

PEPTIDE-MEDIATED SIGNALING PATHWAYS IN THE NEMATODE *ASCARIS SUUM*

By

Jennifer J. Knickelbine

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Zoology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Date of final oral examination: May 15, 2017

The dissertation is approved by the following members of the Final Oral Committee:

Antony O. W. Stretton, Professor, Zoology

Mary C. Halloran, Professor, Zoology

Seth S. Blair, Professor, Zoology

Gail A. Robertson, Professor, Neuroscience

Laura J. Knoll, Associate Professor, Medical Microbiology and Immunology

ACKNOWLEDGEMENTS

In preparing this dissertation, I have been supported by a group of incredible people. My husband has been my lifeline through all the grant applications, Great Floods, cold room/freezer breakdowns, late night/weekend/holiday experiments and trips to the abattoir. When I was exhausted and frustrated, he reminded me that I could handle whatever problem came my way, and helped me keep things in perspective. He also had a joke, quip, or silly reference for nearly occasion. Jacob, I am so lucky to have you in my life. My family has also been very supportive, with phone calls, packages, and YouTube videos when I needed them most. Mom and Dad were always eager to hear about how things were going in the lab, and Matt, Maureen, and Emma were quick to remind me to take a break and do something fun, and made sure I was up to date on all the family pets' antics. My in-laws Therese, Laura, and Megan were always up for a hike or a trip to the zoo when I needed to get out of the lab for a bit. I have also relied heavily on my core group of friends for daily reminders of all things happy. Ashley, Kati, Laura, and Nicole, you are amazing, and I don't know what I would have done without you. Thank you all so much!

I am also extremely grateful to my advisor, Tony Stretton, who was the biggest influence in my decision to go to graduate school and has remained an excellent mentor through it all. He was quick to give me a lot of freedom in the lab, and his trust and confidence in my abilities as a scientist and a teacher have allowed me to grow so much in my time here. He and his wife Philippa welcomed me into their home and their lives, and I am deeply grateful for their openness and generosity. Philippa has also put hours into editing our manuscripts and grant applications, and she is the best person to have at a practice talk because of the unique

perspective she brings to the table. Tony and Philippa, I cannot thank you enough for all your support over the years.

One of the best parts of working in the Stretton laboratory has been the people in the lab group. Catharine Reinitz, India Viola, Kari Andersen, Jessie Jarecki, and Chris Konop were all very supportive when I joined the lab, and helped me learn new techniques and talk about my work. I always enjoyed running into Andy Miller, Anna Merg, Alison Crane, Megan Ramaker, and Molly Sygulla when our class schedules allowed. Chris Konop has been a mentor and friend throughout my time with the lab, and together we answered the major question that this lab has been working on for two decades. Then, as we branched out to neuropeptide receptors and new techniques, we spent a lot of time bouncing ideas around or troubleshooting over morning coffee. While he will probably always joke about me being “his undergrad”, we have had a very effective partnership in science, and I am very grateful for that. I was also lucky to mentor 8 undergraduates in the lab, some of whom contributed to this dissertation, and are acknowledged as authors. Colin Wruck, Becca Garry, Colette Rogers, Heather Malecki, Samantha Miller, Shirley Hu, Gretchen Westphal, and Gwen DeBoth, working with you has helped me become a better teacher and scientist, and I am very glad to see so many of you doing well after leaving the lab.

I have also been very fortunate in collaborations and interactions with other researchers. Martha Vestling in Chemistry was instrumental in performing the mass spectrometry experiments in Chapters 2 and 4. Jacob Hornick of James Broach’s laboratory at Penn State provided me with the yeast strains needed for Chapter 5, and helped me set up the assay. Jennifer Yao of Sam Gellman’s laboratory in Chemistry collaborated with us to design and

synthesize peptide analogs of two bioactive peptides for Chapter 5. It has been a pleasure working with all of you, and I am grateful for all you have taught me.

My committee has also been hugely influential. Mary Halloran and Seth Blair have been advising me since I started graduate school as a master's student, and have supported my transition to the Ph. D. program, as well as personal fellowship and grant applications. You have seen me at my most timid, and helped me grow into someone who is far more comfortable and confident in putting my work out there. Gail Robertson has always encouraged me to branch out and reach high in whatever I am working on, and is always excited to hear how things are going. Laura Knoll always made sure to ask how I was doing at PVB seminars, helped me gain so much perspective on how my work affects the fields of parasitology and global health, and had lots of helpful advice for career and professional development. All of you have taken time out of your insanely busy schedules to meet with me and offer thoughtful advice, and I am so grateful for your support!

Finally, while the work described here was funded in part by the NIH and the NSF, this dissertation would not have been possible without the support of the UW-Madison Graduate School, Department of Zoology, Parasitology and Vector Biology Training Program, and the Graduate Women in Science. I was very fortunate to receive stipend support, research supply funding, and travel funding from these groups while our laboratory was applying for external funding, and this support allowed me to complete these projects and generate preliminary data for our laboratory's grant applications. The Department of Zoology has also been very good to our laboratory in my time here. Jeff Hardin, Peggy Nowicki, Kelly Layton Kaap, Carol Cooley, Julie Lindsey, Lynn Thiele, and Zach Rusch have been instrumental in keeping our lab running

and hiring specimen collectors. Joan Ersland and Kayla Pelland have helped me navigate the formal requirements of the Graduate School. Diane Mockrud, Roger Larson, Duane Barnes, and Mike Mallone helped keep our laboratory stocked and in working order. Bill Feeny's artistic eye helped us turn data into publication-quality figures. And finally, I am very grateful to Jenya Grinblat, Kurt Amann, Stephen Gammie, and Seth Blair for allowing me to use equipment in their laboratories. This department has been incredibly supportive and collaborative, and I am very lucky to have found a home here for the past 6 years. Thank you all so much!

TABLE OF CONTENTS

Introduction	1
Chapter 1: Complexity in a “Simple” System: A Review of Neuropeptide Signaling in Nematodes	
Introduction	4
Figures	17
References	18
Chapter 2: Mass Spectrometry of Single GABAergic Somatic Motorneurons Identifies a Novel Inhibitory Peptide, As-NLP-22, in the Nematode <i>Ascaris suum</i>	
Introduction	28
Methods	32
Results and Discussion	42
Conclusions and Future Directions	55
Figures	57
References	63
Chapter 3: Neuropeptides Related to As-NLP-22 Have Different Effects on <i>A. suum</i> Muscle Contraction	
Introduction	70
Methods	71
Results and Discussion	72
Conclusions and Future Directions	83
Figures	86
References	100
Chapter 4: Different Neuropeptides Are Expressed in Different Functional Subsets of Cholinergic Excitatory Motorneurons in the Nematode <i>Ascaris suum</i>	
Introduction	102
Methods	107
Results and Discussion	115
Conclusions and Future Directions	127
Figures	129
References	140

Chapter 5: Identification of *A. suum* Neuropeptide Receptors and their Endogenous Ligands by Heterologous Expression and Reverse Pharmacology

Introduction.....	147
Methods.....	150
Results and Discussion.....	155
Conclusions and Future Directions.....	180
Figures.....	184
References.....	226
Summary and Conclusions.....	231
Appendix 1: Supplementary Figures for Chapter 2.....	233
Appendix 2: Supplementary Figures for Chapter 4.....	237
Appendix 3: AF2 Effect on Muscle Contraction.....	249
Appendix 4: List of Abbreviations and Nematode Species.....	250

ABSTRACT

Parasitic nematodes remain a significant burden to human health and the global economy. Over 1 billion people worldwide are infected with the intestinal roundworm *Ascaris lumbricoides*, making Ascariasis the most prevalent of the 17 Major Neglected Tropical Diseases. Further, a large proportion of livestock are infected by the close relative *Ascaris suum*. Treatments have largely focused on affecting the worm's locomotion, particularly through the highly-conserved cholinergic and GABAergic pathways. However, the widespread use of these drugs increases the risk for genetic drug resistance, and new targets are urgently needed. This dissertation reports the identification of neuropeptides expressed by the motorneurons of *A. suum* by single-cell mass spectrometry and *in situ* hybridization, and characterizes their effects on muscle contraction and locomotion. We found that the motorneurons can be grouped into 3 classes based on their neuropeptide expression profiles, with anteriorly-projecting excitatory motorneurons (responsible for forward movement) expressing peptides from the *afp-14* and *nlp-21* transcripts, posteriorly-projecting excitatory motorneurons (controlling backward movement) expressing *afp-1* and *afp-4* peptides, and inhibitory motorneurons expressing a single peptide from the *As-nlp-22* transcript. We also report the identification of 7 neuropeptide G protein-coupled receptors with ties to the locomotory circuit, and identify neuropeptide ligands for 3 of these by heterologous expression in yeast. The expression of these neuropeptides and receptors differs from that of homologs in *C. elegans*, highlighting the need for additional studies in parasitic nematodes. This work significantly advances our understanding of neuropeptide signaling in nematode locomotion, and identifies several potential targets for therapeutic development.

INTRODUCTION

The phylum Nematoda is one of the most successful extant phyla, with a huge range of species occupying nearly every ecological niche. It includes free-living species, both terrestrial and marine, as well as notorious parasites of animals and plants, which remain a huge burden to human health and the global economy. Over 1.2 billion people worldwide are infected by the intestinal roundworm *Ascaris lumbricoides*, which causes malnutrition, intestinal obstruction, breathing difficulties, delayed growth and cognitive development, and even death [1, 2]. The World Health Organization estimates that this parasite causes 60,000 deaths/year, mainly in children, making Ascariasis the most prevalent of the 17 Major Neglected Tropical Diseases, which disproportionately affect the world's poor.

A closely-related species, *Ascaris suum*, infects pigs and is responsible for at least \$60.1 million in economic losses each year in the US alone due to delayed growth rates [3, 4], in addition to increased risk of co-infection and condemnation of organs and meat at slaughter. While it is difficult to determine the total prevalence, which can vary according to farming practices, estimated infection rates are as high as 17-35% in Scandinavian countries [5] (except for Denmark, which had 76% infection rates in 2010 [6]), 17% in Canada [7], and 40% in Uganda [8]. Infection rates are even higher on farms that use free-range or organic husbandry, with 50% and 60-73% of pigs, respectively, infected with *A. suum* on these farms in The Netherlands and Austria [9, 10]. As the demand for free-range and organic products continues to rise, unfortunately, so does the prevalence of parasites which have co-evolved with their hosts for millennia.

Treatments for *Ascaris* parasites have largely focused on affecting the worm's locomotion, often through cholinergic or GABAergic neurotransmission in the motor nervous system. Since these worms do not attach to the walls of the intestine, if they can be paralyzed, the adults can be expelled from the small intestine by natural peristalsis, rather than invasive surgical procedures. However, genetic drug resistance can quickly render whole classes of drugs useless, and given the widespread use of these compounds worldwide, it is vital that we identify new systems for future drug development. This requires a detailed understanding of the circuitry involved in nematode locomotory behavior, as well as relevant chemical signals and their receptors.

This dissertation reports the identification of neuropeptides expressed by the motorneurons of *Ascaris suum*, many of which had profound effects on nematode locomotion, thus achieving the primary goal of our laboratory for the last 25 years. These peptides do not seem to be conserved in vertebrates, which could be advantageous for drug development since compounds targeting these systems may not have such severe side effects for the host. This dissertation also reports the expression and peptide-binding profiles of 3 neuropeptide receptors with ties to the motor circuit, marking the first deorphanization of an *A. suum* neuropeptide receptor. Importantly, the expression of these neuropeptides and receptors differed from that of homologs in the free-living nematode *C. elegans*, stressing the importance of direct studies in nematode parasites. These experiments lay the foundation for future therapeutic development, and potentially, a new class of anti-nematode drugs.

REFERENCES

1. de Silva, N. R., Brooker, S., Hotez, P. J., Montresor, A., Engels, D., Savioli, L. (2003). Soil-transmitted helminth infections: updating the global picture. *Trends Parasitol.* 19:547-51.
2. Keiser, J., Utzinger, J. (2008). Efficacy of current drugs against soil-transmitted helminth infections: systematic review and meta-analysis. *JAMA.* 299:1937-48.
3. Thamsborg, S. M., Nejsum, P., Mejer, H. (2013). Impact of *Ascaris suum* in livestock. 363-81.
4. Stewart, T. B. (2001). Economics of endoparasitism of pigs. *Pigs News and Information.* 22:29N-30N.
5. Roepstorff, A., Nilsson, O., Oksanen, A., Gjerde, B., Richter, S. H., Ortenberg, E., Christensson, D., Martinsson, K. B., Bartlett, P. C., Nansen, P., Eriksen, L., Helle, O., Nikander, S., Larsen, K. (1998). Intestinal parasites in swine in the Nordic countries: prevalence and geographical distribution. *Vet Parasitol.* 76:305-19.
6. Haugegaard, J. (2010). Prevalence of nematodes in Danish industrialized sow farms with loose housed sows in dynamic groups. *Vet Parasitol.* 168:156-9.
7. Wagner, B., Polley, L. (1997). *Ascaris suum* prevalence and intensity: an abattoir survey of market hogs in Saskatchewan. *Vet Parasitol.* 73:309-13.
8. Nissen, S., Poulsen, I. H., Nejsum, P., Olsen, A., Roepstorff, A., Rubaire-Akiiki, C., Thamsborg, S. M. (2011). Prevalence of gastrointestinal nematodes in growing pigs in Kabale District in Uganda. *Trop Anim Health Prod.* 43:567-72.
9. Baumgartner, J., Leeb, T., Gruber, T., Tiefenbacher, R. (2003). Husbandry and animal health on organic pig farms in Austria. *Anim Welfare.* 12:631-5.
10. Eijck, I. A., Borgsteede, F. H. (2005). A survey of gastrointestinal pig parasites on free-range, organic and conventional pig farms in The Netherlands. *Vet Res Commun.* 29:407-14.

CHAPTER 1

Complexity in a “Simple” System: A Review of Neuropeptide Signaling in Nematodes

Jennifer J. Knickelbine and Antony O. W. Stretton

Introduction

The aim of this chapter is to describe the state of the field of nematode neuropeptide signaling, and to set the context for the work that makes up the body of this dissertation. While the neuroanatomy and classical transmitter expression of *C. elegans* and *A. suum* were documented decades ago, the role of neuropeptides has remained largely unknown. This dissertation reports that the *A. suum* motoneurons can be grouped into 3 categories based on neuropeptide expression, achieving the goal of our laboratory for the last 25 years, and pairs some of these to neuropeptide-binding GPCRs in the motoneurons and muscle cells. We also identify important differences in neuropeptide expression and signaling between *A. suum* and *C. elegans*. This work is highly relevant to anti-nematode therapeutic development, and could lay the foundation for development of a new class of anthelmintics.

Nematode Neurobiology

Nematodes are touted as model organisms for their “simple” nervous systems, largely because their neurons are small in number and have very simple morphology compared to those of vertebrates. Two of the most widely studied nematodes, free-living *Caenorhabditis elegans* and the parasite *Ascaris suum*, have only 302 [1] and 298 [2] neurons, respectively, and their anatomy is so similar (apart from size) that they share naming conventions. This is

remarkable, considering that these species are estimated to have shared a last common ancestor 550 million years ago [3], and inhabit vastly different ecological niches, one free-living and the other parasitic. There is also an enormous difference in their sizes (*C. elegans* grows to 1 mm in length, while *A. suum* can be up to 45 cm long!). The expression of classical neurotransmitters is also conserved, with the excitatory motoneurons producing acetylcholine (ACh) and the inhibitory motoneurons producing γ -aminobutyric acid (GABA) [4-8]. In *A. suum*, each class of somatic motoneuron has also been electrophysiologically profiled [9, 10].

With all this information available, the next logical step was to construct a functional circuit to describe locomotory behavior. *Ascaris suum* moves around in the host intestine by forming pseudo-sinusoidal waveforms that brace against the intestinal walls. These waves propagate anteriorly or posteriorly to control forward or backward movement, respectively. The functional circuit proposed by Stretton *et al.* (1985) predicts how the alternating contraction and relaxation of dorsal and ventral muscle fields, responsible for locomotion, is generated (Figure 1, [11]). For example, a dorsal excitatory (DE) motoneuron releases ACh onto dorsal muscle causing contraction, while simultaneously releasing ACh onto ventral inhibitory (VI) neurons. Excitation of the VI neurons causes them to release GABA onto ventral muscle fields, causing relaxation, as well as on ventral excitatory (VE) neurons, slowing the release of ACh onto ventral muscle and the dorsal inhibitory neurons (DI). In this scenario, the net outcome is contraction of dorsal muscle and relaxation of ventral muscle, producing a dorsal bend in the body wall. A complementary circuit exists for producing ventral muscle contraction and dorsal relaxation, and intrinsic oscillatory activity of the DI and VI neurons,

which oscillate in antiphase, then controls alternation and propagation of these bends to promote forward or backward movement.

While this circuit seemed to explain the formation and propagation of body waveforms, later electrophysiology on semi-intact worms (only a small section of the worm was opened for recording, and the rest was allowed to behave “normally”) revealed that the timing of the inhibitory motoneuron oscillations was completely different from that of the behavioral waveforms observed [12]. Something was missing from the description of the functional circuit, at least part of which was the action of neuromodulators, which can change the functional properties of neurons.

Neuropeptide Identification and Processing

Neuropeptides are the largest class of neuromodulator molecules, with at least 20 families in vertebrates [13] and 3 large families in nematodes (FMRamide-related peptides, FLPs; insulin-like peptides, INS; and neuropeptide-like proteins, NLPs) [14], with classification largely determined by amino acid sequence similarities. Neuropeptides may be involved in essentially all types of behavior, including locomotion, feeding, satiety, metabolism, reproduction, social behaviors, mechano- and chemo-sensation, learning, and memory [15]. The large number of predicted neuropeptides (>5000 secreted peptides in vertebrates [16], and >250 predicted neuropeptides in *A. suum* [17]) allows for a large variety of potential roles.

Neuropeptides are short sequences of amino acids processed from larger precursor molecules called pre-propeptides, which contain a signal peptide that directs the precursor protein into the secretory pathway, and one or more neuropeptide sequences separated by

cleavage sites. Dibasic cleavage sites (KR, RR, KK) are most common [18-25], but monobasic and tribasic cleavages have also been documented [20, 23, 24, 26-31], as well as rarer “unconventional” cleavages at other sites. As the precursor molecule moves through the endoplasmic reticulum, the signal peptide is cleaved by signal peptidases, and the individual neuropeptides are separated by proprotein convertases, proteolytic enzymes that typically split the precursor C-terminal to the basic amino acids at the cleavage sites [32], followed by removal of the basic residues by carboxypeptidase E [33]. Neuropeptides often undergo additional post-translational modifications, such as C-terminal amidation [34, 35], or N-terminal acetylation, to protect them from degradation and alter biological activity. These modifications begin as the peptides pass through the Golgi, and continue even after they are packaged into dense-core vesicles for secretion. The great deal of processing during formation of the bioactive peptides provides many opportunities for regulation by the cell.

New neuropeptides are often discovered by searches of translated genomic or transcriptomic datasets, using the features described above. However, the variety of processing mechanisms makes prediction of neuropeptides from sequence alone a complicated and unreliable task. For this reason, this laboratory and others have emphasized the direct identification of peptides, either by chemical isolation and sequencing by Edman degradation [36-38], or, more recently, by mass spectrometry (MS) [30, 31, 39-41]. Advances in dissection techniques have allowed the isolation of single neurons in *A. suum*, and the levels of peptide in a single neuron are sufficient to permit sequence determination of the endogenous peptides by tandem MS [30, 31, 39-41].

Cellular expression was traditionally determined by immunocytochemistry, but as the number of isolated peptides increased, generating sufficiently specific antibodies became more difficult, especially where different peptides, even those encoded by different genes, were sequence-related. We have also used *in situ* hybridization for localization of peptide-encoding transcripts, but unfortunately this technique has not yet been successful for *C. elegans* neurons. Instead, cellular expression in *C. elegans* has been determined using constructs that tag the expression of peptides with green fluorescent protein (GFP). However, recent improvements in sensitivity have allowed for analysis of both the transcriptome (RNA-seq) and peptidome (MS) of single cells in *A. suum*. Chapters 2 and 4 present the groundbreaking identification of the neuropeptides expressed by each class of somatic motoneuron, thus achieving the major goal that this laboratory has been working toward for 25 years.

Neuropeptide Signaling Mechanisms

Neuropeptides are generally thought to act as neuromodulators; however, they can also act as fast transmitters in some cases. For example, peptide-gated sodium channels have been identified in the *Hydra* polyp [42, 43] and mollusks [44]. More often, neuropeptides act through G protein-coupled receptors (GPCRs), which couple ligand binding to activation of G proteins and numerous downstream signaling pathways. Upon binding of peptide, the 7-transmembrane (TM) receptor undergoes a conformational change, specifically an outward movement of the TM5 and TM6 domains, which allows the receptor to act as a guanine nucleotide exchange factor (GEF). This activates the associated G protein, allowing it to exchange GDP for GTP. The α subunit of the G protein, bound to GTP, dissociates from the $\beta\gamma$ subunits, and is released into

the cytoplasm. The resulting intracellular response is determined by the type of α subunit employed. For neuropeptide receptors, these would most commonly be the α_s , $\alpha_{i/o}$, or α_q subtypes.

GPCRs coupled to α_s subunits stimulate a cAMP-dependent pathway upon activation. The GTP-bound α_s subunit activates adenylyl cyclase (AC), which converts ATP to cAMP. The increased levels of cAMP in the cytoplasm then activate protein kinase A (PKA), which goes on to phosphorylate numerous downstream targets. Neuropeptides that signal through this pathway include the members of the calcitonin and glucagon/secretin families in vertebrates [45, 46]. The *C. elegans* α_s is encoded by *gsa-1*, and is expressed broadly in neurons and muscle [47, 48]. Partial (mosaic) loss or overexpression of *gsa-1* resulted in defects in movement and egg-laying, and expression of a constitutively active form of the α_s subunit under an inducible promoter caused hypercontraction of muscle and rapid degeneration of motoneurons in the ventral nerve cord [47]. This degeneration of neurons was independent of the apoptotic pathway, and resembled the phenotype of degenerin (Na^+) ion channels, though the effect could not be specifically linked to degenerin or Ca^{2+} channels alone. These studies suggest that the α_s pathway has strong ties to the locomotory circuit in nematodes, and could link neuropeptide signaling to ion channels (either direct activation by the G protein [49, 50] or indirectly through cyclic nucleotide-gated channels [51]) and muscle contraction.

Conversely, receptors coupled to $\alpha_{i/o}$ subunits inhibit the cAMP-dependent pathway, resulting in decreased activity of adenylate cyclase and PKA. This pathway is employed by members of the neuropeptide F/Y, somatostatin, and opioid families in vertebrates [52-54], and the NPR-1 receptor in the neuropeptide Y family in *C. elegans* [55, 56]. The $\alpha_{i/o}$ subunit is

encoded by *goa-1* in *C. elegans*, and is expressed broadly in neurons and pharyngeal and intestinal muscle, though expression in body wall muscle was not reported [57]. It has also been linked to locomotion, egg-laying, and male mating behavior. It also appears to negatively regulate the *egl-30* (α_q) pathway [58]. The widespread expression of *goa-1* in neurons and the potential to further control signaling in the α_q pathway make this an interesting component of the locomotory circuit.

In contrast to α_s and $\alpha_{i/o}$ pathways, GPCRs coupled to α_q subunits activate phospholipase C (PLC) upon activation, catalyzing the conversion of phosphatidylinositol 4,5-bisphosphate (PIP₂) into diacyl glycerol (DAG) and inositol trisphosphate (IP₃). Membrane-bound DAG goes on to activate protein kinase C (PKC), resulting in phosphorylation of many downstream targets. Meanwhile, soluble IP₃ binds to IP₃ channels in the membrane of the endoplasmic reticulum, allowing calcium to flow into the cytosol. Elevated calcium levels can also activate Ca²⁺/calmodulin-dependent kinases (CAMKs), causing phosphorylation of numerous downstream targets, and altering transcription. This pathway is utilized by members of the vasopressin/oxytocin and angiotensin families in vertebrates [59, 60], and the NPR-5 receptor (neuropeptide Y family) in nematodes [61]. The α_q subunit is encoded by the *egl-30* gene in *C. elegans*, and is expressed in most neurons and pharyngeal, vulval, and anal sphincter muscle [62, 63]. It has been linked to locomotion, egg-laying, and pharyngeal pumping. At the neuromuscular junction, the *egl-30* pathway can interact with DAG-binding protein UNC-13 to facilitate release of ACh from the presynaptic motorneuron [63]. These findings directly link the α_q pathway to the locomotory circuit, and could be a mechanism through which neuropeptides affect behavior.

The last conserved G protein subtype is the $\alpha_{12/13}$ subunit, which regulate cell processes through RhoGEFs, affecting actin cytoskeletal remodeling and receptor tyrosine kinase activity. In vertebrates, this pathway is used to control migration of fibroblast and endothelial cells. In *C. elegans*, the α_{12} homolog is encoded by *gpa-12*, and was expressed primarily in the pharynx and hypodermis, though some expression was also observed in the head, tail, and ventral cord [64]. Expression of activated GPA-12 inhibited pharyngeal pumping so dramatically that it caused larval growth arrest [65], and *gpa-12* function is required for antifungal innate immunity [66]. However, a link to locomotion has not been established, so it is less likely than the other α subunits to mediate the effects of neuropeptides from the motor circuit.

Finally, *C. elegans* also expresses 17 other $G\alpha$ subunits which are most closely related to the $\alpha_{i/o}$ subunit, but do not appear to be conserved in vertebrates [64]. These subtypes are generally associated with chemosensation and olfaction, though some are expressed more broadly in other neurons and body wall muscle [64].

While the $G\alpha$ subunits are often the focus, $G\beta$ and γ subunits are also important components of G protein signaling. In vertebrates, the β and γ subunits remain in a complex after the α subunit dissociates, and move along the membrane to activate G-protein-regulated inwardly rectifying K^+ channels (GIRKs), P/Q and N-type voltage-gated Ca^{2+} channels, and some isoforms of AC, PLC, and phosphoinositide-3-kinase (PI3K). Two types of $G\beta$ have been identified in *C. elegans*, encoded by *gpb-1* and *gpb-2* [67-70]. However, only *gpb-1* appears to function with the α and γ subunits in the classical sense. It is expressed in nearly all somatic tissues during development, and expression is highest in neurons after differentiation [67]. Inactivation of *gpb-1* led to abnormalities in embryogenesis, and overexpression caused

lethargic locomotion and delayed egg laying in adults. In contrast, *gpb-2* appears to bind G protein γ -like (GGL)-containing RGS proteins, and this complex can promote the GTPase activity of G α subunits [69].

Similarly, there are also two G γ subunits in *C. elegans*, encoded by *gpc-1* and *gpc-2* [71], but one is responsible for most of the G protein signaling in the worm. GPC-2 is expressed ubiquitously throughout the worm, while GPC-1 expression is limited to sensory neurons. Neither of these subunits are clear orthologs of vertebrate G γ s [71].

Regulation of GPCR Signaling

Finally, once a signaling system has been activated, there must be mechanisms in place to turn it off. At the GPCR level, most natural ligands will eventually dissociate from the binding pocket, and the receptor will undergo a conformational change back to the inactive state. However, this may not happen at a time scale that is appropriate for the cell, so these systems often involve feedback loops with downstream effectors. For example, GPCRs can be phosphorylated by the downstream kinases that they activate (e.g. PKA, PKC, G protein-coupled receptor kinases; GRKs) [72]. This allows β -arrestin proteins to bind to the phosphorylated GPCR to block further G protein-mediated signaling, a process called desensitization, as well as target the GPCR for internalization. In some systems, β -arrestins also interact with their own intracellular signaling cascades (e.g. MAPK cascades)[72, 73], and even translocate across the nuclear membrane to affect transcription [74]. *C. elegans* contains a single arrestin gene, *arr-1*, which is expressed throughout the nervous system and interacts with clathrin and β 2-adaptin, which are involved in receptor endocytosis [75]. Reported functions primarily involve olfactory

adaptation and recovery [75], but it is predicted to fulfill a similar role to vertebrate β -arrestins in GPCR desensitization.

Identification of Neuropeptide GPCRs in Nematodes

Over 1000 GPCRs have been predicted in *C. elegans* (as much as 7% of all predicted protein-encoding genes) [76-78]. The vast majority (ca. 1,300) are predicted nematode-specific chemoreceptors [79], but 100-150 GPCRs are predicted to bind neuropeptides [76, 78, 80, 81]. So far, all known nematode neuropeptide GPCRs fall into the rhodopsin and secretin families according to the GRAFS classification system (Glutamate, Rhodopsin, Adhesion, Frizzled, and Secretin families) [82]. Most of these have not yet been paired with their respective ligands. Since the number of putative neuropeptides (>250 in *A. suum* [17]) is much higher than the number of predicted neuropeptide GPCRs, it is highly likely that single GPCRs bind multiple neuropeptides. For example, the C10C6.2 (NPR-3) receptor binds all 3 of the *flp-15* peptides in *C. elegans* [83], and the Y59H11AL.1 (NPR-22) receptor binds 15 peptides from 6 different *flp* genes, with EC₅₀ values ranging from 25 nM to 5 μ M [84].

To identify the ligands for predicted GPCRs, a process called deorphanization, the receptors are often expressed in a heterologous expression system and then screened against a library of compounds for activation of the GPCR. This approach, called reverse pharmacology, often uses mammalian (CHO, HEK) or insect cell lines because of their capability for post-translational modifications [83-86]. However, several groups have also had success using yeast expression systems [87-89]. Activation of the receptors can be measured using calcium mobilization assays with fluorescent indicators, or through the use of chimeric G proteins that

direct downstream signaling to the pathway of choice (e.g. measuring changes in cAMP, or inducing transcription of specific reporter genes) [90, 91]. All of these methods can be implemented as high-throughput screens, allowing for rapid screening of large peptide libraries.

Once candidate GPCRs have been deorphanized, *in vivo* studies of GPCR function can be performed. In nematodes like *C. elegans*, where genetic manipulations are possible, the methods of choice are RNA-interference (RNAi) and mutant analyses [80, 92, 93]. However, in nematodes like *A. suum* that cannot be cultured in the laboratory, GPCR function is more difficult to assess. In prior studies, our laboratory has only attempted to narrow down the type of G α employed by bioactive neuropeptides by measuring changes in cyclic nucleotide (cAMP, cGMP) concentrations upon injection of peptides [94, 95]. Chapter 5 identifies 7 *A. suum* neuropeptide GPCRs expressed in the somatic motorneurons and muscle cells, and reports the deorphanization of 3 of these by heterologous expression in yeast.

Conservation of Neuropeptide Signaling Systems in Nematodes

Neuropeptide sequences are highly conserved among nematodes (see Chapters 2-4 for several examples); a remarkable finding considering that *C. elegans* and *A. suum* shared a last common ancestor ca. 550 million years ago [3]. However, few nematode neuropeptide homologs have been identified in vertebrates [76]. Similarly, nematode neuropeptide GPCRs are less closely related to predicted vertebrate counterparts than other classes of GPCRs [78]. This presents a problem for GPCR and neuropeptide identification, since prediction is often based on sequence similarity to known vertebrate genes. However, in the context of drug

development, these differences are potentially advantageous, since they could provide a way to combat nematode parasites with fewer adverse side effects for the host.

Our work in *A. suum* has also highlighted the need for additional studies in parasites as opposed to free-living species like *C. elegans*. The somatic inhibitory motoneurons of *A. suum* produce neuropeptides whose sequences are strongly conserved among nematodes, but the expression patterns of these peptides differed substantially between *A. suum* and *C. elegans* [39, 40]. For example, *A. suum* As-NLP-22 (SLASGRWGLRPamide) is strongly expressed by the DI and VI inhibitory motoneurons, and had dramatic inhibitory effects on locomotion and muscle contraction (see Chapter 2) [40]. However, *C. elegans* NLP-22 (SIAIGRAGFRPamide) was expressed only in the RIA interneurons of the head, and was reported to be involved in quiescent behavior during larval lethargus [96]. Our work suggests that As-NLP-22 signaling could be an interesting target for anti-nematode drug development, while the work in *C. elegans* gives no indication of NLP-22 signaling in adult locomotion. Working with parasitic nematodes may have its challenges, but in the case of developing anti-nematode therapeutics, there may be no substitute.

Contributions of this Dissertation to the Field

First, this dissertation identifies the complement of neuropeptides produced by each class of *Ascaris* somatic motoneuron (Chapters 2 and 4), many of which had strong effects on nematode locomotion. Our laboratory has been working toward this goal for 25 years, and this is the first step in identifying potential modulators of the locomotory functional circuit.

Importantly, the expression of these peptides often differed from expression of their homologs in *C. elegans*, highlighting the need for additional studies in parasites.

Second, it reports the sequencing of 7 neuropeptide GPCRs expressed by the motoneurons and muscle cells, and identifies bioactive neuropeptide ligands for 3 of these, as well as a muscarinic ACh receptor, by heterologous expression in yeast (Chapter 5). We show that some of the ligands overlap with deorphanization data in *C. elegans*, but others were unique. These differences in neuropeptide and GPCR expression and deorphanization between *A. suum* and *C. elegans* present an important issue for anti-nematode drug development, since compounds developed from *C. elegans* peptide systems may not work as intended in parasites.

Finally, this dissertation examines the effect of amino acid sequence on activity and ligand binding of two of the most bioactive *A. suum* peptides. Chapter 3 characterizes the pharmacological effects of a family of peptides related to As-NLP-22, the only neuropeptide produced by the inhibitory motoneurons, on muscle contraction. Chapter 5 investigates a series of α/β peptide analogs of AF2, the most abundant and potent excitatory neuropeptide, and reports their effects on GPCR binding. Taken together, this dissertation provides the most complete description of the *A. suum* locomotory circuit to date, and lays groundwork for development of an entirely new class of anthelmintic drugs.

FIGURES

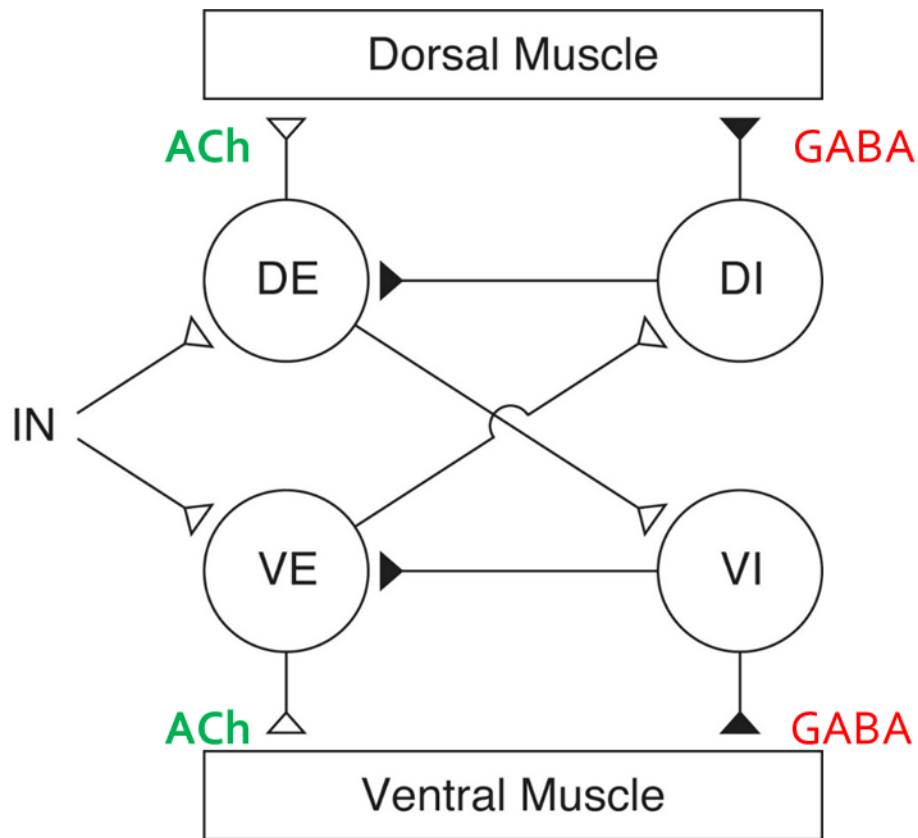


Figure 1. Proposed functional circuit for *A. suum* locomotion. Dorsal excitatory (DE) and dorsal inhibitory (DI) motorneurons synapse onto dorsal muscle fields, while ventral excitatory (VE) and ventral inhibitory (VI) motorneurons synapse onto ventral muscle fields. Excitatory motorneurons release acetylcholine (ACh), and inhibitory motorneurons release γ -aminobutyric acid (GABA). The excitatory motorneurons can receive input from interneurons (IN), and synapse with opposite-side inhibitors to coordinate contraction and relaxation on both sides of the worm. Figure adapted from Stretton *et al.* (1985).

REFERENCES

1. White, J. G., Southgate, E., Thomson, J. N., Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci.* 314:1-340.
2. Stretton, A. O., Fishpool, R. M., Southgate, E., Donmoyer, J. E., Walrond, J. P., Moses, J. E., Kass, I. S. (1978). Structure and physiological activity of the motoneurons of the nematode *Ascaris*. *Proc Natl Acad Sci USA.* 75:3493-7.
3. Holterman, M., van der Wurff, A., van den Elsen, S., van Megen, H., Bongers, T., Holovachov, O., Bakker, J., Helder, J. (2006). Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown Clades. *Mol Biol Evol.* 23:1792-800.
4. Guastella, J., Johnson, C. D., Stretton, A. O. (1991). GABA-immunoreactive neurons in the nematode *Ascaris*. *J Comp Neurol.* 307:584-97.
5. Johnson, C. D., Stretton, A. O. (1987). GABA-immunoreactivity in inhibitory motor neurons of the nematode *Ascaris*. *J Neurosci.* 7:223-35.
6. Johnson, C. D., Stretton, A. O. W. (1985). Localization of choline-acetyltransferase within identified motoneurons of the nematode *Ascaris*. *J Neurosci.* 5:1984-92.
7. Gendrel, M., Atlas, E. G., Hobert, O. (2016). A cellular and regulatory map of the GABAergic nervous system of *C. elegans*. *eLife.* 5:e17686.
8. Pereira, L., Kratsios, P., Serrano-Saiz, E., Sheftel, H., Mayo, A. E., Hall, D. H., White, J. G., LeBoeuf, B., Garcia, L. R., Alon, U., Hobert, O. (2015). A cellular and regulatory map of the cholinergic nervous system of *C. elegans*. *eLife.* 4:e12432.
9. Davis, R. E., Stretton, A. O. (1989). Signaling properties of *Ascaris* motoneurons: graded active responses, graded synaptic transmission, and tonic transmitter release. *J Neurosci.* 9:415-25.
10. Davis, R. E., Stretton, A. O. W. (1989). Passive Membrane-Properties of Motoneurons and Their Role in Long-Distance Signaling in the Nematode *Ascaris*. *J Neurosci.* 9:403-14.

11. Stretton, A. O. W., Davis, R. E., Angstadt, J. D., Donmoyer, J. E., Johnson, C. D. (1985). Neural control of behavior in *Ascaris*. *TRENDS Neurosci.* 8:294-300.
12. Meade, J. (1991). Intracellular recordings from neurons and muscle cells in a semi-intact preparation of the nematode *Ascaris suum* : implications for *Ascaris* locomotion. Ph. D. Thesis. University of Wisconsin-Madison.
13. The Neuropeptide Database (2007). OneStat. <http://www.neuropeptides.nl/>.
14. Li, C., Nelson, L. S., Kim, K., Nathoo, A., Hart, A. C. (1999). Neuropeptide gene families in the nematode *Caenorhabditis elegans*. *Ann N Y Acad Sci.* 897:239-52.
15. Li, C., Kim, K. (2008). Neuropeptides. *WormBook.* 1-36.
16. Li, L. (2017). Personal communication.
17. Yew, J. Y., Kutz, K. K., Dikler, S., Messinger, L., Li, L., Stretton, A. O. (2005). Mass spectrometric map of neuropeptide expression in *Ascaris suum*. *J Comp Neurol.* 488:396-413.
18. Veenstra, J. A. (2000). Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. *Arch Insect Biochem Physiol.* 43:49-63.
19. Fricker, L. D. (2005). Neuropeptide-processing enzymes: applications for drug discovery. *AAPS J.* 7:E449-55.
20. Marks, N. J., Maule, A. G., Geary, T. G., Thompson, D. P., Davis, J. P., Halton, D. W., Verhaert, P., Shaw, C. (1997). APEASPFIRFamide, a novel FMRFamide-related decapeptide from *Caenorhabditis elegans*: structure and myoactivity. *Biochem Biophys Res Commun.* 231:591-5.
21. Marks, N. J., Maule, A. G., Li, C., Nelson, L. S., Thompson, D. P., Alexander-Bowman, S., Geary, T. G., Halton, D. W., Verhaert, P., Shaw, C. (1999). Isolation, pharmacology and gene organization of KPSFVRFamide: a neuropeptide from *Caenorhabditis elegans*. *Biochem Biophys Res Commun.* 254:222-30.

22. Marks, N. J., Sangster, N. C., Maule, A. G., Halton, D. W., Thompson, D. P., Geary, T. G., Shaw, C. (1999). Structural characterisation and pharmacology of KHEYLRFamide (AF2) and KSAYMRFamide (PF3/AF8) from *Haemonchus contortus*. *Mol Biochem Parasitol.* 100:185-94.
23. Marks, N. J., Shaw, C., Halton, D. W., Thompson, D. P., Geary, T. G., Li, C., Maule, A. G. (2001). Isolation and preliminary biological assessment of AADGAPLIRFamide and SVPGVLRFamide from *Caenorhabditis elegans*. *Biochem Biophys Res Commun.* 286:1170-6.
24. L., R. M., Burglin, T. R., Li, C. (1992). Alternatively spliced transcripts of the *flp-1* gene encode distinct FMRFamide-like peptides in *Caenorhabditis elegans*. *J Neurosci.* 12:2356-61.
25. Steiner, D. F. (1998). The proprotein convertases. *Curr Opin Chem Biol.* 2:31-9.
26. Husson, S. J., Clynen, E., Baggerman, G., De Loof, A., Schoofs, L. (2005). Discovering neuropeptides in *Caenorhabditis elegans* by two dimensional liquid chromatography and mass spectrometry. *Biochem Biophys Res Commun.* 335:76-86.
27. Baggerman, G., Boonen, K., Verleyen, P., De Loof, A., Schoofs, L. (2005). Peptidomic analysis of the larval *Drosophila melanogaster* central nervous system by two-dimensional capillary liquid chromatography quadrupole time-of-flight mass spectrometry. *J Mass Spectrom.* 40:250-60.
28. Baggerman, G., Cerstiaens, A., De Loof, A., Schoofs, L. (2002). Peptidomics of the larval *Drosophila melanogaster* central nervous system. *J Biol Chem.* 277:40368-74.
29. Nanda, J. C., Stretton, A. O. (2010). In situ hybridization of neuropeptide-encoding transcripts *afp-1*, *afp-3*, and *afp-4* in neurons of the nematode *Ascaris suum*. *J Comp Neurol.* 518:896-910.
30. Jarecki, J. L., Andersen, K., Konop, C. J., Knickelbine, J. J., Vestling, M. M., Stretton, A. O. (2010). Mapping neuropeptide expression by mass spectrometry in single dissected identified neurons from the dorsal ganglion of the nematode *Ascaris suum*. *ACS Chem Neurosci.* 1:505-19.

31. Jarecki, J. L., Viola, I. R., Andersen, K. M., Miller, A. H., Ramaker, M. A., Vestling, M. M., Stretton, A. O. (2013). Three independent techniques localize expression of transcript *afp-11* and its bioactive peptide products to the paired AVK neurons in *Ascaris suum* : *In situ* hybridization, immunocytochemistry, and single cell mass spectrometry. *ACS Chem Neurosci.* 4:418-34.
32. Thacker, C., Rose, A. M. (2000). A look at the *Caenorhabditis elegans* Kex2/Subtilisin-like proprotein convertase family. *Bioessays.* 22:545-53.
33. Jacob, T. C., Kaplan, J. M. (2003). The EGL-21 carboxypeptidase E facilitates acetylcholine release at *Caenorhabditis elegans* neuromuscular junctions. *J Neurosci.* 23:2122-30.
34. Eipper, B. A., Milgram, S. L., Husten, E. J., Yun, H. Y., Mains, R. E. (1993). Peptidylglycine alpha-amidating monooxygenase: a multifunctional protein with catalytic, processing, and routing domains. *Protein Sci.* 2:489-97.
35. Han, M., Park, D., Vanderzalm, P. J., Mains, R. E., Eipper, B. A., Taghert, P. H. (2004). *Drosophila* uses two distinct neuropeptide amidating enzymes, dPAL1 and dPAL2. *J Neurochem.* 90:129-41.
36. Cowden, C., Stretton, A. O. (1995). Eight novel FMRFamide-like neuropeptides isolated from the nematode *Ascaris suum*. *Peptides.* 16:491-500.
37. Cowden, C., Stretton, A. O. (1993). AF2, an *Ascaris* neuropeptide: isolation, sequence, and bioactivity. *Peptides.* 14:423-30.
38. Cowden, C., Stretton, A. O., Davis, R. E. (1989). AF1, a sequenced bioactive neuropeptide isolated from the nematode *Ascaris suum*. *Neuron.* 2:1465-73.
39. Konop, C. J., Knickelbine, J. J., Sygulla, M. S., Vestling, M. M., Stretton, A. O. (2015). Different neuropeptides are expressed in different functional subsets of cholinergic excitatory motoneurons in the nematode *Ascaris suum*. *ACS Chem Neurosci.* 6:855-70.
40. Konop, C. J., Knickelbine, J. J., Sygulla, M. S., Wruck, C. D., Vestling, M. M., Stretton, A. O. (2015). Mass spectrometry of single GABAergic somatic motoneurons identifies a novel inhibitory peptide, As-NLP-22, in the nematode *Ascaris suum*. *J Am Soc Mass Spectrom.* 26:2009-23.

41. Jarecki, J. L., Frey, B. L., Smith, L. M., Stretton, A. O. (2011). Discovery of neuropeptides in the nematode *Ascaris suum* by database mining and tandem mass spectrometry. *J Proteome Res.* 10:3098-106.
42. Gründer, S., Assmann, M. (2015). Peptide-gated ion channels and the simple nervous system of *Hydra*. *J Exp Biol.* 218:551-61.
43. Cottrell, G. A. (1997). The first peptide-gated ion channel. *J Exp Biol.* 200:2377-86.
44. Lingueglia, E., Champigny, G., Lazdunski, M., Barbry, P. (1995). Cloning of the amiloride-sensitive FMRFamide peptide-gated sodium channel. *Nature.* 378:730-33.
45. Pondel, M. (2001). Calcitonin and calcitonin receptors: bone and beyond. *International Journal of Experimental Pathology.* 81:405-22.
46. Berg, J. M., Tymoczko, J. L., Stryer, L. (2002). Section 21.3, Epinephrine and Glucagon Signal the Need for Glycogen Breakdown.
47. Korswagen, H. C., Park, J. H., Ohshima, Y., Plasterk, R. H. (1997). An activating mutation in a *Caenorhabditis elegans* Gs protein induces neural degeneration. *Genes Dev.* 11:1493-503.
48. Park, J.-H., Ohshima, S., Tani, T., Ohshima, Y. (1997). Structure and expression of the gsa-1 gene encoding a G protein $\alpha(s)$ subunit in *C. elegans*. *Gene.* 194:183-90.
49. Dascal, N. (2001). Ion-channel regulation by G proteins. *Trends in Endocrinology and Metabolism.* 12:391-98.
50. Wickman, K., Clapham, D. E. (1995). Ion channel regulation by G proteins. *Physiol Rev.* 75:865-85.
51. Sunahara, R. K., Dessauer, C. W., Gilman, A. G. (1996). Complexity and diversity of mammalian adenylyl cyclases. *Annu Rev Pharmacol Toxicol.* 36:461-80.
52. Brothers, S. P., Wahlestedt, C. (2010). Therapeutic potential of neuropeptide Y (NPY) receptor ligands. *EMBO Mol Med.* 2:429-39.

53. Murthy, K. S., Coy, D. H., Makhlouf, G. M. (1996). Somatostatin receptor-mediated signaling in smooth muscle: Activation of phospholipase C-beta 3 by Gbeta gamma and inhibition of adenylyl cyclase by Galpha i1 and Galpha o. *J Biol Chem.* 271:23458-63.
54. Al-Hasani, R., Bruchas, M. R. (2011). Molecular mechanisms of opioid receptor-dependent signaling and behavior. *Anesthesiology.* 115:1363-81.
55. Kubiak, T. M., Larsen, M. J., Nulf, S. C., Zantello, M. R., Burton, K. J., Bowman, J. W., Modric, T., Lowery, D. E. (2003). Differential Activation of "Social" and "Solitary" Variants of the *Caenorhabditis elegans* G Protein-coupled Receptor NPR-1 by Its Cognate Ligand AF9. *J Biol Chem.* 278(36):33724-29.
56. Rogers, C., Reale, V., Kim, K., Chatwin, H., Li, C., Evans, P., de Bono, M. (2003). Inhibition of *Caenorhabditis elegans* social feeding by FMRFamide-related peptide activation of NPR-1. *Nat Neurosci.* 6:1178-85.
57. Mendel, J., Korswagen, H., Liu, K., Hajdu-Cronin, Y., Simon, M., Plasterk, R., Sternberg, P. (1995). Participation of the protein Go in multiple aspects of behavior in *C. elegans*. *Science.* 267:1652-55.
58. Miller, K. G., Emerson, M. D., Rand, J. B. (1999). Go α and diacylglycerol kinase negatively regulate the Gq α pathway in *C. elegans*. *Neuron.* 24:323-33.
59. Holmes, C. L., Landry, D. W., Granton, J. T. (2003). Science review: Vasopressin and the cardiovascular system part 1--receptor physiology. *Crit Care.* 7:427-34.
60. Harris, D. M., Chen, X., Pesant, S., Cohn, H. I., MacDonnell, S. M., Boucher, M., Vinge, L. E., Raake, P., Moraca, S. R., Li, D., Most, P., Houser, S. R., Koch, W. J., Eckhart, A. D. (2009). Inhibition of angiotensin II Gq signaling augments beta-adrenergic receptor mediated effects in a renal artery stenosis model of high blood pressure. *J Mol Cell Cardiol.* 46:100-7.
61. Kubiak, T. M., Larsen, M. J., Bowman, J. W., Geary, T. G., Lowery, D. E. (2008). FMRFamide-like peptides encoded on the flp-18 precursor gene activate two isoforms of the orphan *Caenorhabditis elegans* G-protein-coupled receptor Y58G8A.4 heterologously expressed in mammalian cells. *Peptide Science.* 90:339-48.

62. Bastiani, C. A., Gharib, S., Simon, M. I., Sternberg, P. (2003). *Caenorhabditis elegans* Gαq regulates egg-laying behavior via a PLCβ-independent and serotonin-dependent signaling pathway and likely functions both in the nervous system and in muscle. *Genetics*. 165:1805-22.
63. Lackner, M. R., Nurrish, S. J., Kaplan, J. M. (1999). Facilitation of synaptic transmission by EGL-30 Gqα and EGL-8 PLCβ. *Neuron*. 24:335-46.
64. Jansen, G., Thijssen, K. L., Werner, P., van der Horst, M., Hazendonk, E., Plasterk, R. H. (1999). The complete family of genes encoding G proteins of *Caenorhabditis elegans*. *Nat Genet*. 21:414-9.
65. van der Linden, A. M., Moorman, C., Cuppen, E., Korswagen, H. C., Plasterk, R. H. A. (2003). Hyperactivation of the G12-mediated signaling pathway in *Caenorhabditis elegans* induces a developmental growth arrest via protein kinase C. *Current Biology*. 13:516-21.
66. Ziegler, K., Kurz, C. L., Cypowyj, S., Couillault, C., Pophillat, M., Pujol, N., Ewbank, J. J. (2009). Antifungal innate immunity in *C. elegans*: PKCdelta links G protein signaling and a conserved p38 MAPK cascade. *Cell Host Microbe*. 5:341-52.
67. Zwaal, R. R., Ahringer, J., van Luenen, H. G. A. M., Rushforth, A., Anderson, P., Plasterk, R. H. A. (1996). G Proteins are required for spatial orientation of early cell cleavages in *C. elegans* embryos. *Cell*. 86:619-29.
68. Zwaal, R. R., Mendel, J. E., Sternberg, P. W., Plasterk, R. H. (1997). Two neuronal G proteins are involved in chemosensation of the *Caenorhabditis elegans* Dauer-inducing pheromone. *Genetics*. 145:715-27.
69. Robatzek, M., Niacaris, T., Steger, K., Avery, L., Thomas, J. H. (2001). eat-11 encodes GPB-2, a Gβ5 ortholog that interacts with Goα and Gqα to regulate *C. elegans* behavior. *Curr Biol*. 11:288-93.
70. Chase, D. L., Patikoglou, G. A., Koelle, M. R. (2001). Two RGS proteins that inhibit Gao and Gαq signaling in *C. elegans* neurons require a Gβ5-like subunit for function. *Current Biology*. 11:222-31.
71. Jansen, G., Weinkove, D., Plasterk, R. H. (2002). The G-protein gamma subunit gpc-1 of the nematode *C.elegans* is involved in taste adaptation. *EMBO J*. 21:986-94.

72. Lefkowitz, R. J., Shenoy, S. K. (2005). Transduction of receptor signals by beta-arrestins. *Science*. 308:512-7.
73. Luttrell, L. M., Ferguson, S. S. G., Daaka, Y., Miller, W. E., Maudsley, S., Della Rocca, G. J., Lin, F. T., Kawakatsu, H., Owada, K., Luttrell, D. K., Caron, M. G., Lefkowitz, R. J. (1999). β -Arrestin-dependent formation of β 2 adrenergic receptor-Src protein kinase complexes. *Science*. 283:655-61.
74. Kang, J., Shi, Y., Xiang, B., Qu, B., Su, W., Zhu, M., Zhang, M., Bao, G., Wang, F., Zhang, X., Yang, R., Fan, F., Chen, X., Pei, G., Ma, L. (2005). A nuclear function of beta-arrestin1 in GPCR signaling: regulation of histone acetylation and gene transcription. *Cell*. 123:833-47.
75. Palmitessa, A., Hess, H. A., Bany, I. A., Kim, Y. M., Koelle, M. R., Benovic, J. L. (2005). *Caenorhabditis elegans* arrestin regulates neural G protein signaling and olfactory adaptation and recovery. *J Biol Chem*. 280:24649-62.
76. Bargmann, C. I. (1998). Neurobiology of the *Caenorhabditis elegans* Genome. *Science*. 282:2028-33.
77. Fredriksson, R., Schiöth, H. B. (2005). The repertoire of G-protein-coupled receptors in fully sequenced genomes. *Mol Pharmacol*. 67:1414-25.
78. Frooninckx, L., Van Rompay, L., Temmerman, L., Van Sinay, E., Beets, I., Janssen, T., Husson, S. J., Schoofs, L. (2012). Neuropeptide GPCRs in *C. elegans*. *Front Endocrinol*. 3:167.
79. Thomas, J. H., Robertson, H. M. (2008). The *Caenorhabditis* chemoreceptor gene families. *BMC Biol*. 6:42.
80. Keating, C. D., Kriek, N., Daniels, M., Ashcroft, N. R., Hopper, N. A., Siney, E. J., Holden-Dye, L., Burke, J. F. (2003). Whole-Genome Analysis of 60 G Protein-Coupled Receptors in *Caenorhabditis elegans* by Gene Knockout with RNAi. *Current Biology*. 13:1715-20.
81. Hobert, O. (2013). The neuronal genome of *Caenorhabditis elegans*. *WormBook*. 1-106.

82. Schioth, H. B., Fredriksson, R. (2005). The GRAFS classification system of G-protein coupled receptors in comparative perspective. *Gen Comp Endocrinol.* 142:94-101.
83. Kubiak, T. M., Larsen, M. J., Zantello, M. R., Bowman, J. W., Nulf, S. C., Lowery, D. E. (2003). Functional annotation of the putative orphan *Caenorhabditis elegans* G-protein-coupled receptor C10C6.2 as a FLP15 peptide receptor. *J Biol Chem.* 278:42115-20.
84. Mertens, I., Clinckspoor, I., Janssen, T., Nachman, R., Schoofs, L. (2006). FMRFamide related peptide ligands activate the *Caenorhabditis elegans* orphan GPCR Y59H11AL.1. *Peptides.* 27:1291-96.
85. Mertens, I., Meeusen, T., Janssen, T., Nachman, R., Schoofs, L. (2005). Molecular characterization of two G protein-coupled receptor splice variants as FLP2 receptors in *Caenorhabditis elegans*. *Biochemical and Biophysical Research Communications.* 330:967-74.
86. Mertens, I., Vandingenen, A., Meeusen, T., Janssen, T., Luyten, W., Nachman, R. J., De Loof, A., Schoofs, L. (2004). Functional characterization of the putative orphan neuropeptide G-protein coupled receptor C26F1.6 in *Caenorhabditis elegans*. *FEBS Letters.* 573:55-60.
87. Kimber, M. J., Sayegh, L., El-Shehabi, F., Song, C., Zamanian, M., Woods, D. J., Day, T. A., Ribeiro, P. (2009). Identification of an *Ascaris* G protein-coupled acetylcholine receptor with atypical muscarinic pharmacology. *Int J Parasitol.* 39:1215-22.
88. Larsen, M. J., Lancheros, E. R., Williams, T., Lowery, D. E., Geary, T. G., Kubiak, T. M. (2013). Functional expression and characterization of the *C. elegans* G-protein-coupled FLP-2 Receptor (T19F4.1) in mammalian cells and yeast. *Int J Parasitol.* 3:1-7.
89. Tehseen, M., Dumancic, M., Briggs, L., Wang, J., Berna, A., Anderson, A., Trowell, S. (2014). Functional Coupling of a Nematode Chemoreceptor to the Yeast Pheromone Response Pathway. *PLOS ONE.* 9:e111429.
90. Milligan, G., Rees, S. (1999). Chimaeric G alpha proteins: their potential use in drug discovery. *Trends Pharmacol Sci.* 20:118-24.
91. Wang, Z.-x., Broach, J. R., Peiper, S. C. (2006). Functional Expression of CXCR4 in *Saccharomyces cerevisiae* in the Development of Powerful Tools for the Pharmacological Characterization of CXCR4. 332:115-27.

92. Ceron, J., Rual, J. F., Chandra, A., Dupuy, D., Vidal, M., van den Heuvel, S. (2007). Large-scale RNAi screens identify novel genes that interact with the *C. elegans* retinoblastoma pathway as well as splicing-related components with synMuv B activity. *BMC Dev Biol.* 7:30.
93. Rual, J. F., Ceron, J., Koreth, J., Hao, T., Nicot, A. S., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S. H., Hill, D. E., van den Heuvel, S., Vidal, M. (2004). Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library. *Genome Res.* 14:2162-8.
94. Reinitz, C. A., Herfel, H. G., Messinger, L. A., Stretton, A. O. (2000). Changes in locomotory behavior and cAMP produced in *Ascaris suum* by neuropeptides from *Ascaris suum* or *Caenorhabditis elegans*. *Mol Biochem Parasitol.* 111:185-97.
95. Reinitz, C. A., Pleva, A. E., Stretton, A. O. (2011). Changes in cyclic nucleotides, locomotory behavior, and body length produced by novel endogenous neuropeptides in the parasitic nematode *Ascaris suum*. *Mol Biochem Parasitol.* 180:27-34.
96. Nelson, M. D., Trojanowski, N. F., George-Raizen, J. B., Smith, C. J., Yu, C. C., Fang-Yen, C., Raizen, D. M. (2013). The neuropeptide NLP-22 regulates a sleep-like state in *Caenorhabditis elegans*. *Nat Commun.* 4:2846.

CHAPTER TWO

Mass Spectrometry of Single GABAergic Somatic Motorneurons Identifies a Novel Inhibitory Peptide, As-NLP-22, in the Nematode *Ascaris suum*

Christopher J. Konop*, Jennifer J. Knickelbine*, Molly S. Sygulla, Colin D. Wruck, Martha S. Vestling, and Antony O. W. Stretton

This work is published in: Journal of the American Society for Mass Spectrometry (2015), Volume 26, Pages 2009-2023. Doi: 10.1007/s13361-015-1177-z

*These authors made equal contributions to the work reported here.

INTRODUCTION

In the mid-20th century there was a resurgence in interest in simple nervous systems, typically found in invertebrates, where the numbers of neurons are low [1, 2]. The hope was that the low numbers would allow anatomical descriptions of the individual neurons and their synaptic connections to be combined with electrophysiological measurements to give a “functional wiring diagram”, and that this would give us an understanding of the way the nervous system controls behavior. Progress has been slow: it is now recognized that the simplicity of these nervous systems is more apparent than real [3]. A large part of the complexity of even these “simple nervous systems” comes from the existence of a large number of modulating signaling molecules, typically peptides and amines, that affect the neurons and synapses in subtle and complicated ways, each of which needs to be worked out. The description of the functional wiring diagram was incomplete because it did not include the role of these neuromodulators.

The inadequacy of the naïve “functional wiring diagram” approach and the importance of the role of neuropeptides and other modulators in affecting circuit properties have been particularly well described in the stomatogastric ganglion (STG) of Crustacea [4-12]. In this ganglion, the cell numbers are small ($N = 30$) and despite the complicated morphology of the neurons, their synaptic connectivity has been determined, both anatomically and physiologically [11, 12]. However, enormous complexity is introduced by the large number of modulators that reach STG neurons from neurons in other ganglia [7, 13-16]. These modulators can dramatically change the properties of the neurons and synapses in the STG, and each one may produce complex and different responses on different targets, reconfiguring the network into different functional circuits. This is the important concept of “circuit switching” [12, 17, 18]. The analysis of the system is ongoing as more modulators are discovered [4-7, 19, 20].

Our attempt to simplify the analysis of circuitry has taken advantage of the properties of the nervous system of the nematode *Ascaris suum*. Like that of other nematodes, its nervous system is extremely simple, both numerically (adult female *A. suum* has a total of 298 neurons [21]; the *Caenorhabditis elegans* hermaphrodite has 302 neurons [22]) and morphologically (typical neurons have no more than 2 branch points); yet they have an extensive array of neuropeptides [23-29]. In order to understand a particular neuronal circuit at the cellular level, the salient peptides must be identified, and their biological activities analyzed. These are crucial parts of the structure-function description of a circuit that are needed to create a meaningful functional wiring diagram. Similarly, the identification of the neurons that express each neuropeptide, i.e. the cellular expression pattern, is functionally important.

We are concentrating on one aspect of the behavior of *A. suum*, the control of locomotion. *A. suum* exhibits a simple locomotory behavior characterized by a propagating two-dimensional waveform produced by alternating contractions and relaxations of dorsal and ventral muscles; these muscles are controlled by serial sets of dorsal and ventral excitatory and inhibitory motoneurons distributed along the length of the body [30]. The cell bodies of these motor neurons are relatively large and are found in a highly reproducible ordered array in the ventral cord; this has allowed for extensive biochemical and physiological characterization of these neurons [30-37]. After locating the classical neurotransmitters in individual motoneurons and investigating their function, we have taken several paths to identifying endogenous peptides. Initially, we used chemical isolation by multi-step high pressure liquid chromatography (HPLC), followed by the use of specific antibodies and *in situ* hybridization [27, 28, 38, 39]. More recently we have used mass spectrometry (MS), first on isolated ganglia, and then on single identified neurons [23, 40, 41], and have identified and sequenced over 100 peptides from *A. suum*; this is a fraction of the over 250 putative peptides that are predicted [29, 40, 42]. Already it is clear that most of the neuropeptides that have been tested affect the physiological properties of the motoneurons and/or muscle cells, although many of these peptides are expressed in neurons that are not an obvious part of the locomotory circuitry. We have now improved the dissection techniques so that we can isolate individual identified motoneurons from the ventral nerve cord, and isolate and sequence their endogenous neuropeptides.

As in other organisms, peptides in nematodes are generated from precursor proteins by proteolysis and other post-translational processing steps [29, 43]. Nematode peptides and

genes are highly conserved within the phylum and many are grouped into three large families, originally described in *C. elegans* [44-46]: the peptides in two of these families, FMRFamide-like peptides (*flps*) and insulin-like peptides (*ins*), are sequence-related within their respective families. A third family consists of a more diverse group of neuropeptide-like proteins (*nmps*); in this group there are subfamilies of related peptides, but the subfamilies have little sequence similarity to each other [29, 44].

In order to identify the endogenous peptides of motoneurons, we have dissected single motoneurons and analyzed them by single-cell mass spectrometry. This was made possible by a necessary refinement, described in this paper, of the single cell dissection technique previously developed in *A. suum* [23, 41], and used in conjunction with matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS. The quantities of peptide present in single cells are sufficient not only to identify individual peptides in individual identified neurons by their molecular mass, but also to determine their sequence by tandem MS (MS/MS). We confirm the *de novo* sequence determination by tandem MS of synthetic peptide, showing that the fragmentation pattern is the same. Determination of the peptide sequence by MS is what leads this work; the sequence is enabling, since it permits searching of nucleic acid databases for the encoding transcripts which can be used for the synthesis of riboprobes for *in situ* hybridization; it also is essential for making synthetic peptide for functional studies, since the quantities of peptide readily available from biological material are not sufficient (unless heroic purifications are carried out).

We report here the identification, localization, and bioactivity of a peptide expressed in the GABAergic VI and DI inhibitory ventral cord motoneurons. This peptide, As-NLP-22

(SLASGRWGLRPamide), is encoded by the *A. suum* neuropeptide-like protein gene *As-nlp-22*, and is cleaved from the precursor protein encoded by this gene. We have confirmed the cellular expression of this gene by *in situ* hybridization; this is an important validation of the adequacy of the single neuron dissection technique, since it shows that the mass spectra are not massively contaminated by peptides present in processes (e.g. synaptic terminals) from neighboring neurons. Pharmacological experiments with synthetic As-NLP-22 show that it is a potent inhibitor of acetylcholine (ACh)-induced contraction of *A. suum* muscle, and that it induces paralysis in intact behaving worms.

A search of nematode expressed sequence tag (EST) libraries for As-NLP-22-like peptides identified predicted homologous peptides in several parasitic and free-living species, including *C. elegans*. Because of its inhibitory effects on muscle, and its apparent ubiquity among nematodes, it promises to be an excellent candidate for the development of novel anthelmintic drugs.

METHODS

Animals

Live *A. suum* were collected from pig intestines at a slaughterhouse, and maintained in phosphate-buffered saline (PBS, 140 mM sodium chloride, 10 mM sodium phosphate, pH 6.8-7.5) at 37°C. The PBS was changed daily and worms were used within 3 days of collection.

Sample Preparation for Mass Spectrometry

Adult female *A. suum* were injected with 0.1-0.3 mL of 2 mg/mL collagenase (Sigma Blend H) in *Ascaris* saline (4 mM sodium chloride, 125 mM sodium acetate, 24.5 mM potassium chloride, 5.9 mM calcium chloride, 4.9 mM magnesium chloride, 5 mM 3-(N-morpholino) propanesulfonic acid (MOPS) buffer, pH 6.8) and incubated for 1.5- 2 hours at 37°C to dissociate the muscle tissue. A 6-7 cm portion of the worm anterior to the gonopore was removed and transferred to a Sylgard-lined dish. After cutting longitudinally between the dorsal nerve cord and the right lateral line, the preparation was pinned open, the lips were cut off, and the pharynx was removed. The preparations were rinsed 3x with 170 mM ammonium acetate before the isotonic glycerol dissecting solution was added (170 mM ammonium acetate in 30% glycerol [47-50]). Dissecting needles made from electrolytically-sharpened tungsten wire were used to isolate a desired neuron from the many neuronal processes and hypodermis in which it is embedded. For better visualization, some preparations were bathed in 0.8 mM Methylene Blue in 170 mM ammonium acetate for 20-30 seconds, then cleared with 170 mM ammonium acetate and dissected as described above.

Isolated cells were transferred to a Bruker MTP stainless steel MALDI target. Cells were cleared with 0.5 μ L isopropanol [51] to remove the glycerol and any remaining extracellular salts. Individual neurons were spotted with 60 - 100 nL of saturated α -cyano-4-hydroxycinnamic acid (CHCA; Sigma, St Louis, MO) in 70% acetonitrile (ACN; Fisher Scientific, Waltham, MA), 1% trifluoroacetic acid (TFA; Sigma) or saturated 2,5-dihydroxybenzoic acid (DHB; Sigma) in 1:1 HPLC grade methanol and water (Fisher Scientific) using a Nanoliter Cool Wave Syringe II [52].

Mass Acquisition

A Bruker Ultraflex III MALDI-TOF/TOF MS (Bruker Daltonics, Billerica, MA) equipped with a Smartbeam laser and LIFT-TM cell was used to obtain MS and MS/MS spectra with Bruker Compass v. 1.2 software. Because the cells contain relatively small quantities of peptide, all spectra were obtained from 50 laser shots per acquisition. MS spectra were obtained in positive ion reflectron mode, with an m/z range of 500-4000 Da. Instrument settings were optimized for maximum detection sensitivity using the Bruker Flex Control software: ion source 1 voltage, 25.0 kV; ion source 2 voltage, 20.5 kV; reflector 1 voltage, 26.5 kV; reflector 2 voltage, 14.5 kV; lens voltage, 9.5 kV. Synthetic peptide standards were not suitable for calibration of the instrument because the laser intensity needed for analysis of single cells was often far too high when analyzing pure standard, making it impossible to identify the monoisotopic peak of the ion needed for accurate calibration. Therefore, a reference spectrum from a freshly dissected *A. suum* AVK neuron, with known peptide peaks [23], was used for external calibration of the instrument.

Acetylation of Peptides

Following deposition of the cell onto the target surface, 0.5 μ L of isopropanol was applied to each cell to wash away the glycerol solution, followed by 0.5 μ L of methanol/ acetic anhydride (3:1). The cells were allowed to dry before each wash. Cells were then covered with matrix as described above.

Oxidation of Peptides

In cells exposed to Methylene Blue, methionine residues were partially oxidized to the sulfoxide, with a mass shift of +16 Da, and tryptophan residues gave rise to +16 and +32 adducts [41].

Assignment of Peaks and Interpretation of Mass Spectra

Spectra were analyzed using Bruker Daltonics flex-Analysis 3.0 software. The software automatically assigned masses to peaks in each MS spectrum. Each MS/MS spectrum underwent background subtraction and smoothing before the automatic assignment of masses. In some cases, peak m/z values were added manually. Peaks were considered significant if they were twice the intensity of the baseline noise.

Molecular masses and ion fragmentation patterns were calculated by Protein Prospector MS-Product (<http://prospector.ucsf.edu>). Spectra containing peaks with m/z values corresponding to masses ± 0.2 Da of known *A. suum* peptides were temporarily assigned. Confirmation of the assignments was carried out by MS/MS and chemical modifications to the peptides.

De novo sequencing of unidentified peaks was carried out by hand. To establish the b series (N-terminal) and y series (C-terminal) ions for each MS/MS spectrum, candidate b_2 and y_1 ions were chosen as a starting point for sequencing. To grow each ion series, the annotation function in the flex-Analysis software was used to calculate the distance between adjacent peaks to identify each amino acid until the full sequence was deduced in both directions. The sequence was entered into Protein MS-Prospector and the spectrum was analyzed for the complete set of sequence ions, including internal fragments and immonium ions. Verification of

peptide sequence was carried out by comparing the experimental MS/MS spectrum to that of synthetic peptide. Further verification was carried out by comparing sequences to tBLASTn searches and cloning of *A. suum* gene sequences. Raw spectral image files were transferred to Adobe Illustrator (San Jose, CA) for annotation.

Database Searches

Database searches were conducted using methods described in recent publications [23, 41, 42]. Briefly, all predicted peptide sequence assignments were searched using tBLASTn (National Center for Biotechnology Information [NCBI], Bethesda, MD; <http://www.ncbi.nlm.nih.gov/BLAST/>) against *A. suum* ESTs and a library of 447,546 genomic survey sequences (GSS). For all searches, program settings were modified for searching short sequences using a word size of 2, an E value of 20,000, and a PAM30 matrix. Search results were translated using the ExPASy Translate tool (<http://au.expasy.org/tools/dna.html>) and examined for putative peptide cleavage sites, and then for C-terminal glycine for C-terminal amidation of peptides. Genomic sequences were analyzed for mRNA splice sites directly by PCR cloning of the mRNA sequences and comparison with genomic sequences [53]. Signal peptides were predicted using SignalP 4.0 (<http://www.cbs.dtu.dk/services/SignalP/>).

Peptide Synthesis

Peptides were synthesized (Fmoc chemistry) by the University of Wisconsin-Madison Biotechnology Center (UWBC). The integrity of the synthesis was monitored by MALDI-TOF MS and HPLC.

RNA Isolation and cDNA Preparation

Twenty-thirty *A. suum* heads were flash-frozen in liquid nitrogen and ground to a fine powder. Total RNA was isolated using a Nucleospin Nucleic Acid Purification Kit (Clontech, Mountain View, CA). First strand cDNA was generated for PCR using a Superscript First Strand Synthesis System for RT-PCR (Life Technologies, Grand Island, NY). Rapid Amplification of cDNA Ends (RACE)-ready cDNA for 5'- and 3'-RACE was created from total RNA using a SMARTer RACE cDNA amplification kit (Clontech) according to the manufacturer's instructions.

Transcript Identification by PCR and RACE

To obtain information about the transcript that encoded this peptide, primers were designed for PCR and RACE from sequences in the *A. suum* EST library (accession no. BI593877). RACE reactions were performed using an Advantage 2 PCR Kit (Clontech) according to the manufacturer's instructions with the 5'RACE primer nlp22-5RACE3 (5'-GTGCACACGATCCACAAAGCCTTCGAAA-3') and 3'RACE primer nlp22-3RACE (5'-CAAAACGCTCCTTAGCTAGCGGTCGTTG -3'). The PCR reaction was performed using the forward primer nlp22-3RACE2 (5'-GCTCGTTGCTTGCGGTGCTCTTCGTTTCGAT -3') and the reverse primer nlp22-5RACE (5'-CCCATCGAGGCCAAGTTCGTCAGTATAAATCG-3') in a 50 µL reaction as follows: 2 µL of *A. suum* cDNA or 2 µL of cDNA reaction solution with no reverse transcriptase as a control, 5 µL of 10X PCR Gold buffer (Applied Biosystems, Foster City, CA), 2-8 µL of 25 mM MgCl₂, 1 µL each of 10 mM dNTPs, 1 µL of each primer (20 µM), and 0.25 µL of AmpliTaq Gold Polymerase (Applied Biosystems; 5 U/µL). The PCR conditions were programmed into the

Eppendorf Mastercycler Gradient as follows: 95°C for 3 min, followed by 40 cycles of 94°C for 30 seconds, 66°C or 68°C for 30 seconds, 72°C for 1 min.

PCR and RACE products were run on a 1% agarose gel, and novel bands were excised and purified using a Qiaquick Gel Extraction Kit (Qiagen, Chatsworth, CA). The purified products were cloned into *E. coli* using a TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA), and the plasmid DNA was isolated with a Qiagen Miniprep Kit. Automated sequencing was carried out by the DNA Sequencing Facility at the UWBC. Sequence electropherograms were viewed on Chromas Lite software (Technelysium Pty Ltd). Sequences were analyzed using the ExPASy Translate Tool and T-Coffee Multiple Sequence Alignment Tool (Swiss Institute of Bioinformatics), and signal peptide sequences were identified using SignalP 4.0.

Riboprobe Synthesis

The *As-nlp-22* specific riboprobe was created using primers *nlp22-3RACE2* and *nlp22-5RACE* for PCR as described above. Products were cloned and sequenced to confirm the fidelity of the sequence and to determine the orientation of the insert in the vector. The target sequences of the riboprobe are shown in Figure 3b. The constructs were linearized using restriction enzymes NotI and SpeI (New England Biolabs, Beverly, MA). Linearized plasmids were used as a template to synthesize an antisense (experimental) and a sense (negative control) digoxigenin-labeled riboprobe (Maxiscript SP6/T7 kit, Ambion, Austin, TX; digoxigenin-11-dUTP, Roche Applied Science, Indianapolis, IN) as previously described [54]. To remove unincorporated nucleotides the reactions were run through NucAway Spin Columns (Ambion). Probe integrity and concentration were determined by gel electrophoresis with Sybr Gold

staining (Molecular Probes, Eugene, OR) and nylon membrane dot blots (Roche Applied Science protocol).

Localization of As-nlp-22 by in situ Hybridization

A. suum whole mount preparations were prepared by injecting large female worms with 0.1 - 0.3 mL of 2 mg/mL collagenase in autoclaved *Ascaris* saline and incubating in a beaker of PBS at 37°C for 1.5 – 2 hrs. The anterior 6 cm of the worm was taken and rinsed in autoclaved 100 mM phosphate buffer (pH 7.4) and the heads were transferred to a Sylgard-coated dissecting dish. The heads were cut longitudinally between the dorsal nerve cord and the right lateral line, the lips were removed, and the heads were pinned flat. Any remaining muscle cells were gently removed with forceps. The preparations were rinsed several times in 100 mM phosphate buffer, and fixed in 1% paraformaldehyde in 100 mM phosphate buffer (pH 7.4) on a rocking table at 4°C overnight.

After fixation, the head preparations were processed as previously described [54], with the following modifications: incubation in proteinase K solution was increased to 30 min at 37°C, and staining was allowed to occur for up to 16 hours in the dark, or until staining of the cells was evident and background staining began to occur. The staining reaction was stopped with Milli-Q water, and the heads were rinsed with Milli-Q water (2 x 10 min). The preparations were mounted in Clear-Mount (Electron Microscopy Sciences, Hatfield, PA) and allowed to dry. Microphotographs were taken with a Zeiss AxioCam MRc camera on a Zeiss Universal microscope.

Characterization of As-NLP-22 Bioactivity in Muscle Strips

Strips of dorsal muscle were obtained by cutting large female *A. suum* longitudinally along the lateral lines and isolating a 2 cm strip approximately 1 cm posterior to the head. Each preparation contained dorsal muscle cells (all longitudinal) and the motor axons of the dorsal nerve cord. The ends of the preparation were tied with silk thread. One end was attached to a fixed hook in a 7 mL chamber containing normal *Ascaris* saline stirred by bubbling nitrogen gas, and the other end was tied to a FORT25 force transducer (World Precision Instruments, Sarasota, FL). The output of the transducer was routed through a TBM4M transbridge (WPI) and recorded on a computer using Data-Trax or LabScribe2 software (WPI), and the data were analyzed for tension/time relationships.

Baseline contractions were measured by adding 5 μ M ACh to the chamber by micropipette, followed by rinses of *Ascaris* saline. The preparation was then exposed to 10 μ M peptide solution in *Ascaris* saline for 10 min. The chamber was rinsed, fresh peptide was added, and ACh-induced contraction was measured by adding 5 μ M ACh immediately after exposure to peptide solution, followed by rinses with peptide-free *Ascaris* saline. Muscle strip contraction in response to ACh was measured at 10, 20, 30, 40, and 50 min after exposure to peptide, with rinses of peptide-free *Ascaris* saline between doses of ACh. For comparison, control muscle strips were subjected to ACh-induced contraction in peptide-free *Ascaris* saline at the same time points. Responses of the muscle strips were reported as a percentage of the maximum ACh contraction observed for each muscle strip. A dose-response curve was created by repeating the procedure described above with peptide concentrations of 10 μ M, 1 μ M, 100 nM, 10 nM, 1 nM, 0.1 nM, 0.01 nM and comparing the results to control muscle strips.

Statistical Analysis

Contractile responses and baseline tension were reported as the mean $\pm$ S.E. for each peptide concentration and displayed in Prism 6 software (GraphPad, San Diego, CA). Statistical significance between peptide and control trials was determined using unpaired *t*-tests, with significance levels set at $p < 0.05$ or $p < 0.01$. The IC_{50} was calculated using the four-parameter logistic (4PL) curve and Prism 6 software, with constants estimated by nonlinear regression.

Characterization of NLP Effects in Intact Worms

To examine the broader behavioral responses associated with neuropeptides, we used the whole worm injection assay [55]. Worms used in this assay are allowed to move freely inside an 18 mm diameter x 41 cm long glass tube filled with PBS at 37°C and closed off with rubber stoppers. This tube is similar in diameter to the porcine small intestine, which is the natural habitat of these worms. The worms were allowed to acclimate to the tube for 5 minutes, followed by 5 minutes of observation prior to injection to establish a baseline for locomotory activity. Then, the worms were partially removed from the tube and injected with 0.1 mL of 10 μ M peptide solution in *Ascaris* saline, and placed back inside the tube. Handling of the worms was kept to a minimum. Locomotory activity was observed for 60 minutes post-injection, with recorded observations grouped into 5-minute increments. Digital photos were captured with a Nikon D50 camera at 5 minutes pre-injection, immediately before and immediately after injection, and at 5 minute intervals post-injection.

Multiple Alignment

BLAST searches of the nematode EST library were performed using both As-NLP-22 and Ce-NLP-22 as queries. Selected sequences were imported into MEGA5.1 [56] for alignment by MUSCLE [57]. The alignments were imported into Jalview [58] for display.

RESULTS AND DISCUSSION

Anatomical Background

The basic morphological and physiological properties of adult *A. suum* motorneurons have been well described [30], as has their classical neurotransmitter phenotype [36, 37]. The somatic motorneurons, which innervate the dorsal and ventral musculature, comprise seven morphological and physiological types, four (DE1, DE2, DE3, and DI) controlling dorsal muscle, and three (VE1, VE2, and VI) controlling ventral muscle (Figure 1). The cell bodies of all these neurons are located in the ventral nerve cord, and dorsal muscle is innervated via neural processes (commissures) that extend from the ventral to the dorsal nerve cord. These seven types of motorneurons are analogous to seven classes of motorneurons found in *C. elegans* [22, 30]. The dorsal and ventral exciters (DE and VE) are cholinergic [36] and the corresponding inhibitors (DI and VI) are GABAergic [34, 37]. These motorneurons occur in repeating patterns along the length of the worm, such that each of the five “segments” includes 11 motorneurons, four of which (DE1, VE1, VE2, and VI) are present in two copies, with only a single copy of DE2, DE3 and DI [30].

Cell identification relies on several features (see Figure 1 a, b). First, the DE1, DE2 and DE3 neurons have commissures that emerge from the cell body; in the case of the DE1 and DE2 neurons, the identification is reinforced by the diameter of the commissure, which is larger

than that of the accompanying commissure from an inhibitory neuron [36]. Second, the cell body of the VI neuron is the next posterior cell body to the DE1/VI commissure pair. The cell body of the DI motorneuron is posterior to the DE2/DI commissure pair; the relative position of the DE3 and DI cell bodies is variable, but DE3 is recognized by the commissure emerging from its cell body. Five of these types (DE1, DE2, DE3, DI and VI) have been investigated electrophysiologically, and their properties are unusual [59]. None of these neurons conducts action potentials; instead they have high membrane resistance, which enables them to signal passively over long distances. They have long processes (axons) within the nerve cords, and synapse onto muscle cells along the length of the axon, providing synaptic input to muscle cells over a distance of several centimeters [32].

Dissection of Single Motorneurons

Initially, the dissection of single motorneurons from the ventral nerve cord was difficult because of the close association with hypodermis and the other processes within the nerve cords. The use of an isotonic dissecting solution containing glycerol has permitted the isolation of all the cell bodies of the motorneurons in any chosen length of the ventral nerve cord with minimal contamination from surrounding tissue [47]. The mass spectra of different individual cells are highly distinctive, but they fall into classes that correspond to morphological and physiological subtypes. The peptides present are few in number, and individual peptides can readily be isolated as single peaks for subsequent sequencing by tandem MS. Once the sequence has been established, the peptide can be synthesized and tested for its bioactivity. Sequence information also enables searches of EST databases for the encoding transcript,

which can then be cloned and used to design and synthesize gene-specific mRNA probes for *in situ* hybridization (ISH). In all cases, ISH has confirmed the expression of the peptide-encoding transcript in the identified motoneuron that contains the peptide itself. ISH also determines the expression of the transcript in other neurons throughout the entire nervous system.

Peptide Characterization by MS of Single Inhibitory Ventral Cord Motorneurons

From the ventral cord (VC), the cell bodies of individual VI and DI neurons were dissected from the anterior 3 segments of the worm (Figure 1a, b) and analyzed by MALDI-TOF MS for their peptide content. Spectra from VI and DI were virtually identical, each containing only 2 intense to moderate peaks, which is relatively few compared to other cell types analyzed in previous studies [23, 41]. All spectra from both VI and DI neurons contained an intense peak with m/z of 1198.7. In 11/17 VI and 12/15 DI spectra, there were also smaller peaks with m/z 1256.7 (Figure 2a, b). Interestingly, direct tissue analysis by MS on a dissected portion of the *A. suum* ventral nerve cord, which included ventral cord inhibitory motor neuron cell bodies and their neurites, indicated the presence of both peaks, among many others, but they had not been previously sequenced [40].

In MALDI-TOF MS, certain on-target chemical modifications are straightforward, and can identify functional groups in peptides. For example, in some preparations, cells were briefly stained with Methylene Blue to aid in cell identification during dissection (see Methods). Exposure to Methylene Blue has been shown to partially oxidize methionine residues to the sulfoxide, causing a +16 shift in m/z ; to a lesser extent, tryptophan residues may be oxidized to give +16 Da or +32 Da adducts [41, 60]. In 3 of 17 VI spectra and 1 of 15 DI spectra we observed

a small peak with an m/z of 1214.7 ($1198.7 + 16$) and another at 1230.7 ($1198.7 + 2 \times 16$) (data not shown). These minor peaks suggested the presence of a tryptophan residue. On-target acetylation produced a mass shift of +42 Da for both the 1198.7 and 1256.7 peaks consistent with the presence of a single amino group in each peptide (Figure 2c, d).

The 1198.7 ion was selected for sequencing by tandem MS. *De novo* interpretation of spectra was done by hand and yielded the sequence S(I/L)ASGRWG(I/L)RP-amide (Figure 3a). At this time, the MALDI-TOF/TOF technology is unable to resolve the isoleucine (I) / leucine (L) ambiguity because the two amino acids are isobaric. A BLAST search of *A. suum* EST library produced the sequence SLASGRWGRLPG (GenBank Accession no. BI593877), which is orthologous to the Ce-NLP-22 peptide found in *C. elegans* [44]. Because it is the *A. suum* version we refer to it as As-NLP-22. The peptide is flanked by dibasic putative cleavage sites and the C-terminal glycine is known to be a substrate for posttranslational amidation, thus the predicted peptide product from the precursor protein is SLASGRWGRLPamide. The MS/MS fragmentation pattern contained nearly the full set of b- and y-ions with an additional set of internal ion fragments. MS/MS spectra of synthetic As-NLP-22 peptide produced a fragmentation pattern nearly identical to that of the native peptide (Figure 3). In MS/MS spectra from both VI and DI neurons, we noticed a single intense peak at m/z 1156.7 ($MH^+ - 42Da$). Typically, the loss of 42Da is due to loss of an acetyl group from a peptide that has been post-translationally acetylated. Knowing this was not a plausible explanation in this case, we hypothesized that the peak was due to the neutral loss of the C-terminal proline side chain ($CH_2CH_2CH_2$). To test this hypothesis we performed MS/MS on 4 additional synthetic peptides with C-terminal prolines, in both the amidated and carboxyl forms, and with variable N-terminal extensions (Appendix 1,

Supplementary Figure 1). In all cases, an intense peak at $MH^+ - 42Da$ was observed, supporting our hypothesis.

In comparing the MS/MS spectra of natural and synthetic As-NLP-22, there was a striking difference, the presence in the natural peptide spectrum of a strong peak at m/z 184.1, and a somewhat less intense peak at m/z 104.1 (Figure 3a). Such peaks have been observed previously in MS/MS spectra of peptides from single neurons [23, 41]. In each case, including the present example of As-NLP-22, the MS/MS spectra of synthetic peptides lacked these peaks. Our present hypothesis is that this peak is phosphocholine (PC: $[M+H]^+$ 184.073 predicted; 184.069 observed), and that the peak at m/z 104.1 is choline. This is supported by the following data. First, the isotopic patterns of this peak are identical with those predicted for PC, and distinct from those of peptides. Second, when synthetic peptide was applied to a MALDI-TOF target plate in the presence of a single dissected neuron (AVK; [23]), MS/MS spectra were obtained both from the zone with pure peptide, and from the region that included the dissected neuron. The results showed that the m/z 184.1 and 104.1 peaks only occurred in the region adjacent to the neuron. The differences in the spectra were just like those illustrated in Figure 3, showing the same distinctive peaks in the natural and synthetic peptides. At present, it is unclear why PC is found in MS/MS spectra, since it would be expected that any ions but those chosen from the MS spectrum for fragmentation would be selectively removed. However, the fact that these peaks are routinely found in spectra of natural peptides, but not of their synthetic versions, suggests that this is a common problem. PC is an abundant naturally occurring component of nematode cells, since it is used directly as a precursor to phosphatidylcholine [61].

The most likely interpretation of the presence of a peptide with m/z 1256.7 is that it is a processing intermediate in the formation of mature As-NLP-22. Its mass is exactly 58 Da more than As-NLP-22, which is explained by the addition of 57 Da from the glycine residue and 1 Da from the mass difference between a C-terminal carboxylic acid and a C-terminal amide. The presence of processing intermediates is not surprising, since the MS analyses are carried out on neuronal cell bodies, where the rough endoplasmic reticulum and the Golgi apparatus, the first sites of processing, are located; processing is completed within the secretory vesicles [29, 62, 63]. C-terminal amidation is the last step in the processing of these peptides, following proteolytic cleavage, trimming by carboxypeptidase, and finally, oxidation of C-terminal glycine residues, leading to the removal of both the carboxyl and α -carbon of glycine, and leaving the amino group as a C-terminal amide. Perhaps it is surprising that these intermediates are not seen more often. In many neurons previously analyzed, no trace of the glycine form of a processed peptide has been detected, although in a few cases, namely the glycine-adduct of AF2 and AF8, the intermediates were detected by MS [42].

A possible explanation of the detection of the glycine adduct in inhibitory motoneurons is that the peptide complement of these cells is very simple, essentially peptide As-NLP-22 and its G adduct; perhaps this simplicity might affect the ionizability of the peptides present during the MALDI process. Unrefined MALDI-TOF MS, such as we have used, is far from being a quantitative technique, and the presence or even the absence of a peptide peak depends not just on the ionizability and concentration of the peptide, but also on the complexity of the mixture of peptides present, as manifested in the phenomenon called ion suppression [64]. However, the AF2 and AF8 adducts were detected in much more complex mixtures than that

found in the inhibitory motoneurons, making it less likely that suppression is affecting the detectability of G-adducts in general. Perhaps there is an effect of sequence on the efficiency of the amidating enzyme complex, or a difference in the expression levels of enzyme and substrate in different neurons affecting the efficiency of amidation. A thorough analysis of the kinetic properties of the *A. suum* amidating enzyme complex and its activity on different substrates would be useful.

Another example of incomplete processing observed in *A. suum* neurons is incomplete proteolysis, often involving KK cleavage sites [40-42], which apparently are cleaved more slowly than KR or R sites. Proteolytic cleavage is known to be more complex than amidation: in *C. elegans*, there are multiple proteases involved [65], and the cellular expression of these enzymes is not known. Local sequence around the cleavage sites and the three-dimensional structure of the precursor protein are also possible factors that may affect cleavage specificity and efficiency.

Cloning of the As-nlp-22 Transcript

To clone the peptide-encoding transcript we initially used primers specific to SL1, a splice leader sequence on ca. 80% of the known *A. suum* transcripts [66], together with a gene-specific reverse primer, but were unsuccessful. However, the use of 5' and 3' RACE with 5'RACE primer *nlp22*-5RACE3 and 3'RACE primer *nlp22*-3RACE allowed us to identify the 5' and 3' regions of the transcript, and to identify a 385 base product. To fuse the two RACE product sequences and confirm the validity of the overlapping regions, a full-length PCR reaction was performed using the forward primer *nlp22*-3RACE2 and the reverse primer *nlp22*-5RACE. The

deduced amino acid sequence contains an 80 amino acid open reading frame, complete with an initiating methionine N-terminal to a predicted signal peptide, a single copy of the peptide flanked by dibasic cleavage sites, and a 3' polyA tail (GenBank Accession no. KF483662) (Figure 3b). Analysis of genomic sequences showed that the *As-nlp-22* gene is composed of two exons. Exon 1 comprises 104 bases and includes the entire signal peptide region, while exon 2 comprises 206 bases and includes the peptide-encoding region. The two exons are separated by an intron of 5073 bases (Figure 3c). There is no evidence that this transcript is trans-spliced to SL1.

Localization of the As-nlp-22 Transcript by in situ Hybridization

To confirm the presence of the *As-nlp-22* mRNA in GABAergic VI and DI ventral cord motorneurons, we used primers *nlp22-3RACE2* and *nlp22-5RACE* to synthesize a 179 base antisense riboprobe (Figure 3b). As a negative control, a sense strand probe was also synthesized. In the first two to three segments of the 30 worms analyzed, 96% of the DI and VI motorneurons present showed moderate to intense staining when exposed to the antisense probe (Figure 4b, c). No staining was observed in the negative control preparations (data not shown). Interestingly, we also observed a pair of bilaterally symmetrical stained cells in the posterior portion of the ventral ganglion (Appendix 1, Supplementary Figure 2a), thought from their position and size to be either AIY or AIM (these neurons are immediate neighbors in the ventral ganglion and have indistinguishable morphology by light microscopy). MS spectra of individual cells dissected from this pair contain peaks at m/z 1198.7 and 1256.7, suggesting that they contain authentic As-NLP-22 and the As-NLP-22+G adduct (Appendix 1, Supplementary

Figure 2b). As in the VI and DI neurons, strong GABA-like immunoreactivity has been detected in one of the AIY/AIM pair [34], but again, they have not been definitively distinguished from each other. We speculate that GABA and As-NLP-22 co-localize in AIY or AIM as they do in the inhibitory ventral cord motoneurons, but more experiments will be needed to test this. If confirmed, it would strengthen the argument that As-NLP-22 and GABA are often co-expressed, and that As-NLP-22 may be important in the GABA signaling pathway.

Bioactivity

a) Muscle Strips

We submitted the peptide As-NLP-22 to functional analysis by measuring its effects on ACh-induced contraction in strips of *A. suum* dorsal muscle, using standard protocols from this laboratory [67, 68]. Exposure to a single dose of 10 μ M peptide practically abolishes any contractile response to ACh ($2.1 \pm 0.4\%$ of the pre-peptide response, $n = 8$ preparations, $p < 0.01$). Once the peptide was washed out, a gradual partial recovery ($\sim 15\%$) was observed over 50 minutes (Figure 5a). Additional concentrations (1 μ M, 100 nM, 10 nM, 1 nM, 0.1 nM, 0.01 nM) of As-NLP-22 were tested to determine the IC_{50} . A dose-dependent inhibition was observed at all concentrations, followed by a variable extent of recovery once peptide was washed out. A dose-response curve was made by plotting the response in the presence of As-NLP-22 at all 7 concentrations as a percentage of the control (Figure 5b). As-NLP-22 exhibits a strong and prolonged inhibitory effect on ACh-induced dorsal muscle strip contraction with an IC_{50} of 8.3×10^{-9} M.

We also observed dose-related peptide effects on baseline muscle tension, observed at the end of the 10 min exposure to peptide. The threshold for this effect was between 10^{-9} and 10^{-8} M; the reduction in tension was 1.2 ± 0.2 g, 1.9 ± 0.4 g, 2.2 ± 0.3 g, and 2.6 ± 0.3 g at peptide concentrations of 10^{-8} , 10^{-7} , 10^{-6} , and 10^{-5} M, respectively. After the peptide was washed out, there was no significant reversal of baseline tension to control levels.

b) *Behavioral Effects of Injection of Peptides into Intact Worms*

To determine the effects of As-NLP-22 on gross locomotory behavior, we injected intact worms with synthetic peptide and observed their behavior in a glass tube with diameter comparable to that of the porcine small intestine, which is the natural habitat of the adult worms. *A. suum* maintain their position in the small intestine by bracing against the walls with pseudo-sinusoidal body waveforms (produced by alternating dorsal and ventral contraction and relaxation of the somatic musculature). These waveforms propagate in the anterior or posterior direction to control forward or backward locomotion, respectively. Anteriorly-propagating waves can be induced by tying surgical thread tightly around the head, producing a stereotypical behavioral response called Head-Restricted Behavior [55, 69, 70]. Peptide solutions can then be injected into the pseudocoelomic cavity posterior to the ligature, and any changes in locomotory behavior recorded.

Injection of 10 μ M As-NLP-22 into ligatured worms abolished all locomotory activity posterior to the ligature and decreased body tonus, resulting in flaccid paralysis (n=10; Figure 5c). These worms did not recover appreciably even after 60 minutes post-injection. This response is in stark contrast to control worms, injected with *Ascaris* saline, which produced

normal anteriorly-propagating waves for the duration of the experiment (n=13). Similarly, worms injected with 0.1 μM As-NLP-22 displayed a reduction in the number of anteriorly-propagating waveforms in the body of the worm, and in their amplitude, although the results were more variable than the 10 μM trials. Seven of ten worms injected with 0.1 μM As-NLP-22 produced no anteriorly-propagating waves; of these, 5 exhibited paralysis within 5-10 minutes post-injection. The remaining 3/10 worms showed no obvious impairment in locomotory behavior. Previous experiments showed that dye solution is diluted ca. 10-fold upon injection into the pseudocoelomic fluid [55], so the peptide concentration resulting from a 0.1 μM As-NLP-22 injection approximates the IC_{50} concentration ($8.3 \times 10^{-9} \text{ M}$) deduced from the muscle strip assay. Since the IC_{50} is the concentration that produces 50% of the maximal response, this could account for the variable behavioral responses observed in these trials.

We also made injections of As-NLP-22 into unligatured worms, at concentrations of 10 μM , 1 μM , and 0.1 μM , with similar results (not shown), suggesting that the neurons of the head and nerve ring do not play an important role in the peptide response. In these worms, as in the ligatured worms, no effects on the three-dimensional head-searching movements were observed.

Comparison with Other Nematodes

It is estimated that the origin of nematode speciation was about 500 million years ago [71]. Nematodes vary dramatically in size and in the ecological niche that each inhabits, yet their basic body plans are highly conserved and the general architecture of nematode nervous systems is very similar. The detailed cellular morphology of neurons has only been determined

in two species, *A. suum* and *C. elegans*, which have vastly different life styles: *A. suum* is an intestinal parasite about 30 cm in length, while *C. elegans* is a free-living soil nematode, ca. 1 mm in length. The morphology of homologous neurons is virtually identical for both the motorneurons (Figure 1c, d), as well as the other neurons in the nervous system; they differ in size. The expression of classical transmitters in the motor neurons is also highly conserved, with exciters using ACh and inhibitors using GABA.

Structural conservation also extends to neuropeptide sequences. For example, there are 4 readily identifiable families of peptides, found in 22 other nematodes [72], that share a C-terminal -RPamide sequence. As-NLP-22 belongs to one of these, encoded by the *As-nlp-22* gene. The 3 other families are encoded by the *nlp-2*, *-23*, and *-46* genes. Confirmation of the peptide sequence of As-NLP-22 and cloning of its peptide-encoding transcript suggests that *A. suum* SLASGRWGLRPamide is the homolog of Ce-NLP-22 and not Ce-NLP-2 as was previously reported [42]. Homologous *nlp-22* genes from other nematodes encode a predicted precursor protein that includes a single copy of the peptide with the characteristic motif S[A/M/L]A[I/N/S]GR[A/M/W][G/Q][M/F/L]RPG, flanked by dibasic or monobasic putative cleavage sites.

We conducted BLAST searches of nematode ESTs using both As-NLP-22 and Ce-NLP-22 as queries and identified homologous peptides from 10 other parasitic and free-living nematodes (Appendix 1, Supplementary Figure 3), doubling the number of nematode NLP-22 encoding transcripts previously reported by McVeigh *et al.* (2008). When sequences were aligned, they fell into two groups differing slightly in their sequence within the conserved nematode NLP-22 motif (Figure 6). Group 1, from clades IVa and V [71, 73] was typified by a C-

terminal –FRPamide while group 2, from clades III and IVb, were all from parasitic species and shared a C-terminal –LRPamide sequence. Sequences within members of a group were highly conserved (Group 1: SIAIGR[A/S]GFRPamide; Group 2: SLA[S/N]GRW[Q/G]LRPamide). The peptide from parasitic nematode *S. stercoralis* is an exception; it appears to be an amalgam of the two, sharing an N-terminal SL- with Group 2 and C-terminal –AIGR[X]GFRPamide with Group 1. Eventually, it will be interesting to see if these sequence differences are correlated with functional differences, or if they are simply markers of the phylogeny, with shared function.

Little is known about the specific cellular expression of NLP-22 in other nematodes, with the exception of *C. elegans*. *Ce-nlp-22::GFP* was shown to be expressed in the pair of RIA neurons, head interneurons that regulate a nematode sleep-like state known as lethargus [74]. Over-expression of *Ce-nlp-22* in adult *C. elegans* causes a decrease in locomotion and feeding behavior. Although the structure of *As-nlp-22* and *Ce-nlp-22* are highly similar [72], their reported cellular expression is completely different in the two species, with no expression of *Ce-nlp-22* in inhibitory motoneurons [74].

We have previously suggested that differences in cellular expression of homologous peptides could be at least part of the explanation of how organisms with virtually the same nervous systems can generate different behaviors that enable them to inhabit such different ecological niches. However, a note of caution must be added, since in most cases the cellular localization in *C. elegans* was determined by using GFP constructs that included the promoter and various lengths of 5' sequence. It is known that sequences controlling gene expression can be remote from the promoter and may also include sequences 3' to the gene, or intronic

regions [75, 76], so it is hard to know whether the control sequences of the constructs are complete. Corroboratory methods are badly needed. In *A. suum*, we have used *in situ* hybridization to confirm the localizations observed by MS. Importantly, both methods directly detect the gene products, either the peptides themselves or the peptide-encoding transcript, expressed in the normal worm, without any possible interference in gene expression that could be induced by the introduction of additional gene copies. In addition, in cases where we could raise monospecific antibodies recognizing a single peptide [23], we have used immunocytochemistry as a third technique that also validates the cellular expression. For all three techniques, the large size of *A. suum* neurons is an enormous advantage.

CONCLUSIONS AND FUTURE DIRECTIONS

This work identifies the first peptide expressed in *A. suum* inhibitory motoneurons. As-NLP-22 is found in the VI and DI inhibitory motoneurons of the ventral cord. We have demonstrated that As-NLP-22 has a potent inhibitory effect on ACh-induced muscle contraction and locomotory behavior. At the moment, the cellular basis for the inhibitory action of As-NLP-22 is unknown. Further investigation of As-NLP-22 by electrophysiological methods, to determine its mechanism of action and postsynaptic partners, will add to the description of the circuitry that controls locomotion and provide valuable information on developing targets for effective drug treatments of nematode parasites. We will also investigate potential interactions between GABA and this peptide, since they are co-expressed in the inhibitory motoneurons.

AUTHOR CONTRIBUTIONS

C. J. Konop was responsible for single cell dissection, MS, bioinformatics, data analysis, transcript identification, and manuscript preparation. J. J. Knickelbine was responsible for transcript identification, ISH, muscle tension assays, bioinformatics, and manuscript preparation. M. S. Sygulla was responsible for single cell dissections. C. D. Wruck was responsible for whole worm injection experiments. M. M. Vestling was responsible for MS and data analysis. A. O. W. Stretton was responsible for data analysis, bioinformatics, and manuscript preparation.

FIGURES

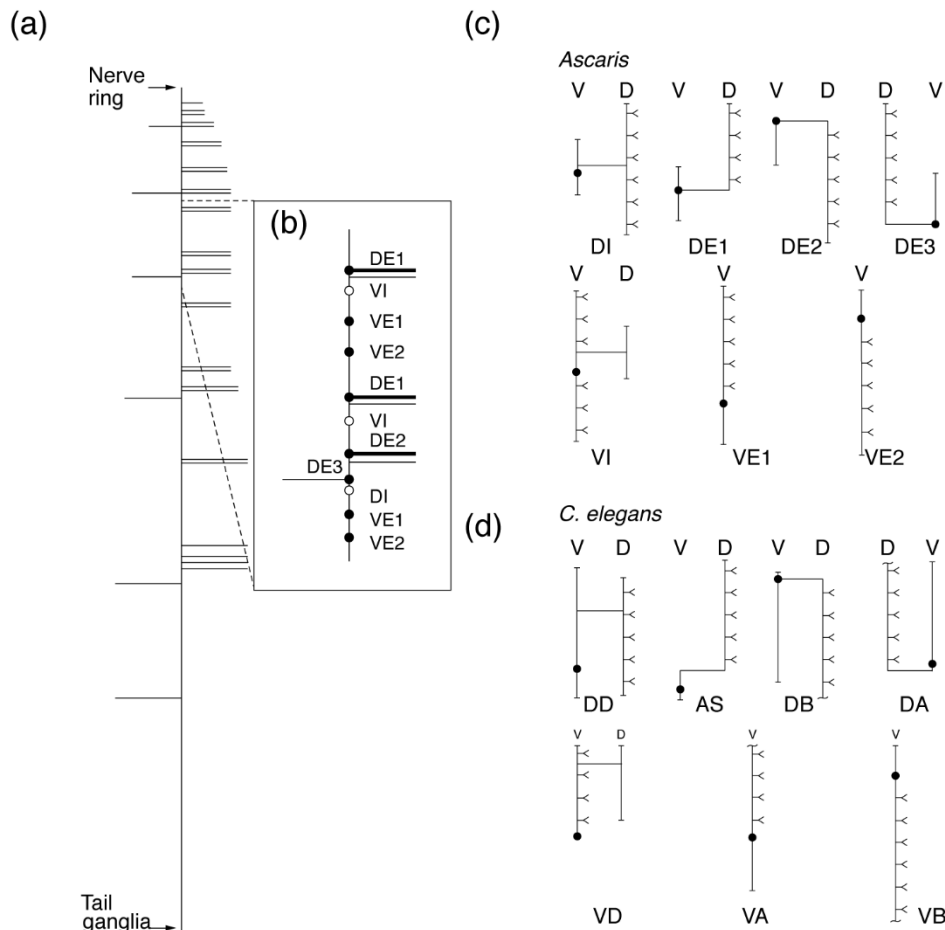


Figure 1. Schematic diagram of the ventral cord motorneurons of *A. suum*. (a) Diagram of the ventral nerve cord adapted from Stretton et al. (1978) showing the position and handedness of the motorneuron commissures. The ventral nerve cord is represented by the vertical line. Right- and left-handed commissures leave the cord in an ordered array, indicated by the horizontal lines. Arrows show the positions of the nerve ring and tail ganglia. Some of the commissures at the anterior end of the ventral cord have been omitted for clarity. (b) Stereotypical pattern of motorneuron cell bodies found in each segment (segments 2-5) and their associated commissures. This figure shows the third segment. Filled circles indicate cell bodies of excitatory motorneurons. Open circles indicate cell bodies of inhibitory motorneurons. DE1 and DE2 commissures have a larger diameter than the inhibitory commissures with which they are paired [36]. (c, d) Schematic diagram of the seven types of (c) *A. suum* and (d) *C. elegans* segmental motorneurons. Filled circles indicate cell bodies, all found in the ventral nerve cord (V). Vertical lines emanating from the cell body represent the ventral nerve cord processes, and horizontal lines represent commissures that connect dorsal (D) and ventral processes. Forked projections represent axonal synapses onto muscle.

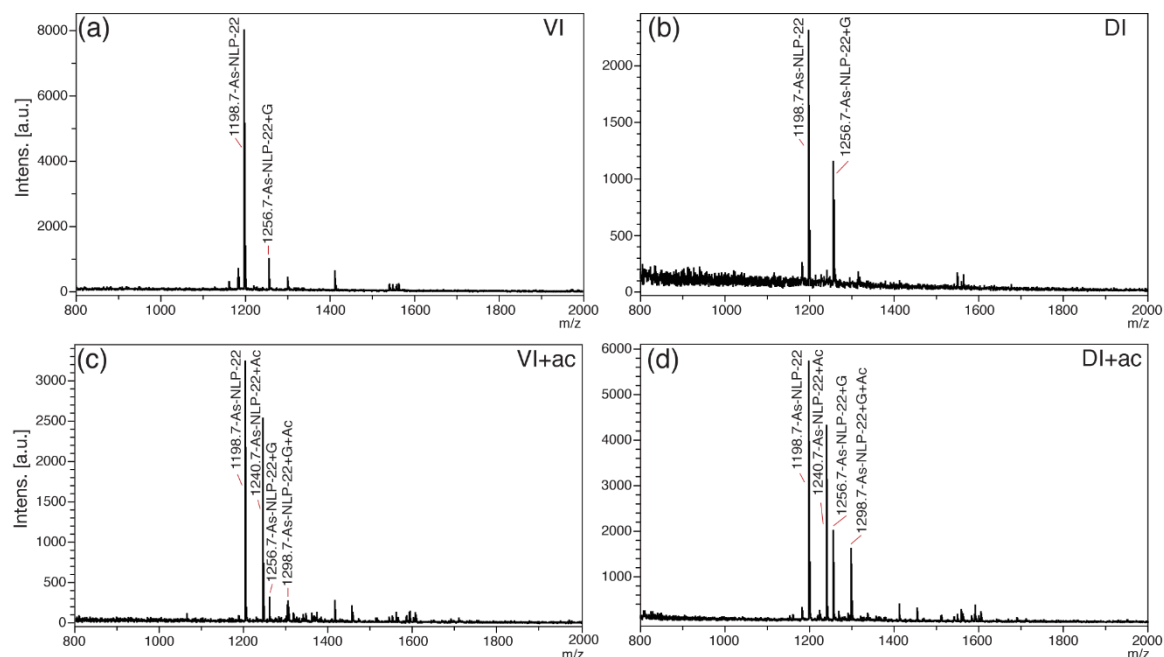


Figure 2. Neuropeptides from single ventral cord inhibitory motoneurons. (a, b) Mass spectra from individual (a) VI and (b) DI neurons. (c, d) Spectra from individual (c) VI and (d) DI neurons treated with acetic anhydride each causing a +42 Da mass shift. In these peptides, only the N-terminal α -amino group is substituted.

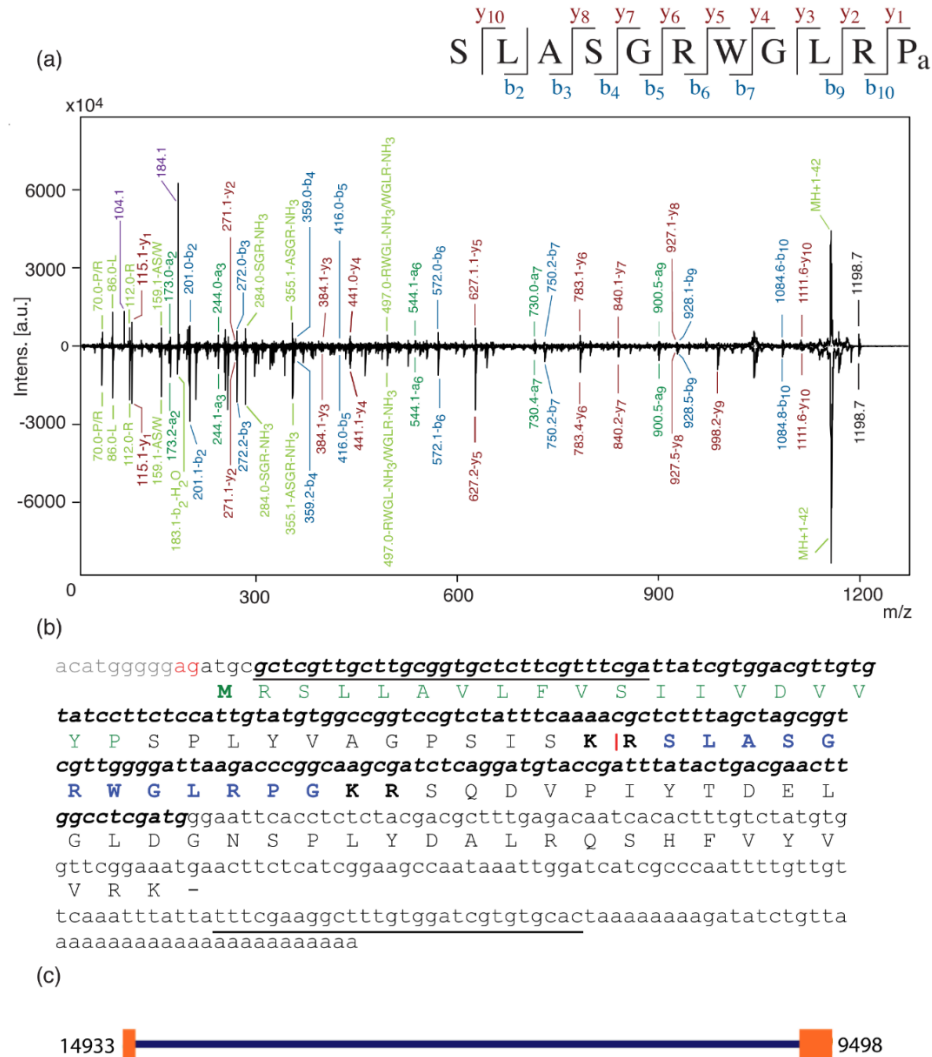


Figure 3. Identification of novel peptide As-NLP-22 in the dorsal and ventral inhibitory motorneurons. (a- upper spectrum) MS/MS of the m/z 1198.7 peak from single DI/VI motorneurons. (a- lower spectrum) MS/MS of synthetic form of As-NLP-22 reveals nearly identical fragmentation pattern. Peaks representing a- (dark green), b- (blue), y- ions (red), and internal fragments and immonium ions (light green) are labeled. b and y-ions are summarized in the sequence at the top of the spectrum. Peaks with the m/z 104.1 and 184.1 from the fragmentation of a PC head group are seen in MS/MS spectrum of native As-NLP-22, but not that from the synthetic peptide (purple). (b) Nucleotide sequence of the *As-nlp-22* transcript, deduced by 5' and 3' RACE and confirmed by PCR using unique primers underlined in black (GenBank Accession no. KF483662). Green amino acids indicate the signal peptide, with the putative start site in bold. Blue amino acid sequence indicates the encoded peptide, flanked by dibasic cleavage sites in bold. Bolded nucleotide sequence is the portion of the transcript targeted by the *As-nlp-22* riboprobe. Red bases indicate start of exon 1, and the vertical red line shows where two exons are joined together. (c) *As-nlp-22* gene model. The gene consists of 2 exons of 104 and 206 bases, respectively. Exons are separated by an intron 5,073 bases in length.

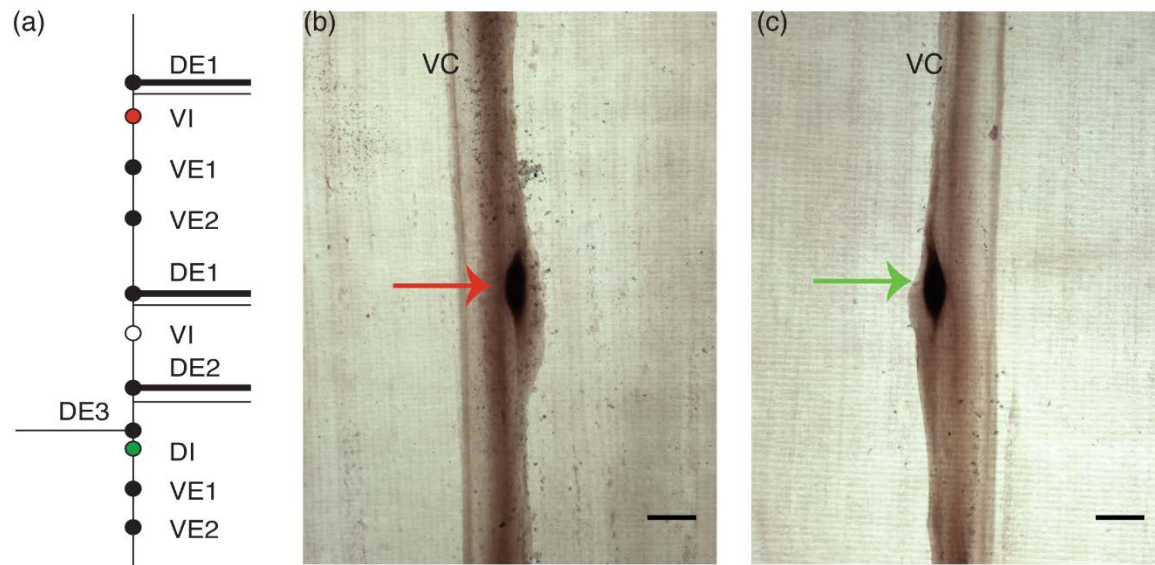


Figure 4. Expression of *As-nlp-22* as determined by ISH. (a) Diagram of *A. suum* ventral cord motorneurons in one of the repeating units. Indicated in red (VI) and green (DI) are the cell bodies of neurons shown in b and c. (b) VI neuron cell body stained with the *As-nlp-22* riboprobe (red arrow). (c) DI neuron cell body stained with the *As-nlp-22* riboprobe (green arrow). Scale bars: 100 μ m.

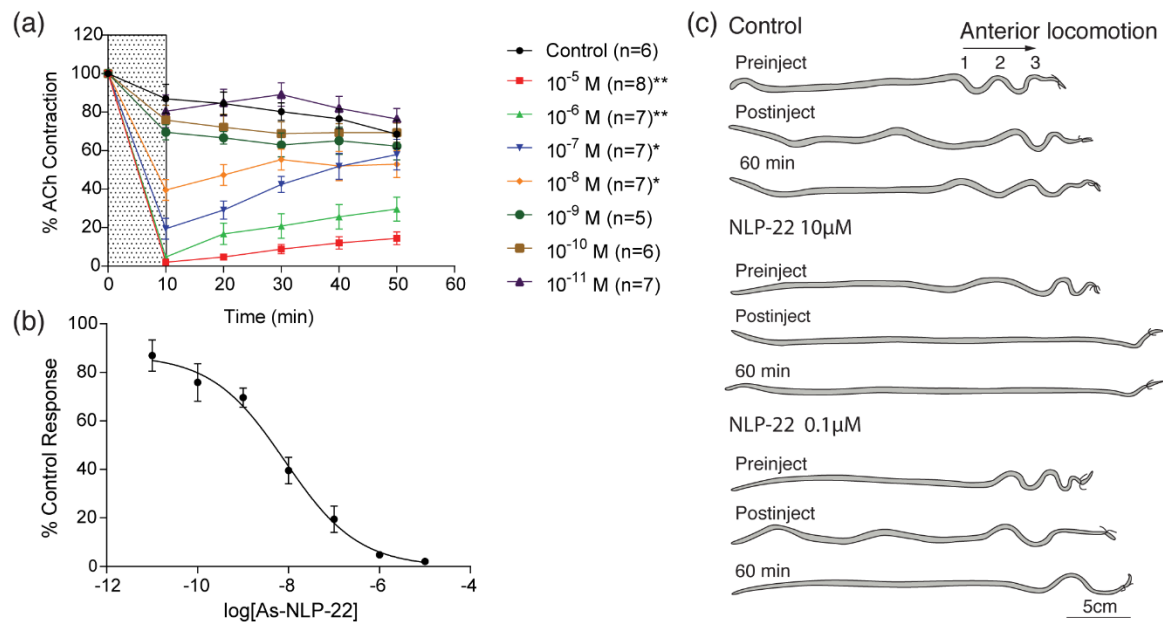


Figure 5. Effects of As-NLP-22 on ACh-induced muscle contraction and gross locomotion. (a) Effects of single exposure to synthetic As-NLP-22 on ACh-induced muscle contraction. Peptide was introduced at 0 min and washed out at 10 min, depicted by the shaded region. Strength of contraction is normalized to the initial contraction prior to the application of peptide for each individual worm. Error bars represent standard error of the mean. At concentrations denoted with asterisks, responses were significantly different from control worms at all time points (** $p < 0.01$, * $p < 0.05$). (b) Dose-response of As-NLP-22. Each data point represents the response in the presence of peptide, expressed as a percentage of the control response. $IC_{50} = 8.3 \times 10^{-9}$ M. (c) Injection of As-NLP-22 into intact worms causes reduced locomotory behavior. Control worms injected with *Ascaris* saline showed no discernible impairment in locomotory activity (n=13). Worms injected with 10 μ M As-NLP-22 displayed paralysis for the duration of the experiment (n=10). Worms injected with 0.1 μ M As-NLP-22 (approximates IC_{50} *in vivo*) displayed a general reduction in locomotory behavior (n=10), though the behavioral responses were more varied than the 10 μ M trials.

Figure 6. Multiple alignment of nematode *nlp-22* transcripts by MEGA5.1 [56]. Results from 11 nematode species, organized alphabetically by clade, were presented in Jalview [56, 58], and show strong conservation of the predicted encoded peptide and signal peptide. Asu = *Ascaris suum* (Clade III); Sst = *Strongyloides stercoralis* (IVa); Aav = *Aphelenchus avenae* (IVb); Gro = *Globodera rostochiensis* (IVb); Mar = *Meloidogyne arenaria* (IVb); Mch = *Meloidogyne chitwoodi* (IVb); Min = *Meloidogyne incognita* (IVb); Cbg = *Caenorhabditis briggsae* (V); Cbr = *Caenorhabditis brenneri* (V); Cel = *Caenorhabditis elegans* (V); Cre = *Caenorhabditis remanei* (V).

REFERENCES

1. Wiersma, C. A. G., California Institute of Technology. (1967). Invertebrate nervous systems, their significance for mammalian neurophysiology; [papers]. University of Chicago Press. Chicago,.
2. Getting, P. A. (1989). Emerging principles governing the operation of neural networks. *Annu Rev Neurosci.* 12:185-204.
3. Bargmann, C. I. (2012). Beyond the connectome: how neuromodulators shape neural circuits. *Bioessays.* 34:458-65.
4. Ma, M., Szabo, T. M., Jia, C., Marder, E., Li, L. (2009). Mass spectrometric characterization and physiological actions of novel crustacean C-type allatostatins. *Peptides.* 30:1660-8.
5. Hui, L., Zhang, Y., Wang, J., Cook, A., Ye, H., Nusbaum, M. P., Li, L. (2011). Discovery and functional study of a novel crustacean tachykinin neuropeptide. *ACS Chem Neurosci.* 2:711-22.
6. Hui, L., Xiang, F., Zhang, Y., Li, L. (2012). Mass spectrometric elucidation of the neuropeptidome of a crustacean neuroendocrine organ. *Peptides.* 36:230-9.
7. Ye, H., Hui, L., Kellersberger, K., Li, L. (2013). Mapping of neuropeptides in the crustacean stomatogastric nervous system by imaging mass spectrometry. *J Am Soc Mass Spectrom.* 24:134-47.
8. Turrigiano, G. G., Selverston, A. I. (1989). Cholecystokinin-like peptide is a modulator of a crustacean central pattern generator. *J Neurosci.* 9:2486-501.
9. Turrigiano, G. G., Van Wormhoudt, A., Ogden, L., Selverston, A. I. (1994). Partial purification, tissue distribution and modulatory activity of a crustacean cholecystokinin-like peptide. *J Exp Biol.* 187:181-200.
10. Selverston, A. I., Moulins, M. (1987). The Crustacean Stomatogastric System. Springer-Verlag. Berlin. 1-338
11. Selverston, A. I., Russell, D. F., Miller, J. P. (1976). The stomatogastric nervous system: structure and function of a small neural network. *Prog Neurobiol.* 7:215-90.
12. Bucher, D., Johnson, C. D., Marder, E. (2007). Neuronal morphology and neuropil structure in the stomatogastric ganglion of the lobster, *Homarus americanus*. *J Comp Neurol.* 501:185-205.
13. Ma, M., Wang, J., Chen, R., Li, L. (2009). Expanding the Crustacean neuropeptidome using a multifaceted mass spectrometric approach. *J Proteome Res.* 8:2426-37.

14. Chen, R., Ouyang, C., Xiao, M., Li, L. (2014). In situ identification and mapping of neuropeptides from the stomatogastric nervous system of *Cancer borealis*. *Rapid Commun Mass Spectrom.* 28:2437-44.
15. Harris-Warrick, R. M., Marder, E. (1991). Modulation of neural networks for behavior. *Annu Rev Neurosci.* 14:39-57.
16. Marder, E., Bucher, D. (2007). Understanding circuit dynamics using the stomatogastric nervous system of lobsters and crabs. *Annu Rev Physiol.* 69:291-316.
17. Johnson, B. R., Schneider, L. R., Nadim, F., Harris-Warrick, R. M. (2005). Dopamine modulation of phasing of activity in a rhythmic motor network: contribution of synaptic and intrinsic modulatory actions. *J Neurophysiol.* 94:3101-11.
18. Sigvardt, K. A., Mulloney, B. (1982). Sensory alteration of motor patterns in the stomatogastric nervous system of the spiny lobster *Panulirus interruptus*. *J Exp Biol.* 97:137-52.
19. Sturm, R. M., Greer, T., Woodards, N., Gemperline, E., Li, L. J. (2013). Mass Spectrometric Evaluation of Neuropeptidomic Profiles upon Heat Stabilization Treatment of Neuroendocrine Tissues in Crustaceans. *J Proteome Res.* 12:743-52.
20. Williams, A. H., Calkins, A., O'Leary, T., Symonds, R., Marder, E., Dickinson, P. S. (2013). The neuromuscular transform of the lobster cardiac system explains the opposing effects of a neuromodulator on muscle output. *J Neurosci.* 33:16565-75.
21. Stretton, A., Donmoyer, J., Davis, R., Meade, J., Cowden, C., Sithigorngul, P. (1992). Motor behavior and motor nervous system function in the nematode *Ascaris suum*. *J Parasitol.* 78:206-14.
22. White, J. G., Southgate, E., Thomson, J. N., Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci.* 314:1-340.
23. Jarecki, J. L., Viola, I. R., Andersen, K. M., Miller, A. H., Ramaker, M. A., Vestling, M. M., Stretton, A. O. (2012). Three independent techniques localize expression of transcript *afp-11* and its bioactive peptide products to the paired AVK neurons in *Ascaris suum* : in situ hybridization, immunocytochemistry, and single cell mass spectrometry. *ACS Chem Neurosci.* 4:418-34.
24. Yew, J. Y., Davis, R., Dikler, S., Nanda, J., Reinders, B., Stretton, A. O. (2007). Peptide products of the *afp-6* gene of the nematode *Ascaris suum* have different biological actions. *J Comp Neurol.* 502:872-82.

25. Husson, S. J., Clynen, E., Baggerman, G., De Loof, A., Schoofs, L. (2005). Discovering neuropeptides in *Caenorhabditis elegans* by two dimensional liquid chromatography and mass spectrometry. *Biochem. Biophys. Res. Commun.* 335:76-86.
26. Nassel, D. R. (2000). Functional roles of neuropeptides in the insect central nervous system. *Naturwissenschaften.* 87:439-49.
27. Cowden, C., Sithigorngul, P., Brackley, P., Guastella, J., Stretton, A. O. (1993). Localization and differential expression of FMRFamide-like immunoreactivity in the nematode *Ascaris suum*. *J Comp Neurol.* 333:455-68.
28. Cowden, C., Stretton, A. O. (1995). Eight novel FMRFamide-like neuropeptides isolated from the nematode *Ascaris suum*. *Peptides.* 16:491-500.
29. Li, C., Kim, K. (2008). Neuropeptides. *WormBook.* 1-36.
30. Stretton, A. O., Fishpool, R. M., Southgate, E., Donmoyer, J. E., Walrond, J. P., Moses, J. E., Kass, I. S. (1978). Structure and physiological activity of the motoneurons of the nematode *Ascaris*. *Proc Natl Acad Sci U S A.* 75:3493-7.
31. Davis, R. E., Stretton, A. O. (2001). Structure-activity relationships of 18 endogenous neuropeptides on the motor nervous system of the nematode *Ascaris suum*. *Peptides.* 22:7-23.
32. Walrond, J. P., Kass, I. S., Stretton, A. O., Donmoyer, J. E. (1985). Identification of excitatory and inhibitory motoneurons in the nematode *Ascaris* by electrophysiological techniques. *J Neurosci.* 5:1-8.
33. Angstadt, J. D., Donmoyer, J. E., Stretton, A. O. W. (2001). The number of morphological synapses between neurons does not predict the strength of their physiological synaptic interactions: A study of dendrites in the nematode *Ascaris suum*. *The Journal of Comparative Neurology.* 432:512-27.
34. Guastella, J., Johnson, C. D., Stretton, A. O. (1991). GABA-immunoreactive neurons in the nematode *Ascaris*. *J Comp Neurol.* 307:584-97.
35. Guastella, J., Stretton, A. O. (1991). Distribution of 3H-GABA uptake sites in the nematode *Ascaris*. *J Comp Neurol.* 307:598-608.
36. Johnson, C. D., Stretton, A. O. W. (1985). Localization of choline-acetyltransferase within identified motoneurons of the nematode *Ascaris*. *J Neurosci.* 5:1984-92.
37. Johnson, C. D., Stretton, A. O. (1987). GABA-immunoreactivity in inhibitory motor neurons of the nematode *Ascaris*. *J Neurosci.* 7:223-35.

38. Cowden, C., Stretton, A. O. (1993). AF2, an *Ascaris* neuropeptide: isolation, sequence, and bioactivity. *Peptides*. 14:423-30.
39. Cowden, C., Stretton, A. O., Davis, R. E. (1989). AF1, a sequenced bioactive neuropeptide isolated from the nematode *Ascaris suum*. *Neuron*. 2:1465-73.
40. Yew, J. Y., Kutz, K. K., Dikler, S., Messinger, L., Li, L., Stretton, A. O. (2005). Mass spectrometric map of neuropeptide expression in *Ascaris suum*. *J Comp Neurol*. 488:396-413.
41. Jarecki, J. L., Andersen, K., Konop, C. J., Knickelbine, J. J., Vestling, M. M., Stretton, A. O. (2010). Mapping neuropeptide expression by mass spectrometry in single dissected identified neurons from the dorsal ganglion of the nematode *Ascaris suum*. *ACS Chem Neurosci*. 1:505-19.
42. Jarecki, J. L., Frey, B. L., Smith, L. M., Stretton, A. O. (2011). Discovery of neuropeptides in the nematode *Ascaris suum* by database mining and tandem mass spectrometry. *J Proteome Res*. 10:3098-106.
43. Li, L., Sweedler, J. V. (2008). Peptides in the brain: mass spectrometry-based measurement approaches and challenges. *Annu Rev Anal Chem (Palo Alto Calif)*. 1:451-83.
44. Nathoo, A. N., Moeller, R. A., Westlund, B. A., Hart, A. C. (2001). Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. *Proc Natl Acad Sci U S A*. 98:14000-5.
45. Pierce, S. B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S. A., Buchman, A. R., Ferguson, K. C., Heller, J., Platt, D. M., Pasquinelli, A. A., Liu, L. X., Doberstein, S. K., Ruvkun, G. (2001). Regulation of DAF-2 receptor signaling by human insulin and ins-1, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev*. 15:672-86.
46. Nelson, L. S., Kim, K., Memmott, J. E., Li, C. (1998). FMRFamide-related gene family in the nematode, *Caenorhabditis elegans*. *Brain Res Mol Brain Res*. 58:103-11.
47. Rubakhin, S. S., Greenough, W. T., Sweedler, J. V. (2003). Spatial profiling with MALDI MS: Distribution of neuropeptides within single neurons. *Anal Chem*. 75:5374-80.
48. Rubakhin, S. S., Sweedler, J. V. (2007). Characterizing peptides in individual mammalian cells using mass spectrometry. *Nat Protoc*. 2:1987-97.
49. Rubakhin, S. S., Churchill, J. D., Greenough, W. T., Sweedler, J. V. (2006). Profiling signaling peptides in single mammalian cells using mass spectrometry. *Anal Chem*. 78:7267-72.

50. Nemes, P., Knolhoff, A. M., Rubakhin, S. S., Sweedler, J. V. (2012). Single-cell metabolomics: changes in the metabolome of freshly isolated and cultured neurons. *ACS Chem Neurosci.* 3:782-92.
51. Schwartz, S. A., Reyzer, M. L., Caprioli, R. M. (2003). Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation. *J Mass Spectrom.* 38:699-708.
52. Hilker, B., Clifford, K. J., Sauter, A. D., Jr., Sauter, A. D., 3rd, Harmon, J. P. (2009). The measurement of charge for induction-based fluidic MALDI dispense event and nanoliter volume verification in real time. *J Am Soc Mass Spectrom.* 20:1064-7.
53. Jex, A. R., Liu, S., Li, B., Young, N. D., Hall, R. S., Li, Y., Yang, L., Zeng, N., Xu, X., Xiong, Z., Chen, F., Wu, X., Zhang, G., Fang, X., Kang, Y., Anderson, G. A., Harris, T. W., Campbell, B. E., Vlaminc, J., Wang, T., Cantacessi, C., Schwarz, E. M., Ranganathan, S., Geldhof, P., Nejsun, P., Sternberg, P. W., Yang, H., Wang, J., Wang, J., Gasser, R. B. (2011). *Ascaris suum* draft genome. *Nature.* 479:529-33.
54. Nanda, J. C., Stretton, A. O. (2010). In situ hybridization of neuropeptide-encoding transcripts *afp-1*, *afp-3*, and *afp-4* in neurons of the nematode *Ascaris suum*. *J Comp Neurol.* 518:896-910.
55. Reinitz, C. A., Stretton, A. O. (1996). Behavioral and cellular effects of serotonin on locomotion and male mating posture in *Ascaris suum* (Nematoda). *J Comp Physiol A.* 178:655-67.
56. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 28:2731-9.
57. Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792-7.
58. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* 25:1189-91.
59. Davis, R. E., Stretton, A. O. (1996). The motornervous system of *Ascaris*: electrophysiology and anatomy of the neurons and their control by neuromodulators. *Parasitology.* 113 Suppl:S97-117.
60. Murachi, T., Tsudzuki, T., Okumura, K. (1975). Photosensitized inactivation of stem bromelain. Oxidation of histidine, methionine, and tryptophan residues. *Biochemistry.* 14:249-55.

61. Palavalli, L. H., Brendza, K. M., Haakenson, W., Cahoon, R. E., McLaird, M., Hicks, L. M., McCarter, J. P., Williams, D. J., Hresko, M. C., Jez, J. M. (2006). Defining the role of phosphomethylethanolamine N-methyltransferase from *Caenorhabditis elegans* in phosphocholine biosynthesis by biochemical and kinetic analysis. *Biochemistry*. 45:6056-65.
62. Kreil, G. (1984). Occurrence, detection, and biosynthesis of carboxy-terminal amides. *Methods Enzymol*. 106:218-23.
63. Bradbury, A. F., Smyth, D. G. (1987). Biosynthesis of the C-terminal amide in peptide hormones. *Biosci Rep*. 7:907-16.
64. Annesley, T. M. (2003). Ion suppression in mass spectrometry. *Clin Chem*. 49:1041-4.
65. McVeigh, P., Leech, S., Mair, G. R., Marks, N. J., Geary, T. G., Maule, A. G. (2005). Analysis of FMRFamide-like peptide (FLP) diversity in phylum Nematoda. *Int J Parasitol*. 35:1043-60.
66. Maroney, P. A., Denker, J. A., Darzynkiewicz, E., Laneve, R., Nilsen, T. W. (1995). Most mRNAs in the nematode *Ascaris lumbricoides* are trans-spliced: a role for spliced leader addition in translational efficiency. *RNA*. 1:714-23.
67. Ash, A. S., Tucker, J. F. (1967). The bioassay of gamma-aminobutyric acid using a muscle preparation from *Ascaris lumbricoides*. *J Pharm Pharmacol*. 19:240-5.
68. Kass, I. S., Larsen, D. A., Wang, C. C., Stretton, A. O. (1982). *Ascaris suum*: differential effects of avermectin B1a on the intact animal and neuromuscular strip preparations. *Exp Parasitol*. 54:166-74.
69. Reinitz, C. A., Herfel, H. G., Messinger, L. A., Stretton, A. O. (2000). Changes in locomotory behavior and cAMP produced in *Ascaris suum* by neuropeptides from *Ascaris suum* or *Caenorhabditis elegans*. *Mol Biochem Parasitol*. 111:185-97.
70. Reinitz, C. A., Pleva, A. E., Stretton, A. O. (2011). Changes in cyclic nucleotides, locomotory behavior, and body length produced by novel endogenous neuropeptides in the parasitic nematode *Ascaris suum*. *Mol Biochem Parasitol*. 180:27-34.
71. Vanfleteren, J. R., Van de Peer, Y., Blaxter, M. L., Tweedie, S. A., Trotman, C., Lu, L., Van Hauwaert, M. L., Moens, L. (1994). Molecular genealogy of some nematode taxa as based on cytochrome c and globin amino acid sequences. *Mol Phylogenet Evol*. 3:92-101.
72. McVeigh, P., Alexander-Bowman, S., Veal, E., Mousley, A., Marks, N. J., Maule, A. G. (2008). Neuropeptide-like protein diversity in phylum Nematoda. *Int J Parasitol*. 38:1493-503.

73. Blaxter, M. L., De Ley, P., Garey, J. R., Liu, L. X., Scheldeman, P., Vierstraete, A., Vanfleteren, J. R., Mackey, L. Y., Dorris, M., Frisse, L. M., Vida, J. T., Thomas, W. K. (1998). A molecular evolutionary framework for the phylum Nematoda. *Nature*. 392:71-5.
74. Nelson, M. D., Trojanowski, N. F., George-Raizen, J. B., Smith, C. J., Yu, C. C., Fang-Yen, C., Raizen, D. M. (2013). The neuropeptide NLP-22 regulates a sleep-like state in *Caenorhabditis elegans*. *Nat Commun*. 4:2846.
75. Etchberger, J. F., Lorch, A., Sleumer, M. C., Zapf, R., Jones, S. J., Marra, M. A., Holt, R. A., Moerman, D. G., Hobert, O. (2007). The molecular signature and cis-regulatory architecture of a *C. elegans* gustatory neuron. *Genes Dev*. 21:1653-74.
76. Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., Ye, Z., Lee, L. K., Stuart, R. K., Ching, C. W., Ching, K. A., Antosiewicz-Bourget, J. E., Liu, H., Zhang, X., Green, R. D., Lobanov, V. V., Stewart, R., Thomson, J. A., Crawford, G. E., Kellis, M., Ren, B. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 459:108-12.

CHAPTER THREE

Neuropeptides Related to As-NLP-22 Have Different Effects on *A. suum* Muscle Contraction

Jennifer J. Knickelbine, Christopher J. Konop, and Antony O. W. Stretton

INTRODUCTION

The somatic inhibitory motorneurons of *A. suum* produce a single neuropeptide, As-NLP-22 (SLASGRWGLRPamide), which profoundly affects ACh-induced muscle contraction in strips of dorsal muscle, and induces flaccid paralysis when injected into intact behaving worms [1]. In *C. elegans*, the NLP-22 peptide is a member of the -RPamide family, which includes peptides from the *nlp-2*, *nlp-23*, and *nlp-46* transcripts [2]. Very little is known about the expression and function of peptides in this family, but homologs have also been identified in *A. suum*, which we are calling the *As-nlp-2*, *As-nlp-23*, and *As-nlp-46* peptides. Given the dramatic and prolonged effect of As-NLP-22 and the highly conserved sequences of -RPamide peptides in *A. suum*, we hypothesized that the rest of the -RPamide peptides could also have strong effects on behavior. In this manuscript, we identify the full mRNA transcripts, cellular localization (via *in situ* hybridization), and pharmacological effects of these peptides on muscle contraction in *A. suum*. Despite their sequence relatedness, these peptides show divergent effects on muscle contraction, with the sign of the activity (excitatory vs. inhibitory) correlating to amino acids at positions 4 and 8. This work provides some insight into the action of As-NLP-22, including residues that are crucial for producing its strong inhibitory effect.

MATERIALS AND METHODS

Specimen Collection, RNA Isolation, and Peptide Synthesis

Large female *Ascaris suum* were collected and maintained at 37°C in PBS as previously described [1]. RNA isolation and preparation of cDNA for PCR and RACE was performed as detailed in Konop *et al.* (2015). All peptides were synthesized by the University of Wisconsin Biotechnology Center and stored as lyophilized powders at -20°C.

Identification of the As-nlp-2, As-nlp-23, and As-nlp-46 Transcripts

Primers for PCR and RACE were designed from *Ascaris* EST sequences, obtained through BLAST searches with the *C. elegans* homologs. Accession numbers are provided in the discussion of the individual transcripts. A summary of the primer sequences is provided in Table 1. PCR and RACE experiments were performed as previously described [1]. In the case of the *nlp-23* transcript, nested 5'RACE was performed with the outer primer *nlp23-5RACE* and nested primer *nlp23-5RACE2* to further amplify the product. Cloning, sequencing, and riboprobe generation were performed as outlined by Konop *et al.* (2015), with riboprobes created using the primers denoted with asterisks in Table 1.

Localization by in situ Hybridization

Preparations containing the anterior 2-3 segments of the worm were dissected, fixed, and stored using our routine methods [1]. The *in situ* hybridization protocol was performed as previously described [3-5], but with increased incubation in proteinase K solution and overnight staining [1]. The stained preparations were mounted in Clear Mount (Electron Microscopy

Sciences) and photographs were taken with a Zeiss Universal microscope equipped with a MRc digital camera and Axio Vision AC software.

Dorsal Muscle Strip Assay

Strips of dorsal muscle were dissected from female *A. suum* 27-32 cm in length and attached to the recording apparatus as previously described [1, 5]. Measurements of ACh-induced contraction were measured before, during, and after exposure to peptide. The peptides used for these experiments were the second peptides on the *As-nlp-2* and *As-nlp-23.2* transcripts, As-NLP-2.2 and As-NLP-23.2, respectively, and the As-NLP-46 peptide. Responses were reported as a percentage of the maximum (pre-peptide) response to 5 μ M ACh.

RESULTS AND DISCUSSION

The first -RPamide peptides in *A. suum* were predicted by database mining of the *A. suum* expressed sequence tag (EST) and genomic survey sequence (GSS) libraries, as part of an effort to compile a database of putative neuropeptide sequences for analyzing mass spectra [6]. Jarecki *et al.* (2011) also detected three of these peptides, now known as As-NLP-2.2 (SLALGRVDFRPamide), As-NLP-23.2 (SIALGRFSLRPamide), and As-NLP-46 (NIAIGRGDGFRPamide), by liquid chromatography-tandem mass spectrometry (LC-MS/MS) of peptide extracts from frozen *A. suum* heads. However, no further effort was made to identify the transcripts, precise localization, or functional significance of these peptides. We have since shown that one member of the -RPamide family, As-NLP-22 (SLASGRWGLRPamide), is expressed by the inhibitory motoneurons and dramatically inhibits muscle contraction (IC_{50} =

8.3 nM, [1]). Therefore, we were eager to see whether related peptides from the -RPamide family had similar relevance to nematode locomotion.

Identification of -RPamide Transcripts

(a) As-nlp-2

5'- and 3'-RACE experiments using the *nlp2-5RACE*, *nlp2-5RACE2*, *nlp2-3RACE*, and *nlp2-3RACE2* primers (Table 1) produced a transcript that was 706 base pairs long and contained an open reading frame of 185 amino acids (Figure 1, accession no. MF346532). This transcript encodes a complete signal peptide as predicted by SignalP 4.0 Server, and 6 putative peptide products separated by dibasic cleavage sites. The predicted peptide sequences are likely cleaved from the precursor peptide C-terminally to dibasic KR, KK, or RR residues [7, 8], and these amino acids are later removed from the C-terminus by carboxypeptidase E [8]. In addition, each of the putative peptide products contains a C-terminal glycine, which is a precursor for amidation [9].

There were few differences between the cloned peptide-encoding sequence and sequences from the *Ascaris* EST library (accession no. GO33WK202FR3CJ, GJVAP8F02H4GRO) and genome (accession no. 1997694). The cloned sequence was identical to the EST sequence over the peptide-encoding region, and the genomic sequence differed by only a single nucleotide (T for G at 420 bp). Numerous differences were observed in the 5'- and 3'-UTRs. In the genome, the signal peptide and the putative peptide products are separated by an intron 465 bp in length, with predicted splice sites on either end (not shown).

In comparing the *nlp-2* homologs in 11 nematode species, the number of predicted peptides varied (Figure 2); however, all species had homologs of the first, fourth, fifth, and sixth peptides on the *A. suum* transcript. Since these species are thought to have shared a last common ancestor 550 million years ago [10], the sequence similarity is striking, and could imply a shared function. The third *A. suum* peptide does not have an -RPamide C-terminus, but instead a predicted -SMamide sequence that is conserved in clades III and IVb, but not in clade V. This raises the possibility of a divergent function, although this has not yet been explored.

(b) As-nlp-23

RACE experiments were performed using the *nlp23-5RACE* and *nlp23-3RACE* primers. Nested 5'-RACE experiments were also performed using the *nlp23-5RACE* product with the *nlp23-5RACE2* nested primer to confirm the presence of two methionines at the 5' end of the sequence. The resulting transcript was 491 bp in length, with an open reading frame of either 110 or 111 amino acids depending on which methionine is considered the "initiating" methionine (Figure 3, accession no. MF346533). SignalP 4.0 prediction software predicts a signal peptide for both the single ($p = 0.805$) and the double ($p = 0.809$) methionine sequences, with predicted cleavage between position 21 and 22 (CDA-NL). In contrast to *As-nlp-2* above, in this case the signal peptide cleavage gives rise to the N-terminus of the first predicted peptide (As-NLP-23.1) on the *As-nlp-23* transcript, which is followed by a dibasic cleavage site and the –RPamide-related peptide As-NLP-23.2. Both of these processed peptides end in a C-terminal glycine, which is a precursor for further processing by amidation [9].

The cloned *As-nlp-23* transcript is identical to the sequence in the EST library (accession no. GO56VPC01CTCHG) from the signal peptide through the last dibasic cleavage site, including both peptide products. However, there are many nucleotide differences at the 3' end of the EST sequence, which stops 11 amino acids after the second dibasic cleavage site. Our sequence extends for an additional 50 amino acids before reaching a stop codon. These differences could be explained by looking through the *A. suum* genome: the *As-nlp-23* transcript is split between two fragments, with one containing the signal peptide and the two putative peptide products (accession no. 2071214_5), and the other covering the rest of the open reading frame and the 3' UTR (accession no. 2071214_4). These sequences are identical to their respective sections of the cloned transcript, and contain predicted splice sites at the ends, supporting the idea that these fragments are spliced together to produce the *As-nlp-23* transcript. Therefore, it appears that the EST sequence came from the first fragment, and was not yet spliced to the second fragment, suggesting that there could be alternatively spliced transcripts. Finally, these genomic sequences are on separate scaffolds, which are separated by an intron that is at least 2,811 bp in length.

Homologous peptides were only observed in 3 other nematode species, all from the *Caenorhabditis* genus (Figure 4). Among these species, sequence conservation was observed for the second putative peptide (As-NLP-23.2) from the *As-nlp-23* transcript, particularly at the internal –GR– residues at positions 5-6 and the C-terminal –RPG. Some similarities to the first putative peptide product on the *As-nlp-23* transcript were observed, but this part of the transcript was more variable even among the *Caenorhabditis* genus, and did not contain a

predicted C-terminally amidated peptide, but instead a short peptide with the sequence LVI[N/S]R[G/Q/D]GFRP[A/V], which shares some homologies with As-NLP-23.1.

(c) *As-nlp-46*

RACE experiments were performed using the *nlp46-5RACE*, *nlp46-5RACE2*, *nlp46-3RACE*, and *nlp46-3RACE2*. Two versions of the transcript were cloned, which were identical over the open reading frame and differed in the length of the 3' UTR (Figure 5, accession no. MF346534). The open reading frame was 82 amino acids long, including a predicted signal peptide and a single putative peptide product flanked by dibasic cleavage sites. The peptide also contains a C-terminal glycine, which is a precursor for amidation [9].

There are several differences between the cloned products and the EST library and genomic sequences. The first 9 nucleotides in the EST sequence (accession no. As_L3-42a04.q1kSKPL), corresponding to position 14-22 bp in the cloned sequence, were different from the cloned sequence. The EST sequence does not include an initiating methionine, and since our cloned sequence was identical to the genomic sequence (see below), we attribute the differences to sequencing error. There is also one other single nucleotide substitution in the EST sequence; a G for an A at 79 bp, which translated into Val for Ile. Again, this was not shared by the genomic sequence. This transcript is split among two genomic fragments, one containing the signal peptide sequence (accession no. 2067242) and the other containing the putative peptide product (accession no. 2118465). These fragments were identical to their respective sections of the cloned sequence. Both genomic sequences are flanked by predicted splice sites,

supporting the idea that they are spliced together to produce the complete *As-nlp-46* transcript. These genomic scaffolds do not overlap, making the intron over 4,455 bp in length.

Homologous peptides were identified in 12 species of nematodes, and showed remarkable conservation (Figure 6), sharing the NIA[I/T]GR[G/A]DG[F/L]RPG motif. Given the long evolutionary divergence time (noted above, [10]), this degree of similarity implies a strong selective pressure on the peptide sequence itself.

Localization of -RPamide Transcripts by in situ Hybridization

The preparations used for *in situ* hybridization contained the anterior 2-4 cm of the worm, containing the nerve ring and head ganglia (Figure 7), and segment 1-2 motorneurons. In most cases, expression was limited to cells in the left and right lateral lines (LLL, RLL), the ventral ganglion (VG), and the retrovesicular ganglion (RVG), with the relative position of the cell given in reference to the nerve ring (NR), amphidial commissures (AC), or deirid commissures (DeC).

(a) As-nlp-2

Several pairs of cells were stained with the *As-nlp-2* probe in the head of the worm (Figure 8). In all preparations, a pair of small cells was stained in the lateral lines (9 preparations). These cells were located at the level of the amphidial commissures, and usually stained with moderate intensity. In 3 of 9 preps, a single cell was lightly stained in the ventral ganglion. In 3 of 9 preparations, a pair of cells was also lightly stained in the retrovesicular ganglion (RVG, not shown). No cells were stained in the ventral nerve cord. These results

contrast with data from GFP constructs in *C. elegans*, which showed fluorescence in one head neuron (unidentified), as well as secretory cells near the vulva, and the intestine [2].

The variable staining in the ventral ganglion and retrovesicular ganglion was sometimes difficult to detect given the high background staining in these preparations. It is possible that the *As-nlp-2* transcript is expressed at low levels in these cells, but since *in situ* hybridization is only semi-quantitative, we cannot determine whether the variability was due to true differences in low-level expression or experimental artifact. For example, cells can be damaged or lost during the dissection and washing steps in the protocol, or riboprobe penetration can be hindered in cells buried among the processes and hypodermal wrapping of the nerve cords. Additional studies using a more quantitative technique would be needed to resolve these questions.

(b) As-nlp-23

In all preparations, a single pair of cells was stained in the lateral lines between the amphidial and deirid commissures (10 preparations; Figure 9). These cells usually stained with moderate intensity. No staining was observed in the motor neurons of the ventral nerve cord, or in the other head ganglia. However, staining was observed in the hypodermis, with much greater intensity than the background staining in control preparations. In *A. suum*, the hypodermis is very closely associated with both neurons and muscle cells [11], so it is possible that the hypodermis releases As-NLP-23 onto motorneurons and muscle fields to modulate muscle contraction. The widespread, interconnected nature of the hypodermis could allow for long-range “tuning” of the excitability of the circuit- a concept known as context-setting.

Hypodermal expression was also found with GFP construct experiments of *nlp-23* in *C. elegans*, which showed fluorescence in the dorsal and ventral hypodermis and the tail, but none in neurons [2].

(c) *As-nlp-46*

Several pairs of cells were stained with the *As-nlp-46* probe (Figure 10). One pair of cells was stained with low to moderate intensity in the anterior ventral ganglion, near the midline (17 out of 28 preps; 60.7%). Another small pair of cells in the posterior ventral ganglion stained with slightly lower intensity (51.9% of preparations). In 6 out of 28 preparations, a medium sized cell was also stained with very low intensity (21.4%).

There were also 5 pairs of cells stained in the lateral lines between the nerve ring and the deirid commissures. The first was a small pair of cells, usually stained with moderate to high intensity, just posterior to the nerve ring (93.1% of preparations). In some preparations, a second pair of large round cells, called the “grapefruit cells”, was very lightly stained where the nerve ring meets the lateral lines (20% of preparations). These cells were easily obscured by bundled processes from the nerve cord and lateral lines. A third pair of cells was usually stained with moderate intensity just posterior to the first pair, near the level of the amphidial commissures (93.3% of preparations). These were followed by a larger pair of cells that stained with low to moderate intensity, located between the amphidial and deirid commissures (80% of preparations). Just anterior to the deirid commissures was a fifth pair of small cells, usually stained with moderate to high intensity (93.3% of preparations). No staining was observed in the ventral nerve cord. Unfortunately, more precise identification of the stained cells is difficult

because there is little or no staining of the neural processes in *in situ* hybridization. However, these results clearly contrast with tiling array studies (a type of microarray that focuses on particular regions of the genome, rather than genes that are dispersed throughout) in *C. elegans* [12], which reported that *nlp-46* is enriched in the ventral nerve cord and neurons, but not in the lateral lines or ventral ganglion.

Pharmacological Effects of -RPamide Peptides on A. suum Muscle Contraction

We recently identified a peptide, As-NLP-22 (SLASGRWGLRPamide), in the inhibitory motoneurons. This peptide practically abolished acetylcholine-induced muscle contraction in strips of dissected *A. suum* muscle ($IC_{50} = 8.3$ nM), and caused flaccid paralysis when injected into intact worms [1]. Given the strong sequence similarity between As-NLP-22 and the rest of the -RPamide peptides, we wanted to test whether these peptides had similar effects on muscle contraction.

(a) As-nlp-2

The second peptide on the *As-nlp-2* transcript, As-NLP-2.2 (SLALGRVDFRPamide), was tested in the dorsal muscle strip assay as previously described [1, 5]. At 10 μ M exposure to As-NLP-2.2, ACh-induced muscle contraction did not differ significantly from control for the first 20 min of the assay (Figure 11). However, at the 30-, 40-, and 50- minute time points, preparations exposed to As-NLP-2.2 showed an increased response to ACh. After 50 min, the response was 32 ± 6 % greater than that observed in control preparations. No differences in spontaneous contraction or baseline tension were observed.

The response observed for As-NLP-2.2 (SLALGRVDFRPamide) differed dramatically from the As-NLP-22 (SLASGRWGLRPamide) response (Figure 11), despite similar amino acid sequences. These peptides differ most in the 4th position (Leu vs. Ser) and 8th position (Asp vs. Gly), and in positions 7 and 9 where aromatic amino acids were substituted with non-polar aliphatic amino acids (Trp vs. Val and Leu vs. Phe). This could indicate that either the presence of a hydrophobic residue at position 4 or a negatively charged residue at position 8 are required for the potentiated response in the presence of As-NLP-2.2, while polar/uncharged or uncharged amino acids at these positions are crucial for the dramatic inhibition in the presence of As-NLP-22. These residues could be key for selective binding to different G protein-coupled receptors, with a different receptor responsible for each type of response.

(b) As-NLP-23.2

The second peptide on the *As-nlp-23* transcript, As-NLP-23.2 (SIALGRFSLRPamide), was tested in the dorsal muscle strip assay. At 10 μ M exposure to As-NLP-23.2, a biphasic effect was observed (Figure 12). In the presence of peptide, the elicited contraction was reduced to 50 ± 9 % of the baseline contractions, and 37 ± 9 % less than control preparations. The baseline tension also showed an average decrease of ≈ 1 g over the 10-min incubation time. However, once peptide was washed out of the chamber, the ACh-induced contractions recovered to control levels, and then increased in magnitude up to 39 ± 12 % of control preparations by the 50-min time point. These results differ from *Ascaris* body wall muscle experiments with *C. elegans* peptide NLP-23 (SMAIGRAGMRPamide), which only showed a slow relaxation of -5.1 ± 1.14 g in the presence of 10 μ M peptide [13].

The biphasic response of As-NLP-23.2 (SIALGRFSLRPamide) is particularly interesting, with the early inhibitory phase resembling As-NLP-22 (SLASGRWGLRPamide) and the late potentiated phase resembling As-NLP-2.2 (SLALGRVDFRPamide). As-NLP-23.2 contains Leu at position 4, like As-NLP-2.2, and Ser at position 8, which is more similar to the Gly in As-NLP-22 than the Asp in As-NLP-2.2. This suggests that presence of a hydrophobic amino acid at position 4 is required for the elevated late phase of the response, while a polar/uncharged or uncharged amino acid at position 8 is required for the early inhibitory phase and reduction in baseline muscle tension.

(c) As-NLP-46

The As-NLP-46 peptide (NIAIGRGDGFRPamide) had an excitatory effect on ACh-induced muscle contraction in the dorsal muscle strip assay (Figure 13). At 10 μ M exposure to As-NLP-46, preparations initially contracted at 23 ± 5 % above control preparations, and continued to increase in magnitude even after peptide was washed out. After 50 min, the response was 57 ± 8 % greater than controls.

As-NLP-46 differs the most in sequence from As-NLP-22 and the other -RPamide peptides; however, it contains a hydrophobic amino acid (Ile) at position 4, and a negatively charged amino acid (Asp) at position 8, both similar to As-NLP-2.2. The contractile response in muscle was also most similar to As-NLP-2.2, even though As-NLP-46 lacks the hydrophobic amino acids at positions 7 and 9 which are typical of the other -RPamide peptides, though it does resume the conserved “hydrophobic amino acid-RPamide” after the extra glycine at position 9. Furthermore, the amino acid differences between As-NLP-46 and As-NLP-2.2 (Asn

for Ser at position 1, Ile for Leu at positions 2 and 4, Gly for Val at position 7, and the additional Gly residue at position 9) resulted in an even greater potentiated response than that observed for As-NLP-2.2, although at present we do not have enough information to say which changes are important.

CONCLUSIONS AND FUTURE DIRECTIONS

The work described here identifies three additional transcripts containing *A. suum* - RPamide peptides and reports their localization for the first time. All of these transcripts were expressed in the head of the worm, primarily in the lateral lines, and no expression was observed in the ventral nerve cord, which contains the cell bodies of the motoneurons that control body locomotion. In some cases, staining of cells in the lateral lines was very dark and consistent. However, staining of other cells in the ventral ganglion, retrovesicular ganglion, and lateral lines was more variable. Since *in situ* hybridization is only semi-quantitative, we cannot determine whether these differences are due to low levels of expression in these cells, or were a result of experimental artifact. For example, poor penetration of the riboprobe in cells that are buried within the bundled processes of the lateral lines, or even damage or loss of cells during the dissection and washing process. Interestingly, however, we saw clear expression of one of these transcripts, *As-nlp-23*, in the hypodermis. Since the hypodermis is closely associated with neurons and muscle in *A. suum*, this finding could indicate that As-NLP-23 plays a role in modulating the relative excitability of the motor circuit, also known as context-setting.

In addition, all the peptides we tested (As-NLP-2.2, As-NLP-23.2, As-NLP-46) affected muscle contraction in *A. suum* dorsal muscle strip preparations. Since these preparations

contain motorneuron processes (but not cell bodies) in addition to the muscle and body wall, we cannot yet determine whether the effects on muscle are direct. However, it is possible that these peptides can act on muscle at a distance from the site of expression and release (e.g. are secreted into the pseudocoelomic fluid).

Further, by comparing the peptide sequences and pharmacological responses with that of As-NLP-22, we have identified two key amino acid positions that affect the activity of RPamide peptides. Peptides that contained hydrophobic amino acids (Leu or Ile) at position 4 produced a long-term potentiated response to ACh (as seen in As-NLP-2.2, As-NLP-23.2, and As-NLP-46). Polar/uncharged or uncharged amino acids (or at least a lack of negatively charged Asp) at position 8, in contrast, were critical for the initial inhibition of the ACh response and decrease in baseline tension (as seen in As-NLP-22 and As-NLP-23.2). To further investigate these ideas, a series of synthetic peptides with substitutions at the relevant positions would be needed.

One interpretation of these data is that the early inhibitory phase and the later potentiated phase are mediated by two different receptors with very similar ligand-binding domains, with ligand specificity determined by the amino acids at positions 4 and 8. This model is supported by cases where only the potentiated phase is present, because if all of these peptides were acting on the same receptor, we would expect to see at least some of the early inhibition with any peptide that bound. This is clearly not the case for As-NLP-46, which had significantly stronger contractile responses at all time points. However, these studies do not tell us much more about the mechanism of action and what type of cellular response could be

involved. Electrophysiological studies such as those described in Davis and Stretton (2001) would help to tease apart actions by different neuropeptide receptors.

Since these peptides are not expressed by the ventral cord motoneurons yet induce a strong pharmacological response in muscle tissue, it is possible that they are long-distance modulators of locomotion. To better understand the importance of these peptides *in vivo*, it would be helpful to examine their effects when injected into intact behaving worms [14-16]. The potential role as long-distance modulators would also be greatly influenced by peptide concentration, so dose-response in all of these experiments will be crucial for understanding biases in receptor interaction, signaling pathway, and general behavioral response.

AUTHOR CONTRIBUTIONS

J. J. Knickelbine was responsible for transcript identification, riboprobe preparation, bioinformatics, *in situ* hybridization, muscle contractile assays, and manuscript preparation. C. J. Konop assisted with dissections for *in situ* hybridization. A. O. W. Stretton was responsible for manuscript preparation.

TABLES AND FIGURES

Transcript	Primer Name	Sequence
<i>As-nlp-2</i>	nlp2-5RACE*	5'- CCG GGT CGG AAT CCA ACT CTG CCA AGC GC -3'
	nlp2-5RACE2	5'- CTA TGT CCT TAT CCG CAA TGT CGA CGT TCC -3'
	nlp2-3RACE*	5'- CGA TGG CTC TTG GAC GGC TTG CAT TTC GTC -3'
	nlp2-3RACE2	5'- GGA ACG TCG ACA TTG CGG ATA AGG ACA TAG -3'
<i>As-nlp-23</i>	nlp23-5RACE*	5'- GCC AGG ACG CAA CGA AAA CCT GCC GAG TGC G -3'
	nlp23-5RACE2	5'- CCA CTA GGT CGA AGA GAA GAC CTC CCG ACG -3'
	nlp23-3RACE*	5'- GAT GTG CAC TTG GAA ATG TGT GCT GCT TGT G -3'
<i>As-nlp-46</i>	nlp46-5RACE*	5'- ATT TAC CGG GCC TGA AAC CAT CAC CAC GTC -3'
	nlp46-5RACE2	5'- CCA TCA CCA CGT CCA ATT GCA ATG TTA CGC -3'
	nlp46-3RACE	5'- GTT TCT AAG GGG CGA TGA ATA TGC GGA CTT TA -3'
	nlp46-3RACE2*	5'- GGT GCT ATT AAC GAC AAC CTT GAT AGC GCT TG -3'

Table 1. Summary of primers used for PCR and RACE for the –RPamide family of peptides.
Primers marked with asterisks were used to create the riboprobes.

acatgggggaatgtgtattgcgctatagcgatgacgcagtaagttagcggcactctttcca
M H A K L A A L F P
ttacttgtcgtttccacgttatcattggccgaagaagaacgccgtgcgaaaagatcgata
L L V V S T L S L A E E E R R A K R S I
gtgatcggacgaggaatatatcggtccaacgaattgggaggcaatatccgagagcgaattt
V I G R G I Y R P T N W E A I S E S E F
cggcgaggctgcagaaaagatcgatgggtcttggacgggttgcatcttcgtccaggaaag
P A E L Q K R S M A L G R L A F R P G K
cgctctctagccctcggcagggtagacttccgtcctggtaaacggaacgtcgacattgcg
R S L A L G R V D F R P G K R N V D I A
gataaggacatagaatatctgatggaagattcgatgggcaaacgaagtgcagctttcggg
D K D I E Y L M E D S M G K R S A A F G
cgtttccactttaggccaggcaaaagatcatttagcgctaggccgcagcggattccgcctt
R F H F R P G K R S L A L G R S G F R P
ggaaaacgctcccttgcgcttggcagagttggattccgacccggaaagcgagctagtgtt
G K R S L A L G R V G F R P G K R A S G
gaagtgttgatcgatccatgttctatcgaaatcgattcgctttggggtgacatgtacgac
E V L I D P C S I E I D S L W G D M Y D
attgcacagttgtacgaggaaatctcgaagaattgtggaaaagtttgaagagaacgcac
I A Q L Y E E I S K N C G K V -
gctatgaacaacttcaatggatggagacatgtgaaataacttatatgctgaaaatgcatg
cagaattttgtttatgccatgcataaaattaagcaatccgcgttaataaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaa

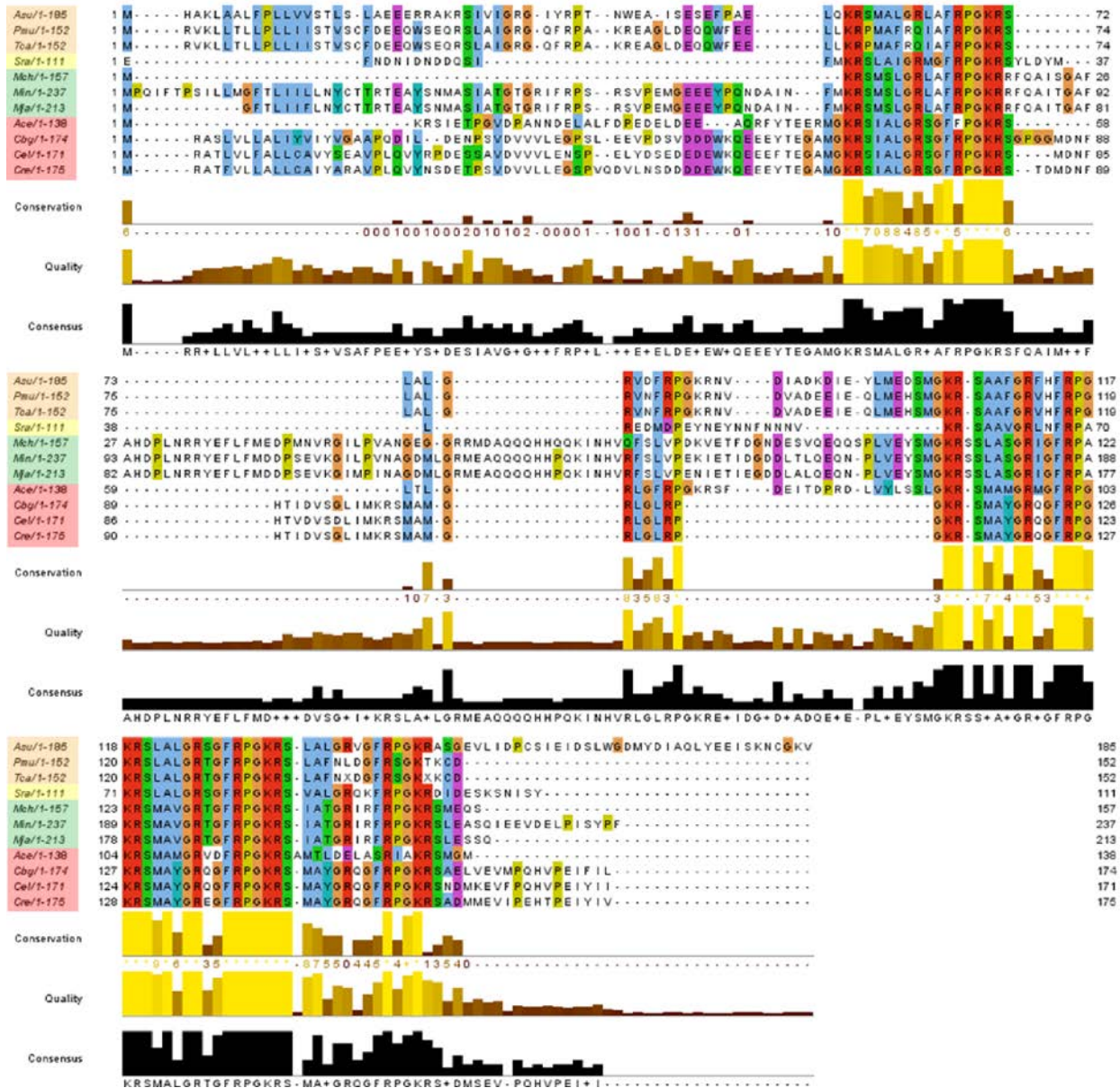


Figure 2. Comparison of *As-nlp-2* homologs in other nematodes. Nematode sequences are organized alphabetically by clade, shaded with orange for clade III, yellow for clade IVa, green for clade IVb, and red for clade V. Note blocks of high sequence similarity within the peptide-encoding region among 11 nematode species. Asu = *Ascaris suum*; Pmu = *Plectus murrayi* (FG621023); Tca = *Toxocara canis* (AA874711.1); Sra = *Strongyloides ratti* (LN609528.1); Mch = *Meloidogyne chitwoodi* (CD418949); Min = *Meloidogyne incognita* (JK281546); Mja = *Meloidogyne javanica* (CK427690); Ace = *Ancylostoma ceylanicum* (CB277108); Cbg = *Caenorhabditis briggsae* (XM_002643839); Cel = *Caenorhabditis elegans* (NM_076025); Cre = *Caenorhabditis remanei* (XM_003092743).

atgatgtgcacttggaatgtgtgctgcttgtgctaattggtctcagtgtgtcgtgac
 M M C T W K C V L L V L M V S V L S C D
 gctaacctcgccgtcgggaggtcttctcttcgacctagtggaaagcgctcaatcgactc
 A N L A V G R S S L R P S G K R S I A L
 ggcagggttttcgttgctcctggc
 G R F S L R P G K R S I E S T A Q S L Q
 acaatactcggggataagagtggatgggactgttcaacacaacaacttgatgatcagt
 T I L G D K S G W D C S T Q Q L A M I S
 gacgttatggaaggattggcctcgcttttggaggactacacctcctatctaggtcgatgc
 D V M E G L A S L L E D Y T S Y L G R C
 tcggcttttagcagatatcgctggcaacttccgatgatcgaatgcgaactgcagcaccatt
 S A L A D I A G N F R -
 aacaataagcctctccaaaatgcccttccaatacgcaggatgcattcggttggttgatg
 ataagattgtttcttaaggaagaggaaaagaaaattgaagaaaaaaaaaaaaaaaaaaaa
 aaaaaaaaaa

Figure 3. *As-nlp-23* transcript. The nucleotide sequence and deduced amino acid sequence of the *As-nlp-23* gene transcript (MF346533) obtained by RACE and PCR. The putative signal peptide is marked in blue. The peptide-encoding amino acid sequences are in red and underlined, with cleavage sites in bold italics. A dash signifies a stop codon. The target region for the *As-nlp-23* riboprobe is indicated by the bolded italicized nucleotides.

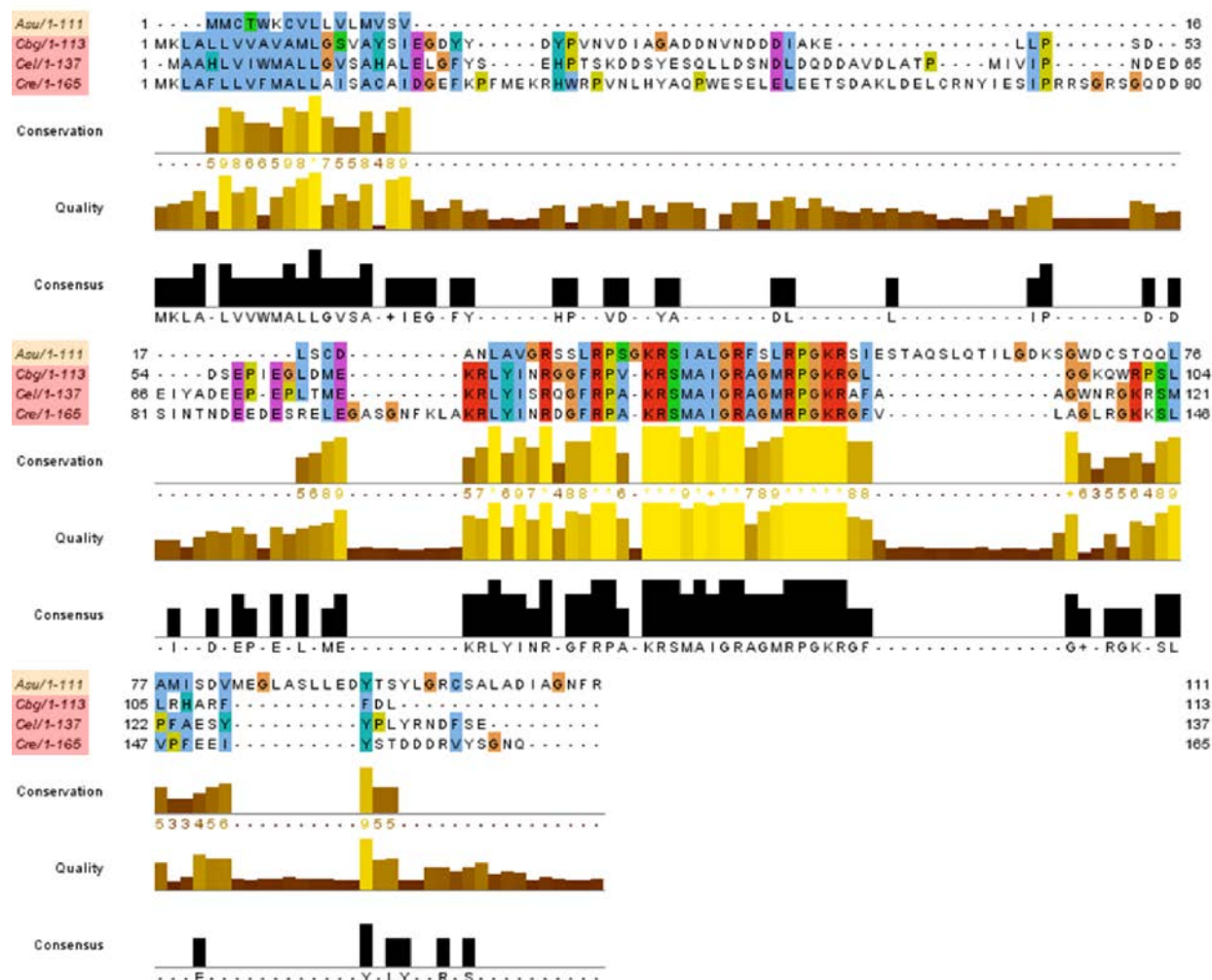


Figure 4. *As-nlp-23* homologs in other nematodes. Nematode species are organized alphabetically by clade, with orange shading for clade III and red for clade V. Note conservation of amino acid sequence in the peptide-encoding region of 4 nematode species. Asu = *Ascaris suum*; Cbg = *Caenorhabditis briggsae* (XM_002643840); Cel = *Caenorhabditis elegans* (NM_076024); Cre = *Caenorhabditis remanei* (XM_003092737).

(a) Short isoform of *As-nlp-46* transcript

gcggacgaaatgcaatacgggttattcgataaa**ggtgctattaacgacaaccttgatagcg**
 M Q Y G Y S I K V L L T T T L I A
cttggttatttcggtgccagttcatgatgcggattatagcagcaaggaccaatacagggca
 L V I S L P V H D A D Y S S K D Q Y R A
cttgaagaaacgcctattctttatcctgcaaacattttgtttctaaggggcgatgaatat
 L E E T P I L Y P A N I L F L R G D E Y
gcggacttttatttcgccacgagccggaagcgtaacattgcaattggacgtggtgatggt
 A D F I S P R A G K R N I A I G R G D G
ttcaggccccggtaaataaaaacgtacacgaaatgtggcaagaacggaagtcgggtggcact
F R P G K -
 ggaaaagaggaaacatccttgcaaaagcaataaagaaaataacacggcagtatttcatat
 atttgcattccgctctattatctaaaggaaaagtcatgaaatttaagaaaaaaaaaaaaa
 aaaaaaaaaaaaaaaaaa

(b) Long isoform of *As-nlp-46* transcript

gcggacgaaatgcaatacgggttattcgataaa**ggtgctattaacgacaaccttgatagcg**
 M Q Y G Y S I K V L L T T T L I A
cttggttatttcggtgccagttcatgatgcggattatagcagcaaggaccaatacagggca
 L V I S L P V H D A D Y S S K D Q Y R A
cttgaagaaacgcctattctttatcctgcaaacattttgtttctaaggggcgatgaatat
 L E E T P I L Y P A N I L F L R G D E Y
gcggacttttatttcgccacgagccggaagcgtaacattgcaattggacgtggtgatggt
 A D F I S P R A G K R N I A I G R G D G
ttcaggccccggtaaataaaacgtacacgaaatgtggcaagaacggaagtcgggtggcact
F R P G K -
 ggaaaagaggaaacatccttgcaaaagcaataaagaaaataacacggcactatttcatat
 atttgcacccgctctattatctaaaggaaaagtcatgaaattta**catacatataatgagc**
agaactcgatttgaaatgcagaaatgttcgattaatgaaactatttctcatgtgtttacggt
cctctttgaattaaaatcgatttcgaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

Figure 5. Two versions of the *As-nlp-46* transcript. The nucleotide sequence and deduced amino acid sequence of the (a) short and (b) long versions of the *As-nlp-46* gene transcript (MF346534), differing in the length of the 3'-UTR (additional nucleotides bolded and underlined in b). The putative signal peptide is marked in blue. The peptide-encoding amino acid sequences are in red and underlined, with cleavage sites in bold italics. A dash signifies a stop codon. The target region for the *As-nlp-46* riboprobe is indicated by the bolded italicized nucleotides.

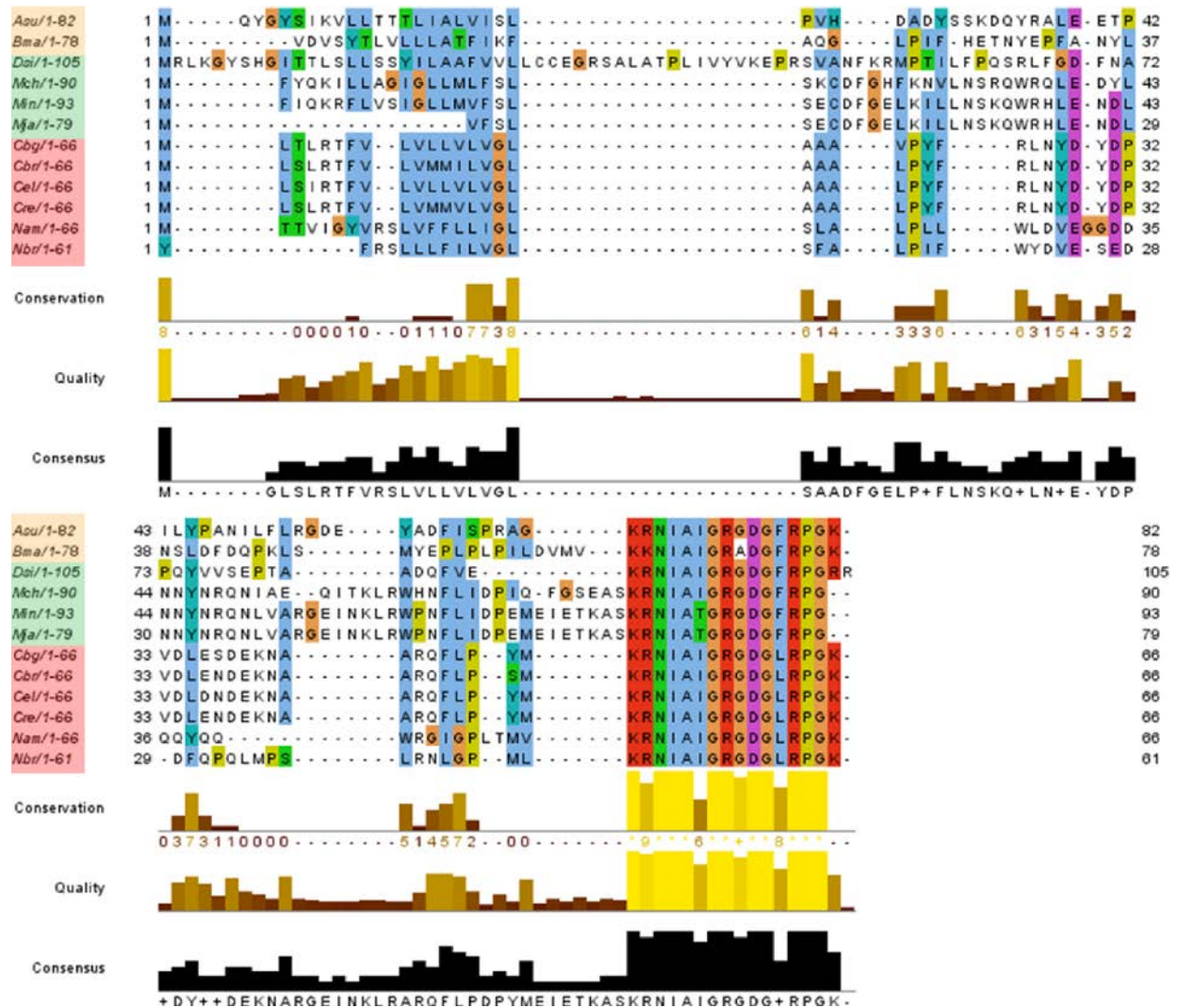


Figure 6. *As-nlp-46* homologs in other nematodes. Nematode sequences are organized alphabetically by clade, with orange shading for clade III, green for clade IVb, and red for clade V. Note high degree of conservation in the peptide-encoding region among 12 nematode species. *Asu* = *Ascaris suum*; *Bma* = *Brugia malayi* (XM_001901865); *Dsi* = *Deladenus siricidicola* (EZ435553); *Mch* = *Meloidogyne chitwoodi* (CB930664); *Min* = *Meloidogyne incognita* (JK298029); *Mja* = *Meloidogyne javanica* (CK427277); *Cbg* = *Caenorhabditis briggsae* (XM_002636732); *Cbr* = *Caenorhabditis brenneri* (FF086526); *Cel* = *Caenorhabditis elegans* (NM_001029005); *Cre* = *Caenorhabditis remanei* (XM_003114233); *Nam* = *Necator americanus* (BU087520); *Nbr* = *Nippostrongylus brasiliensis* (EH362371).

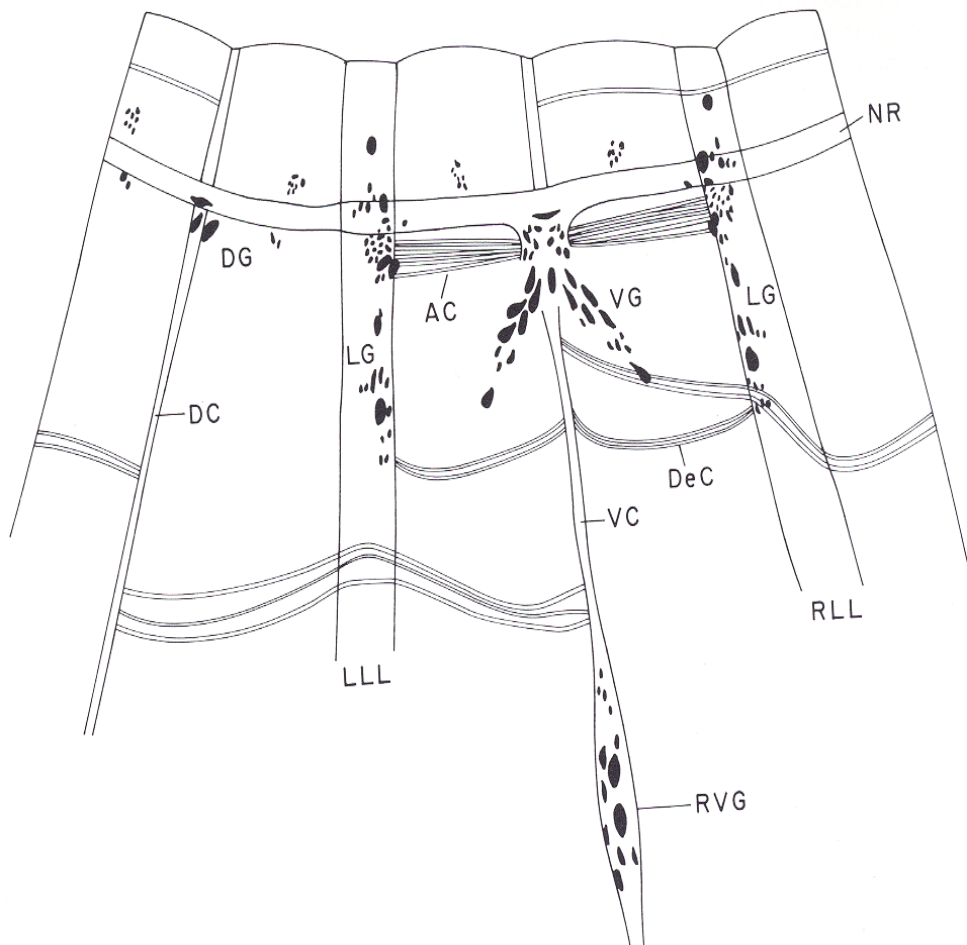


Figure 7. *Ascaris suum* nerve cord and head ganglia. AC = amphidial commissures, NR = nerve ring, DC = dorsal cord, RLL = right lateral line, DeC = deirid commissures, RVG = retrovesicular ganglia, DG = dorsal ganglia, VC = ventral cord, LG = lateral ganglia, VG = ventral ganglia, LLL = left lateral line. Figure adapted from Goldschmidt (1908).

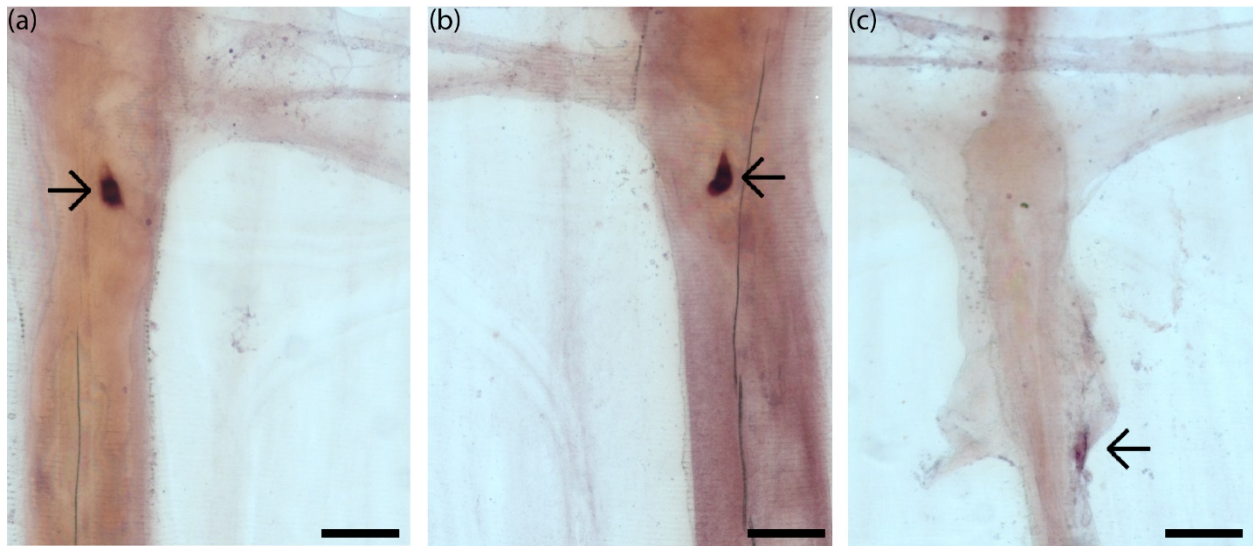


Figure 8. *As-nlp-2* localization by *in situ* hybridization. *Ascaris* head preparation stained with the *As-nlp-2* riboprobe. Two small cell bodies were stained in the (a) left- and (b) right- lateral lines, indicated by arrows. Rarely, a single cell was stained (very lightly) in the ventral ganglion (c), and a pair of cells in the RVG (not shown). No staining was observed in the ventral nerve cord. Scale bar = 100 μ m.

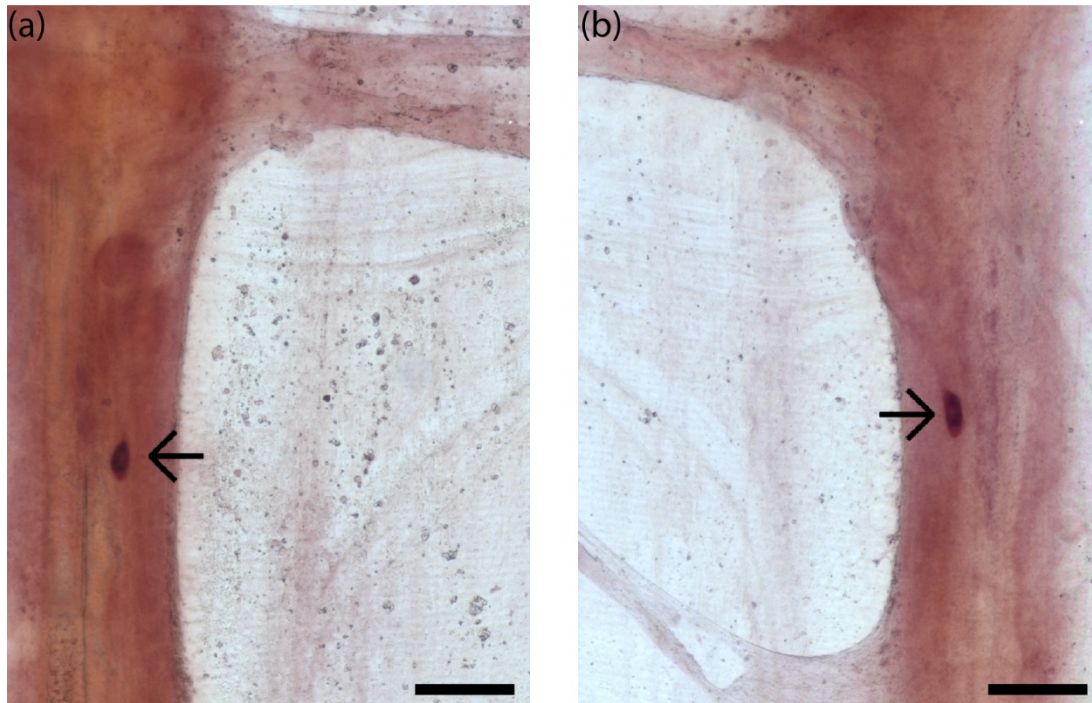


Figure 9. *As-nlp-23* localization by *in situ* hybridization. A single cell body was stained in the (a) left and (b) right lateral lines (indicated by arrows). No staining was observed in the ventral ganglion or the ventral nerve cord. Preparations stained with the *As-nlp-23* probe also showed much higher hypodermal staining than all other probes tested, even with shorter incubation time in visualization buffer, and streaky staining of the hypodermis. Scale bar = 100 μ m.

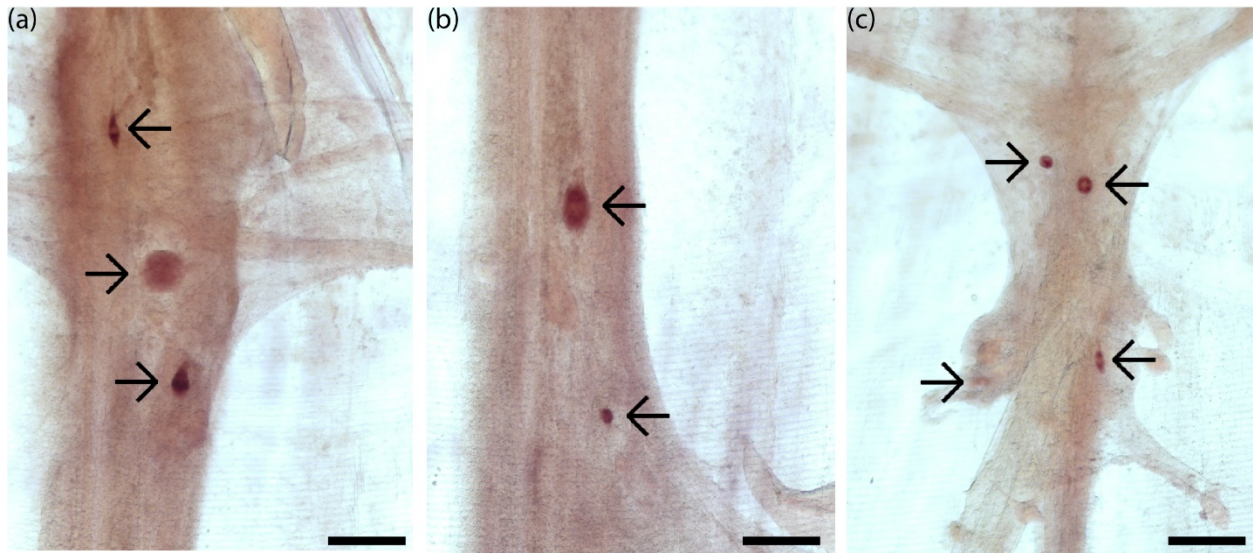


Figure 10. *As-nlp-46* localization by *in situ* hybridization. Five pairs of cells stained in the left and right lateral lines. (a) The first three cells in the left lateral line. (b) The fourth and fifth cells in the left lateral line. The first, third, and fifth pairs were stained with moderate to high intensity, and the second, larger cell body (grapefruit cell) was stained with a lower intensity. (c) Two pairs of cells stained in the ventral ganglion. Scale bar = 100 μm.

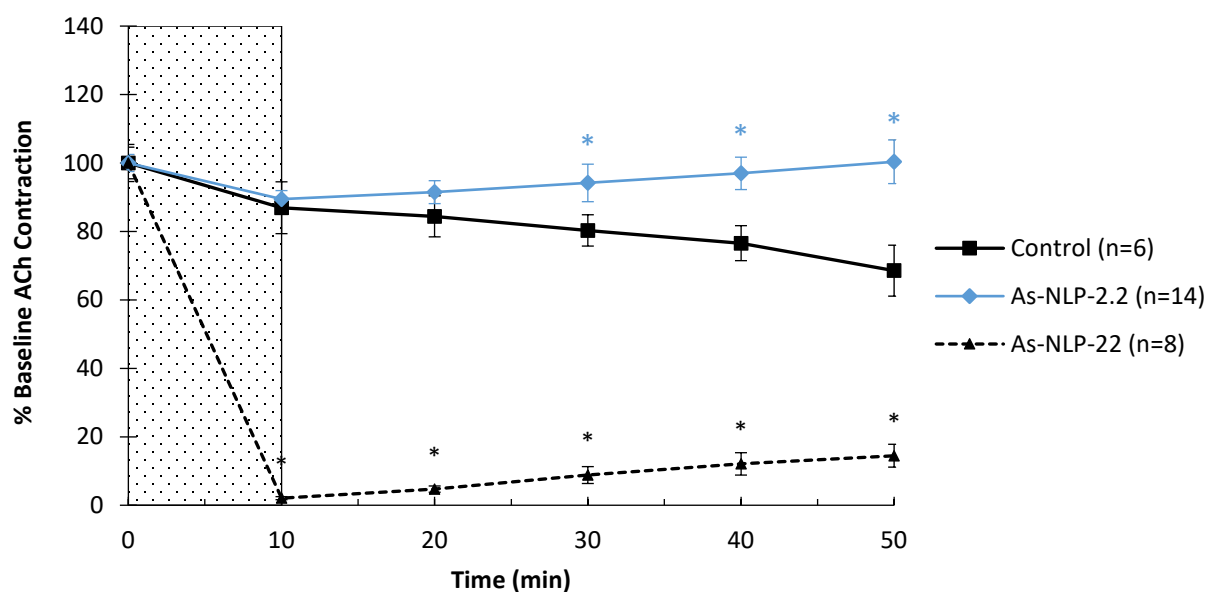


Figure 11. Effect of As-NLP-2.2 on ACh-induced muscle contraction. Baseline contractions in response to 5 μ M ACh are reported at 0 min, and subsequent responses are reported as a percentage of the baseline ACh-induced contraction. Shaded region indicates incubation with 10 μ M As-NLP-2.2 (SLALGRVDLFRPamide) or As-NLP-22 (SLASGRWGLRPamide) peptide solution, which was washed out after the 10-min time point. Error bars represent the standard error of the mean. Data points marked with an asterisk are significantly different from control preparations ($P < 0.05$).

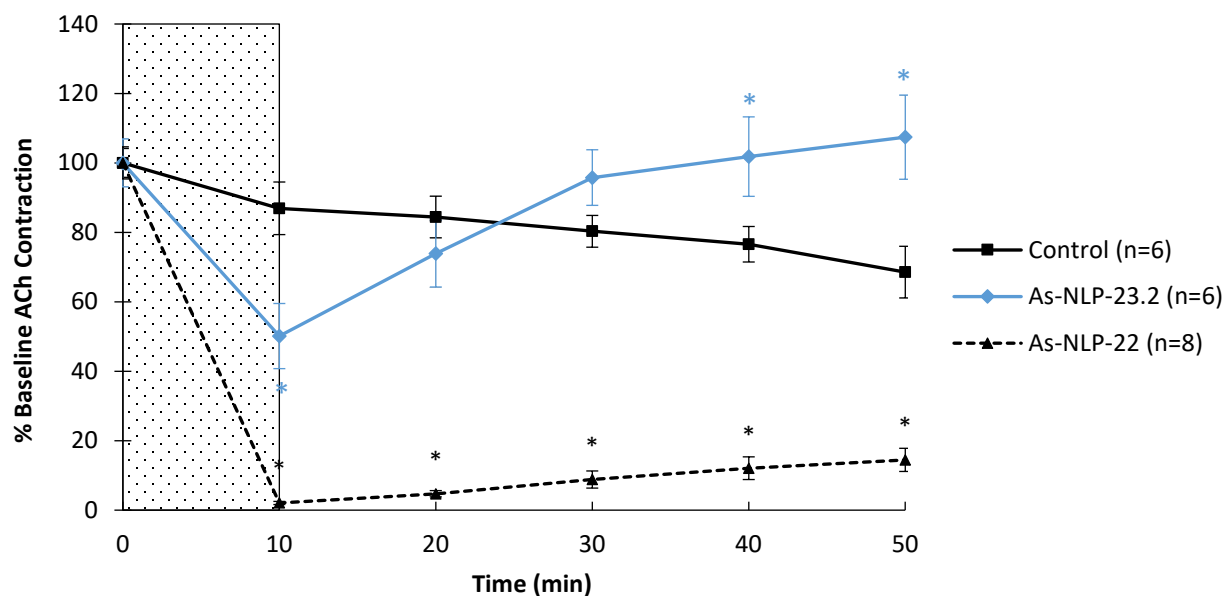


Figure 12. Effect of As-NLP-23.2 on ACh-induced muscle contraction. Baseline contractions in response to 5 μ M ACh are reported at 0 min, and subsequent responses are reported as a percentage of the baseline ACh-induced contraction. Shaded region indicates incubation with 10 μ M As-NLP-23.2 (SIALGRFSLRPamide) or As-NLP-22 (SLASGRWGGLRPamide) peptide solution, which was washed out after the 10-min time point. Error bars represent the standard error of the mean. Data points marked with an asterisk are significantly different from controls ($P < 0.05$).

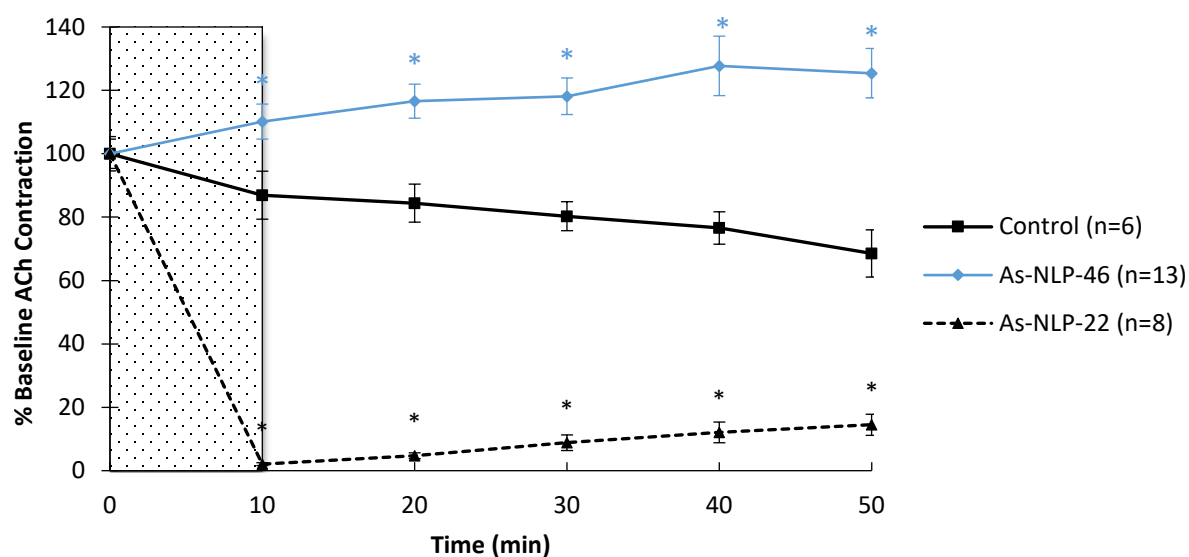


Figure 13. Effect of As-NLP-46 on ACh-induced muscle contraction. Baseline contractions in response to 5 μ M ACh are reported at 0 min, and subsequent responses are reported as a percentage of the baseline ACh-induced contraction. Shaded region indicates incubation with 10 μ M As-NLP-46 (NIAIGRGDGFRPamide) or As-NLP-22 (SLASGRWGLRPamide) peptide solution, which was washed out after the 10-min time point. Error bars represent the standard error of the mean. Data points marked with an asterisk are significantly different from controls ($P < 0.05$).

REFERENCES

1. Konop, C. J., Knickelbine, J. J., Sygulla, M. S., Wruck, C. D., Vestling, M. M., Stretton, A. O. (2015). Mass spectrometry of single GABAergic somatic motoneurons identifies a novel inhibitory peptide, As-NLP-22, in the nematode *Ascaris suum*. *J Am Soc Mass Spectrom.* 26:2009-23.
2. Nathoo, A. N., Moeller, R. A., Westlund, B. A., Hart, A. C. (2001). Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. *Proc Natl Acad Sci U S A.* 98:14000-5.
3. Nanda, J. C., Stretton, A. O. (2010). In situ hybridization of neuropeptide-encoding transcripts *afp-1*, *afp-3*, and *afp-4* in neurons of the nematode *Ascaris suum*. *J Comp Neurol.* 518:896-910.
4. Jarecki, J. L., Andersen, K., Konop, C. J., Knickelbine, J. J., Vestling, M. M., Stretton, A. O. (2010). Mapping neuropeptide expression by mass spectrometry in single dissected identified neurons from the dorsal ganglion of the nematode *Ascaris suum*. *ACS Chem Neurosci.* 1:505-19.
5. Jarecki, J. L., Viola, I. R., Andersen, K. M., Miller, A. H., Ramaker, M. A., Vestling, M. M., Stretton, A. O. (2013). Three independent techniques localize expression of transcript *afp-11* and its bioactive peptide products to the paired AVK neurons in *Ascaris suum* : *In situ* hybridization, immunocytochemistry, and single cell mass spectrometry. *ACS Chem Neurosci.* 4:418-34.
6. Jarecki, J. L., Frey, B. L., Smith, L. M., Stretton, A. O. (2011). Discovery of neuropeptides in the nematode *Ascaris suum* by database mining and tandem mass spectrometry. *J Proteome Res.* 10:3098-106.
7. Thacker, C., Rose, A. M. (2000). A look at the *Caenorhabditis elegans* Kex2/Subtilisin-like proprotein convertase family. *Bioessays.* 22:545-53.
8. Jacob, T. C., Kaplan, J. M. (2003). The EGL-21 carboxypeptidase E facilitates acetylcholine release at *Caenorhabditis elegans* neuromuscular junctions. *J Neurosci.* 23:2122-30.
9. Kreil, G. (1984). Occurrence, detection, and biosynthesis of carboxy-terminal amides. *Methods Enzymol.* 106:218-23.

10. Holterman, M., van der Wurff, A., van den Elsen, S., van Megen, H., Bongers, T., Holovachov, O., Bakker, J., Helder, J. (2006). Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown Clades. *Mol Biol Evol.* 23:1792-800.
11. Stretton, A. O. W., Fishpool, R. M., Southgate, E., Donmoyer, J. E., Walrond, J. P., Moses, J. E. R., Kass, I. S. (1978). Structure and physiological activity of the motoneurons of the nematode *Ascaris*. *Proc Natl Acad Sci USA.* 75:3493-97.
12. Spencer, W. C., Zeller, G., Watson, J. D., Henz, S. R., Watkins, K. L., McWhirter, R. D., Petersen, S., Sreedharan, V. T., Widmer, C., Jo, J., Reinke, V., Petrella, L., Strome, S., Von Stetina, S. E., Katz, M., Shaham, S., Ratsch, G., Miller, D. M., 3rd (2011). A spatial and temporal map of *C. elegans* gene expression. *Genome Res.* 21:325-41.
13. McVeigh, P., Alexander-Bowman, S., Veal, E., Mousley, A., Marks, N. J., Maule, A. G. (2008). Neuropeptide-like protein diversity in phylum Nematoda. *Int J Parasitol.* 38:1493-503.
14. Davis, R. E., Stretton, A. O. W. (2001). Structure-activity relationships of 18 endogenous neuropeptides on the motornervous system of the nematode *Ascaris suum*. *Peptides.* 22:7-23.
15. Reinitz, C. A., Herfel, H. G., Messinger, L. A., Stretton, A. O. (2000). Changes in locomotory behavior and cAMP produced in *Ascaris suum* by neuropeptides from *Ascaris suum* or *Caenorhabditis elegans*. *Mol Biochem Parasitol.* 111:185-97.
16. Reinitz, C. A., Pleva, A. E., Stretton, A. O. (2011). Changes in cyclic nucleotides, locomotory behavior, and body length produced by novel endogenous neuropeptides in the parasitic nematode *Ascaris suum*. *Mol Biochem Parasitol.* 180:27-34.
17. Goldschmidt, R. (1908). Das nervensystem von *Ascaris lumbricoides* und *megalocephala*. *Zeitchrift fur Wissenschaftliche Zoologie.* 90:73-136.

CHAPTER FOUR

Different Neuropeptides Are Expressed in Different Functional Subsets of Cholinergic

Excitatory Motorneurons in the Nematode *Ascaris suum*

Christopher J. Konop, Jennifer J. Knickelbine, Molly S. Sygulla, Martha S. Vestling, and

Antony O. W. Stretton

This work is published in: ACS Chemical Neuroscience (2015), Volume 6, Pages 855-870.

Doi: 10.1021/cn5003623

INTRODUCTION

Nematodes are famous for being the first organisms in which the “complete wiring diagram” of the nervous system was determined. This was reported nearly 30 years ago in the free-living nematode *Caenorhabditis elegans* [1]. More recently, such a description has been labeled a “connectome” [2-4]. Yet, to understand the neural circuits that control behavior, it is not sufficient merely to describe the anatomical connections of the wiring diagram; even when the basic electrophysiological properties of the neurons and their synapses are added to the morphological description, this “functional wiring diagram” does not adequately describe the circuit that generates behavior. It is a “first-pass” description of the system, necessary, of course, but not sufficient. Importantly, it lacks the incorporation of the functional role of neuromodulators into the description of circuitry. The importance of modulators has perhaps been best described in the stomatogastric nervous system of crustaceans [5]; both biogenic amines and neuropeptides play crucial roles in configuring neural circuits. Simple circuits take

on additional layers of complexity, but the full extent of the added complexity is yet to be described.

In the nematode *Ascaris suum*, the inadequacy of the simple connectome approach was demonstrated about 25 years ago [6, 7]. At first this was a surprise, because the *A. suum* nervous system, like that of other nematodes, is remarkably simple, both numerically (the entire nervous system comprises only 298 neurons [8]) and morphologically (almost all neurons have at most two branch points). It seemed like the simplest of the so-called simple nervous systems, and therefore probably to be solved in rapid order. However, the first-pass description is clearly incomplete [9]. It was soon recognized that at least part of what is missing is a description of the activity of neuromodulators. These compounds can affect the activity of the neurons and their synapses, often in intricate ways [10], so it is essential that a complete description of their structure, cellular localization, and action be added to the basic functional circuit. Neuropeptides are among the most numerous and varied of the known neuromodulators. As in *C. elegans* [11], it is estimated that there are at least 250 endogenous peptides in *A. suum* [12, 13]; we have chemically isolated and sequenced ca. 100 so far, and the majority of those tested to date affect the activity of the neurons and/or muscle in the locomotory system, or locomotory behavior itself [10, 14-17]. In retrospect, it is no wonder that a functional circuit description that lacked the role of neuropeptides was a failure.

Initially, our approach to the discovery of endogenous peptides was by chemical fractionation and direct sequence determination by Edman degradation [14, 18, 19]. More recently, we have used mass spectrometry (MS) to detect and then to sequence neuropeptides [12, 13, 17, 20-22]. We have now extended this technique to the single cell level; mass

spectrometric methods are sensitive enough both to detect and to sequence peptides from single dissected neurons [17, 20]. This approach is limited only by how successfully we can dissect single neurons. For some neurons, it is relatively easy, but for others, such as the ventral cord motoneurons, the close association between the neurons and hypodermis, in which most of the neurons are embedded, and the many other neural processes in the nerve cord, complicate the dissection.

In this paper, we report our success in dissecting single motoneurons and in cataloguing the neuropeptidome of the different somatic motoneurons. Although the current matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) instruments are easily capable of detecting levels of peptides in single cells, careful isolation and preparation of the neurons is necessary to make sure that none of the sample is lost. Glycerol has long been used as a cryoprotectant [23, 24], but more recently, glycerol treatment has been shown to help improve the cells' stability during mechanical manipulation without affecting peptide structure or ion suppression during single cell analysis by MS [25-27]. We found that the use of an isotonic dissecting solution that contains 30% glycerol was highly successful. Consistent with other work, we find that glycerol treatment confers several benefits: (i) it prevents damage to the cell membrane during dissection, (ii) it loosens adhering cells from each other, probably by transient shrinkage, (iii) it helps to reduce contamination from damaged surrounding tissue, and (iv) it prevents dehydration of the cell once it is exposed to the atmosphere, allowing for better extraction of the analyte(s) when exposed to matrix on the MS target plate.

Analysis of single cells also has important advantages over the analysis of more complex tissue, such as ganglia or the entire nerve cord, since the complexity of the sample is reduced.

This reduces the effects of ion suppression, thus increasing the probability of isolating a desired peak for sequencing by tandem MS (MS/MS). The improvements in dissection techniques have now enabled the isolation of ventral cord motoneurons, and an enumeration and sequencing of the peptides that each contains.

There are several other important advantages of single cell MS. First and foremost, it identifies the actual peptide product(s) within each cell, rather than relying on predictions from translated nucleic acid sequences; such predictions are unreliable as the cellular expression patterns of the processing enzymes, especially the proteases, controlling the post-translational modifications of the precursor protein are not yet understood. Since our ultimate goal is to carry out physiological experiments with the peptides, it is crucial to find the fully processed peptides that are actually present in each cell; physiological experiments are very time-consuming, so it is pointless to investigate the bioactivity of a non-endogenous peptide. Second, it identifies the cell in which the peptide is expressed, i.e. one half of the neuronal signaling dyad mediated by the peptide. Third, it enables comparison of expression in identified cells from individual animals; we have previously noted significant individual-to-individual variation in the expression of peptides, especially in those peptides expressed at low levels [12, 17, 20].

In this paper, we describe the peptides present in the excitatory ventral nerve cord motoneurons, and we report an interesting difference in the peptide content of different classes of these motoneurons. They can be divided into two groups, the first of which (DE1, DE3, and VE1) contains a set of shared peptides in neurons where the neuromuscular output region of the neuron is anterior to the cell body; these neurons mediate anteriorly-propagating

body waveforms [28] and forward locomotion (Stretton, unpublished). In the other group (DE2 and VE2), in which the output is posterior to the cell body (mediating posteriorly-propagating waveforms and backward locomotion in *A. suum*), there is a different set of peptides. *C. elegans* has motorneurons that are anatomically identical to those of *A. suum* (Figure 1 b, c). In *C. elegans*, the relationship between the direction of waveform propagation and the morphology of the motorneurons controlling each direction has been established by laser ablation experiments in which either the forward-projecting or the backward-projecting neurons were deleted [29]. In both *A. suum* and *C. elegans* the anteriorly-projecting and posteriorly-projecting motorneurons mediate body waveforms that propagate anteriorly or posteriorly, respectively. However, the relationship between the direction of propagation of the body waveform and the direction of locomotion is different in the two worms: in *C. elegans* anteriorly-propagating waves drive the worm backwards and posteriorly-propagating waves drive it forwards [30], whereas in *A. suum* it is anteriorly-propagating waves that drive it forwards [31]. These differences between *A. suum* and *C. elegans* are due to differences in the interaction between the body movements and the physical environment – moving on a surface versus moving in a tube. However, it is important to reiterate that the motorneurons driving the anteriorly- and the posteriorly- propagating waves are homologs in the two species.

The cellular expression patterns observed by single-cell MS have been confirmed by *in situ* hybridization (ISH), using riboprobes targeting the transcripts that encode the different peptides. In all cases, the localization by ISH was identical to that seen with single-cell MS. This is an important result, since it validates the dissection technique: the major peptides seen by MS are not due to contamination from other neurons, and in particular, are not due to

contamination from input synaptic terminals. Taken together, these experiments identify new components of the functional circuit that describes behavior.

MATERIALS AND METHODS

Animals

Live *A. suum* were collected from pig intestines at a slaughterhouse, and maintained in phosphate buffered saline (PBS, 140 mM sodium chloride, 10 mM sodium phosphate, pH 6.8-7.5) at 37°C. The PBS was changed daily and worms were used within 3 days of collection.

Sample Preparation for Mass Spectrometry

Adult female *A. suum* were injected with 0.1-0.3 mL of 2 mg/mL collagenase (Sigma Blend H) in *Ascaris* saline (4 mM sodium chloride, 125 mM sodium acetate, 24.5 mM potassium chloride, 5.9 mM calcium chloride, 4.9 mM magnesium chloride, 5 mM 3-(N-morpholino) propanesulfonic acid (MOPS) buffer, pH 6.8) and incubated for 1.5- 2 hours at 37°C to dissociate the muscle tissue. A 6-7 cm portion of the worm anterior to the gonopore was removed and transferred to a Sylgard-lined dish. After cutting longitudinally between the dorsal nerve cord and the right lateral line, the preparation was pinned open, the lips were cut off, and the pharynx was removed. The preparations were rinsed 3x with 170 mM ammonium acetate before the isotonic glycerol dissecting solution was added (170 mM ammonium acetate in 30% glycerol) [21, 26, 32-34]. Dissecting needles made from electrolytically sharpened tungsten wire were used to isolate a desired neuron from the many neuronal processes and hypodermis in which it is embedded. For better visualization, some preparations were bathed

in 0.8 mM Methylene Blue in 170 mM ammonium acetate for 20-30 seconds, then cleared with 170 mM ammonium acetate and dissected as described above.

Isolated cells were transferred to a Bruker MTP stainless steel MALDI target. Cells were cleared with 0.5 μ L isopropanol [35] to remove the glycerol and any remaining extracellular salts. Individual neurons were spotted with 60 - 100 nL of saturated α -cyano-4-hydroxycinnamic acid (CHCA; Sigma, St Louis, MO) in 70% acetonitrile (ACN; Fisher Scientific, Waltham, MA), 1% Trifluoroacetic acid (TFA; Sigma) or saturated 2,5-dihydroxybenzoic acid (DHB; Sigma) in 1:1 HPLC grade methanol and water (Fisher Scientific) using a Nanoliter Cool Wave Syringe II [36].

Mass Acquisition

A Bruker Ultraflex III MALDI-TOF/TOF MS (Bruker Daltonics, Billerica, MA) equipped with a Smartbeam laser and LIFT-TM cell was used to obtain MS and MS/MS spectra with Bruker Compass v. 1.2 software. Because the cells contain relatively small quantities of peptide, all spectra were obtained from 50 laser shots per acquisition. MS spectra were obtained in positive ion reflectron mode, with an m/z range of 500-4000. Instrument settings were optimized for maximum detection sensitivity using the Bruker Flex Control software: ion source 1 voltage, 25.0 kV; ion source 2 voltage, 20.5 kV; reflector 1 voltage, 26.5 kV; reflector 2 voltage, 14.5 kV; lens voltage, 9.5 kV. Synthetic peptide standards were not suitable for calibration of the instrument because the laser intensity needed for analysis of single cells was often far too high when analyzing pure standard, making it impossible to identify the monoisotopic peak of the ion needed for accurate calibration. Therefore, a reference spectrum from a freshly dissected

A. suum AVK neuron, with known peptide peaks previously described by Jarecki et al. (2012), was used for external calibration of the instrument.

Acetylation of Peptides

Following deposition of the cell onto the target surface, 0.5 μ L of isopropanol was applied to each cell to wash away the glycerol solution, followed by 0.5 μ L of methanol/ acetic anhydride (3:1). The cells were allowed to dry before each wash. Cells were then covered with matrix as described above.

Oxidation of Peptides

In cells exposed to Methylene Blue, methionine residues were partially oxidized to the sulfoxide, with a mass shift of +16 Da, and tryptophan residues gave rise to +16 and +32 adducts [20].

Assignment of Peaks and Interpretation of Mass Spectra

Spectra were analyzed using Bruker Daltonics flex- Analysis 3.0 software. The software automatically assigned masses to peaks in each MS spectrum. Each MS/MS spectrum underwent background subtraction and smoothing before the automatic assignment of masses. In some cases, peak m/z values were added manually. Peaks were considered significant if they were twice the intensity of the baseline noise.

Molecular masses and ion fragmentation patterns were calculated by Protein Prospector MS-Product (<http://prospector.ucsf.edu>). Spectra containing peaks with m/z values

corresponding to masses ± 0.2 Da of *A. suum* peptides were temporarily assigned. Confirmation of the assignments was carried out by MS/MS and chemical modifications to the peptides. *De novo* sequencing of unidentified peaks was carried out by hand. Verification of peptide sequence was carried out by comparing the experimental MS/MS spectrum to that of synthetic peptide. Further verification was carried out by comparing sequences to tBLASTn searches and cloning of *A. suum* gene sequences. Raw spectral image files were acquired with Bruker Daltonics flex- Analysis 3.0 software and transferred to Adobe Illustrator (San Jose, CA) for annotation.

Database Searches

Database searches were conducted using methods described in recent publications [13, 17, 20]. Briefly, all predicted peptide sequence assignments were searched using tBLASTn (National Center for Biotechnology Information [NCBI], Bethesda, MD; <http://www.ncbi.nlm.nih.gov/BLAST/>) against *A. suum* ESTs and a library of 447,546 genomic survey sequences (GSS). For all searches, program settings were modified for searching short sequences using a word size of 2, an E value of 20,000, and a PAM30 matrix. Search results were translated using the ExPASy Translate tool (<http://au.expasy.org/tools/dna.html>) and examined for putative peptide cleavage sites, and C-terminal glycine, a precursor for C-terminal amidation of peptides. Genomic sequences were analyzed for mRNA splice sites directly by PCR cloning of the mRNA sequences and comparison with genomic sequences [37, 38]. Signal peptides were predicted using SignalP 4.0 (<http://www.cbs.dtu.dk/services/SignalP/>).

Multiple Alignment

BLAST searches of the nematode EST library were performed using both *As*-NLP-21 and *Ce*-NLP-21 as queries. Selected sequences were imported into MEGA5.1 [39] for alignment by MUSCLE [40]. The alignments were imported into Jalview [41] for display.

Peptide Synthesis

Peptides were synthesized by the University of Wisconsin Biotechnology Center (UWBC). The integrity of the synthesis was monitored by MALDI-TOF MS and high pressure liquid chromatography (HPLC). Two of the *As-nlp-21* peptides were >95% pure (*As*-NLP-21.4, *As*-NLP-21.5). The remaining peptides were <95% pure by HPLC (Appendix 2, Supplementary Table 1).

RNA Isolation and cDNA Preparation

Twenty-thirty *A. suum* heads were flash-frozen in liquid nitrogen and ground to a fine powder. Total RNA was isolated using a Nucleospin Nucleic Acid Purification Kit (Clontech, Mountain View, CA). First strand cDNA was generated for PCR using a Superscript First Strand Synthesis System for RT-PCR (Life Technologies, Grand Island, NY). Rapid Amplification of cDNA Ends (RACE)-ready cDNA for 5'- and 3'-RACE was created from total RNA using a SMARTer RACE cDNA amplification kit (Clontech) according to the manufacturer's instructions.

Transcript Identification by PCR and RACE

To obtain information about the transcripts that encoded each peptide, primers were designed for PCR and RACE from sequences in the *A. suum* EST library. RACE reactions were performed using an Advantage 2 PCR Kit (Clontech) according to the manufacturer's instructions with the appropriate primers in a 50 μ L reaction as follows: 2 μ L of *A. suum* cDNA or 2 μ L of cDNA reaction solution with no reverse transcriptase as a control, 5 μ L of 10X PCR Gold buffer (Applied Biosystems, Foster City, CA), 2-8 μ L of 25 mM $MgCl_2$, 1 μ L each of 10 mM dNTPs, 1 μ L of each primer (20 μ M), and 0.25 μ L of AmpliTaq Gold DNA polymerase (Applied Biosystems; 5 U/ μ L). The PCR conditions were programmed into the Eppendorf Mastercycler Gradient as follows: 95°C for 3 min, followed by 40 cycles of 94°C for 30 seconds, 66°C or 68°C for 30 seconds, 72°C for 1 min.

PCR and RACE products were run on a 1% agarose gel, and novel bands were excised and purified using a Qiaquick Gel Purification Kit (Qiagen, Chatsworth, CA). The purified products were cloned into *E. coli* using a TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA), and the plasmid DNA was isolated with a Qiagen Miniprep Kit. Automated sequencing was carried out by the DNA Sequencing Facility at the UWBC. Sequence electropherograms were viewed on Chromas Lite software (Technelysium Pty Ltd). Sequences were analyzed using the ExPASy Translate Tool and T-Coffee Multiple Sequence Alignment Tool (Swiss Institute of Bioinformatics), and signal peptide sequences were identified using the SignalP 4.0 Server (Technical University of Denmark).

Riboprobe Synthesis

The primer pairs listed in Appendix 2, Supplementary Table 2 were used to amplify a selected region by PCR. Products were cloned and sequenced to confirm the fidelity of the sequence, and to determine the orientation of the insert in the vector. The target sequences of each riboprobe are shown in Figures 5, 8. The constructs were linearized using restriction enzymes NotI and SpeI (New England Biolabs, Beverly, MA), and purified (Gel Extraction kit, Qiagen). Linearized plasmids were used as a template to synthesize an antisense (experimental) and a sense (negative control) digoxigenin-labeled riboprobe (Maxiscript SP6/T7 kit, Ambion, Austin, TX; digoxigenin-11-dUTP, Roche Applied Science, Indianapolis, IN) as described in Nanda and Stretton (2010). To remove unincorporated nucleotides, the reactions were run through NucAway Spin Columns (Ambion). Probe integrity and concentration were determined by gel electrophoresis with Sybr Gold staining (Molecular Probes, Eugene, OR) and nylon membrane dot blots (Roche Applied Science protocol).

Localization of afp-1, -4, -14, and As-nlp-21 by in situ Hybridization

A. suum whole mount preparations were prepared by injecting large female worms with 0.1 - 0.3 mL of 2 mg/mL collagenase in autoclaved *Ascaris* saline and incubating in a beaker of PBS at 37°C for 1.5 – 2 hrs. The anterior 6 cm of the worm was taken and rinsed in autoclaved 100 mM phosphate buffer (pH 7.4) and the heads were transferred to a Sylgard-coated dissecting dish. The heads were cut longitudinally between the dorsal nerve cord and the right lateral line, the lips were removed, and the heads were pinned flat. Any remaining muscle cells were gently removed with forceps. The preparations were rinsed several times in 100 mM

phosphate buffer, and fixed in 1% paraformaldehyde in 100 mM phosphate buffer (pH 7.4) on a rocking table at 4°C overnight.

After fixation, the head preparations were processed as previously described [42], with the following modifications: incubation in proteinase K solution was increased to 30 min at 37°C, and staining was allowed to occur for up to 16 hours in the dark, or until staining of the cells was evident, and background staining began to occur. The staining reaction was stopped with Milli-Q water, and the heads were rinsed with Milli-Q water (2 x 10 min). The preparations were mounted in Clear-Mount (Electron Microscopy Sciences, Hatfield, PA) and allowed to dry. Photomicrographs were taken with a Zeiss AxioCam MRc camera on a Zeiss Universal microscope.

Characterization of As-NLP-21 Bioactivity

Muscle tension assays were performed using standard protocols [43, 44]. Strips of dorsal muscle were obtained by cutting large female *A. suum* longitudinally along the lateral lines and isolating a 2 cm strip approximately 1 cm posterior to the head. Each preparation contained dorsal muscle cells (all longitudinal) and the motor axons of the dorsal nerve cord. The ends of the preparation were tied with silk thread. One end was attached to a fixed hook in a 7 mL chamber containing normal *Ascaris* saline 37°C and stirred by bubbling nitrogen gas, and the other end was tied to a FORT25 force transducer (World Precision Instruments, Sarasota, FL). The output of the transducer was routed through a TBM4M transbridge (WPI) and recorded on a computer using Data-Trax or LabScribe2 software (WPI), and the data were analyzed for tension/time relationships.

Baseline contractions were measured by adding 5 μ M ACh to the chamber by micropipette and waiting until the contraction reached its maximum tension, followed by rinses of *Ascaris* saline. The preparation was then exposed to 10 μ M peptide solution in *Ascaris* saline for 10 min (in the case of As-NLP-21.1 and As-NLP-21.2, the peptide was first dissolved in 10% DMSO, then diluted into *Ascaris* saline, resulting in a final concentration of 0.001% DMSO in the chamber). The chamber was rinsed, fresh peptide was added, and ACh-induced contraction was measured by adding 5 μ M ACh immediately after exposure to peptide solution, followed by rinses with peptide-free *Ascaris* saline. Muscle strip contraction in response to ACh was measured at 10, 20, 30, 40, and 50 min after exposure to peptide, with rinses of peptide-free *Ascaris* saline between doses of ACh. For comparison, control muscle strips were subjected to ACh-induced contraction in peptide-free *Ascaris* saline (or 0.001% DMSO in *Ascaris* saline) at the same time points. Responses of the muscle strips were reported as a percentage of the maximum ACh contraction observed for each muscle strip.

Responses for each peptide group were reported as the mean $\pm$ S.E., and displayed using Prism 6 software (GraphPad, San Diego, CA). Significance between peptide and control groups was determined using unpaired *t*-tests, with significance levels set as $P < 0.05$.

RESULTS AND DISCUSSION

It is important to know the chemical nature of the intercellular signaling molecules that are used in nervous system circuits. Besides the classical neurotransmitters, there are many neuromodulators that affect the properties of neurons and their synapses in important and varied ways, and knowledge of these effects is an essential part of understanding circuit

properties. The first step is to identify these molecules, and this is a problem for chemists. The majority of known modulators are peptides, and the recent advances in mass spectrometry, particularly the exquisite sensitivity and accuracy of MS techniques, have opened up new possibilities for the rapid exploration of the neuropeptide complement of nervous systems. We are using the simple nervous system of the nematode *Ascaris suum*, which has a total of only 298 neurons, as a model system. We aim to understand the neural basis of the control of locomotory behavior. It is clear from our previous work that the role of modulators in neural circuitry must be added to our description of the system. Despite extensive studies on the motor nervous system, previously we were not able to define the peptides present in individual motoneurons. Recent improvements in dissection techniques have enabled us to isolate single motoneurons and to both isolate and to sequence their complement of peptides.

Anatomical Background

In the ventral and dorsal nerve cords, there are 7 classes of somatic motor neurons that occur in 5 basic repeating units ("segments"), each comprising 11 motoneurons (4 types are present in two copies per repeat, and the other 3 types in one copy per segment) (Figure 1a) [8]. Some innervate dorsal muscle and others innervate ventral muscle. The cell bodies of all 7 classes are found in the ventral nerve cord, and the neurons that innervate dorsal muscle send projections to the dorsal nerve cord via semi-circumferential commissures. The different classes were initially established on the basis of their distinct cellular morphology (Figure 1b). Subsequent electrophysiological experiments distinguished sets of excitatory and inhibitory motoneurons: in each segment, dorsal muscle is innervated by four excitors (two DE1, one

DE2, and one DE3), and one inhibitor (DI), and ventral muscle is innervated by 2 copies each of the excitors (VE1, VE2) and inhibitor (VI). DE1, DE3, and VE1 neurons have their neuromuscular output regions anterior to the cell body, and DE2 and VE2 neurons have their neuromuscular output posterior to the cell body (Figure 1b). The excitors are cholinergic [45, 46], and the inhibitors are GABAergic [47, 48].

MS of Excitatory Motor Neurons in the Ventral Nerve Cord

Individual dissected motoneuron cell bodies from the ventral cord were subjected to MALDI-MS analysis. In addition to familiar peaks with mass-to-charge ratios (m/z) matching those of peptides encoded by two previously described *A. suum* genes, *afp-1* [49] and *afp-4* [42], and the previously identified peptide AF9 [18], there were several novel peptides that had not been identified previously (Figure 2, 3; Table 1). The intensity of the peaks varied; the most intense peaks were consistent for a particular cell type, whereas minor peaks varied from sample to sample of the same cell type.

From the 8 excitatory cells in each segment, two spectral patterns emerged. In spectra from posteriorly-projecting VE2 and DE2, the previously identified peptide AF2 (m/z 991.5), commonly thought to be the most abundant peptide expressed in *A. suum*, is often the most intense peak, and was consistently observed. VE2 and DE2 also contained peaks corresponding to the six peptides encoded by *afp-1*: AF3, m/z 857.5; AF20, m/z 875.5; AF14, m/z 905.5; AF4, m/z 958.5; AF13, m/z 1020.5; and AF10, m/z 1541.7 (Figure 2; Table 1) [12]. In most cases, these peptides are present at moderate to high intensity. The only exception was AF10, which was absent in 5 out of 10 cells in DE2. Interestingly, spectra from VE2 and DE2 were very similar

to spectra from the RID neuron in the dorsal ganglion (DG) [20]; AF10 also was absent in about 30% of the spectra from RID neurons.

In the anteriorly-projecting DE1, DE3, and VE1 motoneurons, an ion with m/z corresponding to AF9 (GLGPRPLRFa, 1011.7 Da) was observed. These neurons also contained large peaks from a group of novel peptides with m/z of 1018.5, 1271.7, 1322.7, 1327.6, 1516.8, and 1666.8 (Figure 3; Table 1).

Identities of certain functional residues in peptides were initially confirmed by chemical modification (Figures 2- 3). Exposure to methylene blue, which oxidizes methionine residues to their sulfoxides, causes a +16 Da mass shift for each oxidized methionine, while on-target acetylation produces a mass shift of +42 Da for acetylation of each N-terminus, as well as an additional +42 Da for each lysine and a fraction of tyrosine residues. In both oxidation and acetylation, modifications are incomplete, which produces spectra that contain both modified and unmodified peaks, allowing for easy comparison between the two. In all cases, peaks expected from the putative sequences were observed in treated cells, confirming the presence of methionine and the number of amino groups in the peptide (Figures 2-3).

Sequence Determination

(a) Sequences of Known Peptides

MS/MS was performed on the peak with m/z 991.5 (AF2), on a peak presumed to be from the *afp-1* transcript (958.5 Da – AF4), and on a peak with m/z 1011.7 (AF9). *De novo* sequencing confirmed the predicted amino acid sequence of each peptide, although MS/MS is not able to resolve the (I/L) ambiguities because the two amino acids are isobaric (Figure 4).

(b) Sequences of As-nlp-21 Encoded Peptides

Sequence determination of the 6 novel peptides from DE1, DE3, and VE1 motoneurons depended on a combination of tandem MS and analysis of sequences from gene transcripts. MS/MS analysis of the peptide with m/z 1271.7 produced the sequence R[I/L]YDAFSRFP (Figure 5a). To resolve the isoleucine/leucine ambiguity, we performed a tBLASTn search of the *A. suum* genomic survey sequence (GSS) database using both RIYDAFSRFP and RLYDAFSRFP as queries. In the translated sequence (GenBank accession no. E194631.1), we identified an N-terminal dibasic putative cleavage site and a C-terminal stop codon, which would predict the processed peptide RIYDAFSRFP (As-NLP-21.6) with a predicted m/z of 1271.7. Further analysis of the translated GSS revealed a second putative peptide sequence GGGRSFRSLSLGE (As-NLP-21.4), which has a m/z of 1322.7, matching that seen in spectra from VE1, DE1, and DE3 neurons. The As-NLP-21.4 sequence corresponds to a peptide first identified by homology from a search of the *A. suum* GSS and EST databases using known *C. elegans* peptides, and then confirmed by LCMS and MS/MS from a peptide extract of *A. suum* heads [13].

MS/MS of synthetic As-NLP-21.6 showed a highly similar fragmentation pattern to that of the endogenous peptide, serving as further confirmation of the peptide sequence (Appendix 2, Supplementary Figure 2a). Despite several attempts, As-NLP-21.4 could not be sequenced by MS/MS.

The *As-nlp-21* gene was cloned by 5' and 3' rapid amplification of cDNA ends (RACE) using the target sequence (5'-CCCAAGCTGCTCAAACGAAACGATCGTCC-3') from the GSS to design primers (Appendix 2, Supplementary Table 2). The product yielded two overlapping

sequences that were fused to produce the full *As-nlp-21* transcript (GenBank accession no. KP311680; Figure 5b), encoding a predicted protein with an initiating methionine and predicted N-terminal signal peptide, 6 putative peptide products (As-NLP-21.1 through As-NLP-21.6), named for the order in which they appear in the transcript, and a C-terminal stop codon. None of these peptides is C-terminally amidated. MS/MS of the peak with m/z 1516.8 confirmed the identity of peptide As-NLP-21.5 (Figure 6a). Besides As-NLP-21.4, As-NLP-21.5, and As-NLP-21.6, the full-length transcript encodes three additional predicted non-amidated peptides (As-NLP-21.1, GGGRPFTLVD; As-NLP-21.2, AGGRFFKIEDVDALE; and As-NLP-21.3, GGGRPFYTQTDE). All have predicted m/z that correspond to intense peaks consistently seen in VE1, DE1, and DE3 motoneurons by MS (Figure 3). MS/MS spectra partially confirmed the peptide sequences (Figure 6). For peptides As-NLP-21.1 and -21.3, a characteristic peak at m/z 328.1 (b_4 ion) was observed in tandem MS spectra (Figure 6; Appendix 2, Supplementary Figure 2). In MALDI, as well as electrospray ionization (ESI), fragmentation of peptides with an internal proline often preferentially cleaves N-terminal to that residue [50]. We also observe $b_4 - NH_3$ (-17 m/z), which we have found to be a typical neutral loss for a fragment ion with an arginine (R) residue.

Localization of the As-nlp-21 Transcript by in situ Hybridization

To confirm the localization of the *As-nlp-21* peptides identified by MS, we synthesized an RNA probe targeting a 431 bp section of the peptide-encoding transcript (Figure 5b). This probe labeled the DE1, DE3, and VE1 neurons in the ventral nerve cord with moderate to high intensity (Figure 7a-c; 10/10 preparations), supporting the results obtained by single-cell MS. No staining was observed in DE2 or VE2 (Figure 7d-e), or negative control preparations using a

sense probe (Appendix 2, Supplementary Figure 3a-c). The antisense probe also labeled many cell bodies in the head of the worm, including 2 pairs of cells associated with the nerve ring, 7 pairs of cells and one unpaired cell in the lateral ganglia, 7 pairs of cells in the ventral ganglion, and 2 pairs of cells in the retrovesicular ganglion (Appendix 2, Supplementary Figure 3d). An additional 4 pairs of cells were sometimes stained in the lateral lines, but staining was weak and highly variable.

*Identification of *afp-1*, *afp-4*, and *afp-14* Peptide-Encoding Gene Transcripts and Localization by *in situ* Hybridization*

*(a) *afp-14**

Several previous attempts to clone the transcript that encodes AF9 (*afp-14*) using primers designed to SL1 (a 5' splice leader sequence found in ca. 80% of *A. suum* transcripts [51]) and a gene-specific primer were unsuccessful. However, by using primers for 5' and 3' RACE, designed from AF9 GSS information (GenBank accession no. KP311679), we successfully amplified overlapping 3' and 5' sequences of the *afp-14* transcript. The deduced sequence contained an N-terminal signal peptide with an initiating methionine, the AF9 peptide-encoding region with C-terminal glycine (typically the precursor for C-terminal amidation), immediately followed by a stop codon, and an N-terminal dibasic cleavage site (Figure 8). The full-length transcript was confirmed by PCR. The genomic structure, showing introns and exons, is shown in Figure 8b. To synthesize an antisense riboprobe complementary to a 266-nucleotide portion of the *afp-14* transcript for ISH, we used specific primers AF9FWDPROBE1 (5'-GCGTACCGTTTGTACGACCT-3') and AF9REVPROBE2 (5'-CATTAGCGAGAAGCCGAGAG-3'). We

observed medium to intense staining in DE1 and DE3 motorneurons (Figure 9), but no staining in DE2, VE2, or the inhibitory motorneurons DI and VI (data not shown). VE1 neurons were also stained, but this staining is highly variable and always weak (Figure 9c). This could be due to lower expression levels in this cell type or insufficient penetration of the probe. VE1 is a relatively small cell and buried deep in the cord, which might make access of the probe more difficult. This variability is consistent with the MS of VE1 cells. In 3 out of 11 cells, we saw moderate to intense peaks for AF9 (m/z 1011.7), in 4/11 cells the AF9 peak was of low intensity, and in 4/11 cells the AF9 peak was not detected. No staining was observed in DE2 or VE2 (Figure 9d, e), or negative control preparations using a sense probe (Appendix 2, Supplementary Figure 4a-c). ISH of *afp-14* using the antisense probe also detected two other cells in the head. These are a bilateral pair believed to be either the ADA, FLP, or RMG neurons because of their position in the lateral lines; all of these send projections into the deirid commissure bundle which is observable in these preparations (Appendix 2, Supplementary Figure 4d).

(b) afp-1 and afp-4

Previous experiments in this laboratory investigated the localization of *afp-1* and *afp-4* peptide-encoding transcripts in the head of *A. suum* by ISH [42]; however, these preparations did not include the segment 2 motorneurons analyzed by MS. To corroborate our single cell MS data by ISH, we made *afp-1* and *afp-4* riboprobes that targeted the same region of the transcripts used in those experiments [42], and confirmed the expression of both peptide-encoding mRNAs in VE2 and DE2 motorneurons (Figure 10). We saw no expression in any other

ventral cord motorneurons. Negative controls with the *afp-1* and *afp-4* sense probes showed no staining (Appendix 2, Supplementary Figure 5).

Comparison with Other Nematodes

Nematodes are the most prevalent phylum on the planet and inhabit a wide variety of ecological niches, yet their nervous systems are highly conserved. The chemical messenger systems used to control behaviors like locomotion are also highly conserved; this includes the structure of neuropeptides expressed by individual nematode species. Previous work suggests that cellular expression of orthologous peptides is in most cases quite different between *A. suum* and *C. elegans* [17, 20, 42, 52, 53]. Our current investigation finds some instances of overlap in cellular expression of peptide-encoding genes in the ventral cord motorneurons. We suspect that, in some cases, there are true differences in expression of homologous genes between species, but at this point it is too early to be completely certain about the comparisons made between the two species because the techniques used to identify expression of peptide-encoding genes in the two species are different. In *A. suum*, the techniques we used, namely single-cell MS, ISH and immunocytochemistry directly detect the peptides or the transcripts that encode them, whereas in *C. elegans*, GFP constructs are used to detect peptide-encoding genes [11]. The interpretation of the results from using these constructs relies on adequate knowledge of the *cis*-regulatory domains that interact with the numerous transcription factors and regulators to control expression of different genes. The use of SAGE (Serial Analysis of Gene Expression, which detects mRNA) on distinct populations of *C. elegans* neurons identified significant differences from expression data obtained from GFP reporter strains [54], calling

into question, in each particular case, the completeness of the control sequences used with GFP constructs. With a better understanding of the regulatory mechanisms that control gene expression, reporter constructs are becoming more reliable in *C. elegans* [54-56].

We show that the posteriorly-projecting VE2 and DE2 motor neurons express *afp-1* (AF3, 4, 10, 13, 14, 20) and AF2. *Afp-1* is orthologous to the *C. elegans* gene *flp-18*, while AF2 (encoded by *afp-4*) is orthologous to the *C. elegans* gene *flp-14*. Both genes are among the most broadly conserved genes in nematodes [57-59], with the sequence of AF2 being completely conserved, and peptides encoded by *afp-1* orthologs maintaining a conserved PNFLRFamide C-terminus. GFP reporter constructs for *Ce-flp-14* did not identify any cells in *C. elegans*, while ISH experiments in *G. pallida* using a *Gp-flp-14* specific probe did not identify any of the ventral cord motorneurons, but instead the RME motorneurons that innervate muscles of the head [60].

ISH for *flp-18* failed in *C. elegans*, but *flp-18*:GFP-constructs in *C. elegans* do show expression in excitatory ventral cord motorneurons (Table 2) [61]. *Flp-18* has been shown to work in a homeostatic manner to regulate excitation/inhibition balances in locomotion via two G-protein coupled receptors, NPR-1 and NPR-5 [62]. Both AF9 and the *flp-18* peptides were identified as ligands for NPR-1 [63-65].

AF9 is also highly conserved and found broadly in nematodes; we show that it is expressed in cholinergic motorneurons VE1, DE1, and DE3 in *A. suum*; however, the *C. elegans* ortholog *Ce-flp-21* is reported to be expressed in ADL, ASE, and ASH sensory neurons, the URA motorneurons, the MC, M2, and M4 pharyngeal neurons [11], but not in any of the motorneurons in the ventral nerve cord.

The neuropeptide-like protein (NLP) gene *As-nlp-21* is expressed in VE1, DE1, and DE3 motorneurons. Searches of nematode ESTs identified orthologous *nlp-21* peptides in 11 species from 3 different clades (Appendix 2, Supplementary Figure 6) [66, 67]. The alignment of peptide sequences predicted from nematode *nlp-21* precursor proteins suggests that they share a common motif: [G/A][G/A][A/G]RXF[X₀₋₁₉] (Figure 11, Table 2). Predicted nematode *nlp-21* peptides have a conserved N-terminal sequence and a C-terminus that is highly variable, both in sequence and length. Currently, NLP localization data are restricted to *A. suum* and *C. elegans*. Expression of *nlp-21* is observed in both nematodes in VA, AS, and DA (VE1, DE1, and DE3 in *A. suum*), but in *C. elegans*, *nlp-21*:GFP is also reported in VB and DB [67] (VE2 and DE2) whereas in *A. suum* there is no expression of the orthologous transcript in these neurons. Both nematodes also show a number of unidentified cells in the head, and *Ce-nlp-21*:GFP was also expressed in the pharynx, intestine, and tail.

Bioactivity

To date, over 100 peptides have been isolated and sequenced from *A. suum* [10, 12, 13, 17-20, 22]. Each new peptide needs to be studied for its biological function. The functions of AF2, AF9 and the 6 peptides encoded on the *afp-1* transcript (AF3, 4, 10, 13, 14, 20) have been studied extensively by electrophysiological recording from individual motorneurons [10]. These peptides have been shown to cause a variety of effects in both the motorneurons and/or the muscle they innervate.

We show that AF9 is expressed in DE1, VE1, and DE3. From our ISH experiments, there is no evidence to suggest expression of AF9 in any other ventral cord neurons (i.e. interneurons).

An interesting observation is that AF9 causes a strong depolarization of DE2 and a weak but significant hyperpolarization of DI. This may explain the dorsal coiling posture shown by Davis and Stretton (2001). A few synapses have been shown between DE1 and DE2. Possibly these mediate the effect of AF9 on DE2, although it is also possible that the interaction is non-synaptic, involving diffusion of the peptide from its presumptive release site from DE1, DE3, and/or VE1 to receptors on DE2.

AF2 elicits a strong depolarization of DE2 motoneurons, and a biphasic change in input resistance [10]. In muscle tension experiments, AF2 produces a biphasic response, with an initial relaxation followed by contraction and rhythmic activity [19]. The explanation for these effects is not yet clear. Perhaps there is an autoreceptor on DE2 axons, although it is not certain that the electrophysiological effects of AF2 on DE2 neurons are direct. When AF2 was injected into whole worms, it caused a cessation of body waveform propagation; this was accompanied by a more than 100-fold increase in cAMP in muscle [15]. AF2 has also been shown to potentiate the response of nicotinic ACh receptors in *A. suum* [68, 69].

The *afp-1* peptides produced small depolarizations in DE2 motoneurons, and had little effect on DI motoneurons, yet when injected into whole worms, these peptides caused an increase in the total number of body waveforms, and a dramatic effect on their amplitude [10]. We are still seeking a better understanding of these results.

Preliminary experiments to characterize the bioactivity of the 6 novel peptides encoded by the *As-nlp-21* gene were carried out using the dorsal muscle strip assay. No significant effect by any of the peptides was observed on acetylcholine (ACh)-induced muscle contraction when compared to controls (Appendix 2, Supplementary Figure 7). The *C. elegans* ortholog, Ce-NLP-

21, has been shown to be trafficked in dense core vesicles in the cholinergic ventral cord motorneurons, and secreted into the extracellular fluid, but no postsynaptic target or primary function has been identified [70].

CONCLUSIONS AND FUTURE DIRECTIONS

By single-cell MS, we have shown that there are 2 functional classes of *A. suum* excitatory somatic motorneurons, which contain distinct subsets of neuropeptides. The anteriorly-projecting excitatory motorneurons (DE1, DE3, and VE1) express 6 novel peptides encoded by the *As-nlp-21* transcript, and AF9, encoded by the *afp-14* transcript. The posteriorly-projecting excitatory motorneurons (DE2 and VE2) express AF2 (encoded by *afp-4*) and the six related peptides encoded by the *afp-1* transcript. Cellular localization was confirmed by *in situ* hybridization. The biological activity of most of these previously identified peptides has been reported [10, 19]. Here we report preliminary experiments on the biological activity of the novel peptides we discovered. All of the biological effects described in this paper involve responses to single peptides. Now that we know that the excitatory motorneurons contain multiple peptides, it will be important in the future to examine possible interactions between them, using a combination of behavioral, muscle contractile, and electrophysiological assays. Initially, the interest will be in pairwise interactions, but we realize that there may be differences in time course of the effects of different peptides. It may be that some peptides set a background context against which other peptides (or neurotransmitters) act. There is overlap between the output regions of different motorneurons, i.e. the muscle cells at a particular position in the worm receive input from all the motor axons in that region. Furthermore, the

muscle cells in the dorsal or ventral fields are interconnected with electrical synapses. Thus, any context setting might be produced not only by interaction between the peptides in each cell, but also by interactions between the two distinct sets of excitatory neurons (the anterior-projectors and the posterior-projectors) that we have identified by their function in this paper.

AUTHOR CONTRIBUTIONS

C. J. Konop was responsible for single cell dissection, MS, bioinformatics, data analysis, transcript identification, and manuscript preparation. J. J. Knickelbine was responsible for transcript identification, ISH, muscle tension assays, bioinformatics, and manuscript preparation. M. S. Sygulla was responsible for single cell dissections. M. M. Vestling was responsible for MS and data analysis, A. O. W. Stretton was responsible for data analysis, bioinformatics, and manuscript preparation.

FIGURES

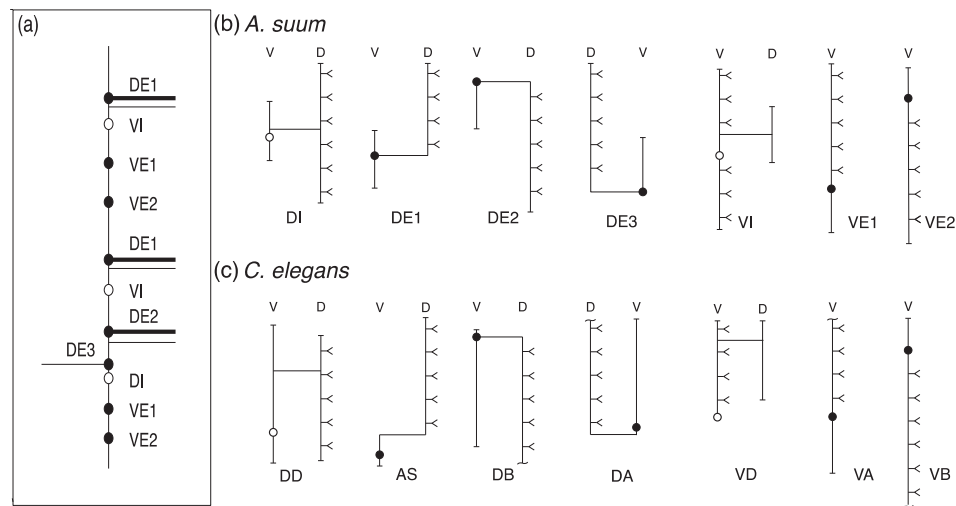


Figure 1. Schematic diagram of the ventral cord motorneurons of *A. suum*. (a) Stereotypical pattern of motorneuron cell bodies found in each segment (segments 2-5) and their associated commissures adapted from Stretton *et al.* (1978). Filled circles indicate cell bodies of excitatory motorneurons. Open circles indicate cell bodies of inhibitory motorneurons. DE1 and DE2 commissures have a larger diameter (thicker line) than the inhibitory commissures with which they are paired. (b, c) Schematic diagram of the seven types of (b) *A. suum* and (c) *C. elegans* segmental motorneurons. Circles indicate cell bodies, all found in the ventral nerve cord (V). Filled circles, excitatory neurons; open circles, inhibitory. Vertical lines emanating from the cell body represent the ventral nerve cord processes, and horizontal lines represent commissures that connect dorsal (D) and ventral processes. Forked projections represent axonal synapses onto muscle. Note that the synaptic output from DE1, VE1, and DE3 is anterior to the cell bodies, while output from DE2 and VE2 is posterior to the cell bodies.

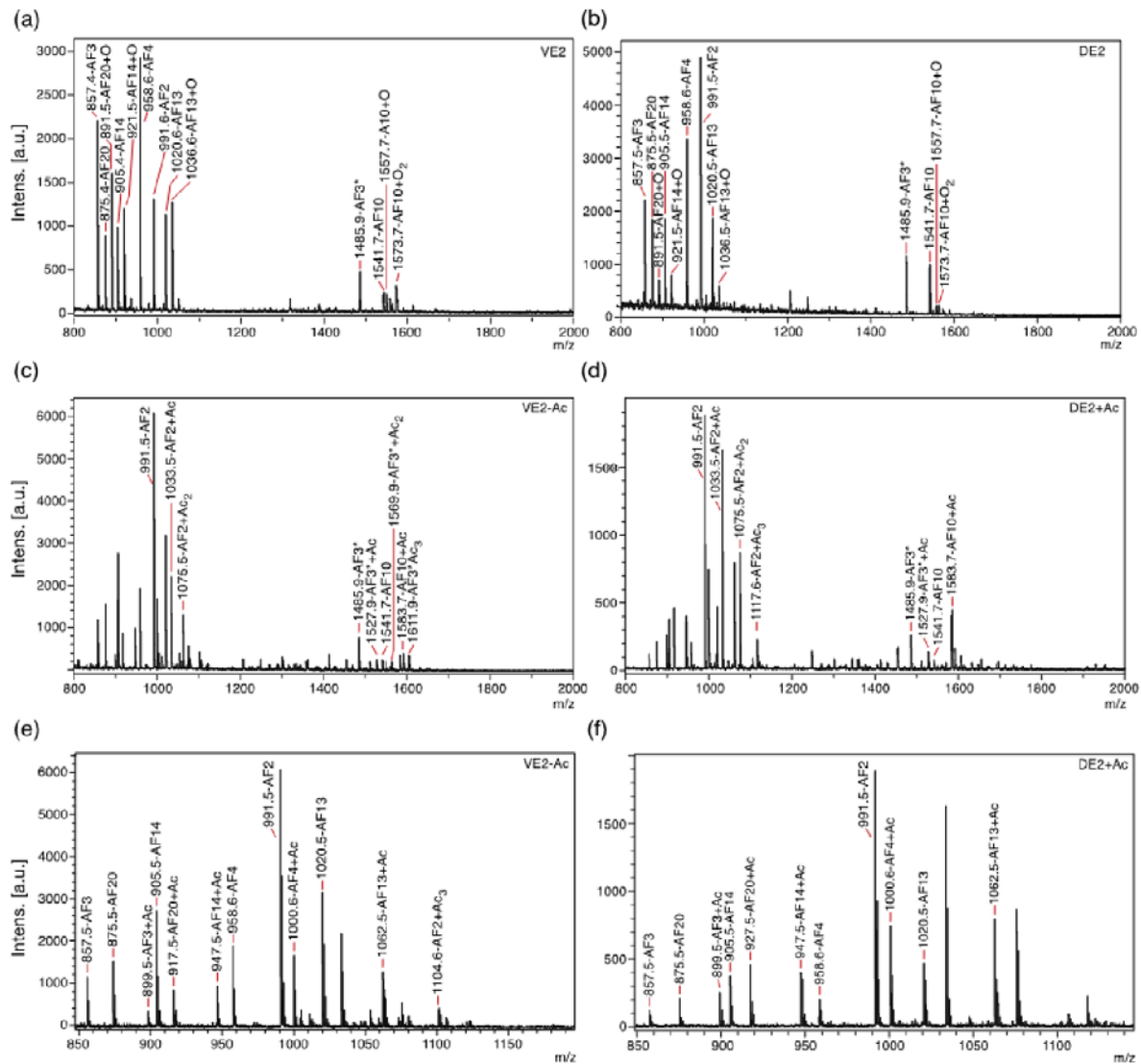


Figure 2. Neuropeptides from single excitatory motoneurons in unmodified and acetylated forms. Mass spectra from individual (a) VE2, (b) DE2 motoneurons with no chemical modifications. Spectra from individual (c) VE2, (d) DE2 motoneurons treated with acetic anhydride adding 42 mass units for each acetylation. (e) and (f) are expansions of (c) and (d), respectively, between 850 and 1150-2000 m/z . X-axis: m/z is mass-to-charge ratio. Y-axis is intensity of MS signal in arbitrary units, a.u.

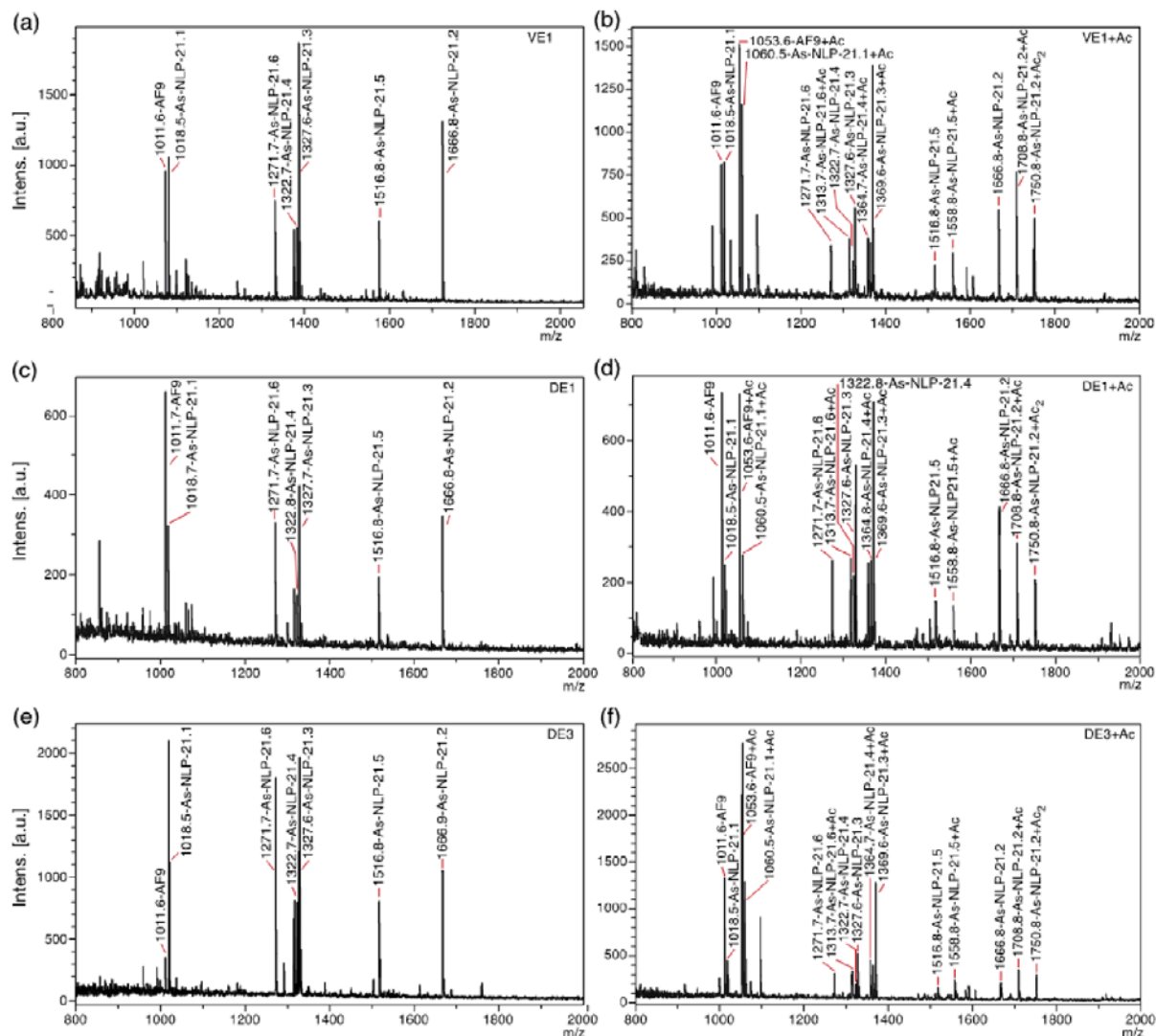


Figure 3. Neuropeptides from single excitatory motoneurons in unmodified and acetylated forms. Mass spectra from individual (a) VE1, (c) DE1, (e) DE3 motoneurons with no chemical modifications. Spectrum from an individual (b) VE1, (d) DE1, (f) DE3 treated with acetic anhydride, which acetylates free amines (N-terminus, K) and to a lesser extent Y, adding 42 mass units for each acetylation. X axis: m/z is mass-to-charge ratio. Y axis is intensity of MS signal in arbitrary units, a.u. * indicates extended peptide due to incomplete cleavage.

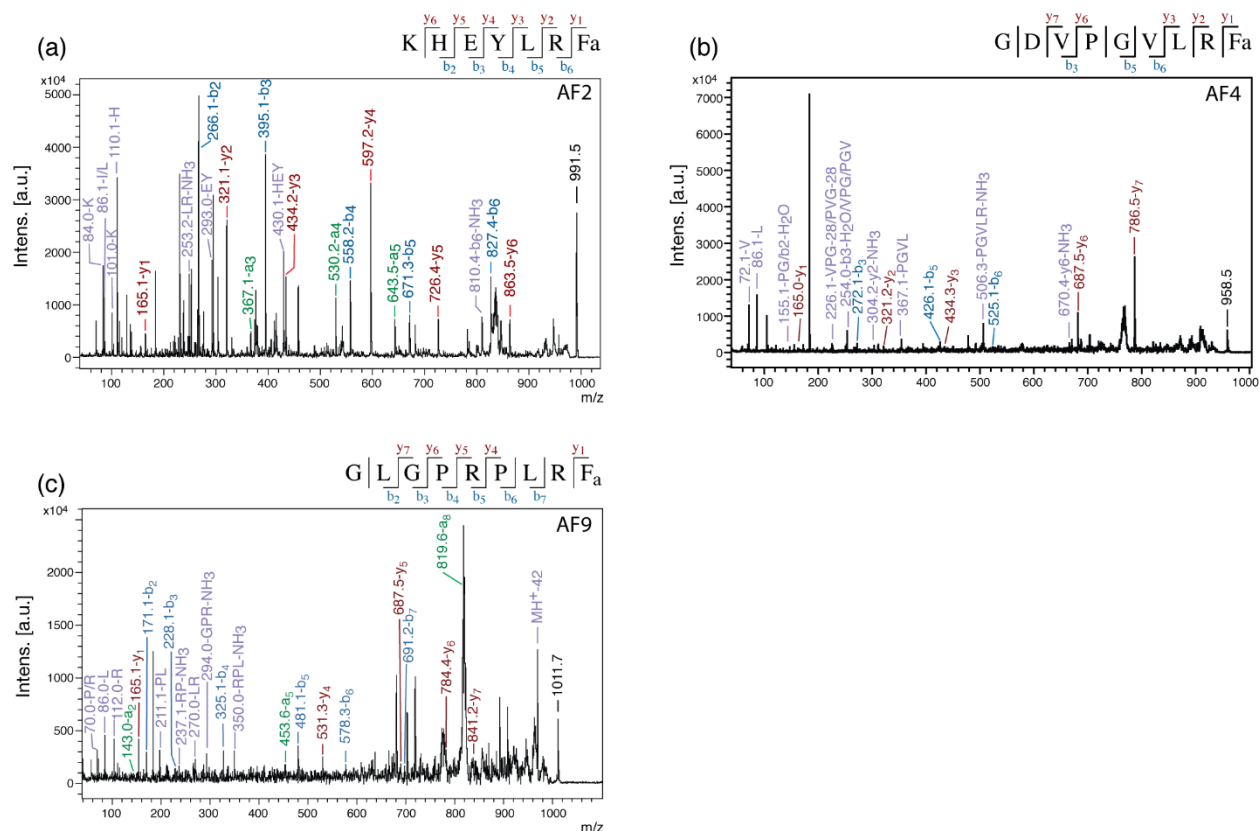
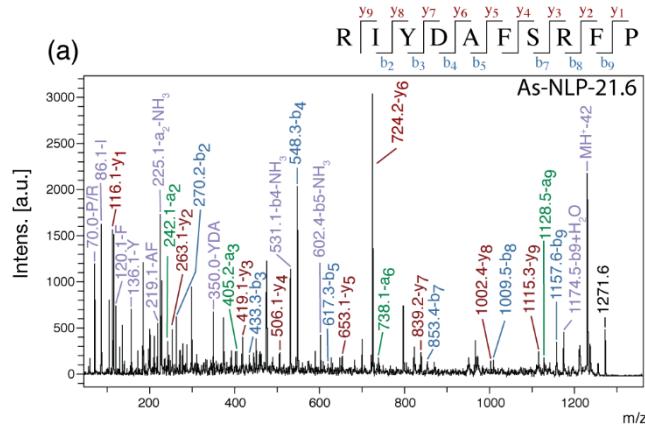


Figure 4. MS/MS spectra of known peptides from the excitatory ventral cord motoneurons. Peaks representing a (green), b (blue), y (red), and high intensity internal fragment and immonium (purple) ions are labeled, and b and y ions are summarized in the sequence at the top of each spectrum. (a) MS/MS spectrum from DE2 of AF2 (*afp-4*), *m/z* 991.6. (b) MS/MS spectrum from DE2 of AF4 (*afp-1*), *m/z* 958.5 (c) MS/MS spectrum from DE1 of AF9 (*afp-14*), *m/z* 1011.7.

(b) *As-nlp-21* transcript

acatgggggcaggttgacagccgagtgctactcaaggcgctttacgcattcgaacatgcatt
 accgtcaacttcgctctcatacttcttcttgttcactacggaatcgcacaggaaccagat
 T V N F A L I L L L V H Y G I A Q E P D
 gtattggagaaacgtggtggtggccgtccgtttacggttagttgacaaacgagctggtggc
 V L E K R G G G R P F T L V D K R A G G
 agatttttcaagattgaagatggtgatgcacttgagaagcgtggcggaggaggcctttc
R F F K I E D V D A L E K R G G G R P F
 tatactcaaacggacgaaaaacgaggaggtggacgatcgtttcgtttgagcagcttgggt
Y T Q T D E K R G G G R S F R L S S L G
 gagaagagaatggctgcacaaaaaatcgcgatgccagagctggatttttatgaaggagct
E K R M A A Q K I A M P E L D F Y E G A
 ctgcgtgaaattggcaactggccatttatcgagaaacgtgcaggtgcgagatcatttccg
 L A E I G N W P F I E K R A G A R S F P
 gtgctggaggacttactcgagaaacgaagaatttacgatgccttctctaggttcccttga
V L E D L L E K R R I Y D A F S R F P -
 agagtggatgaacacatctcgaatttcttctgtgaataccgtatgcagctcagtaaaacgaa
 attgctcaaaagaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

(c) *As-nlp-21* Gene Model

Figure 5. Sequencing As-NLP-21 peptides. (a) MS/MS sequencing of As-NLP-21.6. Peaks representing a (green), b (blue), y (red), and high-intensity internal fragment (purple) ions are labeled, and b and y ions are summarized in the sequence at the top of the spectrum. (b) *As-nlp-21* transcript; bold nucleotides indicate portion of the transcript targeted by the *As-nlp-21* anti-sense riboprobe. Blue and underlined indicates encoded peptide while bold amino acids indicate basic cleavage sites. Signal peptide start site is bolded and underlined. (c) *As-nlp-21* gene model. The gene consists of 3 exons of 93, 149, and 298 bases separated by 5459 and 3007 base introns.

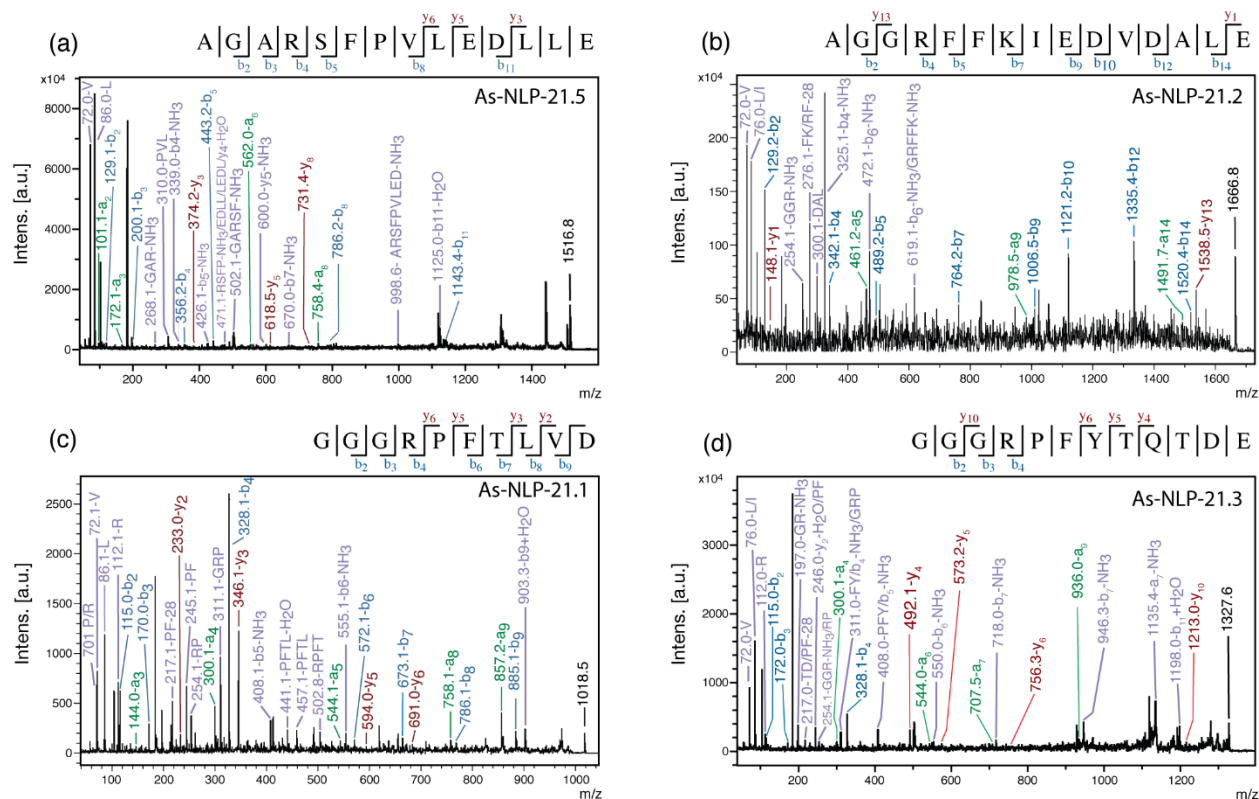


Figure 6. MS/MS of peptides encoded by *As-nlp-21*. Peaks representing a (green), b (blue), y (red), and high-intensity internal fragment and immonium (purple) ions are labeled, and b and y ions are summarized in the sequence at the top of each spectrum. The names of the peptides are below the sequences. (a) MS/MS spectrum from DE1 of As-NLP-21.5, m/z 1516.8. (b) MS/MS spectrum from DE1 of As-NLP-21.2, m/z 1666.8. (c) MS/MS spectrum from DE1 of As-NLP-21.1, m/z 1018.5. (d) MS/MS spectrum from DE1 of As-NLP-21.3, m/z 1327.6.

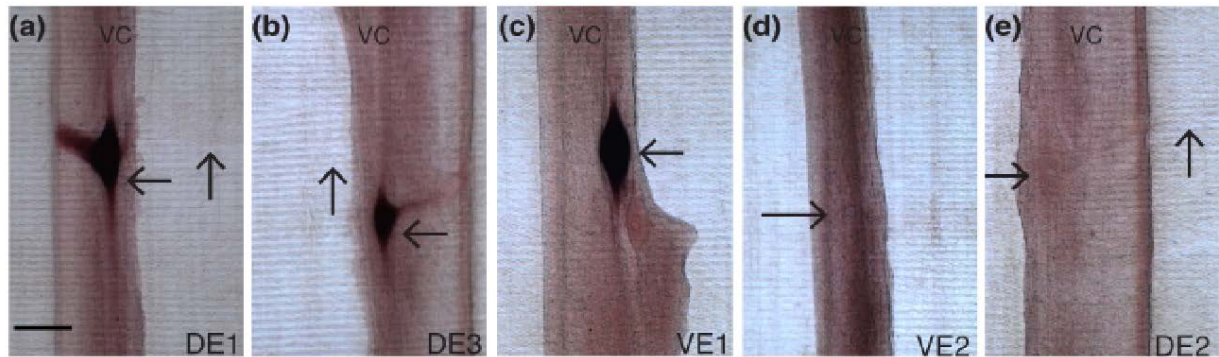


Figure 7. Confirmation of *As-nlp-21* expression in DE1, VE1, and DE3 motorneurons in *A. suum* by *in situ* hybridization. Positive *As-nlp-21 in situ* hybridization staining in DE1 (a), DE3 (b), and VE1 (c). No staining was observed in VE2 (d) or DE2 (e) VC= Ventral nerve cord. Vertical arrows indicate commissures and horizontal arrows indicate cell bodies. Scale bar = 100 μ m.

(a) *afp-14* Transcript

ctaatacgactcactatagggcaagcagtggtatcaacgcagaggtacatggggaccattt
 aagcgtgcatctgcgat**gcgtaccgtttgtacgaccttcgtgatcggtccccctgctactt**
 M R T V C T T F V I V P L L L
gtcgcaatatatttgacgcatctcttccgctttgccatcttcagatgctgatgtcgtttac

 V A I F G R I S S A L P S S D A D V V Y
agacttctgagtcagtacgggagccgagcaaattttccggatgctgatctctccaatttg
 R L L S Q Y G S R A N F P D A D L S N L
ctggaagaacagaggtcgatgaagcgaggtcttgggccacgcccactacgattcggttaa
 L E E Q R S M K R **G L G P R P L R F G** -
gcggtgttgtcgctgtcttcacactctcggcttctcgctaatgcttctcattctatcctg
 cagtgttaacctgttcgtttaatgaataaaatattcatcgaaaaaaaaaaaaaaaaaaaa
 aa

(b) *afp-14* Gene Model

Figure 8. Nucleotide sequence and deduced amino acid sequence of *afp-14*. (a) The *afp-14* transcript was found by using unique gene specific primers for 5'RACE, underlined with a solid black line, and 3'RACE, dotted underlined. Bolded nucleotides indicate the portion of the transcript targeted by the *afp-14* riboprobe. Blue, bold, and underlined indicates the encoded peptide with an N-terminal dibasic cleavage site in bold. Putative start site of the signal peptide is bolded and underlined. (b) The structure of the *afp-14* gene consists of 3 exons 134, 67, and 144 bp in length. Exons are separated by introns of 2429 and 337 bp.

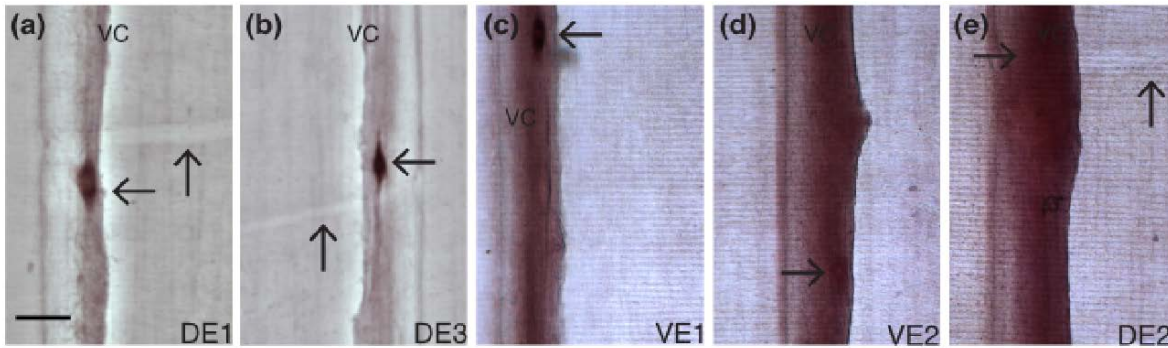


Figure 9. Confirmation of *afp-14* (AF9) expression in DE1, VE1, and DE3 motorneurons in *A. suum* by *in situ* hybridization. Positive *afp-14 in situ* hybridization staining in DE1 (a), DE3 (b), and VE1 (c). No staining was observed in VE2 (d) or DE2 (e) VC= Ventral nerve cord. Vertical arrows indicate commissures and horizontal arrows indicate cell bodies. Scale bar = 100 μ m.

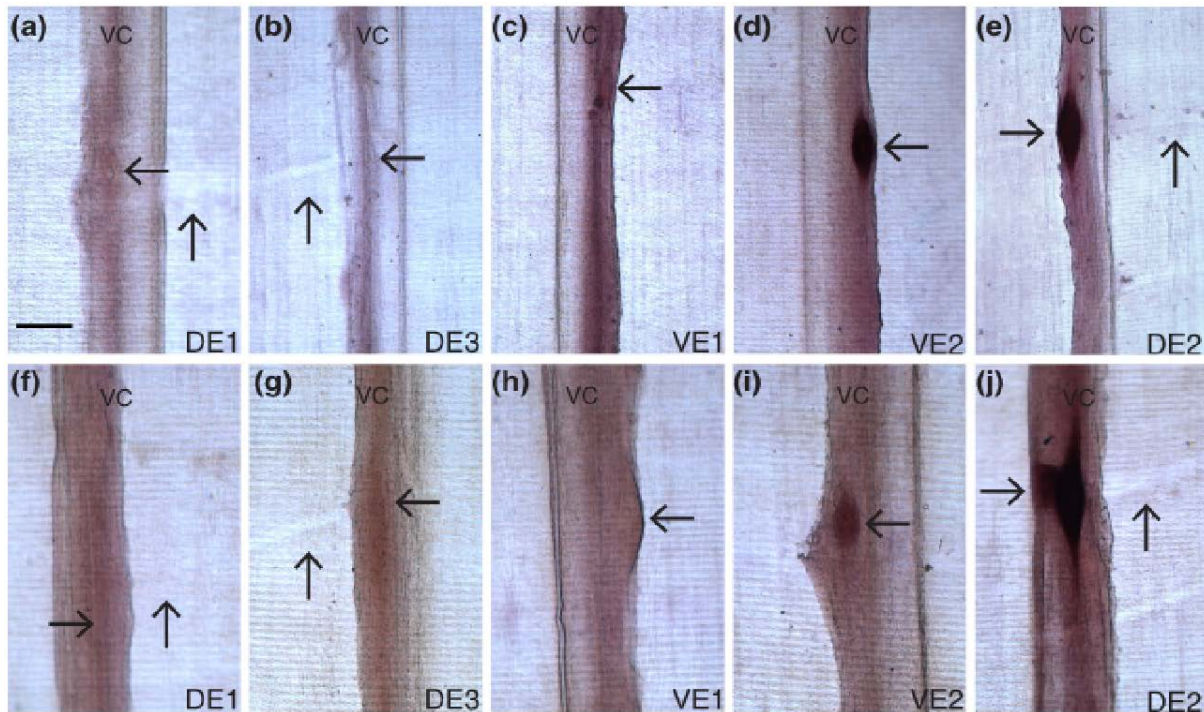


Figure 10. Confirmation of *afp-1* (AF3, 4, 10, 13, 14, 20) and *afp-4* (AF2) expression in VE2 and DE2 motorneurons in *A. suum* by *in situ* hybridization. Positive *afp-1* ISH staining in (d) VE2 and (e) DE2. Positive *afp-4* ISH staining in (i) VE2 and (j) DE2. No staining was observed in DE1, DE3, or VE1 for neither *afp-1* (a-c) nor *afp-4* (f-h) VC= Ventral nerve cord. Arrows indicate commissures. Scale bar = 100 μ m.

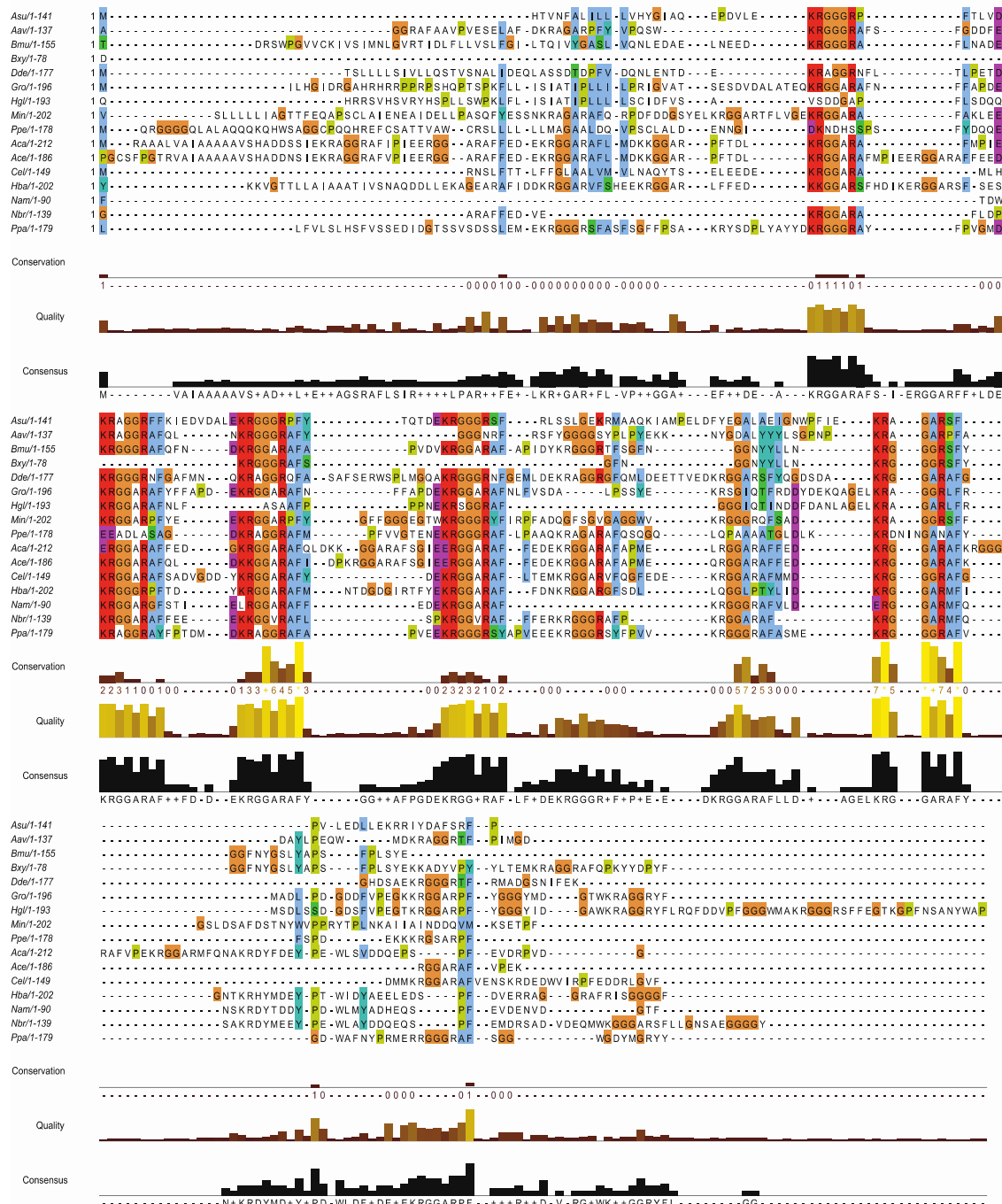


Figure 11. *As-nlp-21* multiple alignment by MEGA5.1 [39] of *nlp-21*-like transcripts from different nematodes. Results were presented in Jalview [39, 41] and show strong conservation of the predicted encoded peptides. Asu = *Ascaris suum* (Clade III), Aav = *Aphelenchus avenae* (IVb), Bmu = *Bursaphelenchus mucronatus* (IVb), Bxy = *Bursaphelenchus xylophilus* (IVb), Dde = *Ditylenchus destructor* (IVb), Gro = *Globodera rostochiensis* (IVb), Hgl = *Heterodera glycine* (IVb), Min = *Meloidogyne incognita* (IVb), Ppe = *Pratylenchus penetrans* (IVb), Aca = *Ancylostoma caninum* (V), Ace = *Ancylostoma ceylonicum* (V), Cel = *Caenorhabditis elegans* (V), Hba = *Heterorhabditis bacteriophora* (V), Nam = *Necator americanus* (V), Nbr = *Nippostrongylus brasiliensis* (V), Ppa = *Pristionchus pacificus* (V).

REFERENCES

1. White, J. G., Southgate, E., Thomson, J. N., Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci.* 314:1-340.
2. Towlson, E. K., Vertes, P. E., Ahnert, S. E., Schafer, W. R., Bullmore, E. T. (2013). The rich club of the *C. elegans* neuronal connectome. *J Neurosci.* 33:6380-7.
3. Kleinfeld, D., Bharioke, A., Blinder, P., Bock, D. D., Briggman, K. L., Chklovskii, D. B., Denk, W., Helmstaedter, M., Kaufhold, J. P., Lee, W. C., Meyer, H. S., Micheva, K. D., Oberlaender, M., Prohaska, S., Reid, R. C., Smith, S. J., Takemura, S., Tsai, P. S., Sakmann, B. (2011). Large-scale automated histology in the pursuit of connectomes. *J Neurosci.* 31:16125-38.
4. Bargmann, C. I. (2012). Beyond the connectome: how neuromodulators shape neural circuits. *Bioessays.* 34:458-65.
5. Marder, E., Calabrese, R. L., Nusbaum, M. P., Trimmer, B. (1987). Distribution and partial characterization of FMRFamide-like peptides in the stomatogastric nervous systems of the rock crab, *Cancer borealis*, and the spiny lobster, *Panulirus interruptus*. *J Comp Neurol.* 259:150-63.
6. Stretton, A., Donmoyer, J., Davis, R., Meade, J., Cowden, C., Sithigorngul, P. (1992). Motor behavior and motor nervous system function in the nematode *Ascaris suum*. *J Parasitol.* 78:206-14.
7. Meade, J., Stretton, A. (1989). Activity of motoneurons during *Ascaris* locomotion. *Soc Neurosci Abs.* 15:1299.
8. Stretton, A. O., Fishpool, R. M., Southgate, E., Donmoyer, J. E., Walrond, J. P., Moses, J. E., Kass, I. S. (1978). Structure and physiological activity of the motoneurons of the nematode *Ascaris*. *Proc Natl Acad Sci USA.* 75:3493-7.
9. Davis, R. E., Stretton, A. O. W. (1996). The motornervous system of *Ascaris*: Electrophysiology and anatomy of the neurons and their control by neuromodulators. *Parasitology.* 113:S97-S117.
10. Davis, R. E., Stretton, A. O. W. (2001). Structure-activity relationships of 18 endogenous neuropeptides on the motornervous system of the nematode *Ascaris suum*. *Peptides.* 22:7-23.
11. Kim, K., Li, C. (2004). Expression and regulation of an FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. *J Comp Neurol.* 475:540-50.

12. Yew, J. Y., Kutz, K. K., Dikler, S., Messinger, L., Li, L., Stretton, A. O. (2005). Mass spectrometric map of neuropeptide expression in *Ascaris suum*. *J Comp Neurol.* 488:396-413.
13. Jarecki, J. L., Frey, B. L., Smith, L. M., Stretton, A. O. (2011). Discovery of neuropeptides in the nematode *Ascaris suum* by database mining and tandem mass spectrometry. *J Proteome Res.* 10:3098-106.
14. Cowden, C., Stretton, A. O., Davis, R. E. (1989). AF1, a sequenced bioactive neuropeptide isolated from the nematode *Ascaris suum*. *Neuron.* 2:1465-73.
15. Reinitz, C. A., Herfel, H. G., Messinger, L. A., Stretton, A. O. (2000). Changes in locomotory behavior and cAMP produced in *Ascaris suum* by neuropeptides from *Ascaris suum* or *Caenorhabditis elegans*. *Mol Biochem Parasitol.* 111:185-97.
16. Reinitz, C. A., Pleva, A. E., Stretton, A. O. (2011). Changes in cyclic nucleotides, locomotory behavior, and body length produced by novel endogenous neuropeptides in the parasitic nematode *Ascaris suum*. *Mol Biochem Parasitol.* 180:27-34.
17. Jarecki, J. L., Viola, I. R., Andersen, K. M., Miller, A. H., Ramaker, M. A., Vestling, M. M., Stretton, A. O. (2013). Three independent techniques localize expression of transcript *afp-11* and its bioactive peptide products to the paired AVK neurons in *Ascaris suum* : *In situ* hybridization, immunocytochemistry, and single cell mass spectrometry. *ACS Chem Neurosci.* 4:418-34.
18. Cowden, C., Stretton, A. O. (1995). Eight novel FMRFamide-like neuropeptides isolated from the nematode *Ascaris suum*. *Peptides.* 16:491-500.
19. Cowden, C., Stretton, A. O. (1993). AF2, an *Ascaris* neuropeptide: isolation, sequence, and bioactivity. *Peptides.* 14:423-30.
20. Jarecki, J. L., Andersen, K., Konop, C. J., Knickelbine, J. J., Vestling, M. M., Stretton, A. O. (2010). Mapping neuropeptide expression by mass spectrometry in single dissected identified neurons from the dorsal ganglion of the nematode *Ascaris suum*. *ACS Chem Neurosci.* 1:505-19.
21. Jarecki, J. L. (2010). Identification and Localization of Neuropeptides in *Ascaris suum* By Mass Spectrometry. *Ph. D. Thesis.* University of Wisconsin-Madison.
22. Yew, J. Y., Davis, R., Dikler, S., Nanda, J., Reinders, B., Stretton, A. O. (2007). Peptide products of the *afp-6* gene of the nematode *Ascaris suum* have different biological actions. *J Comp Neurol.* 502:872-82.

23. Miyamoto, M. D. (1975). Binomial Analysis of Quantal Transmitter Release at Glycerol Treated Frog Neuromuscular-Junctions. *Journal of Physiology-London*. 250:121-42.
24. Pegg, D. E. (1976). Long-term preservation of cells and tissues: a review. *J Clin Pathol*. 29:271-85.
25. Nemes, P., Rubakhin, S. S., Aerts, J. T., Sweedler, J. V. (2013). Qualitative and quantitative metabolomic investigation of single neurons by capillary electrophoresis electrospray ionization mass spectrometry. *Nat Protoc*. 8:783-99.
26. Rubakhin, S. S., Greenough, W. T., Sweedler, J. V. (2003). Spatial profiling with MALDI MS: Distribution of neuropeptides within single neurons. *Anal Chem*. 75:5374-80.
27. Rubakhin, S. S., Sweedler, J. V. (2008). Quantitative measurements of cell-cell signaling peptides with single-cell MALDI MS. *Anal Chem*. 80:7128-36.
28. Walrond, J. P., Kass, I. S., Stretton, A. O., Donmoyer, J. E. (1985). Identification of excitatory and inhibitory motoneurons in the nematode *Ascaris* by electrophysiological techniques. *J Neurosci*. 5:1-8.
29. Chalfie, M., Sulston, J. E., White, J. G., Southgate, E., Thomson, J. N., Brenner, S. (1985). The neural circuit for touch sensitivity in *Caenorhabditis elegans*. *J Neurosci*. 5:956-64.
30. Wen, Q., Po, M. D., Hulme, E., Chen, S., Liu, X. Y., Kwok, S. W., Gershow, M., Leifer, A. M., Butler, V., Fang-Yen, C., Kawano, T., Schafer, W. R., Whitesides, G., Wyart, M., Chklovskii, D. B., Zhen, M., Samuel, A. D. T. (2012). Proprioceptive coupling within motor neurons drives *C. elegans* forward locomotion. *Neuron*. 76:750-61.
31. Johnson, C., Stretton, A. (1980). Neural control of locomotion in *Ascaris*: anatomy, electrophysiology, and biochemistry. 1:159-95.
32. Nemes, P., Knolhoff, A. M., Rubakhin, S. S., Sweedler, J. V. (2012). Single-cell metabolomics: changes in the metabolome of freshly isolated and cultured neurons. *ACS Chem Neurosci*. 3:782-92.
33. Rubakhin, S. S., Churchill, J. D., Greenough, W. T., Sweedler, J. V. (2006). Profiling signaling peptides in single mammalian cells using mass spectrometry. *Anal Chem*. 78:7267-72.
34. Rubakhin, S. S., Sweedler, J. V. (2007). Characterizing peptides in individual mammalian cells using mass spectrometry. *Nat Protoc*. 2:1987-97.

35. Schwartz, S. A., Reyzer, M. L., Caprioli, R. M. (2003). Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation. *J Mass Spectrom.* 38:699-708.
36. Hilker, B., Clifford, K. J., Sauter, A. D., Jr., Sauter, A. D., 3rd, Harmon, J. P. (2009). The measurement of charge for induction-based fluidic MALDI dispense event and nanoliter volume verification in real time. *J Am Soc Mass Spectrom.* 20:1064-7.
37. Jex, A. R., Liu, S., Li, B., Young, N. D., Hall, R. S., Li, Y., Yang, L., Zeng, N., Xu, X., Xiong, Z., Chen, F., Wu, X., Zhang, G., Fang, X., Kang, Y., Anderson, G. A., Harris, T. W., Campbell, B. E., Vlaminck, J., Wang, T., Cantacessi, C., Schwarz, E. M., Ranganathan, S., Geldhof, P., Nejsun, P., Sternberg, P. W., Yang, H., Wang, J., Gasser, R. B. (2011). *Ascaris suum* draft genome. *Nature.* 479:529-33.
38. Wang, J., Czech, B., Crunk, A., Wallace, A., Mitreva, M., Hannon, G. J., Davis, R. E. (2011). Deep small RNA sequencing from the nematode *Ascaris* reveals conservation, functional diversification, and novel developmental profiles. *Genome Res.* 21:1462-77.
39. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 28:2731-9.
40. Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792-7.
41. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* 25:1189-91.
42. Nanda, J. C., Stretton, A. O. (2010). In situ hybridization of neuropeptide-encoding transcripts *afp-1*, *afp-3*, and *afp-4* in neurons of the nematode *Ascaris suum*. *J Comp Neurol.* 518:896-910.
43. Ash, A. S., Tucker, J. F. (1967). The bioassay of gamma-aminobutyric acid using a muscle preparation from *Ascaris lumbricoides*. *J Pharm Pharmacol.* 19:240-5.
44. Kass, I. S., Larsen, D. A., Wang, C. C., Stretton, A. O. (1982). *Ascaris suum*: differential effects of avermectin B1a on the intact animal and neuromuscular strip preparations. *Exp Parasitol.* 54:166-74.
45. Johnson, C. D., Stretton, A. O. W. (1985). Localization of choline-acetyltransferase within identified motoneurons of the nematode *Ascaris*. *J Neurosci.* 5:1984-92.

46. Angstadt, J. D., Donmoyer, J. E., Stretton, A. O. (1989). Retrovesicular ganglion of the nematode *Ascaris*. *J Comp Neurol*. 284:374-88.
47. Guastella, J., Johnson, C. D., Stretton, A. O. (1991). GABA-immunoreactive neurons in the nematode *Ascaris*. *J Comp Neurol*. 307:584-97.
48. Johnson, C. D., Stretton, A. O. (1987). GABA-immunoreactivity in inhibitory motor neurons of the nematode *Ascaris*. *J Neurosci*. 7:223-35.
49. Edison, A. S., Messinger, L. A., Stretton, A. O. (1997). *afp-1*: A gene encoding multiple transcripts of a new class of FMRFamide-like neuropeptides in the nematode *Ascaris suum*. *Peptides*. 18:929-35.
50. Khatun, J., Ramkissoon, K., Giddings, M. C. (2007). Fragmentation characteristics of collision-induced dissociation in MALDI TOF/TOF mass spectrometry. *Anal Chem*. 79:3032-40.
51. Maroney, P. A., Denker, J. A., Darzynkiewicz, E., Laneve, R., Nilsen, T. W. (1995). Most mRNAs in the nematode *Ascaris lumbricoides* are trans-spliced: a role for spliced leader addition in translational efficiency. *RNA*. 1:714-23.
52. Sithigorngul, P., Jarecki, J. L., Stretton, A. O. (2011). A specific antibody to neuropeptide AF1 (KNEFIRFamide) recognizes a small subset of neurons in *Ascaris suum*: differences from *Caenorhabditis elegans*. *J Comp Neurol*. 519:1546-61.
53. Sithigorngul, P., Nanda, J. C., Stretton, A. O. (2003). A strategy for isolating rare peptides: isolation and sequencing of a large peptide present in a single neuron of the nematode *Ascaris suum*. *Peptides*. 24:1025-33.
54. Etchberger, J. F., Lorch, A., Sleumer, M. C., Zapf, R., Jones, S. J., Marra, M. A., Holt, R. A., Moerman, D. G., Hobert, O. (2007). The molecular signature and cis-regulatory architecture of a *C. elegans* gustatory neuron. *Genes Dev*. 21:1653-74.
55. Zhang, F., O'Meara, M. M., Hobert, O. (2011). A left/right asymmetric neuronal differentiation program is controlled by the *Caenorhabditis elegans* *lsy-27* zinc-finger transcription factor. *Genetics*. 188:753-9.
56. Kratsios, P., Stolfi, A., Levine, M., Hobert, O. (2012). Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. *Nat Neurosci*. 15:205-14.

57. McCoy, C. J., Atkinson, L. E., Zamanian, M., McVeigh, P., Day, T. A., Kimber, M. J., Marks, N. J., Maule, A. G., Mousley, A. (2014). New insights into the FLPerGic complements of parasitic nematodes: Informing deorphanisation approaches. *EuPA Open Proteom.* 3:262-72.
58. McVeigh, P., Geary, T. G., Marks, N. J., Maule, A. G. (2006). The FLP-side of nematodes. *Trends Parasitol.* 22:385-96.
59. McVeigh, P., Leech, S., Mair, G. R., Marks, N. J., Geary, T. G., Maule, A. G. (2005). Analysis of FMRFamide-like peptide (FLP) diversity in phylum Nematoda. *Int J Parasitol.* 35:1043-60.
60. Kimber, M. J., Fleming, C. C., Prior, A., Jones, J. T., Halton, D. W., Maule, A. G. (2002). Localisation of *Globodera pallida* FMRFamide-related peptide encoding genes using *in situ* hybridisation. *Int J Parasitol.* 32:1095-105.
61. Li, C., Kim, K., Nelson, L. S. (1999). FMRFamide-related neuropeptide gene family in *Caenorhabditis elegans*. *Brain Res.* 848:26-34.
62. Stawicki, T. M., Takayanagi-Kiya, S., Zhou, K., Jin, Y. (2013). Neuropeptides function in a homeostatic manner to modulate excitation-inhibition imbalance in *C. elegans*. *PLoS Genet.* 9:e1003472.
63. Kubiak, T. M., Larsen, M. J., Nulf, S. C., Zantello, M. R., Burton, K. J., Bowman, J. W., Modric, T., Lowery, D. E. (2003). Differential Activation of "Social" and "Solitary" Variants of the *Caenorhabditis elegans* G Protein-coupled Receptor NPR-1 by Its Cognate Ligand AF9. *J Biol Chem.* 278(36):33724-29.
64. Rogers, C., Reale, V., Kim, K., Chatwin, H., Li, C., Evans, P., de Bono, M. (2003). Inhibition of *Caenorhabditis elegans* social feeding by FMRFamide-related peptide activation of NPR-1. *Nat Neurosci.* 6:1178-85.
65. Dossey, A. T., Reale, V., Chatwin, H., Zachariah, C., deBono, M., Evans, P. D., Edison, A. S. (2006). NMR analysis of *Caenorhabditis elegans* FLP-18 neuropeptides: implications for NPR-1 activation. *Biochemistry.* 45:7586-97.
66. McVeigh, P., Alexander-Bowman, S., Veal, E., Mousley, A., Marks, N. J., Maule, A. G. (2008). Neuropeptide-like protein diversity in phylum Nematoda. *Int J Parasitol.* 38:1493-503.
67. Nathoo, A. N., Moeller, R. A., Westlund, B. A., Hart, A. C. (2001). Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. *Proc Natl Acad Sci U S A.* 98:14000-5.

68. Puttachary, S., Robertson, A. P., Clark, C. L., Martin, R. J. (2010). Levamisole and ryanodine receptors. II: An electrophysiological study in *Ascaris suum*. *Mol Biochem Parasitol.* 171:8-16.
69. Robertson, A. P., Clark, C. L., Martin, R. J. (2010). Levamisole and ryanodine receptors. I: A contraction study in *Ascaris suum*. *Mol Biochem Parasitol.* 171:1-7.
70. Goodwin, P. R., Juo, P. (2013). The scaffolding protein SYD-2/Liprin-alpha regulates the mobility and polarized distribution of dense-core vesicles in *C. elegans* motor neurons. *PLoS One.* 8:e54763.

CHAPTER FIVE

Identification of *A. suum* Neuropeptide Receptors and their Endogenous Ligands by Heterologous Expression and Reverse Pharmacology

Jennifer J. Knickelbine, Zhihui Yao, Rebecca J. Garry, Samuel H. Gellman, and Antony O. W.
Stretton

INTRODUCTION

Neuropeptides can dramatically affect nematode locomotory behavior. In *A. suum*, each somatic motoneuron expresses a small complement of neuropeptides: DE1, DE3, and VE1 anteriorly-projecting excitatory neurons express peptides from the *afp-14* and *nlp-21* transcripts [1], DE2 and VE2 posteriorly-projecting excitatory motoneurons express peptides encoded by *afp-1* and *afp-4* [1], and the DI and VI inhibitory motoneurons express a single peptide encoded by *nlp-22* [2]. Nearly all of these peptides have profound effects on behavior; for example, AF2 (encoded by *afp-4*) induced jerky movement and quasi-static waves when injected into intact worms [3], while As-NLP-22 (encoded by *nlp-22*) induced persistent flaccid paralysis [2]. A great deal of effort has gone into identifying the neuropeptides that may affect behavior, but very little is known about the signaling mechanisms that they employ, and the receptors that mediate the effects.

The majority of known peptide-mediated responses involve signaling through GPCRs. Over 1000 GPCRs have been predicted in *C. elegans*, at least 128 of which are predicted to bind neuropeptides [4-6]. Most of these have not yet been paired with their respective ligands. In *A. suum*, the number of putative neuropeptides (>250 [7]) is much higher than the number of

predicted neuropeptide GPCRs (ca. 62, [8, 9]), so it is highly likely that single GPCRs bind multiple neuropeptides. Already, there are examples of such promiscuity in *C. elegans*; for example, the C10C6.2 (NPR-3) receptor binds all 3 of the *flp-15* peptides [10], and the Y59H11AL.1 (NPR-22) receptor binds 15 peptides from 6 different *flp* genes, with EC₅₀ values ranging from 25 nM to 5 μ M [11].

To deorphanize such receptors, a high-throughput screen of potential ligands is needed. Several expression systems are commonly used to deorphanize GPCRs. Bacterial expression systems are cheap and generally easy to work with; however, bacteria have limited capability for post-translational modifications (PTMs), making this system unsuitable for overexpression of many eukaryotic membrane proteins. Insect and mammalian cell lines allow for high levels of expression, as well as multiple PTMs, but these systems are more costly and labor-intensive than bacterial systems. In contrast, yeast expression systems are highly amenable to genetic transformation, capable of important post-translational modifications, and are much less expensive than working with insect or mammalian cell lines. Yeast systems also have the potential for high-throughput screening of potential ligands using simple growth assays. For these reasons, yeast systems can be a good compromise between cost, labor, and cell capabilities.

Nematode GPCRs have been difficult to study, in part because they do not always express well in insect and mammalian cell lines [12, 13]. However, several groups have had success using yeast expression systems [12, 14, 15]. We have cloned several candidate receptors into *S. cerevisiae* using the *FUS1-HIS3* reporter gene system [16]. We used 3 strains of yeast (CY13393, CY13399, and YEX108), which differ in their expression of G protein α subunits

(the last 5 amino acids of the single endogenous yeast G_α were replaced with those of rat G_i , G_s , and G_q ; these amino acids are also identical in *A. suum* G_s and G_q) [12, 14]. This increases the likelihood of getting good coupling between the candidate nematode receptor and the G_α subunit, since some mammalian and nematode receptors do not interact well with the endogenous yeast G_α . This can also provide some information about the downstream signaling mechanisms involved. This system ties receptor activation to a pheromone-responsive growth pathway [12, 16], such that activation of the receptor causes increased growth, which can be measured using the fluorescent cell growth indicator Alamar Blue. In this manner, we can pair *Ascaris* GPCRs with their neuropeptide ligands.

A wealth of sequence information is now available with publication of the *A. suum* genome [8] and transcriptome [9], including 279 predicted *A. suum* GPCRs, 62 of which are predicted to bind peptides. Combining this information with single-cell RNA-seq data collected by our laboratory, we have initially selected 7 candidate GPCRs that are expressed in motoneurons, muscle cells, or both (Table 1). Here, we clone, for the first time, 3 of these receptors into *Saccharomyces cerevisiae* and determine the endogenous neuropeptide ligands by reverse pharmacology.

We now know which neuropeptides are produced by MNs [1, 2], and many of these have significant effects on locomotory behavior. However, peptides make poor drugs, in large part due to their sensitivity to proteases *in vivo*. A viable drug candidate must be more resistant to endogenous proteases, and ideally be more tissue-permeable than natural peptides. We are collaborating with Zhihui Yao (Gellman laboratory, Department of Chemistry, UW-Madison) to design α/β -peptide analogs of our bioactive peptides. In α/β -peptides, one or more α -amino

acids are replaced with a β -amino acid, which is elongated by one CH_2 group within the polymeric chain. It has been shown that this makes these compounds more resistant to proteolysis (e.g. the α/β -Bim BH3 analog was >100-fold more resistant to proteinase K treatment than the parent peptide [17]), and can improve tissue- and cell-permeability, all while maintaining the bioactivity of the parent peptide. The Gellman lab has extensive experience in designing, synthesizing, and analyzing these molecules [17, 18], and our laboratory can now screen for binding to neuropeptide GPCRs using the yeast assay, and examine effects on locomotory behavior of *A. suum* through muscle contractile assays and injections into intact worms [1-3, 19]. With our newfound knowledge of the neuropeptides and GPCRs (from single-cell RNA-seq) that are expressed by the MNs, we are now in prime position to deorphanize these receptors and begin testing chemical analogs of bioactive peptides as a first step toward an entirely new class of anti-nematode drugs.

MATERIALS AND METHODS

Identification of GPCR Transcripts

BLAST searches were conducted as previously described [1, 2] to identify sequences of GPCRs in the *A. suum* genome [8] and transcriptome [20]. Several of these were partial transcripts with only 4-6 predicted transmembrane domains, instead of the typical 7 found in GPCRs. These sequences were used to design primers for PCR and RACE, which were performed as described in Chapters 2-3 to extend the transcripts. The full list of primers used in these studies is presented in Table 2.

Phylogenetic Analysis of GPCR Sequences in Nematodes

BLAST searches were conducted as previously described [1, 2] using nucleotide collections, transcriptomic and genomic libraries, and non-redundant protein databases to identify GPCR homologs in additional nematode species. Alignments were generated using MEGA7 [21] and displayed in Jalview [22], as previously described [1, 2]. Phylogenetic analysis was conducted using the Neighbor-Joining method [23] in MEGA7, with clustering determined by bootstrap analysis (500 replicates), and evolutionary distances calculated using the Poisson correction method [24]. We also created trees using the Maximum Likelihood method [25], but the trees were nearly identical, and as such were not included in this manuscript.

Preparation of GPCR Constructs

A. suum GPCR transcripts were PCR-amplified using the primers labeled GENE-ExpF and GENE-ExpR (Table 2), which were designed to include restriction enzyme cut sites (NcoI for forward primers, XbaI for reverse primers) flanking the gene. PCR products were purified by running on a 1% agarose gel, and extracting the bands using the QIAquick Gel Extraction Kit (Qiagen). The ends of the sequences were digested with NcoI and XbaI (New England Biolabs) at 37°C for 1 hour. Vector plasmids (Cp1651 or Cp4258, obtained from Dr. James Broach, Penn State College of Medicine) were similarly digested with NcoI and XbaI, followed by treatment with shrimp alkaline phosphatase (New England Biolabs) at 37°C for 1 hour to remove the 3' phosphate groups at the ends of the linear plasmid, preventing recircularization. Both the GPCR inserts and the vector were purified by running on a 1% agarose gel as described above. Purified vector and insert were combined in 1:1, 1:3, or 1:5 molar ratios and ligated together

using T4 DNA ligase (Invitrogen) according to the manufacturer's protocols. The reactions were then diluted 1:5 and used directly for transformation into JM109 competent *E. coli* (Promega) using standard protocols. Individual clones were selected and cultured in Luria-Bertani (LB) + ampicillin medium. Plasmid DNA from the clones was isolated using the QIAprep Spin Miniprep Kit (Qiagen), and sequenced by the University of Wisconsin Biotechnology Center to confirm receptor orientation and fidelity.

In cases where the cloned transcript was incomplete or contained intervening sequences that resulted in premature stop codons (e.g. *npr-1* and *frpr-6* in Results and Discussion), we had the GPCR constructs (longer receptor isoforms, where applicable) synthesized by GenScript (Piscataway, NJ), and optimized the codon usage for expression in yeast. The inserts were then incorporated into the Cp1651 plasmid for expression.

Yeast Expression Assay

This assay is adapted from the pheromone-responsive growth assay developed by Wang *et al.* (2006) [16], with adaptations similar to those in Kimber *et al.* (2009) [12] and Larsen *et al.*, (2013) [14] to characterize an *A. suum* muscarinic acetylcholine receptor and *C. elegans* FMRFamide receptor, respectively. The *Saccharomyces cerevisiae* strains used in these experiments (CY13393, CY13399, YEX108) were developed and kindly provided by Dr. James Broach (Penn State University College of Medicine).

Competent yeast cells were prepared from mid- to late-log phase CY13393, CY13399, and YEX108 cultures using the EZ Yeast Transformation Kit (Zymo Research, Irvine, CA),

according to the manufacturer's specifications. Competent cells were stored at -80°C for 24 hours to 3 months before transformation.

For transformation, the cells were thawed at room temperature and combined with approximately 1 µg receptor-vector construct or empty vector as a mock-transfected control. The mixture was incubated with EZ3 solution from the kit for 1-3 hours at 30°C, with brief mixing 3-4 times during the incubation. Then 100-150 µL of the transformation mixture was plated onto pre-warmed leucine-deficient plates [1X YNB (Difco), 1X yeast synthetic drop-out medium supplement without leucine, histidine, tryptophan, and uracil (Sigma), supplemented with 1X histidine, tryptophan, and uracil, 2% dextrose, 2% bacto-agar (Difco)], and incubated at 30°C for 2-4 days.

Positive transformants were grown in leucine-deficient liquid medium [1X YNB (Difco), 1X yeast synthetic drop-out medium supplement without leucine, histidine, tryptophan, and uracil (Sigma), supplemented with 1X L-histidine, L-tryptophan, and uracil, 2% dextrose] in a 30°C shaking incubator until mid-log phase. Plasmids were isolated using the Zymoprep Yeast Plasmid Miniprep II Kit (Zymo Research), and verified by PCR to confirm construct fidelity.

For the growth assay, 4 mL of leucine-deficient medium was inoculated with receptor-expressing or mock-transfected yeast and grown at 30°C until mid-log phase. The cells from 1 mL of culture were pelleted and washed 2X with leucine and histidine-deficient medium [1X YNB (Difco), 1X yeast synthetic drop-out medium supplement without leucine, histidine, tryptophan, and uracil (Sigma), supplemented with 1X L-tryptophan and uracil, 10 mM ammonium sulfate (Sigma), 2% dextrose, 50 mM 4-Morpholinepropanesulfonic acid (MOPS), pH 6.8], followed by one wash with leucine and histidine-deficient medium containing 1.5 mM

3-amino-1,2,4-triazole (Sigma), a competitive inhibitor of the His3 product (required for histidine biosynthesis and growth on histidine-deficient medium) and an additional challenge to the yeast. The cells were then diluted to $OD_{600} \approx 0.01$ in the same medium, and 95 μL of this culture was added to each well of a black-walled, clear bottom 96-well plate. Test agonists (A. *suum* neuropeptides, Table 3) were applied at 10 μM , and the volume in each well was adjusted to 100 μL with leucine and histidine-deficient medium containing 1.5 mM 3-amino-1,2,4-triazole. Each experimental condition was performed in 3-6 replicates per plate. The plates were incubated at 30°C, with shaking, for 24 hours.

The next day, 10 μL of Alamar Blue (Invitrogen) was added to each well. Alamar Blue is a fluorescent cell growth indicator whose reduction results in a blue to pink color change. It can be fully and irreversibly converted to the reduced form by autoclaving a sample in blank solution, and this was used as an assay control along with blanks.

After addition of Alamar Blue, the plates were transferred to a Cytation3 Multimode Plate Reader (BioTek, Winooski, VT). The reader was programmed using Gen5 software (Version 2.04.11, BioTek) to measure fluorescence intensity (560 nm excitation, 590 nm emission, gain 68) at 6.5 mm read height every 30 min for up to 24 hours, with orbital shaking at 30°C between reads. This provided initial growth curves, and was used to choose a single endpoint (48 hours) for the experiments shown below. Data was reported as the average response of 12-24 replicates from at least 3 different plates, after blank subtraction and conversion to percent of the no-agonist control. Error was reported as the standard error of the mean, and significance was determined using unpaired *t*-tests with a significance threshold of $P < 0.05$ or $P < 0.01$.

For dose-response experiments, the 48-hour single endpoint was too late for the higher ligand concentrations, so we measured fluorescence every 30 min after addition of Alamar Blue for up to 18 hours. Dose-response curves and EC₅₀ calculations were generated using Prism 7 software (GraphPad Software, Inc, La Jolla, CA), with constants estimated by nonlinear regression.

Design and Synthesis of Peptide Analogs

We collaborated with Zhihui Yao and the laboratory of Dr. Samuel Gellman (UW Chemistry) to design and synthesize α/β -peptide analogs of one of the bioactive neuropeptides produced by the motoneurons AF2 (KHEYLRFamide). In α/β -peptides, one or more α -amino acids are replaced with a β -amino acid, which is elongated by one -CH₂ group within the polymeric chain. These compounds can be screened for binding to neuropeptide GPCRs using the yeast assay described above, and examined for effects on locomotory behavior through muscle contractile assays [1-3, 19]. The chemical structures and purities of the α/β -peptides used in these experiments are displayed in Table 4.

RESULTS AND DISCUSSION

Potential Modes of Neuropeptide Action Through GPCRs

When characterizing neuropeptide signaling through GPCRs, we must consider two general types of action: direct effects of peptides through dedicated GPCRs, or binding of peptide to a GPCR that is part of another signaling system. In the first case, the neuropeptide is the primary ligand of the receptor, and binding of peptide triggers a cascade that has

immediate effects on the cell. For example, a neuropeptide binding to a GPCR expressed on muscle cell membranes triggers conversion of ATP to cAMP, an increase in kinase activity, and ultimately, contraction of the muscle spindle. In the second scenario, binding of peptide alone does not necessarily activate the GPCR, but rather modifies the GPCR response to its normal ligand. An example of this would be a receptor that binds acetylcholine to produce muscle contraction, but activation can be “tuned” in the presence of a peptide so that the response or the time course is altered. In this way, a cell or network can use allosteric modulation to dial in the appropriate amount of activity like setting a thermostat.

While it is easy to imagine direct effects of peptides on muscle contraction, indirect effects could also greatly impact the locomotory circuit. One example could be the interaction of neuropeptides with the classical transmitter signaling systems, either as additional ligands or allosteric modulators of classical transmitter receptors or through their own dedicated GPCRs. Release of peptide from the motoneurons could modulate the release or response of the classical transmitter (ACh for excitors and GABA for inhibitors) at the neuromuscular junction. G_q activity has been linked to facilitation of ACh release from presynaptic motoneurons in *C. elegans* [26, 27], providing a precedent for this type of action. Alternatively, since several types of motoneuron are known to synapse with other motoneurons in the ventral nerve cord [28], the peptides that they express could be responsible for “tuning” the circuit and managing responses between the dorsal and ventral sides of the worm. Finally, since peptides are not degraded as quickly as ACh, and can act at a distance from the site of release, these neuropeptides could be acting to set the relative excitability of the system. This concept is

known as “context setting”, and could determine the magnitude or the time scale of the cellular responses.

All of these potential modes of action will be greatly affected by the peptide concentration *in vivo*, and this is unknown. Ignorance of the ongoing peptide concentration in time and space during relevant behavior is a major gap in our knowledge. Even the complement and concentration of neuropeptides in the pseudocoelomic fluid has not yet been adequately described. This is an issue that will need attention in the future.

At present, we only have dose-response data for one *A. suum* bioactive peptide, As-NLP-22, which inhibited muscle contraction with an IC_{50} of 8.3 nM [2]. Further, many of the neuropeptide transcripts we have identified encode more than one peptide, and it is not unusual for these peptides to be unrelated in sequence. Since peptides can act at a distance, there is also a great potential for interactions between neuropeptides released from different cells. Some nematode GPCRs are highly specific for peptides from a single transcript [10], while others have been shown to bind up to 15 peptides from 6 *flp* genes [11]. These results highlight the need for both screening of all known neuropeptides against candidate GPCRs, and determining relative potency of the potential ligands compared with *in vivo* concentration/localization to determine which effects are relevant to the locomotory circuit.

Identification of GPCR Transcripts in A. suum

To identify candidate GPCRs that might bind to the neuropeptides from the motoneurons, we used a two-pronged approach. First, we took advantage of deorphanization efforts in *C. elegans* and sought out receptors that bound to homologs of *A. suum* peptides we

have found in the motoneurons. Neuropeptide sequences are strongly conserved between these species in most cases, despite differences in cellular expression of peptides. We hypothesized that the strong conservation might extend to ligand binding in GPCRs. Second, we are analyzing each class of motoneuron and muscle using single-cell RNA-seq. These experiments are ongoing, but preliminary results have shown that many of these GPCR transcripts are expressed by the motoneurons (Table 1). Through RNA-seq, we are also identifying receptors that have not yet been characterized in *C. elegans*. While we are eager to characterize these receptors in the future, that is beyond the scope of this study.

(a) npr-1

The NPR-1 receptor has been shown to bind *flp-21* and *flp-18* peptides in *C. elegans*, employing an α_i/o pathway to affect food-dependent behaviors (social versus solitary feeding), thermal avoidance, ethanol tolerance, and innate immunity [6, 29-31]. FLP-21 is homologous to AF9 in *A. suum*, which we have shown to be expressed by the DE1, DE3, and VE1 excitatory motoneurons [1]. Furthermore, preliminary RNA-seq analysis of single motoneurons identified *npr-1* transcript in DE1 neurons (C. Konop, personal communication). This made NPR-1 a good candidate for investigating the actions of our neuropeptides.

We cloned the *npr-1* transcript from mRNA extracted from *A. suum* heads (Figure 1, MF319187). The resulting product was 1,690 bp long (8 exons), which translated into a 434-amino acid receptor that included a 25 bp intervening sequence between exons 7-8, resulting in a premature stop codon. This intervening sequence is immediately adjacent to exon 7 in the *A. suum* genome (AEUI02000547.1), with putative splice sites on either end. It also falls after what

would be TM7, though the TMHMM transmembrane domain prediction tool (version 2.0) only predicted 6 transmembrane domains for the cloned sequence in both the short and long isoforms (probability for TM7 is ca. 0.6). Since this intervening sequence was consistently present in PCR products amplified from extracted head mRNA, it is possible that this sequence acts to regulate translation and expression/function of *npr-1*. For example, the shortening of the C-terminal tail could affect coupling between the NPR-1 receptor and its associated G protein. However, we did not investigate the functionality of the short isoform of this receptor in this study. Instead, since our goal was deorphanization of the receptor and we did not expect ligand binding to be disrupted in either isoform, we synthesized the coding regions of the longer isoform to create a full-length (554 amino acid) NPR-1 receptor construct for the yeast assay. Our sequence also included several single nucleotide substitutions that resulted in amino acid changes (see Figure 1) from the sequence reported in the *A. suum* transcriptome (JI169959, JI171129): Ile-179 for Val-179; Ala-301 for Arg-301; Gly-310 for Ser-310; Asp-406 for His-406; Tyr-447 for His-447; and Gly-539 for Asp-539.

In *C. elegans*, a single amino acid change (Phe-215 to Val-215) in the NPR-1 receptor has been shown to control social versus solitary feeding behavior [29, 31]. The *A. suum* NPR-1 receptor that we cloned contains phenylalanine at position 215, which was associated with social feeding in *C. elegans*. While we have observed that *A. suum* survive longer in the lab when densely packed in saline solution, this was not a behavior we investigated further in this study.

Our cloned NPR-1 receptor (long isoform) shared 39% identical residues and 53% similar residues with the *C. elegans* receptor. An alignment of these receptors and 9 other nematode

NPR-1 homologs is included in Figure 2. We identified several differences in amino acid sequence between animal parasites, plant parasites, and free-living species, which are indicated by arrows in Figure 2. We also constructed a phylogenetic tree with these sequences using the Neighbor-Joining method in MEGA7 (Figure 3), which did not group the receptors by nematode clade as expected; rather, *Haemonchus contortus* NPR-1 branches off before other nematodes in clade V, with *Ascaris suum* and *Toxocara canis* of clade III and *Strongyloides ratti* of clade IVa clustering more closely to the rest of the clade V nematodes. Similarly, *Brugia malayi* NPR-1 was grouped with *Meloidogyne incognita* (clade IVb) rather than the rest of clade III. The groupings also cannot be explained by lifestyle of the species (animal/human parasite, plant parasite, or free-living). We also constructed a tree using the Maximum Likelihood method to see whether this changed the grouping, but the results were highly similar (tree not shown). This prompted us to construct a phylogenetic tree with all of the nematode sequences discussed in this study to see whether the receptors had been improperly categorized (discussed further in section (h) below, and Figure 22). Indeed, while all of the sequences in question were annotated as “NPR-1”, we found that the *Brugia malayi* sequence clustered with the NPR-4 family, and the *Meloidogyne incognita* sequence clustered with the NPR-10 family. The *Haemonchus contortus* sequence clustered most closely with NPR-2 sequences in *Caenorhabditis* nematodes, though this part of the tree was at least much closer to the rest of the NPR-1 family. Unfortunately, our BLAST searches did not uncover sequences that were a better match for NPR-1 in these nematodes, so we will have to revisit these classifications as additional data becomes available, and take great care in inferring pharmacological or functional information from NPR-1 receptors in other nematodes. This

analysis raises the all-important and ever-present issue of the physiological relevance and importance of sequence relatedness.

(b) npr-2

NPR-1 and NPR-2 are highly related receptors, with 37% identical residues and 53% similar residues in *C. elegans*. The NPR-2 receptor has been implicated in fat storage and locomotion through RNAi [5, 32], but no neuropeptide ligands have been reported so far. Given the strong conservation of receptor sequence, we predicted that this receptor could also bind to the *flp-18* and *flp-21* peptides that bind to NPR-1. These peptides are also homologous to the *afp-1* and *afp-14* peptides found in *A. suum* motorneurons (DE2 and VE2 for *afp-1*, and DE1, DE3, and VE1 for *afp-14* [1]). Preliminary single cell RNA-seq of *A. suum* motorneurons also showed *npr-2* transcript in DE2 and VI (C. Konop, personal communication), suggesting that this receptor could affect motorneuron activity, making it a good candidate for exploration.

The cloned *npr-2* transcript was 1,118 bp (8 exons) long, and encoded a 372-amino acid receptor (Figure 4, MF319188). Our sequence has a single nucleotide substitution (G instead of C) at position 667, resulting in an Glu instead of Gln at position 223 compared to sequence reported in the *A. suum* transcriptome (JI172860). TMHMM analysis predicted 6 transmembrane domains, with the probability of TM7 being only 0.4. The site of TM7 is right at the end of the sequence, suggesting that we might be missing some of the C-terminal tail. Additional 3'-RACE experiments will be needed to extend this sequence out to the stop codon.

The cloned *A. suum* NPR-1 and NPR-2 receptors were highly related, with 31% identical residues and 46% similar residues. The similarity was almost as high as NPR-2 homologs in *A.*

suum and *C. elegans*, which shared 37% identical residues and 61% similar residues. An alignment of these and 7 other nematode NPR-2 homologs is included in Figure 5. We identified several differences in amino acid sequence between animal parasites, plant parasites, and free-living species, which are indicated by arrows in Figure 5. We also constructed a phylogenetic tree with these sequences (Figure 6) as described above. In this case, nematode NPR-2 sequences seemed to correlate with lifestyle of the nematode (animal/human parasite, plant parasite, free-living) rather than evolutionary clade. This suggests that the amino acid differences noted in Figure 5 could be functionally important for different types of nematodes, and could be relevant to deorphanization and drug development efforts, but would require much more information to substantiate.

(c) *npr-4*

The NPR-4 receptor has been assigned a number of putative functions in *C. elegans*, including implication in fat storage, olfaction, foraging behavior, and reproduction [5, 6, 32]. For example, NPR-4 activation in the intestine led to fat accumulation, while activation in RIV and AVA interneurons modulated behavioral responses to odors [32]. It has also been deorphanized, binding most strongly to *C. elegans* -RFamide neuropeptides from the *flp-18* and *flp-4* transcripts, and to a lesser extent, *flp-1* [5, 32]. These peptides are homologous to *A. suum* peptides on the *afp-1*, *afp-2*, and *afp-11* transcripts, respectively, and all these *A. suum* peptides have effects on behavior. Since *afp-1* is expressed by the *A. suum* DE2 and VE2 motorneurons [1], and *afp-2* is expressed by the RME inhibitory motorneurons in the head, this receptor was included in our list of initial GPCR candidates.

The cloned *npr-4* transcript was 1,195 bp (7 exons) long, and encoded a 397-amino acid receptor with 7 putative transmembrane domains (Figure 7, MF319189). Our sequence contains 2 additional exons (exons 5-6) compared to the transcriptomic sequence that was used to design primers (PRJNA80881:GS_14587, [8]). All 7 exons map to the same scaffold of the *A. suum* genome (AMPH01002059.1).

Our cloned NPR-4 receptor shared 57% identical residues and 72% similar residues with the *C. elegans* receptor. This was the most highly conserved receptor identified in this study. We also found homologs in 11 other nematode species, and an alignment is presented in Figure 8. There were few clear differences between parasitic and free-living nematodes, and they are indicated by arrows in Figure 8. We also constructed a phylogenetic tree with these sequences (Figure 9), and found that NPR-4 sequences grouped by nematode clade.

(d) *npr-8*

No information on function or predicted ligand is available for NPR-8 in *C. elegans* or in *A. suum*. However, preliminary single-cell RNA-seq analysis identified *npr-8* transcript in all of the cell types tested so far (DE1, DE2, DE3, VI, and muscle cells; C. Konop, personal communication). The prevalence of this receptor could indicate that it plays a role in context-setting of the locomotory circuit, and makes this receptor a candidate of interest.

The cloned *npr-8* transcript was 1255 bp (7 exons) long, and encodes a 417-amino acid receptor with 6 putative transmembrane domains, plus a weak prediction for TM7 (Figure 10, MF319190). Our sequence is identical to PRJNA80881:GS_23888 [8] from the *A. suum*

transcriptome, which was used to design primers. All 7 exons map to the same scaffold (AMPH01006138.1) of the *A. suum* genome.

The cloned NPR-8 receptor shared 44% identical residues and 60% similar residues with *C. elegans* NPR-8. An alignment of these and homologs from 7 other nematodes is presented in Figure 11. We identified several differences in amino acid sequence between animal parasites and free-living species, which are indicated by arrows in Figure 11. We also constructed a phylogenetic tree (Figure 12), which grouped NPR-8 sequences by nematode clade.

(e) *npr-10*

The function of NPR-10 has not been established in *C. elegans*; however, neuropeptides from the *flp-3* and to a lesser degree, *flp-18* transcripts did activate the receptor in heterologous expression systems [6, 33, 34]. No *flp-3* homologs have been identified in *A. suum* to date, but *flp-18* peptides are homologous to *afp-1* peptides in *A. suum*, which are expressed by the DE2 and VE2 motoneurons [1]. We have also found *npr-10* transcript in DE2 motoneurons by single-cell RNA-seq (C. Konop, personal communication), making NPR-10 a possible candidate for the actions of *afp-1* peptides released from the motoneurons.

The cloned *npr-10* transcript was 1,158 bp (7 exons) long, and encoded a 386-amino acid receptor with 7 putative transmembrane domains (Figure 13, MF319191). Our sequence had two single nucleotide differences that resulted in different amino acids from sequence reported in the *A. suum* transcriptome (PRJNA80881:GS_17528, [8]): A for G at position 394, resulting in an Ile rather than Val at position 132; and C for T at position 764, resulting in Phe

for Leu at position 255. All 7 exons mapped onto the same scaffold in the *A. suum* genome (ANBK01002419.1).

The cloned NPR-10 receptor shared 51% identical residues and 67% similar residues with *C. elegans* NPR-10. An alignment of these and homologs from 7 other nematodes is presented in Figure 14. We identified several differences in amino acid sequence between animal parasites, plant parasites, and free-living species, which are indicated in Figure 14. We also constructed a phylogenetic tree (Figure 15), which grouped NPR-10 sequences by nematode clade.

(f) *npr-11*

In *C. elegans*, the NPR-11 receptor has been shown to affect local searching behaviors, olfactory adaptation, and reproduction [5, 6, 34, 35]. For example, when worms are removed from a food source, they undergo a period of increased activity for the first 15 min, characterized by increased turning rates [6]. This behavior is modulated by a feedback loop between AWC sensory neurons and AIA interneurons in the head [35], in which odor detection by AWC causes it to release NLP-1 peptide. The peptide binds to NPR-11 receptors on AIA neurons, causing release of INS-1 peptide and inhibition of AWC activity. Mutants lacking *nlp-1*, *npr-11*, or *ins-1* display the same increase in turning rate upon removal of food. However, of particular interest to us are the other neuropeptides that bind to NPR-11: *flp-1*, *flp-5*, *flp-14*, *flp-18*, and *flp-21* [5, 6, 34]. These peptides are homologous to peptides on the *A. suum* *afp-11*, *afp-15*, *afp-4*, *afp-1*, and *afp-14* transcripts, and all have been shown to affect locomotory behavior [1, 3, 19, 36, 37]. Furthermore, except for *afp-11*, all of these peptides are produced

by *A. suum* motorneurons: *afp-1* and *afp-4* are expressed by DE2 and VE2, *afp-14* is expressed by DE1, DE3, and VE1, and *afp-15* is expressed by all 4 of the RME inhibitory motorneurons of the head. We also saw *npr-11* transcript in preliminary single-cell RNA-seq analysis of DE2, VI, and muscle cells (C. Konop, personal communication). This suggests that NPR-11 could be an important component of the locomotory circuit in *A. suum*.

The cloned *npr-11* transcript was 1,261 bp (7 exons) long (Figure 16, MF319192). It encoded a receptor that was 419 amino acids in length, with 7 predicted transmembrane domains. Our sequence was identical to one reported in the *A. suum* transcriptome (PRJNA80881:GS_14579, [8]). All 7 exons mapped to a single contig of the *A. suum* genome (ANBK01000995.1).

The cloned NPR-11 receptor shared 54% identical residues and 70% similar residues with *C. elegans* NPR-11. An alignment of these and homologs from 10 other nematodes is presented in Figure 17. We identified several differences in amino acid sequence between animal parasites and free-living species, which are indicated in Figure 17. We also constructed a phylogenetic tree (Figure 18), which grouped NPR-11 sequences by nematode clade.

(g) *frpr-6*

Very little is known about the FRPR-6 receptor in *C. elegans* and *A. suum*; no predicted function, ligand, or expression patterns have been identified. It is a member of the somatostatin class of GPCRs, which are poorly characterized compared to the neuropeptide Y/FF/AF classes that include most of the receptors discussed here. However, preliminary single-cell RNA-seq identified *frpr-6* transcript in DE1 and DE2 motorneurons (C. Konop, personal

communication). This suggests that it could play a role in the locomotory circuit, and we included it in our candidate list.

The cloned *frpr-6* transcript was 1,126 bp (7 exons) long (Figure 19, MF319193). It included a 16 bp stretch of nucleotides at the end of exon 6 that was not observed in the *A. suum* transcriptome (PRJNA80881:GS_11869, [8]), and produced a stop codon prior to exon 7 (though exon 7 was still amplified by PCR). The TMHMM transmembrane domain prediction tool still predicted 7 transmembrane domains with the shorter product, but the length of the C-terminal tail was shortened by 14 residues. The 16 bp intervening sequence was flanked by putative splice sites in the genome (ANBK01002449.1), so it is possible that two versions of this GPCR (355 or 369 amino acids) could be produced by alternative splicing. The C-terminal tail of GPCRs hangs inside the cell, and can be involved in coupling to the associated G protein. Therefore, it is possible that these two isoforms may have different activities inside the cell. Additional studies expressing both isoforms of this GPCRs in yeast will be needed to investigate potential changes in function.

Another interesting structural consideration arose from the TMHMM analysis. When we analyzed the receptor beginning with the first methionine residue (MIIMCMC...), which we presumed to be the initiating methionine, only 6 transmembrane domains were predicted (lacking TM1 in Figure 19b). However, repeating the inquiry with the additional valine residues preceding the methionine produced 7 putative transmembrane domains. To extend the sequence N-terminally, we searched through the *A. suum* genome looking for additional methionine residues preceding exon 1. We found 20 amino acids (LRNISDSELIENASTICVAV) in-frame with the first methionine in exon 1; however, none of these were methionines. Including

these residues in the TMHMM analysis increased the probability of a transmembrane helix at TM1, with the predicted transmembrane domain starting at TICVAVMIIMCMC-. This suggests that we might be missing some N-terminal sequence prior to TM1, or that there are alternative start sites for this gene. Additional experiments will be needed to extend the cloned sequence by 5' RACE.

The cloned FRPR-6 receptor shared 47% identical residues and 64% similar residues with *C. elegans* FRPR-6. An alignment of these and homologs from 8 other nematodes is presented in Figure 20. We identified several differences in amino acid sequence between animal parasites, plant parasites, and free-living species, which are indicated by arrows in Figure 20. We also constructed a phylogenetic tree (Figure 21), which grouped FRPR-6 sequences by nematode clade.

(h) Phylogenetic Analysis of All Nematode GPCRs Discussed Above

Since some of the GPCRs discussed here are highly related in *C. elegans* and *A. suum* (e.g. NPR-1 and NPR-2), we were interested to see how all of the nematode GPCR sequences identified so far were related. We constructed a phylogenetic tree using the Neighbor-Joining method as described above (Figure 22). As expected, we saw the FRPR-6 family break off from the NPY/FF family very early on. Most receptor families grouped nicely, but annotated NPR-1 and NPR-2 sequences from several nematodes (*Brugia malayi*, *Haemonchus contortus*, and *Meloidogyne incognita*) were scattered throughout the NPY family. It is likely that these sequences have been mislabeled, since naming is usually based on sequence conservation, and the GPCRs discussed here usually had 30-50% identity across species rather than conservation

of 70-100% in nematode neuropeptide sequences [1, 2, 38]. However, in the process of conducting BLAST searches for each GPCR, we did not identify a better candidate for NPR-1 or NPR-2 homologs in the nematodes in question. At present, we must conclude that there is a greater degree of variability in these two receptor sequences than typically seen in other GPCR families.

Identification of GPCR Neuropeptide Ligands Using Yeast Growth Assay

In cases where the complete GPCR transcript was identified, we cloned the ORFs into the Cp4258 or Cp1651 plasmids for expression in yeast. Both of these plasmids carry a LEU2 selectable marker, which allows yeast to grow in leucine-deficient medium [16]. Cp1651 differs from Cp4258 in that it does not have an alpha factor signal, which resulted in better expression for some constructs (Jacob Hornick, Broach Lab, Penn State College of Medicine, personal communication). The constructs were then transformed into CY13393, CY13399, or YEX108 *Saccharomyces cerevisiae* strains. These strains are based on the strain used in Wang *et al.* (2006) [16], except that the last 5 amino acids of the endogenous yeast $G\alpha$ have been replaced with those of rat G_i , G_s , or G_q subunits, respectively [12]. These amino acids are also identical to those in *A. suum* and *C. elegans* G_s and G_q [14] (no *A. suum* homolog has been identified yet for G_i). These strains express *HIS3* under control of the *FUS1* promoter such that activation of the candidate GPCR by ligand binding induces transcription of *FUS1-HIS3*, allowing the yeast to grow in histidine-deficient medium (Figure 23).

(a) AsGAR-1

To validate our assay and experimental protocol, we started by attempting to replicate results from Kimber *et al.* (2009) [12], who expressed an *A. suum* muscarinic ACh receptor in yeast to examine its pharmacology. We synthesized the AsGAR-1 ORF and subcloned it into the Cp1651 plasmid for expression in yeast. Like Kimber *et al.*, we saw dose-dependent activation of AsGAR-1 in the YEX108 (G_q) strain in the presence of ACh (Figure 24). Our calculated EC₅₀ was $0.9 \pm 0.3 \mu\text{M}$, which was a full order of magnitude more potent than the $20.3 \pm 1.7 \mu\text{M}$ reported previously [12], possibly as a result of optimizing the codon usage for expression in yeast when we generated the construct. It should also be noted that EC₅₀ concentrations generated using the yeast assay may differ from *in vivo* concentrations by 1-2 orders of magnitude [12, 16, 39], presumably due to the yeast cell wall or an imperfect interaction between the receptor of interest and the yeast signaling system. However, since Kimber *et al.* were unable to express functional AsGAR-1 receptor in CHO, HEK-293, and insect Sf9 cell lines, this decrease in ligand potency may be a necessary concession. Overall, our preliminary results for AsGAR-1 were encouraging, and demonstrated that the changes in experimental protocol and equipment did not prevent us from detecting potential ligands.

We also expanded upon the work described by Kimber *et al.* and looked for potential influence of 52 of the most bioactive neuropeptides from *A. suum* (Table 3). We saw significantly increased growth responses when exposed to peptides from the *afp-1*, *afp-4*, *afp-10*, *afp-11*, *nlp-8*, and *nlp-46* transcripts, which the greatest increase in growth observed in the presence of *afp-4* (Figure 24). We tested 6 concentrations of the *afp-4* peptide (AF2) to investigate dose-response, but have not yet been able to reach the upper plateau. It appears that the EC₅₀ for AF2 will be ca. 2 orders of magnitude greater than that for ACh; however, the

magnitude of the growth response appears to be much greater for AF2 (the ACh response topped out at ca. 130% of control at this time point, while the AF2 response from the same stock culture and experimental plate has surpassed 200% without leveling off. This could suggest that while the EC₅₀ for AF2 is much higher than ACh, a smaller amount of ligand could produce a physiologically relevant response. However, until we determine the *in vivo* concentration of AF2 in the pseudocoelomic fluid, and investigate the behavioral response to lower concentrations of peptide, this suggestion remains speculative.

The *afp-1* and *afp-4* peptides are expressed by the DE2 and VE2 motoneurons [1], and have strong effects on behavior: worms injected with *afp-1* peptides exhibited exaggerated waveforms so strong that these worms were nicknamed “super worms”, while exposure to *afp-4* peptide (AF2) caused jerky movement and quasi-static waves [3, 19, 40]. Interestingly, injection of the *afp-1* peptides did not cause changes in cAMP or cGMP concentrations [3, 19], which would be consistent with a G_q-modulated pathway as proposed by Kimber *et al.* (2009). Conversely, injection of AF2 caused a 175-fold increase in cAMP levels [3, 19], and was also shown to activate a voltage-gated calcium channel in *A. suum* muscle [41], although how AF2 causes this effect remains to be elucidated.

It is also important to note that expression of AsGAR-1 was found in the head and tail of the worm using RT-PCR, but not in the dorsal or ventral body wall muscle [12]. We tested the effect of AF2 on ACh-induced contraction in muscle strips and found that it was greatly increased in the presence of 10 μ M AF2 (60-75% greater than the control response for 30 min after exposure to AF2, $P < 0.001$, Appendix 3), and was accompanied by increases in the frequency and amplitude of spontaneous muscle contractions (after ACh has been washed out

of the chamber). These results, combined with the large increase in cAMP concentrations observed after injection of AF2 in intact worms [3], make it likely that AF2 can act directly on muscle cells to induce contraction, both in the presence and absence of ACh. These studies suggest that binding to the AsGAR-1 receptor is not the only action of this peptide, but it could play a role in potentiating the ACh response in the head or tail of the worm.

(b) NPR-1

We synthesized the long isoform of the *A. suum* NPR-1 receptor transcript and subcloned it into the Cp1651 plasmid for expression in yeast. We expressed the construct in YEX108 (G_q) yeast and screened for growth in the presence of 52 *A. suum* neuropeptides (Figure 25). Significant increases in growth were observed in the presence of peptides from the *afp-1*, *afp-4*, *afp-10*, *afp-14* and *nlp-22* transcripts, with the largest response upon exposure to the *afp-14* peptides. The *afp-1* and *afp-14* results agree with deorphanization of NPR-1 in *C. elegans* (homologs *flp-18* and *flp-21* were ligands) [6, 29].

Our single-cell RNA-seq experiments showed expression of NPR-1 in the DE1 excitatory motorneurons (C. Konop, personal communication). In addition, nearly all of the peptides that bound to NPR-1 in yeast were expressed by the motorneurons: *afp-1* and *afp-4* are expressed by DE2 and VE2, *afp-14* is expressed by DE1, DE3, and VE1, and *nlp-22* is expressed by the DI and VI inhibitory motorneurons [1, 2]. However, a significant response was also observed in the presence of *afp-10* (AF30), which was only expressed in a small subset of neurons in the head (2 pairs of cells in the lateral lines, a single cell in the ventral ganglion, and the RIG neurons in the ventral nerve cord) [42]. No pharmacological or functional data is available for *afp-10*, but the

rest of the peptides that bound to NPR-1 have a variety of effects: *afp-1* peptides caused quasi-static, exaggerated locomotory waveforms, weak depolarization and increased input resistance of DE2 neurons [40], and no change in cAMP or cGMP concentrations [3]; *afp-4* caused a strong depolarization of DE2 with biphasic effects on input resistance, erratic and jerky movements in intact worms [40], and a 175-fold increase in cAMP concentrations [3]; *afp-14* had biphasic effects on membrane potential in DE2 with a 3.5-fold increase in EPSP frequency, pronounced body bends, coils, corkscrews and decreased locomotion in intact worms [40], and decreased cAMP concentrations [3]; and *nlp-22* inhibited muscle contraction and caused flaccid paralysis in intact worms [2]. Given the expression of NPR-1 in the DE1 motorneurons and its ability to bind neuropeptides from all 3 classes of ventral cord motorneurons, perhaps NPR-1 is functioning as a sensor of motorneuron activity in the ventral nerve cord, allowing DE1 neurons to modulate their activity as necessary.

So far, we have only been able to express the NPR-1 construct in the G_q strain of yeast (YEX108). Since we cannot compare deorphanization across strains, it is not yet clear whether the NPR-1 receptor preferentially couples to the G_q subunit. If it can utilize the G_q pathway in *A. suum*, this could be a way for NPR-1 activity to affect ACh release from DE1. In *C. elegans*, diacylglycerol (DAG) interacts with *unc-13* (DAG receptor protein) to facilitate synaptic vesicle fusion at active sites [26, 27]. This would also fit with the lack of effect on cyclic nucleotide concentrations in the presence of *afp-1*. However, *afp-4* and *afp-14* caused significantly increased and decreased changes in cAMP, respectively, so this is clearly not the only action of these peptides.

(c) NPR-4

We subcloned the NPR-4 receptor into the Cp4258 plasmid, and expressed the construct in CY13393 (G_i), CY13399 (G_s), and YEX108 (G_q) yeast. We then screened for binding of 52 *A. suum* neuropeptides (Figure 26). The strongest response was observed in the presence of *afp-4* (AF2, $\approx 120\%$ of the no-agonist control in all 3 strains), and a smaller yet significant response was observed in the presence of *afp-5*, *afp-10*, *afp-13*, and *afp-16* peptides in all strains. We also saw small but significant growth increases for a number of the peptides we tested in 1 strain; however, since the effect is only a $\approx 5\%$ increase compared to control, it remains unclear whether these effects are functionally relevant. Most of the AF peptides share the C-terminal -R_{Fa} sequence (*afp-5* is -H_{Fa}), so perhaps that is enough to produce low-level activation of the NPR-4 receptor.

In addition, the AF17 (*afp-5*, DFDRDFMH_{Fa}amide) and AF19 (*afp-13*, AEGLSSPLIR_{Fa}amide) peptides have very different sequences, but highly similar effects in the worm- both caused hyperpolarization and decreased input resistance of DE2 and DI neurons, decreases in EPSP amplitude, and rapidly abolished locomotion and head searching with flaccid tone in intact animals [40]. In fact, these responses were so similar that Davis & Stretton (2001) suggested that they might be utilizing the same receptor and/or downstream signaling systems in DE2 and VI. While the short length of these peptides has led some to believe that they cannot maintain much three-dimensional structure, perhaps there is a shared secondary structure between these peptides that receptors can recognize.

Interestingly, while a small increase in growth was observed in the presence of *afp-1* in 1 strain (CY13393), and *afp-11* in 1 strain (CY13399), we did not see significantly increased

growth when exposed to *afp-2*. Homologs of all of these peptides are reported as ligands for the NPR-4 receptor in *C. elegans* [6]. Our results suggest that AF2, which shares a [L/I]RFa C-terminus with these peptides, has a much stronger effect on the *A. suum* receptor. While these receptors did cluster in the NPR-4 branch of the phylogenetic tree (Figure 22), they were in separate branches associated with nematode clade. It appears that in this case, the differences in sequence between clade III and clade V nematodes are important for ligand binding.

Since we do not yet have localization data for NPR-4 in *A. suum*, it is difficult to determine what role it might play. However, the AF2-NPR-4 interaction produced one of the largest growth increases of the GPCRs and *A. suum* peptides tested so far. Given the dramatic behavioral effects (discussed above) and sheer abundance of AF2, additional studies are warranted. Interestingly, however, the other putative ligands for which we have functional data were all associated with inhibition of locomotion [40]. While we have not seen the NPR-4 transcript in the motorneurons or muscle tested so far, it could be acting in the head. The *afp-2*, *afp-4*, *afp-15*, *afp-16*, and *nlp-48* peptides are expressed by the RME inhibitory motorneurons in the head, and a member of the *afp-13* transcript was identified by MS/MS of head extract [43] and single cell MS/MS of the ALA neuron in the dorsal ganglion of the head [38]. The NPR-4 receptor could be a mediator of RME and ALA signaling in the head to control searching behavior of the worm.

We also tested 7 concentrations of AF2 to develop a dose response curve (Figure 26c), which placed the EC₅₀ at 37 ± 10 μ M. Since decreases in potency of up to 1-2 orders of magnitude have been reported for the yeast expression system as opposed to other functional expression systems (presumably due to the yeast cell wall or an imperfect interaction between

the receptor and downstream signaling pathway) [12, 16, 39], we predict that this corresponds to an *in vivo* EC₅₀ in the 370 nM – 3.7 μ M range. While we do not yet have dose-response information for AF2 in electrophysiological or behavioral assays, this EC₅₀ is much higher than the IC₅₀ of 8.3 nM that we reported for As-NLP-22 in muscle strip preparations [2], and the EC₅₀ of 57 nM for AF2 in a [³⁵S]GTP γ S binding assay on muscle membrane [44]. However, AF2 is the most abundant *A. suum* neuropeptide [45], and since other neuropeptide receptors have been shown to bind ligands with EC₅₀ values of up to 5 μ M [11], the AF2-NPR-4 pairing could still be biologically relevant.

(d) NPR-11

We synthesized the NPR-11 receptor and subcloned it into the Cp1651 plasmid for expression in CY13393 (G_i), CY13399 (G_s), and YEX108 (G_q) yeast. We then screened for binding of 52 *A. suum* neuropeptides, and found relatively small but significant responses to peptides from the *afp-1*, *afp-4*, *afp-10*, and *afp-11* transcripts, and the largest response was observed for *afp-16* peptides but was strain-specific (Figure 27). The NPR-11 receptor was activated by *afp-16* peptides in CY13393 (G_i) and to a lesser extent, CY13399 (G_s), but in the YEX108 (G_q) strain the response was not significant. Generally, the trend of greater growth in the CY13393 or CY13399 strains held for most of the peptides tested, with the notable exception being AF2 (*afp-4*). The results for *afp-4* and *afp-11* agree with deorphanization data in *C. elegans*, which identified homologs *flp-14* and *flp-1* as well as *flp-5*, *flp-18*, *flp-21*, and *nlp-1* as ligands [6], though in our hands exposure to *afp-1* (*flp-18* homolog) produced a significant increase in

growth in only 1 out of 3 strains, and exposure to *afp-14* and *afp-15* peptides did not significantly increase growth compared to the no-agonist control.

The strain-specific changes in response to the *afp-16* peptides are particularly interesting, and could hint at the importance of differential coupling to G proteins in different tissue types. NPR-11 was notable in that its transcript was found in muscle, as well as DE2 and VI motoneurons (Table 1), which made this the most broadly-expressed receptor that we attempted to express in yeast so far. All of the putative ligands identified here are expressed in the head of the worm, with the two strongest responses produced in the presence of peptides from the inhibitory RME neurons: *afp-4* is expressed in RMEV and RMED, in addition to DE2 and VE2 motoneurons in the ventral cord, while *afp-16* is expressed primarily in RMER and RMEL, with more variable expression in RMEV/D. Since the RME neurons are inhibitory motoneurons that control head searching behavior, perhaps NPR-11 receptor expressed in muscle could mediate/modulate these effects. AF2 (*afp-4*) has been shown to stimulate [³⁵S]GTPγS binding in muscle membranes [44], and cause abrupt jerky, thrashing movements and head tremor with a 175-fold increase in cAMP concentrations upon injection [3]. NPR-11 coupling to the G_s pathway in head muscle could explain these effects.

Another possible function could be mediation of AF2 (*afp-4*) signaling in the body locomotory circuit. AF2 is also expressed in DE2 and VE2 ventral cord motoneurons [1], and since NPR-11 is expressed in muscle, DE2, and VI, there are several potential modes of action. First, NPR-11 in muscle could mediate direct effects of AF2 on muscle contraction in the same manner described above for the head. Second, NPR-11 in the DE2 and VI motoneurons could mediate the AF2-induced changes in membrane potential and input resistance observed by

Davis & Stretton (2001). Exposure to AF2 caused a strong depolarization (the strongest response of all peptides tested) of DE2 neurons that was not Ca^{2+} -dependent, indicating that its receptor is on DE2 neurons. This was accompanied by a biphasic change in input resistance, with an early transient increase followed by a sustained decrease. This could indicate that there are multiple mechanisms of AF2 action at play, perhaps mediated by different receptors or even the same receptor coupled to different $\text{G}\alpha$ subunits. No change in EPSP frequency was observed, suggesting that presynaptic input to DE2 from central spiking neurons was unaffected. Since AF2 is released from DE2 and VE2 neurons, it appears that an autoreceptor on the DE2 membrane is at least partially responsible for these effects. Given the expression of NPR-11 in DE2 and its ability to bind AF2, this is the first receptor identified to date that could mediate these effects.

The input resistance of DI neurons was also affected (increased) by AF2 superfusion; however, this was not accompanied by a change in membrane potential [40]. This suggests that there could be a compensatory mechanism in DI to prevent AF2-related changes in membrane potential, or perhaps that different receptors/downstream signaling pathways mediate these effects. While we do not yet have an electrophysiological analysis of AF2 effects in VI, we have shown that DI and VI have nearly identical resting membrane potentials, input resistance, and cable properties [46], as well as intrinsic oscillatory activity [47]. Further, they have the same neuropeptide expression profile [2]. Given the similarities between these cells, we predict that AF2 could have similar effects in VI, where NPR-11 is expressed (we have not yet analyzed DI neurons by RNA-seq). If this is true, then AF2 binding to NPR-11 could be responsible for the early increase in input resistance in DE2 and VI neurons.

Effects of AF2 α/β -Peptide Analogs on GPCR Binding

Of the *A. suum* GPCRs investigated so far, nearly all have shown activation in the presence of AF2. Since AF2 is the most abundant neuropeptide in the animal [45], it is possible that it could be acting (at least in part) in a context-setting capacity. To determine which amino acids are important for GPCR binding, we designed and synthesized a series of α/β -peptide analogs for AF2. In α/β -peptides, one or more α -amino acids are replaced with a β -amino acid, which is elongated by one $-\text{CH}_2$ group within the polymeric chain. Previous studies have shown that this extra $-\text{CH}_2$ group can confer resistance to proteolysis (e.g. the α/β -Bim BH3 analog was >100-fold more resistant to proteinase K treatment than the parent peptide [17]), and in some cases, improves tissue- and cell-permeability while maintaining the bioactivity of the parent peptide. We began by altering one amino acid at a time (Table 4), scanning the peptide sequence for important residues in GPCR binding and functional activity.

Since the AsGAR-1 and NPR-4 receptors bound AF2 in the yeast growth assay, we also tested the effects of AF2 α/β -peptides on growth of the strains (Figure 28). We found that NPR-4 expressing strains exposed to the AF2-Y, and AF2-F analogs exhibited similar growth to culture exposed to endogenous AF2 peptide ($P < 0.01$ in all 3 strains, Figure 28a), but modification of the K residue produced growth rates comparable to no-agonist control in 2 strains, and a significant but small response in YEX108. When compared to the endogenous AF2 response in NPR-4-expressing strains, a significant decrease was observed for AF2-K (all strains), AF2-E (CY13393, CY13399), and AF2-R (CY13399, YEX108), and a significant increase was observed for AF2-F (all strains). These results suggest that the K, E, and R residues could be important for binding of peptide to the receptor, and that the flexibility introduced by the

additional -CH₂ group is detrimental at these positions. In contrast, modification of the F residue produced a stronger growth response than that observed for endogenous AF2, suggesting that flexibility at the C-terminus allowed the compound to bind better. The same general trend in growth was observed with the AsGAR-1 receptor (Figure 28b), however, when compared to endogenous AF2, only the potentiated growth in the presence of AF2-F was significant. These results suggest that activation of both of these receptors is hindered by modification of the N-terminus and 5th position R, while increased flexibility at the C-terminus promoted activation.

CONCLUSIONS AND FUTURE DIRECTIONS

This work represents the first steps toward neuropeptide GPCR identification and characterization in the parasitic nematode *A. suum*. Here, we have cloned 7 putative GPCRs that are expressed within the locomotory circuit, and in some cases, extended sequences reported in the *A. suum* transcriptome. The gene families for genomic and transcriptomic sequences were not previously reported, so here we identified receptor homologs from other nematode species to determine receptor identity and phylogeny. We found that in several cases, receptors that were annotated as “NPR-1” and “NPR-2” in other nematodes did not cluster with their respective gene families, indicating that these sequences may have been mis-categorized, or at least they are not as strongly conserved. The NPR-2 family was also noteworthy in that the sequences we found clustered better with nematode lifestyle (animal/human parasites versus free-living species) than by evolutionary clade. This could be relevant for future therapeutic development, perhaps affecting ligand specificity and providing

a way to treat parasitic infections without affecting free-living populations. In the cases of NPR-1 and FRPR-6, we also identified potential sites for alternative splicing of the transcripts to produce truncated receptors with shorter C-terminal tails. These isoforms could be functionally important, since the C-terminus of 7-transmembrane receptors hangs inside the cell, and can interact with the associated G protein.

We also report the first attempt to heterologously express and deorphanize neuropeptide GPCRs from a parasitic nematode. Most deorphanization efforts have been confined to *C. elegans* receptors, and the only *A. suum* GPCR studied previously was a muscarinic ACh receptor [12]. Since we have shown that neuropeptide expression differs substantially between *A. suum* and *C. elegans* in the motorneurons [1, 2], we expected to see differences in GPCR expression and/or neuropeptide binding as well. Indeed, the receptors we identified in preliminary single-cell RNA-seq experiments had not been previously localized to the motorneurons in *C. elegans*, and while there was some overlap in ligand binding, we also saw differences in peptide binding using the yeast assay. Although these efforts just scratch the surface of GPCR signaling in *A. suum*, they are an important first step in understanding how neuropeptides can elicit such potent and far-reaching effects on nematode locomotion.

We also found that most of the receptors we were able to express in yeast showed increased activation in the presence of AF2. This is the most abundant neuropeptide in the worm, and has a variety of effects on locomotion, including changes in membrane potential and input resistance of specific motorneurons, increased contraction and opening of ion channels in muscle, increased cAMP concentrations, and jerky, erratic movement of the animal. This variety of effects suggests that AF2 acts on multiple signaling pathways, so it is not entirely surprising

to find binding to so many receptors. However, the relevance of these effects will be determined by the EC_{50} for each receptor and the actual concentration of AF2 in the pseudocoelomic fluid, or in the intercellular spaces between neurons or between neurons and muscle, none of which are known. The EC_{50} concentrations reported in the yeast assay are likely to differ from *in vivo* concentrations by 1-2 orders of magnitude, which is an inherent pitfall of this assay; however, since the *A. suum* muscarinic ACh receptor could not be expressed in mammalian or insect cell lines, we chose to start with an expression system that had worked for our organism. We have also not yet tested for effects of the remaining known neuropeptides in *A. suum*, or the potential interactions between neuropeptides. We attempted to address this in part by testing all neuropeptides from a given transcript at once, since they are likely packaged into dense core vesicles together and co-released from the cell. However, most neurons produce multiple peptides from multiple transcripts, and potential interactions between them have yet to be investigated.

We have also found some interesting differences in peptide binding to the same receptor expressed in yeast strains carrying different G_α subunits, most notably in the case of NPR-11 and the *afp-16* peptides. While the cellular distribution and relative expression levels of the natural G_α subunits are not yet known, it would be interesting to compare expression among different neurons and muscle cells of *A. suum*.

Finally, we attempted to identify the important amino acids for AF2 binding to these receptors by synthesizing a series of α/β peptide analogs. We found that modification of the N-terminal K, the 2nd position E, and the 5th position R significantly reduced growth of NPR-4-expressing yeast compared to endogenous AF2, and modification of the C-terminal F produced

a greater growth response. Additional studies will be needed to test these analogs for effects on muscle contraction and behavior in the worm and confirm that these findings are functionally relevant *in vivo*. It would also be interesting to synthesize additional analogs that vary in stereochemistry (D-amino acid substitutions), as these compounds can also be more resistant to protease activity and have great potential for therapeutic development.

AUTHOR CONTRIBUTIONS

J. J. Knickelbine was responsible for bioinformatics, cloning of GPCR transcripts, GPCR construct design and preparation, growth and transformation of yeast cultures, yeast growth assays, α/β -peptide testing in yeast, and manuscript preparation. Z. Yao was responsible for α/β -peptide synthesis. R. J. Garry assisted with cloning of GPCR transcripts. S. H. Gellman was responsible for α/β -peptide design. A. O. W. Stretton was responsible for bioinformatics, α/β -peptide design, and manuscript preparation.

TABLES AND FIGURES

GPCR	Predicted Function (<i>C. elegans</i>) ¹	Peptide Ligand (<i>C. elegans</i>) ¹	<i>A. suum</i> Peptide Homolog	Peptide Expression in <i>A. suum</i> (MS, ISH, ICC)	GPCR Expression in <i>A. suum</i> (RNA-seq) ²
NPR-1	Feeding behavior, aerotaxis, thermal avoidance, ethanol tolerance, innate immunity	<i>flp-18</i>	<i>afp-1</i>	DE2, VE2	DE1
		<i>flp-21</i>	<i>afp-14</i>	DE1, DE3, VE1	
NPR-2	Locomotion, fat storage	N/A	N/A	N/A	DE2, VI
NPR-4	Foraging behavior, reproduction, fat storage, olfaction	<i>flp-1</i>	<i>afp-11</i>	AVK, ALA (head)	
		<i>flp-4</i>	<i>afp-2</i>	N/A	
		<i>flp-18</i>	<i>afp-1</i>	DE2, VE2	
NPR-8	N/A	N/A	N/A	N/A	DE1, DE2, DE3, VI, muscle
NPR-10	N/A	<i>flp-3</i>	N/A	N/A	DE2
		<i>flp-18</i>	<i>afp-1</i>	DE2, VE2	
NPR-11	Local search behavior, olfactory adaptation, reproduction	<i>flp-1</i>	<i>afp-11</i>	AVK, ALA (head)	DE2, VI, muscle
		<i>flp-5</i>	<i>afp-15</i>	RME (head)	
		<i>flp-14</i>	<i>afp-4</i>	DE2, VE2, RMED/V	
		<i>flp-18</i>	<i>afp-1</i>	DE2, VE2	
		<i>flp-21</i>	<i>afp-14</i>	DE1, DE3, VE1	
		<i>nlp-1</i>	<i>nlp-1</i>	N/A	
FRPR-6	N/A	N/A	N/A	N/A	DE1, DE2

Table 1. Candidate GPCRs as determined by ligand homology and single-cell RNA-seq. ¹GPCR function and ligands in *C. elegans* are reviewed in Frooninckx *et al* (2012). ²Preliminary single-cell RNA-seq performed by Christopher J. Konop.

GPCR	Category	Primer	Sequence
NPR-1	Neuropeptide FF Receptor	NPR1-3RACE	ATGAAGCGAATGCTAAATTATGGCACAACCAACGG
		NPR1-3RACE2	GTGTATACTTTCGGTGCTGTAACCTCCGTAACG
		NPR1-3RACE3	CGTCATCATCTACTGAAACAGACACGACGTACGAC
		NPR1-ExpF	GCCATACCATGGAGCGAATGCTAAATTATGGCACAACC
		NPR1-5RACE	CTAAACAAAGATATCCGTATCGTGAACGGTCAACATCGTAGCG
		NPR1-5RACE2	GTCGTACGTCGTGTCTGTTTCAGTAGATGATGACG
		NPR1-5RACE3	CGTTACGGGAGTTACAGGCACCGAAAGTATACAC
		NPR1-ExpR	GCCATATCTAGACTAAACAAAGATATCCGTATCGTGAACG
NPR-2	Neuropeptide Y Receptor	NPR2-3RACE	ATGGGAATGACTATCGCAACAACAATAACGCATTATCG
		NPR2-3RACE2	GGAACGTTCACTCTTTGTGCTATTTCAATCGATAGGTATC
		NPR2-ExpF	GCCATAGCCGGCATGGGAATGACTATCGC
		NPR2-5RACE	CGGTGGAAAAAGTATGAATGAAAGTCGTTCTGTAATGTG
		NPR2-5RACE2	GATACCTATCGATTGAAATAGCACAAAGACTGAACGTTCC
		NPR2-ExpR	GCCATATCTAGACAATTGAATGAAGCACAATAAATATCACAGTATATTTG
NPR-4	Neuropeptide FF/Neuropeptide AF Receptor	NPR4-3RACE	ATGAGCAATGAATCATGTTTGGATATGAACGATGAATTATGG
		NPR4-3RACE2	CACATGCACCTATTCTAATTGCGGGGATCTGTGC
		NPR4-ExpF	GCCATACCATGGGCAATGAATCATGTTTGGATATGAACG
		NPR4-5RACE	TCAAGAACCGCAGCTCGCGCAACTCAGCTCTGTGAC
		NPR4-5RACE2	GCACAGATCCCGCAATTAGAATAAGTGCATGTG
NPR-8	Apelin Receptor	NPR8-3RACE	ATGGATTTATCTTGAGGAGGTACAGTGAATGGGAGG
		NPR8-3RACE2	GGGTATGGAAGACCCCAATGAAGCTGCC
		NPR8-5RACE	TCAACGGAAGTTCACATATTGTGAGATGAATGAAATAAATACG
		NPR8-5RACE2	GGCAGCTTCATTGGGGTCTTCCATACCC
NPR-10	Neuropeptide Y Receptor	NPR10-3RACE	ATGAGCAATACAACGCCAGAGTGCATCGACATAAAGG
		NPR10-3RACE2	GGTACATGCGTTATGGATGATTGCTTTGAATATTACACTCGCGAC
		NPR10-ExpF	GCCATACCATGGGCAATACAACGCCAGAGTGC
		NPR10-5RACE	CTATAATGAACTGGATACGGCCGTGTAATCACCATGC
		NPR10-5RACE2	GTCGCGAGTGTAATTCAAAGCAATCATCCATAACGCATGTACC
		NPR10-ExpR	GCCATATCTAGACTATAATGAACTGGATACGGCCGTGTAATC
NPR-11	Neuropeptide Y Receptor	NPR11-3RACE	ATGCCAAACAGTACGCTAGCAGAAAAAGACTG
		NPR11-3RACE2	GGCTTATTTCAATTGCACTCGCGTTCCG
		NPR11-ExpF	GCCATACCATGGCAAACAGTACGCTAGCAGAAAAAG
		NPR11-5RACE	TCAAACAGTGTTACCATATTAATCCTTAAGATAAACTCGCTC
		NPR11-5RACE2	CGGAAGCGCGAGTGCAATTGAAATAAGC
		NPR11-ExpR	GCCATATCTAGATCAAACAGTGTTACCATATTAATCCTTAAGATAAAC
FRPR-6	Somatostatin Receptor	FRPR6-3RACE	GCGTTGTTGTAATGATTATAATGTGTATGTGTGGTGTAGTTGG
		FRPR6-3RACE2	GAGGCTGGTGTGCAACGCCGTG
		FRPR6-5RACE	TCACCATATACGTTTGAGAGATGGGATGGGGGAC
		FRPR6-5RACE2	CACGGCGTTGCACACCAGCCTC

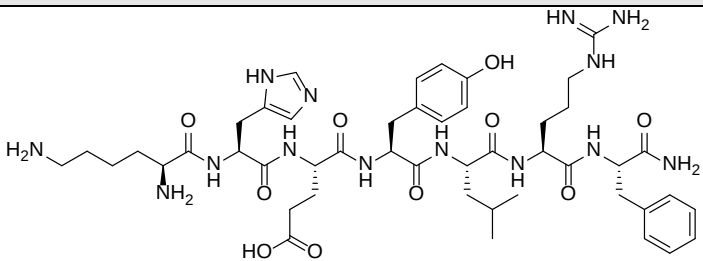
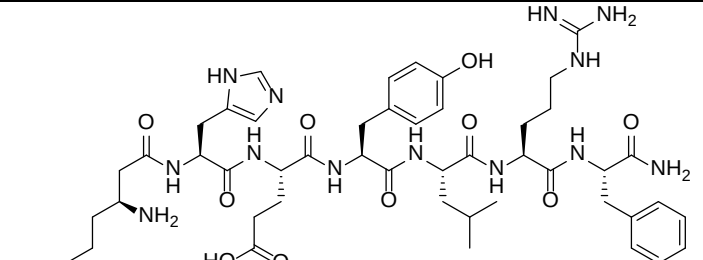
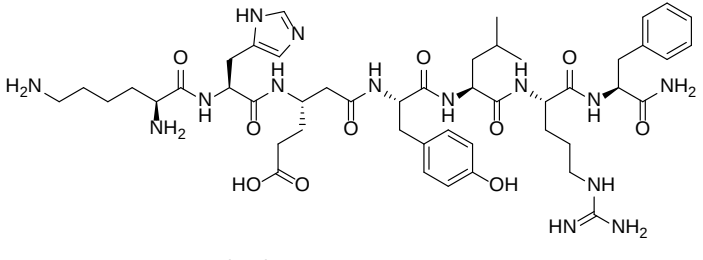
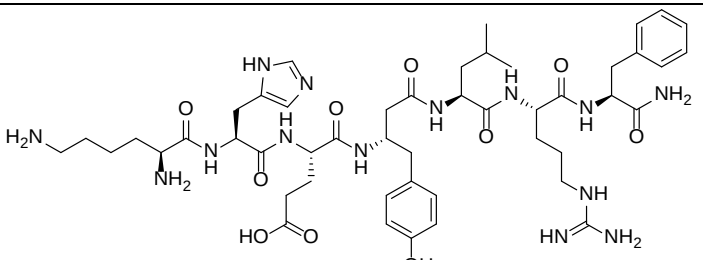
Table 2. Primer sequences. Primers labeled “GENE-ExpF/R” were used to generate full length GPCR inserts for expression in yeast. Restriction enzyme cut sites (NcoI, XbaI) are underlined.

Transcript	Peptide	Sequence	Mass	Tested	<i>C. elegans</i> homolog
<i>afp-1</i>	AF3	AVPGVLRFa	857.5	X	<i>flp-18</i>
	AF4	GDVPGVLRFa	958.5	X	
	AF10	GFGDEMSMPGVLRFa	1541.7	X	
	AF13	SDMPGVLRFa	1020.5	X	
	AF14	SMPGVLRFa	905.5	X	
	AF20	GMPGVLRFa	875.5	X	
<i>afp-2</i>	AF5	SGKPTFIRFa	1051.6	X	<i>flp-4</i>
<i>afp-3</i>	AF8	KSAYMRFa	901.5	X	<i>flp-6</i>
<i>afp-4</i>	AF2	KHEYLRFa	991.5	X	<i>flp-14</i>
<i>afp-5</i>	AF17	DFDRDFMHFa	1228.5	X	
<i>afp-6</i>	AF21	AMRNALVRFa	1076.6		<i>flp-11</i>
	AF22	NGAPQPFVRFa	1131.6		
	AF23	SGMRNALVRFa	1149.6		
<i>afp-7</i>	AF1	KNEFIRFa	952.5	X	<i>flp-8</i>
<i>afp-8</i>	AF24	RNKFEFIRFa	1255.7	X	<i>flp-12</i>
<i>afp-9</i>	AF15	AQTFVRFa	867.5		<i>flp-16</i>
<i>afp-10</i>	AF16	ILMRFa	678.4		<i>flp-28</i>
	AF30	APNKILMRFa	1088.6	X	
<i>afp-11</i>	AF11	SDIGISEPNFLRFa	1493.8	X	<i>flp-1</i>
	AF25	NNFLRFa	809.4	X	
	AF26	KPNFLRFa	920.5	X	
	AF27	PADPNFLRFa	1075.6	X	
	AF28	SAEPNFLRFa	1079.6		
	AF29	NAEPNFLRFa	1106.6	X	
	AF31	TPSNNFLRFa	1094.6	X	
	AF32	GSDPNFLRFa	1051.5	X	
	AF33	SNQAQNFLRFa	1223.6	X	
	AF41	KPNFIRFa	920.5	X	
	PepTF	TAPPLTFa	745.4	X	
	PepNY	NDQFDREY	1086.4	X	
<i>afp-12</i>	AF36	VPSAADMIRFa	1236.6	X	<i>flp-24</i>
<i>afp-13</i>	AF19	A EGLSSPLIRFa	1188.7	X	<i>flp-13</i>
	AF34	DSKLMDPLIRFa	1333.7	X	
	AF35	DPQQRIVTDETVLRFa	1815.9	X	
	PepGE	GNSYRSFFNPNEYTVVE	2022.9		
	PepTL	TNIMGENRLNRNL	1544.8	X	
	PepTT	TPPEEDLLGRFT	1374.7	X	
<i>afp-14</i>	AF9	GLGPRPLRFa	1011.6	X	<i>flp-21</i>
<i>afp-15</i>	AF7	AGPRFIRFa	962.6	X	<i>flp-5</i>
	AF39	SPKQKFIRFa	1149.7	X	
	AF40	AGSARFIRFa	1023.6	X	

Transcript	Peptide	Sequence	Mass	Tested	<i>C. elegans</i> homolog
<i>afp-16</i>	AF42	SFPNNYLASDLALRFa	1726.9	X	<i>flp-26</i>
	AF43	RIDINDLALRFa	1344.8	X	
	AF44	SSYSFDPSNLRFa	1645.8	X	
	AF45	TFVIPTDLALRFa	1391.8	X	
	AF45*	SFGGMLEGFDEKRTFVIPTDLALRFa	2845.5		
	PepQV	pQEEEEEDGVLPFRV	1658.8	X	
?	AF37	FRGEPIRFa	1020.6		<i>flp-2</i>
	AF38	AQREPIRFa	1015.5		
<i>As-nlp-2</i>	As-NLP-2.1	SMALGRLAFRPa	1217.5		<i>nlp-2</i>
	As-NLP-2.2	SLALGRVDFRPa	1229.5	X	
	As-NLP-2.3	NVDIADKDIEYLMEDSMa	2000.2	X	
	As-NLP-2.4	SAAFGRFHFRPa	1291.5		
	As-NLP-2.5	SLALGRSGFRPa	1159.4		
	As-NLP-2.6	SLALGRVGFRPa	1171.4		
<i>As-nlp-7</i>	As-NLP-7	YDAAFDDPRFLSLAFa	1746.8		<i>nlp-7</i>
	As-NLP-7	FADEWDDPRFLSSAFa	1801.8		
	As-NLP-7	SGMMEFDDPRFLSSAFa	1835.8		
	As-NLP-7	NHLIGFDDPRLFSSSYa	1866.9		
<i>As-nlp-8</i>	As-NLP-8.1	AFDRLDESALGLF	1453.7		<i>nlp-8</i>
	As-NLP-8.2	SFDRIDGSAFGPH	1405.6		
	As-NLP-8.2*	SFDRIDGSAFGPHRH	1698.8		
	As-NLP-8.3	AFDRIDGSDFGLV	1411.7		
	As-NLP-8.4	AFDRIEGAGFGLS	1339.7		
	As-NLP-8.5	AFDRIEGSGFGLD	1383.7		
	As-NLP-8.6	AFDRLEGSDFGVL	1425.7		
	As-NLP-8.7	AFDRLEGSDFGLD	1441.7	X	
	As-NLP-8.8	AFDRVGESNFGFD	1460.6	X	
<i>As-nlp-9</i>	As-NLP-9.1	AGARSFGPRIYDVESYLYPTN	2376.2		<i>nlp-9</i>
	As-NLP-9.2	FGSFSPYYFYQQP	1630.7		
	As-NLP-9.3	GGGRTFASYWIPPSGNGRFYDDSPFF	2898.3		
	As-NLP-9.4	SYTSDDYF	1316.5		
<i>As-nlp-12</i>	As-NLP-12	DYRPLQFa	937.5		<i>nlp-12</i>
<i>As-nlp-13</i>	As-NLP-13	DFDRDFMHFa	1228.5		<i>nlp-13</i>
<i>As-nlp-14</i>	As-NLP-14	ALDGLGDFGFGFT	1298.6		<i>nlp-14</i>
	As-NLP-14	ALDSLEGTDFFa	1328.6		
<i>As-nlp-21</i>	As-NLP-21.1	GGGRPFTLVD	1018.5	X	<i>nlp-21</i>
	As-NLP-21.2	AGGRFFKIEDVDALE	1666.8		
	As-NLP-21.3	GGGRPFTYTQDE	1327.6	X	
	As-NLP-21.4	GGGRSFRLLSLGE	1322.7	X	
	As-NLP-21.5	AGARSFPVLEDLLE	1516.8	X	
	As-NLP-21.6	RIYDAFSRFP	1271.6	X	
<i>As-nlp-22</i>	As-NLP-22	SLASGRWGLRPa	1198.4	X	<i>nlp-22</i>

Transcript	Peptide	Sequence	Mass	Tested	<i>C. elegans</i> homolog
<i>As-nlp-23</i>	As-NLP-23.1	NLAVGRSSLRPSa	1255.5		<i>nlp-23</i>
	As-NLP-23.2	SIALGRFSLRPa	1215.5		
<i>As-nlp-37</i>	As-NLP-37	NNAEVVNHILKNFGALDRLGDVa	2407.3		<i>nlp-37</i>
<i>As-nlp-38</i>	As-NLP-38	GHSWQTANGLWa	1255.6		<i>nlp-38</i>
<i>As-nlp-43</i>	As-NLP-43.1	KFYAWAa	784.4		<i>nlp-43</i>
	As-NLP-43.2	KFYQWAa	841.4		
<i>As-nlp-46</i>	As-NLP-46	NIAIGRGDGFRPa	1271.5	X	<i>nlp-46</i>
<i>As-nlp-48</i>	As-NLP-48.1	AALLSRYa	792.6	X	<i>nlp-48</i>
	As-NLP-48.2	AVLSRYa	707.5	X	

Table 3. *A. suum* neuropeptide sequences. Peptides screened against GPCRs in yeast are marked with “X”. The remainder are ongoing.

Name	Sequence	Structure	Purity
AF2	KHEYLRF-amide	 Molecular Weight: 991.17	99.8%
AF2-K	K(beta-3)HEYLRF-amide	 Molecular Weight: 1005.19	80%
AF2-E	KHE(beta-3)YLRF-amide	 Molecular Weight: 1005.19	96.7%
AF2-Y	KHEY(beta-3)LRF-amide	 Molecular Weight: 1005.19	97.5%

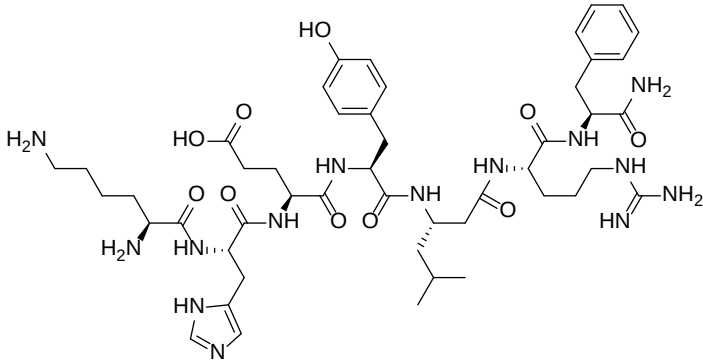
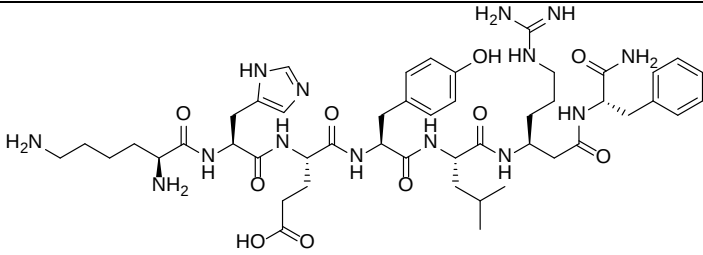
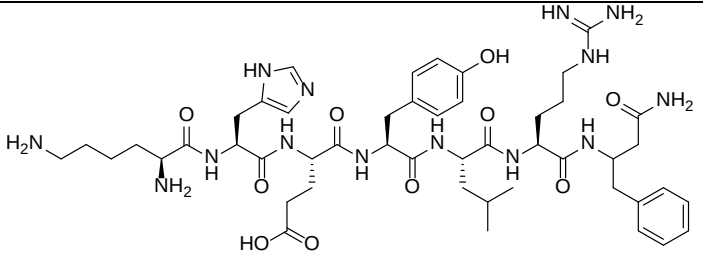
Name	Sequence	Structure	Purity
AF2-L	KHEYL(beta-3)RF-amide	 Molecular Weight: 1005.19	97.4%
AF2-R	KHEYLR(beta-3)F-amide	 Molecular Weight: 1005.19	93.2%
AF2-F	KHEYLRF(beta-3)-amide	 Molecular Weight: 1005.19	91.2%

Table 4. Chemical structures of the endogenous AF2 peptide and α/β -peptide analogs used in these experiments.

(a) *npr-1* transcript

atgaagcgaatgctaaattatggcacaaccaaacggatgaaaacgattgtcgacattacg
 M K R M L N Y G T T K R M K T I V D I T
 tttgcatcgatctgcaaacgtgtaataaaacgacgtatatcgacacgatcgatcgttggct
 F A S I C K R V I K R R I S H D R S L A
 ctgcttcagttaccgcaaatatcgctgctccgctgacaataatgtcatcgacacaac
 L A S V T R K Y R C A P L T I M S S H N
 ggatgcgacgtttatttcgcaactttatccggatccgctcgagtagtccctctagctattgga
 G C D V Y S Q L Y P D P S S S P L A I G
 acattcacgatttctctatgcgcttattcttctgtatgggtattatcggcaattcagcgctc
 T F T I L Y A L I F L M G I I G N S A L
 atttatgtcactctgcaaaatcgacactacaacacgtacagaatatgttcattttaaat
 I Y V T L Q N R T L Q T V Q N M F I L N
 ctggctgctagtgtatattgtaattgtgtatactttcgggtgctgttaactcccgtaacgaat
 L A A S D I V M C I L S V P V T P V T N
 atttataaaaattgggttttttgggtgcaatattatgtcgtcttattccatatgtacaagga
 I Y K N W F F G A I L C R L I P Y V Q G
 gtttcgggttttcacgtacgttcagtccttgggtgcaatcgacttgatcgatataattg
 V S V F T I S T F S L G A I A L D R Y I L
 gtgggtacatccacatacacgctctctaagcaaacatggcgctatgactgttactacgatt
 V V H P H T R S L S K H G A M T V T T I
 ctatggacgttatccatatattgttacattaccctatgcatttttcacgactattgaatct
 L W T L S I F V T L P Y A F E M T I E S
 tatccgggggtatgcggagagttctgtacagagaaatggccgaatgcgtcaagtcgacgt
 Y P G V C G E F C T E K W P N A S S R R
 gcttacacaatggctgctgtagtggcacaattcgtgattccggttattgtgatggcattt
 A Y T M V V L V A Q F V I P F I V M A F
 tgctatgcgggtgatatttgcacgactccgtactagggcacgtttacgactacaacgtatc
 C Y A V I F A R L R T R A R L R L Q R I
 aacgaacgctgctcatttgtgctcgagaaaaccaactcaccatccagcatgcgtcatcct
 N E R C S F V L E K T N S P S S M R H P
 gcttttaaggctgtcagcagtgagatcggtgacatttgtatgacgcaagagaaagtctgt
 A F K A V S S E I G D I C M T Q E K V R
 catcatctactgaaacagacacgacgtacgacaatgattcttgaagtatgggtgctttt
 H H L L K Q T R R T T M I L V S M V L F
 ttgcgtctcacctgggtgccacacaacatcgctctccctaatacatcgaatacgcgaatcg
 F A L T W L P H N I V S L I I E Y D E S
 aaccgattttttcgcacgttcggcatcgacaatatgagcttatcttatctcattaacctc
 N R F F T T F G I D N M S L S Y L I N L
 ttcacacatagtgtggcaatgacaaataatatcgcaaatccgggtattgtacgcttggctc
 F T H S V A M T N N I A N P V L Y A W L
 aatccgacatttctgtgatttgggtacatacaaacatgttgtagaagatatcaacctattcaa
 N P T F R D L V I Q T C C R R Y Q P I Q
 cgcaacaatgtgatttccttgaatatctttgcgaaatgaaatgtgatgcgaatacgtcaaaa
 R N N V Y S L K S L R N E M -
 K C D A N T S K
 atttcgaacaatgctgcaatcaattgtattgggtgaaatttcttatgatagtcaccaagaa
 I S N N A A I N C I G E I S Y D S P Q E
 attggacgtattccacagttgaaaacagcaacacatgaggacaaaattcattcgactat
 I G R I P Q L K T A T H E D K I H S H Y
 cgtcaaaagtatgacgatgaggatgtctcatgtgctgatgaatcgaacgaggcgaaacat
 R Q K Y D D E V S C A D E S N E A K H
 ccgaatgttctacatcgctatcgagttccaacaagtcacaaaaacgacggcaacactgaa
 P N V L H R Y R V P T S H Q N D G N T E
 aacccaaatgagaatgctacggcttattccaccaataacatgatgcttatttcgaggata
 N P N E N A T A L F T N N M M L I S R I
 tcagaaaaaggcaagttggccaagacgctacgatgttgaccgttcacgatacggatatac
 S E K G K V G Q D A T M L T V H D T D I
 tttgttttag
 F V -

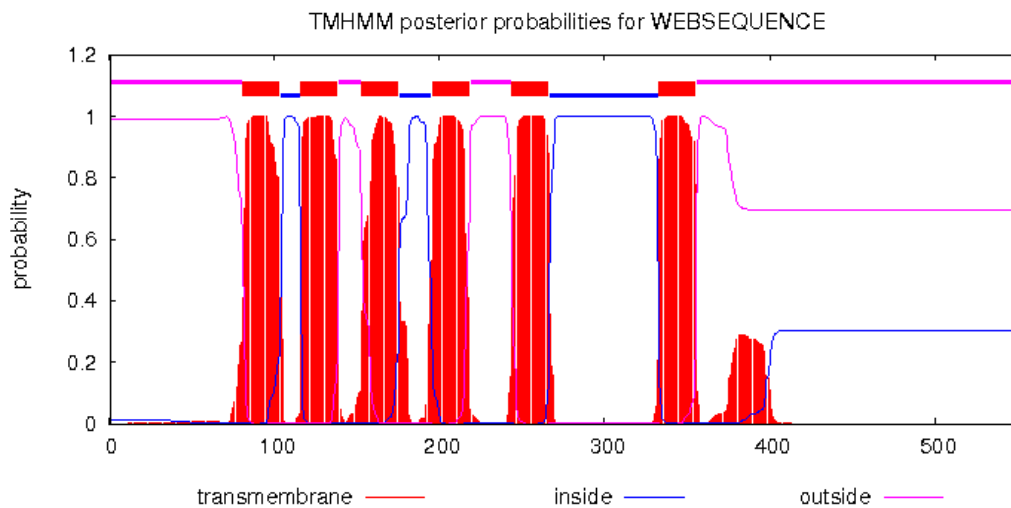
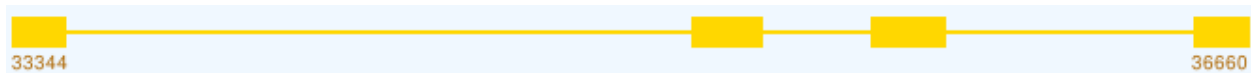
(b) Transmembrane domain prediction**(c) Gene model for exons 1-4 (AEUI02001223.1)****(d) Gene model for exons 5-8 (AEUI02000547.1)**

Figure 1. *npr-1* transcript. (a) Cloned *npr-1* transcript (MF319187) contains a 25 bp insertion (bold, underlined), which shifts the reading frame resulting in a premature stop codon (gray, italicized). This insertion is adjacent to exon 7 in the genome, and contains putative splice sites. The cloned transcript also contains 7 nucleotide changes (bolded) that code for different amino acids (bolded) than those reported in the *A. suum* transcriptome (JI169959). (b) TMHMM analysis predicts that NPR-1 contains 6 transmembrane domains. (c-d) The *npr-1* gene contains 8 exons split between two scaffolds (AEUI02001223.1 and AEUI02000547.1) in the *A. suum* genome.

Figure 2. Comparison of NPR-1 receptor homologs in nematodes. *Ace* = *Ancylostoma ceylanicum* (EYB97833.1), *Asu* = *Ascaris suum*, *Bma* = *Brugia malayi* (XM_001897956.1), *Cbn* = *Caenorhabditis brenneri* (EGT29983.1), *Cbr* = *Caenorhabditis briggsae* (XM_002644546.1), *Cel* = *Caenorhabditis elegans* (NM_076415.4), *Cre* = *Caenorhabditis remanei* (XM_003117529.1), *Hco* = *Haemonchus contortus* (CDJ80286.1), *Min* = *Meloidogyne incognita* (FJ456886.1), *Sra* = *Strongyloides ratti* (CEF71417.1), *Tca* = *Toxocara canis* (KHN84095.1). Free-living nematodes are shaded in yellow, plant parasites in green, and human/animal parasites in orange. Red arrows indicate residues that differ between groups.

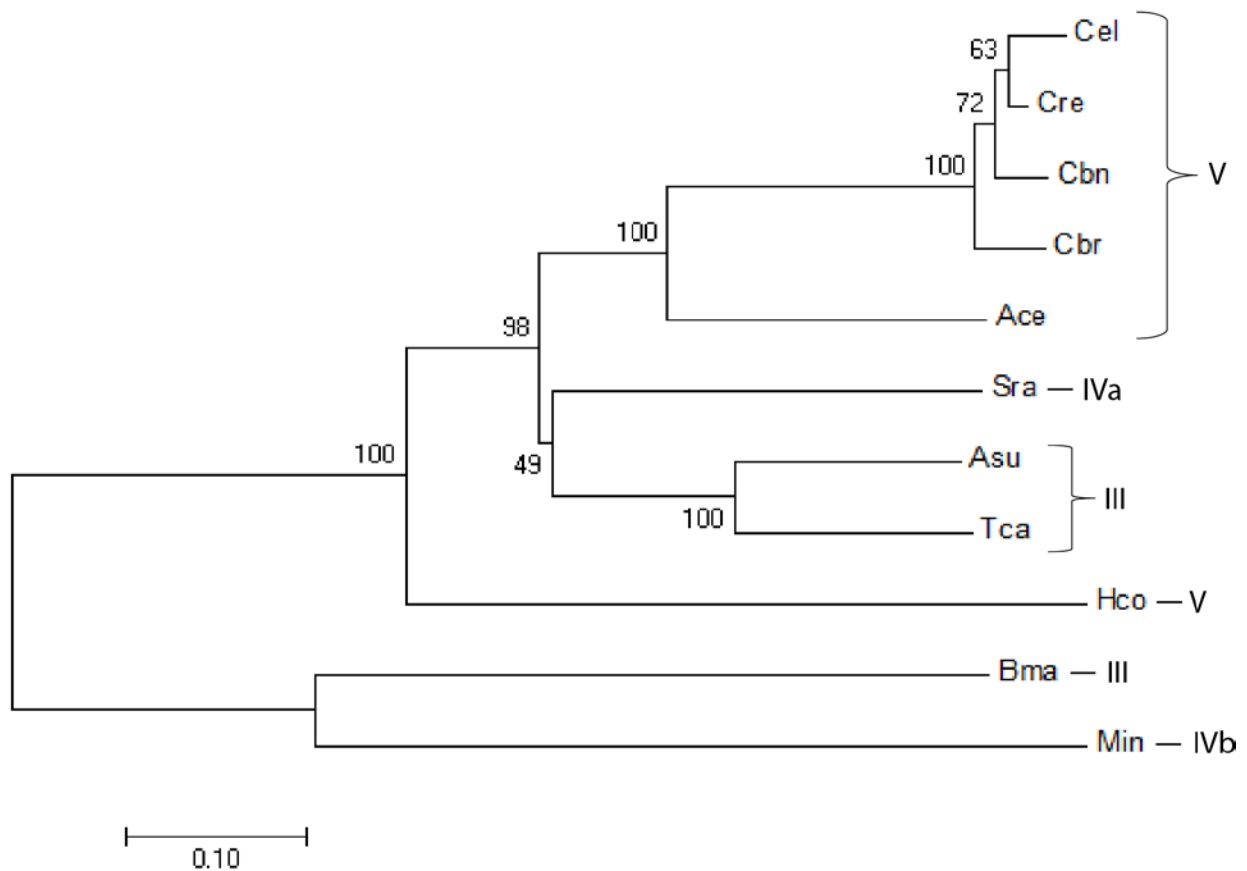
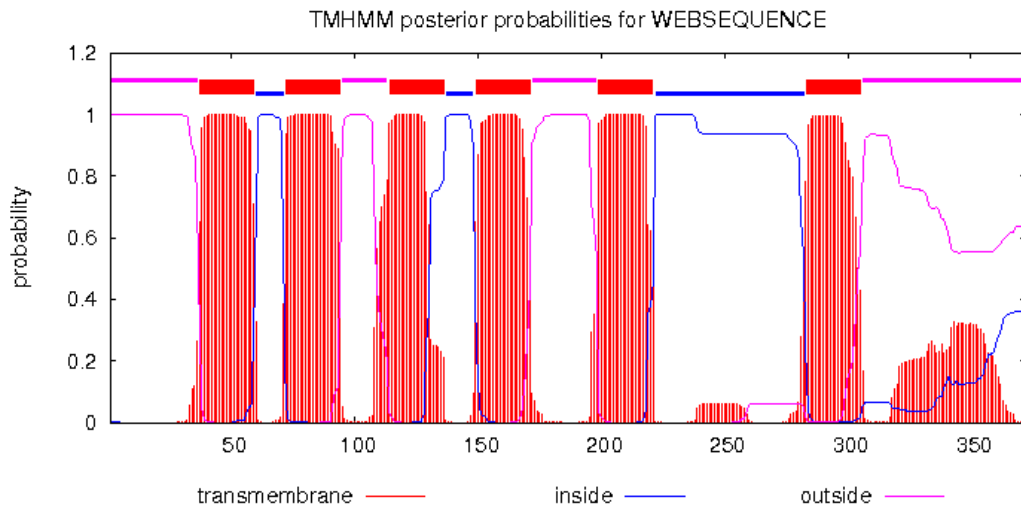


Figure 3. Phylogenetic tree of NPR-1 receptor homologs in nematodes. Nematode sequences and accession numbers are the same as reported in Figure 2, and clade divisions are indicated in Roman numerals. Tree was constructed using the Neighbor-Joining method [23]. The optimal tree with the sum of branch length = 2.79622725 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [48]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [24] and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA7 [21].

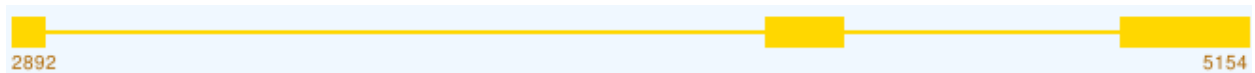
(a) *npr-2* transcript

atgggaatgactatcgcaacaacaactaacgcattatcgaacgatgattgcatcgtatat
 M G M T I A T T T N A L S N D D C I V Y
 atagaggagtttcaagatccatcgatgaaactcgaagtgatcgcttcatttgcactcatc
 I E E F Q D P S M K L E V I A S F A L I
 tattcgggtgatttttgtgcttggcgttatcggaataactttcgttatattttcttacattg
 Y S V I F V L G V I G N T F V I F L T L
 aggcacgcgaaacttcaaacagtacagaatatgttcattttgaatttggctggtgcagat
 R H R K L Q T V Q N M F I L N L A G A D
 cttgtcgtatgtatattttccacttccaattacaccgatcacaaatgtctataagaattgg
 L V V C I F S L P I T P I T N V Y K N W
 tacttcgggtcgacaatgtgccatggattaccgtggatgcaaggcattgcagtttttatt
 Y F G S T M C H G L P W M Q G I A V F I
 ggaacgttcagtcctttgtgctatttcaatcgataggtatctcatgggttactactaccgaca
 G T F S L C A I S I D R Y L M V T L P T
 aagaaaataatcgacaaaactcgctcaatatggattactgcgatcctatggacactttcg
 K K I I D K T R S I W I T A I L W T L S
 attatttttactctaccgtatgcaatgtacatggatggttattgttgttagacgggtatatgt
 I I F T L P Y A M Y M D V I V V D G I C
 ggacaattttgcacggaaaaatggcctgattctggtttaaggcgtagatatacgatattc
 G Q F C T E K W P D S G L R R T Y T I F
 gtttttattgtacaatttattataaccattaataataatgtttatatgttactataaaatt
 V F I V Q F I I P L I I M F I C Y Y K I
 ttcgatgaactcaagaaaaggacaactaaaaaatcaaacaataagagaaagggtcggtg
 F D E L K K R T T K K I K Q I R E R S V
 gttttggcatctgcacttgccttgacacctaatacaagttgcgagtcatgcgctggcagag
 V L A S A L A C T P N Q V A S H A L A E
 aaagataccaaggaatttcaggatcgacttagatgtgagctagttgcgaatgcacgtcgg
 K D T K E F Q D R L R C E L V A N A R R
 aatacaattttattagtgagtatgggtggtgatgtttgcatcgcatggttaccacacaac
 N T I L L V S M V V M F A I A W L P H N
 attgtctcacttttgatggaattcgatgagcgaatattcgaaagggatgacggcggttaat
 I V S L L M E F D E R I F E R D D G V N
 attattttatctctttaatctttttacgcatacggttgcaatgacaaataacgtatcgaat
 I I Y L F N L F T H S V A M T N N V S N
 ccaatttttatatgcattattgaatccgttattttattgagttaatcgccgaatcgattaca
 P I L Y A L L N P L F I E L I A E S I T
 catttacgaacgactttcattcatacttttttccaccg
 H L R T T F I H T F F H

(b) Transmembrane domain prediction



(c) Gene model for exons 1-3 (AMPH01007011.1)



(d) Gene model for exons 4-8 (AEUI02000255.1)

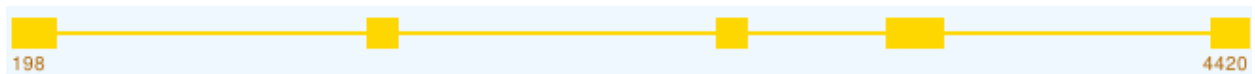


Figure 4. *npr-2* transcript. (a) Cloned *npr-2* transcript (MF319188). Single nucleotide substitution (G for C at position 667) compared to transcriptome (JI172860) is indicated in bold. (b) TMHMM predicts 6 transmembrane domains. (c-d) Gene models for exons 1-3 and exons 4-8, which were found in separate scaffolds of the *A. suum* genome (AMPH01007011.1 and AEUI02000255.1, respectively).

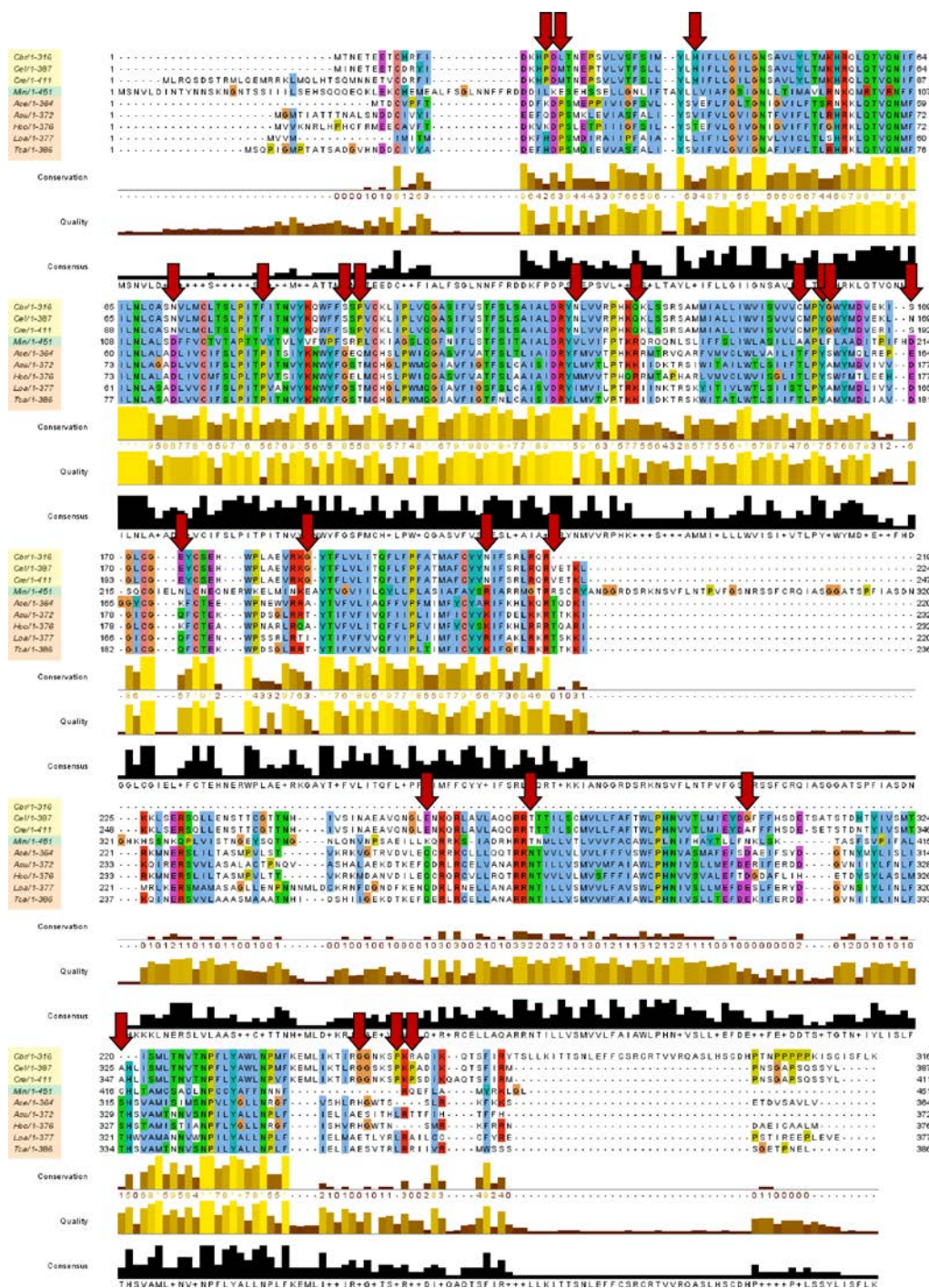


Figure 5. Comparison of NPR-2 homologs in nematodes. *Ace* = *Ancylostoma ceylanicum* (EPB72771.1 7), *Asu* = *Ascaris suum*, *Cbr* = *Caenorhabditis briggsae* (XM_002633911.1), *Cel* = *Caenorhabditis elegans* (NM_182289.3), *Cre* = *Caenorhabditis remanei* (XM_003093586.1), *Hco* = *Haemonchus contortus* (CDJ82824.1), *Loa* = *Loa loa* (XP_020303059.1), *Min* = *Meloidogyne incognita* (FJ456887.1), *Tca* = *Toxocara canis* (KHN76374.1). Free-living nematodes are shaded in yellow, plant parasites in green, and human/animal parasites in orange. Red arrows indicate residues that differ between groups.

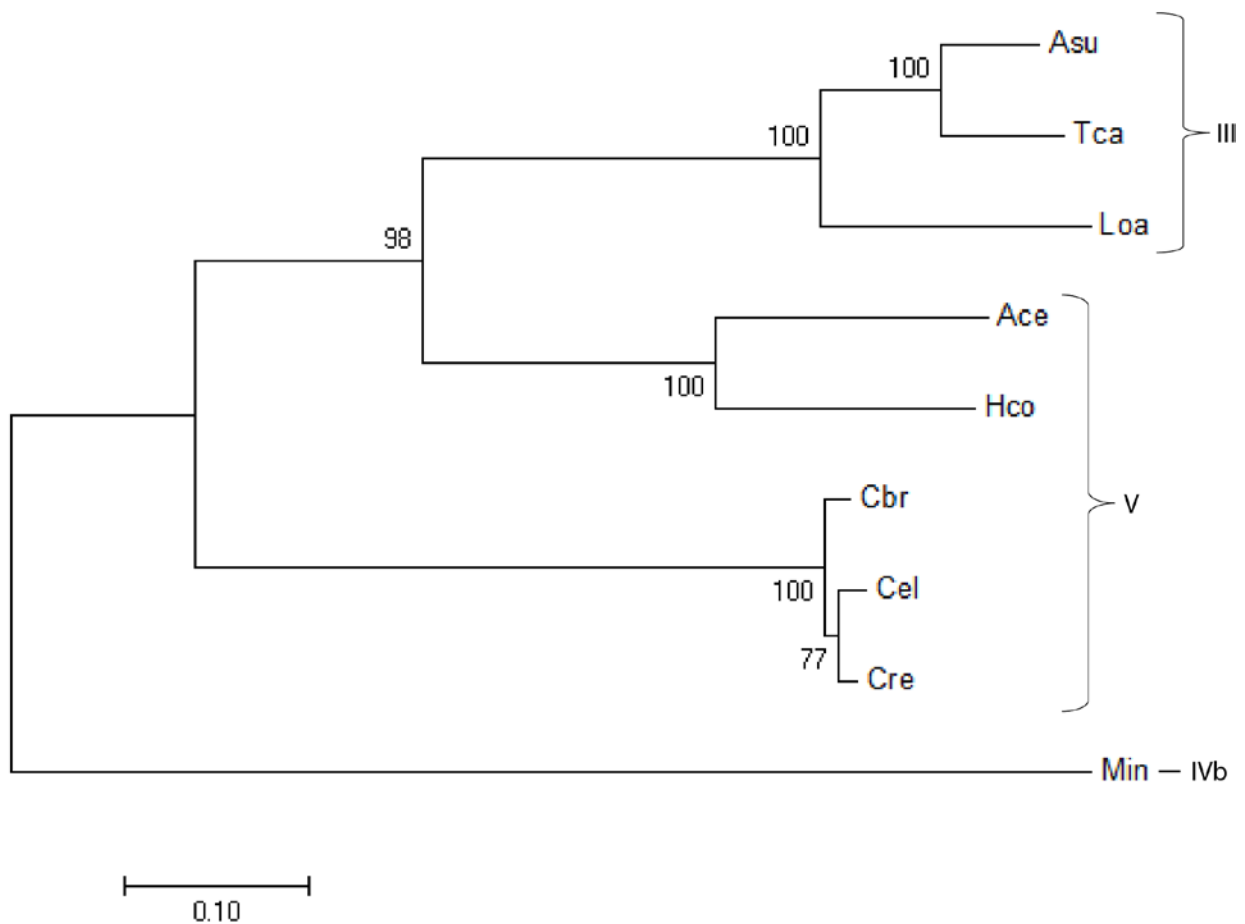


Figure 6. Phylogenetic tree of NPR-2 receptor homologs in nematodes. Nematode sequences and accession numbers are the same as reported in Figure 5, and clade divisions are indicated in Roman numerals. Tree was constructed using the Neighbor-Joining method [23]. The optimal tree with the sum of branch length = 2.20206590 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [48]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [24] and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA7 [21].

(a) *npr-4* transcript

tatgagcaatgaatcatgtttggatatgaacgatgaattatggcgtttttcgtaacgatctc
 M S N E S C L D M N D E L W R F R N D L
 actacacaaccagctgtgatggccattttcgcgttactctatgcgacgatcatcgactt
 T T Q P A V M A I F A L L Y A T I I V L
 ggcatatttggcaatcctttgtgttatttttagcgatcgcaaggacaagggtcattacaaacg
 G I F G N L C V I L A I A R T R S L Q T
 gtgccaaattttattcatattttcattgagttgttcggatgtcgttgtatgttgcataatcg
 V P N L F I F S L S C S D V V V C C I S
 gccactatcacaccaataactgccttcaaaaaagaatggctcttcggagttacactttgc
 A T I T P I T A F K K E W L F G V T L C
 tccatagcacctttcattgcggggatttcactatgtttctcaacgttcacattgactgcc
 S I A P F I A G I S L C F S T F T L T A
 atatcagtggtatcgatttttgcctaatttgcctttccaatgaaaaatgcgctaacaaaatca
 I S V D R F L L I C F P M K N A L T K S
 catgcacttatttctaattgcggggatctgtgcaatggccgcatgcctttcggcaccaata
 H A L I L I A G I C A M A A C L S A P I
 atgttcaagcagcgtctacaaacatttcgagaatttttgtggtaaattttgtaccgaggat
 M F K Q R L Q T F E N F C G K F C T E D
 tggggaatggatcagagtggaagacgtatctatggtgccataatggttggtgtacaattt
 W G M D Q S G R R I Y G A I M V G V Q F
 gttgtaccggttaacgattattgtcattttgctatacggcaatagcgggtgaaacttggtcag
 V V P L T I I V I C Y T A I A V K L G Q
 ggtatgctaataaaaaggaaaaaatagcaactacgattggcaagtccagatgaccgatcag
 G M L I K G K N S N Y D W Q V Q M T D Q
 caaagagcagccgtacgaaggcgacaacgaacaaatcgtatgttaatttgtatggtagtg
 Q R A A V R R R Q R T N R M L I C M V V
 gcattttccgctagttggctatggagtgtattattcaatgttttacgtgattatgatcga
 A F S A S W L W S V L F N V L R D Y D R
 ttacctgattggatgaaaattcaagaatatctttatggaattgcaactcattgtattgct
 L P D W M K I Q E Y L Y G I A T H C I A
 atgacatccacagtttggaaatcctttattatacgtgtattaaatccacaattacgcgt
 M T S T V W N P L L Y A V L N P Q L R A
 gcatttatacgccttatgcctagttgtctacaatcaacgactacatttgagaggtgtgat
 A F I R L M P S C L Q S T T T F E R C D
 caattaactcgaaatgggtactctactcgaaacgggtgacaggcgagattcgacaaaaatg
 Q L T R N G T L L E T G D R R D S T K M
 caacaaaacgtgaggcaagccatatcatctgcaccgggacagtatggagcgacagaggtt
 Q Q N V R Q A I S S A P G Q Y G A T E V
 agggctatctggagtgcgtcacagagctgagttgcgcgagctgcggttcttgaa
 R A I W S D V T E L S C A S C G S -

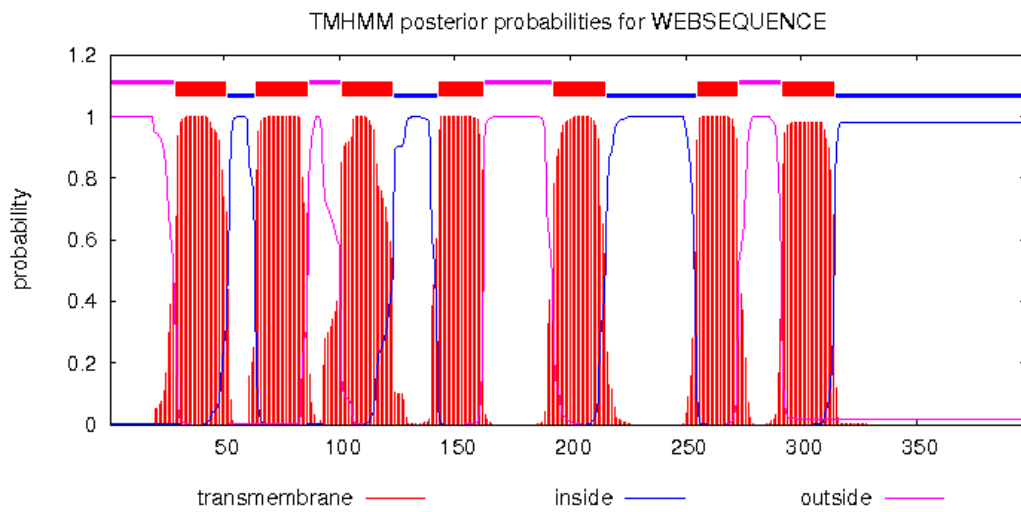
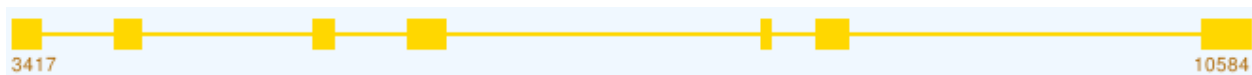
(b) Transmembrane domain prediction**(c) Gene model, 7 exons (AMPH01002059.1)**

Figure 7. *npr-4* transcript. (a) Cloned *npr-4* transcript (MF319189). (b) TMHMM predicts 7 transmembrane domains. (c) All 7 exons map to the same scaffold of the *A. suum* genome (AMPH01002059.1).

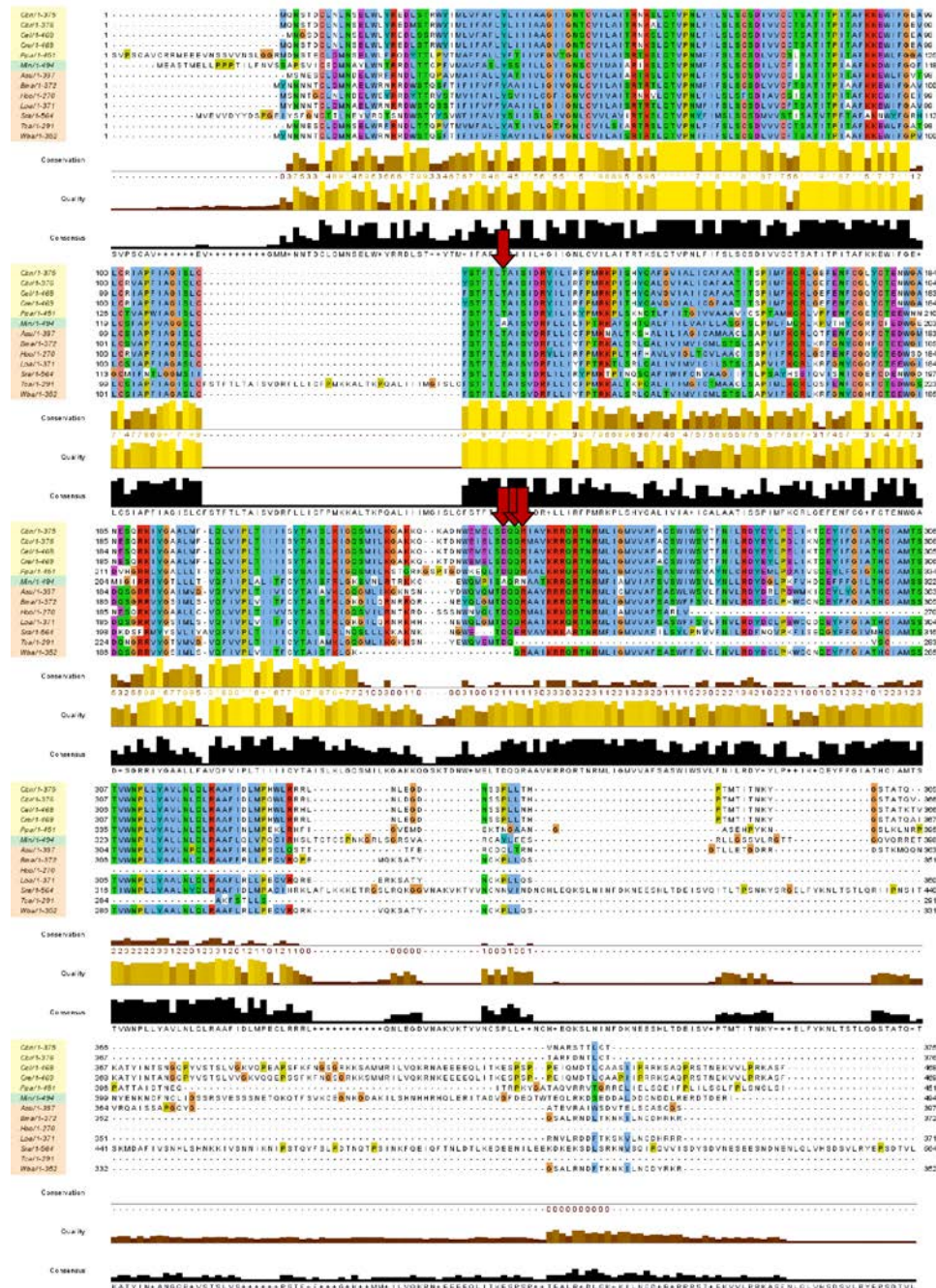


Figure 8. Comparison of *npr-4* transcripts in nematodes. *Asu* = *Ascaris suum*, *Bma* = *Brugia malayi* (CDP94294.2), *Cbn* = *Caenorhabditis brenneri* (EGT46658.1), *Cbr* = *Caenorhabditis briggsae* (XM_002645190.1), *Cel* = *Caenorhabditis elegans* (NM_001307873.1), *Cre* = *Caenorhabditis remanei* (XM_003101112.1), *Hco* = *Haemonchus contortus* (CDJ91479.1), *Loa* = *Loa loa* (XP_020302592.1), *Min* = *Meloidogyne incognita* (AY882441.1), *Ppa* = *Pristionchus pacificus* (KKA78805.1), *Sra* = *Strongyloides ratti* (CEF70404.1), *Tca* = *Toxocara canis* (KHN76189.1), *Wba* = *Wuchereria bancrofti* (EJW83575.1). Free-living nematodes are shaded in yellow, plant parasites in green, and human/animal parasites in orange. Red arrows indicate residues that differ between groups.

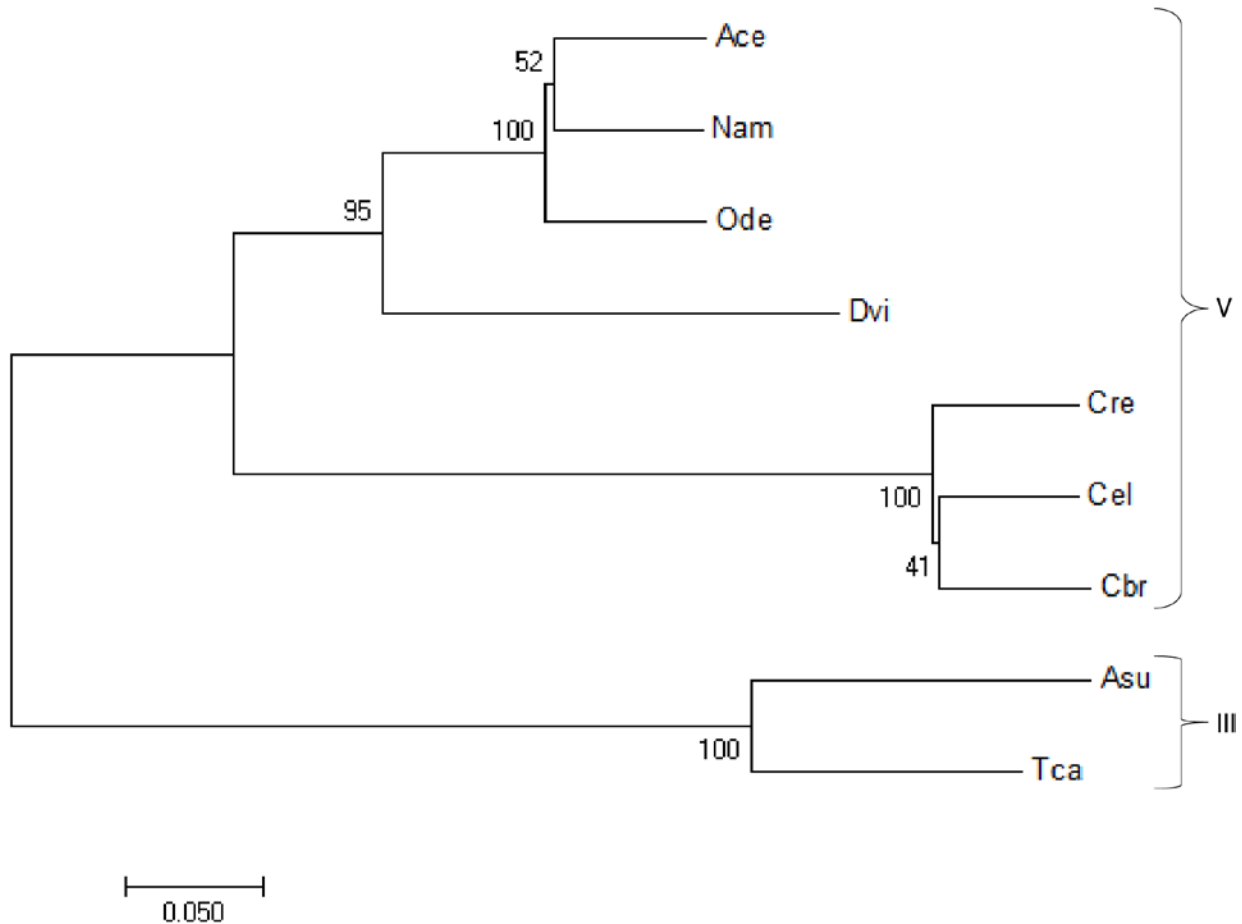
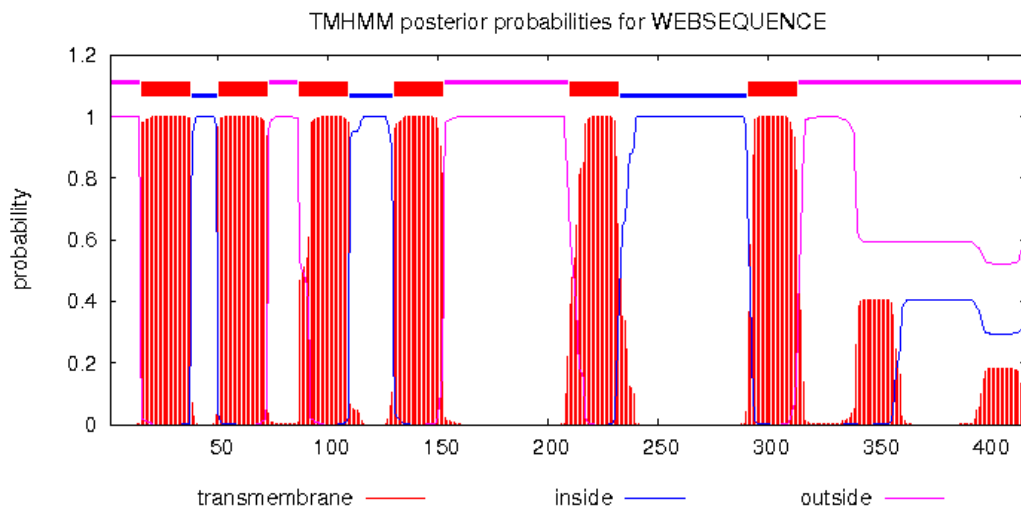


Figure 9. Phylogenetic tree of NPR-4 receptor homologs in nematodes. Nematode sequences and accession numbers are the same as reported in Figure 8, and clade divisions are indicated in Roman numerals. Tree was constructed using the Neighbor-Joining method [23]. The optimal tree with the sum of branch length = 1.44110411 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [48]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [24] and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA7 [21].

(a) *npr-8* transcript

atggattttatttcttgaggaggtacagtgaatgggaggaagttattttatacggcaatttac
 M D L F L R R Y S E W E E V I Y T A I Y
 atgggtcatatcggtgctcgccgtcattgggaatggcgtagtaatattagcagtcattcgg
 M V I S V L A V I G N G V V I L A V I R
 aaaagggaaatgaggacgaatcgaaatgtattaatcgtaatttggcactatcgaatttg
 K R E M R T N R N V L I V N L A L S N L
 atgcttgcgctcactaatattccttttctgtggcttccgtctattgacttcgaatttctg
 M L A L T N I P F L W L P S I D F E F L
 tattcgcggttttttctgcaaatttgctaacgcgctacccggaagtaacatctattgctca
 Y S R F F C K F A N A L P G S N I Y C S
 acgctgactatttctgtgatggcgattgatcgctattataaccgtgaaaaagatgtacgca
 T L T I S V M A I D R Y Y T V K K M Y A
 agctctcatcgacgtcagtggtttacgtgctgtatgtatttccatattgatttggatcatc
 S S H R R Q C L R A V C I S I L I W I I
 tcattcttgctgagcttttcttctcgtttactacgacactaccatgctctacgtattc
 S F L L S F P L L V Y Y D T T M L Y V F
 aaggatgtgactgtattttcgtctggaaatgcgacgaagttaagcagttatgggttggcgg
 K D V T V F S S G N A T K L S S Y G W R
 cagtgcaggctggcgccaactcaaagaaaggtatggaagacccaatgaagctgccaaa
 Q C R L A P T Q R K G M E D P N E A A K
 cgtgaaagcgacgcacaaaaagtacagctttttatgtcgggtgatgcaagcgctatttctt
 R E S D A Q K V Q L F M S V M Q A L F L
 tatgtggtaccgcttatcggttcttttgattttcaacgcaaagttgacacgctttttgcaa
 Y V V P L I V L L I F N A K L T R F L Q
 gaaaattcacggcaaatgcatcgatcgctcagcgatggtcgcagtgacgatgcgggtaca
 E N S R Q M H R S L S D G R S D D A G T
 ttatgcatcgtgaaatcgtccgatcttggacagacggccgctgttcggaacttgccggcg
 L C I V K S S D L G Q T A A V R N L P A
 aggaggatgaatgtgcgcaaaagaaatagaacaactgcattgctaattgcatggctgct
 R R M N V R K R N R T T A L L I A M A A
 tcttatgctgtgctttggctaccggtttaccggttgtttcaatgctccttgacctacgaatt
 S Y A V L W L P F T V V S M L L D L R I
 ctctatatggaaggcaatggagcggcagtaattgagcgtatcgatcaatcatgtaaaatg
 L Y M E G N G A A V I E R I D Q S C K M
 atatcgataatgtcaatttgtgtgaaccattttctatacgggttttctcaacacaaacttc
 I S I M S I C V N P F L Y G F L N T N F
 cgacatgaattcacagacattttcaatcaatggatatgctacaaattacgcacaccgaca
 R H E F T D I F N Q W I C Y K L R T P T
 aacgtccaacaagagtattcaacgatcgcttctcctcgaagggttcaccttctttttttt
 N V Q Q E Y S T I A S P R R F T F F F F
 catcgctatcgatatttatttcatattcatctcacaatatgtgaacttccgttgaa
 H R Y R I Y F I F I S Q Y V N F R -

(b) Transmembrane domain prediction



(c) Gene model exons (AMPH01006138.1)



Figure 10. *npr-8* transcript. (a) Cloned *npr-8* transcript (MF319190). (b) TMHMM predicts 6 transmembrane domains, with weaker prediction of TM7. (c) Gene model of 7 exons mapped onto *A. suum* scaffold (AMPH01006138.1).

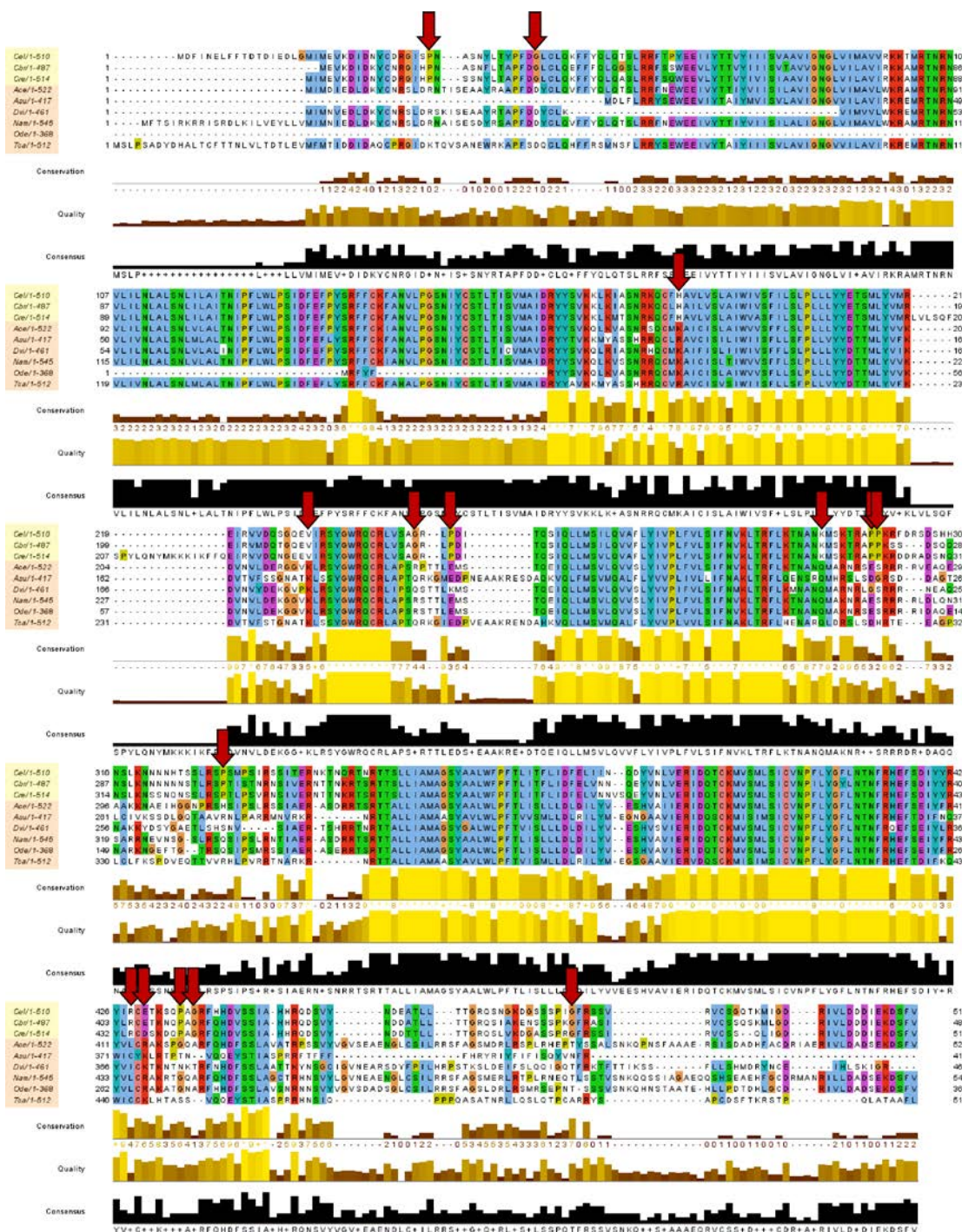


Figure 11. Comparison of NPR-8 homologs in nematodes. *Ace* = *Ancylostoma ceylanicum* (EYC20064.1), *Asu* = *Ascaris suum*, *Cbr* = *Caenorhabditis briggsae* (XM_002644982.1), *Cel* = *Caenorhabditis elegans* (NM_171723.4), *Cre* = *Caenorhabditis remanei* (XM_003118371.1), *Dvi* = *Dictyocaulus viviparus* (KJH40534.1), *Nam* = *Necator americanus* (XM_013450376.1), *Ode* = *Oesophagostomum dentatum* (KHJ95503.1), *Tca* = *Toxocara canis* (KHN80044.1). Free-living nematodes are shaded in yellow, and human and animal parasites in orange. Red arrows indicate residues that differ between groups.

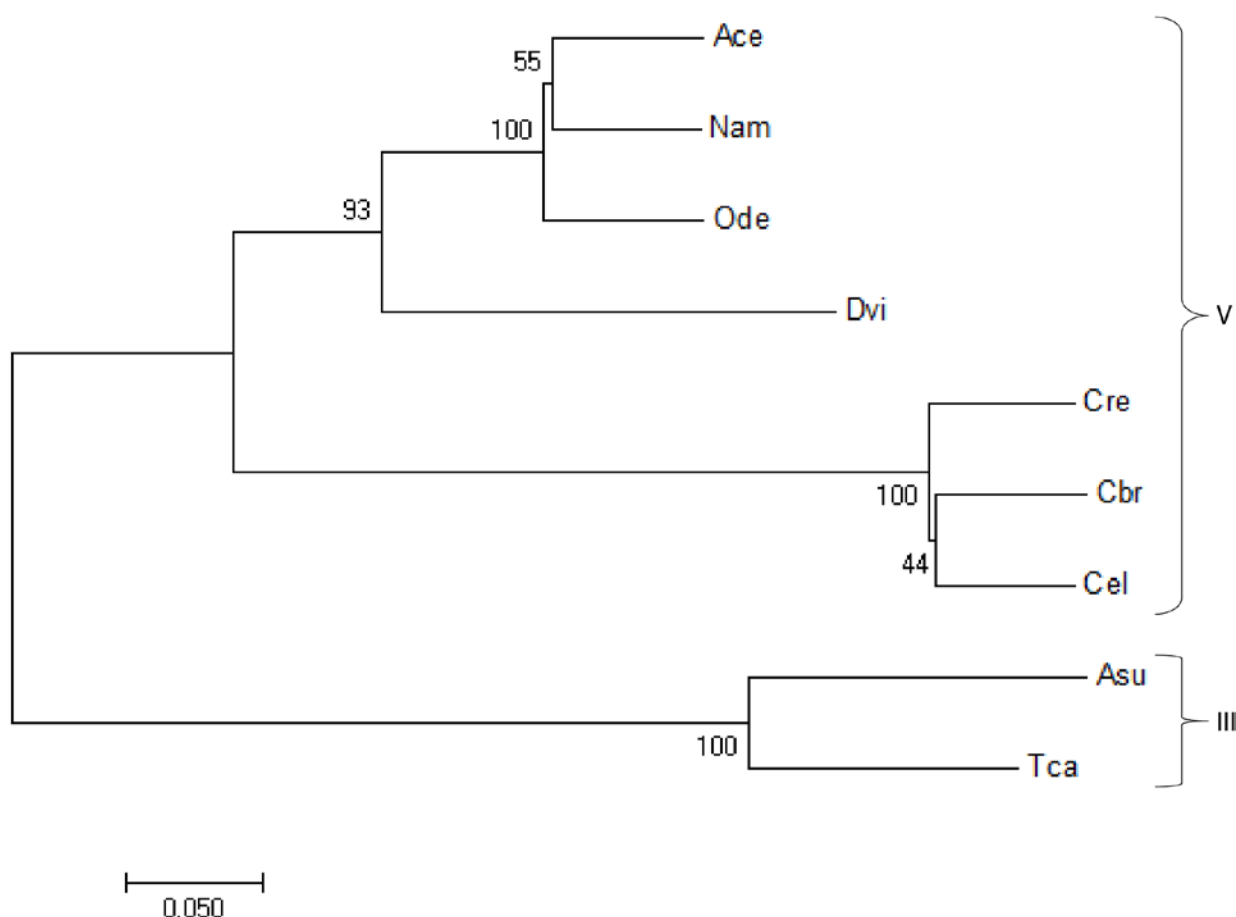


Figure 12. Phylogenetic tree of NPR-8 receptor homologs in nematodes. Nematode sequences and accession numbers are the same as reported in Figure 11, and clade divisions are indicated in Roman numerals. Tree was constructed using the Neighbor-Joining method [23]. The optimal tree with the sum of branch length = 1.44110411 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [48]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [24] and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA7 [21].

(a) *npr-10* transcript

tatgagcaataacaacgccagagtgcacatcgacataaaggaatacctatggaaaaatcagtcc
 M S N T T P E C I D I K E Y L W K N Q S
 gaccttacttccttgccatccacaatgagtgtattcgccgtcatctactcgttcattcatt
 D L T S L P S T M S V F A V I Y S F I I
 acgttaggcgttctcggcaacacattggtggtgttatcggtgatacgacacaaatcgtta
 T L G V L G N T L V V L S V I R H K S L
 cagtcagtcgcaacatgttcattcgtcagtcctgtcagtcgtcggacatcgttatctcgatt
 Q S V R N M F I V S L S V S D I V I S I
 gtatccgggtcaataacaccgatcactgctttcacaaagatatggctcttcggcggaattc
 V S G S I T P I T A F T K I W L F G E F
 ctctgctatttgggtgccactttttcaggggtgcaagtctttgcttttcaacgttaacacta
 L C Y L V P L F Q G A S L C F S T L T L
 atggcgattgctatcgatcgctttattctcattatatatccaatgaaaaggcccattcaa
 M A I A I D R F I L I I Y P M K R P I Q
 aaggtacatgcgttatggatgattgctttgaataattacactcgcgacggcaatctcaatg
 K V H A L W M I A L N I T L A T A I S M
 ccaatgtttctgatgcagaaactgggtcgactacggcaacttttgtggacaattttgcacc
 P M F L M Q K L V D Y G N F C G Q F C T
 gaagattggggcaataatgaatttggccgaagtacatatggaacggtcgttttcatattg
 E D W G N N E F G R S T Y G T V V F I L
 caattcgttgttccattgatcgtaatcaccttctgctacacaatgatctcactgaaactt
 Q F V V P L I V I T F C Y T M I S L K L
 aataaaggctctgctgggtgaaacaggcatcgaagcatgcacgggtctaggaaccgagcaa
 N K G L L V K Q A S K H A S G L G T E Q
 cgtcgtttggcactgcaacggcgcttacgaacgaatcgatatgccaatggcgatgggttgg
 R R L A L Q R R L R T N R M P M A M V G
 gtatttttgtgctgttggtatgcctactgtcgcattcaactttcttagagactacagatgg
 V F L C C W M P T V A F N F L R D Y R W
 ctgccatctttcgtcgcacaacaggaatatctcttttggtattattacacactgcattcta
 L P S F V A Q Q E Y L F G I I T H C I S
 atgagttcgacagtatggaatccgtgtctttatgcgttgcttaacgaacaatttcggcta
 M S S T V W N P C L Y A L L N E Q F R L
 gcatatgtcgatctactgcagttattgtcgaagcggacagaagagatgtggtactaggtat
 A Y V D L L Q Y C R S G Q K R C G T R Y
 accgatgcacgttggttcacgacttgtctcgaatgcttttcttgctaataacagtatcact
 T D A R C S R L V S N A F L A N N S I T
 gaagagaatatataaattgcctaactcagtaaggaggagtaagcatgggtgattacacggcc
 E E N I K L P N S V R R S K H G D Y T A
 gtatccagttcattataa
 V S S S L -

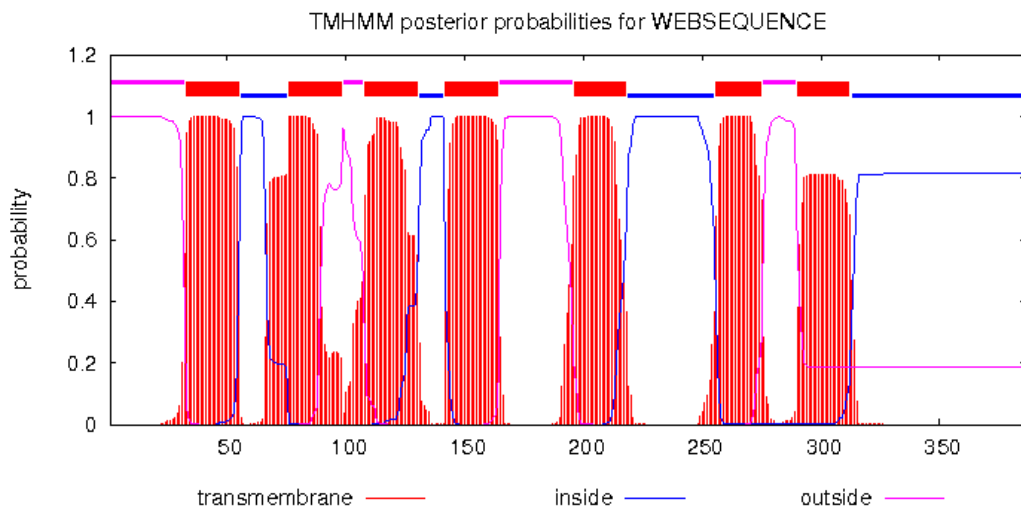
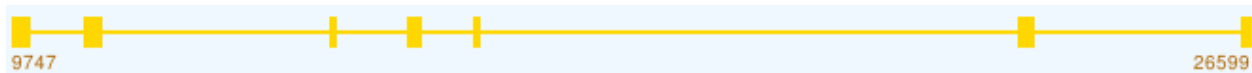
(b) Transmembrane domain prediction**(c) Gene model of exons 1-7 (ANBK01002419.1)**

Figure 13. *npr-10* transcript. (a) Cloned *npr-10* transcript (MF319191). Single nucleotide polymorphisms compared to sequences in the *A. suum* transcriptome are bolded. (b) TMHMM predicted 7 transmembrane domains. (c) All 7 exons mapped onto the same scaffold in the *A. suum* genome (ANBK01002419.1).

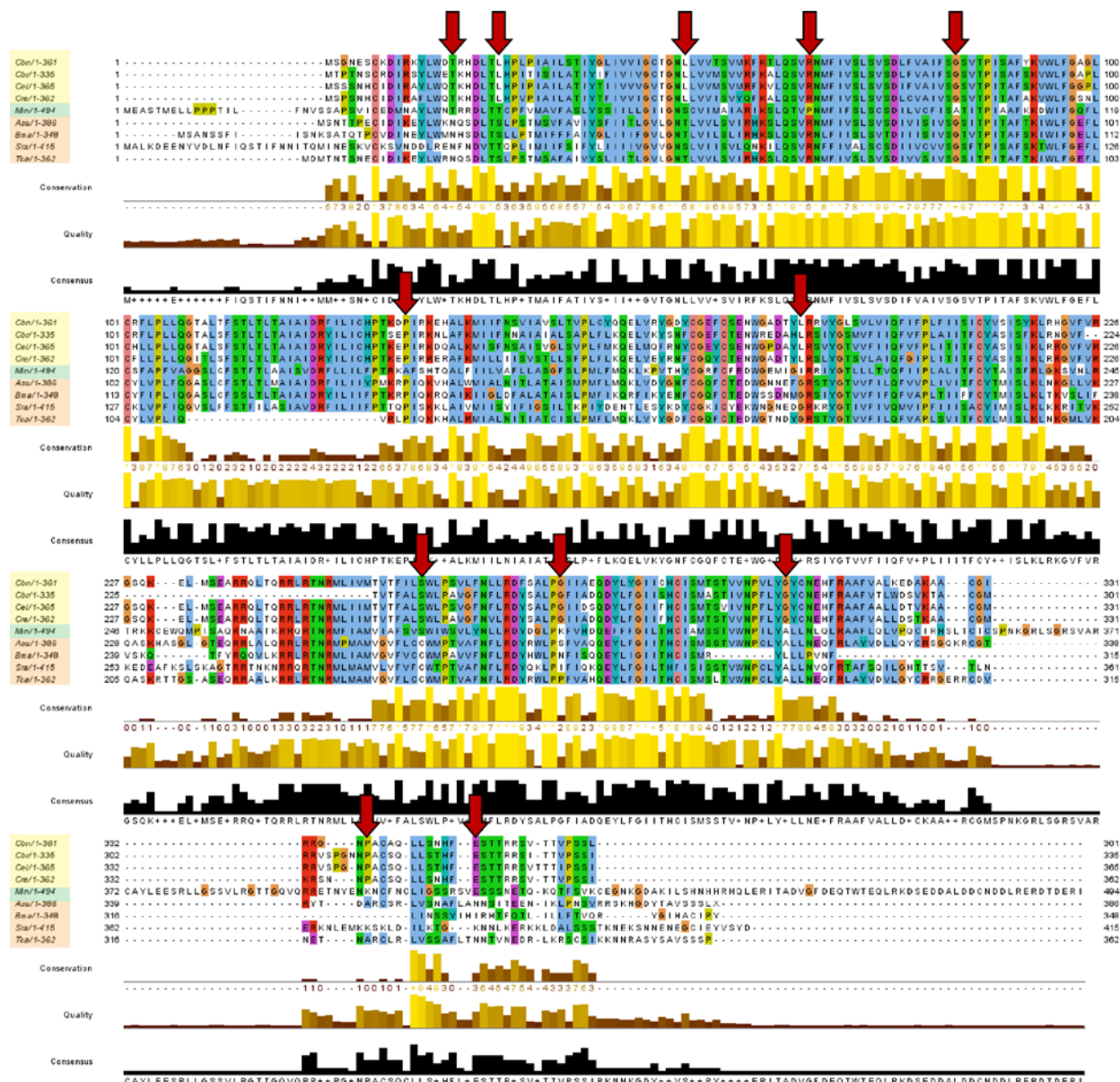


Figure 14. Comparison of NPR-10 receptor homologs in nematodes. *Asu* = *Ascaris suum*, *Bma* = *Brugia malayi* (CDP96008.1), *Cbn* = *Caenorhabditis brenneri* (EGT46754.1), *Cbr* = *Caenorhabditis briggsae* (CAP34742.2), *Cel* = *Caenorhabditis elegans* (NM_001029316.1), *Cre* = *Caenorhabditis remanei* (XM_003099783.1), *Min* = *Meloidogyne incognita* (AY882441.1), *Sra* = *Strongyloides ratti* (CEF64234.1), *Tca* = *Toxocara canis* (KHN75272.1). Free-living nematodes are shaded in yellow, plant parasites in green, and human/animal parasites in orange. Red arrows indicate residues that differ between groups.

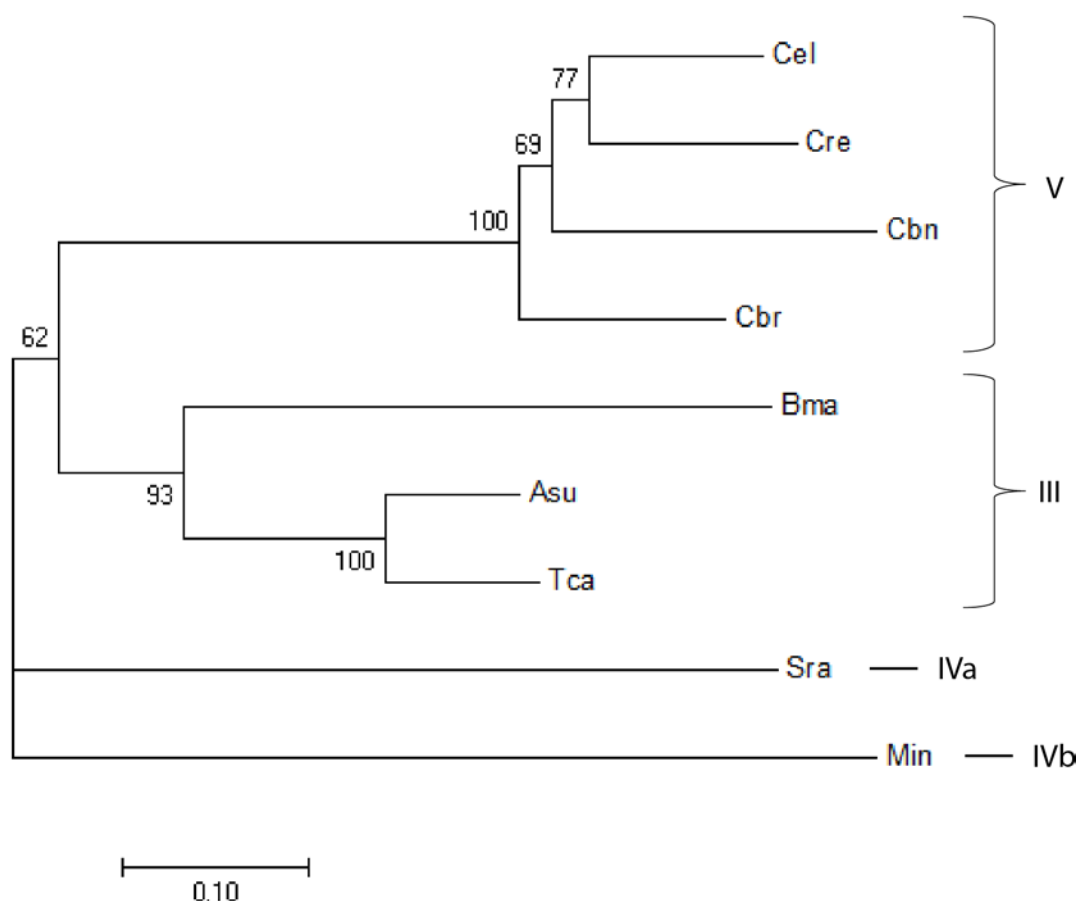


Figure 15. Phylogenetic tree of NPR-10 receptor homologs in nematodes. Nematode sequences and accession numbers are the same as reported in Figure 14, and clade divisions are indicated in Roman numerals. Tree was constructed using the Neighbor-Joining method [23]. The optimal tree with the sum of branch length = 2.31040756 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [48]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [24] and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA7 [21].

(a) *npr-11* transcript

atgccaaacagtagcctagcagaaaaagactgtaaccaatggcagcatatgttctcagaa
 M P N S T L A E K D C N Q W Q H M F S E
 atcaactcctacttccggaacgaaattgtgctggaaggtagctgacatgcttcagttgag
 I N S Y F R N E I V L E G T D H A S V E
 ctcggttggtgcatcacattcgccattgttttgtgatcgtgttcggtgcgctcggcaat
 L G W C I T F A Y C F V I V F G A L G N
 atgctcacctgtgttggtgtcatcacgaaatccacaaatgcgaacaacaagaaactttttt
 M L T V L A V I R N P Q M R T T R N F F
 attgtcaatttagctctatccgatttttttatctgtacagttacagcacctatcactctc
 I V N L A L S D F F I C T V T A P I T L
 tacacggttttatatacgttttggccatttgaacagcgacgtgtaaaattgctggctcg
 Y T V L Y T F W P F G T A T C K I A G S
 cttcagggattcggcatttttttatcaactttttctattgcagcgattgcgctcgaataga
 L Q G F G I F L S T F S I A A I A L D R
 tatgttttagttatatttccaacgaaacggaaaaggcagcaccaatctttcattgcttttt
 Y V L V I F P T K R K R Q H N L S L L F
 ttctcatthaatttggcttatttcaattgcactcgcgcttccggtattaattgcatcggac
 F S L I W L I S I A L A L P L L I A S D
 ttgaatacgaatattcgaagatcaccaattgcgacatttctactgcaaatttggcacgaaaaa
 L N T I F E D H N C D I S L Q I C H E K
 aatgaaatttggcaacagatgccgattggttaaggaagcatatactttggcggtagtagta
 N E I W Q Q M P I G K E A Y T L A V L V
 acgcaatatgctttaccactaatttcaattgtatttgcgtacactcggatagcacacagg
 T Q Y A L P L I S I V F A Y T R I A H R
 atgcaaggacaatttgcgattcgcgatgagctcgtgcgcaaactcgctcgccactatgagt
 M Q G Q F A I R M S S C A N S L A T M S
 catcgacgaaggtagttgtggaacgtcagaggcgaactcatttacttttgatgtgtgtg
 H R R R S V V E R Q R R T H L L L M C V
 gttattatgtttgcggttgcatggctgccactcaatgtgttccacgtgttgaacacattc
 V I M F A V A W L P L N V F H V L N T F
 ggaatcgtaagggtttctcgggtgccctatttgcactatgtcatgtgatagcgatggga
 G I V K G F S V P L F A L C H V I A M G
 tcagcgtgtctcaatcccgtctcatatgccttcttcaatcaccaattttcgacagcagttc
 S A C L N P V S Y A F F N H N F R Q Q F
 gtcgctatgttgcgaacaagctggactaagccgcctaagagggtatattgcttcaccttcga
 V A M F E Q A G L S R L R G I L L H L R
 aagaagtgcattgtaggtggacgaagaaagtacgctgcaacgtcatgtgcagtggcatca
 K K C M L G G R R K Y A A T S C A V A S
 agtaagtcgccacggatgttgccgaccactgtacacaatgaagaatttaatgaagaacac
 S K S P R M L P T T V H N E E F N E E H
 gatgcagcaccaaaagagagcgagtttatcttaaggattaatatgggtgaacactgtttgaa
 D A A P K E S E F I L R I N M V N T V -

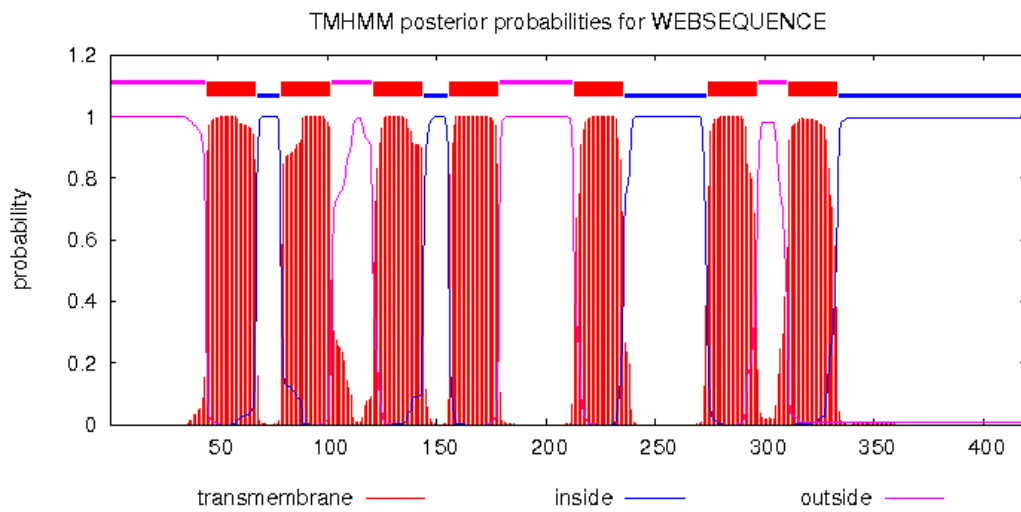
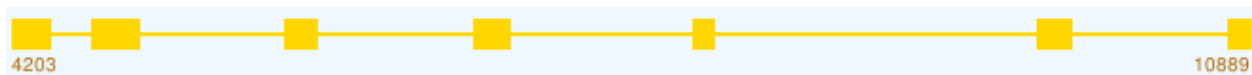
(b) Transmembrane domain prediction**(c) Gene model of exons 1-7 (ANBK01000995.1)**

Figure 16. *npr-11* transcript. (a) Cloned *npr-11* transcript (MF319192). (b) TMHMM predicts 7 transmembrane domains. (c) All 7 exons mapped to a single scaffold of the *A. suum* genome (ANBK01000995.1).

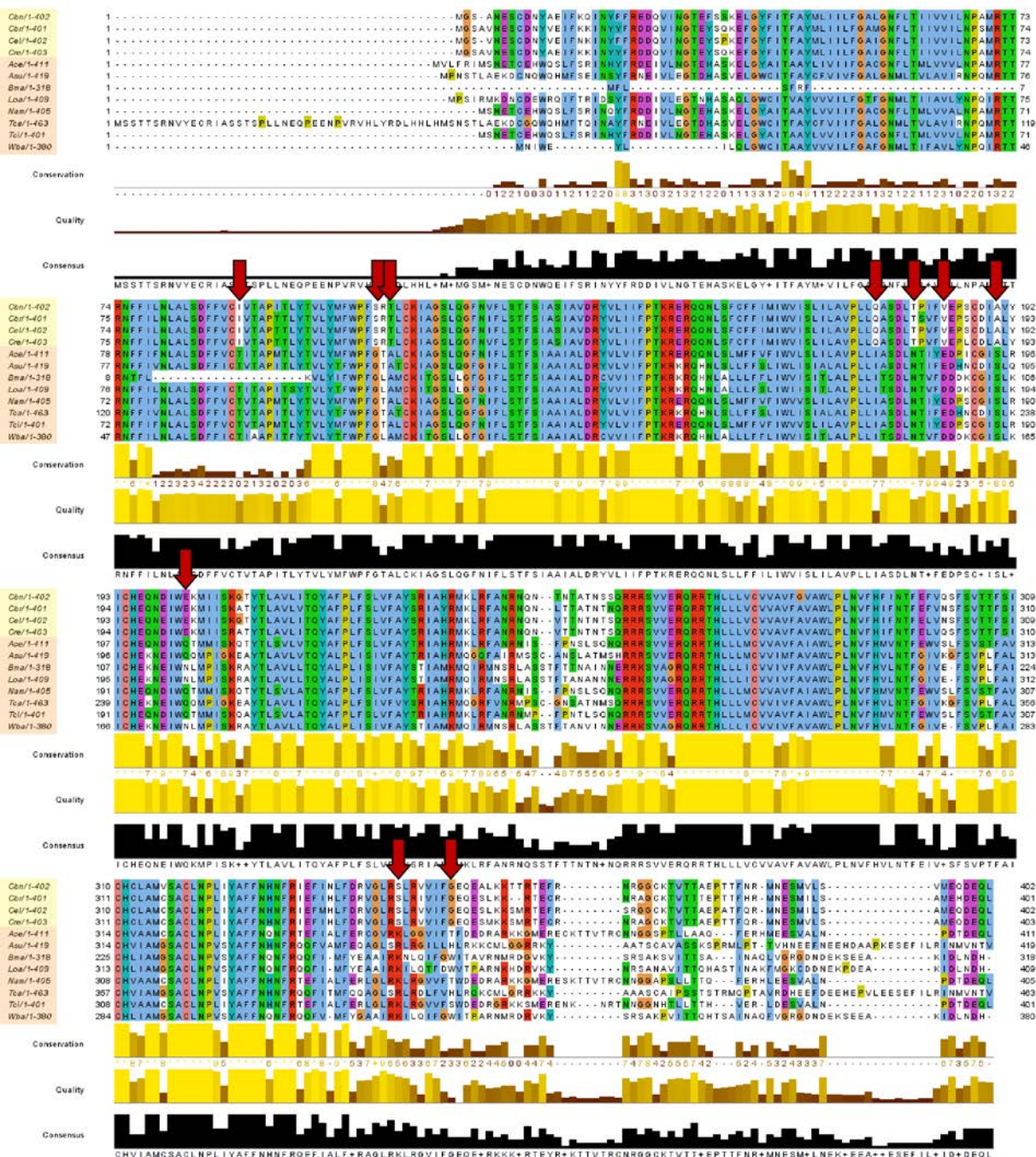


Figure 17. Comparison of NPR-11 receptor homologs in nematodes. *Ace* = *Ancylostoma ceylanicum* (EYC24047.1), *Asu* = *Ascaris suum*, *Bma* = *Brugia malayi* (XM_001897640.1), *Cbn* = *Caenorhabditis brenneri* (EGT54821.1), *Cbr* = *Caenorhabditis briggsae* (XM_002643146.1), *Cel* = *Caenorhabditis elegans* (NM_075833.2), *Cre* = *Caenorhabditis remanei* (XM_003106875.1), *Loa* = *Loa loa* (XP_020302033.1), *Nam* = *Necator americanus* (XM_013445476.1), *Tca* = *Toxocara canis* (KHN76098.1), *Tci* = *Teladorsagia circumcincta* (DQ840477.1), *Wba* = *Wuchereria bancrofti* (EJW87503.1). Free-living nematodes are shaded in yellow, and human/animal parasites in orange. Red arrows indicate residues that differ between groups.

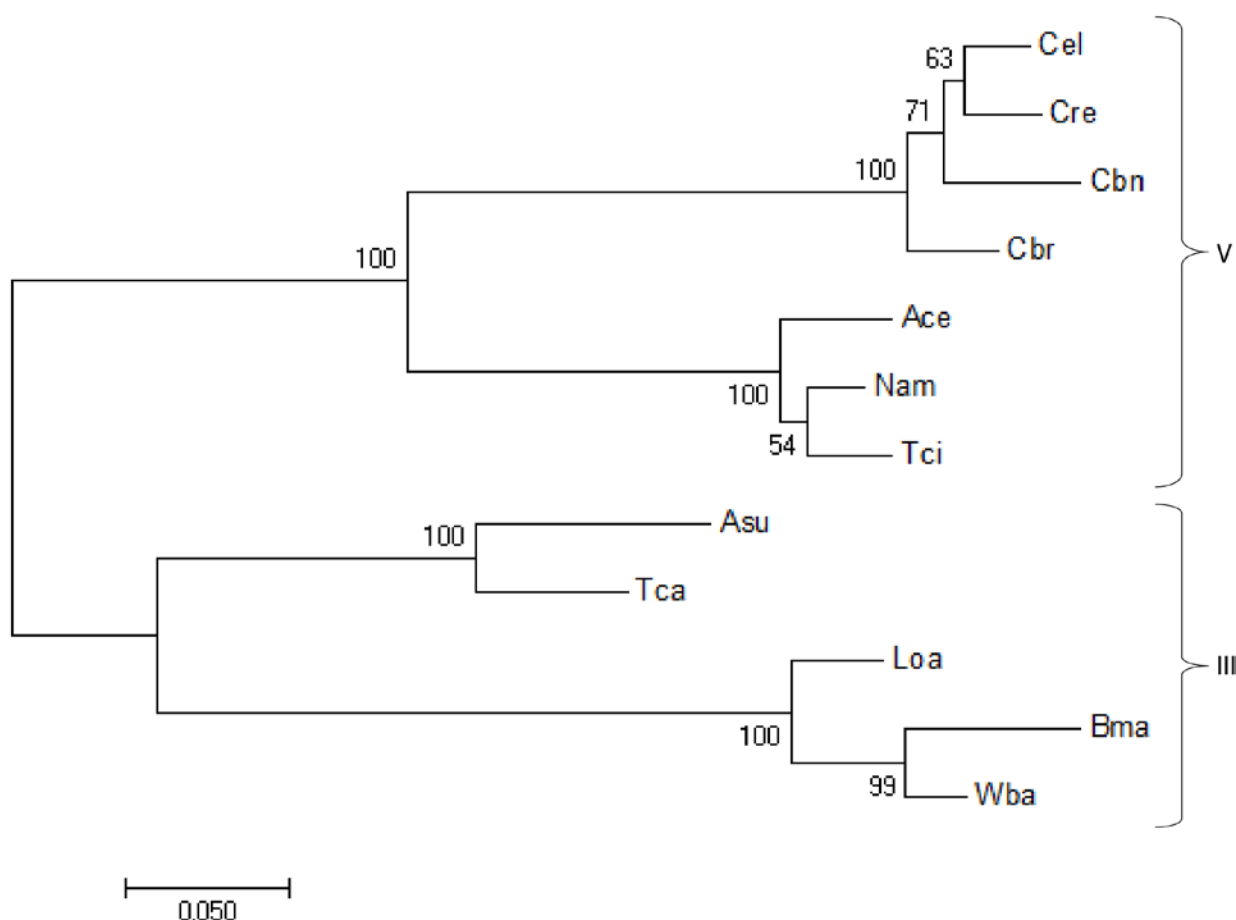
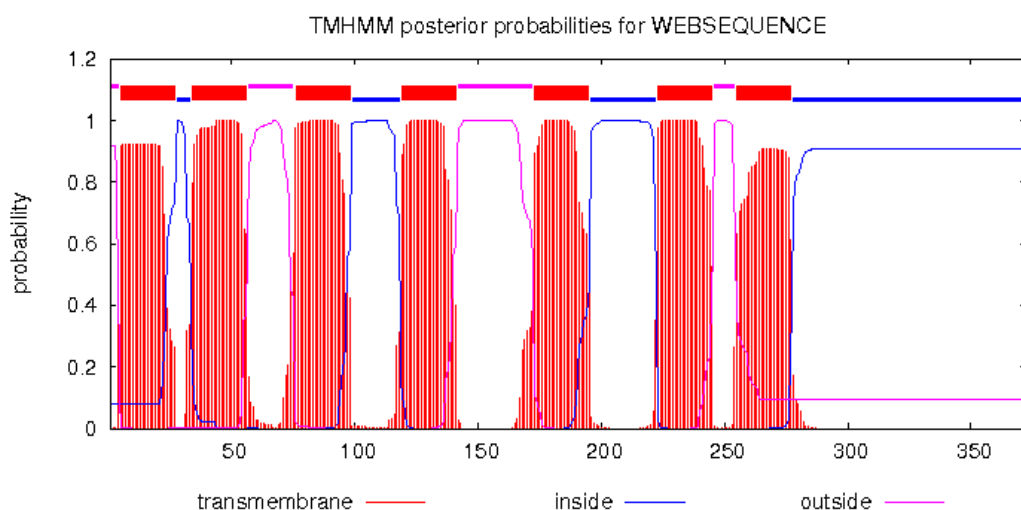


Figure 18. Phylogenetic tree of NPR-11 receptor homologs in nematodes. Nematode sequences and accession numbers are the same as reported in Figure 17, and clade divisions are indicated in Roman numerals. Tree was constructed using the Neighbor-Joining method [23]. The optimal tree with the sum of branch length = 1.19100089 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [48]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [24] and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA7 [21].

(a) *frpr-6* transcript

gcgttggtgtaatgattataatgtgtatgtgtggtgtagttggtaattcactctcactctac
 V V V M I I M C M C G V V G N S L S L Y
 gtctattggacaccaacatttcgtaaagcttcgataaatctcctattggctgcgttgagt
 V Y W T P T F R K R S I N L L A A L S
 gcctctgattttattcttatgccttctagcgatgccagttttctcaataacgcaaattcaa
 A S D L F L C L L A M P V F S I T Q I Q
 agttttctacctggcctatcaaagcatttaaccgggtcatattcttctttttgcttattcct
 S F L P G L S K H L T G H I L L F A Y P
 gttacactaatgctacaaagcatctcagtatggttattagtaagtattacaattgatcgt
 V T L M L Q S I S V W L L V S I T I D R
 tatttggctgtttggtatcctttcaaagtacgcgcctattgtacaagaactcgagcaacg
 Y L A V C Y P F K V R A Y C T R T R A T
 ataactataatattgatagatttcttttcgatcgtctataattttgtacgcttttgggag
 I T I I L I V F F S I V Y N F V R F W E
 ttacttttagacgatttcgggcggaacatcatcaatcgaggaatcaatcgtaccattatta
 F T L D D S G G T S S I E E S I V P L L
 aggggtgacccccctatttatgcttctctatcagaatatcgccactctaatacacacaattc
 R G D P L F M L L Y Q N I A T L I T Q F
 tttttacctcttattgtactttgcttcttaaatattacatgttgcgaggacaattctagag
 F L P L I V L C F L N L H V A R T I L E
 gctgggtgtgcaacgccgtgaactagtagcatcagaaaggcgcggaatatcgtagtgcacaaa
 A G V Q R R E L V A S E R R E Y R T A K
 atgatgctttatgttggttattgtattttatattttgctatgctcttagtatgattctcaat
 M M L Y V V I V F I F C Y A L S M I L N
 cttgctgagattttggatgaagattttatttcgacaaccaattggttatcttctcaacgat
 L A E I L D E D L F R Q P I G Y L L N D
 gttaataacattcttgctcgttattaattcgtcatcatcgttcatattttacacaaaatat
 V N N I L V V I N S S S S F I F Y T K Y
 tcaacacgattccgtgcacaactaaggaacgtatgctgtattttcattttgcaatggacgt
 S T R F R A Q L R N V C C I S F C N G R
 ctctcatgctcttacaatgccaactgccaagatcaagctgcactgaagccgcgtactagc
 L S C S Y N A N C Q D Q A A L K P R T S
 tttgatgtaagtacactgcgagcaatctaccttctcagaaacgtagctcaagctgcagc
 F D V S T L R S N L P S Q K R S S S C S
 gaaaagaaggcctcacgactcgacctcttcactaagctcattgaatgtgagcgctagtca
 E K K A S R L D L F T K L I E C **E R** -
aggtatcgacatgcgccttttcgtcccccattcccatctctcaaacgtatatgggtga
 I D M R L S S P I P S L K R I W -

(b) Transmembrane domain prediction



(c) Gene model for exons 1-7 (ANBK01002449.1)



Figure 19. *frpr-6* transcript. (a) Cloned *frpr-6* transcript (MF319193). Cloned sequence contained a 16 bp insertion (bolded nucleotides) compared to sequence reported in the *A. suum* transcriptome (PRJNA80881:GS_11869). This insert results in a stop codon prior to exon 7, which would shorten the C-terminal tail by 14 amino acids. The insert is flanked by putative splice sites in the genome, allowing for the possibility of alternative splicing to produce 2 versions of the FRPR-6 receptor. (b) TMHMM predicts 7 transmembrane domains, but only with inclusion of the three Val residues preceding the Met. (c) Exons 1-7 mapped onto the same scaffold of the *A. suum* genome (ANBK01002449.1).

Figure 20. Comparison of FRPR-6 receptor homologs in nematodes. *Ace* = *Ancylostoma ceylanicum* (EYB84317.1), *Asu* = *Ascaris suum*, *Cbn* = *Caenorhabditis brenneri* (EGT43739.1), *Cbr* = *Caenorhabditis briggsae* (XP_002635595.1), *Cel* = *Caenorhabditis elegans* (NP_505004.1), *Cre* = *Caenorhabditis remanei* (XP_003115478.1), *Gpa* = *Globodera pallida* (AFJ93333.1), *Hco* = *Haemonchus contortus* (CDJ92216.1), *Sra* = *Strongyloides ratti* (CEF62596.1), *Tca* = *Toxocara canis* (KHN73880.1). Free-living nematodes are shaded in yellow, plant parasites in green, and human/animal parasites in orange. Red arrows indicate residues that differ between groups.

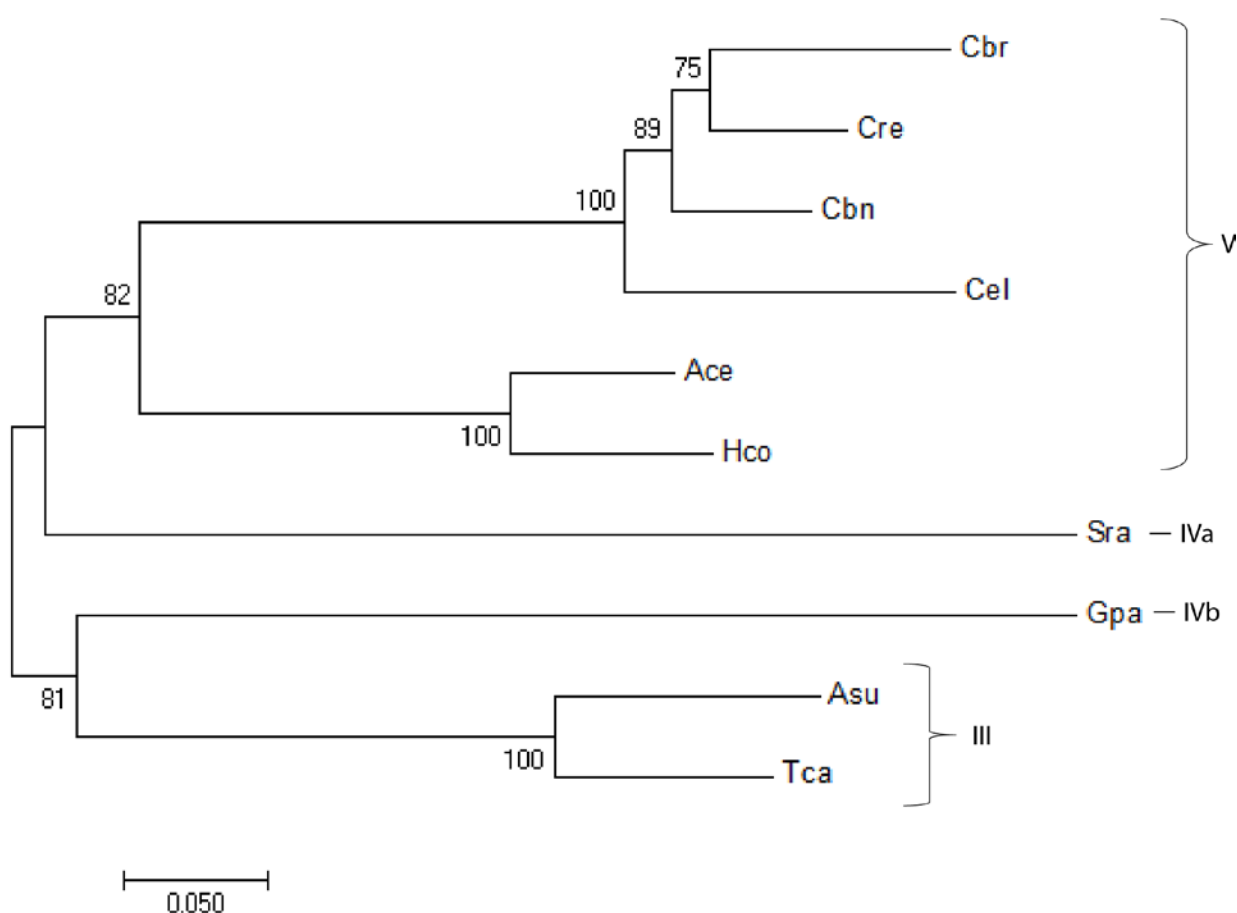


Figure 21. Phylogenetic tree of FRPR-6 receptor homologs in nematodes. Nematode sequences and accession numbers are the same as reported in Figure 20, and clade divisions are indicated in Roman numerals. Tree was constructed using the Neighbor-Joining method [23]. The optimal tree with the sum of branch length = 1.84974341 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [48]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [24] and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA7 [21].

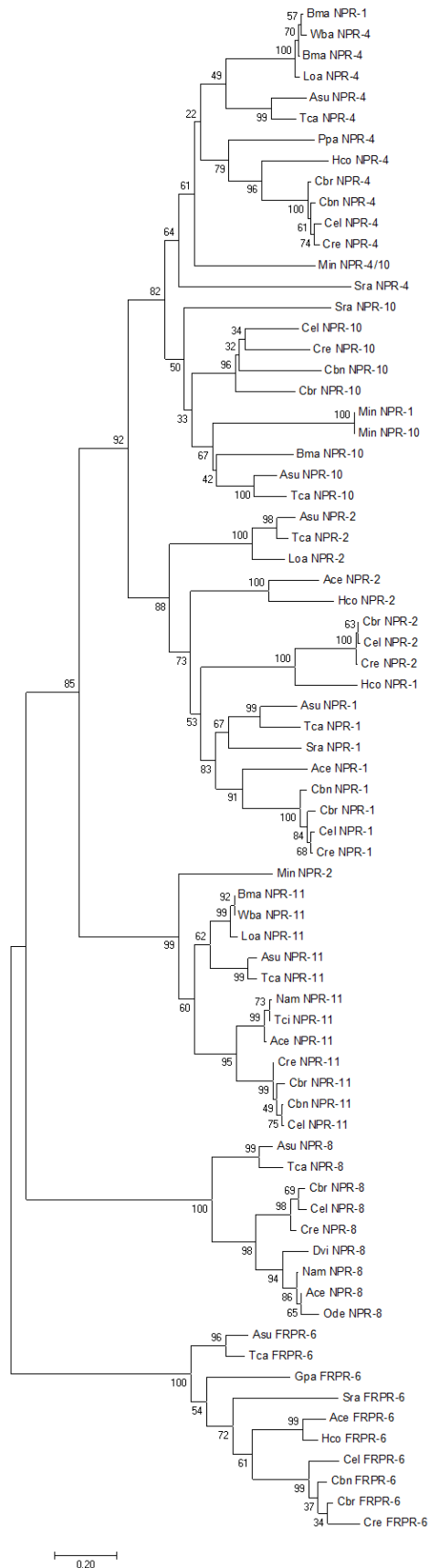


Figure 22. Phylogenetic tree of all receptor homologs in nematodes. Tree was constructed using the Neighbor-Joining method [23]. The optimal tree with the sum of branch length = 14.65885092 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [48]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [24] and are in the units of the number of amino acid substitutions per site. The analysis involved 73 amino acid sequences. Evolutionary analyses were conducted in MEGA7 [21].

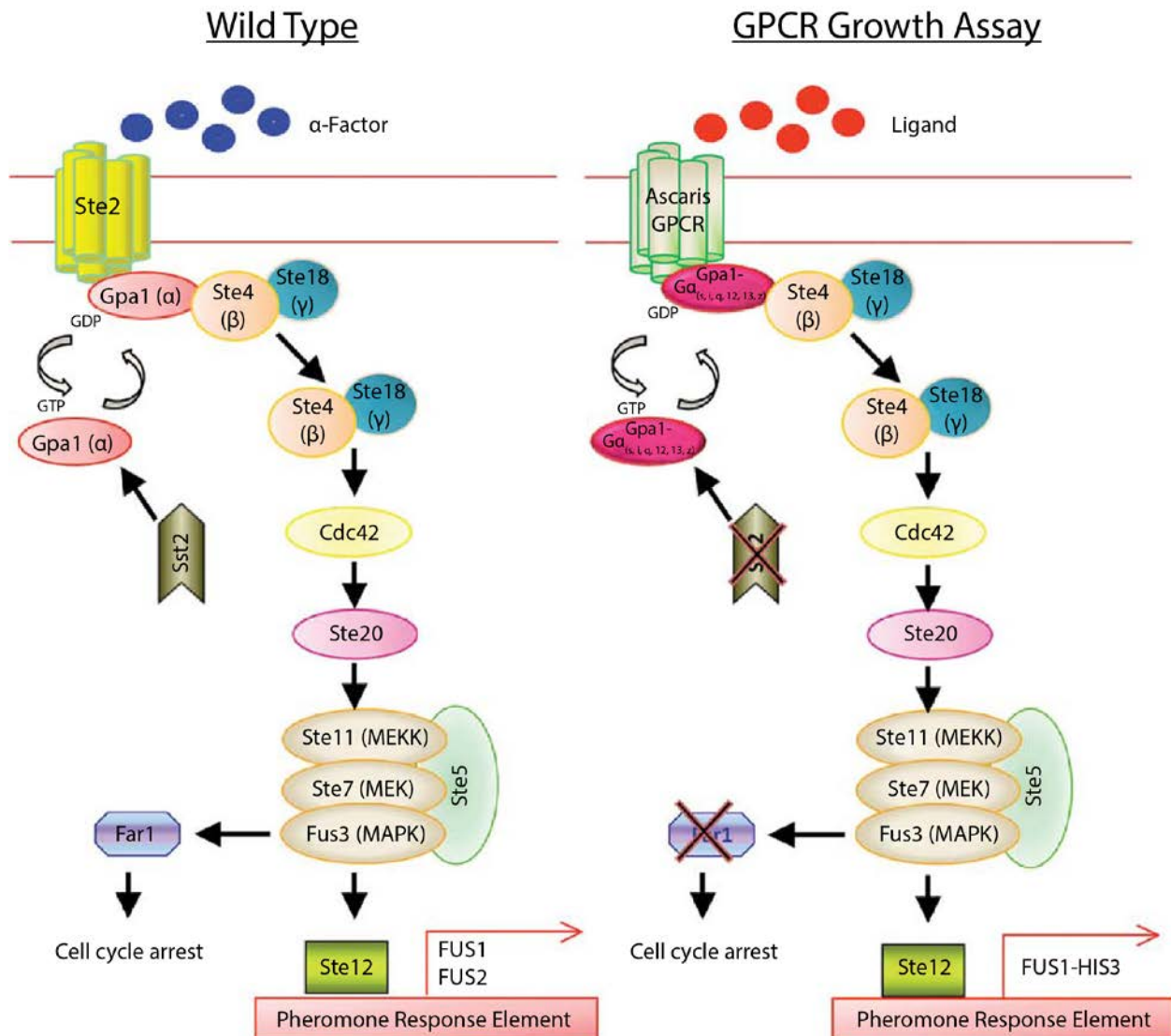


Figure 23. Yeast pheromone response pathway. The endogenous mating pheromone (peptide) GPCR (Ste2) in *Saccharomyces cerevisiae* is replaced with the *A. suum* candidate receptor of interest. The receptor couples to a chimeric G α subunit (type of G α is determined by yeast strain). Ligand binding to the candidate GPCR activates a MAP kinase signaling cascade, resulting in transcription of *FUS1-HIS3* to allow growth in histidine-deficient medium. Null mutations have been introduced into regulatory elements Sst2 and Far1 that would normally desensitize the signaling cascade. Figure adapted from Dong *et al.* (2010) [49].

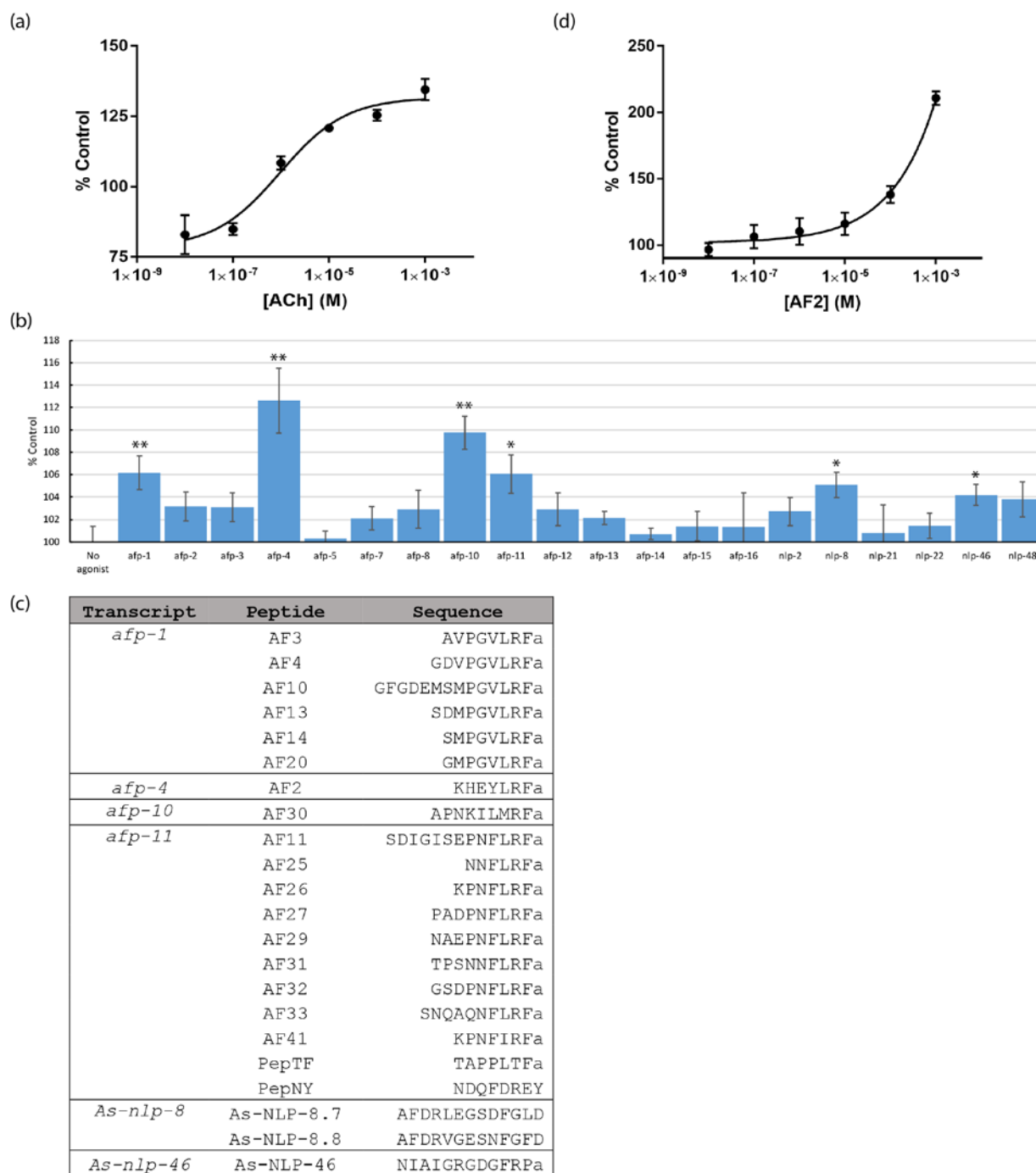


Figure 24. Proof-of-concept of yeast assay with *A. suum* muscarinic acetylcholine receptor AsGAR-1. (a) The AsGAR-1 receptor expressed in YEX108 (G_q) displayed increased growth in the presence of ACh (36 hrs), in agreement with Kimber *et al.* (2009). Dose response placed the EC_{50} at $0.9 \pm 0.3 \mu M$. (b) AsGAR-1-expressing yeast displayed significantly increased growth when exposed to *afp-1*, *afp-4*, *afp-10*, *afp-11*, *nlp-8*, and *nlp-46* peptides (48 hrs). * $P < 0.05$, ** $P < 0.01$. (c) Table of peptide sequences from significant trials in (b). (d) Preliminary dose response of AF2 (36 hrs).

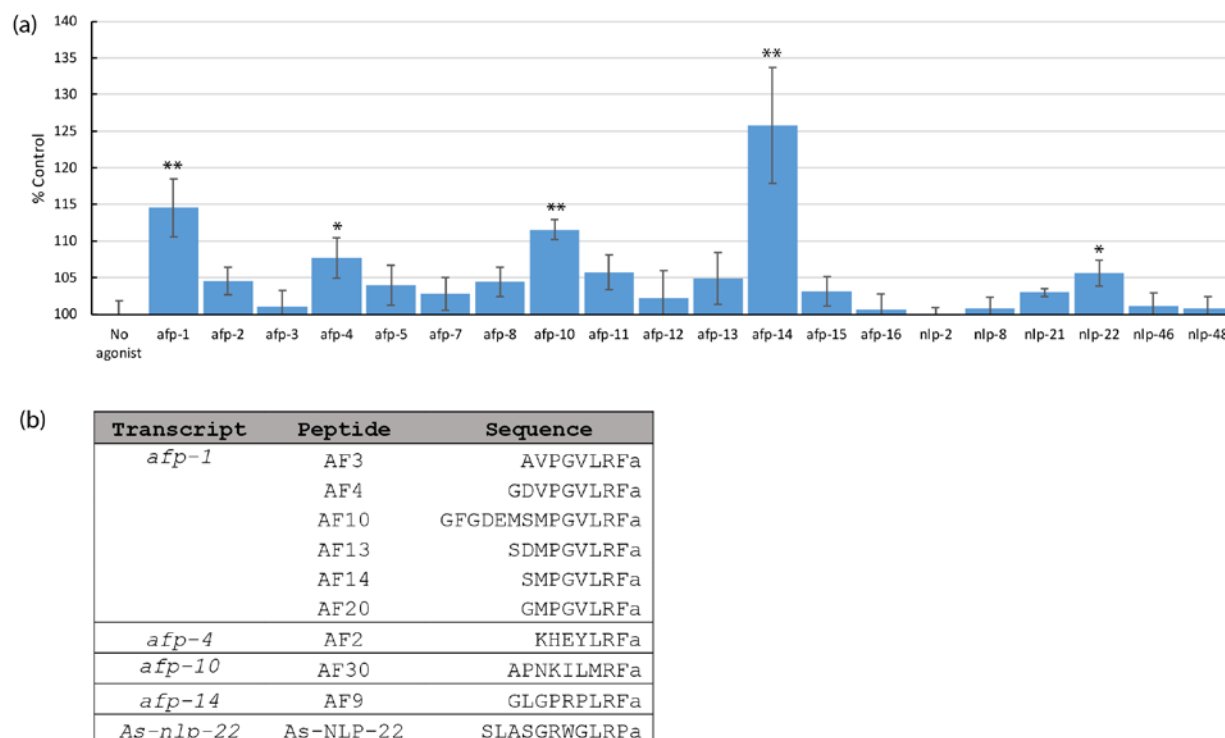


Figure 25. Deorphanization of NPR-1 receptor in yeast. (a) NPR-1 expressed in YEX108 (G_q) displayed increased growth in the presence of *afp-1*, *afp-4*, *afp-10*, *afp-14*, and *As-nlp-22* peptides (48 hrs). * $P < 0.05$, ** $P < 0.01$. (b) Table of peptide sequences from significant transcripts in (a).

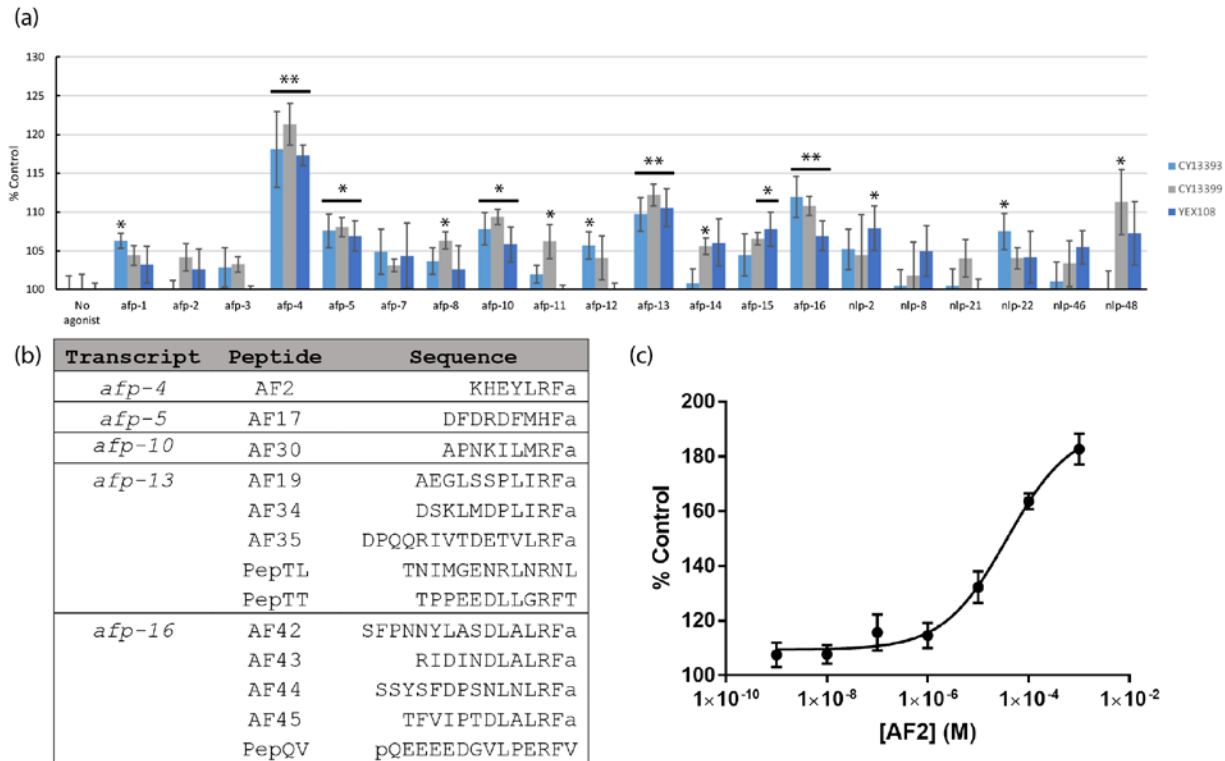


Figure 26. Deorphanization of NPR-4 receptor in yeast. (a) NPR-4 expressed in all strains (CY13393, G_i ; CY13399, G_s ; YEX108, G_q) exhibited increased growth in the presence of *afp-4*, *afp-5*, *afp-10*, *afp-13*, and *afp-16* peptides (48 hrs). A number of other peptides increased growth in a single strain. $*P < 0.05$, $**P < 0.01$. (b) Table of peptide sequences from transcripts that produced significantly increased growth in all strains. (c) Dose response of AF2 (*afp-4*) in CY13393 places EC_{50} at $37 \pm 10 \mu M$ (42 hrs).

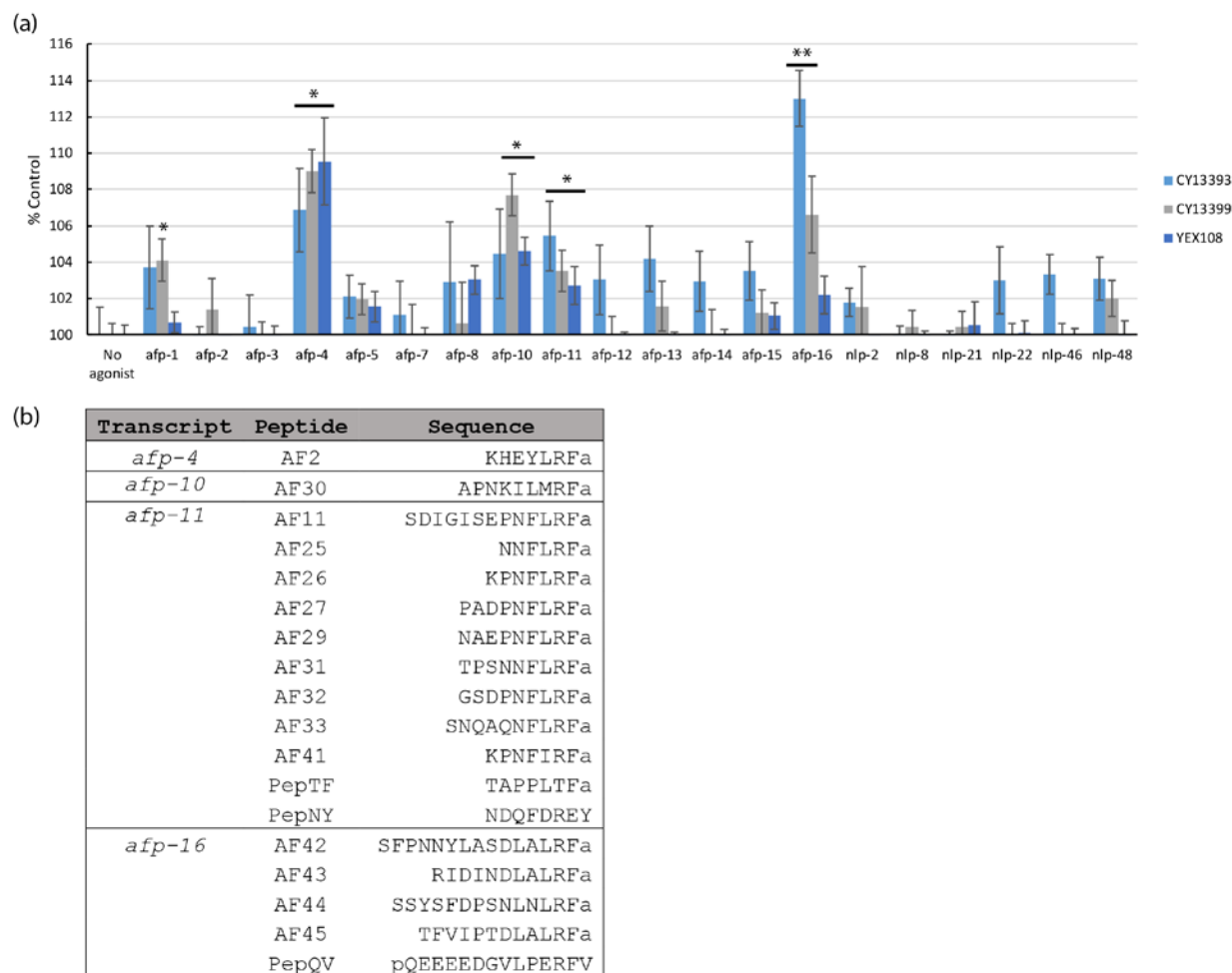


Figure 27. Deorphanization of NPR-11 receptor in yeast. (a) NPR-11 expressed in all strains (CY13393, G_i; CY13399, G_s; YEX108, G_q) exhibited increased growth in the presence of *afp-4*, *afp-10*, and *afp-16* peptides (48 hrs). * $P < 0.05$, ** $P < 0.01$. (b) Table of peptide sequences from significant transcripts in (a).

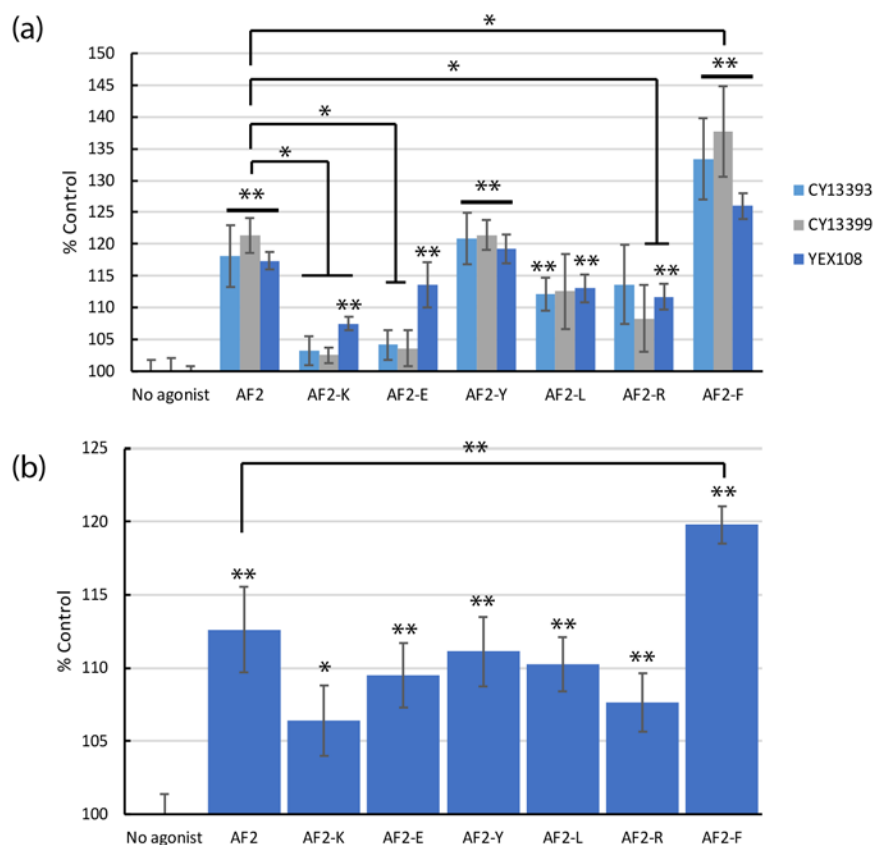


Figure 28. Chemical analogs of AF2 had varying effects based on the position of the β -amino acid substitution. AF2 (KHEYLRFa) and analogs with β -amino acid substitution of the indicated residue were tested in (a) NPR-4 and (b) AsGAR-1- expressing yeast.

REFERENCES

1. Konop, C. J., Knickelbine, J. J., Sygulla, M. S., Vestling, M. M., Stretton, A. O. (2015). Different neuropeptides are expressed in different functional subsets of cholinergic excitatory motoneurons in the nematode *Ascaris suum*. *ACS Chem Neurosci*. 6:855-70.
2. Konop, C. J., Knickelbine, J. J., Sygulla, M. S., Wruck, C. D., Vestling, M. M., Stretton, A. O. (2015). Mass spectrometry of single GABAergic somatic motoneurons identifies a novel inhibitory peptide, As-NLP-22, in the nematode *Ascaris suum*. *J Am Soc Mass Spectrom*. 26:2009-23.
3. Reinitz, C. A., Pleva, A. E., Stretton, A. O. (2011). Changes in cyclic nucleotides, locomotory behavior, and body length produced by novel endogenous neuropeptides in the parasitic nematode *Ascaris suum*. *Mol Biochem Parasitol*. 180:27-34.
4. Bargmann, C. I. (1998). Neurobiology of the *Caenorhabditis elegans* Genome. *Science*. 282:2028-33.
5. Keating, C. D., Kriek, N., Daniels, M., Ashcroft, N. R., Hopper, N. A., Siney, E. J., Holden-Dye, L., Burke, J. F. (2003). Whole-Genome Analysis of 60 G Protein-Coupled Receptors in *Caenorhabditis elegans* by Gene Knockout with RNAi. *Current Biology*. 13:1715-20.
6. Frooninckx, L., Van Rompay, L., Temmerman, L., Van Sinay, E., Beets, I., Janssen, T., Husson, S. J., Schoofs, L. (2012). Neuropeptide GPCRs in *C. elegans*. *Front Endocrinol*. 3:167.
7. Yew, J. Y., Kutz, K. K., Dikler, S., Messinger, L., Li, L., Stretton, A. O. (2005). Mass spectrometric map of neuropeptide expression in *Ascaris suum*. *J Comp Neurol*. 488:396-413.
8. Jex, A. R., Liu, S., Li, B., Young, N. D., Hall, R. S., Li, Y., Yang, L., Zeng, N., Xu, X., Xiong, Z., Chen, F., Wu, X., Zhang, G., Fang, X., Kang, Y., Anderson, G. A., Harris, T. W., Campbell, B. E., Vlaminck, J., Wang, T., Cantacessi, C., Schwarz, E. M., Ranganathan, S., Geldhof, P., Nejsun, P., Sternberg, P. W., Yang, H., Wang, J., Gasser, R. B. (2011). *Ascaris suum* draft genome. *Nature*. 479:529-33.
9. McCoy, C. J., Atkinson, L. E., Zamanian, M., McVeigh, P., Day, T. A., Kimber, M. J., Marks, N. J., Maule, A. G., Mousley, A. (2014). New insights into the FLPerGic complements of parasitic nematodes: Informing deorphanisation approaches. *EuPA Open Proteom*. 3:262-72.
10. Kubiak, T. M., Larsen, M. J., Zantello, M. R., Bowman, J. W., Nulf, S. C., Lowery, D. E. (2003). Functional annotation of the putative orphan *Caenorhabditis elegans* G-protein-coupled receptor C10C6.2 as a FLP15 peptide receptor. *J Biol Chem*. 278:42115-20.

11. Mertens, I., Clinckspoor, I., Janssen, T., Nachman, R., Schoofs, L. (2006). FMRFamide related peptide ligands activate the *Caenorhabditis elegans* orphan GPCR Y59H11AL.1. *Peptides*. 27:1291-96.
12. Kimber, M. J., Sayegh, L., El-Shehabi, F., Song, C., Zamanian, M., Woods, D. J., Day, T. A., Ribeiro, P. (2009). Identification of an *Ascaris* G protein-coupled acetylcholine receptor with atypical muscarinic pharmacology. *Int J Parasitol*. 39:1215-22.
13. Zhang, Y., Chou, J. H., Bradley, J., Bargmann, C. I., Zinn, K. (1997). The *Caenorhabditis elegans* seven-transmembrane protein ODR-10 functions as an odorant receptor in mammalian cells. *Proc Natl Acad Sci U S A*. 94:12162-67.
14. Larsen, M. J., Lancheros, E. R., Williams, T., Lowery, D. E., Geary, T. G., Kubiak, T. M. (2013). Functional expression and characterization of the *C. elegans* G-protein-coupled FLP-2 Receptor (T19F4.1) in mammalian cells and yeast. *Int J Parasitol*. 3:1-7.
15. Tehseen, M., Dumancic, M., Briggs, L., Wang, J., Berna, A., Anderson, A., Trowell, S. (2014). Functional Coupling of a Nematode Chemoreceptor to the Yeast Pheromone Response Pathway. *PLOS ONE*. 9:e111429.
16. Wang, Z.-x., Broach, J. R., Peiper, S. C. (2006). Functional Expression of CXCR4 in *Saccharomyces cerevisiae* in the Development of Powerful Tools for the Pharmacological Characterization of CXCR4. 332:115-27.
17. Checco, J. W., Lee, E. F., Evangelista, M., Sleebs, N. J., Rogers, K., Pettikiriachchi, A., Kershaw, N. J., Eddinger, G. A., Belair, D. G., Wilson, J. L., Eller, C. H., Raines, R. T., Murphy, W. L., Smith, B. J., Gellman, S. H., Fairlie, W. D. (2015). Alpha/beta-Peptide Foldamers Targeting Intracellular Protein-Protein Interactions with Activity in Living Cells. *J Am Chem Soc*. 137:11365-75.
18. Cheloha, R. W., Maeda, A., Dean, T., Gardella, T. J., Gellman, S. H. (2014). Backbone modification of a polypeptide drug alters duration of action in vivo. *Nat Biotechnol*. 32:653-5.
19. Reinitz, C. A., Herfel, H. G., Messinger, L. A., Stretton, A. O. (2000). Changes in locomotory behavior and cAMP produced in *Ascaris suum* by neuropeptides from *Ascaris suum* or *Caenorhabditis elegans*. *Mol Biochem Parasitol*. 111:185-97.
20. Ma, X., Zhu, Y., Li, C., Shang, Y., Meng, F., Chen, S., Miao, L. (2011). Comparative transcriptome sequencing of germline and somatic tissues of the *Ascaris suum* gonad. *BMC Genomics*. 12:481.

21. Kumar, S., Stecher, G., Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol.* 33:1870-4.
22. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* 25:1189-91.
23. Saitou, N., Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 4:406-25.
24. Zuckerkandl, E., Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In: Bryson, V. and Vogel, H.J., Eds., *Evolving Genes and Proteins*, Academic Press, New York, 97-166.
25. Jones, D. T., Taylor, W. R., Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci.* 8:275-82.
26. Richmond, J. E., Davis, W. S., Jorgensen, E. M. (1999). UNC-13 is required for synaptic vesicle fusion in *C. elegans*. *Nat Neurosci.* 2:959-64.
27. Richmond, J. E., Weimer, R. M., Jorgensen, E. M. (2001). An open form of syntaxin bypasses the requirement for UNC-13 in vesicle priming. *Nature.* 412:338-41.
28. Stretton, A. O., Fishpool, R. M., Southgate, E., Donmoyer, J. E., Walrond, J. P., Moses, J. E., Kass, I. S. (1978). Structure and physiological activity of the motoneurons of the nematode *Ascaris*. *Proc Natl Acad Sci USA.* 75:3493-7.
29. Kubiak, T. M., Larsen, M. J., Nulf, S. C., Zantello, M. R., Burton, K. J., Bowman, J. W., Modric, T., Lowery, D. E. (2003). Differential Activation of "Social" and "Solitary" Variants of the *Caenorhabditis elegans* G Protein-coupled Receptor NPR-1 by Its Cognate Ligand AF9. *J Biol Chem.* 278(36):33724-29.
30. de Bono, M., Bargmann, C. I. (1998). Natural Variation in a Neuropeptide Y Receptor Homolog Modifies Social Behavior and Food Response in *C. elegans*. *Cell.* 94:679-89.
31. Rogers, C., Reale, V., Kim, K., Chatwin, H., Li, C., Evans, P., de Bono, M. (2003). Inhibition of *Caenorhabditis elegans* social feeding by FMRFamide-related peptide activation of NPR-1. *Nat Neurosci.* 6:1178-85.
32. Cohen, M., Reale, V., Olofsson, B., Knights, A., Evans, P., de Bono, M. (2009). Coordinated regulation of foraging and metabolism in *C. elegans* by RFamide neuropeptide signaling. *Cell Metab.* 9:375-85.
33. Hobert, O. (2013). The neuronal genome of *Caenorhabditis elegans*. *WormBook.* 1-106.

34. Janssen, T., Lindemans, M., Meelkop, E., Temmerman, L., Schoofs, L. (2010). Coevolution of neuropeptidergic signaling systems: from worm to man. *Ann N Y Acad Sci.* 1200:1-14.
35. Chalasani, S. H., Kato, S., Albrecht, D. R., Nakagawa, T., Abbott, L. F., Bargmann, C. I. (2010). Neuropeptide feedback modifies odor-evoked dynamics in *Caenorhabditis elegans* olfactory neurons. *Nat Neurosci.* 13:615-21.
36. Jarecki, J. L., Viola, I. R., Andersen, K. M., Miller, A. H., Ramaker, M. A., Vestling, M. M., Stretton, A. O. (2013). Three independent techniques localize expression of transcript *afp-11* and its bioactive peptide products to the paired AVK neurons in *Ascaris suum* : *In situ* hybridization, immunocytochemistry, and single cell mass spectrometry. *ACS Chem Neurosci.* 4:418-34.
37. Davis, R. E., Stretton, A. O. (2001). Structure-activity relationships of 18 endogenous neuropeptides on the motor nervous system of the nematode *Ascaris suum*. *Peptides.* 22:7-23.
38. Jarecki, J. L., Andersen, K., Konop, C. J., Knickelbine, J. J., Vestling, M. M., Stretton, A. O. (2010). Mapping neuropeptide expression by mass spectrometry in single dissected identified neurons from the dorsal ganglion of the nematode *Ascaris suum*. *ACS Chem Neurosci.* 1:505-19.
39. Park, Y. S., Lee, Y. S., Cho, N. J., Kaang, B. S. (2000). Alternative splicing of GAR-1a, a *Caenorhabditis elegans* G protein-linked acetylcholine receptor gene. *Biochem Biophys Res Commun.* 268:354-58.
40. Davis, R. E., Stretton, A. O. W. (2001). Structure-activity relationships of 18 endogenous neuropeptides on the motor nervous system of the nematode *Ascaris suum*. *Peptides.* 22:7-23.
41. Verma, S., Robertson, A. P., Martin, R. J. (2007). The nematode neuropeptide, AF2 (KHEYLRN-NH₂), increases voltage-activated calcium currents in *Ascaris suum* muscle. *Br J Pharmacol.* 151:888-99.
42. Viola, I. R. (2011). Comparing the cellular localization of *Ascaris* neuropeptides using immunocytochemistry, molecular biology, and mass spectrometry. Ph. D. Thesis. University of Wisconsin-Madison.
43. Jarecki, J. L., Frey, B. L., Smith, L. M., Stretton, A. O. (2011). Discovery of neuropeptides in the nematode *Ascaris suum* by database mining and tandem mass spectrometry. *J Proteome Res.* 10:3098-106.

44. Kubiak, T. M., Larsen, M. J., Davis, J. P., Zantello, M. R., Bowman, J. W. (2003). AF2 interaction with *Ascaris suum* body wall muscle membranes involves G-protein activation. *Biochem Biophys Res Comm.* 301:456-59.
45. Cowden, C., Stretton, A. O. (1993). AF2, an *Ascaris* neuropeptide: isolation, sequence, and bioactivity. *Peptides.* 14:423-30.
46. Davis, R. E., Stretton, A. O. W. (1989). Passive Membrane-Properties of Motorneurons and Their Role in Long-Distance Signaling in the Nematode *Ascaris*. *J Neurosci.* 9:403-14.
47. Davis, R. E., Stretton, A. O. (1989). Signaling properties of *Ascaris* motorneurons: graded active responses, graded synaptic transmission, and tonic transmitter release. *J Neurosci.* 9:415-25.
48. Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution.* 39:783-91.
49. Dong, S., Rogan, S. C., Roth, B. L. (2010). Directed molecular evolution of DREADDs: a generic approach to creating next-generation RASSLs. *Nat Protoc.* 5:561-73.

SUMMARY AND CONCLUSIONS

This work describes significant advances in the field of nematode neuropeptide signaling. In chapters 2 and 4, we identified the unique complement of neuropeptides produced by each class of *A. suum* somatic motoneuron using two independent techniques: single-cell mass spectrometry and *in situ* hybridization. We found that the motoneurons could be grouped into 3 categories based on neuropeptide expression, and these groups also correlated with the anatomy and function of the motoneurons: anteriorly-projecting excitatory motoneurons, responsible for forward locomotion, expressed peptides from the *afp-14* and *nlp-21* transcripts; posteriorly-projecting excitors, responsible for backward locomotion, expressed *afp-1* and *afp-4*; and inhibitory motoneurons expressed *nlp-22*. Importantly, the expression of these neuropeptides differed from what has been reported in *C. elegans*, suggesting that parasitic and free-living nematodes may use conserved neuropeptide sequences for different functions. Further, nearly all the peptides found in these neurons had strong effects on muscle contraction and/or locomotory behavior. These experiments have achieved our laboratory's goal for the last 25 years, and provide a defined set of neuromodulators that are likely to be part of the locomotory functional circuit.

In chapter 3, we further explored the function of NLP-22 (the only neuropeptide expressed in the inhibitory motoneurons) by testing the effects of sequence-related peptides from the -RPamide family. We found that differences in two amino acid positions (4 and 8) seemed to correlate with changes in function: the presence of hydrophobic Leu or Ile (instead of Ser) at position 4 was associated with later potentiation of ACh-induced muscle contraction, while small uncharged amino acids (Ser, Gly) instead of negatively charged Asp at position 8

was associated with strong early inhibition of muscle contraction. These results suggest that there could be two receptors with very similar ligand-binding domains mediating these effects, with specificity determined by positions 4 and 8.

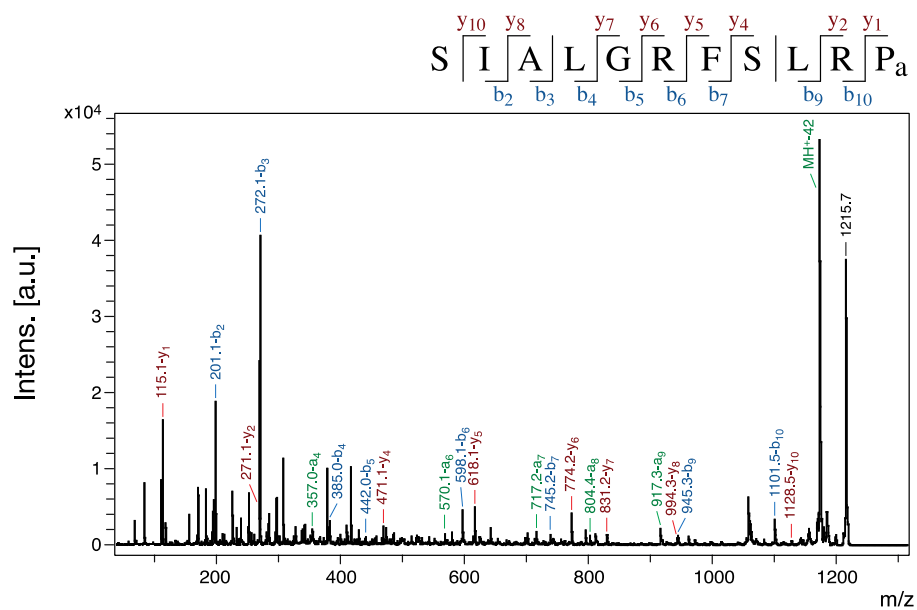
Chapter 5 reported the transcript identification and phylogeny of 7 *A. suum* neuropeptide GPCRs that were tied to the motoneurons by preliminary single-cell RNA-seq analysis. We heterologously expressed 3 of these, in addition to a muscarinic ACh receptor, in yeast to determine their neuropeptide ligands by reverse pharmacology. We found that all receptors bound multiple neuropeptides, with some being specific to one receptor/strain and others binding to all receptors tested. Again, the expression and deorphanization data reported here differs from that of homologs in *C. elegans*, highlighting the need for additional studies in parasitic nematodes. These experiments are the first to deorphanize a neuropeptide receptor from *A. suum*, and provide a protocol for testing additional GPCRs as they are identified.

Finally, this dissertation adds many possible components to the nematode functional circuit for locomotion. There is a great potential for some of these to become useful targets for therapeutic development, since the neuropeptides discussed here have dramatic effects on the worm but do not seem to be conserved in vertebrates. Drugs that affect these signaling pathways might provide relief to human and animal hosts without the harmful side effects of drugs that target more strongly-conserved cholinergic and GABAergic pathways. Much more work will be needed to develop such compounds, but as we continue to piece together the behavioral circuit, we lay the foundation for these efforts.

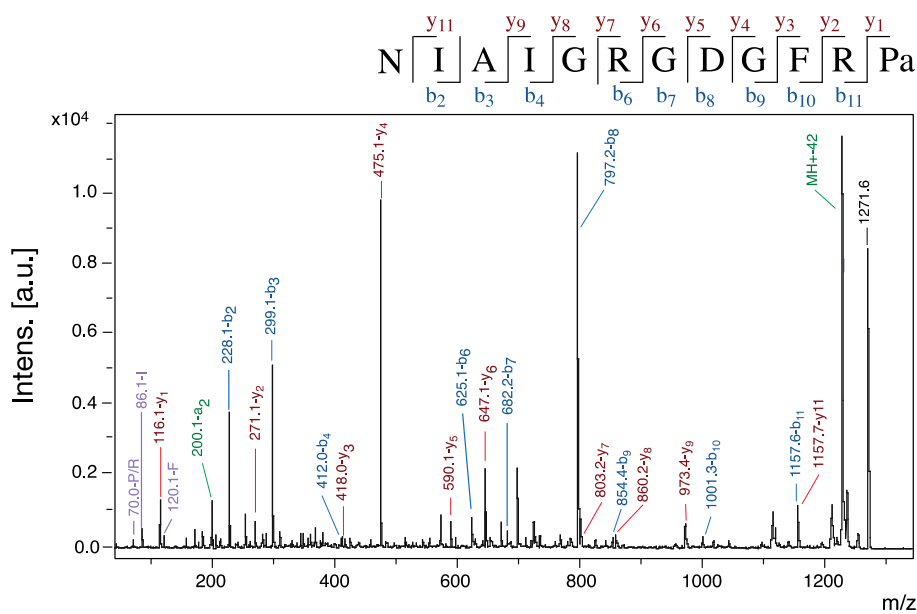
APPENDIX 1

Supplementary Figures for Chapter 2

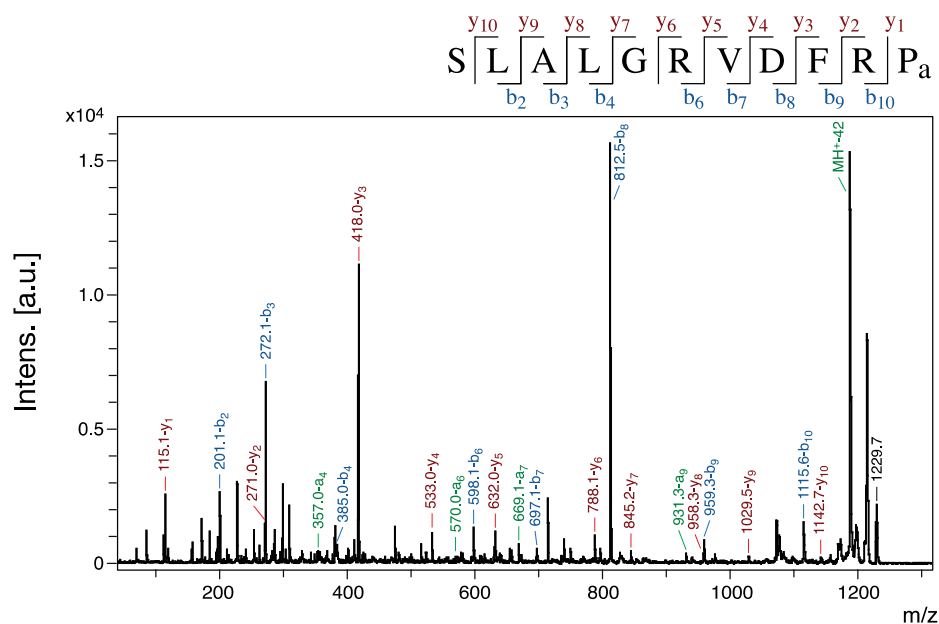
(a)



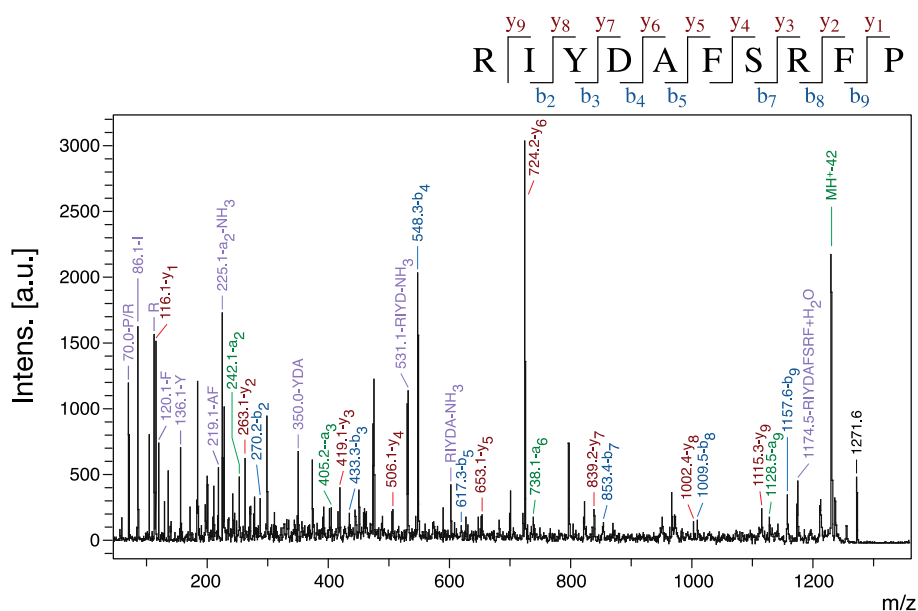
(b)



(c)



(d)

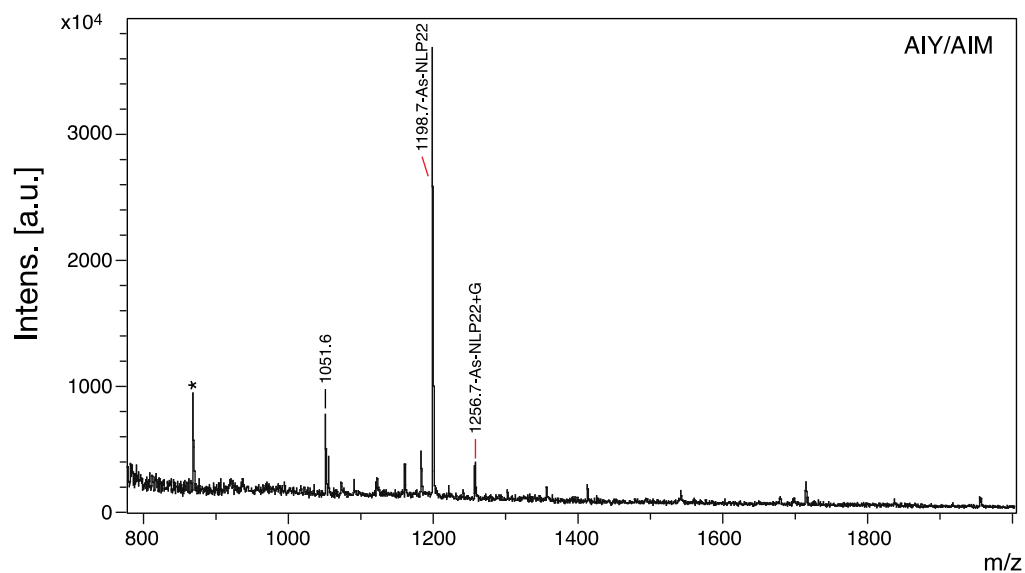


Supplementary Figure 1. MS/MS of synthetic peptides with a common C-terminal proline residue similar to As-NLP-22. Peaks representing a (green), b (blue), y (red), and high-intensity internal fragment (purple) ions are labeled, and b and y ions are summarized in the sequence at the top of each spectrum. (a) MS/MS of synthetic As-NLP-23. (b) MS/MS of synthetic As-NLP-46. (c) MS/MS of synthetic As-NLP-2.2. (d) MS/MS of synthetic As-NLP-21.6. Analysis of these spectra identified a common intense ion corresponding to MH+42 (green).

(a)



(b)



Supplementary Figure 2. Expression of *As-nlp-22* in the ventral ganglion. (a) Pair of cells in the posterior ventral ganglion, stained with the *As-nlp-22* riboprobe. These cells are either the AIY or AIM neurons, which are morphologically indistinguishable in *A. suum*. Scale bar: 100 μm . (b) Mass spectrum from a single AIY or AIM neuron from the ventral ganglion containing peaks with m/z 1198.7 and 1256.7.

C.e. BLAST results

>Cbg *Caenorhabditis briggsae* CGG79409_1

MAQLVVFVCIMAAFAFVASDQSIYDQEEGTNNFRFVGFGGPERVPPFESLEGVLRRLHLRALPMM**KRS**IAIGRSGFRPG
KRSMEIFAF

>Cbr *Caenorhabditis brenneri* CBG91628_1

MNSLLLLFLCVAILAVKAVPIREQISGEDTGTPSVQGEFEFYNDHRVFDAESDQKDLLHFTDLQVLPMM**KK**SIAIGRAGF
RPGKRAVIDISDF

>Cel *Caenorhabditis elegans* CEG82481_1

MRSIIVFIGLTIFALDILLVQTSALGLQGGIDVFRGLGVVDQVDFNQILHRANYLRNTREGRLRYWRLRTLPI**MM**SIAIGR
AGFRPG**KRT**DELTFPIGV

>Cre *Caenorhabditis remanei* CRG82931_1

MPRLMLFFCIALIANQGIALHQPVYDQDDDINVVQSAGGFHLVNEDRVISVEPVEARLSFERLRALPIM**KK**SIAIGRAGF
RPG**KRT**VDIYDF

>Sst *Strongyloides stercoralis* SSP05227_1

MVRSFISIFILFITMYCEEENDNLESDDPSIFM**KRS**LAIGRMGFRP

A.s. BLAST results

>Aav *Aphelenchus avenae* GO481772

MSFVLSNVLLRYLLCAVVVCRALPVNPSEPKLFAILERDPFTMPTSPEPEEYADAAVHQPSYAAADLTSRV**RR**SLASGR
WGLRPG**KR**SSYLPSPTWEEEQGPVPDSVGIRLKRSGSDQKCICFFCDKCSC

>Asu *Ascaris suum* BI593877

MRSLAVLVFSIIVDVVYPSPLYVAGPSISK**RS**LASGRWGLRPG**KRS**QDVPIYTDELGLDGNSPYDALRQSHFVYVVRK

>Gro *Globodera rostochiensis* EE266657

MTSLPLLTLSVLLALQLIGQSFADFEQVEDQDLLNPQQLTIDSNRN**IR**SLANGRWQLRPG**KR**ASLIDFHPMDGLGER
ARRFRNLYALLPPLNNWN

>Mar *Meloidogyne arenaria* CF358247

MRFSIALIVFAFILGIFNCAESKDEGGNQEEIEQQYQNRRL**RS**LANGRWQLRPG**KR**FVPENYYYQMMLDN

>Mch *Meloidogyne chitwoodi* CB930549

MRFLTIFVLILVFITILGVQLKEENNVIKSWRVEQQLYQNRRL**RS**LANGRWQLRPG**KR**FSPENYYYQMMLDN

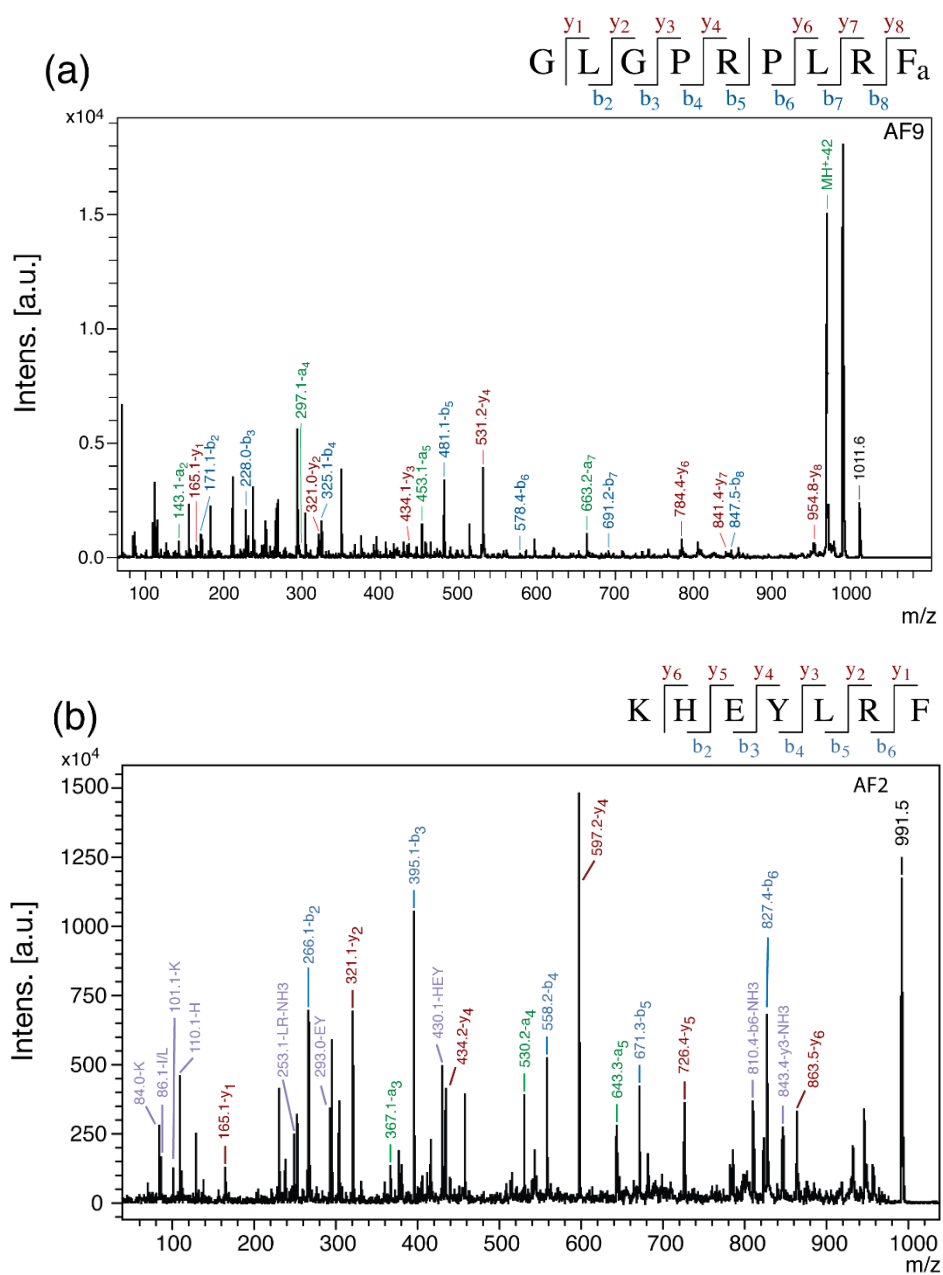
>Min *Meloidogyne incognita* AW783047

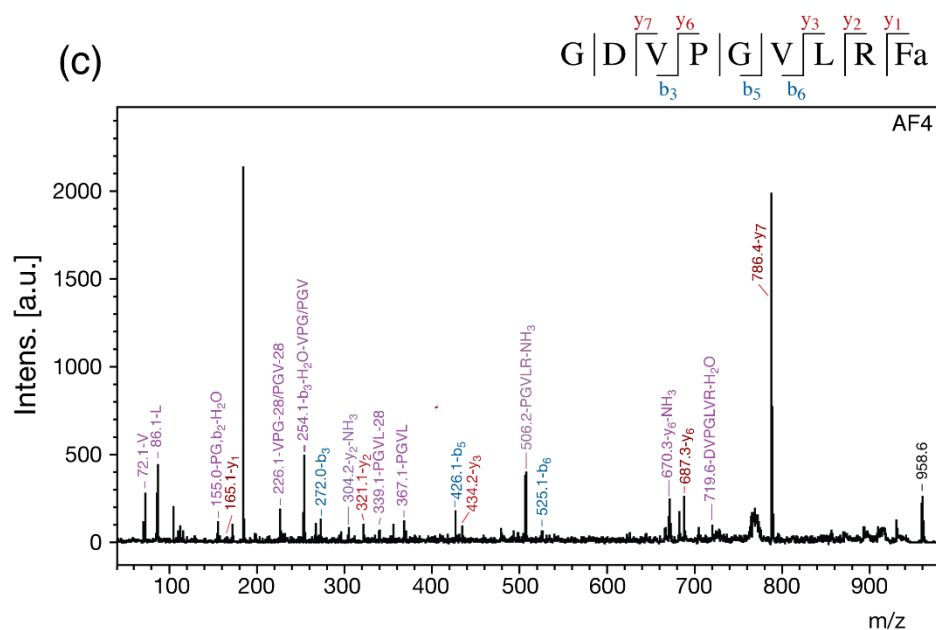
MRFSIALIVFAFILVIFNCAESKDEGGNQEEIEQQYQNRRL**RS**LANGRWQLRPG**KR**FVPENYYYQMMLDN

Supplementary Figure 3. Results from BLAST searches of nematode EST databases using both Ce-NLP-22 and As-NLP-22 as queries. In the title line for each sequelog is the three-letter identifier for each species and the GenBank accession number. Green represents a predicted signal peptide (SignalP 4.0) and processed peptides are in blue. Putative basic cleavage residues are indicated in bold.

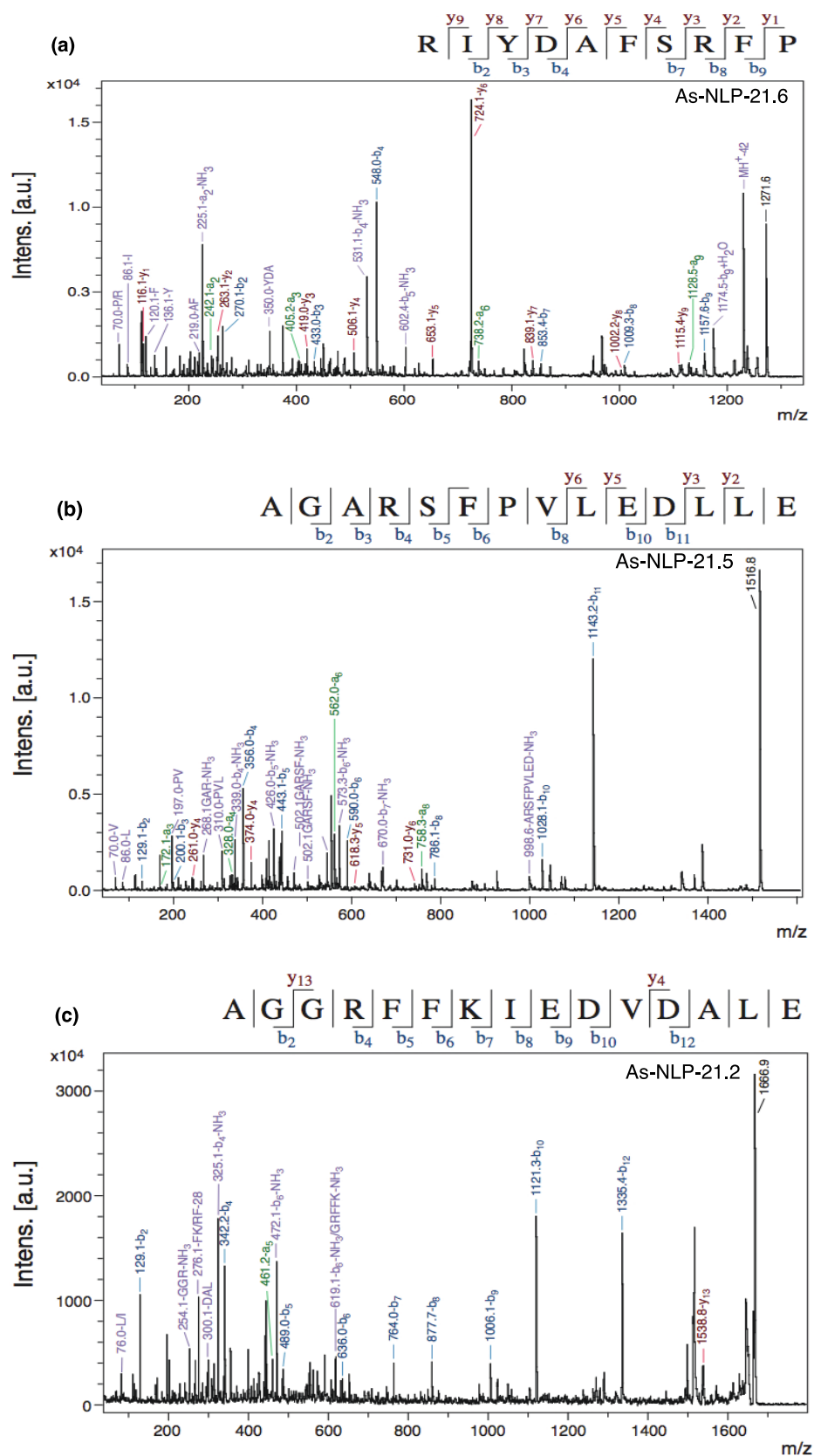
APPENDIX 2

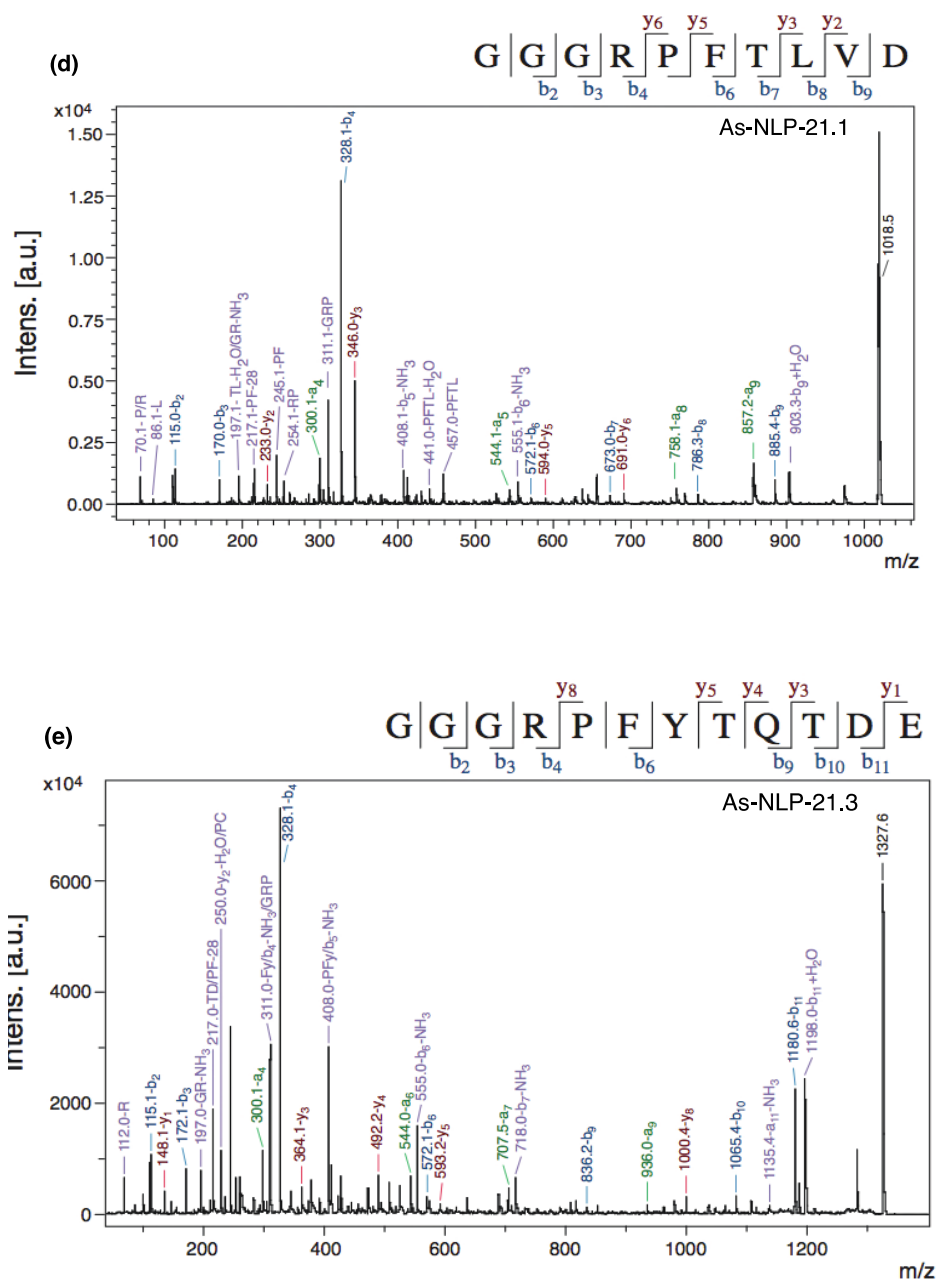
Supplementary Figures for Chapter 4



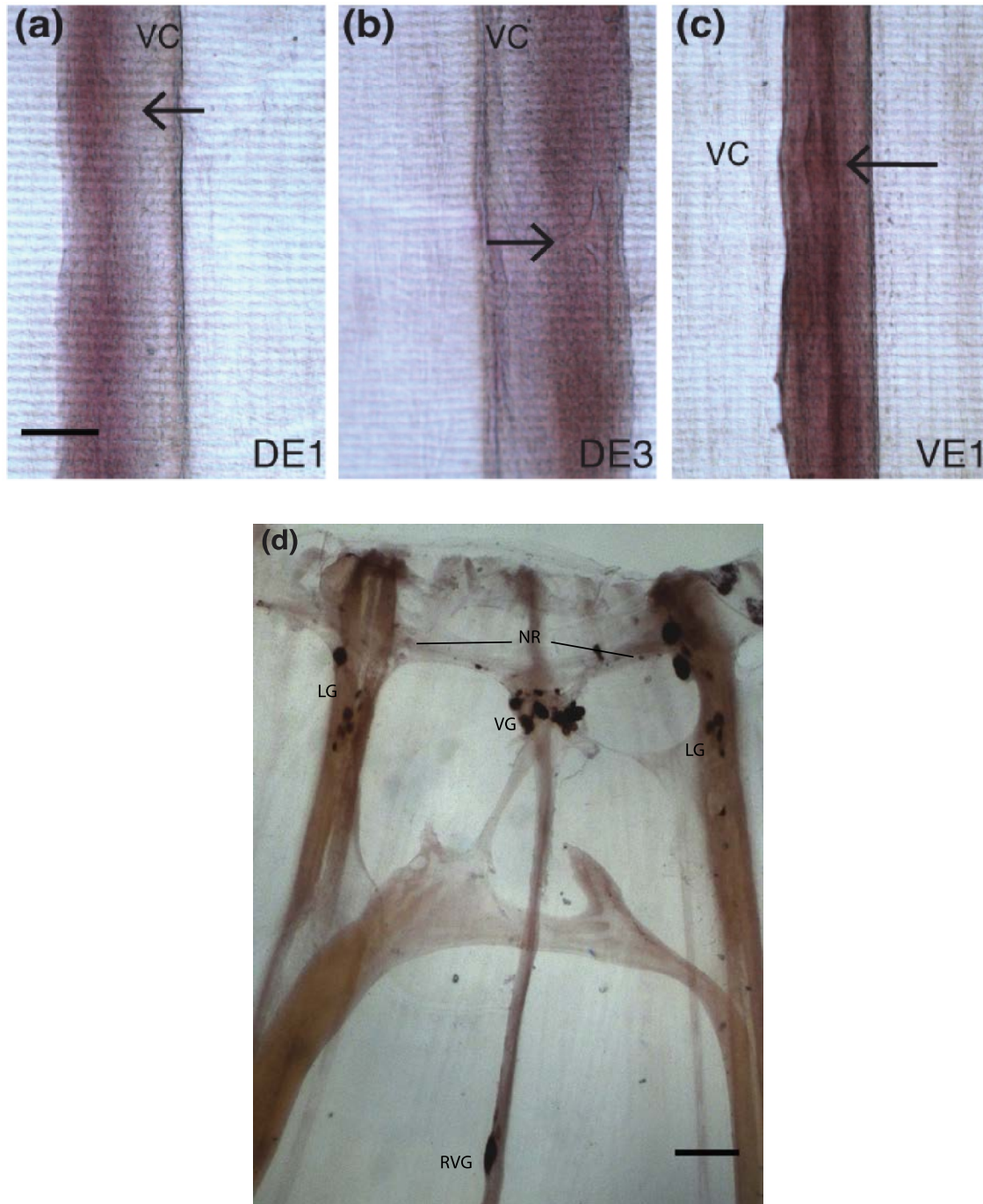


Supplementary Figure 1. MS/MS spectra of synthetic peptides for AF9, AF2, and AF4. Peaks representing a (green), b (blue), y (red), and high-intensity internal fragment and immonium ions (purple) ions are labeled, and b and y ions are summarized in the sequence at the top of each spectrum (a) MS/MS spectrum of AF9 (*afp-14*), m/z 1011.6. (b) MS/MS spectrum of synthetic AF2 (*afp-4*), m/z 991.5. (c) MS/MS spectrum of synthetic AF4 (*afp-1*), m/z 958.6.

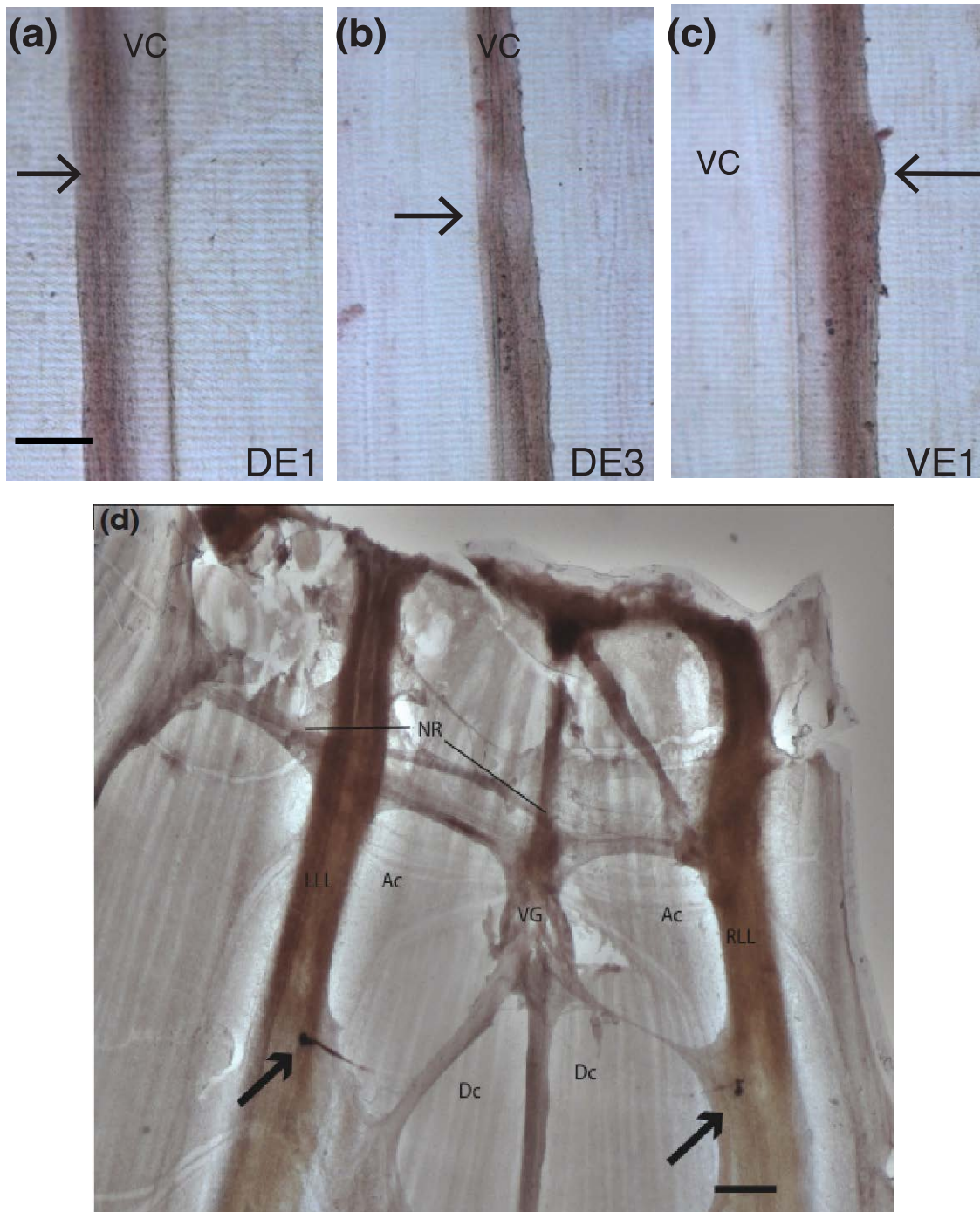




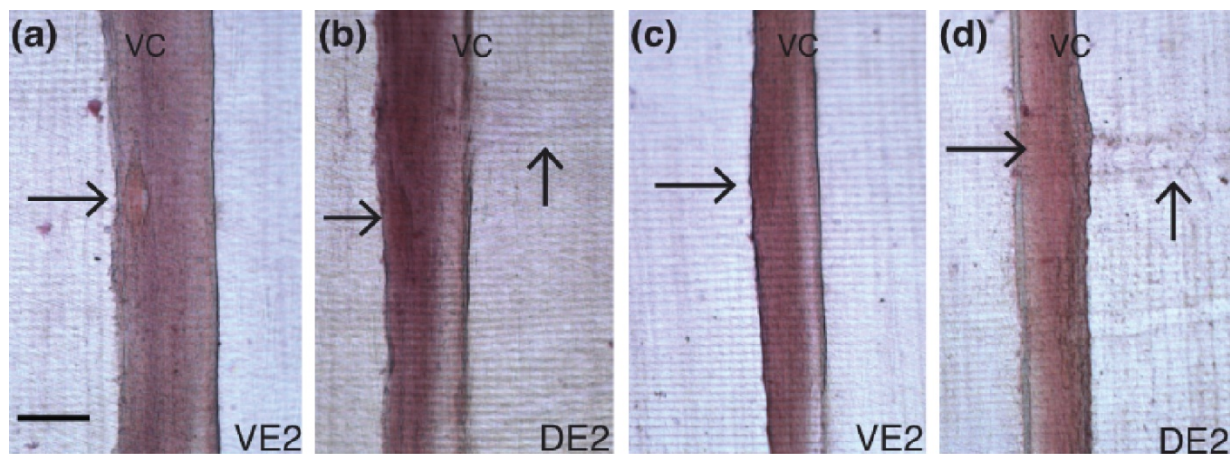
Supplementary Figure 2. MS/MS of synthetic peptides encoded by *As-nlp-21*. Peaks representing a (green), b (blue), y (red), and high-intensity internal fragment and immonium ions (purple) ions are labeled, and b and y ions are summarized in the sequence at the top of each spectrum. The names of the peptides are below the sequences. (a) As-NLP-21.6, m/z 1271.6; (b) As-NLP-21.5, m/z 1516.8; (c) As-NLP-21.2, m/z 1666.9; (d) As-NLP-21.1, m/z 1018.5; (e) As-NLP-21.3, m/z 1327.6.



Supplementary Figure 3. *In situ* hybridization of *As-nlp-21* using a sense probe (negative control, a-c) and antisense probe (experimental, d). No staining is observed in the (a) DE1, (b) DE3, and (c) VE1 motorneurons. Horizontal arrows indicate non-stained cell bodies. VC= Ventral cord. Scale= 100 μ m (d) *In situ* hybridization in the head using an *As-nlp-21* antisense probe. LLL = Left lateral line; RLL = Right lateral line, Dc = Deirid commissure, Ac = Amphidial commissure, NR= Nerve ring, VG= Ventral ganglion. Scale bar= 250 μ m.



Supplementary Figure 4. *In situ* hybridization of *afp-14* using a sense probe (negative control, a-c) and antisense probe (experimental, d). No staining is observed in the (a) DE1, (b) DE3, and (c) VE1 motorneurons. Horizontal arrows indicate non-stained cell bodies. VC= Ventral cord. Scale bar= 100 μ m. (d) *In situ* hybridization in the head using an *afp-14* antisense probe. Arrows indicate stained cell bodies. LLL = Left lateral line; RLL = Right lateral line, Dc = Deirid commissure, Ac = Amphidial commissure, NR= Nerve ring, VG= Ventral ganglion. Scale bar= 250 μ m.



Supplementary Figure 5. *In situ* hybridization of *afp-1* (a,b) and *afp-4* (c,d) using sense probes as negative controls. No staining of VE2 or DE2 motorneurons was observed. Horizontal arrows indicate non-stained cell bodies and vertical indicate commissure of DE2. VC= Ventral cord. Scale=100 μ m.

>Asu *Ascaris suum*; ED417768

MHTVNFALLLVHYGIAQEPDVLEKRGGGRPFTLVDKRAGGRFFKIEDVDALEKRGGGRPFYTQTDEKRGGGRSFRLS
SLGEKRMAAQKIAMPELD FYEGALAEIGNWPFIEKRAGARSPVLEDLLEKRRIYDAFSRFP

>Aav *Aphelenchus avenae*; GO483031.1

AGGRAFAAVPVESELAFDKRAGARPFYVPQSWKRGGGRAFSFGDDFEKRAGGRAFQLNKRGGGRAFYGGGNRFRSF
YGGGGSYPLPYEKKNYGDALYYLSGPNPKRAGARPFADAYLPEQWMDKRAGGRTPIMGD

>Bmu *Bursaphelenchus mucronatus*; CJ977836.

TDRSWPGVVCKIVSI MNLGVRTIDLFLVSLFGILTQIVYGASLVQNLEDAELNEEDKRGGGRAFLNADEKRGGGRAFQF
NDKRGGARAFAPVDVKRGGARAFAPIDYKRGGGRTFSGFNNGNYLLNKRGGGRSFYGGFNYSYAPSFPLSYE

>Bxy *Bursaphelenchus xylophilus*; CJ989167.1

DKRGGGRAFSGFNGGNYLLNKRGGGRSFYGGFNYSYAPSFPLSYEKKADYVPYYLTEMKRAGGRAFQPKYYDPYF

>Dde *Ditylenchus destructor*; JZ126856.1

MTSLLLLSIVLLQSTVSNALIDEQLASSDTPFVDQNLNLENTDEKRAGGRNFLTLPETDKRGGGRNFGAFMNQKRAGGR
QFASAFSERWSPLMGQA KRGGGRNFGEMLDEKRAGGRGFQMLDEETTVEDKRGGARSFYQGSDAKRGGARAFG
GHDSAEKRGGGRTRFMADGSNIFEK

>Gro *Globodera rostochiensis*; BM345727.1

MILHGIDRGHRHRRPPRPSHQPTSPKFLLSIATIPLLILPRIGVATSESDVDALATEQKRGGARAFNFFAPDEKRGGAR
AFYFFAPDEKRGGARAFNFFAPDEKRGGARAFNLFVSDALPSSYEKRSGIQTFRDDYDEKQAGELKRAGGRFLRMADL
PDGDDFVPEGK KRGGARPFYGGGYMDGTWKRAGGRYF

>Hgl *Heterodera glycine*; CB934882.1

QHRRSVHSVRYHSPLLSWPKFLLSIATIPLLLLSCIDFVSAVSDDGAPFLSDQQKRGGARAFNLFASAAFPPNEKRSGG
RAFGGGIQTINDDFDANLAGELKRAGARLFMSDLSSDGDSFVPEGT KRGGARPFYGGGYIDGAWKRAGGRYFLRQF
DDVPFGGGWMAKRGGGRSFEGTKGPFNSANYWAP

>Min *Meloidogyne incognita*; JK304458.1

VSLLLLLIAGTTFEQAPSCLAIENEAIDELLASQFYESSNKRAGARAFQRPDFDDGSYELKRGGARTFLVGEKRGGARAF
AKLEEKRGGARPFYEEKRGGARPFYGGFGGEGTWKRGGGRYFIRPFADQGFSGVGAGGWVKRGGGRQFSADKRA
GGRSFFGSLDSAFDSTNYWVPPRYTPLNKAIIAINDDQVMKSETPF

>Ppe *Pratylenchus penetrans*; BQ626931.1

MQRGGGGQLALAAQQKQHWSSAGGCPQQHREFCSATTVAWCRSLLLLLMAGAA LDQVPSCALDENNGIDKNDHS
SPSFYDQQEEEADLASAGDKRAGGRAFMFPVVGTENEKRGGGRAFLPAAQKRAGARAFQSQGQLQPAATGLDLK
KRDNINGANAFYFSPDEKKRGSARPF

>Aca *Ancylostoma caninum*; ES672867.1

MRAAALVAIAAAAVSHA DDSSIEKRAGGRAFIPIEERGGARAFFEDEKRGGARAFMDKKGGARPFDTLKRGGARAF
MPIEERGGARAFFEDGKRGGARAFQLDKKGGARAFSGIEERGGARAFFEDEKRGGARAFAPMELRGGARAFFEDKR
GGARAFKRGGGRAFVPEKRGGARMFQNAKRDFDEYPEWLSVDDQEPSPFEVDRPVDG

> Ace *Ancylostoma ceylanicum*; CB276019.1

PGCSFPGTRVAIAAAAAVSHA DDNSIEKRAGGRAFPVPIEERGGARAFFEDEKRGGARAFLMDKKGGARPFTDLKRGG
ARAFMPIEERGGARAFFEEDKRGGARAFQLDKKGGARAFIDPKRGGARAFSGIEERGGARAFFEDEKRGGARAFPM
EQRGGARAFFEDKRGGARAFKRGGARAFVPEK

> Cel *Caenorhabditis elegans*; EC035782

MRNSLFTTLFFGLAALVMVLNAQYTSELEEDEKRGGARAMLHKRGGARAFSADVGGDDYKRGGARAFYDEKRGGARA
FLTEMKRGGARVFQGFEDKRGGARAFMMDKRGGGRAFGDMMKRGGARAFVENSKRDEDWVIRPFEDDRLGVF

> Hba *Heterorhabditis bacteriophora*; EX913335.1

YKKVGTLLAIAAATIVSNAQDDLLEKAGEARAFIDDKRGGARVFSHEEKRGGARLFFEDKKGGARSFHDIKRGGARSF
SESKRGGGRPFDTYKRGGARAFMNTDGDGIRTFYEKRGGARAFFDNKRGGARGFSDLLQGGLPTYLIDKRGGARMFI
GNTKRHYMDEYPTWIDYAELEDSFPDVERRAGGRAFRISGGGGF

> Nam *Necator americanus*; BU087648.1

FTDWKRGGARGFSTIELRGGARAFFEDEKRGGARAFKRGGGRAFVLDERGGARMFQNSKRDYTDYDPDWLWYADH
EQSPFEVDENVDTGTF

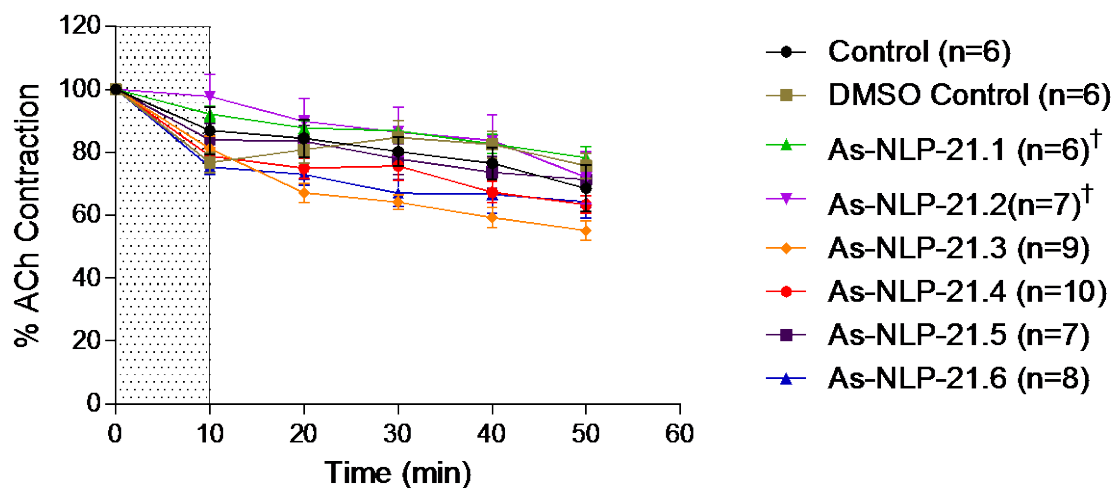
> Nbr *Nippostrongylus brasiliensis*; FE928565.1.

GARAFFEDVEKRGGARAFLDPKRGGARAFFEEKKGGVRAFLSPKRGGVRAFFFERKRGGGRAFPKRGGARAFKRGG
ARMFQSAKRDYMEEYPEWLAYDDQEQSPFEMDRSADVDEQMWKGGGARSFLLGNSAEGGGGY

> Ppa *Pristionchus pacificus*; FG098688.1

LLFVLSLHSFVSEIDIGTSSVSDSLEMEKRGGGRSFASFSGFFPSAKRYSPLYAYYDKRGGGRAYFPVGMDKRAGGR
AYFPTDMDKRAGGRAFPVEEKRGGGRSYAPVEEEKRGGGRSYFPVVKRGGGRAFASMEKRGGGRAFVGDWAFNY
PRMERRGGGRAFSGGWGDYMGRIY

Supplementary figure 6. Results from BLAST searches of nematode EST databases using both Ce-NLP-21 and As-NLP-21. *nlp-21* sequences are predicted from nematode *nlp* precursors. In the title line for each sequelog is the three-letter identifier for each species and the GenBank accession number. Green represents a predicted signal peptide (Signal P 4.1) and processed peptides are in blue. Asu = *Ascaris suum* (Clade III), Aav = *Aphelenchus avenae* (IVb), Bmu = *Bursaphelenchus mucronatus* (IVb), Bxy = *Bursaphelenchus xylophilus* (IVb), Dde = *Ditylenchus destructor* (IVb), Gro = *Globodera rostochiensis* (IVb), Hgl = *Heterodera glycine* (IVb), Min = *Meloidogyne incognita* (IVb), Ppe = *Pratylenchus penetrans* (IVb), Aca = *Ancylostoma caninum* (V), Ace = *Ancylostoma ceylanicum* (V), Cel = *Caenorhabditis elegans* (V), Hba = *Heterorhabditis bacteriophora* (V), Nam = *Necator americanus* (V), Nbr = *Nippostrongylus brasiliensis* (V), Ppa = *Pristionchus pacificus* (V).



Supplementary figure 7. Effects of peptide on ACh-induced muscle contraction. Strength of contraction is normalized to the initial contraction prior to the application of peptide for each individual worm. Error bars show standard error of the mean. Effects due to single exposure to different peptides from the *As-nlp-21* transcript (10 μ M). Peptide was introduced at time 0 and washed out at 10 min (shaded region). The remaining time points show recovery after peptide was washed out. [†] = peptides initially dissolved in 10% DMSO.

Peptide	Purity by HPLC (%)
AF2	93
AF4	99
AF9	84
As-NLP-21.1	91
As-NLP-21.2	85
As-NLP-21.3	82
As-NLP-21.4	95
As-NLP-21.5	95
As-NLP-21.6	90

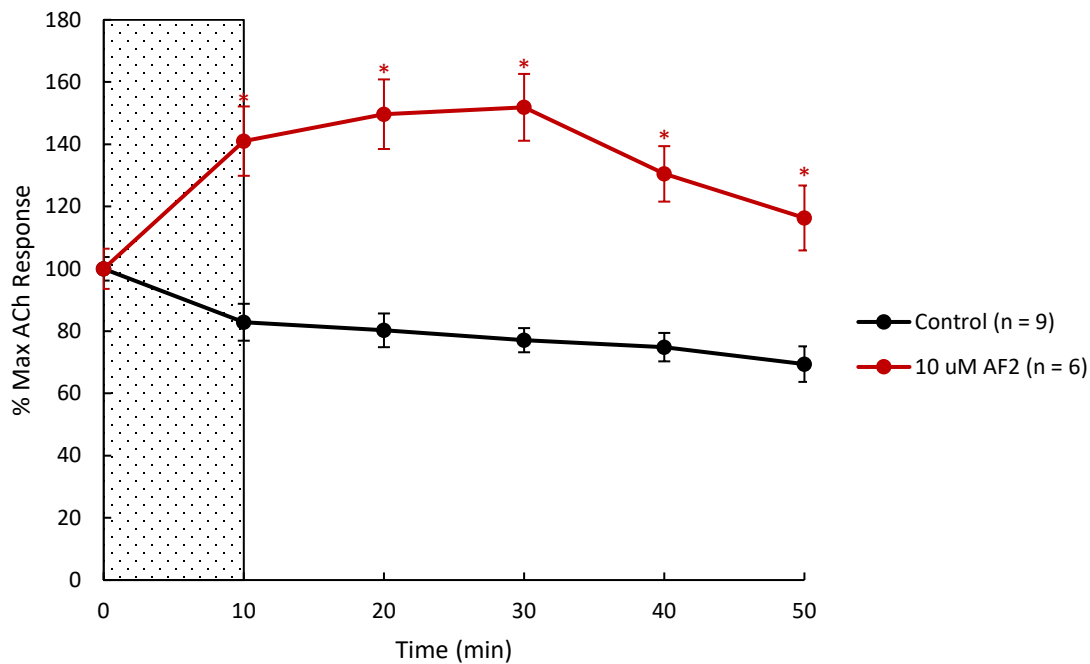
Supplementary table 1. Synthetic peptide purities as assessed by HPLC.

Transcript	Forward Primer	Reverse Primer
<i>afp-1</i>	SL1 (GGTTTAATTACCCAAGTTTGAG)	afp1 REV PROBE (5'-CGACCAAATCGGAGAACACCCGGCATA-3')
<i>afp-4</i>	afp4fwdprobe (5'-GTCATCATCGCTTGTGTCTCCGCTGATG-3')	afp4revprobe (5'-CAATGCTTTGCCCTTCTCTTCGGCGC-3')
<i>afp-14</i>	AF9FWDPROBE1 (5'-GCGTACCGTTTGTACGACCT-3')	AF9REVPROBE2 (5'-CATTAGCGAGAAGCCGAGAG-3')
<i>As-nlp-21</i>	As-nlp21-5'RACE (5'-CCCAAGCTGCTCAAACGAAACGATCGTCC-3')	As-nlp21-3'RACE GGACGATCGTTTCGTTTGAGCAGCTTGGG

Supplementary table 2. Primer sequences used to clone the *afp-1*, *afp-4*, *afp-14*, *As-nlp-21* transcripts.

APPENDIX 3

AF2 Effects on Muscle Contraction



Effect of AF2 on ACh-induced muscle contraction. Responses are reported as a percentage of the maximum contraction recorded before exposure to peptide. Error bars represent standard error of the mean. The shaded region indicates presence of 10 μM AF2 in the chamber. Exposure to AF2 also increased baseline tension and spontaneous contractile activity.

APPENDIX 4

List of Abbreviations and Nematode Species

Abbreviation	Full Name	Description
AC	Adenylyl cyclase	Enzyme that converts ATP into cAMP, major effector of G protein -s and -i pathways
ACh	Acetylcholine	Excitatory neurotransmitter used by DE1, DE2, DE3, VE1, and VE2 type motor neurons
AF	<i>Ascaris</i> FMRFamide-like peptide	Terminology for FMRFamide-like peptides in <i>Ascaris</i>
<i>afp</i>	Gene encoding AF peptide	Genes encoding <i>Ascaris</i> FMRFamide-like peptides are denoted by <i>afp-x</i>
DAG	Diacyl glycerol	Effector of G _q signaling, formed by hydrolysis of PIP ₂ by PLC, activates PKC
DC	Dorsal nerve cord	
DE1, DE2, DE3	Excitatory motor neurons (types 1-3)	Excitatory motor neurons with cell bodies in the ventral nerve cord, innervate dorsal muscle
DG	Dorsal ganglion	
DI	Dorsal inhibitory motor neuron	Inhibitory motor neuron with cell body in the ventral nerve cord, innervates dorsal muscle
EST	Expressed sequence tag	Sequence resulting from one-shot sequencing of a cloned cDNA
FLP	FMRFamide-like peptide	Class of neuropeptides sharing the C-terminal -Phe-Met-Arg-Phe-amide motif
<i>flp</i>	Gene encoding FLPs	Genes encoding FMRFamide-like peptides are denoted by <i>flp-x</i>
GABA	γ -aminobutyric acid	Inhibitory neurotransmitter used by DI and VI type motor neurons
GEF	Guanine nucleotide exchange factor	Protein or protein domain that activate GTPases to stimulate the release of GDP and binding of GTP
GPCR	G protein-coupled receptor	Receptor that converts extracellular ligand binding into an intracellular signal transduction pathway through associated G proteins. These receptors are comprised of 7 transmembrane-spanning domains, and the type of intracellular response is determined by the subtype of the associated G protein.
GRK	G protein-coupled receptor kinases	Enzymes that phosphorylate active GPCRs on Ser and Thr residues to block signaling through internalization or arrestin binding
GSS	Genomic survey sequence	Short, often single-pass sequencing of genomic DNA
HPLC	High pressure liquid chromatography	Technique used to separate sample components based on their size

Abbreviation	Full Name	Description
IN	Interneuron	
IP ₃	Inositol trisphosphate	Effector of G _q signaling formed by hydrolysis of PIP ₂ by PLC, diffuses through the cytosol to bind IP ₃ receptors and open Ca ²⁺ channels in the endoplasmic reticulum, resulting in increased Ca ²⁺ levels in the cytosol
ISH	<i>In situ</i> hybridization	Technique that uses labeled RNA probe to label cells that express the target mRNA
LLL, RLL	Left/right lateral line	
MALDI	Matrix-assisted laser desorption/ionization	Type of mass spectrometry that uses organic acid matrices to ionize sample
MN	Motorneuron	
MS	Mass spectrometry	Technique used to determine the peptide content of single-dissected neurons
MS/MS	Tandem mass spectrometry	Technique used to sequence peptide peaks identified by MS
NLP	Neuropeptide-like protein	Class of neuropeptides that do not share a common motif, but includes several sequence-related families
PIP ₂	Phosphatidylinositol 4,5-bisphosphate	Effector of G _q signaling pathway, cleaved by PLC to form DAG and IP ₃
PKA	Protein kinase A	Enzyme that phosphorylates numerous downstream targets upon activation by cAMP
PKC	Protein kinase C	Enzyme the phosphorylates numerous downstream targets upon activation by DAG and intracellular Ca ²⁺
PLC	Phospholipase C	Major effector of G _q signaling pathway, cleaves PIP ₂ into DAG and IP ₃
RACE	Rapid amplification of cDNA ends	Type of PCR that uses gene-specific and universal primers to amplify cDNA 5'- and 3'- to the target sequence
RGS	Regulator of G protein signaling	Structural domain of G α subunit that regulate intrinsic GTPase activity
RIA	Radioimmunoassay	Technique used to measure concentrations of antigens (in this case cAMP and cGMP) with antibodies
-RPamide	Family of NLPs	Family of neuropeptide-like proteins sharing the C-terminal -Arg-Pro-amide, includes peptides from the <i>nlp-2</i> , <i>nlp-22</i> , <i>nlp-23</i> , and <i>nlp-46</i> transcripts
RVG	Retrovesicular ganglion	
TM	Transmembrane domain	Alpha-helical domain that spans the membrane, named 1-7 for G protein-coupled receptors

Abbreviation	Full Name	Description
UTR	Untranslated region	Sequence adjacent to the open reading frame (either 3' or 5'), that is not translated into protein
VE1, VE2	Excitatory motor neurons (types 1-2)	Excitatory motor neurons with cell bodies in the ventral nerve cord, innervate ventral muscle
VG	Ventral ganglion	
VI	Ventral inhibitory motor neuron	Inhibitory motor neuron with cell body in the ventral nerve cord, innervates ventral muscle

Abbreviation	Nematode	Clade	Type
Aav	<i>Aphelenchus avenae</i>	IVb	fungi parasite
Aca	<i>Ancylostoma caninum</i>	V	dog parasite
Ace	<i>Ancylostoma ceylanicum</i>	V	human/hamster parasite
Adu	<i>Ancylostoma duodenale</i>	V	human parasite
Alu	<i>Ascaris lumbricoides</i>	III	human parasite
Asi	<i>Anisakis simplex</i>	III	human parasite
Asu	<i>Ascaris suum</i>	III	pig parasite
Bma	<i>Brugia malayi</i>	III	human parasite
Bmu	<i>Bursaphelenchus mucronatus</i>	IVb	plant parasite (pine wood)
Bxy	<i>Bursaphelenchus xylophilus</i>	IVb	plant parasite (pine wood)
Cbn	<i>Caenorhabditis brenneri</i>	V	free living
Cbr	<i>Caenorhabditis briggsae</i>	V	free living
Cel	<i>Caenorhabditis elegans</i>	V	free living
Cja	<i>Caenorhabditis japonica</i>	V	free living
Cre	<i>Caenorhabditis remanei</i>	V	free living
Dde	<i>Ditylenchus destructor</i>	IVb	plant parasite (potato rot)
Dsi	<i>Deladenus siricidicola</i>	IVb	wasp parasite
Dvi	<i>Dictyocaulus viviparus</i>	V	cow parasite
Gpa	<i>Globodera pallida</i>	IVb	plant parasite (potato cyst)
Gro	<i>Globodera rostochiensis</i>	IVb	plant parasite (potato cyst)
Hba	<i>Heterorhabditis bacteriophora</i>	V	insect parasite
Hco	<i>Haemonchus contortus</i>	V	sheep/goat parasite
Hgl	<i>Heterodera glycines</i>	IVb	plant parasite (soybean cyst)
Loa	<i>Loa loa</i>	III	human parasite
Mar	<i>Meloidogyne arenaria</i>	IVb	plant parasite (root knot)
Mch	<i>Meloidogyne chitwoodi</i>	IVb	plant parasite (root knot)
Min	<i>Meloidogyne incognita</i>	IVb	plant parasite (root knot)
Mja	<i>Meloidogyne javanica</i>	IVb	plant parasite (root knot)
Nam	<i>Necator americanus</i>	V	human parasite
Nbr	<i>Nippostrongylus brasiliensis</i>	V	rat parasite
Ode	<i>Oesophagostomum dentatum</i>	V	pig parasite
Pmu	<i>Plectus murrayi</i>	III	free living
Ppa	<i>Pristionchus pacificus</i>	V	free living
Ppe	<i>Pratylenchus penetrans</i>	IVb	plant parasite
Sra	<i>Strongyloides ratti</i>	IVa	rat parasite
Sst	<i>Strongyloides stercoralis</i>	IVa	human parasite
Tca	<i>Toxocara canis</i>	III	dog parasite
Tci	<i>Teladorsagia circumcincta</i>	V	sheep parasite
Wba	<i>Wuchereria bancrofti</i>	III	human parasite