-NM600 uptake was not observed over three consecutive days of contrast delivery.

**Discussion**

Currently, targeted MRI contrast agents have demonstrably failed due to low relaxivities, and insufficient ability to obtain sufficient concentrations of Gd in cancers. This challenge is further complicated by deposition of Gd in normal human tissues, which is found more often with the linear Gd chelates. This report intends to introduce and validate a new small molecule, macrocyclic Gd contrast agent that exhibits high relaxivity properties, and broad-spectrum, tumor-targeting. Gd-NM600 can image multiple cancer phenotypes with higher contrast than commercial Gd chelates. Gd-NM600 demonstrates a high longitudinal relaxivity in water of $16.4 s^{-1}$/mM at 1.5T, which is approximately 4.2-fold higher relaxivity compared to Gd-DOTA[37, 38] that has a longitudinal relaxivity of $3.9 s^{-1}$/mM at 1.5T.[37, 38] The significantly higher relaxivities of Gd-NM600 coupled with its cancer selectivity and prolonged retention may improve the diagnostic accuracy for soft tissue cancers and metastases over current commercial Gd contrast agents. Previous crossover studies have shown improved cancer visualization, definition, and contrast enhancement when MRI agents with higher relaxivities are used instead of lower relaxivity agents.[10, 11, 39-41] The correlation of higher tumor conspicuity with the use of high relaxivity agents is promising as Gd-NM600 demonstrates impressive longitudinal relaxivities in plasma, water, and human serum albumin at 1.5T and 3.0T, which are 4.2 times higher than even the best-performing commercial macrocyclic Gd MRI agents.[9] Moreover, enhanced tumor selectively should result in use of lower Gd-NM600 doses and minimize off-target gadolinium burden and deposition.

Furthermore, Gd-NM600 appears to exhibit high capacity uptake and minimal saturability by cancer cells at clinically relevant dose levels. Unlike many other targeting which bind to cell receptors, thus limiting their sensitivity, our APC analogs enter cells via lipid rafts overexpressed

in cancer and are subsequently incorporated into endosomes as seen by TEM imaging. This high-capacity mechanism allows for high concentrations of Gd to accumulate within cancer cells, which translates into elevated and persistent tumor contrast enhancement. This distinctive mechanism of cell entry through lipid rafts and endosomal targeting circumvents traditional receptor-based approaches, which have failed due to inadequate concentrations of even the highest-expressed cell surface receptors to achieve T1-weighted contrast enhancement.[3, 4] Importantly, the high capacity uptake and prolonged tumor retention (exceeding a week) enable some additional clinical applications of this agent, such as motion management for MRI-guided radiation therapy or neutron-capture therapy.

Long-term retention of Gd MRI agents in patients after multiple Gd exposures remains a contentious issue given the observed deposition of Gd in the globus pallidus and dentate nucleus after multiple long term exposures.[42, 43] Macrocyclic Gd-NM600 was designed to offer better thermodynamic and kinetic stability relative to linear chelates.[37] As the debate continues on the clinical relevance of Gd deposition in normal brain tissues, we have demonstrated the utility of MALDI-MSI to enhance our understanding of the chemical speciation and spatial distribution of Gd-NM600 and other Gd MRI agents.

MALDI-MSI mapping of Gd-NM600 in the highest uptake tissues demonstrates some notable breakdown products at 7 days post-injection, some of which are present in only the tumors, and some of which appear to be broken down and metabolized by only the liver. To the best of our knowledge, this is the first time the chemical speciation has been characterized for Gd-containing compounds and offers proof-of-concept methodologty to understand the controversial phenomenon of Gd deposition. Further work is underway to characterize the breakdown products

as well as any Gd species that are deposited at time points beyond a week in order to understand how Gd-chelates interact with normal tissue.

## Conclusion

In this study, we report a new cancer-targeted MRI contrast agent that demonstrates high relaxivity, and tumor-specific uptake and retention in all tested models, including six different flank xenograft cancer models and three orthotopic cancer models. This novel targeted macrocyclic agent NM600 for cancer outperformed commercial Gd contrast agents in our studies by providing higher relaxivities, superior tumor conspicuity, and prolonged retention. Its advantages include a 4.2-fold higher longitudinal relaxivity compared to Gd-DOTA and higher cancer specificity due to the APC targeting moiety, while maintaining the stability of a macrocyclic DOTA chelate. The higher relaxivities and tumor specificity may enable detection of smaller soft-tissue tumors and metastases throughout a patient compared with current clinical Gd chelates. We have observed a distinct cancer-targeting mechanism of these APC chelates that is not receptor-mediated, has high capacity, and appears broadly selective for a broad array of cancers. Additionally, we demonstrate a specific and sensitive method to identify and spatially map the presence and the chemical form of Gd containing compounds, including Gd-NM600, in tissue samples through MALDI-MSI analysis.

## Acknowledgements

R.R.Z. is partially supported by the University of Wisconsin MD/PhD program via T32 GM008692. We are grateful to the University Of Wisconsin Department Of Radiology Research and Development Pilot Grant program, Dept. of Neurological Surgery, and the University of Wisconsin Carbone Cancer Center Support Grant (P30 CA014520) for supporting this work. PAC

and JSK were partly supported by the Roger Loff Memorial Fund for GBM Research. We also thank Dr. Caroline Keller, Dr. Lingjun Li for their mass spectrometry work and use of their core facility, Dr. Martin Schafer at the Wisconsin State Lab of Hygiene for use of Magnector sector instrumentation and analysis, and Dr. Vincent Cryns for his labs triple negative breast cancer model. The MALDI LTQ Orbitrap XL instrument was purchased through an NIH shared instrument grant (NCRR S10RR029531 to LL). L.L. acknowledges a Vilas Distinguished Achievement Professorship and the Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy. We also thank Dr. Thomas Grist (UW), and Dr. Scott Reeder (UW) for early discussions and feedback as well as manuscript preview by Dr. Michael Tweedle (OSU).

**References**

1. Lohrke, J.; Frenzel, T.; Endrikat, J.; Alves, F. C.; Grist, T. M.; Law, M.; Lee, J. M.; Leiner, T.; Li, K. C.; Nikolaou, K.; Prince, M. R.; Schild, H. H.; Weinreb, J. C.; Yoshikawa, K.; Pietsch, H., 25 Years of Contrast-Enhanced MRI: Developments, Current Challenges and Future Perspectives. *Advances in therapy* **2016,** *33* (1), 1-28.

2. Zhou, Z.; Lu, Z. R., Gadolinium-based contrast agents for magnetic resonance cancer imaging. *Wiley interdisciplinary reviews. Nanomedicine and nanobiotechnology* **2013,** *5* (1), 1-18.

3. Caravan, P.; Zhang, Z., Targeted MRI Contrast Agents. In *The Chemistry of Contrast Agents in Medical Magnetic Resonance Imaging*, 2 ed.; Merbach, A.; Helm, L.; Toth, E., Eds. John Wiley & Sons: 2013; pp 311-342.

4. Caravan, P., Strategies for increasing the sensitivity of gadolinium based MRI contrast agents. *Chemical Society Reviews* **2006,** *35* (6), 512-523.

5. Aime, S.; Caravan, P., Biodistribution of gadolinium-based contrast agents, including gadolinium deposition. *Journal of magnetic resonance imaging : JMRI* **2009,** *30* (6), 1259-67.

6. Zhou, Z.; Han, Z.; Lu, Z. R., A targeted nanoglobular contrast agent from host-guest self-assembly for MR cancer molecular imaging. *Biomaterials* **2016,** *85*, 168-79.

7. Mullins, M. E.; Barest, G. D.; Schaefer, P. W.; Hochberg, F. H.; Gonzalez, R. G.; Lev, M. H., Radiation necrosis versus glioma recurrence: conventional MR imaging clues to diagnosis. *American Journal of Neuroradiology* **2005,** *26* (8), 1967-1972.

8. Raymond, K. N.; Pierre, V. C., Next generation, high relaxivity gadolinium MRI agents. *Bioconjugate chemistry* **2005,** *16* (1), 3-8.

9. Rohrer, M.; Bauer, H.; Mintorovitch, J.; Requardt, M.; Weinmann, H.-J., Comparison of magnetic properties of MRI contrast media solutions at different magnetic field strengths. *Investigative radiology* **2005,** *40* (11), 715-724.

10. Maravilla, K. R.; Maldjian, J. A.; Schmalfuss, I. M.; Kuhn, M. J.; Bowen, B. C.; Wippold, F. J., 2nd; Runge, V. M.; Knopp, M. V.; Kremer, S.; Wolansky, L. J.; Anzalone, N.; Essig, M.; Gustafsson, L., Contrast enhancement of central nervous system lesions: multicenter intraindividual crossover comparative study of two MR contrast agents. *Radiology* **2006,** *240* (2), 389-400.

11. Rowley, H. A.; Scialfa, G.; Gao, P. Y.; Maldjian, J. A.; Hassell, D.; Kuhn, M. J.; Wippold, F. J., 2nd; Gallucci, M.; Bowen, B. C.; Schmalfuss, I. M.; Ruscalleda, J.; Bastianello, S.; Colosimo, C., Contrast-enhanced MR imaging of brain lesions: a large-scale intraindividual crossover comparison of gadobenate dimeglumine versus gadodiamide. *AJNR. American journal of neuroradiology* **2008,** *29* (9), 1684-91.

12. Gulani, V.; Calamante, F.; Shellock, F. G.; Kanal, E.; Reeder, S. B., Gadolinium deposition in the brain: summary of evidence and recommendations. *The Lancet Neurology* **2017,** *16* (7), 564-570.

13. Pinchuk, A. N.; Rampy, M. A.; Longino, M. A.; Skinner, R. W.; Gross, M. D.; Weichert, J. P.; Counsell, R. E., Synthesis and structure-activity relationship effects on the tumor avidity of radioiodinated phospholipid ether analogues. *Journal of medicinal chemistry* **2006,** *49* (7), 2155-65.

14. Weichert, J. P.; Pinchuk, A.; Kozak, K.; Longino, M.; Grudzinski, J.; Titz, B.; Pak, C.; Stehle, N., Phospholipid ether analogs as cancer-targeting drug vehicles. Google Patents: 2016.

15. Weichert, J. P.; Clark, P. A.; Kandela, I. K.; Vaccaro, A. M.; Clarke, W.; Longino, M. A.; Pinchuk, A. N.; Farhoud, M.; Swanson, K. I.; Floberg, J. M., Alkylphosphocholine analogs for broad-spectrum cancer imaging and therapy. *Science translational medicine* **2014,** *6* (240), 240ra75-240ra75.

16. Grudzinski, J. J.; Titz, B.; Kozak, K.; Clarke, W.; Allen, E.; Trembath, L.; Stabin, M.; Marshall, J.; Cho, S. Y.; Wong, T. Z.; Mortimer, J.; Weichert, J. P., A phase 1 study of 131I-CLR1404 in patients with relapsed or refractory advanced solid tumors: dosimetry, biodistribution, pharmacokinetics, and safety. *PloS one* **2014,** *9* (11), e111652.

17. Lubner, S. J.; Mullvain, J.; Perlman, S.; Pishvaian, M.; Mortimer, J.; Oliver, K.; Heideman, J.; Hall, L.; Weichert, J.; Liu, G., A Phase 1, Multi-Center, Open-Label, Dose-Escalation Study of 131I-CLR1404 in Subjects with Relapsed or Refractory Advanced Solid Malignancies. *Cancer investigation* **2015,** *33* (10), 483-489.

18. Swanson, K. I.; Clark, P. A.; Zhang, R. R.; Kandela, I. K.; Farhoud, M.; Weichert, J. P.; Kuo, J. S., Fluorescent Cancer-Selective Alkylphosphocholine Analogs for Intraoperative Glioma Detection. *Neurosurgery* **2015,** *76* (2), 115-124.

19. Zhang, R. R.; Schroeder, A. B.; Grudzinski, J. J.; Rosenthal, E. L.; Warram, J. M.; Pinchuk, A. N.; Eliceiri, K. W.; Kuo, J. S.; Weichert, J. P., Beyond the margins: real-time detection of cancer using targeted fluorophores. *Nat Rev Clin Oncol* **2017,** *advance online publication*.

20. Torigian, D. A.; Zaidi, H.; Kwee, T. C.; Saboury, B.; Udupa, J. K.; Cho, Z.-H.; Alavi, A., PET/MR imaging: technical aspects and potential clinical applications. *Radiology* **2013,** *267* (1), 26-44.

21. Antoch, G.; Vogt, F. M.; Freudenberg, L. S.; Nazaradeh, F.; Goehde, S. C.; Barkhausen, J.; Dahmen, G.; Bockisch, A.; Debatin, J. F.; Ruehm, S. G., Whole-body dual-modality PET/CT and whole-body MRI for tumor staging in oncology. *Jama* **2003,** *290* (24), 3199-3206.

22. Boss, A.; Bisdas, S.; Kolb, A.; Hofmann, M.; Ernemann, U.; Claussen, C. D.; Pfannenberg, C.; Pichler, B. J.; Reimold, M.; Stegger, L., Hybrid PET/MRI of intracranial masses: initial experiences and comparison to PET/CT. *Journal of Nuclear Medicine* **2010,** *51* (8), 1198-1205.

23. Yang, Y.; He, M. Z.; Li, T.; Yang, X., MRI combined with PET-CT of different tracers to improve the accuracy of glioma diagnosis: a systematic review and meta-analysis. *Neurosurg. Rev.* **2017**.

24. Clark, P. A.; Bhattacharya, S.; Elmayan, A.; Darjatmoko, S. R.; Thuro, B. A.; Yan, M. B.; van Ginkel, P. R.; Polans, A. S.; Kuo, J. S., Resveratrol targeting of AKT and p53 in glioblastoma and glioblastoma stem-like cells to suppress growth and infiltration. *J. Neurosurg.* **2017,** *126* (5), 1448-1460.

25. Clark, P. A.; Iida, M.; Treisman, D. M.; Kalluri, H.; Ezhilan, S.; Zorniak, M.; Wheeler, D. L.; Kuo, J. S., Activation of multiple ERBB family receptors mediates glioblastoma cancer stem-like cell resistance to EGFR-targeted inhibition. *Neoplasia* **2012,** *14* (5), 420-8.

26. Zorniak, M.; Clark, P. A.; Leeper, H. E.; Tipping, M. D.; Francis, D. M.; Kozak, K. R.; Salamat, M. S.; Kuo, J. S., Differential expression of 2',3'-cyclic-nucleotide 3'-phosphodiesterase and neural lineage markers correlate with glioblastoma xenograft infiltration and patient survival. *Clin. Cancer Res.* **2012,** *18* (13), 3628-36.

27. Malin, D.; Strekalova, E.; Petrovic, V.; Deal, A. M.; Al Ahmad, A.; Adamo, B.; Miller, C. R.; Ugolkov, A.; Livasy, C.; Fritchie, K.; Hamilton, E.; Blackwell, K.; Geradts, J.; Ewend, M.; Carey, L.; Shusta, E. V.; Anders, C. K.; Cryns, V. L., alphaB-crystallin: a novel regulator of breast cancer metastasis to the brain. *Clin. Cancer Res.* **2014,** *20* (1), 56-67.

28. Yarnykh, V. L., Actual flip-angle imaging in the pulsed steady state: a method for rapid three-dimensional mapping of the transmitted radiofrequency field. *Magnetic resonance in medicine* **2007,** *57* (1), 192-200.

29. Robichaud, G.; Garrard, K. P.; Barry, J. A.; Muddiman, D. C., MSiReader: An Open-Source Interface to View and Analyze High Resolving Power MS Imaging Files on Matlab Platform. *Journal of the American Society for Mass Spectrometry* **2013,** *24* (5), 718-721.

30. Moore, D. A., Selective Trialkylation of Cyclen with tert‐Butyl Bromoacetate. *Organic Syntheses* **2008**, 10-14.

31. El‐Faham, A.; Albericio, F., COMU: A third generation of uronium‐type coupling reagents. *Journal of Peptide Science* **2010,** *16* (1), 6-9.

32. Gajate, C.; Mollinedo, F., Edelfosine and perifosine induce selective apoptosis in multiple myeloma by recruitment of death receptors and downstream signaling molecules into lipid rafts. *Blood* **2007,** *109* (2), 711-719.

33. Gajate, C.; Mollinedo, F., The antitumor ether lipid ET-18-OCH3 induces apoptosis through translocation and capping of Fas/CD95 into membrane rafts in human leukemic cells. *Blood* **2001,** *98* (13), 3860-3863.

34. Malin, D.; Strekalova, E.; Petrovic, V.; Deal, A. M.; Ahmad, A. A.; Adamo, B.; Miller, C. R.; Ugolkov, A.; Livasy, C.; Fritchie, K.; Hamilton, E.; Blackwell, K.; Geradts, J.; Ewend, M.; Carey, L.; Shusta, E. V.; Anders, C. K.; Cryns, V. L., αB-crystallin: a Novel Regulator of Breast Cancer Metastasis to the Brain. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2014,** *20* (1), 56-67.

35. Cornett, D. S.; Reyzer, M. L.; Chaurand, P.; Caprioli, R. M., MALDI imaging mass spectrometry: molecular snapshots of biochemical systems. *Nature methods* **2007,** *4* (10), 828.

36. Tweedle, M. F., Gadolinium deposition: Is it chelated or dissociated gadolinium? How can we tell? *Magnetic resonance imaging* **2016,** *34* (10), 1377-1382.

37. Frenzel, T.; Lengsfeld, P.; Schirmer, H.; Hütter, J.; Weinmann, H.-J., Stability of gadolinium-based magnetic resonance imaging contrast agents in human serum at 37 C. *Investigative radiology* **2008,** *43* (12), 817-828.

38. Shen, Y.; Goerner, F. L.; Snyder, C.; Morelli, J. N.; Hao, D.; Hu, D.; Li, X.; Runge, V. M., T1 relaxivities of gadolinium-based magnetic resonance contrast agents in human whole blood at 1.5, 3, and 7 T. *Investigative radiology* **2015,** *50* (5), 330-338.

39. Colosimo, C.; Knopp, M.; Barreau, X.; Gerardin, E.; Kirchin, M.; Guezenoc, F.; Lodemann, K. P., A comparison of Gd-BOPTA and Gd-DOTA for contrast-enhanced MRI of intracranial tumours. *Neuroradiology* **2004,** *46* (8), 655-665.

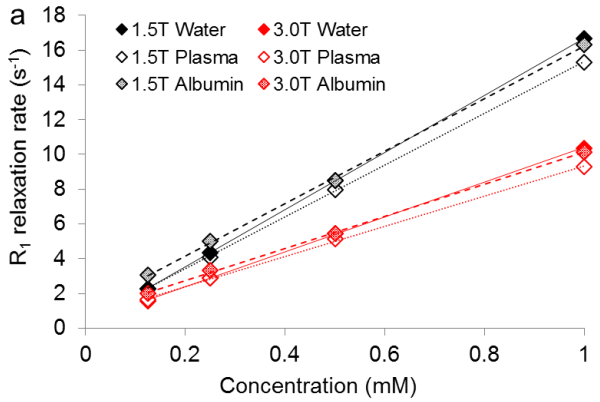
40. Rumboldt, Z.; Rowley, H. A.; Steinberg, F.; Maldjian, J. A.; Ruscalleda, J.; Gustafsson, L.; Bastianello, S., Multicenter, double‐blind, randomized, intra‐individual crossover comparison of gadobenate dimeglumine and gadopentetate dimeglumine in MRI of brain tumors at 3 tesla. *Journal of Magnetic Resonance Imaging* **2009,** *29* (4), 760-767.

41. Maravilla, K. R.; Smith, M.; Vymazal, J.; Goyal, M.; Herman, M.; Baima, J.; Babbel, R.; Vaneckova, M.; Žižka, J.; Colosimo, C., Are there differences between macrocyclic gadolinium contrast agents for brain tumor imaging? Results of a multicenter intraindividual crossover comparison of gadobutrol with gadoteridol (the TRUTH study). *American Journal of Neuroradiology* **2015,** *36* (1), 14-23.

42. McDonald, R. J.; McDonald, J. S.; Kallmes, D. F.; Jentoft, M. E.; Murray, D. L.; Thielen, K. R.; Williamson, E. E.; Eckel, L. J., Intracranial gadolinium deposition after contrast-enhanced MR imaging. *Radiology* **2015,** *275* (3), 772-782.

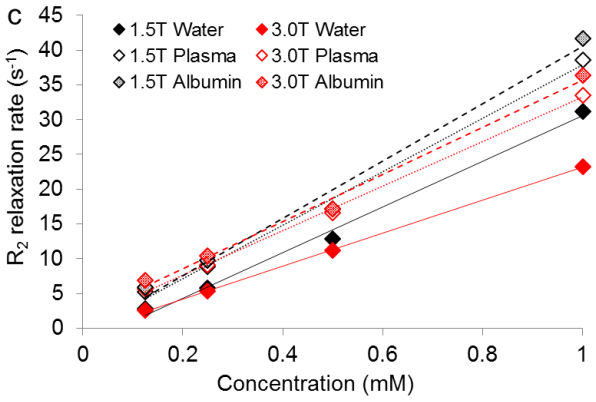
43. Kanda, T.; Ishii, K.; Kawaguchi, H.; Kitajima, K.; Takenaka, D., High signal intensity in the dentate nucleus and globus pallidus on unenhanced T1-weighted MR images: relationship with increasing cumulative dose of a gadolinium-based contrast material. *Radiology* **2013,** *270* (3), 834-841.

## Figures




b

| $r_1$ ($s^{-1}$/mM) | 1.5T | 3.0T |
|---|---|---|
| Gd-NM600 in Water | 16.4 | 10.0 |
| Gd-NM600 in Albumin | 15.1 | 9.2 |
| Gd-NM600 in Plasma | 14.9 | 8.7 |




d

| $r_2$ ($s^{-1}$/mM) | 1.5T | 3.0T |
|---|---|---|
| Gd-NM600 in Water | 32.8 | 23.7 |
| Gd-NM600 in Albumin | 41.2 | 33.2 |
| Gd-NM600 in Plasma | 38.4 | 31.1 |

**Figure 1.** Longitudinal relaxivity ($r_1$) and transverse relaxivity ($r_2$) of Gd-NM600. **(a, c)** Relaxivity was determined by linear regression of relaxation rate ($R_1$ or $R_2$) versus sample concentration at each field strength and for each solvent. **(b)** $r_1$ values at 1.5T and 3.0T in water, albumin, and plasma. Gd-NM600 was found to have greater $r_1$ values at 1.5T than at 3.0T, and when prepared in water compared to plasma and albumin. The longitudinal relaxivity in water (16.4 s-1/mM) compares favorably against Gd-BOPTA (MultiHance), with the highest relaxivity (6.3-7.9 s-1/mM) of the commercial agents. **(d)** $r_2$ values at 1.5T and 3.0T in water, albumin, and plasma. Gd-NM600 was found to have greater $r_2$ values when prepared in plasma and albumin compared to water**.**

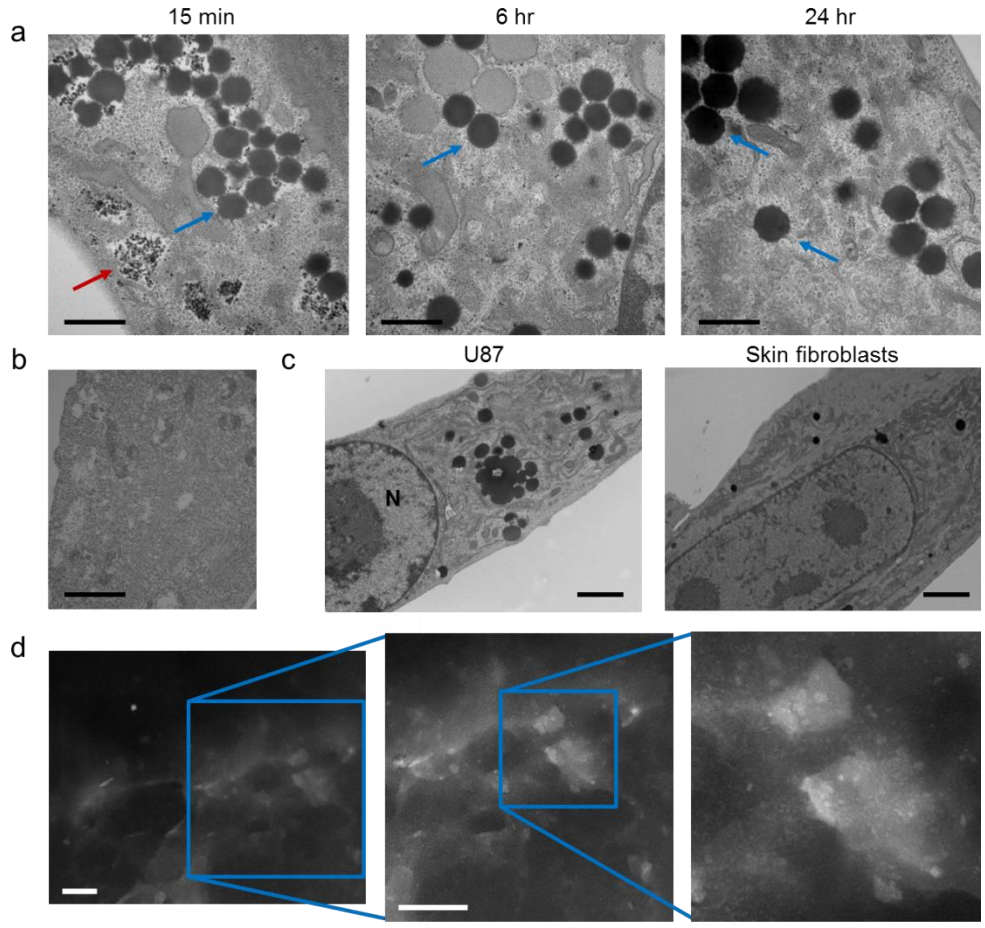



**Figure 2.** Uptake of Gd-NM600 in U87 glioblastoma cells and SK fibroblasts *in vitro*. **(a)** Transmission electron microscopy images of a U87 cell treated with 1µM Gd-NM600. Accumulation in early and mature endosomes (red/blue arrows) leads to electron dense spheres corresponding to intracellular Gd-NM600. Scale bar = 1µm. **(b)** Untreated control U87 cell shows no electron dense regions. Scale bar = 1µm. **(c)** Uptake of Gd-DOTA-APC in U87 after 24 hour incubation compared to minimal uptake in SK fibroblasts at the same time point. No nuclear uptake is observed (N denotes nucleus). Scale bar = 2µm. **(d)** Scanning transmission electron microscopy of U87 cells treated with 1µM Gd-NM600 confirms the presence of the heavy metal Gd within cells. Scale bar = 200nm.

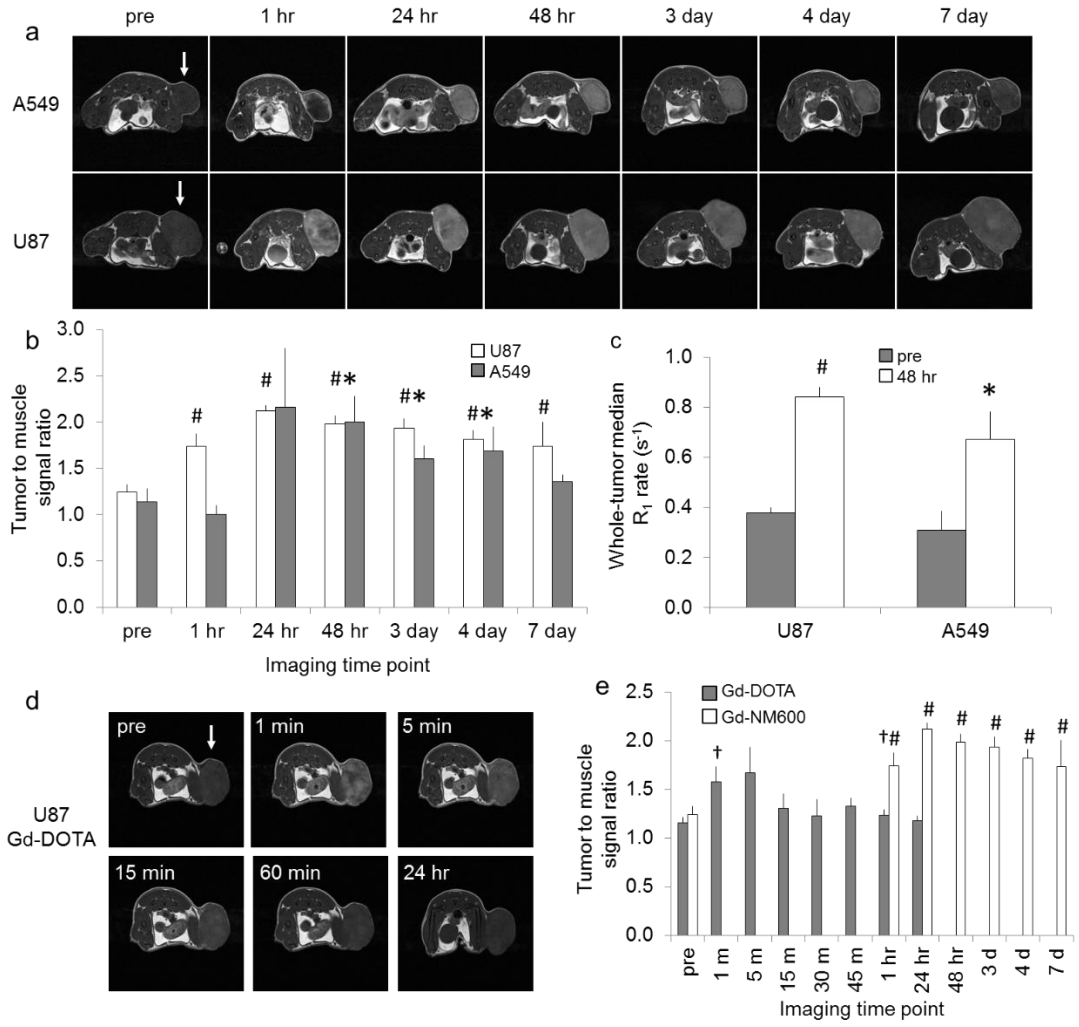



**Figure 3.** Tumor uptake of Gd-NM600 in U87 and A549 and comparative uptake dynamics of Gd-DOTA. **(a)** Representative $T_1$-weighted images of two tumor models pre-contrast and at multiple time points up to 7 days following delivery of Gd-NM600. Tumor location indicated by white arrow. **(b)** Tumor to muscle ratio of $T_1$-weighted signal increased over the course of 24-48 hours and remained significantly enhanced compared to pre-contrast for up to 4 and 7 days for A549 and U87 xenografts, respectively. Error bars reflect standard deviation among three subjects (*$p$<0.05 compared to pre-contrast, A549; $^{\#}p$<0.05 compared to pre-contrast, U87). **(c)** Quantification of $R_1$ relaxation rate ($R_1$=1/$T_1$) revealed a significant increase of greater than 2x in whole-tumor median $R_1$ rate in both tumor models 48 hours after contrast administration. **(d)** Representative T1-weighted images of Gd-DOTA uptake in a single subject with flank U87 xenograft model. **(e)** Time course of $T_1$-weighted signal enhancement after delivery of Gd-NM600 and Gd-DOTA in U87 (N=3 for each model) shows that Gd-NM600 enhancement is greater and more prolonged, indicating that uptake reflects specific targeting and incorporation of the contrast agent in cancer cells. (†$p$<0.05 compared to pre-contrast, Gd-DOTA; $^{\#}p$<0.05 compared to pre-contrast, Gd-NM600)

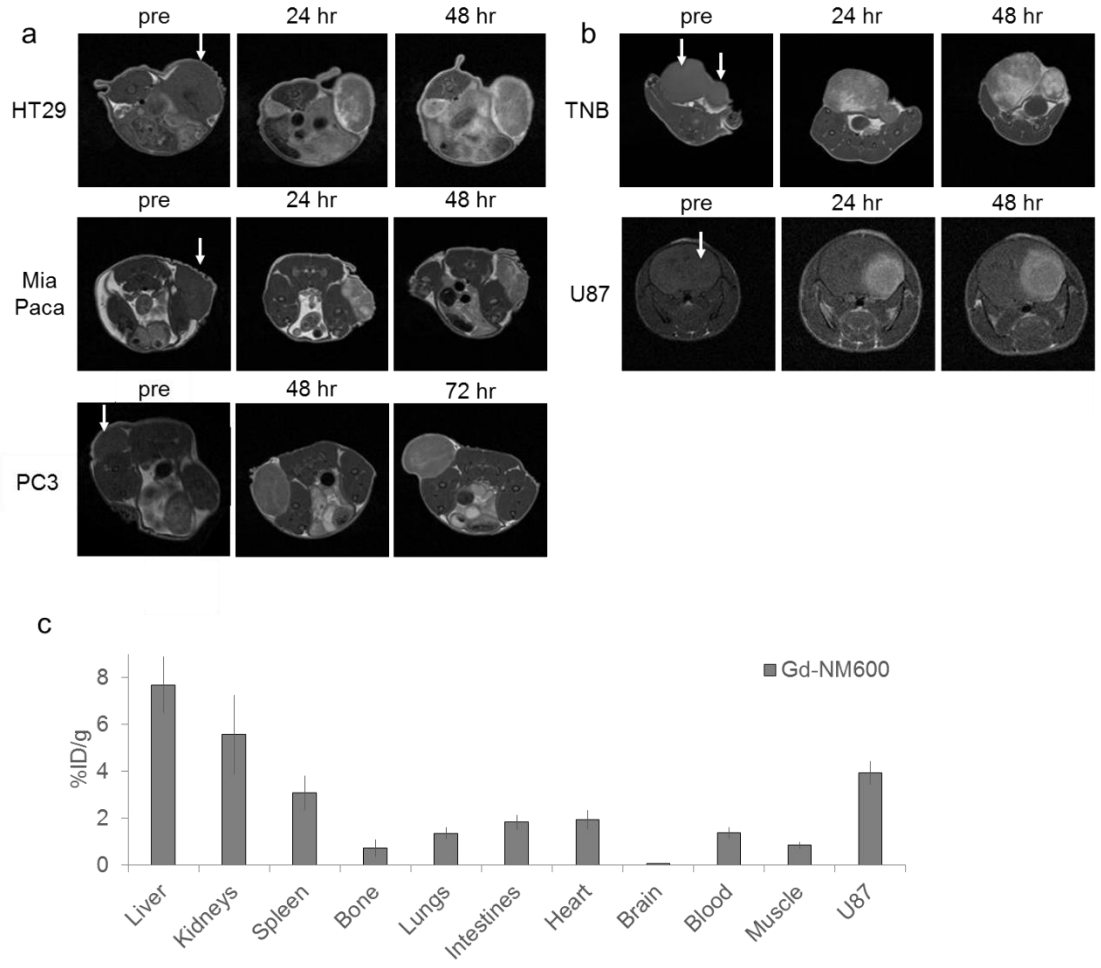



**Figure 4.** Tumor-specific uptake of Gd-NM600 in multiple human cancer models. **(a)** Representative axial $T_1$-weighted images of three flank xenograft models including HT29 (colorectal), MiaPaca (pancreatic), and PC3 (prostate) at various time points pre-contrast and up to three days post-contrast. White arrows indicate the location of the cancer. **(b)** $T_1$-weighted images of two orthotopic tumor models, a triple-negative breast cancer (TNB) implanted in the mammary fat pad and a U87 glioblastoma in the brain. Uptake was observed at 24 and 48 hours, and U87 tumors were confirmed with H&E staining. **(c)** Perfused tissues from nude athymic mice bearing U87 flank xenografts (N=3) were collected 48 hours following Gd-NM600 administration. Gd uptake was measured with high-resolution (magnetic-sector) inductively-coupled plasma mass spectrometry (ICP-MS) following acid digestion. Other than the organs of clearance, the tumor had the highest uptake of any organ.

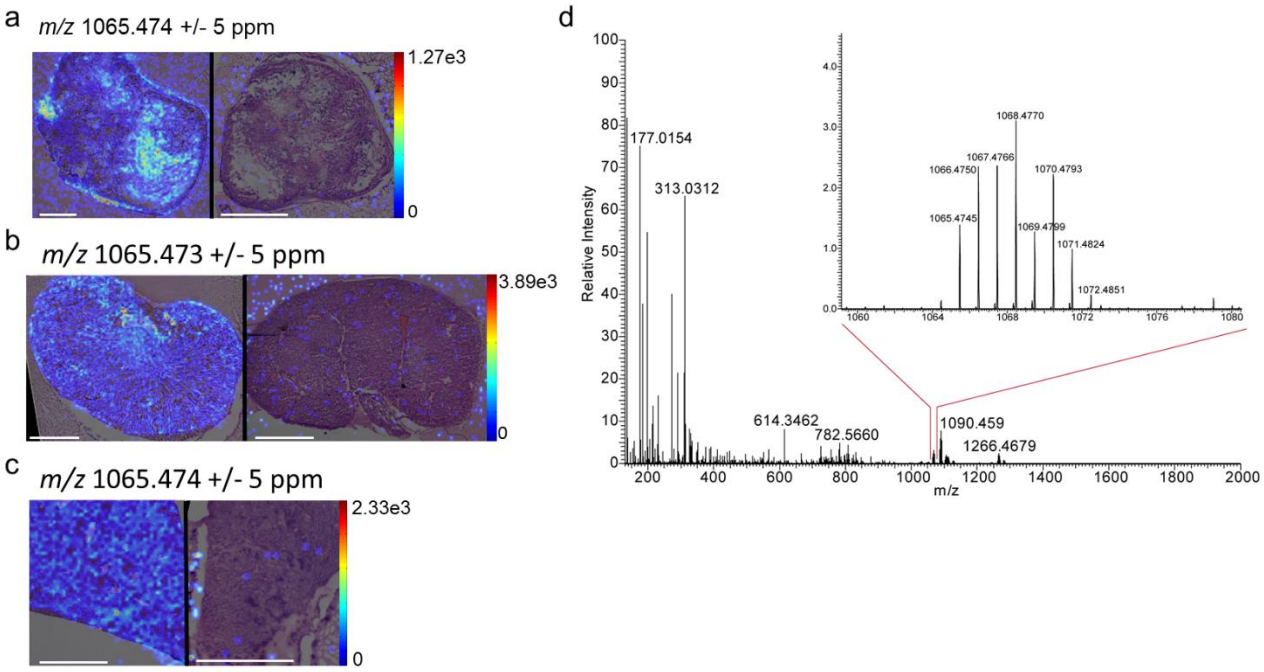



**Figure 5.** MALDI-MSI images of Gd-NM600 in mice bearing A549 flank xenografts. MALDI-MSI and histological overlays of Gd-NM600 in whole A549 tumor **(a)**, whole kidney **(b)**, and liver tissue **(c)** in an animal injected with Gd-NM600 (*left*), and an animal injected with vehicle (*right*). The white bar indicates 2 mm for **(a)** and **(b)** and 1 mm for **(c)**. The color bar maximum was set at 50% of the maximum normalized intensity of all three biological replicates (only biological replicate 2 is shown here). **(d)** An averaged mass spectra from the tumor region is shown, with the Gd peak corresponding to the mass peaks of the compound Gd-NM600 highlighted. The isotopic distribution of the Gd provides a specific Gd mass signature.

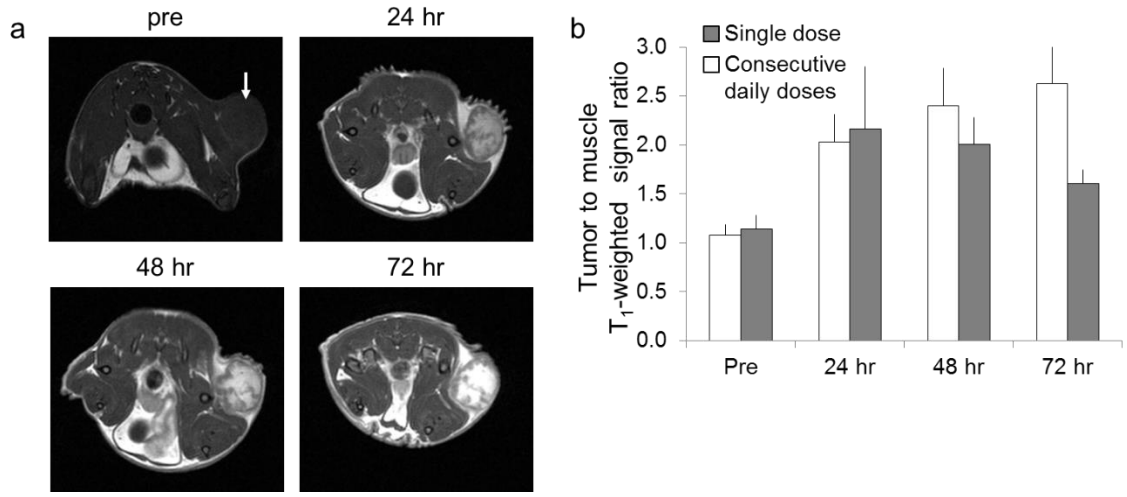



**Figure 6.** Continuous uptake of multiple Gd-DOTA-APC doses. **(a)** In a representative subject bearing an A549 flank xenograft, non-saturating tumor enhancement after consecutive Gd-NM600 administration (i.v. after pre-contrast imaging and after imaging at 24 and 48 hours) was observed over the course of three days on T1-weighted axial imaging. **(b)** Tumor to muscle T1-weighted signal ratio continually increased in subjects delivered consecutive daily doses of Gd-NM600 (N=6), while contrast enhancement slowly decreased after 24 hours in subjects administered a single dose (N=3). White arrow highlights location of the tumor.

## Supplemental Information

## Supplemental Experimental Methods

Chemical synthesis

All chemicals were purchased from Aldrich Chemical Co. (Milwaukee, WI) except for cyclene, which was from Strem Chemicals (Newburyport, MA). Thin-layer chromatography was performed using DC-Alufolien Kieselgel 60 $F_{254}$ plates (EMD, Germany). Visualization was achieved by UV light (254 nm) and ceric-molybdenum stain. NMR data were collected on a Varian Unity-Inova 500 MHz NMR Spectrometer. Chemical shifts are reported in parts per million (ppm) relative to tetramethylsilane (TMS), and spin multiplicities are given as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), or br m (broad multiplet). High-resolution mass spectra were obtained on a Bruker MaXis Ultra-High Resolution Quadrupole Time-of-Flight MS.

*1,4,7-Tris(benzyloxycarbonylmethyl)-1,4,7,10-tetraazacyclododecane hydrobromide (DO3A tri-benzyl ester) (Compound **2**)*

Cyclen **1** (2.00 g, 11.6 mmol) and sodium acetate trihydrate (5.22 g, 38.3 mmol) were suspended in *N,N*-dimethylacetamide (24 mL) and cooled to -15°C in a methanol-ice bath. A solution of benzyl bromoacetate (6 mL, 38.3 mmol) in *N,N*-dimethylacetamide (12 mL) was added via cannula within 10 min. Upon completion of addition, the cooling bath was removed and the mixture was stirred at RT for 72 h. Water (120 ml) was added to the reaction mixture. Solid potassium bicarbonate (6 g, 60 mmol) was added portion wise with stirring, and the product separated as an oil. The reaction mixture was placed in an ice bath for 1 h, during which time the oily precipitate solidified. The clear solution was decanted before the solid material was dissolved in ethyl acetate, transferred to a separatory funnel and washed with aqueous ammonium bromide

solution and water. The extract was dried over $Na_2SO_4$ and concentrated. The compound was purified by column chromatography on silica gel. Sequential elution of the column with $CHCl_3$-hexane (1:1), $CHCl_3$-hexane (2:1), $CHCl_3$, $CHCl_3$-MeOH (97:3), and finally with $CHCl_3$-MeOH (97:3) afforded pure compound **2** as a thick yellow oil. Yield: 4.866 g (60%). $^1$H-NMR (500 MHz, $CDCl_3$): 7.39-7.32 (m, 15H, Ar*H*$_5$), 5.13 (s, 6H, OC*H*$_2$Ar), 3.49 (s, 4H NC*H*$_2$CO$_2$), 3.42 (s, 2H NC*H*$_2$CO$_2$), 3.09 (br s, 4H, (cyclene C*H*$_2$), 2.92-2.80 (m, 12H, cyclene C*H*$_2$). HRMS: calculated for $C_{35}H_{45}N_4O_6$ $(M+H)^+$ 617.33336, found 617.33427, calculated for $C_{35}H_{44}N_4O_6Na$ $(M+Na)^+$ 639.315306, found 639.31488.

*1,4,7-Tris(benzyloxycarbonylmethyl)-7-(t-butyoxycarbonylmethyl)-1,4,7,10-tetraazacyclododecane (DOTA t-butyl tri-benzyl ester) (**3**)*

DO3A tri-benzyl ester hydrobromide **2** (4.82 g, 6.90 mmol) was dissolved in acetonitrile (70 mL). *tert*-Butyl-bromoacetate (1.21 mL, 8.28 mmol) and powdered potassium carbonate (2.38 g, 17.3 mmol) were added to the solution. The reaction mixture was stirred at room temperature for 24 h and then filtered. The filtrate was concentrated and the residue was purified on a silica gel column eluted with a step gradient of $CHCl_3$-MeOH (99:1), (98:2), (95:5) and (90:10) to provide compound **3** as a hydrobromide (thick yellow oil). Yield: 5.28 g (94%). This compound was dissolved in chloroform (250 ml), transferred to a separatory funnel and washed with aqueous potassium carbonate (5 g) in water (150 ml) and then water (150 ml). The chloroform layer was dried over $Na_2SO_4$ and concentrated to yield compound **3** as a free base: 4.35 g (86%). $^1$H-NMR for hydrobromide of **3** (500 MHz, $CDCl_3$): 7.38-7.30 (m, 15H, Ar*H*$_5$), 5.14 (s, 4H, OC*H*$_2$Ar), 5.12 (s, 2H, OC*H*$_2$Ar), 3.92 (s, 2H NC*H*$_2$CO$_2$), 3.76 (br s, 4H NC*H*$_2$CO$_2$), 3.71 (s, 2H NC*H*$_2$CO$_2$), 3.40 (br s, 4H, cyclene C*H*$_2$) , 3.16-3.00 (m, 12H, cyclene C*H*$_2$), 1.45 (s, 9H, O(C*H*$_3$)$_3$). HRMS:

calculated for $C_{41}H_{55}N_4O_8$ $(M+H)^+$ 731.40144, found 731.40000, calculated for $C_{41}H_{54}N_4O_8Na$ $(M+Na)^+$ 753.38339, found 753.38284.

*1,4,7-Tris(benzyloxycarbonylmethyl)-7-(carboxymethyl)-1,4,7,10- tetraazacyclododecane hydrochloride (DOTA tri-benzyl ester) (**4**)*

DOTA *t*-butyl tri-benzyl ester **3** (4.33 g, 5.92 mmol) was dissolved in ethanol-free chloroform (28 mL), followed by addition of 4M HCl in dioxane (10 mL, 40 mmol). The mixture was stirred for 24 h at room temperature. A white precipitate formed as the reaction proceeded. Chloroform (25 mL) and hexane (25 mL) were added with stirring. The mixture was filtered and the solid white material was collected on a filter, washed with hexane and dried under high vacuum to constant weight. Yield: 4.15 g (98%). $^1$H-NMR (500 MHz, DMSO-d6): 7.40-7.34 (m, 15H, Ar*H*$_5$), 5.17 (s, 2H, OC*H*$_2$Ar), 5.11 (br s, 4H, OC*H*$_2$Ar), 4.30-3.80 (m, 14H, NC*H*$_2$CO$_2$Bn and cyclene C*H*$_2$), 3.55-3.10 (m, 10H, cyclene C*H*$_2$). HRMS: calculated for $C_{37}H_{47}N_4O_8$ $(M+H)^+$ 675.33884, found 675.33713, calculated for $C_{37}H_{46}N_4O_8Na$ $(M+Na)^+$ 697.32079, found 697.31994.

*DOTA tri-benzyl ester - 18-(p-aminophenyl)octadecyl phosphocholine conjugate (**6**)*

DOTA tri-benzyl ester (1.48 g, 2.08 mmol) and 18-(*p*-aminophenyl)octadecyl phosphocholine **5** (914 mg, 1.74 mmol) were suspended in ethanol-free chloroform (50 mL)(13). Triethyl amine (866 µL, 6.25 mmol) and the coupling agent COMU (891 mg, 2.08 mmol) were added to the reaction mixture. The reaction mixture was stirred at room temperature for 24 h by which time the reaction mixture became a clear yellow-orange solution. The following solvents were added to the reaction mixture: chloroform (30 mL), methanol (80 mL) and water (70 mL). The mixture was vigorously stirred for 5 min and transferred to a separatory funnel. Chloroform layer was removed, and extraction was repeated (2 × 80 mL of chloroform). The combined extracts

were dried over $Na_2SO_4$, filtered and concentrated *in vacuo*. The product was purified by column chromatography on silica gel eluted sequentially with $CHCl_3$-MeOH (9:1), (8:2), (5:5) and then $CHCl_3$-MeOH-$H_2O$ (65:25:3), (65:25:4). After removal of the solvent in vacuo, the purified compound **6** was dried under high vacuum to constant weight. Yield: 1.28 g (60%) of yellow amorphous solid. $^1$H-NMR (500 MHz, $CDCl_3$-$CD_3OD$ 1:1) 7.51 (d, *J* = 8.5 Hz, 2H, Ar*H*), 7.37-7.28 (m, 15H, benzyl Ar*H*), 7.02 (d, *J* = 8.5 Hz, 2H, Ar*H*), 5.28-5.12 (m, 6H, OC$H_2$Ar), 4.70-4.10 (m, 2H, $PO_4CH_2CH_2NMe_3$), 3.87 (q, *J* = 6.4 Hz, 2H, $(CH_2)_{17}CH_2PO_4$), 3.62-3.59 (m, 2H, $PO_4CH_2CH_2NMe_3$), 3.50-2.80 (br m, 24H, $NCH_2CO_2$ and cyclene $CH_2$), 3.23 (s, 9H, $N(CH_3)_3$), 2.54 (t, *J* = 7.5 Hz, 2H, $ArCH_2(CH_2)_{17}$), 1.66-1.60 (m, 2H, $(ArCH_2CH_2(CH_2)_{16}$), 1.60-1.52 (m, 2H, $(CH_2)_{16}CH_2CH_2PO_4$), 1.40-1.24 (br m, 28H, $CH_2(CH_2)_{14}CH_2$). HRMS: calculated for $C_{66}H_{100}N_6O_{11}P$ $(M+H)^+$ 1183.71822, found 1183.71879; calculated for $C_{66}H_{99}N_6O_{11}PNa$ $(M+Na)^+$ 1205.70017, found 1205.70204.

*DOTA - 18-(p-aminophenyl)octadecyl phosphocholine conjugate (7)*

Compound **6** (1.22 g, 1.03 mmol) was dissolved in 25 mL of EtOH and transferred into the Parr hydrogenation bottle. Palladium on carbon (10 wt. %) (280 mg) was added to the bottle, and the mixture was hydrogenated at 52 psi of hydrogen with shaking for 24hr at RT. Completion of the reaction was determined by TLC analysis. The mixture was filtered through 0.45 micron PTFE syringe filter. The filtrate was concentrated *in vacuo*, and the product was precipitated by the addition of acetone. Precipitated compound **6** was dried under high vacuum to constant weight. Yield: 613 mg (65%) of amorphous white solid. $^1$H-NMR (500 MHz, $CDCl_3$-$CD_3OD$ 1:1) 7.46 (d, *J* = 8.5 Hz, 2H, Ar*H*), 7.12 (d, *J* = 8.5 Hz, 2H, Ar*H*), 4.30-4.22 (m, 2H, $PO_4CH_2CH_2NMe_3$), 3.92 (q, *J* = 6.4 Hz, 2H, $(CH_2)_{17}CH_2PO_4$), 3.56 (m, 2H, $PO_4CH_2CH_2NMe_3$), 4.00-3.00 (br m, 24H, $NCH_2CO_2$ and cyclene $CH_2$), (3.49 (s, 4H $NCH_2CO_2$), 3.15 (s, 9H, $N(CH_3)_3$), 2.57 (t, *J* = 7.5 Hz,

2H, ArC*H*$_2$(CH$_2$)$_{17}$), 1.68-1.54 (br m, 4H, ArCH$_2$C*H*$_2$(CH$_2$)$_{16}$ and C*H*$_2$CH$_2$OP), 1.40-1.25 (m, 28H, CH$_2$(C*H*$_2$)$_{14}$CH$_2$). HRMS: calculated for $C_{45}H_{81}N_6O_{11}P+H^+$ ($M^+$+H) 913.57737, found 913.57710, calculated for $C_{45}H_{81}N_6O_{11}P+Na^+$ ($M^+$+Na) 935.55932, found 935.55795, calculated for $[C_{45}H_{81}N_6O_{11}P+2H]^{2+}$ ([M+2H]$^{2+}$/2) 457.29232, found 457.29208.

*Gd-DOTA - 18-(p-aminophenyl)octadecyl phosphocholine conjugate (**8**)*

Compound **7** (513 mg, 0.562 mmol) was dissolved in the mixture of pyridine (5 mL) and methanol (7 mL) before a solution of gadolinium (III) chloride hexahydrate (250 mg, 0.672 mmol) in water (0.7 mL) was added to the reaction mixture. A white precipitate formed immediately. The mixture was slightly warmed with a heat gun while stirring to dissolve the precipitate. The slightly cloudy mixture was stirred at room temperature for 24 h. The reaction mixture was concentrated on the rotary evaporator, and the residue was kept under high vacuum for 2 h to remove residual solvents. The residue was purified by column chromatography on silica gel, eluted with CHCl$_3$-MeOH (8:2), (5:5), (10:8) and then with CHCl$_3$-MeOH-conc. NH$_4$OH (10:8:2). Fractions containing **8** were combined and concentrated *in vacuo* before the residue was dried under high vacuum to constant weight. Purified compound **8** was dissolved in CHCl$_3$-MeOH (2 mL), and acetone (50 mL) was added with stirring. The mixture was evaporated to yield a white powder. Yield: 450 mg (75%). HRMS: calculated for $C_{45}H_{78}GdN_6O_{11}P+H^+$ ($M^+$+H) 1068.47912, found 1068.47533, calculated for $C_{45}H_{78}GdN_6O_{11}P+Na^+$ ($M^+$+Na) 1090.46106, found 1090.45729, calculated for $C_{45}H_{78}GdN_6O_{11}P+K^+$ ($M^+$+K) 1106.43479, found 1106.43135.

Orthotopic Rat Model Preparation and Simultaneous PET/MR Imaging

An RNU nude athymic rat bearing a flank glioblastoma (U87, n=1). NM600 was chelated with both Gd and $^{64}$Cu, and both were injected via tail vein into the rat at 0.12g/kg=0.11mmol/kg

and 5mCi (20mCi/kg) respectively. Simultaneous PET/MRI was acquire using the GE Signa PET/MR (Ge Healthcare, Waukesha, WI).

Tissue Biodistribution

U87 flank-bearing mice (n=3) were injected with 0.12g/kg of Gd-NM600. Seventy-two hours following Gd-NM600 administration, animals were humanely euthanized and perfused with saline before organs were harvested. Tissue samples were sectioned and homogenized using ceramic blades and Teflon implements, and 20-25mg of homogenized wet tissue was transferred to a Teflon digestion vessel. Organic matrix removal and complete solubilization of gadolinium and other trace major elements in the tissues was performed by microwave-assisted acid digestion in Teflon digestion-bombs. Gadolinium along with eight additional elements in the tissue digests were quantified using high-resolution (magnetic-sector) inductively-coupled plasma mass spectrometry (ICP-MS). For comparison, $^{64}$Cu-DOTA-APC biodistribution was also measured in U87 flank-bearing mice (n=3). A bolus injection of 2.5 nmol (approximately 250 μCi) was intravenously delivered to each of three flank U87-bearing mice. Forty-eight hours later, animals were humanely euthanized and organ biodistribution was measured by gamma-counting on a PerkinElmer Wizard 2480 (Waltham, MA).

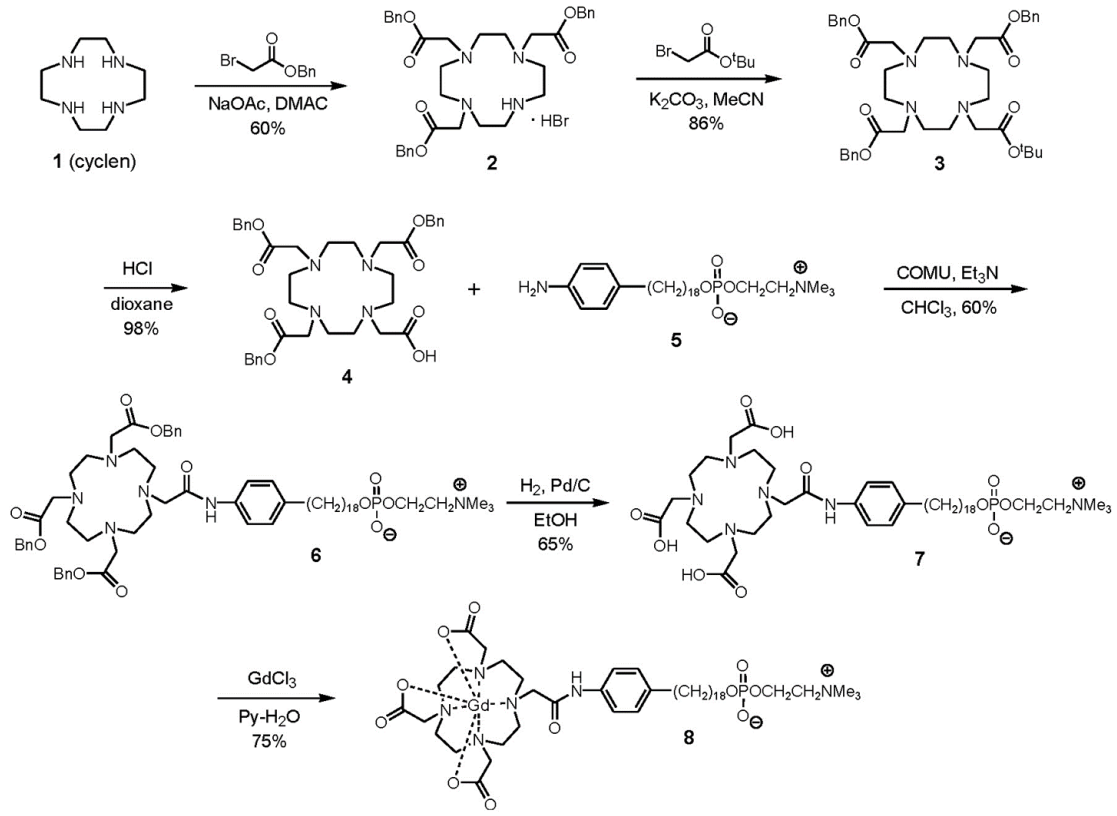
Transmission Electron Microscopy and Surface Transmission Electron Microscopy

To observe uptake of Gd-NM600 in cancer and control cells with transmission electron microscopy (TEM), SK fibroblasts and U87 cells were plated on coverslips in DMEM, 10% FBS, and 1% P/S overnight. Tumor cells were untreated or treated for 0.25hr, 0.5hr, 1hr, 6hr, or 24hr with 1μM of Gd-NM600, and then fixed in 2.5% glutaraldehyde, 2.0% paraformaldehyde in 0.1M sodium phosphate buffer (PB), pH 7.4 for ~1 hrs. SK fibroblast samples were prepared in the same way after 24hr treatment with 1μM of Gd-NM600. Cells were then rinsed and then post-fixed in

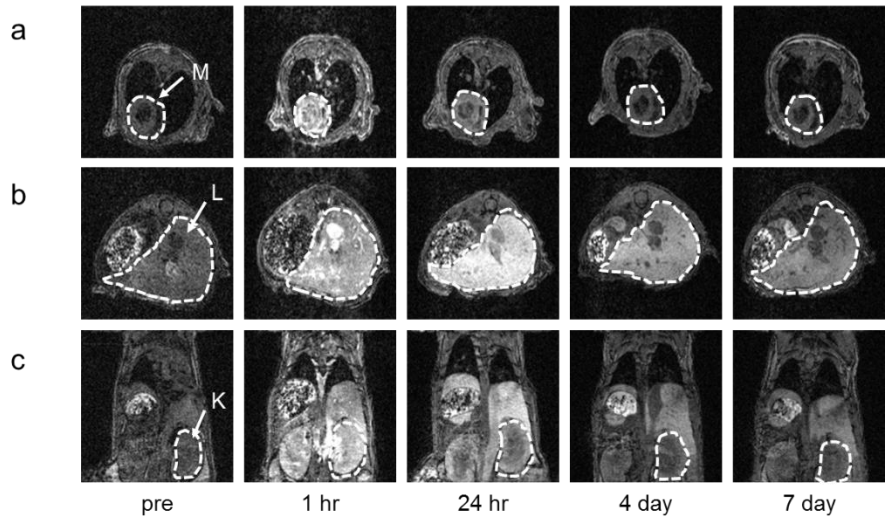
1% osmium tetroxide, 1% potassium ferrocyanide in PB for ~1hr. After fixation, cells were dehydrated in a graded ethanol series (35, 50, 70, 80, 90% for 5 minutes, 95% for 10 minutes, 100% for 3 x 10 minutes) then transitioned in propylene oxide (PO) 2 x 7 minutes. Durcupan ACM resin (Fluka AG, Switzerland) was used during infiltration and embedding. Increasing concentrations of accelerated Durcupan (10ml A/M, 10ml B, 300ul C, 100ul D components) were used for infiltration. The cultures were embedded in open aluminum weighing dishes at 60$^{o}$C in a drying oven overnight until polymerized. The samples were then treated with concentrated hydrofluoric acid, glass side down, for approximately 15 minutes to etch off the glass, revealing the embedded cell culture. Five mm x 5 mm sample pieces were sawed out and glued to a prepared blank using a Leica EM UC6 ultra microtome. Ultrathin sections (100 nm) were prepared on 300 mesh Cu thin-bar grids and contrasted with Reynolds lead citrate and 8% uranyl acetate in 50% EtOH. The sections were observed with a Philips CM120 transmission electron microscope, and images were collected with an AMT BioSprint12 (Advanced Microscopy Techniques Corp., Woburn, MA) side mounted digital camera.

For scanning transmission electron microscopy (STEM), U87 cells were prepared in the same fashion, with the exception of treatment with osmium tetroxide to eliminate potential heavy metal contamination. STEM experiments were performed on a Technai F30 (Fei Company, Hillsboro, Oregon) with a CEOS probe aberration-corrector operated at 200 keV, as previously reported(27). STEM images were collected with a resolution of 0.8 A˚ using a 24.5 mrad probe semi-angle and 25 pA probe current. High angle annular dark field (HAADF) images were acquired with a detector subtending 54 to 270 mrad.

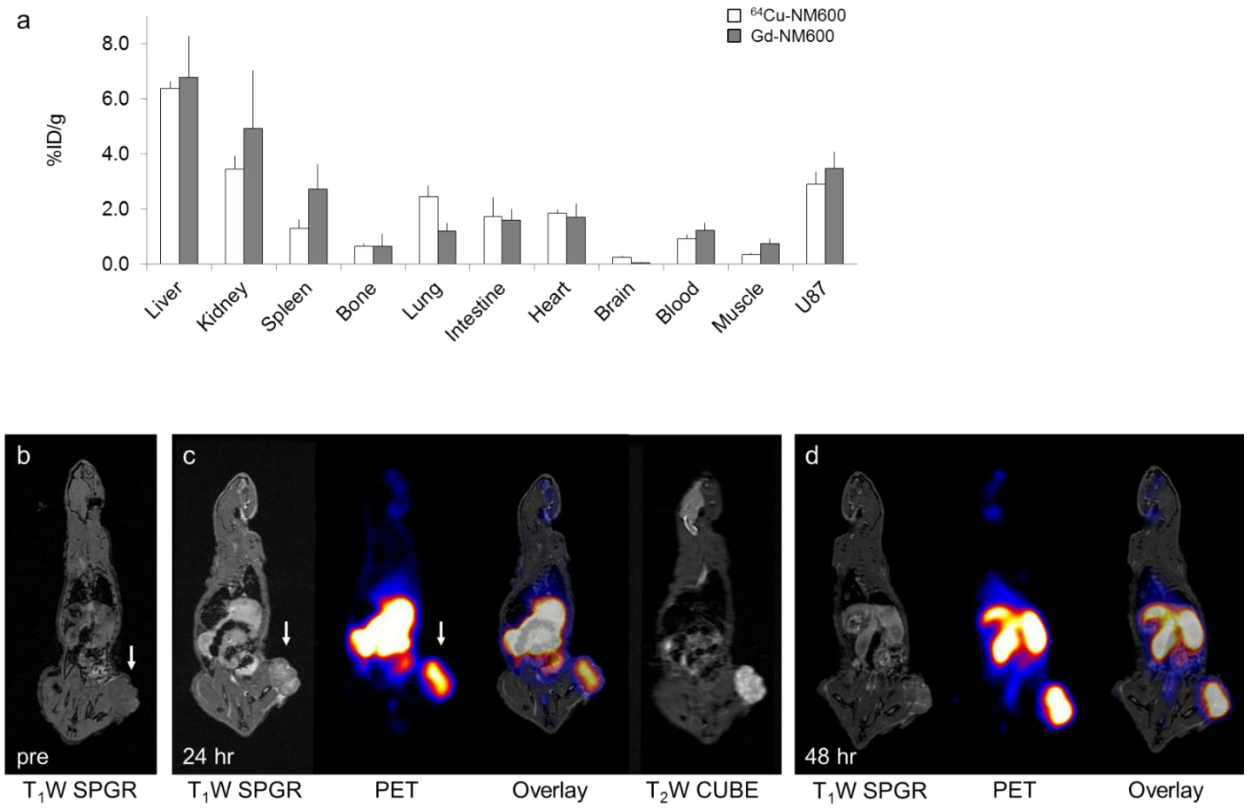
**Supplemental Figures**



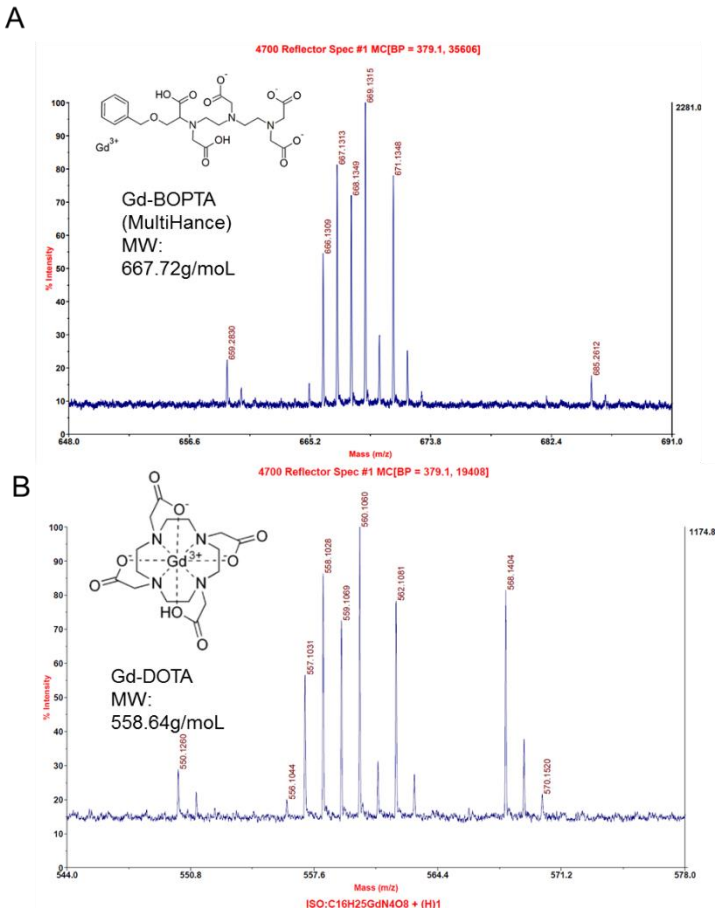

**Supplemental Figure S1** Chemical synthesis of Gd-NM600. Tri-benzyl ester 2 was prepared from commercially available cyclen 1 by treatment with 3.3 equivalents of benzyl bromoacetate and sodium acetate in N,N-dimethylacetamide. N-Alkylation of 2 with tert-butyl bromoacetate in acetonitrile using K2CO3 as a base resulted in DOTA t-butyl tri-benzyl ester 3. Deprotection of the tert-butyl ester in 3 with HCl in dioxane yielded DOTA tri-benzyl ester 4. This compound was coupled with 18-(p-aminophenyl)octadecyl phosphocholine 5 using the coupling agent COMU to obtain conjugate 6. The simultaneous deprotection of the three benzyl ester groups in 6 by the catalytic hydrogenation resulted in DO3A derivative 7, which upon complexation with gadolinium (III) chloride produced the Gd chelate 8**.**



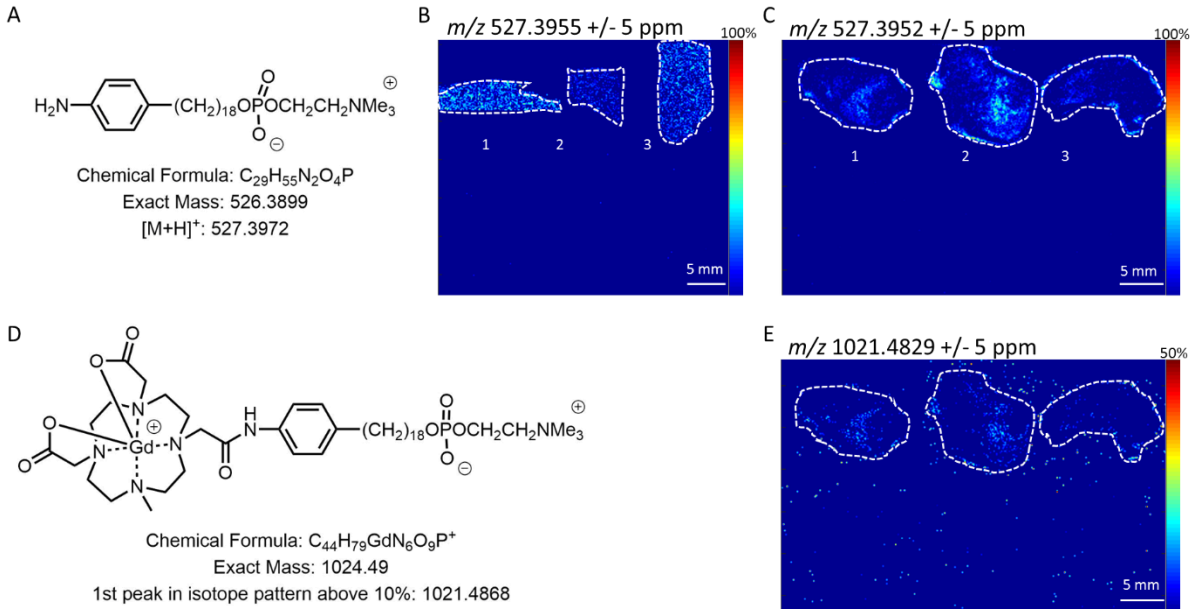

**Supplemental Figure S2.** Tissue Biodistribution of Gd-NM600 and $^{64}$Cu-NM600 surrogate. **(a-c)** *In vivo* $T_1$-weighted axial SPGR images of organs of clearance in a nude athymic mouse showed peak $T_1$-weighted enhancement of the heart (blood pool), liver, and *in vivo* T1-weighted coronal SPGR images of kidneys 1-24 hour post-administration that subsequently decreased over the course of 7 days. $T_1$ enhancement of the liver was the highest and most prolonged suggesting hepatobiliary clearance. White dashed lines outline organs of interest. (Figure legend M = myocardium, L = liver, K = kidney).




**Supplemental Figure S3.** Biodistribution of $^{64}$Cu-NM600 in a mouse model and simultaneous PET/MRI of Gd-NM600/$^{64}$Cu-NM600 uptake at 3.0T in a flank U87 xenograft rat model. **(a)** Organ and tumor biodistribution of micromolar doses of Gd-NM600 and $^{64}$Cu-NM600 in an U87 flank xenograft model. $^{64}$Cu-DOTA-APC was measured by gamma counting. **(b)** Pre-contrast T1-weighted (T1W) coronal images of rat with flank U87 xenograft. Arrow denotes location of the flank xenograft. **(c)** Coronal T1W MRI, PET, overlap images of the same tumor y, and T2W MRI images of a rat 24 hours following co-injection of 5mCi (20mCi/kg) 64Cu-DOTA-APC and Gd-DOTA-APC. **(d)** Coronal T1W MRI, PET, and overlay images 48 hours following contrast delivery.




**Supplemental Fig S4.** MALDI-MSI of commercial chelates Gd-BOPTA and Gd-DOTA. A mixture of 1µmol of Gd-DOTA and Gd-BOPTA were mixed with 40mg/mL of 2,5-dihydroxybenzoic acid (DHB) in distilled water to obtain the mass spectra. Chemical structures and molecular mass of Gd-DOTA and Gd-BOPTA are shown.




**Supplemental Figure S5.** MALDI-MSI of breakdown products in different tissues 7 days after injection of Gd-NM600. **(A)** Cleavage of the amide bond results in the structures shown in A. The distribution of the breakdown product in the liver sections **(B)** and tumor sections **(C)** in all 3 mice injected with Gd-NM600. The breakdown product is uniformly distributed throughout the liver. The compound was not observed in the kidneys. **(D)** Loss of the carboxyl group (-$CO_2$) is shown in **(D)** and the distribution in the tumor is shown in **(E)**. The compound in D was only observed in the tumor sections. White dotted lines represent the borders of the different sections of tissues.

# Appendix X

# Mass Spectrometry Imaging Investigation of Lipid Insulation of Electrocytes in the Electric Eel

Adapted from: Johnson J, Keller C, Gemperline E, Traeger LL, Albert JS, Sussman MR, Li L. Mass Spectrometry Imaging Investigation of Lipid Insulation of Electrocytes in the Electric Eel. In preparation.

Keyword: *Electrophorus electricus*, Electric Eel, Main Electric Organ, MALDI, Mass Spectrometry, Imaging, Lipids

**Abstract**

The electric eel (*Electrophorus electricus*) is a freshwater fish native to South America that produces an electric discharge for predation, defense, and navigation. This unique ability of the electric eel to generate an electric discharge has fascinated scientists with regards to why the electric eel is able to produce electric charges without damaging its own tissues. We hypothesize that the presence of lipid species surrounding the electrocytes, the cells that produce the electric charge, in the main electric organ, help to insulate the eel from the electric shock. To better understand insulation mechanisms in the eel, we conducted spatial, untargeted lipidomics of the main electric organ. The main electric organ is responsible for generating a strong electric charge used to stun prey and is responsible for the largest source of electric shock, delivering up to 600 volts of electricity. By performing mass spectrometry imaging of the main electric organ, we can discriminate lipid patterns present in the insulating areas of the electric organ. Lipid classes localized to insulating regions of the electric organ resemble the lipid composition of the myelin sheath and include lipids such as cerebrosides, sphingomyelins, and sulfatides. Additionally, we detected triacylglycerides and diacyglycerols localized to the insulation regions, likely functioning for energy storage. These results reveal previously unexplored spatial lipidomics of the *E. electricus* main electric organ. This spatial analysis of the main electric organ of the electric eel allows us to hypothesize further biological mechanisms for understanding insulation, generation of shock, and lumen development in the electric eel.

**Introduction**

*Electrophorus electricus*, also known as the electric eel, is a freshwater fish that inhabits rivers and streams in the northern South America, and is well-known for its ability to generate

volleys of high voltage electric discharges of up to 600V used for defense and predation.[1] [2] Although it is called an eel, *E. electricus* is actually a member of a taxonomic order of Neotropical knifefishes called Gymnotiformes. The electric eel has three electric organs: the main organ, Sach's organ, and Hunter's organ. The main electric organ (also called the high or strong voltage organ) is responsible for generating electricity to stun or kill prey and defensive mechanisms, while the Sach's organ and Hunter's organ are weakly electric organs (low voltage organs) primarily used for navigation.[3, 4] The amazing ability of the electric eel to produce large voltages to stun and kill prey has many people wondering how the electric eel is able to produce such large voltages without causing physical harm.[5] Here, we investigate the main electric organ to understand the composition of insulating regions present between electrocytes in the electric eel. Electrocytes, the cellular units of the electric organ, are arranged in longitudinal columns forming insulated tubes along the length of the electric organ.[6] The electrocytes are aligned so that $Na^+$ and $K^+$ ions can pass between the cells to generate a flow of electrical current along the length of the eel's body.[7, 8] We hypothesize that lipid insulation within and among electrocyte tubes channels the flow of this electrical current and contributes to the physiological ability of the electric eel to withstand the production of large voltages.

In biology, lipids can form bilayers to function as essential components of biological membranes, contain hydrocarbon chains that function as energy storage and insulation, and also for inter- and intra-cellular signaling. Only a few studies have looked at the lipid compositions of electric fishes and report on the lipids present in the electric organs; however, these studies have relied on lipid detection via liquid chromatography tandem mass spectrometry or thin layer chromatography analysis in electric organ tissue homogenates.[9-11]Therefore, information concerning the spatial distribution of these detected lipids, possibly functioning as key insulators

between the electrocytes, is a key missing component to better hypothesize how these electric fishes do not shock themselves. Additionally, one study indicated that specifically phenolic lipids directly affect the activity and confirmation of acetylcholinesterases, which are enzymes that transmit a key neurotransmitter, acetylcholine, between neuron and electrocytes to allow for contraction and conductivity in the electric eel.[12] Therefore, the spatial resolution of lipids in the electric organ is very important for understanding lipid functionality in the tissue. Here, we apply mass spectrometry imaging to understand lipid distribution, with a specific interest in lipids localized to the insulating regions of the electric organ.

Mass spectrometry imaging (MSI) is a unique analytical technique that allows us to look at the spatial distribution of lipids in the electric organ. In this study, we use matrix-assisted laser desorption/ionization (MALDI)-MSI to examine insulator regions present in the main electric organ of the electric eel. By performing MSI on a Thermo MALDI LTQ Orbitrap XL mass spectrometer, we can take advantage of the high-resolution, accurate-mass capabilities, which maintains <5 ppm mass accuracy. This high mass accuracy in conjunction with tandem MS (MS/MS) helps us to more confidently identify the detected lipids.[13] In this study, we describe the first mass spectrometry imaging analysis of the electric eel, or any electric fish organ, to reveal the distinct molecular composition of insulating regions located between stacks of electrocytes in the electric eel.

Understanding lipid insulation of electrocytes in electric eel is very relevant for the design and construction of electric-eel inspired bio-compatible electricity sources.[14, 15] In one recent application, engineers are developing hydrogel technologies to mimic the electrocyte design of the electric eel to be used to power pacemakers or other biomedical devices that require an electricity source implanted in the body.[16] Understanding how to make these devices so that this technology

can be used as power source without causing electrical harm to the rest of the body is a fascinating idea,[14] although currently a long way from commercialization. Insulation mechanisms will be needed in order to develop safe bioelectric batteries based on the electric eel's insulation mechanisms. Additionally, very little biological research is currently being done on insulation mechanisms in the electric eel. The electric eel genome was only first sequenced in 2009,[17] and recent transcriptomics[18, 19] and phosphoproteomics[3] studies revealed many key biochemical players involved in how electrocytes are able to generate electricity. However, because these multi-omics analyses are still very recent, most of the current studies focus on key relevant biochemical process that allow the electric eel to generate powerful voltages;[6] little work has been done in understanding the insulation mechanisms needed for electric eel to survive the production of these high voltages. Here, we use analytical chemistry techniques to further understand the spatial localization of lipids in these insulation regions present between stacks of electrocytes. These identified lipid species could be useful for designing artificial electrocytes for the development of bioelectric batteries.

## Materials and Methods

### Electric Eel Dissection of the Main Electric Organ

The protocol governing the animal care and usage of *E. electricus* for this study was approved by the University of Wisconsin Animal Care and Use Committee (Protocol Number M01657). Two ~2 foot long electric eel specimens were used in this study. The main organ of the electric eel tissue was dissected, flash-frozen, and stored in 50mL conical tubes at −80°C until use.

### Sample Preparation for MALDI MSI

Horizontal cross sections of electric eel are dissected off from the main electric organ and embedded in gelatin (100mg/mL in double distilled water) and flash frozen on dry ice. The frozen tissue is then sectioned into 20 μm sections using a cryostat set at -20°C. Tissue sections were thaw-mounted onto standard glass microscope slides. Two matrices were applied in this study. 10 mg/mL α-cyano-4-hydroxycinnamic acid (CHCA) (in 50:50 acetonitrile:water and 0.1% TFA) was applied using an automated TM Sprayer (HTX Technologies, LLC, Carrboro, NC, USA). The matrix TM Sprayer conditions for CHCA were: nozzle temperature: 90°C, flow rate: 0.2 mL/min, 8 passes, 3 mm track spacing (rotate and offset) at a velocity of 1100 mm/min. 10mg/mL 1,5 diaminonapthalene (DAN) matrix (in 90:10 acetone:water and 0.1% TFA) was also applied using an automated TM Sprayer. The matrix TM Sprayer conditions for DAN were: 30°C, flow rate: 0.1 mL/min, 4 passes, 2.5 mm spacing (rotate and offset) at a velocity 1300 mm/min.

MALDI-Orbitrap MSI

A MALDI-Orbitrap mass spectrometer (Thermo Scientific, Waltham, MA, USA), equipped with an $N_2$ laser (spot diameter of 75 μm) was used in positive ion mode for imaging of the main electric organ sections. Imaging was performed using a mass range of *m/z* 500-1200, a mass resolution of 60,000 (at *m/z* 400). The tissue region to be imaged and the raster step size were controlled using the LTQ Tune software (Thermo Scientific, Waltham, MA, USA). The instrument methods were created using Xcalibur (Thermo Scientific, Waltham, MA, USA). To generate images, the spectra were collected at 75 μm intervals in both the x and y dimensions across the surface of the sample. The raw data was converted into imzmL using ImageQuest (Thermo Scientific, Waltham, MA, USA) prior to data processing.

MALDI MSI Data Processing and Statistical Analysis

MS data was processed using Xcalibur (Thermo Scientific, Bremen, Germany) and ImageQuest (Thermo Scientific, Bremen, Germany). MS data (in .raw format) was exported into .imzmL format prior to uploading into MSiReader.[20, 21] MSiReader was used to create a list of compounds of interest by selecting the eel tissue as the interrogated zone and subtracting the matrix peaks as the reference zone. Images were then filtered based on spatial distribution into 3 categories of interest: electrocyte, lumen, and insulator regions. A list of *m/z* values from all 3 regions generated from this list was used for targeted MS/MS analysis. Images were normalized to total ion current (TIC) and the mass tolerance window for image construction is 5 ppm. MS data as a non-centroided .imzmL format is loaded into SCiLS software 2019 (SCiLS Lab Software, Bruker Daltonics, Bremen, Germany) for statistical analysis. Using SCiLS, we performed Principal Component Analysis on individual spectra from imported peaks using 5 principal components. SciLs software was also used for spatial analysis by co-registration of the histology to the TIC of the electric eel and annotating the electrocytes and lumen based on the histology.

Histology Processing

CHCA matrix was removed using by washing the slide with 50:50 acetonitrile:water, and the slides were stained with Hematoxylin and Eosin (H&E) to reveal staining of the cellular regions of the electric eel main organ. Images of the H&E stained tissues were acquired using a Meyer Instruments PathScan Enabler IV (Meyer Instruments) at 4X magnification.

Lipid Extraction

Lipids were extracted from the electric organ. Briefly, the main electric organ tissue was ground using mortar and pestle in liquid nitrogen to freeze the tissue. The powder was transferred to a 50 mL conical tube and then 50 mL of extraction solvent composed of 40% methanol, 40% acetonitrile, and 20% water was used to extract the metabolites and lipids from the electric organ.

The conical tube was centrifuged at 7,000 x g for 10 min at 4°C. Following this, the supernatant was removed and dried down in multiple aliquots the SpeedVac under medium heat. Samples were resuspended in 75% methanol and 25% water until dissolved for use prior to in a 3kDa molecule weight cutoff (MWCO) clean-up (Amicon Ultra, Millipore). The MWCO was rinsed with 0.2 mL of 0.1 M NaOH, and then rinsed with 500 µL 75% methanol and 25% water. Rinses were centrifuged at 14,000 x g until the rinse was through the cartridge. Aliquots of 500 µL of extract was added per MWCO device. The sample was then loaded and centrifuged at 14,000g for 1 hour. The MWCO device containing the extract was rinsed with 0.2 mL 75% methanol and centrifuged at 14,000 x g until the rinse was through the cartridge. The flow through was saved for mass spectrometry analysis and the solvent was evaporated using a SpeedVac and then stored at -80°C.

LC-ESI-MS

To acquire LC-ESI-MS, eel extracts were resuspended at 10 mg/mL in 50% acetonitrile/ 50% water with 0.1% FA. The eel extracts were separated on a Kinetix C18 column (2.1-mm internal diameter × 150-mm length, 1.7-µm particle size; Phenomenex, Torrance, CA, USA), equipped with a corresponding guard column, and heated to 35 °C. The mobile phases were (A) optima grade water with 0.1% formic acid and (B) optima grade acetonitrile with 0.1% formic acid. The eel samples were separated within 35 min under the following conditions: 0-5 min, isocratic hold at 1% B; 5-10 min, linear gradient from 1-3% B; 10-18 min, linear gradient from 3-40% B; 16-22 min, linear gradient from 40-80% B; 22-27 hold at 95% B and finally re-equilibration of the system at 1% B for 8 min. The flow rate was 0.3 mL/min and the injection volume was 3 µL. The samples were kept at 10°C during the analysis. MS and MS/MS data were acquired on a Q-Exactive instrument (Thermo Scientific, Waltham, MA, USA) that was equipped with an ESI source operated in positive ion mode. The MS scan range was from *m/z* 500–1200. A

targeted MS/MS list was created in MSiReader from the MALDI LTQ Orbitrap XL imaging data to include MS peaks found in regions of interest from the mass spectrometry imaging experiments. The MS/MS scan range was adjusted depending on the parent mass and high-energy collision dissociation (HCD). The MS/MS data were collected for the targeted lipids using a stepped collision energy set-up with collision energies of 20, 30, and 40 eV.

Lipid Identification

The target list of putative identifications were identified by searching the accurate mass obtained through imaging using Metlin,[22] Metabosearch,[23] and Metaspace.[24] Some of these tentative identifications were confirmed with LC-MS/MS spectra. LC-MS/MS spectra were analyzed with MetFrag *in silico* fragmentation tool.[25] For parent ion searching, a 5 ppm tolerance was used. The $[M]^+$, $[M+H]^+$, $[M+NH_4]^+$, $[M+Na]^+$, and $[M+K]^+$ adducts of the parent ion were searched against the KEGG and LipidMaps databases. For MS/MS searching, fragmentation ion tolerances of 5 Mzppm and 0.01 Mzabs were used and the fragmentation mode was set to match the parent ion adduct. The three spectra from the *m/z* 544. 33566 peaks (RT 22.02 and 22.15 min) were averaged as the MS/MS was very similar between the three and the two peaks were the main peaks in the extracted ion chromatogram. Putative identifications were made if all major peaks were explained by the MetFrag *in silico* fragmentation.

## Results and Discussion

Histological Features of Interest in the Main Electric Organ

For mass spectrometry imaging, thin horizontal cross sections of the main electric organ were taken for analysis. **Figure 1(A)** shows the in-tact main electric organ dissected out from the electric eel. In **Figure 1(A)**, clear striation patterns present between electrocytes can be seen, which

function as the insulation regions in the tissue. Here, multiple electrocytes are stacked together, with these striation patterns insulating each stack of electrocytes from the next stack. **Figure 1(B)** shows the electric organ tissue and sections from a horizontal cross sectioned used for the MSI analysis. The Hematoxylin & Eosin (H&E) stain was performed on these thin tissue sections (shown on the bottom row of **Figure 1(B)**). The H&E stains the cellularized regions of the tissue in a pink color with cell's nuclei shown in a slightly darker purple. The staining of cellularized regions in the H&E stain revealed the presence of lumen (unstained holes) approximately 1mm in diameter inside of the electrocytes of the electric eels. These lumen have never been reported in the electric eel, although they have been observed during development in other electric fishes.[2, 26] **Supplemental Figure 1** also shows MSI data detecting the presence of a chemically distinct inner lining in these lumen. The H&E also shows the expected multi-nucleated portions in the electrocytes of electric eel (**Figure 1(B)**). The multi-nucleated cell regions are consistent with previous reports of electric eel histological features.[6] From this feature analysis, we can visually discriminate the specific regions localized between stacks of the electrocytes to understand the lipid composition of these regions and identify the presence of lumen inside the electrocytes.

Unsupervised Data Clustering Reveals Distinct Insulator, Electrocyte, and Lumen as Key Spectral Features

While histological inspection of the tissue reveals features, unsupervised data clustering can also be used to reveal regions with distinct chemical compositions in the tissue. Following mass spectrometry data acquisition, we applied principal component analysis (PCA) to perform unsupervised data clustering to reveal regions that are causing the greatest variance in the data.[27] The principal components (PC) are arranged so that PC 1 has the largest variance. Each succeeding component accounts for less variance within the dataset. For example, here in as shown in **Figure**

**2** and **Supplemental Figure S2**, PC1 accounts for 12%, PC 2 accounts for 4.3%, PC 3 accounts for 2.5%, PC 4 accounts for 2%, and PC 5 accounts for 1% of the total data. Together PC1-5 showcase over 20% of the variance in the total sample (**Supplemental Figure S2**). These components show us the ranked importance of tissue regions in terms of how much they contribute to the overall variance in the tissue. PC1 distinguishes tissue signal from the background. PC2 distinguished the lumen in the electrocytes and PC3 and PC4 distinguish the electrocytes from the insulator regions in the tissue as shown in **Figure 2**. PC 5 demonstrates that most of the tissue feature variance can be accounted for in PC1-4. PCA provides an unsupervised processing method to segment out distinct regions of chemical complexity in the electric organ and to look at the main underlying trends and spectral features. This method provides an unbiased way for us to classify different regions of the tissues, such as the insulator regions, electrocytes, and lumen for further investigation.

Lipids Localized to the Insulator Regions

As shown in the PCA results, the MALDI-MSI results show lipids with distinct localization patterns to the insulator regions, electrocyte regions, and the electrocyte lumen regions. **Supplemental Figure S3** shows example MS spectra for each of the three regions and **Supplemental Tables S1-S3** list detected *m/z* values unique to each of these classified regions. From some of the images shown in **Figure 3** and from PC4, we understand that the distinct striation patterns present in the data correspond with the insulator regions located between arranged stacks of electrocytes in the electric organ. The serial arrangement of electrocytes between insulating regions allows current to easily pass to opposite ends of the eel at high voltages. However, because the electric eel is also able to laterally give off and sense electricity, likely these insulating regions would need to be imperfect to allow electricity to create transdermal electric fields at all regions

of the eel.[28] **Figure 3** shows some of the *m/z* values of lipids detected that are specifically localized to these insulating regions. Many of the striation lines in the patterns shown in these *m/z* values appear to have small pockets of high relative concentrations of lipids, followed by what appears to be small breaks in the striation line patterns. We hypothesize that these patterns seen in the insulator regions could allow for electricity to pass through small gaps in the insulating barrier, however these insulating regions would still allow enough insulation to for the delivery of $Na^+$ and $K^+$ ions to generate large voltages through the summation of the stacks of electrocytes throughout the main electric organ.

*The Lipid Composition of Insulator Regions is Similar to the Lipid Composition in Myelin Sheath of Neurons*

Common lipids specific to the insulator regions include cerebrosides, sulfatides, and spingomyelins, based on accurate mass matching with high mass resolution mass spectrometry. The distributions of examples from these lipid classes that show spatial localization to the insulator regions are shown in **Figure 3**. What is particularly interesting about the identification of these lipids as localized to these insulator regions is that many of these lipids also match key lipid components of the myelin sheath. This is possibly because innervation along the insulator regions can function to stimulate electrical function and to simultaneously insulate stacks of electrocytes from each other. The myelin sheath is a thin layer that acts as an insulator for nerves in both the central nervous system and peripheral nervous system.[29] The myelin sheath by composition is very lipid rich, with about 70% of its composition due to lipid insulation to protect the nerves. While there are no specific lipids unique to myelin, cerebrosides are the most typical lipid found in myelin.[29, 30] Cerebrosides consist of a ceramide lipid with a single sugar residue at the 1-hydroxy moiety of the lipid, typically either glucose or galactose. Here we see a galactylceramide

cerebroside, shown in **Figure 3(A)** as *m/z* 788.5574. Additionally, the 3-hydroxyl moiety of the galactose moiety of a cerebroside is also typically sulfated in myelin, a group of lipids commonly referred to as sulfatides.[31] Here, we tentatively identify *m/z* 796.5288 (**Figure 3(B)**) as a sulfatide, 3'sulfo Galbeta-Cer (d18:/16:0(2OH)), localized to the striation patterns in the insulator regions of the electric eel tissue. Previous work in identifying lipids in the electric organ of the ray, *Torpedo marmorata*, an electric fish, found that sulfatides were unique to the electric organ. The authors hypothesize that the function of the sulfatides could be related to the selection of $K^+$ by the sulfate group and the interaction of these lipids with the ATPase pole.[10] Here, the sulfatide localization to the insulator region is an important addition combined with previous literature as a unique lipid to the electric organ and could possible lead to new biological insights for the role of the sulfatides with the ATPase function in the electric eel. We also tentatively identified precursors to cerebroside synthesis, including sphingomyelin (*m/z* 833.6510) and ceramide 1-phosphate (*m/z* 766.5512) (**Figure 3(C-D)**), localized to the insulator regions. **Supplemental Table S4** provides a complete list of the insulator tentative identifications, with many examples of myelin similar lipids. The co-localization of several lipid species found prominently in myelin, including the most myelin specific lipids, cerebrosides, sulfatides, and sphingomyelins, with the insulator regions of the electric organ tissue supports the hypothesis that myelin-similar lipid rich insulator regions allow for increased flow of $Na^+$ and $K^+$ ions through the eel. Additionally, we hypothesize that the small gaps in the insulator regions seen spatially in the mass spectrometry imaging data could allow for electrical flow out horizontally from the electric eel, although further investigation with functional studies would be needed to confirm this hypothesis. There is some additional literature that supports the idea that other biomolecules, such as collagen, can also act as electrical insulators that may be localized to the insulator regions.[32-34] To probe the hypothesis that collagen could be

co-acting as an insulator, we performed picrosirus red staining on the electric organ tissue sections and used polarized light to visualize the collagen distribution in the electric organ, shown in **Supplemental Figure S4**. The microscopy results do not show a specific spatial localization of fibrillar collagen to the insulators region, but rather show collagen fibrils present in the both the electrocyte and insulator regions of the tissue. This collagen staining further supports our hypothesis that these myelin-specific lipid species, rather than fibrillar collagen, showing specific localization to insulator regions, might be key to the regulation in electric eel insulation.

*Energy Reserves Stored in Insulator Regions*

The electric eel is shockingly able to deliver up to 600V of electricity in predation or defense, with suddenly being able to amp up its base-line electricity levels on the time order of milliseconds. This rapid need for electrical conductivity suggests that energy repositories must immediately be available to create this sudden electric discharge. In the insulator regions, surrounding the electrocytes, we also find the localization of triacylglycerols and their precursors lipids, diacyglycerols. Triacyglycerols are fatty acid triester glycerols, and are typically known for their role in energy storage in adipose tissues.[35] Figure 3E shows *m/z* 779.6134 tentatively identified as a triacyglycerol and *m/z* 729.4839 as a diacylglycerol, a precursor to triacyglycerol synthesis. Other diacyglycerol identifications can be found in **Supplemental Table S4**. It is interesting that we also find the energy storage reserves, in the form of triacygrlycerols and diacyglycerols, localized in the insulator regions. We hypothesize that the localization of these lipids in the insulators is necessary for rapid energy transport in the eel to create large electric discharges in response to threat or attack.

Lipids Localized to the Electrocyte Tissues

Although this paper mostly focuses on the insulation regions of the main electric organ, we also detect several lipids with distinct distributions in the cellularized region of the electrocyte and lipids also specific to the lumen inside of the electrocytes. Here, in **Figure 4**, we highlight one example of the cellularized electrocyte region, *m/z* value 844.5180, and one example showing localization to the lumen inside of the electrocytes, *m/z* value 543.8911. For a more comprehensive list of the *m/z* values found localized to both the lumen and the electrocyte, refer to **Supplemental Table S2** and **Supplemental Table S3**, respectively. Of particular interest to electric eel physiology, inside the lumen, *m/z* 543.9811 tentatively matches using Metlin[22] database to a Methylene ATP analog, either *β,γ*-methylene ATP (*βγ*-meATP) or *α,β*-methylene ATP (*αβ*-meATP), which has a methylene group substitute for the oxygen is the phosphodiester bridge between the phosphate moieties of ATP. This methylene ATP modulates extracellular ATP metabolism and accumulation and is likely involved in mechanisms to help generate electricity.[36] Literature also links methylene ATP with a direct ability to causes contractile response in the bladder,[37] which in the eel could help accelerate the contractile response and curling behavior of the eel when it delivers electricity.[38] Its high relative concentration localized in the lumen of the electrocyte is particularly interesting, as these lumen have never previously been identified in the electric eel development. We hypothesize from this localization that the lumen could be functioning as a reservoir for these methylene ATP analogs to help modulate the surrounding electrocytes for ATP controlled biochemical mechanisms. Further investigation into lumen development and possible lumen disappearance in the adult electric eel could lead to new insight to correlate these biochemical changes in eel development and the unknown role of the electrocyte lumen.

Only a few MS/MS spectra were able to be obtained from the target list of *m/z* observed in the MALDI-MSI experiments during the LC-MS/MS experiments. From the MS/MS spectra, *m/z* 544.33566, which is localized in the electrocyte region of the main electric organ, was identified as a PC lipid. **Figure 5** shows the results for the analysis of the MS/MS spectrum obtained from the LC-MS/MS experiments. The MetFrag analysis for this *m/z* revealed no matches to the KEGG databases and the only matches to the LipidMaps databases were with the $[M+Na]^+$ adduct.[25] All of the six matches to the LipidMaps database were from PC lipids. The *in silico* fragmentation results (**Figure 5(B)**) show that 4 of the 6 matches are most likely as the last two options are much lower scoring. Further experiments would be necessary to narrow down the putative identification, but the 184.07233 fragment, which is indicative of a PC head group, suggests that this *m/z* is a PC lipid.

In this manuscript, we describe the application of mass spectrometry imaging to tentatively identify lipids in the insulation patterns in the electric eel. From this analysis, we notice similarities in lipid composition in these insulator regions to the myelin sheath. We hypothesize that these insulation regions could be multi-functional: to allow innervation of the electric organ and to serve as both insulators and energy storage to support electrocyte functionality. The results from this study can be used in conjugation with recent other multi-omics analyses of including genomics,[17] transcriptomics,[18] and proteomics[3] to help chemical biologists better understand the amazing electrical power of the electric eel and its ability to withstand such high voltages. This work merges mass spectrometry imaging with electric fish lipid biology, as many of these lipids had previously been identified in the electric organ,[10, 11] but the localization of these lipids has led to new important biological insights into electrical eel biochemical mechanisms.

**Acknowledgements**

L.L. acknowledges a Vilas Distinguished Achievement Professorship and a Janis Apinis Professorship with funding provided by the WARF and UW-Madison School of Pharmacy. The MALDI-Orbitrap was purchased through a National Institutes of Health shared instrument grant (NCRR S10RR029531).

**References**

1. KEYNES, R. D.; MARTINS-FERREIRA, H., Resting and action potentials in the electric organ. *J Physiol* **1952,** *116* (3), 26p-7p.

2. Albert, J. S.; Crampton, W. G. R., Electrorecepton and electrogenesis. The physiology of fishes: 2005; Vol. 3.

3. Traeger, L. L.; Sabat, G.; Barrett-Wilt, G. A.; Wells, G. B.; Sussman, M. R., A tail of two voltages: Proteomic comparison of the three electric organs of the electric eel. *Sci Adv* **2017,** *3* (7), e1700523.

4. Cox, R. T.; Coates, C. W.; Brown, M. V., ELECTRIC TISSUE : RELATIONS BETWEEN THE STRUCTURE, ELECTRICAL CHARACTERISTICS, AND CHEMICAL PROCESSES OF ELECTRIC TISSUE. *J Gen Physiol* **1945,** *28* (3), 187-212.

5. Catania, K., The shocking predatory strike of the electric eel. *Science* **2014,** *346* (6214), 1231-4.

6. Gotter, A. L.; Kaetzel, M. A.; Dedman, J. R., Electrophorus electricus as a model system for the study of membrane excitability. *Comp Biochem Physiol A Mol Integr Physiol* **1998,** *119* (1), 225-41.

7. WHITTAM, R.; GUINNEBAULT, M., The efflux of potassium from electroplaques of electric eels. *J Gen Physiol* **1960,** *43*, 1171-91.

8. Markham, M. R., Electrocyte physiology: 50 years later. *J Exp Biol* **2013,** *216* (Pt 13), 2451-8.

9. Rotstein, N. P.; Arias, H. R.; Barrantes, F. J.; Aveldaño, M. I., Composition of lipids in elasmobranch electric organ and acetylcholine receptor membranes. *J Neurochem* **1987,** *49* (5), 1333-40.

10. Hansson, G. C.; Heilbronn, E.; Karlsson, K. A.; Samuelsson, B. E., The lipid composition of the electric organ of the ray, Torpedo marmorata, with specific reference to sulfatides and Na+-K+-ATPase. *J Lipid Res* **1979,** *20* (4), 509-18.

11. Supernovich, C.; Crain, R.; Rosenberg, P., Phosphatidylcholine asymmetry in electroplax from the electric eel: use of a phosphatidylcholine exchange protein. *J Neurochem* **1991,** *57* (2), 575-84.

12. Stasiuk, M.; Janiszewska, A.; Kozubek, A., Phenolic lipids affect the activity and conformation of acetylcholinesterase from Electrophorus electricus (Electric eel). *Nutrients* **2014,** *6* (5), 1823-31.

13. Palmer, A.; Phapale, P.; Chernyavsky, I.; Lavigne, R.; Fay, D.; Tarasov, A.; Kovalev, V.; Fuchser, J.; Nikolenko, S.; Pineau, C.; Becker, M.; Alexandrov, T., FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry. *Nat Methods* **2017,** *14* (1), 57-60.

14. Sun, H.; Fu, X.; Xie, S.; Jiang, Y.; Peng, H., Electrochemical Capacitors with High Output Voltages that Mimic Electric Eels. *Adv Mater* **2016,** *28* (10), 2070-6.

15. Xu, J.; Lavan, D. A., Designing artificial cells to harness the biological ion concentration gradient. *Nat Nanotechnol* **2008,** *3* (11), 666-70.

16. Schroeder, T. B. H.; Guha, A.; Lamoureux, A.; VanRenterghem, G.; Sept, D.; Shtein, M.; Yang, J.; Mayer, M., An electric-eel-inspired soft power source from stacked hydrogels. *Nature* **2017,** *552* (7684), 214-218.

17. Gallant, J. R.; Traeger, L. L.; Volkening, J. D.; Moffett, H.; Chen, P. H.; Novina, C. D.; Phillips, G. N.; Anand, R.; Wells, G. B.; Pinch, M.; Güth, R.; Unguez, G. A.; Albert, J. S.; Zakon, H. H.; Samanta, M. P.; Sussman, M. R., Nonhuman genetics. Genomic basis for the convergent evolution of electric organs. *Science* **2014,** *344* (6191), 1522-5.

18. Traeger, L. L.; Volkening, J. D.; Moffett, H.; Gallant, J. R.; Chen, P. H.; Novina, C. D.; Phillips, G. N.; Anand, R.; Wells, G. B.; Pinch, M.; Güth, R.; Unguez, G. A.; Albert, J. S.; Zakon, H.; Sussman, M. R.; Samanta, M. P., Unique patterns of transcript and miRNA expression in the South American strong voltage electric eel (Electrophorus electricus). *BMC Genomics* **2015,** *16*, 243.

19. Ching, B.; Woo, J. M.; Hiong, K. C.; Boo, M. V.; Wong, W. P.; Chew, S. F.; Ip, Y. K., Voltage-Gated Na+ Channel Isoforms and Their mRNA Expression Levels and Protein Abundance in Three Electric Organs and the Skeletal Muscle of the Electric Eel Electrophorus electricus. *PLoS One* **2016,** *11* (12), e0167589.

20. Bokhart, M. T.; Nazari, M.; Garrard, K. P.; Muddiman, D. C., MSiReader v1.0: Evolving Open-Source Mass Spectrometry Imaging Software for Targeted and Untargeted Analyses. *J Am Soc Mass Spectrom* **2018,** *29* (1), 8-16.

21. Robichaud, G.; Garrard, K. P.; Barry, J. A.; Muddiman, D. C., MSiReader: an open-source interface to view and analyze high resolving power MS imaging files on Matlab platform. *J Am Soc Mass Spectrom* **2013,** *24* (5), 718-21.

22. Guijas, C.; Montenegro-Burke, J. R.; Domingo-Almenara, X.; Palermo, A.; Warth, B.; Hermann, G.; Koellensperger, G.; Huan, T.; Uritboonthai, W.; Aisporna, A. E.; Wolan, D. W.; Spilker, M. E.; Benton, H. P.; Siuzdak, G., METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Anal Chem* **2018,** *90* (5), 3156-3164.

23. Zhou, B.; Wang, J.; Ressom, H. W., MetaboSearch: tool for mass-based metabolite identification using multiple databases. *PLoS One* **2012,** *7* (6), e40096.

24. C Silva, A. S.; Palmer, A.; Kovalev, V.; Tarasov, A.; Alexandrov, T.; Martens, L.; Degroeve, S., Data-Driven Rescoring of Metabolite Annotations Significantly Improves Sensitivity. *Anal Chem* **2018,** *90* (19), 11636-11642.

25. Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S., MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform* **2016,** *8*, 3.

26. Rodríguez-Cattáneo, A.; Aguilera, P.; Cilleruelo, E.; Crampton, W. G.; Caputi, A. A., Electric organ discharge diversity in the genus Gymnotus: anatomo-functional groups and electrogenic mechanisms. *J Exp Biol* **2013,** *216* (Pt 8), 1501-15.

27. Buchberger, A. R.; DeLaney, K.; Johnson, J.; Li, L., Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal Chem* **2018,** *90* (1), 240-265.

28. Stoddard, P. K.; Markham, M. R., Signal Cloaking by Electric Fish. *Bioscience* **2008,** *58* (5), 415-425.

29. Morell, P.; Quarles, R., The Myelin Sheath, Basic Neurochemistry: Molecular, Cellular, and Medical Aspects 6th Edition ed.; Lippincott Raven Philadelphia, 1999.

30. Aggarwal, S.; Yurlova, L.; Simons, M., Central nervous system myelin: structure, synthesis and assembly. *Trends Cell Biol* **2011,** *21* (10), 585-93.

31. Grassi, S.; Prioni, S.; Cabitta, L.; Aureli, M.; Sonnino, S.; Prinetti, A., The Role of 3-O-Sulfogalactosylceramide, Sulfatide, in the Lateral Organization of Myelin Membrane. *Neurochem Res* **2016,** *41* (1-2), 130-43.

32. Cho, Y.; Borgens, R. B., Electrically controlled release of the nerve growth factor from a collagen-carbon nantotube composite for supporting neuronal growth. Journal of Materials Chemistry, Vol. 1, pp 4166-4170.

33. Tzukert, A.; Leviner, E.; Mahler, Y.; Shoshan, S., Electroconductivity of collagen in vitro and in vivo. Biochimica et Biphysica Acta- General Subjects: 1980; Vol. 3, pp 276-280.

34. Kaur, G.; Adhikari, R.; Cass, P.; Bown, M.; Gunatillake, P., Electrically conductive polymers and composites for biomedical applications. RSC Advances, 2015; Vol. 5, pp 37553-37567.

35. Ahmadian, M.; Duncan, R. E.; Jaworski, K.; Sarkadi-Nagy, E.; Sul, H. S., Triacylglycerol metabolism in adipose tissue. *Future Lipidol* **2007,** *2* (2), 229-237.

36. Joseph, S. M.; Pifer, M. A.; Przybylski, R. J.; Dubyak, G. R., Methylene ATP analogs as modulators of extracellular ATP metabolism and accumulation. *Br J Pharmacol* **2004,** *142* (6), 1002-14.

37. Burnstock, G., Purinergic signaling in the gastrointestinal tract. *World J Gastrointest Pathophysiol* **2011,** *2* (2), 31-4.

38. Catania, K. C., Electric Eels Concentrate Their Electric Field to Induce Involuntary Fatigue in Struggling Prey. *Curr Biol* **2015,** *25* (22), 2889-98.

**Figures**

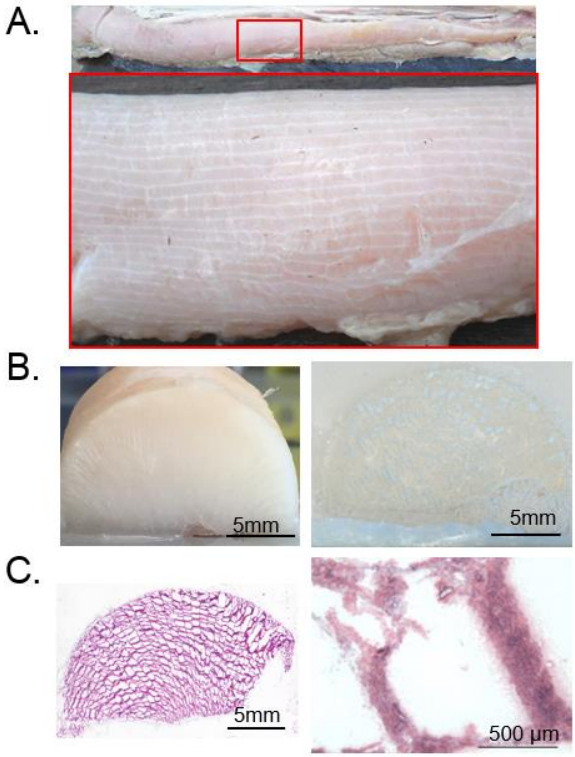



**Figure 1.** Histological Features of the Main Electric Organ. **(A)** Dissected electric eel main electric organ contains white striation patterns present in the morphology, visually showing the insulating regions located between electrocytes **(B)** Thin horizontal section of the main electric organ are taken for matrix application and mass spectrometry analysis. **(C)** The Hematoxylin & Eosin Stain reveals the presence of lumen, located between electrocytes, and upon closer examination of the electrocytes, multi-nucleation in the tissues is observed.

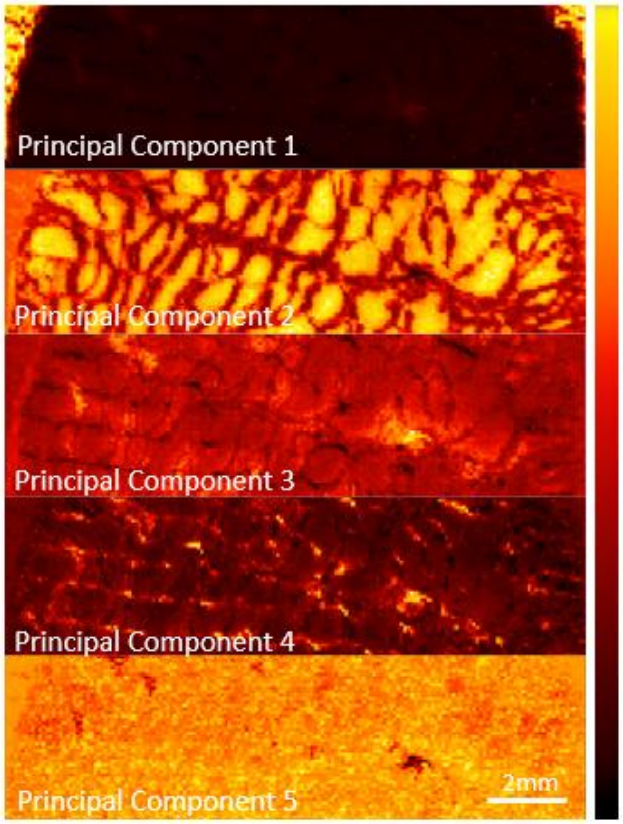



**Figure 2.** Principal Components 1-5 of a small section of electric eel main electric organ reveals key features in the tissue including tissue lumen (PC2), electrocytes (PC3), and insulator regions (PC4) based on variance in the tissue.

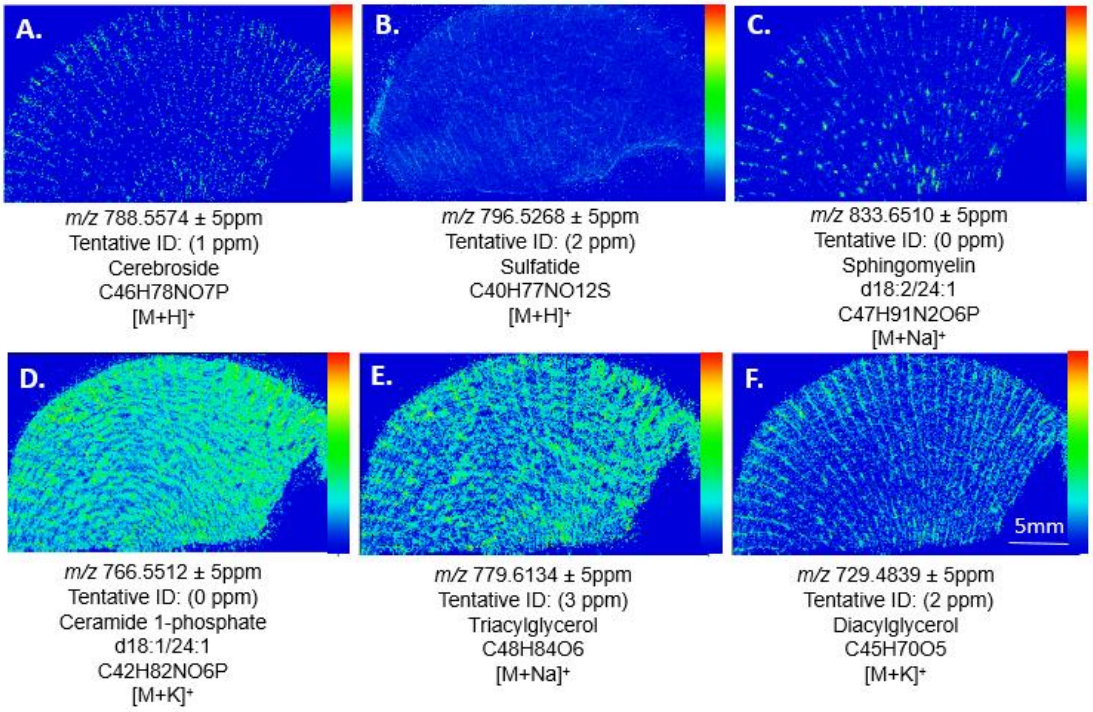



**Figure 3.** Lipids Localized Insulator Regions. **(A)** *m/z* 788.5574, tentatively identifying as a Cerebroside **(B)** 796.5268, tentatively identifying as a Sulfatide **(C)** *m/z* 833.6510, tentatively identifying as a Sphingomyelin **(D)** *m/z* 766.5512 tentatively identifying to a Ceramide. **(E)** *m/z* 779.6134 tentatively identifying as a Triacylglycerol **(F)** *m/z* 729.4839, tentatively identifying as a Diacylglycerol.

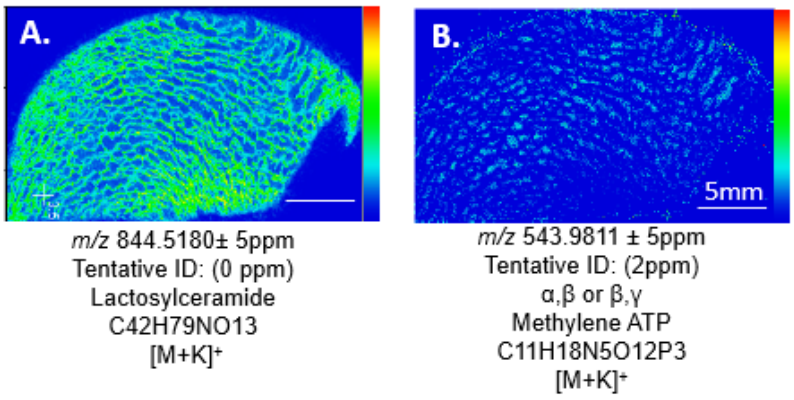
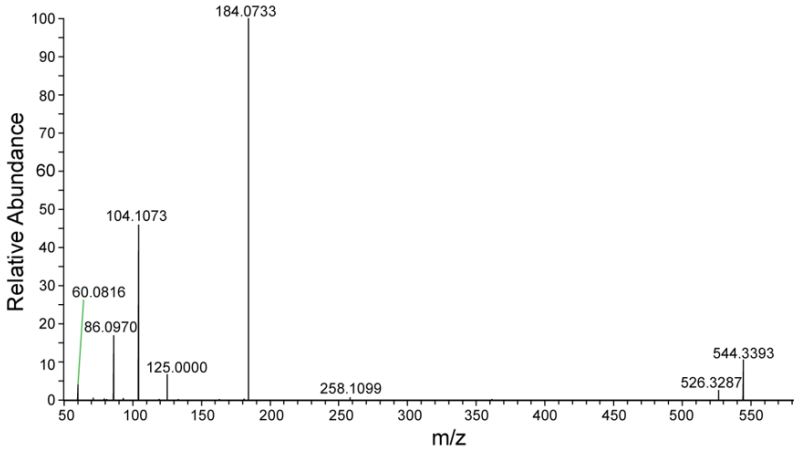



**Figure 4.** Lipids localized to the Electrocyte Tissue. **(A)** *m/z* 844.5180, tentatively identifying to Lactosylceramide, is localized to the electrocyte tissue. **(B)** *m/z* 543.9811, tentatively identifying to methylene ATP, is localized to lumen between electrocytes.

A. MSMS *m/z* 544.33566 RT 22.02/22.15




B.

| Compound Name |
|---|
| PC(18:1(9Z)\0:0) |
| PC(18:1(6Z)\0:0) |
| PC(0:0\18:1(9Z)) |
| PC(0:0\18:1(6Z)) |
| PC(P-16:0\2:0) |
| PC(O-16:1(11Z)\2:0) |

C. *m/z* 544.33566 +/- 5 ppm

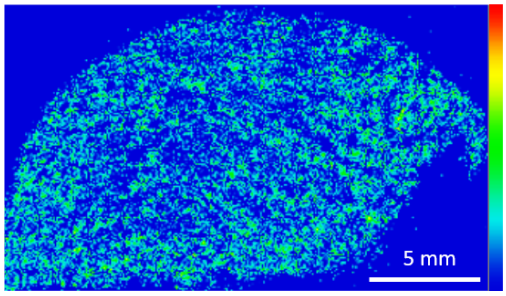



**Figure 5.** PC lipid identification. **(A)** MS/MS Spectra from LC-MS/MS of *m/z* 544.33566 used to identify the *m/z* as a PC lipid with associated image **(B)** Table showing MetFrag results from the LipidMaps database to the $[M+Na]^+$ adduct **(C)** Localization in the electrocyte regions of the tissue

## Supplemental Information

## Supplemental Materials and Methods

Supporting Information for Supplemental Figure 4.

Electric Organ was embedded in Optimal Cutting Temperature (OCT) and 10um flash frozen sections were mounted onto glass slides for staining. The Picrosirus Red staining solution was made by dissolving 0.1g Fast Green FCF (F7258; Sigma-Aldrich, St. Louis, MO) and 0.1 g Sirius Red F3B (S03695; Pfaltz & Bauer, Waterbury, CT) in 100 ml of saturated picric acid. Images of the Picrosirus Red stain were obtained on a traditional transmission pathology microscope (BX53, Olympus Corp) using a 10X objective without polarizers and with linear polarizers.

Supporting Information for Supplemental Figure 5 On-Tissue MS/MS

A target list of *m/z* observed in imaging experiments on the Rapiflex TOF/TOF mass spectrometer (Bruker Daltonics) was also used for on-tissue MS/MS experiments. The parent ion spectrum was summed with the fragment ion spectrum (1,000 laser shots). Precursor ions were isolated with a window of 0.60% of the precursor, and a laser boost of 60% was used for the fragment ions. The initial precursor ion laser energy was optimized for each precursor to maximize parent ion signal without fragmenting the precursor. FlexAnalysis software (Bruker Daltonics) was used to visualize MS/MS spectra and create peak lists for the spectra.

## Supplemental Results and Discussion

Supporting Information for Supplemental Figure 4

Picrosirius Red stains connective tissue with a rich red color but can also be combined with conventional widefield polarization microscopy to greatly enhance the birefringence of collagen fibers in the electric organ tissue, as shown in **Supplemental Figure 4 (B**). In supplemental figure 3, we demonstrate collagen content present in both the electrocyte tissue and in the insulator tissue in the main electric organ. No fibrillar collagen is detected in lumen in between electrocytes. Some hypothesize that collagen could contribute to electrical insulation in tissues. While we detect the collagen in the insulator region, it does not exhibit regional specificity to the insulator regions in the electric eel. This evidence makes us hypothesize that regional specific lipid insulators may be more likely contribute to insulation mechanisms in the electric eel main organ.

Supporting Information for Supplemental Figure 5 On-Tissue MS/MS

Due to the low number of MALDI-MSI target *m/z* observed in the ESI experiments, on-tissue MS/MS was also attempted to get MS/MS of the MALDI-MSI targets. On-tissue MS/MS was conducted on the Rapiflex instrument in Lift mode due to its good MS/MS sensitivity. Good MS/MS spectra was obtained for multiple *m/z* with the Rapiflex. Co-isolation issues, along with the lower mass resolution of the Rapiflex, however, prevented the acquisition of on-tissue MS/MS spectra for most MALDI-MSI targets. While the quality of the obtained MS/MS spectra was good, the lack of MS1 matches to the KEGG and LipidMaps databases to the target *m/z* from the MALDI Orbitrap analysis impeded identification of the MS/MS spectra. The 184 fragment ion, indicative of the PC head group, did appear in the MS/MS spectra, so it is possible that the unknown *m/z* values shown in **Supplemental Figure S5** are also PC or LysoPC lipids.

## Supplemental Tables

**Supplemental Table S1.** This table shows the detected peaks from the different matrix conditions of the main electric eel tissue localized to the insulator regions. The values shown below are detected in both biological replicates of the electric eels, matched within 5ppm of the peak detected, and different peaks detected between the two matrix conditions: CHCA and DAN matrix

**Insulator Regions of Tissue**

| CHCA Matrix | |
|---|---|
| Bioreplicate 1 *m/z* values | Bioreplicate 2 *m/z* value |
| 542.4907 | 542.4903 |
| 666.4842 | 666.4826 |
| 682.4581 | 682.4574 |
| 725.5574 | 725.5571 |
| 741.5313 | 741.5312 |
| 766.5512 | 766.5504 |
| 767.5002 | 767.4989 |
| 768.5036 | 768.5022 |
| 776.594 | 776.5928 |
| 778.6105 | 778.6084 |
| 779.4402 | 779.4387 |
| 789.4847 | 789.4831 |
| 796.5268 | 796.5254 |
| 825.6259 | 825.6256 |
| 826.5734 | 826.5729 |
| 827.5771 | 827.576 |
| 839.6403 | 839.6403 |
| 851.6424 | 851.6408 |
| 853.6574 | 853.6568 |

| DAN Matrix | |
|---|---|
| Bioreplicate 1 *m/z* values | Bioreplicate 2 *m/z* values |
| 542.4905 | 542.4911 |
| 624.5695 | 624.5696 |
| 626.5854 | 626.5855 |
| 640.6010 | 640.6003 |
| 652.6000 | 652.6007 |
| 654.6155 | 654.6167 |
| 666.4837 | 666.4839 |
| 667.4878 | 667.4876 |
| 680.4989 | 680.4992 |
| 682.4572 | 682.4582 |
| 722.5477 | 722.5472 |
| 725.5575 | 725.5572 |
| 726.5607 | 726.5608 |
| 729.4839 | 729.4837 |
| 731.4991 | 731.4999 |
| 732.5023 | 732.5023 |
| 739.573 | 739.5731 |
| 747.4748 | 747.4755 |
| 748.5643 | 748.5626 |
| 750.5789 | 750.5786 |
| 751.5809 | 751.5815 |
| 757.5154 | 757.5146 |
| 778.6089 | 778.6092 |
| 779.6134 | 779.6137 |
| 781.5928 | 781.5928 |
| 781.6197 | 781.62 |
| 788.5574 | 788.5569 |
| 789.5603 | 789.5597 |
| 790.5727 | 790.5733 |
| 794.6056 | 794.6059 |
| 795.6089 | 795.608 |
| 795.6348 | 795.6352 |
| 801.6853 | 801.4227 |
| 807.6356 | 807.6358 |
| 809.6515 | 809.6518 |
| 810.6536 | 810.6544 |
| 823.6089 | 823.6094 |
| 823.6674 | 823.6677 |
| 825.6249 | 825.6263 |
| 826.6288 | 826.6291 |
| 833.651 | 833.6519 |
| 835.6671 | 835.6671 |
| 836.6697 | 836.6712 |
| 837.682 | 837.6834 |
| 838.6858 | 838.6865 |
| 839.64 | 839.641 |
| 840.644 | 840.6448 |
| 851.6406 | 851.641 |
| 852.6449 | 852.6452 |
| 853.6571 | 853.6575 |
| 854.6604 | 854.6597 |

**Supplemental Table S2.** This table shows the detected peaks from the different matrix conditions of the main electric eel tissue localized to the lumen regions. The values shown below are detected in both biological replicates of the electric eels, matched within 5ppm of the peak detected, and different peaks detected between the two matrix conditions: CHCA and DAN matrix.

**Lumen Regions of Tissue**

| CHCA Matrix | | DAN Matrix | |
|---|---|---|---|
| Bioreplicate 1 m/z values | Bioreplicate 2 m/z value | Bioreplicate 1 m/z values | Bioreplicate 2 m/z values |
| 535.9934 | 535.9933 | 689.3254 | 689.3249 |
| 543.982 | 543.9811 | 728.3506 | 728.3495 |
| 559.9546 | 559.9544 | 856.5016 | 856.5009 |
| 569.0284 | 569.0262 | | |
| 585.0013 | 585.0015 | | |
| 592.0362 | 592.0344 | | |
| 601.9883 | 601.9909 | | |
| 602.9524 | 602.9525 | | |
| 608.0095 | 608.008 | | |
| 609.0122 | 609.0118 | | |
| 618.9273 | 618.925 | | |
| 623.9819 | 623.9847 | | |
| 624.9877 | 624.9866 | | |
| 626.0257 | 626.0246 | | |
| 632.3702 | 631.0031 | | |
| 641.9991 | 641.9982 | | |
| 645.4086 | 646.978 | | |
| 650.0439 | 650.0419 | | |
| 656.064 | 656.063 | | |
| 657.0672 | 657.0663 | | |
| 666.0192 | 666.018 | | |
| 672.0379 | 672.0364 | | |
| 673.0414 | 673.0406 | | |
| 697.4786 | 697.4778 | | |
| 698.4822 | 698.4813 | | |
| 770.5682 | 770.567 | | |
| 813.017 | 813.0148 | | |
| 828.9924 | 828.9899 | | |
| 836.5179 | 836.5163 | | |
| 861.0691 | 861.068 | | |
| 862.0718 | 862.0708 | | |
| 877.0451 | 877.043 | | |
| 878.0472 | 878.0456 | | |

**Supplemental Table S3:** This table shows the detected peaks from the different matrix conditions of the main electric eel tissue localized to the electrocyte regions. The values shown below are detected in both biological replicates of the electric eels, matched within 5ppm of the peak detected, and different peaks detected between the two matrix conditions: CHCA and DAN matrix.

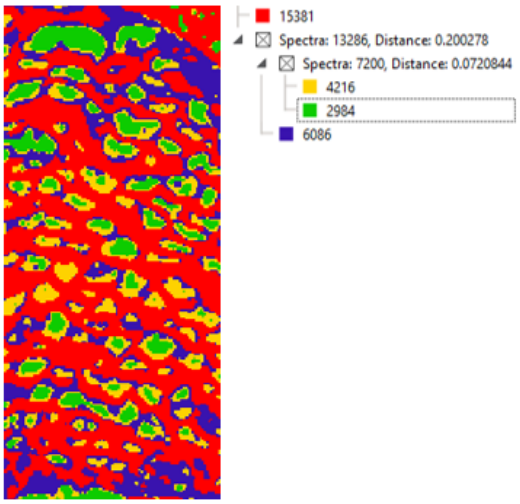
**Electrocyte Regions of Tissue**

| CHCA Matrix | | DAN Matrix | |
|---|---|---|---|
| Bioreplicate 1 m/z values | Bioreplicate 2 m/z value | Bioreplicate 1 m/z values | Bioreplicate 2 m/z values |
| 503.9482 | 503.9478 | 619.4697 | 619.4697 |
| 505.9525 | 505.9526 | 621.485 | 621.4853 |
| 507.9201 | 507.9201 | 622.4881 | 622.4882 |
| 541.9042 | 541.9045 | 623.5033 | 623.5031 |
| 542.9094 | 542.9084 | 624.5074 | 624.5066 |
| 543.9034 | 543.9024 | 645.4853 | 645.4853 |
| 565.0277 | 565.0261 | 646.4882 | 646.4887 |
| 581.9782 | 581.978 | 647.5006 | 647.5012 |
| 583.9163 | 583.9145 | 650.5207 | 650.5206 |
| 597.9531 | 597.9514 | 651.5342 | 651.5343 |
| 613.9273 | 613.9256 | 673.517 | 673.5168 |
| 643.8541 | 643.8531 | 694.5153 | 694.5149 |
| 655.9384 | 655.936 | 743.4622 | 743.462 |
| 717.9142 | 717.9125 | 745.4784 | 745.4774 |
| 745.802 | 745.8016 | 761.453 | 761.4509 |
| 802.9346 | 802.9325 | 769.4779 | 769.478 |
| 854.4876 | 854.4863 | 770.4821 | 770.4816 |
| 868.4668 | 868.4657 | 785.4522 | 785.4513 |
| 869.471 | 869.4693 | 786.4555 | 786.455 |
| 912.4562 | 912.4544 | 787.47 | 787.4691 |
| | | 788.4721 | 788.4706 |
| | | 793.4761 | 793.4754 |
| | | 796.5841 | 796.5824 |
| | | 797.5287 | 797.527 |
| | | 803.5385 | 803.5383 |
| | | 804.5513 | 804.551 |
| | | 805.5548 | 805.5548 |
| | | 813.4879 | 813.484 |
| | | 820.5849 | 820.5851 |
| | | 821.5293 | 821.5281 |
| | | 822.4814 | 822.4815 |
| | | 824.4605 | 824.4602 |
| | | 829.5554 | 829.5548 |
| | | 830.5684 | 830.5686 |
| | | 832.5831 | 832.5836 |
| | | 838.5127 | 838.5119 |
| | | 844.4666 | 844.4655 |
| | | 844.5256 | 844.525 |
| | | 845.5295 | 845.5286 |
| | | 846.5429 | 846.5437 |
| | | 847.5455 | 847.5451 |
| | | 852.4922 | 852.4917 |
| | | 857.5863 | 857.5861 |
| | | 868.4661 | 868.4657 |
| | | 872.557 | 872.5569 |
| | | 873.5611 | 873.5601 |

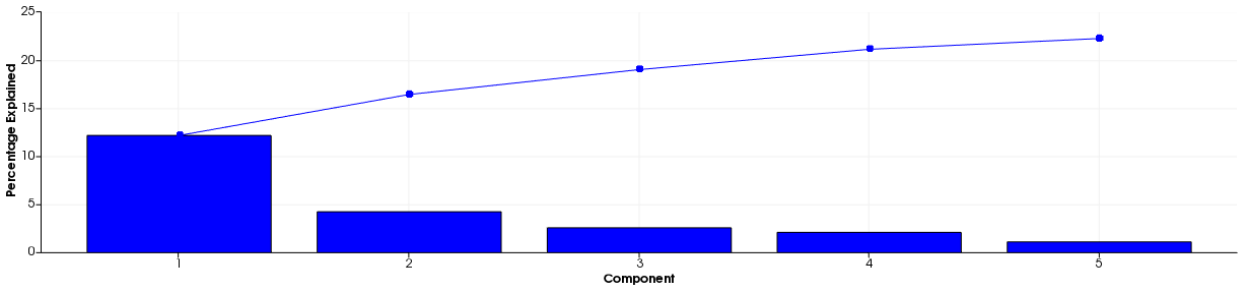
**Supplemental Table S4:** Tentative Identification of Lipids Localized to the Insulator Regions Based on Accurate Mass Matching to Metlin Database.

| **Tentative Lipid Identification** | **Ion** | **Molecular Formula** | **Detected *m/z*** | **Δ ppm** | **Detected *m/z*** | **Δ ppm** |
|---|---|---|---|---|---|---|
| Cerebroside | $[M+Na]^+$ | C44H84NO6P | 776.594 | 1 | 776.5928 | 0 |
| Cerebroside | $[M+H]^+$ | C46H78NO7P | 788.5574 | 1 | 788.5569 | 2 |
| Cerebroside | $[M+Na]^+$ | C42H80NO6P | 748.5643 | 3 | 748.5626 | 1 |
| Cerebroside | $[M+NH_4]^+$ | C48H88NO6P | 823.6674 | 1 | 823.6677 | 1 |
| Sulfatide | $[M+H]^+$ | C40H77NO12S | 796.5268 | 2 | 796.5254 | 1 |
| Ceramide-Phosphate | $[M+K]^+$ | C42H82NO6P | 766.5512 | 0 | 766.5504 | 0 |
| Ceramide-Phosphate | $[M+Na]^+$ | C42H82NO6P | 750.5789 | 2 | 750.5786 | 2 |
| Cholesteryl ester | $[M+NH_4]^+$ | C42H72O2 | 626.5854 | 2 | 626.5855 | 2 |
| Cholesteryl ester | $[M+NH_4]^+$ | C43H74O2 | 640.601 | 2 | 640.6003 | 3 |
| Cholesteryl ester | $[M+NH_4]^+$ | C44H76O2 | 654.6155 | 4 | 654.6167 | 2 |
| Sphingomyelin | $[M+K]^+$ | C39H79N2O6P | 741.5313 | 0 | 741.5312 | 0 |
| Sphingomyelin | $[M+Na]^+$ | C43H87N2O6P | 781.6197 | 0 | 781.62 | 0 |
| Sphingomyelin | $[M+Na]^+$ | C40H81N2O6P | 739.573 | 0 | 739.5731 | 0 |
| Sphingomyelin | $[M+Na]^+$ | C47H91N2O6P | 833.651 | 0 | 833.6519 | 1 |
| Sphingomyelin | $[M+Na]^+$ | C47H93N2O6P | 835.6671 | 0 | 835.6671 | 0 |
| Sphingomyelin | $[M+Na]^+$ | C47H95N2O6P | 837.682 | 0 | 837.6834 | 1 |
| Sphingomyelin | $[M+K]^+$ | C45H91N2O6P | 825.6259 | 1 | 825.6256 | 1 |
| Sphingomyelin | $[M+K]^+$ | C46H93N2O6P | 839.6403 | 0 | 839.6403 | 0 |
| Sphingomyelin | $[M+K]^+$ | C47H93N2O6P | 851.6424 | 2 | 851.6408 | 0 |
| Sphingomyelin | $[M+K]^+$ | C47H95N2O6P | 853.6574 | 1 | 853.6568 | 1 |
| Sphing0myelin | $[M+Na]^+$ | C39H79N2O6P | 725.5574 | 0 | 725.5571 | 0 |
| Sphingomyelin | $[M+K]^+$ | C39H79N2O6P | 741.5313 | 0 | 741.5312 | 0 |
| Diacylglycerol | $[M+K]^+$ | C45H70O5 | 729.4839 | 2 | 729.4837 | 2 |
| Diacylglycerol | $[M+K]^+$ | C45H72O5 | 731.4991 | 2 | 731.4999 | 0 |
| Diacylglycerol | $[M+K]^+$ | C47H74O5 | 757.5154 | 1 | 757.5146 | 2 |
| Triacyglycerol | $[M+Na]^+$ | C48H84O6 | 779.6134 | 3 | 779.6137 | 2 |
| Phosphatidylglycerol | $[M+H]^+$ | C35H71O9P | 667.4878 | 4 | 667.4876 | 4 |
| Phosphatidylglycerol | $[M+NH_4]^+$ | C46H93O9P | 838.6858 | 4 | 838.6865 | 4 |
| Phosphatidylglycerol | $[M+Na]^+$ | C44H85O10P | 827.5771 | 0 | 827.576 | 1 |
| Phosphatidylinositol | $[M+NH_4]^+$ | C34H68NO13P | 747.4748 | 2 | 747.4755 | 1 |
| Phosphatidylcholine | $[M+H]^+$ | C46H80NO7P | 790.5727 | 2 | 790.5733 | 1 |
| Phosphatidylcholine | $[M+H]^+$ | C46H84NO7P | 794.6056 | 0 | 794.6059 | 0 |
| Phosphatidylcholine | $[M+Na]^+$ | C46H92NO8P | 840.6440 | 1 | 840.6448 | 3 |
| Phosphatidic acid | $[M+K]^+$ | C41H77O8P | 767.5002 | 1 | 767.4989 | 0 |
| Phosphatidic acid | $[M+K]^+$ | C43H75O8P | 789.4840 | 2 | 789.4831 | 0 |

**Supplemental Figures**

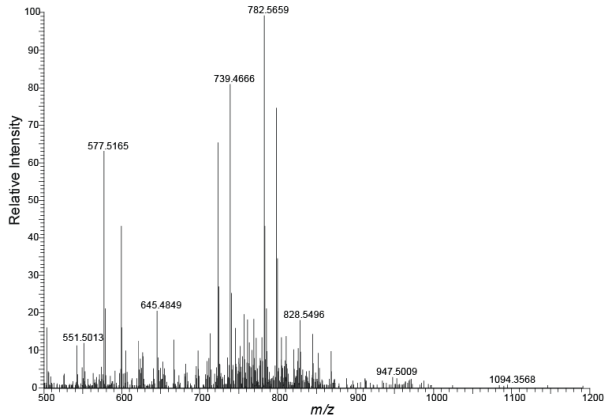



**Supplemental Figure S1.** High spatial resolution analysis and statistical segmentation analysis reveals important lining of the Electrocyte Lumen. Segmentation using bisecting k-means clustering of main electric organ was performed to show key features of the tissue. This section of the electric organ was sectioned at 20µm, coated with 40mg/mL DHB and then analyzed on Bruker Rapiflex from m/z 500-1200. The spatial resolution reveals distinct chemical composition of the inner lining of the lumen. This high spatial resolution may show a small epithelial boundary that are found in other types of electric fishes.

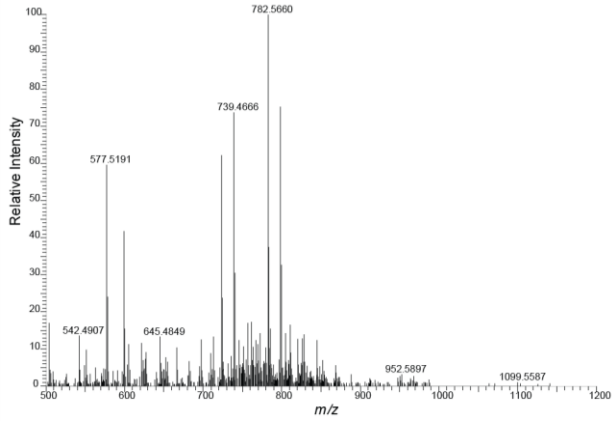



**Supplemental Figure S2**. Principal Component Analysis Supplemental plot of percentage of data explained by the analysis.

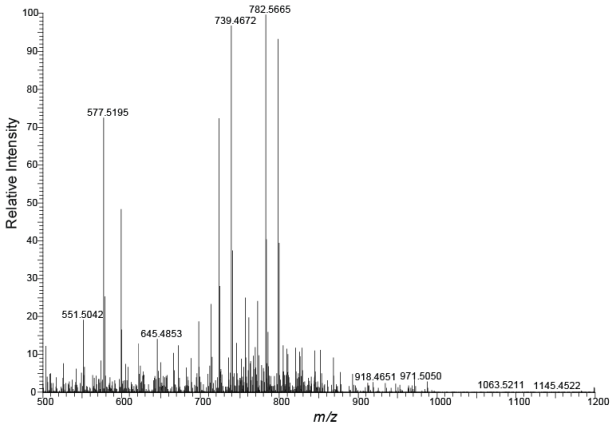
Insulator Spectra:



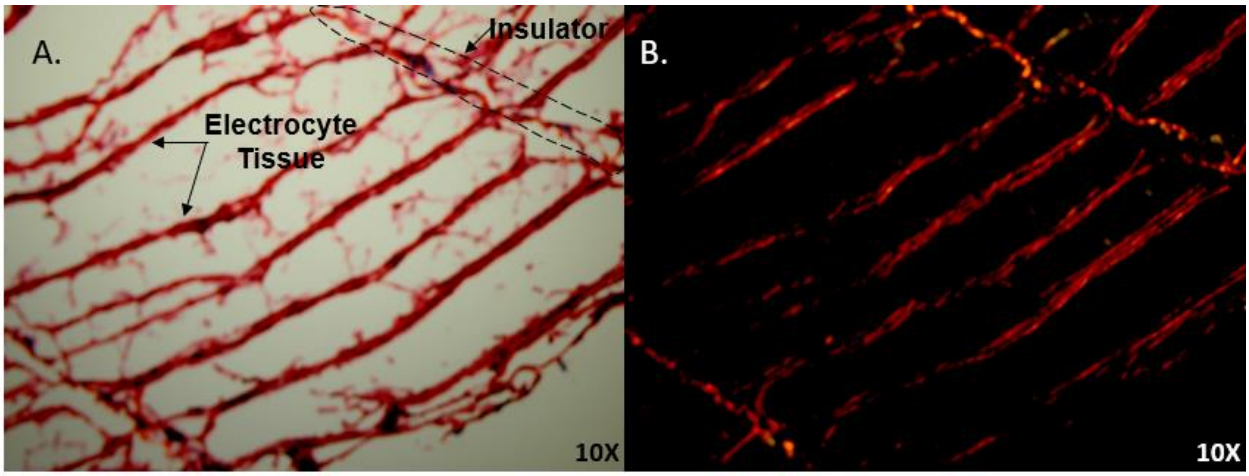

Electrocyte Spectra:




Lumen Spectra:



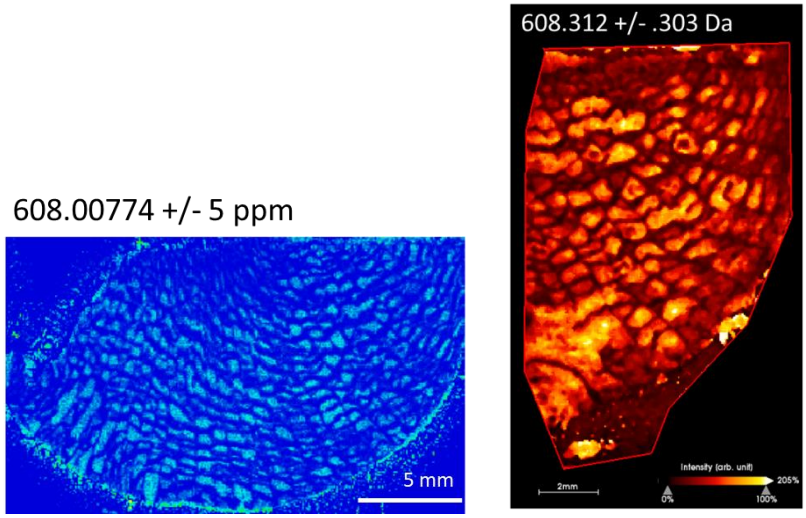

**Supplemental Figure S3**. Example spectra for electric eel main electric organ with CHCA as the matrix.

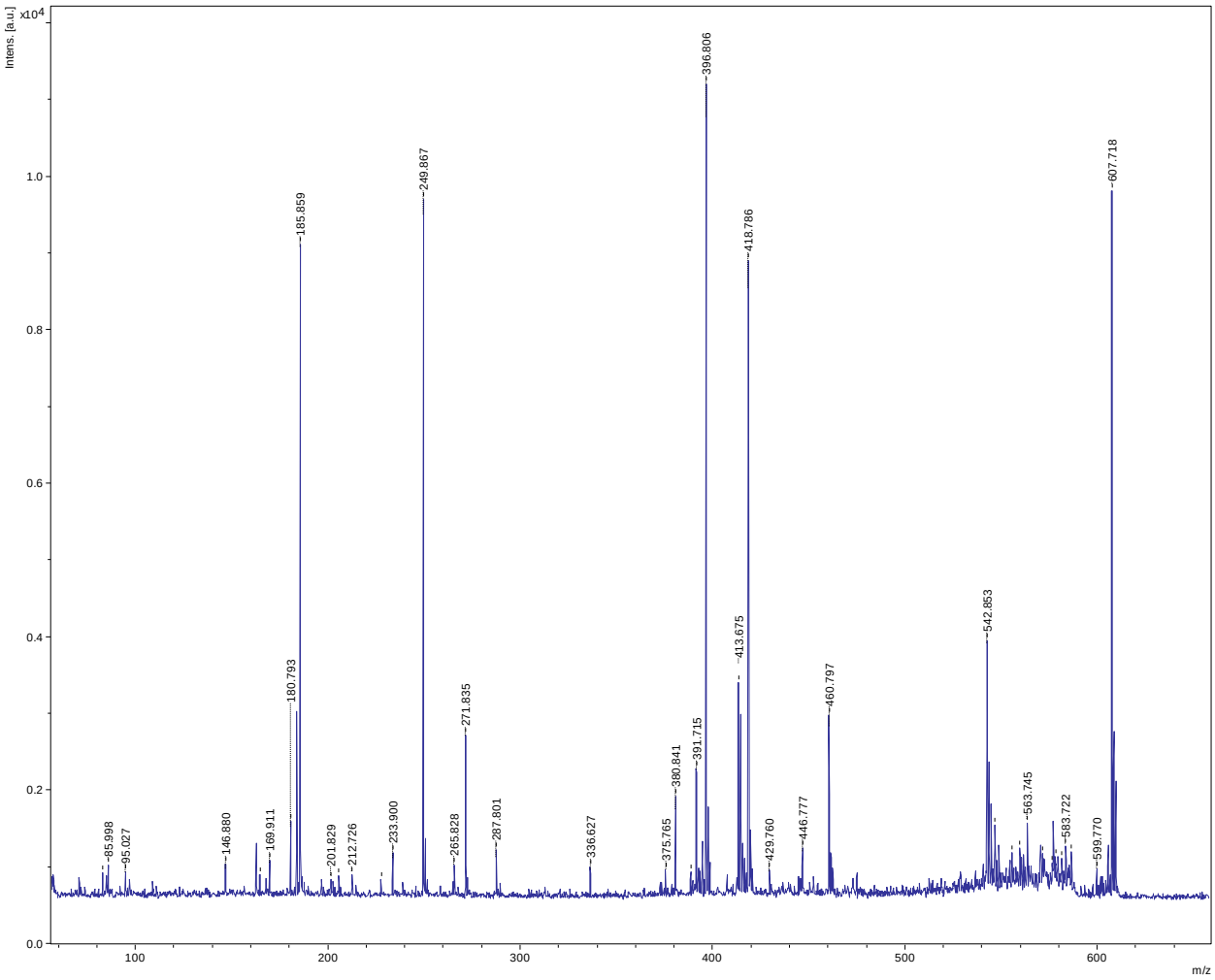



**Supplemental Figure S4.** Picrosirius Red Stain of the main electric organ visualized **(A)** without polarizer and **(B)** with polarizer

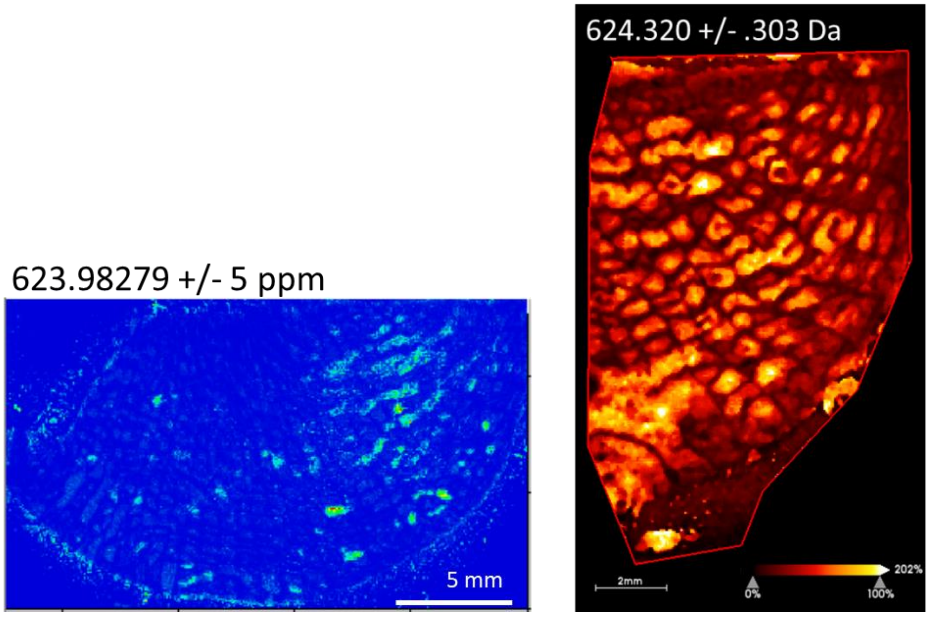
**Supplemental Figure S5**. Targets with MSMS from the Rapiflex.

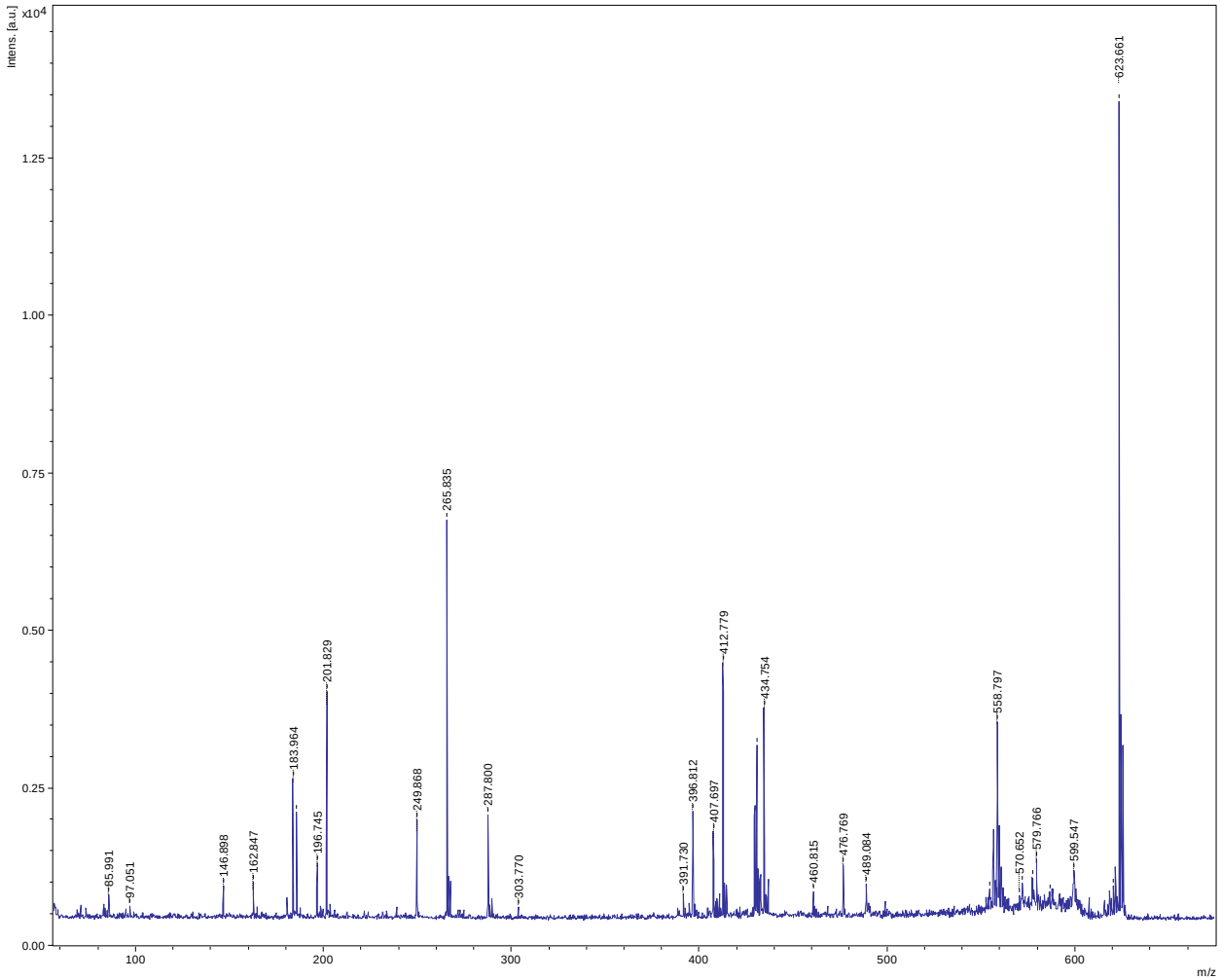
A



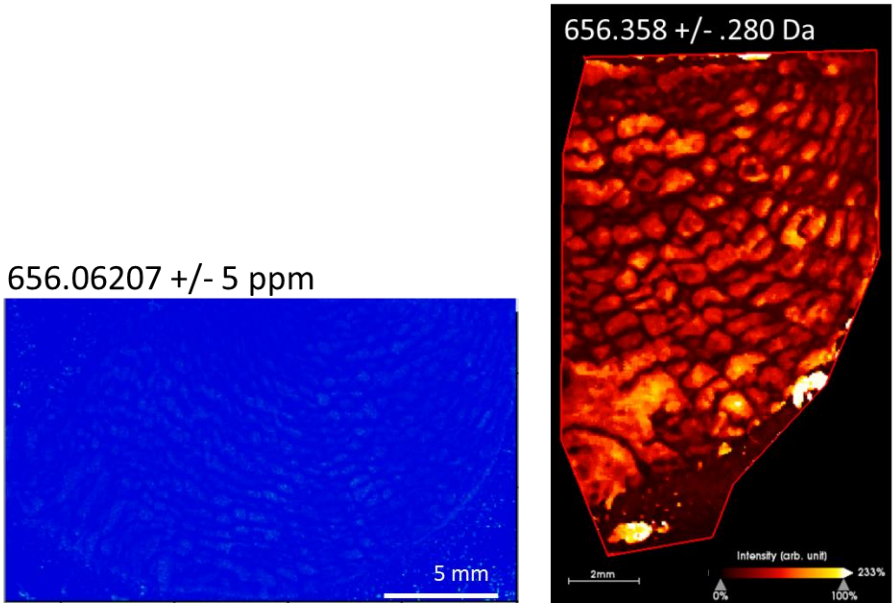



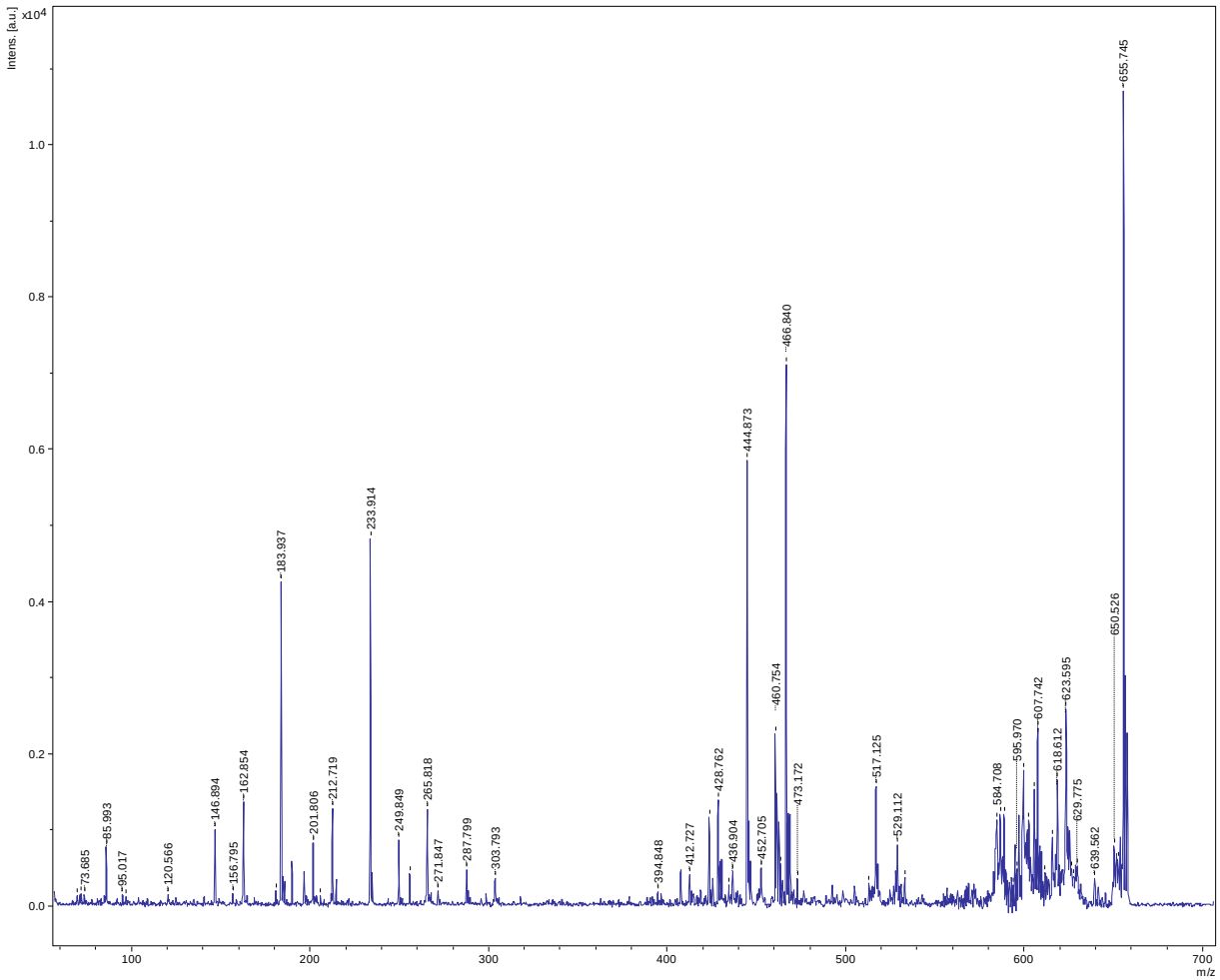
B


623.98279 +/- 5 ppm
5 mm
624.320 +/- .303 Da
2mm
0%
100%
202%

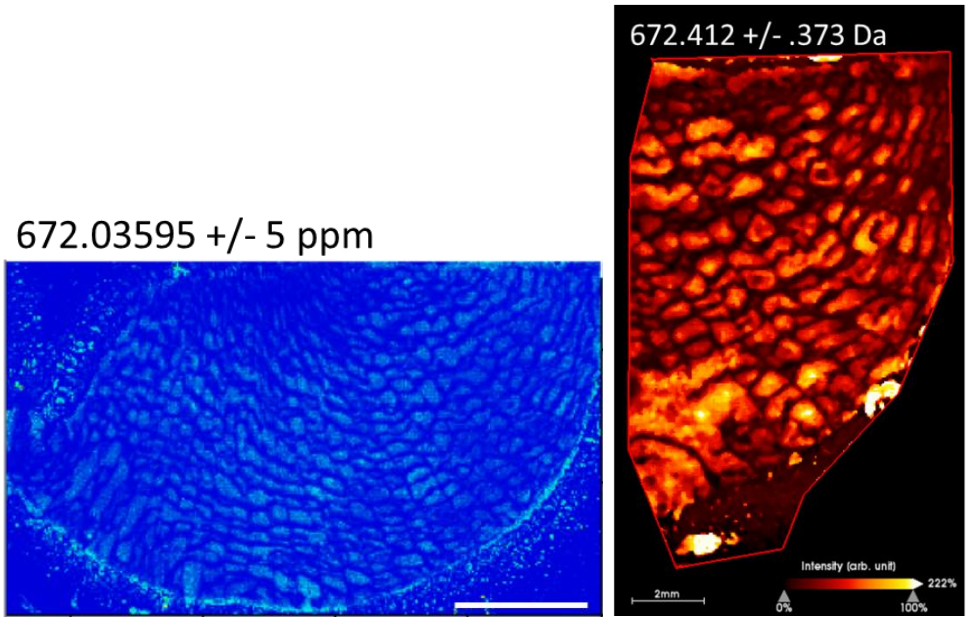

Intens. [a.u.]
x10^4
1.25
1.00
0.75
0.50
0.25
0.00
100
200
300
400
500
600
m/z
85.991
97.051
146.898
162.847
183.964
196.745
201.829
249.868
265.835
287.800
303.770
391.730
396.812
407.697
412.779
434.754
460.815
476.769
489.084
558.797
570.652
579.766
599.547
623.661

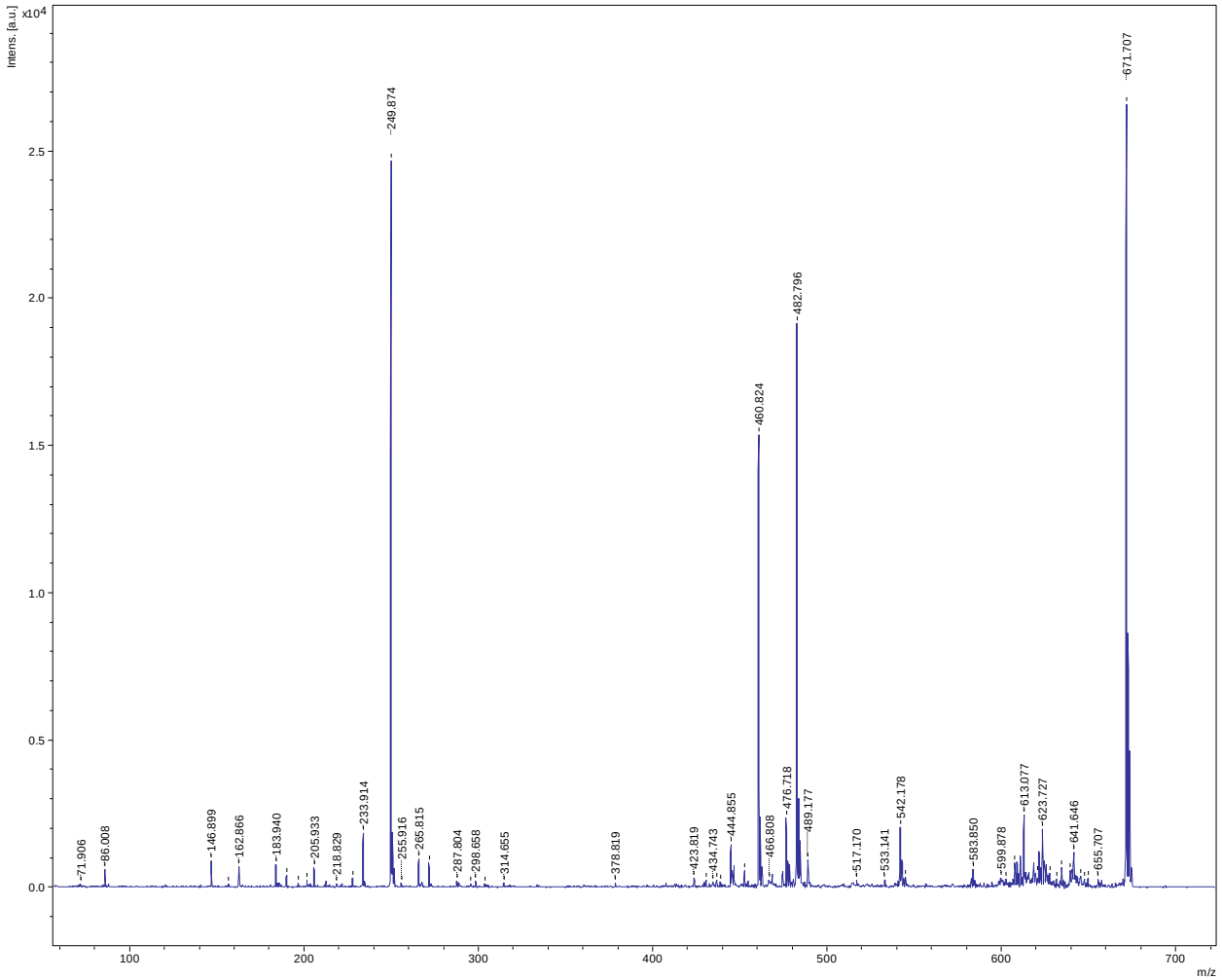
C


656.06207 +/- 5 ppm
5 mm
656.358 +/- .280 Da
Intensity (arb. unit)
233%
0%
100%
2mm


Intens. [a.u.]
x10^4
1.0
0.8
0.6
0.4
0.2
0.0
100
200
300
400
500
600
700
m/z
655.745
650.526
466.840
444.873
460.754
623.595
607.742
618.612
595.970
584.708
517.125
473.172
529.112
428.762
233.914
183.937
212.719
249.849
265.818
201.806
162.854
146.894
85.993
73.685
95.017
120.566
156.795
271.847
287.799
303.793
394.848
412.727
436.904
452.705
629.775
639.562

D


672.03595 +/- 5 ppm
672.412 +/- .373 Da
Intensity (arb. unit)
222%
0%
100%
2mm


Intens. [a.u.]
x10^4
m/z
671.707
249.874
482.796
460.824
476.718
489.177
444.855
542.178
613.077
623.727
641.646
233.914
255.916
265.815
146.899
162.866
183.940
205.933
218.829
71.906
86.008
287.804
298.658
314.655
378.819
423.819
434.743
466.808
517.170
533.141
583.850
599.878
655.707