Exploration of Novel Phenotyping Techniques and Identification of Quantitative Trait Loci for Chip Processing Potatoes

By

Curtis M. Frederick

A dissertation submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

(Plant Breeding & Plant Genetics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Date of final oral examination: 02/14/2017

The dissertation is approved by the following members of the Final Oral Committee:
Paul Bethke, Associate Professor, Horticulture
Shelley Jansky, Professor, Horticulture
Jeffrey Endelman, Assistant Professor, Horticulture
Shawn Kaeppler, Professor, Agronomy
Phil Townsend, Professor, Forestry and Wildlife Ecology

Dedication

A doctorate in science, is not a small feat,
This marks its completion, it feels bittersweet.
I'd like to say thank you, for the help that I've gotten,
With coursework and research, when tubers were rotten.

First and foremost, thanks to Rebecca my wife,
Her presence beside me, gave balance to life.
Thanks also to family, who planted the seed,
Of growing potatoes and the research it needs.

To all my lab members who helped me along,
And to all the grad students, who'd argue who's wrong,
Without all these lessons, who knows what I'd write,
I'd mess up my LOD scores, and surely not cite.

To all those who read this, please forgive the mistakes,
I did my best for the program, UW, and lakes.
And if there's a question on data or R scripts,
I'll always field questions on potato chips.

Abstract

Processing potatoes (*Solanum tuberosum*) are an important market segment for North American potato producers and much of this crop is kept in storages over winter. There has been demand for potato varieties that exhibit high dry matter content and the ability to fry out of cold storage conditions. Lenape is a notable variety release in the chip processing potato market segment and is well represented in the pedigree of many commercially grown cultivars. The precise genetic regions that confer improved attributes are unknown, so a Wauseon x Lenape (WxL) population consisting of 191 individuals was evaluated for traits related to processing quality under storage conditions and used to map QTL.

In Chapter 2, metabolite and vacuolar acid invertase (*Vinv*) relative expression was evaluated after two weeks in abrupt transition, 3^0C storage. There were 33 QTL detected in this study. The largest effect QTL (R^2=14.4) was for fructose and glucose on chromosome XII and mapped with 2.5 Mb of an invertase inhibitor gene.

In Chapter 3, chip color and stem end chipping defect (SECD) were evaluated at three processing intervals across multiple years. Sixty significant QTL were detected: 9 for SECD, 28 for chip color, 17 for chip color change in storage, 2 for maturity, and 4 for specific gravity. The largest effect QTL for SECD was on chromosome III and mapped to a position within 2 Mb of vacuolar acid invertase.

In Chapter 4, hyperspectral prediction models were developed for estimating dry matter, sucrose, glucose and fructose contents in raw potato tissue from chipping potatoes. We evaluated this technology on individuals from the WxL cross under different tuber dissection techniques, spectral subsets, and spectra transformations using a repeated, randomly sampled cross validation approach. The best prediction was for dry matter using 10 mm slab samples and 700-1450 nm or 700-1800 nm spectra with a mean R^2 = 0.757. External application of tissue and spectra parameters to a short wave infrared reflectance camera displayed highly accurate performance for dry matter prediction and has potential as a research device or quality assurance tool in industry.Table of Contents

List of Tables

List of Figures

Chapter 1: Literature Review

1.1 Storage of Processing Potatoes

Fall harvested potatoes account for approximately 90 percent of the planted potato acreage in the United States (Wells 2012). Across the 2012, 2013, and 2014 harvests, this amounted to an annual average of 20,321,000 metric tons (448,000,000 cwt) of potatoes produced nationwide and approximately 60 percent of this was still held by growers or processors in December (National Potato Council 2015)(National Potato Council 2013). Storing potatoes is primarily important for northern growing regions; 50 percent of Northern production was held into February of 2014 and 2015 (Mertz et al. 2016). Processing potatoes, which are generally used for French fries, potato chips, or other fried products, constitute the largest part of the stored potato segment. Approximately 40 percent of the total US crop is consistently delivered to processers between December and May (USDA-NASS 2016).

Wisconsin growers, in general, follow this regional trend in potato storage. In the state, growers carry approximately 65 percent of their harvested crop into the month of December ("North American Potato Market News December 21, 2016", 2016)(USDA-NASS 2016). Wisconsin grows approximately half of its potatoes for processing and the other half for fresh market, a market segmentation not common in other major production states (USDA-NASS 2016). Wisconsin's potato industry has a large economic effect with 270 million dollars in direct effect and 522 million dollars in total effect. For every one dollar generated by sales of the state's potatoes, another 93 cents of economic activity is created through purchases of goods and job creation (Kashian et al. 2014). The value of the crop is higher during the winter storage season both in the state and throughout the country (USDA-NASS 2016). There is strong economic incentive to increase the storage length and quality of potatoes both for Wisconsin growers and other stakeholders in the economy.

To realize the potential value of storing potatoes several months after harvest, producers of processing potatoes continuously monitor the quality of tubers and use the data acquired to market their crop (Bussan et al. 2009). One driver for vigilance over the stored crop is potential for loss caused by disease. The other driver is the Maillard reaction; a non-enzymatic process that causes the browning of cooked products. This reaction occurs during the frying process when the nucleophilic amino group of either an amino acid or protein interacts with the carbonyl group of a reducing sugar (Hodge 1953). Besides forming dark-colored pigments, these reactions create complex odors and flavors that give a distinct quality to a cooked food (Martins et al. 2000). Glucose and fructose are the reducing sugars that accumulate in stored potato tubers and vacuolar invertase activity plays an important role in this process. Invertase hydrolyzes one molecule of sucrose to produce one molecule each of glucose and fructose. When tuber reducing sugar concentrations are elevated, finished products become excessively dark and accumulate flavor compounds that are undesirable to consumers (Habib & Brown 1957)(Shallenberger et al. 1959). Many of the traits of interest for processing potatoes involve the biochemical pathways between stored starch, accumulation of sucrose, glucose and fructose, and respiration. Consequently, quantifying the levels of sugars in potato tubers has been an area of great interest to the chipping and fry potato industry (Sowokinos & Preston 1988). Failing to monitor sugar content in the context of both production and variety development can result in unacceptable defects for the consumer and financial losses due to quality issues for the producer (Sowokinos 2001a).

Acrylamide, a neurotoxin and suspected carcinogen (Costa et al. 1992)(LoPachin Jr. & Lehning 1994) is another product of the Maillard reaction (Rosén & Hellenäs 2002)(Tareke et al. 2002). Safety concerns about potato chips have been raised because they have a high level of acrylamide after cooking (Rosén & Hellenäs 2002)(Tareke et al. 2002). Acrylamide forms when the substrates of reducing sugars and asparagine are present during high temperature cooking (Gökmen & Palazoğlu 2008). There have been efforts to reduce asparagine and lower the levels of reducing sugars because both have a beneficial effect in lowering acrylamide production during the frying process (Bethke 2013)(FDA 2013). Newly

developed varieties have been identified that produce products with less acrylamide than standard varieties (Wang et al. 2016) and transgenic modifications have been used successfully to lower production of acrylamide in processed potato as well (Chawla et al. 2012)(Ye et al. 2010). It is also possible to lower acrylamide production by decreasing invertase gene expression, using conventional breeding, and suitable germplasm exists in wild species relatives of potato (Bethke 2013).

1.2 Traits of Interest in Chip Processing

Cold Induced Sweetening

A notable trait of interest for processing potatoes is cold-induced sweetening (CIS), in which a cold-mediated biological pathway accelerates the conversion of starch to sugars, generally at temperatures less than 10° C (Müller-Thurgau 1882)(Burton 1969)(Coffin et al. 1987). Cold storage is beneficial in that it allows for less respiration, less loss due to disease, and longer dormancy (Brown et al. 1990). CIS is a complex trait and many of the enzymes involved have been characterized at the molecular and biochemical levels (Kumar et al. 2004)(Sowokinos 2001a)(Chen et al. 2001). It is primarily comprised of two metabolic processes; the conversion of starch into sucrose and the hydrolysis of sucrose into the six carbon reducing sugars glucose and fructose (Sowokinos 2001a). During CIS, the synthesis of sucrose is accelerated (Isherwood 1973) and much of this sucrose is hydrolyzed in the vacuole to form glucose and fructose (Isla et al. 1992)(Blenkinsop et al. 2004). Most of the variation in fry color for potatoes out of cold storage is highly associated with the concentration of these reducing sugars that accumulate in the vacuole (Pritchard & Adam 1994).

Many enzymes influence the rate or extent of CIS. These include starch phosphorylase, UDP glucose pyrophosphorylase, sucrose phosphate synthase, β – amylase, acid invertase and others (Zrenner et al. 1993)(McKenzie et al. 2005)(Sowokinos 2001b)(Hill et al. 1996)(Nielsen et al. 1997)(Lorberth et al. 1998)(Richardson et al. 1990)(Zrenner et al. 1996). The production of reducing sugars has been strongly associated with the activity of vacuolar acid invertase (VInv) (Pressey 1969)(Richardson et al. 1990).

Transgenic plants expressing anti-sense or RNA interference (RNAi) constructs of *VInv* have provided evidence for the key role this enzyme in CIS. For example, Katahdin lines transformed with an RNAi silencing construct resulted in tubers that had minimal Maillard reaction browning and low reducing sugar (<0.3 mg g-1 FW) concentration after storage at 4 degrees C. Severity of CIS decreased as the level of *VInv* expression decreased (Bhaskar et al. 2010). Expression profiles show that *VInv* is upregulated in tuber tissue after exposure to cold temperatures (Matsuura-Endo et al. 2006)(Matsuura-Endo et al. 2006)(Liu et al. 2011)(Wiberley-Bradford et al. 2014).

Understanding the processes that regulate the abundance of *VInv* mRNA and activity could open new ways for incorporating additional levels of control over CIS. Temperatures that promote CIS also promote the up-regulation of *VInv* transcription (Zrenner et al. 1996)(Liu et al. 2011). This increase in transcript number has been shown to occur within 7 days when tuber tissue is suddenly exposed to storage temperatures of 4 degrees C (Zrenner et al. 1996). This large increase in transcript amount then gradually decreased over the storage period (Wiberley-Bradford et al. 2014)(Zrenner et al. 1996). Interestingly, differences in *VInv* gene expression in long term storage between wild type and transgenic RNAi lines are more pronounced during the beginning period of the storage period. This may be attributed to the persistence in VInv metabolic activity through the storage period regardless of transcript levels (Wiberley-Bradford et al. 2014). There is also evidence for native post-translational inhibitors of invertase activity (Pressey 1969). These small, 17 kDa proteins bind irreversibly to invertase in vivo and form an inactive complex (Rausch & Greiner 2004). Inhibitors of VInv from tobacco have been overexpressed in transgenic potato and resulted in short term CIS resistance (Greiner et al. 1999). It has been shown that cultivars that have a higher level of translation of native inhibitor gene *INH2* have both a lower level of reducing sugar accumulation and a suppression of *VInv* expression (Liu et al. 2010)(Brummell et al. 2011)(McKenzie et al. 2013). It is not known how CIS-resistant, North American chipping germplasm regulates the expression of *VInv*. It would be valuable to understand more about the regulation of this gene under cold storage stress.

Stem-End Chip Defect

A common tuber quality defect in chipping potatoes is stem-end chip defect (SECD), a localized, post-fry accumulation of dark color near the vasculature at the basal (stem end) of the tuber (Figure 1.1 from Wang, 2012). SECD can be caused by either early to mid-season heat stress and is associated with infection by *Verticillium* species. Increased severity of SECD is accompanied by increased amounts of VInv activity (Wang et al. 2012)(Wang & Bethke 2013). This defect has not had its economic value quantified but has been reported to affect between 5 to 25% of the chip processing crop annually (Dickman 2016). Occurrence of this defect is sporadic and afflicts growing regions to differing extents depending on the year. The lack of a reliable environmental stress makes studying the mechanisms of SECD and selecting against it difficult (Dickman 2016)(Wang et al. 2015).

SECD is comparable to sugar end defect in fry processing potato varieties in that a localized dark pattern occurs at the stem end of the fried tuber (Kleinkopf 1976)(Iritani & Weller 1973). Sugar end defects are also referred to as jelly ends, dark ends, translucent ends, and glassy ends (Thompson et al. 2008). Sugar end defect develops in response to moderate heat and drought stress during early and middle bulking stages (Eldredge et al. 1996)(Bethke et al. 2009). There are instances where the incidence of sugar end in russet class potatoes is expressed differentially across abiotic stress treatments and growing environments (Thompson et al. 2008). Enzymatic changes are associated with the development of sugar end defects, with vacuolar acid invertase having a sizable increase in activity in sugar end defect tissue (Sowokinos et al. 2000)(Zhu et al. 2014). It has been postulated that an environmental stress event results in one or a combination of the following: a disruption of carbohydrate production from the vines (Iritani & Weller 1973), translocation of carbohydrates from the basal end to the apical end of the tuber (Shekhar & Iritani 1979), a shift in sink strength that favors the vines (Iritani & Weller 1973)(Lugt 1960), or starch degradation in the basal end as a result of the disruption of carbohydrate flow (Krauss & Marschner 1984)(Hiller et al. 1985). The starch degradation due to carbohydrate flow interruption theory is further

supported by a shift of enzymes in the basal tissue away from starch accumulation toward glycolysis and starch mobilization (Sowokinos et al. 2000).

There is a genetic component to sugar end defect and SECD expression. Large differences have been observed among russet cultivars in their susceptibility to sugar end defects (Shock et al. 1993)(Wharton et al. 2015) and among chipping cultivars in severity of SECD (Wang et al. 2015)(Dickman 2016). Additionally, a half-diallel experiment conducted between 5 parents displayed a mean broad sense heritability of 0.76 for the sugar end defect trait (Thompson-Johns 1998). Genetic components to SECD rating at harvest and the level of reconditioning that occurs in storage have also been observed (Wang et al. 2015).

Although similar in many ways, there are differences between SECD and sugar end defect. Sugar ends, when present, have a reduction in specific gravity at the tuber stem end and the location of the defect is not closely associated with the vasculature (Kleinkopf 1976)(Sowokinos et al. 2000). Comparable reductions in dry matter content have not been observed in SECD tissue and the defect is found in close proximity to the vasculature (Wang et al. 2015). SECD has been observed to recondition in storage, a process whereby reducing sugar content decreases as a result of respiration or starch synthesis, and the defect becomes less severe with time in the cultivar Snowden (Wang et al. 2015). However, there is evidence that this does not occur in all years or locations (Dickman 2016). Sugar end defects, for the most part, do not improve with preconditioning or longer storage (Thompson et al. 2008). Another difference is that SECD is not always present at harvest and glucose, fructose, and sucrose assays taken at that time do not reliably predict the severity of the trait (Wang et al. 2015)(Dickman 2016). Sugar ends may be apparent at harvest and sugar sampling at harvest readily predicts the severity of the trait (Sowokinos et al. 2000). The exact relationship between these two traits in these two market categories has not been thoroughly explained. Understanding more about the genetics of SECD or sugar end defect could help breeders and researchers develop improved varieties more quickly.

Tuber Specific Gravity

Another major tuber quality trait that impacts the processing potato industry is tuber specific gravity. This measurement, used commonly in North America, is a simple estimate for dry matter content (Schippers 1976)(Simmonds 1977). Dry matter content is strongly related to tuber performance during the cooking process; higher dry matter results in more cooked product and less fry oil absorbed per unit of raw product (Lulai & Orr 1979). Commonly, potato dry matter content is quantified by measuring the weight of potatoes in air and when submerged in water (Nissen 1955). The buoyancy of the tubers is dependent on density, and density is linearly related to dry matter content (Schippers 1976; Simmonds 1977). The weight in air, weight in water method has limitations in that data collection on individual tubers is rarely performed due to time constraints. Thus, variation of specific gravity among tubers within a treatment is not measured routinely. Errors can be introduced when potatoes contain a hollow center or other defects that impact buoyancy. This technique also does not permit convenient measurement of within tuber variation of dry matter (Reeve et al. 1969)(Peiris et al. 1999).

Tuber Maturity at Harvest

Tuber maturity at harvest is an important trait because it determines a cultivar's suitability for a growing environment. Tuber maturity at harvest impacts the aforementioned traits of sugar accumulation and specific gravity, as well as processing quality out of storage (Heltoft et al. 2016). The term maturity is not described by a single parameter, however, but rather by multiple terms (Bussan et al. 2009). It is measured in four ways: chemically, physiologically, physically, and in terms of haulm maturity. Chemical maturity is a term that the potato industry and researchers use to describe sugar concentrations in the potato (Driskill et al. 2007)(Pritchard & Adam 1992) and harvest is best timed when sucrose levels reach a minimum (Kolbe & Stephan-beckmann 1997)(Sowokinos 1978). Physiological maturity is measured by the concentration of dry matter in the tuber; ideally with the highest levels at the time of harvest (Sabba et al. 2007). Physical maturity involves the development of a mature periderm which is important for skin-set (Lulai & Orr 1993). Haulm maturity, or vine maturity, is commonly measured in research programs as

a visual score of vine vigor, or greenness approximately 14 days before harvest (Bussan et al. 2009). Each of these maturity measurements can be affected by management decisions, environmental factors, and genotype.

A single, subjective haulm maturity ratings does not necessarily correspond with the other tuber maturity measurements, but it has advantages over numerous tuber samplings (Sabba et al. 2007)(Heltoft et al. 2016)(Kumar et al. 2004)(Bussan et al. 2009). There has been a trend in selecting varieties in North America for shorter season haulm maturity indicating a preference for varieties that mature earlier (Douches et al. 1996).

1.3 Phenotyping Methods for Processing Potatoes

Chip processing potatoes have many traits that are economically important, challenging for breeders to improve upon, and difficult to measure. The latter two attributes are interconnected. In order to improve breeding progress for these traits, one can either improve plant measurement speed and accuracy or invest energy into developing a reliable molecular marker that is linked to a particular trait (Bernardo 2010).

Tuber biochemical trait phenotyping

The primary biochemical traits measured in processing potatoes are sucrose, glucose, fructose, and dry matter. The techniques used to measure these traits have limitations. Specific gravity is used as a proxy for dry matter and limitations relating to measurements within tubers, on individual tubers in a plot, and the effect that defects have on a tuber's buoyancy were described above (Nissen 1955)(Schippers 1976). The quantification of sucrose, glucose and fructose is important in processing potatoes and there are several methods in use. The most common are high performance liquid chromatography (Nikolov et al. 1984)(Wilson et al. 1981)(Bethke & Busse 2008) and an enzyme-based assay commonly implemented on instruments from YSI (Yellow Springs Instrument)(Mason 1983)(Sowokinos et al. 1985). These methods require expensive equipment, a sizable amount of time to prepare samples, and results may not

be available for several hours or days. Other methods include portable blood glucose monitors that have been adapted for horticultural use. These have the disadvantages of not being very accurate, and they can only measure glucose directly (Helgerud et al. 2016). None of these methods can easily assess within tuber heterogeneity.

VIS-NIR-SWIR reflectance spectroscopy

VIS-NIR-SWIR reflectance spectroscopy is a technology that has the potential to improve phenotyping of processing potatoes for dry matter and sugar concentration measurements. This technology is employed as an alternative to labor intensive or slow wet-chemistry analysis and has been adopted for use in phenotyping other crops (Nicolaï et al. 2007). Full-range hyperspectral reflectance spectroscopy captures the reflectance of light from a specimen at 400-2500 nm at narrow intervals (1-10 nm). The spectral signature of the reflectance is dependent on the chemical composition of the sample. This is attributed to the fact that molecular bonds within the compounds that make up a biological sample absorb energies at discreet wavelengths, and the proportion of these bonds depends on the composition of the sample. A predictive model in the form of an algebraic equation can be developed by relating spectral signatures to measured physical properties. Such equations have a measureable degree of accuracy, and only when the accuracy of the estimate is acceptable should the method be used to replace laboratory techniques (Osborne 2000).

Partial Least Squares Regression (PLSR) is the method most commonly used to construct a mathematical relationship between the spectral reflectance values and chemical trait values (Wold et al. 1984). PLSR is a generalization of multiple linear regression that creates a relationship between highly collinear predictor variables (spectral reflectance) and one or more response variables (traits of interest). PLSR has advantages over multiple linear regression in that PLSR avoids over fitting a model by reducing the number of predictor variables to a small number of orthogonal factors known as latent variables (Wold et al. 2001). Additionally, PLSR is advantageous over principal component analysis (PCA) in the context of dimensional reduction for use in prediction modeling because it maximizes the

covariance between predictor and response variables (Maitra & Yan 2008)(Wold et al. 2001)(Tobias 1995).

In the absence of an external dataset, a PLSR model is evaluated for model accuracy by using cross validation methods. During a cross validation procedure a calibration model is constructed using a subset of the full data set. This model is used to predict values for the hold out data set (which can range in size from one sample to a large proportion of the complete dataset) using only spectra reflectance data. This cross validation method is commonly repeated multiple times on a random split of the dataset (Osten 1988). Model accuracy is measured by comparing predicted trait values to observed trait values, and it is desirable to minimize predicted residual sum of squares (PRESS) when developing calibration models (Wold et al. 2001)(Wakeling & Morris 1993)(Chen et al. 2004). Calibration equations built using PLSR can be optimized by adjusting variables important to creating the prediction equation. These include the number of selected latent variables, spectral value transformation, removal of irrelevant spectral bands, data smoothing algorithms, and weighting of factor loadings (Wold et al. 2001)(Chen et al. 2004)(Chunhua et al. 2013). Additionally, robust PLSR calibration equations must consider the distribution of the laboratory trait values as there have been situations when transformation of the data improves model prediction accuracy (Osborne 2000)(Mehmood et al. 2012).

Hyperspectral reflectance-based estimation of dry mater and sugars in potato has been evaluated previously for a small number of varieties. Dry matter has been estimated in several studies with correlation coefficients for model validation ranging from 0.90 - 0.98 and a root mean squared error of prediction (RMSE) between 0.0130 – 0.0169 g dry matter per g fresh weight (w/w) (Haase 2004)(Dull et al. 1989)(Subedi & Walsh 2009)(Singh et al. 2004). Reflectance spectroscopy has been less accurate for the prediction of glucose, fructose, and sucrose. Calibration correlation coefficients have ranged from 0.38 – 0.95 and RMSEs were between 0.280 – 3.259 mg g-1 fresh weight (Rady et al. 2014)(Rady & Guyer 2015a)(Hartmann & Büning-Pfaue 1998). Researchers developing methods to estimate potato processing traits using reflectance spectroscopy have evaluated a range of tissue sample preparation

techniques including raw whole tubers, peeled whole tubers, raw tissue dissection, homogenized tissue, and freeze dried tissue (Dull et al. 1989)(Haase 2011)(Scanlon et al. 1999)(Hartmann & Büning-Pfaue 1998)(Haase 2004)(Singh et al. 2004). These studies and other hyperspectral estimation studies suggest that prediction accuracy varies with sample preparation method (López et al. 2013)(Rady & Guyer 2015b).

In some crops, PLSR equations have been developed based on a large number of varieties grown in multiple environments and years. These global equations have permitted the use of hyperspectral reflectance in breeding (De Leon 2006)(Osborne 2000)(Nilsson 2011). One study in potato evaluated hyperspectral reflectance estimation for sugars on a broad range of genotypes using freeze dried samples. Sucrose, glucose, and fructose had prediction correlation coefficients between 0.95 – 0.97 and root mean squared error of prediction (RMSE) of 0.32 -1.15 mg g-1 fresh weight in a single hold out, external validation method (Ayvaz et al. 2015). To our knowledge, there are no previous studies that have investigated the accuracy of a population-based model on sugar or dry matter estimation where spectra were collected from raw potato tissue.

1.4 Potato Genetics

Cultivated potato is a clonally propagated autotetraploid that generally outcrosses with limitations on interspecies combination and occasional fertility issues (Spooner et al. 2014)(Jansky 2006). Cultivated potato has a basic chromosome number of 48 (2n = 4x = 48). Autotetraploid species are a class of crops also occupied by alfalfa, sour cherry, blueberry, leek, and some ornamental roses (Quiros 1982)(Tavaud et al. 2004)(Gar et al. 2011)(Jones & Khazanehdari 1996)(Boches et al. 2006). Autotetraploids are defined by the presence of four homologous or non-preferentially pairing chromosomes. Two chromosome copies come from each parent and at meiosis, any two of the four copies can pair. This results in there being up to 4 alleles at any locus and segregation patterns that depend on the mode of inheritance and allelic dosage present in the parents at a given locus. In order to determine the inheritance of an allele *in situ*, more progeny must be evaluated in autotetraploids compared to diploids because there

are a greater number of allelic combinations produced from heterozygous parents (Table 1.1)(J. E. Bradshaw 2007).

Observed segregation ratios in autopolyploid progeny are influenced by a process called double reduction. Double reduction occurs when quadrivalents form rather than non-preferential bivalents (J. E. Bradshaw 2007). A single cross over occurs between homologous chromosomes and then through co-segregation in anaphase I and II of meiosis, a gamete receives alleles that were originally from sister chromatids (Haynes & Douches 1993)(Comai 2005). Double reduction products in potato occur more frequently as distance from gene of interest to the centromere increases (Bourke et al. 2015). Double reduction introduces complications when using marker data for linkage mapping, but can also be beneficial for crop improvement in that it can enrich progeny in useful, rare alleles.

The potato genome is estimated to consist of 40,000 protein encoding genes and has a size of 844 megabases (Xu et al. 2011). Genetic analysis of the potato genome has been facilitated by the development of the SolCAP SNP array, which contained 8303 SNP markers from expressed genes (Hamilton et al. 2011). This array has been used in QTL mapping studies and in a study examining the diversity of modern potato germplasm (Hirsch et al. 2013). More recently, a 20,000 SNP array was developed in Wageningen and a 12,000 SNP array is being developed by Douches (Vos et al. 2015)(NCCC215: Potato Breeding and Genetics Technical Committee, 2015). These genetic resources have permitted genetic populations to be genotyped with high marker density coverage and screened for quantitative trait loci (QTL) that are associated with important traits in potato.

1.5 Breeding for Potato Market classes

Breeding progress in potato is not a clearly illustrated success story. Critics will highlight the dominance of Russet Burbank, a sport of a variety introduced in 1876 that still accounts for approximately 40% of the potato acreage planted in the U. S. (Ortiz, 2001)(Van Berloo et al. 2007)(Bethke et al. 2014). Its persistence highlights the difficulty in replacing varieties for the frozen process industry. The low rate of new variety approval is tied to difficulties in achieving culinary traits

such as flavor and texture that processors and customers demand (Gould & Plimpton 1985). Although Russet Burbank represents an extreme example, potato across all market segments is a difficult crop to breed because of the number of quality traits a new variety must meet. Fresh market potato varieties must focus on consumer acceptability attributes (McGregor 2007). This includes flavor compounds, texture, size, and pigmentation (Stushnoff et al. 2008)(Jansky 2008)(Jansky 2010). The fresh potato market class segment has shown evidence of selection among skin and tuber flesh pigmentation pathways (Hirsch et al. 2013). In all potato market segments, the combination of desired customer quality traits with the requirements for agronomic traits and disease resistance along with the low proportion of desirable gametophytes produced by autotetraploids makes the likelihood of a releasing a new variety low (John E Bradshaw 2007). Given these constraints, comparing this breeding segment to grain crops is not appropriate but improvements in efficiency would be welcomed.

The chip processing segment has been more successful than the frozen process segment in replacing commonly grown varieties. Storage traits for chipping potatoes have been purposefully targeted by breeders for the past 75 years. Primarily, these traits include maturity, dry matter, and fry color out of cold storage (Stevenson 1957). Progress was aided by an understanding of how sugars, amino acids and proteins impacted fry color for chips and other fried products (Habib & Brown 1957). University and private breeding programs began breeding for lighter fry color in the 1950's for an emerging potato chip processing industry. Red Dot foods, a prominent potato chip manufacturer and grower, developed its own breeding program in 1947 in Rhinelander, Wisconsin (Stevenson 1957). This breeding program established the quality standards and germplasm pool that was eventually acquired by Lay's potato chips in 1961. While the factories and distribution network were sold off, Frito-Lay kept the breeding program and is still developing proprietary varieties in Rhinelander (Wisconsin Historical Society 2003). Part of the breeding program's value to Frito-Lay was based on the corporate strategy of focusing on quality; minimizing variation in its product both within a bag and through the calendar year (Schreiber 1997). In both public and private breeding releases, there has been improvement for processing quality out of long

term storage. However, these improvements do not exhibit smooth, deliberate gains for chip color over time. Lenape displays notable improvement for processing traits and variation in performance for cultivars released thereafter, and in many cases these appear to be associated with the presence of Lenape in the pedigree (Figures 1.2 and 1.3) (Love et al. 1998).

Potato varieties in cultivation have clearly separated into population structure groups that correspond to their market class categories (Hirsch et al. 2013).While there are many areas of the genome that do not have selection signatures, the pathways involving pigmentation and carbohydrate metabolism have evidence of selection (Hirsch et al. 2013). Two era studies comparing major varieties released from the 1870s through the 1990s suggested that the main traits of improvement have been tuber dry matter and chip quality out of cold storage (Love et al. 1998)(Douches et al. 1996). One contributor to the improvement of the North American germplasm is thought to be the cultivar Lenape, which showed notable resistance to cold-induced sweetening and a large increase of dry matter in comparison to earlier cultivars (Love et al. 1998). Lenape was released in 1968 and was derived from a cross between USDA Seedlings B3672-3 and 47156. B3672-3 contained a *S. chacoense* accession in its background that is thought to have conferred cold-induced sweetening resistance (Akeley et al. 1968). While not widely adopted itself due to concerns over moderately high tuber glykoalkaloid content, Lenape is prominent in the pedigree of numerous successful chipping and fry processing cultivars including Snowden, Atlantic, Belchip, Gemchip, Pike, NorValley, Tundra, Accumulator, Lamoka, Highland Russet, Blazer Russet, and Alturas, amongst others (Van Berloo et al. 2007). For future breeding efforts, it would be useful to understand the mode of inheritance for the beneficial traits provided by Lenape so that selections might be made more efficiently using marker technology.

Breeding schemes in potato typically rely on recurrent phenotypic selection of progeny resulting from crosses between two elite parents. Clonal propagation is used in typical breeding programs because parents are heterozygous and it is important to preserve the genetic architecture of favorable offspring. Wide-area testing is delayed because clonal propagation limits the rate of increase of selected material;

one tuber planted produces 8-12 tubers (John E Bradshaw 2007). This is in contrast to other widely grown crops in which hybrid development, family development, and inbreeding has been possible and much higher seed multiplication ratios permit rapid testing in more locations (Fehr 1991). These limitations shape the methodology and pace of evaluating and developing material in a breeding program.

Variety development by a typical North American breeding program begins with families produced from crosses between pairs of elite parents. In year one, 40,000 – 100,000 single hills may be evaluated in the field. Approximately 2,000 offspring are planted the next year in larger plots, 800 are selected for the third year, and these numbers are reduced gradually until 5-20 clones remain in year 5 (John E Bradshaw 2007)("NCCC215: Potato Breeding and Genetics Technical Committee", 2015).These lines are then tested in regional trials for years 6-10. In each generation, some tubers become infected with virus and may be excluded from planting the following year. In many, but not all cases, favorable lines are entered into a tissue culture disease clean-up program to remove accumulated virus before they are entered into national agronomic and on-farm field trials. Overall, it often takes 10-14 years from the time a cross is made until a variety is released. Parent selections occur no earlier than year 6 (NCCC215: Potato Breeding and Genetics Technical Committee, 2015)(Bradshaw & Mackay 1994).

Multiple strategies that might improve the efficiency of tetraploid potato breeding have been suggested. Several studies suggest it is more effective to reduce clonal selection in the first generation and focus more on family selection, and that identification of favorable parents can result in a higher success rate in progeny (Tai & Young 1984)(Brown & Caligari 1989)(Bradshaw & Mackay 1994). The Scottish Crop Research Institute uses this strategy for variety development (John E Bradshaw 2007).

Regardless of breeding strategy used, aids in parental selection and efficiency improvements for selection years beyond field year 3 are desired. Research has been conducted on marker-based tools to achieve both of these gains (Gebhardt et al. 2005)(Hamilton et al. 2011)(Li et al. 2013)(Ramakrishnan et al. 2015). Genomic selection is the development of a prediction model that uses the marker and phenotypic data for a particular population of related individuals, known as a training population. Trait

performance on a separate individual is predicted by using its genetic marker information and coefficients supplied from the prediction model. The coefficients are used to weight each marker along the genome and are based on that particular marker's contribution to the trait. Genomic selection has been successful in other crop and animal species but for use in potato requires validation, a large robust training population dataset, and a lower cost of genotyping before it can be used in a breeding pipeline (Bernardo 2010)(Lorenz et al. 2011). Marker assisted selection, which uses markers linked to large effect alleles, has been used for selection in traits that are difficult to screen (Ribaut & Hoisington 1998)(Bernardo 2010). Markers currently being used in potato breeding programs include markers for resistance to Potato Virus Y, Potato Virus X, Potato Leaf Roll Virus, Golden Nematodes and Late Blight with differing use depending on region (personal communication with Walter De Jong, Cornell; Dave Douches, Michigan State University; Craig Yencho, NC State University; Jeff Endelman, University of Wisconsin-Madison)(Dale et al. 2016). A marker for UGPase isozymes has been used to predict cold-induced sweetening resistance (Sowokinos 2001b)(McKenzie et al. 2005). Breeders are interested in additional markers that can assist them in selecting for hard to measure traits.

1.6 QTL Mapping and Genetics in Autotetraploids

Most traits are inherited and controlled by multiple loci in a quantitative manner. Known as quantitative trait loci (QTL), these genomic regions are often gateways to understanding the inheritance of important traits (Bernardo 2010)(Falconer & Mackay 1996). Researchers often want to know the location of a QTL within the genome, and they find that location by using genetic marker maps. First reported by Sax in 1923 in his pioneering work on the use of seed pigmentation to identify genetic contributors to seed weight (Sax 1923), QTL mapping is performed by finding a relationship between trait means and genotype at genetic marker loci. Genetic marker technology and statistical analysis have become more advanced since Sax did his research, however, QTL mapping still relies on the development of a population consisting of related individuals and statistical techniques to find the relationships between phenotypes and the genetic markers along the genome.

Developing populations

Methods used to develop a linkage mapping population depend on the biology of the organism being studied. The strategy is often dictated by whether the organism permits inbreeding and the nature of the trait being studied (Semagn et al. 2010). For inbreeding organisms, the most commonly developed populations are F_2, recombinant inbred lines (RIL), backcross (BC), and near-isogenic lines (NIL). To create a F_2 population, two inbred parents are crossed to create an F_1 that is then selfed to create the F_2 generation. This population has the ability to estimate dominance effects of a locus (Broman & Sen 2009). The RIL population is developed by either selfing the F_2 for over 6 generations or by producing doubled haploids. Both methods result in an inbred population that permits seed propagation and wide area testing but does not produce heterozygotes for the estimation of dominance effects (Paterson & Wing 1993). Other strategies for RIL development include intermating between each generation, which has the advantage of breaking up linkage blocks. A BC population is developed by crossing the F_1 with one of the two parents (Bernardo 2010). This can be continued to a NIL population by using markers to create an inbred genome amongst the progeny with the exception of a particular genetic region of interest, where there are different linkage blocks inherited from the original two parents (Tanksley et al. 1996). For outcrossing species, the populations developed are generally half or full sib populations. Genetic analysis for populations developed in an outcrossing species face more challenges because the parents are typically heterozygous. For example, segregation patterns and number of marker alleles vary from locus to locus. There may be complications when parents have markers in common at QTL, or if they share QTL alleles but in different linkage phases (Bernardo 2010)(Lynch & Walsh 1998). Additionally, linkage phases among markers are usually not known for each of the parents in advance and must be determined prior to linkage analysis (Lu et al. 2004).

QTL Mapping techniques

After the population is developed, phenotyped, genotyped, and had its markers grouped and ordered in a linkage map, QTL are mapped for the traits evaluated. Many statistical approaches have been used for QTL mapping in the past several decades. Here we review some of the earlier approaches and conclude with the more recent and computationally intensive methods. Single factor analysis independently tests each marker for significant associations between marker genotype and the trait in an analysis of variance (ANOVA) approach (Lynch & Walsh 1998). This has limitations in that the location of the QTL relative to the marker is not determined and two or more markers can detect the same QTL (Doerge 2002). Interval mapping is used to overcome the position issue by using flanking markers; information provided by the linkage map (Soller et al. 1976). Interval mapping uses a maximum likelihood approach to estimate the trait locus. This involves the calculation of the logarithm of odds (LOD) score, and the marker interval with the highest LOD score is assigned as the most likely position of the QTL. A limitation of interval mapping is that it is unable to accurately estimate the effect size or the interactions between multiple QTL (Broman & Sen 2009)(Bernardo 2010). Multiple regression is used to find which marker intervals contain QTL, and does so by using the trait value as the y variable and the number of alleles from a given parent at each marker as the X variables. The regression coefficient is significant for a marker with QTL in one of its intervals (Wright & Mowers 1994). Composite interval mapping uses both the multiple regression and interval mapping methodologies. Interval mapping finds the location of the QTL while multiple regression accounts for effects caused by QTL located in other areas of the genome (Zeng 1993). There is also Multiple Interval Mapping which considers multiple marker intervals at a time and models multiple QTL, their effect on the trait, and any interactions they might have (Kao et al. 1999).

Autotetraploid QTL Mapping

The methods described above were developed for diploid organisms and, unfortunately, are not readily transferrable to an outcrossing, heterozygous autopolyploid. For this reason, many of the early potato

genetic populations were developed from diploid material (Gebhardt et al. 1989)(Tanksley et al. 1992)(Hämäläinen et al. 1997)(Schäfer-Pregl et al. 1998). However, this approach has disadvantages in that results are not immediately transferable to breeding programs focused on the cultivated tetraploid crop (Milbourne et al. 2009). For potato, initial genetic maps were constructed using only single dose (simplex) and nulliplex markers that segregated in a 1:1 ratio (Bonierbale et al. 1988)(Gebhardt et al. 1989)(Meyer et al. 1998). This mapping approach had limitations in that the number of markers was reduced and it ignored co-dominant markers as well as additive allelic dosage effects. Advances that permitted the use of codominant markers included the development of methods to detect double reduction rates, predict segregation ratios, predict parental genotypes, and identify crossing over events (Wu et al. 2001)(Luo et al. 2000). Linkage mapping using duplex codominant marker classes became possible through the use of an expectation-maximization (EM) algorithm that allowed recombination frequencies between all marker classes to be calculated (Wu et al. 2001) . Multiallelic markers also permitted a maximum likelihood approach towards parental phase estimation between pairs of markers which could then be extended to all markers in a linkage group (Luo et al. 2001). An interval mapping approach was developed for a full-sib population using the linkage map construction method, and a branch and bound method for configuring chromosome configurations for each offspring (Hackett et al. 2001). However, without enough multiallelic markers, separate maps had to be constructed for each parent (Hackett et al. 2001)(Hackett & Luo 2003).

Once SNP markers became a viable option for genotyping in potato, the challenge was calling dosage classes in tetraploids based on theta values from the SNP array. This has been addressed through a highly conservative manual approach, a mixture model approach, a maximum likelihood approach based on parental and progeny marker segregation, and a hierarchical clustering method (Douches et al. 2014)(Voorrips et al. 2011)(Hackett et al. 2013)(Schmitz-Carley et al. 2017). Dosage call information for each marker is then used to model the additive effect of each of the 8 inherited parental alleles using the genotypic probabilities (GP) model. This method requires that identity by descent be known for each

allele across the genotyped set of loci within an individual. To do this, the linkage phase of each marker in the parents must be determined and the approach described by Luo (2001) applied this to SNP markers. Hackett et al. (2013) provides a hidden Markov model routine that interpolates the probabilities that homologous chromosome segments were inherited from each parent for each individual in the population (Hackett et al. 2013). This information is then used in an interval mapping approach for QTL discovery (Hackett et al. 2014).The mapping platform that brought many of these statistical tools together was TetraploidMap 2.0 and besides being used in this project, this software has been used to map QTL in recent tetraploid population studies (Rak 2015)(Massa et al. 2015)(Hackett et al. 2014).

There has been success in finding QTL in autotetraploids for a wide range of traits including resistance to diseases, nematodes, yield, nutrient content, and processing quality (Bradshaw et al. 2004)(Meyer et al. 1998)(Zorrilla et al. 2014)(Massa et al. 2015)(Khu et al. 2008)(Groth et al. 2013). There have been QTL studies in processing potatoes for chipping color (Bradshaw et al. 2008)(Rak 2015), cold chipping tolerance (Bradshaw et al. 2008)(Rak 2015), and dry matter content (Bradshaw et al. 2008)(Per H. McCord et al. 2011). Yield traits have been mapped, including those for tuber size (Bradshaw et al. 2008)(Rak 2015) and tuber weight (Rak 2015). There have been studies on physiological traits including maturity (Hackett et al. 2014)(Bradshaw et al. 2008)(P. H. McCord et al. 2011)(Massa et al. 2015)(Per H. McCord et al. 2011), area under a senescence curve (Per H. McCord et al. 2011), crop emergence (Bradshaw et al. 2008), height (Bradshaw et al. 2004), internal calcium concentration (Zorrilla et al. 2014), and internal heat necrosis (P. H. McCord et al. 2011). Many of the yield and late blight resistance traits map to the same arm as maturity on chromosome V. This is the same location of both canopy development traits (Khan 2012) and a transcription factor that regulates tuberization in cultivated potato (Kloosterman et al. 2013), indicating possible pleiotropic control.

Limitations of Autotetraploid QTL Mapping

While QTL mapping is a powerful tool to understand genetics in autotetraploid potato, there are limitations to this approach. To detect small effect QTL, population sizes must be larger than those used

when working with diploid organisms in order to achieve the same power (Hackett et al. 2014)(Bourke 2014). This has a negative consequence in that it raises phenotyping costs. This method also has limitations regarding the analysis of the genotype data. TetraploidMap 2.0 removes double reduction products because they violate the segregation patterns used in its algorithms. Doing this simplifies haplotyping, but it falsely suggests that there are systematic errors in the markers that may reduce the accuracy of the linkage map (Li et al. 2010). Actual frequencies of double reduction are low in potato for the simplex by nulliplex markers, but more understanding of this effect on other markers would be useful (Bourke et al. 2015). Copy number variation (CNV) is highly abundant in the potato genome and has an impact on transcriptome levels and plant vigor (Iovene et al. 2013). This has been shown to affect approximately 30% of the genes in a diploid set of genotypes and there are instances of both duplication and deletion events (Hardigan et al. 2016). CNV can have effects on mapping tools if the SNP itself was duplicated. This may result in segregation patterns of SNP markers that either result in mapping error or the exclusion of potentially informative SNPs. If the marker is adjacent to a functional gene that has CNV, the allelic interaction effect could be misconstrued. Additionally, the current linkage phase calling routine for the each of the 12 linkage groups is mostly a manual process. While there is a simplex structure base used to guide this phasing, it remains a laborious task. This is likely to improve in the near future with the development of an accurate and automated method for haplotyping (Zheng et al. 2016). Finally, TetraploidMap uses an interval mapping approach that does not account for multiple QTL and their interactions within the chromosome and between linkage groups. Additionally, the size of linkage groups and marker coverage impact the resolution of QTL position estimation (Doerge 2002). It is important to understand these limitations when interpreting results from TetraploidMap.

1.7 Thesis Objectives

This research has the intended purpose of understanding the genetic causes behind key quality traits in processing potatoes. In addition, it aims to provide insight into novel phenotyping methodologies and

provide information to guide their implementation for use with a diverse set of potato genotypes. The approaches used to address these objectives are as follows:

- A bi-parental population was developed using a Wauseon x Lenape cross, prominent in current chip processing germplasm. A linkage map was developed using this SNP genotyped population and that map was used for QTL analysis.
- The mapping population was evaluated for chip processing traits throughout the storage season. Replicate tuber samples were measured for chip color and ranked for stem-end chip defect (SECD) across years and data were used to identify QTL that are related to these traits.
- The mapping population was evaluated in a separate study under extreme cold storage stress for two weeks. Sugars, organic acids, and RT-qPCR based relative gene expression data were collected to map regions involved in the regulation of these traits under early storage cold stress across three years of field grown material.
- Hyperspectral reflectance collection and data processing methods were evaluated on a panel of genotypes from a breeding population. Prediction models were evaluated using 500 permutations of a randomly sampled cross validation method to determine accurate trait models. The results were then used for application on a hyperspectral imaging platform.

1.8 References

Akeley RV, Mills WR, Cunningham CE, Watts J. 1968. Lenape: A new potato variety high in solids and chipping quality. Am Potato J. 45:142–145.

Ayvaz H, Santos AM, Moyseenko J, Kleinhenz M, Rodriguez-Saona LE. 2015. Application of a portable infrared instrument for simultaneous analysis of sugars, asparagine and glutamine levels in raw potato tubers. Plant Foods Hum Nutr. 70:215–220.

Van Berloo R, Hutten RCB, Van Eck HJ, Visser RGF. 2007. An online potato pedigree database resource. Potato Res. 50:45–57.

Bernardo R. 2010. Breeding for quantitative traits in plants. Second. Woodbury, Minnesota: Stemma Press.

Bethke P. 2013. Reducing the acrylamide content of processed potato products through germplasm improvement: Opportunities, challenges and progress. Asp Appl Biol. 116:79–87.

Bethke PC, Busse JC. 2008. Validation of a simple, colorimetric, microplate assay using amplex red for the determination of glucose and sucrose in potato tubers and other vegetables. Am J Potato Res. 85:414–421.

Bethke PC, Nassar AMK, Kubow S, Leclerc YN, Li XQ, Haroon M, Molen T, Bamberg J, Martin M, Donnelly DJ. 2014. History and origin of Russet Burbank (Netted Gem) a sport of Burbank. Am J Potato Res. 91:594–609.

Bethke PC, Sabba RP, Bussan AJ. 2009. Tuber water and pressure potentials decrease and sucrose contents increase in response to moderate drought and heat stress. Am J Potato Res. 86:519–532.

Bhaskar PB, Wu L, Busse JS, Whitty BR, Hamernik AJ, Jansky SH, Buell CR, Bethke PC, Jiang J. 2010. Suppression of the vacuolar invertase gene prevents cold-induced sweetening in potato. Plant Physiol. 154:939–948.

Blenkinsop RW, Yada RY, Marangoni AG. 2004. Metabolic control of low-temperature sweetening in potato tubers during postharvest storage. Hortic Rev (Am Soc Hortic Sci). 30:317–354.

Boches P, Bassil N V., Rowland L. 2006. Genetic diversity in the highbush blueberry evaluated with microsatellite markers. J Am Hortic Soc. 131:674–686.

Bonierbale MW, Plaisted RL, Tanksley SD. 1988. RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics. 120:1095–103.

Bourke P. 2014. QTL analysis in polyploids QTL analysis in polyploids: Model testing and power calculations. Wageningen, The Netherlands: Wageningen University.

Bourke PM, Voorrips RE, Visser RGF, Maliepaard C. 2015. The double-reduction landscape in tetraploid potato as revealed by a high-density linkage map. Genetics. 201:853–863.

Bradshaw JE. 2007. The canon of potato science: 4. Tetrasomic inheritance. Potato Res. 50:219–222.

Bradshaw JE. 2007. Potato-Breeding strategy. In: Potato Biol Biotechnol Adv Perspectices. First. Amsterdam: Elsevier Ltd; p. 157–174.

Bradshaw JE, Hackett CA, Pande B, Waugh R, Bryan GJ. 2008. QTL mapping of yield, agronomic and quality traits in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*). Theor Appl Genet. 116:193–211.

Bradshaw JE, Mackay GR. 1994. Potato genetics. Wallingford: CAB International.

Bradshaw JE, Pande B, Bryan GJ, Hackett CA, McLean K, Stewart HE, Waugh R. 2004. Interval mapping of quantitative trait loci for resistance to late blight [*Phytophthora infestans* (Mont.) *de bary*], height and maturity in a tetraploid population of potato (*Solanum tuberosum* subsp. *tuberosum*). Genetics. 168:983–995.

Broman K, Sen S. 2009. A guide to QTL mapping with R. J Stat Software, B Rev. 32:1–3.

Brown J, Caligari PDS. 1989. Cross prediction in a potato breeding programme by evaluation of parental material. Theor Appl Genet. 77:246–252.

Brown J, Mackay GR, Bain H, Griffith DW, Allison MJ. 1990. The processing potential of tubers of the cultivated potato, *Solanum tuberosum L.*, after storage at low temperatures. 2. Sugar concentration. Potato Res. 33:219–227.

Brummell DA, Chen RKY, Harris JC, Zhang H, Hamiaux C, Kralicek A V., McKenzie MJ. 2011. Induction of vacuolar invertase inhibitor mRNA in potato tubers contributes to cold-induced sweetening resistance and includes spliced hybrid mRNA variants. J Exp Bot. 62:3519–3534.

Burton WG. 1969. The sugar balance in some British potato varieties during storage. II. The effects of tuber age, previous storage temperature, and intermittent refrigeration upon low-temperature sweetening. Eur Potato J. 12:81–95.

Bussan AJ, Sabba RP, Drilias MJ. 2009. Tuber maturation and potato storability : Optimizing skin set, sugars, and solids (A3884-02). Division of Cooperative Extension of the University of Wisconsin--Extension.

Chawla R, Shakya R, Rommens CM. 2012. Tuber-specific silencing of asparagine synthetase-1 reduces the acrylamide-forming potential of potatoes grown in the field without affecting tuber shape and yield. Plant Biotechnol J. 10:913–924.

Chen S, Hong X, Harris CJ, Sharkey PM. 2004. Sparse modeling using orthogonal forward regression with PRESS statistic and regularization. IEEE Trans Syst Man, Cybern Part B Cybern. 34:898–911.

Chen X, Salamini F, Gebhardt C. 2001. A potato molecular-function map for carbohydrate metabolism and transport. Theor Appl Genet. 102:284–295.

Chunhua L, Du Jianqiang JG, Chunlei C. 2013. Improved partial least squares regression recommendation algorithm. In: Int Conf Adv Inf Eng Educ Sci. Beijing, China.

Coffin RH, Yada RY, Parkin KL, Grodzinski B, Stanley DW. 1987. Effect of low temperature storage on sugar concentrations and chip color of certain processing potato cultivars and selections. J Food Sci. 52:639–645.

Comai L. 2005. The advantages and disadvantages of being polyploid. Nat Rev Geenetics. 6:836–846.

Costa LG, Deng H, Gregotti C, Manzo L, Faustman EM, Bergmark E, Calleman CJ. 1992. Comparative studies on the neuro- and reproductive toxicity of acrylamide and its epoxide metabolite glycidamide in the rat. Neurotoxicology. 13:219–24.

Dale MFB, Sharma SK, Bryan GJ. 2016. Potato breeding now and into the genomics era. Acta Hortic. 1118:1–10.

Dickman L V. 2016. Stem end chipping defect incidence and severity in potatoes (*Solanum tuberosum*). Madison, Wisconsin: University of Wisconsin-Madison.

Doerge RW. 2002. Multifactorial genetics mapping and analysis of quantitative trait loci in experimental populations. Nat Rev Genet. 3:43–52.

Douches D, Hirsch CN, Manrique-Carpintero NC, Massa AN, Coombs J, Hardigan M, Bisognin D, De Jong W, Buell CR. 2014. The contribution of the solanaceae coordinated agricultural project to potato breeding. Potato Res. 57:215–224.

Douches DS, Jastrzebski K, Maas D, Chase RW. 1996. Assessment of potato breeding over the past century. Crop Sci. 36:1544–1552.

Driskill EP, Knowles LO, Knowles NR. 2007. Temperature-induced changes in potato processing quality during storage are modulated by tuber maturity. Am J Potato Res. 84:367–383.

Dull G, Birth G, Leffler R. 1989. Use of near infrared analysis for the nondestructive measurement of dry matter in potatoes. Am Potato J. 66.

Eldredge EP, Holmes ZA, Mosley a. R, Shock CC, Stieber TD. 1996. Effects of transitory water stress on potato tuber stem-end reducing sugar and fry color. Am Potato J. 73:517–530.

Falconer DS, Mackay T. 1996. Introduction to quantitative genetics. Fourth Edi. Essex, England: Pearson Education Limited.

FDA. 2013. Guidance for industry: Acrylamide in foods. FDA Food Guid [Internet]. [cited 2017 Feb 1]. Available from: http://www.fda.gov/FoodGuidances

Fehr WR. 1991. Principles of Cultivar Development: Volume 1. Fehr E, Jessen H, editors. Ames, Iowa: Macmillan Publishing Company.

Gar O, Sargent DJ, Tsai CJ, Pleban T, Shalev G, Byrne DH, Zamir D. 2011. An autotetraploid linkage map of rose (Rosa hybrida) validated using the strawberry (*Fragaria vesca*) genome sequence. PLoS One. 6.

Gebhardt C, Menendez C, Chen X, Li L, Schäfer-Pregl R, Salamini F. 2005. Genomic approaches for the improvement of tuber quality traits in potato. Acta Hortic. 684:85–91.

Gebhardt C, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F. 1989. RFLP analysis and linkage mapping in *Solanum tuberosum*. Theor Appl Genet. 78:65–75.

Gökmen V, Palazoğlu TK. 2008. Acrylamide formation in foods during thermal processing with a focus on frying. Food Bioprocess Technol. 1:35–42.

Gould WA, Plimpton S. 1985. Quality evaluation of potato cultivars for processing. North Cental Regional Research Publication 305. The Ohio State University.

Greiner S, Rausch T, Sonnewald U, Herbers K. 1999. Ectopic expression of a tobacco invertase inhibitor homolog prevents cold-induced sweetening of potato tubers. Nat Biotechnol. 17:708–11.

Groth J, Song Y, Kellermann A, Schwarzfischer A. 2013. Molecular characterisation of resistance against potato wart races 1, 2, 6 and 18 in a tetraploid population of potato (*Solanum tuberosum* subsp. *tuberosum*). J Appl Genet. 54:169–178.

Haase NU. 2004. Estimation of dry matter and starch concentration in potatoes by determination of under-water weight and near infrared spectroscopy. 46:117–127.

Haase NU. 2011. Prediction of potato processing quality by near infrared reflectance spectroscopy of

ground raw tubers. J Near Infrared Spectrosc. 19:37–45.

Habib AT, Brown HD. 1957. Role of reducing sugars and amino acids in the browning of potato chips. Food Technol. 11:85–89.

Hackett C, Bradshaw JE, Bryan GJ. 2014. QTL mapping in autotetraploids using SNP dosage information. Theor Appl Genet. 127:1885–1904.

Hackett C, Bradshaw JE, McNicol JW. 2001. Interval mapping of quantitative trait loci in autotetraploid species. Genetics. 159:1819–1832.

Hackett C, Luo ZW. 2003. TetraploidMap: construction of a linkage map in autotetraploid species. J Hered. 94:358–359.

Hackett C, McLean K, Bryan GJ. 2013. Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. PLoS One. 8:e63939.

Hämäläinen JH, Watanabe KN, Valkonen JPT, Arihara A, Plaisted RL, Pehu E, Miller L, Slack SA. 1997. Mapping and marker-assisted selection for a gene for extreme resistance to potato virus Y. Theor Appl Genet.:192–197.

Hamilton JP, Hansey CN, Whitty BR, Stoffel K, Massa AN, Van Deynze A, De Jong WS, Douches DS, Buell CR. 2011. Single nucleotide polymorphism discovery in elite North American potato germplasm. BMC Genomics. 12:302-313.

Hardigan MA, Crisovan E, Hamiltion JP, Kim J, Laimbeer P, Leisner CP, Manrique-Carpintero NC, Newton L, Pham GM, Vaillancourt B, et al. 2016. Genome reduction uncovers a large dispensable genome and adaptive role for copy number variation in asexually propagated *Solanum tuberosum*. Plant Cell. 28:388-405.

Hartmann R, Büning-Pfaue H. 1998. NIR determination of potato constituents. Potato Res. 41:327–334.

Haynes KG, Douches DS. 1993. Estimation of the coefficient of double reduction in the cultivated tetraploid potato. Theor Appl Genet. 85:857–862.

Helgerud T, Knutsen SH, Afseth NK, Stene KF, Rukke EO, Ballance S. 2016. Evaluation of hand-held instruments for representative determination of glucose in potatoes. Potato Res.:1–14.

Heltoft P, Wold AB, Molteberg EL. 2016. Effect of ventilation strategy on storage quality indicators of processing potatoes with different maturity levels at harvest. Postharvest Biol Technol. 117:21–29.

Hill LM, Reimholz R, Schröder R, Nielsen TH, Stitt M. 1996. The onset of sucrose accumulation in cold-stored potato tubers is caused by an increased rate of sucrose synthesis and coincides with low levels of hexose-phosphates, an activation of sucrose phosphate synthase and the appearance of a new form of amylase. Plant Cell Env. 19:1223–1237.

Hiller LK, Koller DC, Thornton RE. 1985. Physiological disorders of potato tubers. In: Potato Physiol. Orlando, FL: Academic Press Inc.; p. 389–455.

Hirsch CN, Hirsch CD, Felcher K, Coombs J, Zarka D, Van Deynze A, De Jong W, Veilleux RE, Jansky S, Bethke P, et al. 2013. Retrospective view of North American potato (*Solanum tuberosum L.*) breeding in the 20th and 21st centuries. G3 Genes|Genomes|Genetics. 3:1003–13.

Hodge JE. 1953. Browning reactions in model systems. J Agric Food Chem. 1:928–943.

Iovene M, Zhang T, Lou Q, Buell CR, Jiang J. 2013. Copy number variation in potato - An asexually propagated autotetraploid species. Plant J. 75:80–89.

Iritani WM, Weller L. 1973. The development of translucent end tubers. Am Potato J. 50:223–233.

Isherwood FA. 1973. Starch-sugar interconversion in *Solanum tuberosum*. Phytochemistry. 12:2579–2591.

Isla MI, Leal DP, Vattuone MA, Sampietro AR. 1992. Cellular localization of the invertase, proteinaceous inhibitor and lectin from potato tubers. Phytochemistry. 31:1115–1118.

Jansky S. 2006. Overcoming hybridization barriers in potato. Plant Breed. 125:1–12.

Jansky SH. 2008. Genotypic and environmental contributions to baked potato flavor. Am J Potato Res. 85.

Jansky SH. 2010. Potato Flavor. Am J Potato Res. 87:209–217.

Jones GH, Khazanehdari KA. 1996. Meiosis in the leek (*Allium porrum L.*) revisited. II. Metaphase I observations. Heredity (Edinb). 76:186–191.

Kao CH, Zeng ZB, Teasdale RD. 1999. Multiple interval mapping for quantitative trait loci. Genetics. 152:1203–16.

Kashian R, Depas J, Fogarty P, Peterson J. 2014. Potato production in Wisconsin: Analyzing the economic impact. Whitewater, Wisconsin.

Khan MS. 2012. Assessing genetic variation in growth and development of potato. Wageningen, The Netherlands: Wageningen University.

Khu DM, Lorenzen J, Hackett CA, Love SL. 2008. Interval mapping of quantitative trait loci for corky ringspot disease resistance in a tetraploid population of potato (*Solanum tuberosum* subsp. *tuberosum*). Am J Potato Res. 85:129–139.

Kleinkopf GE. 1976. Translucent-end of potatoes. Current Information Series No. 488. Division of Cooperative Extension of the University of Idaho College of Agriculture.

Kloosterman B, Abelenda JA, Gomez M del MC, Oortwijn M, De Boer JM, Kowitwanich K, Horvath BM, Van Eck HJ, Smaczniak C, Prat S, et al. 2013. Naturally occurring allele diversity allows potato cultivation in northern latitudes. Nature. 495:246–250.

Kolbe H, Stephan-beckmann S. 1997. Development, growth and chemical composition of the potato crop (*Solanum tuberosum L.*). II. Tuber and whole plant. Potato Res. 40:135–153.

Krauss A, Marschner H. 1984. Growth rate and carbohydrate metabolism of potato tubers exposed to high temperatures. Potato Res. 27:297–303.

Kumar D, Singh BP, Kumar P. 2004. An overview of the factors affecting sugar content of potatoes. Ann Appl Biol. 145:247–256.

De Leon N. 2006. NIRS Prediction Equations : Development of NIRS prediction equations [Internet]. [cited 2016 Nov 16]. Available from: http://cornbreeding.wisc.edu/analytics/nirs-prediction-equations/

Li J, Das K, Fu G, Tong C, Li Y, Tobias CM, Wu R. 2010. EM algorithm for mapping quantitative trait loci in multivalent tetraploids. Int J Plant Genomics. 2010:216547.

Li L, Tacke E, Hofferbert H-R, Lübeck J, Strahwald J, Draffehn AM, Walkemeier B, Gebhardt C. 2013. Validation of candidate gene markers for marker-assisted selection of potato cultivars with improved tuber quality. Theor Appl Genet. 126:1039–52.

Liu X, Song B, Zhang H, Li XQ, Xie C, Liu J. 2010. Cloning and molecular characterization of putative

invertase inhibitor genes and their possible contributions to cold-induced sweetening of potato tubers. Mol Genet Genomics. 284:147–159.

Liu X, Zhang C, Ou Y, Lin Y, Song B, Xie C, Liu J, Li X-Q. 2011. Systematic analysis of potato acid invertase genes reveals that a cold-responsive member, StvacINV1, regulates cold-induced sweetening of tubers. Mol Genet Genomics. 286:109–18.

LoPachin Jr. RM, Lehning EJ. 1994. Acrylamide-induced distal axon degeneration: a proposed mechanism of action. Neurotoxicology. 15:247–259.

López A, Arazuri S, García I, Mangado J, Jarén C. 2013. A review of the application of near-infrared spectroscopy for the analysis of potatoes. J Agric Food Chem. 61:5413–24.

Lorberth R, Ritte G, Willmitzer L, Kossmann J. 1998. Inhibition of a starch-granule–bound protein leads to modified starch and repression of cold sweetening. Nat Biotechnol. 16:473–477.

Lorenz A, Chao S, Asoro F, Heffner EL, Hayashi T, Iwata H, Smith KP, Sorrells ME, Jannink J-L. 2011. Chapter two - Genomic selection in plant breeding: Knowledge and prospects. Adv Agron. 110:77–123.

Love S, Pavek J, Thompson-Johns A, Bohl W. 1998. Breeding progress for potato chip quality in North American cultivars. Am J Potato Res. 75:27–36.

Lu Q, Cui Y, Wu R. 2004. A multilocus likelihood approach to joint modeling of linkage, parental diplotype and gene order in a full-sib family. BMC Genet. 5:20-34.

Lugt C. 1960. Second-growth phenomena. Eur Potato J. 3:307–324.

Lulai EC, Orr PH. 1979. Influence of potato specific gravity on yield and oil content of chips. Am Potato J. 56:379-390.

Lulai EC, Orr PH. 1993. Determining the feasibility of measuring genotypic differences in skin-set. Am Potato J. 70:599–609.

Luo ZW, Hackett C, Bradshaw JE, McNicol JW, Milbourne D. 2000. Predicting parental genotypes and gene segregation for tetrasomic inheritance. Theor Appl Genet. 100:1067–1073.

Luo ZW, Hackett C, Bradshaw JE, McNicol JW, Milbourne D. 2001. Construction of a genetic linkage map in tetraploid species using molecular markers. Genetics. 157:1369–1385.

Lynch M, Walsh B. 1998. Genetics and analysis of quantitative traits. Vol. 1. Sunderland, MA: Sinauer.

Maitra S, Yan J. 2008. Principle component analysis and partial least squares : Two dimension reduction techniques for regression. Casualty Actuar Soc.:79–90.

Martins SIFS, Jongen WMF, Van Boekel MAJS. 2000. A review of Maillard reaction in food and implications to kinetic modelling. Trends Food Sci Technol. 11:364–373.

Mason M. 1983. Determination of glucose, sucrose, lactose, and ethanol in foods and beverages, using immobilized enzyme electrodes. Journal-Association Off Anal Chem. 66:981–984.

Massa AN, Manrique-Carpintero NC, Coombs JJ, Zarka DG, Boone AE, Kirk WW, Hackett CA, Bryan GJ, Douches DS. 2015. Genetic linkage mapping of economically important traits in cultivated tetraploid potato (*Solanum tuberosum L.*). G3 Genes|Genomes|Genetics. 5:2357–2364.

Matsuura-Endo C, Ohara-Takada A, Chuda Y, Ono H, Yada H, Yoshida M, Kobayashi A, Tsuda S, Takigawa S, Noda T, et al. 2006. Effects of storage temperature on the contents of sugars and free amino acids in tubers from different potato cultivars and acrylamide in chips. Biosci Biotechnol Biochem.

70:1173–80.

McCord PH, Sosinski BR, Haynes KG, Clough ME, Craig Yencho G. 2011. Linkage mapping and QTL analysis of agronomic traits in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*). Crop Sci. 51:771–785.

McCord PH, Sosinski BR, Haynes KG, Clough ME, Yencho GC. 2011. QTL mapping of internal heat necrosis in tetraploid potato. Theor Appl Genet. 122:129–142.

McGregor I. 2007. The fresh potato market. In: Potato Biol Biotechnol Adv Perspectices. First Edition. Amsterdam: Elsevier Ltd; p. 3–25.

McKenzie MJ, Chen RKY, Harris JC, Ashworth MJ, Brummell DA. 2013. Post-translational regulation of acid invertase activity by vacuolar invertase inhibitor affects resistance to cold-induced sweetening of potato tubers. Plant, Cell Environ. 36:176–185.

McKenzie MJ, Sowokinos JR, Shea IM, Gupta SK, Lindlauf RR, Anderson JAD. 2005. Investigations on the role of acid invertase and UDP-glucose pyrophosphorylase in potato clones with varying resistance to cold-induced sweetening. Amer J Potato Res. 82:231–239.

Mehmood T, Liland KH, Snipen L, Sæbø S. 2012. A review of variable selection methods in partial least squares regression. Chemom Intell Lab Syst. 118:62–69.

Mertz C, Matthews V, Losh D, Benz S. 2016. National agricultural statistics service press release. Washington, D.C.

Meyer RC, Milbourne D, Hackett CA, Bradshaw JE, McNichol JW, Waugh R. 1998. Linkage analysis in tetraploid potato and association of markers with quantitative resistance to late blight (*Phytophthora infestans*). Mol Gen Genet. 259:150–160.

Milbourne D, Bradshaw JE, Hackett C. 2009. Molecular mapping and breeding in polyploid crop plants. In: Kole C, Abbott AG, editors. Princ Pract Plant Genomics, vol 2 Mol Breed. New Hampshire, Plymouth, Jersey, USA: Science Publishers; p. 355–394.

Müller-Thurgau H. 1882. Ueber zuckeranhäufung in pflanzentheilen in folge niederer temperatu. Landwirtsch Jahrb. 11:751–828.

National Potato Council. 2013. 2013 potato statistical yearbook. Washington, D.C.

National Potato Council. 2015. Potato statistical yearbook 2015. Washington, D.C.

Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI, Lammertyn J. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biol Technol. 46:99–118.

Nielsen TH, Deiting U, Stitt M. 1997. A beta-amylase in potato tubers is induced by storage at low temperature. Plant Physiol. 113:503–510.

Nikolov L, Jakovljevic JB, Boskov M. 1984. High performance liquid chromatographic separation of oligosaccharides using amine modified silica columns. Starch-Stärke. 36:97–100.

Nilsson T. 2011. Comparison of the FOSS NIR global ANN calibration against reference methods : A five year pan-European study. Hilleroed, Denmark.

Nissen M. 1955. The weight of potatoes in water. Am Potato J. 32:332–339.

North American Potato Market News December 21, 2016. 2016. North Amercian Potato Mark News.

25:1–4.

Ortiz R. 2001. The state of the use of potato genetic diversity. In: Cooper H, Spillane C, Hodgkin T, editors. Broadening Genet Base Crop Prod. New York, NY: CABI, FAO, IPGRI; p. 181–200.

Osborne BG. 2000. Near-infrared spectroscopy in food analysis. In: Meyers R, editor. Encycl Anal Chem. Chichester: John Wiley & Sons Ltd; p. 1–14.

Osten DW. 1988. Selection of optimal regression models via cross-validation. J Chemom. 2:39–48.

Paterson A, Wing R. 1993. Genome mapping in plants. Curr Opin Plant Biol. 4:142–147.

Peiris KHS, Dull G, Leffler RG, Kays SJ. 1999. Spatial variability of soluble solids or dry-matter content within individual fruits, bulbs, or tubers: Implications for the development and use of NIR spectrometric techniques. HortScience. 34:114–118.

Pressey R. 1969. Role of invertase in the accumulation of sugars in cold-stored potatoes. Am Potato J. 46:291–297.

Pritchard MK, Adam LR. 1992. Preconditioning and storage of chemically immature Russet Burbank and Shepody potatoes. Am Potato J. 69:805–815.

Pritchard MK, Adam LR. 1994. Relationships between fry color and sugar concentrations in stored Russet Burbank and Shepody potatoes. Am J Potato Res. 71:59–68.

Quiros CF. 1982. Tetrasomic segregation for multiple alleles in alfalfa. Genetics. 101:117–27.

Rady AM, Guyer DE. 2015a. Evaluation of sugar content in potatoes using NIR reflectance and wavelength selection techniques. Postharvest Biol Technol. 103:17–26.

Rady AM, Guyer DE. 2015b. Rapid and/or nondestructive quality evaluation methods for potatoes: A review. Comput Electron Agric. 117:31–48.

Rady AM, Guyer DE, Kirk W, Donis-González IR. 2014. The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict processing-related constituents of potatoes. J Food Eng. 135:11–25.

Rak K. 2015. Breeding and molecular genetics for the improvement of cold storage potato chip quality. University of Wisconsin-Madion. Ann Arbor, MI: ProQuest LLC.

Ramakrishnan AP, Ritland CE, Blas Sevillano RH, Riseman A. 2015. Review of potato molecular markers to enhance trait selection. Am J Potato Res. 92:455–472.

Rausch T, Greiner S. 2004. Plant protein inhibitors of invertases. Biochim Biophys Acta - Proteins Proteomics. 1696:253–261.

Reeve RM, Hautala E, Weaver ML. 1969. Anatomy and compositional variations within potatoes III. Gross Compositional Gradients. Am Potato J. 47:148–162.

Ribaut J-M, Hoisington D. 1998. Marker-assisted selection: new tools and strategies. Trends Plant Sci. 3:236–239.

Richardson DL, Davies H V, Ross HA, Mackay GR. 1990. Invertase activity and its relation to hexose accumulation in potato tubers. J Exp Bot. 41:95–99.

Rosén J, Hellenäs K-E. 2002. Analysis of acrylamide in cooked foods by liquid chromatography tandem mass spectrometry. Analyst. 127:880–882.

Sabba RP, Bussan AJ, Michaelis B a., Hughes R, Drilias MJ, Glynn MT. 2007. Effect of planting and vine-kill timing on sugars, specific gravity and skin set in processing potato cultivars. Am J Potato Res. 84:205–215.

Sax K. 1923. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus Vulgaris*. Genetics. 8:552–560.

Scanlon MG, Pritchard MK, Adam LR. 1999. Quality evaluation of processing potatoes by near infrared reflectance. 771:763–771.

Schäfer-Pregl R, Ritter E, Concilio L, Hesselbach J, Lovatti L, Walkemeier B, Thelen H, Salamini F, Gebhardt C. 1998. Analysis of quantitative trait loci (QTLs) and quantitative trait alleles (QTAs) for potato tuber yield and starch content. Theor Appl Genet. 97:834–846.

Schippers PA. 1976. The relationship between specific gravity and percentage dry matter in potato tubers. Am Potato J. 53:111–122.

Schmitz-Carley CA, Coombs JJ, Douches DS, Bethke PC, Palta JP, Novy RG, Endelman JB. 2017. Automated tetraploid genotype calling by hierarchical clustering. Theor Appl Genet. 0:0.

Schreiber C. 1997. Declining market share forced Frito Lay to make big changes. Livest Wkly.:1–7.

Semagn K, Bjornstad A, Xu Y. 2010. The genetic dissection of quantitative traits in crops. Electron J Biotechnol. 13:1–45.

Shallenberger R, Smith O, Treadway R. 1959. Food color changes - Role of the sugars in the browning reaction in potato chips. J Agric Food Chem. 7:274–277.

Shekhar VC, Iritani WM. 1979. Influence of moisture stress during growth on $^{14}CO_2$ fixation and translocation in *Solanum tuberosum L.* Am Potato J. 56:307–311.

Shock CC, Holmes ZA, Stieber TD, Eldredge EE, Zhang P. 1993. The effect of timed water stress on quality, total solids, and reducing sugar content of potatoes. Am Potato J. 70:227–241.

Simmonds NW. 1977. Relations between specific gravity, dry matter content and starch content of potatoes. Potato Res. 20:137–140.

Singh B, Wang N, Prasher S, Ngadi M. 2004. A spectroscopic technique for water content determination in potato. ASAE Annu Int Meet 2004.:1–11.

Soller M, Brody T, Genizi A. 1976. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor Appl Genet. 47:35–39.

Sowokinos JR. 1978. Relationship of harvest sucrose content to processing maturity and storage life of potatoes. Am Potato J. 55:333–344.

Sowokinos JR. 2001a. Biochemical and molecular control of cold-induced sweetening in potatoes. Am J Potato Res. 78:221–236.

Sowokinos JR. 2001b. Allele and isozyme patterns of UDP-glucose pyrophosphorylase as a marker for cold-sweetening resistance in potatoes. Am J Potato Res. 78:57–64.

Sowokinos JR, Lulai EC, Knoper JA. 1985. Translucent tissue defects in *Solanum tuberosum L.* Plant Physiol. 78:489–494.

Sowokinos JR, Preston DA. 1988. Maintenance of potato processing quality by chemical maturity monitoring (CMM). Station Bulletin 586-1988. Minnesota Agricultural Experimental Station. St. Paul,

Minnesota.

Sowokinos JR, Shock CC, Stieber TD, Eldredge EP. 2000. Compositional and enzymatic changes associated with the sugar-end defect in Russet Burbank potatoes. Am J Potato Res. 77:47–56.

Spooner DM, Ghislain M, Simon R, Jansky SH, Gavrilenko T. 2014. Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. Bot Rev. 80:283–383.

Stevenson FJ. 1957. Red Dot Foods, Inc., and its potato research program. Am Potato J. 34:136–141.

Stushnoff C, Holm D, Thompson MD, Jiang W, Thompson HJ, Joyce NI, Wilson P. 2008. Antioxidant properties of cultivars and selections from the Colorado potato breeding program. Am J Potato Res. 85:267–276.

Subedi PP, Walsh KB. 2009. Assessment of potato dry matter concentration using short-wave near-infrared spectroscopy. Potato Res. 52:67–77.

Tai GC, Young DA. 1984. Early generation selection for important agronomic characteristics in a potato breeding population. Am Potato J. 61:419–434.

Tanksley SD, Ganal MW, Prince JP, De Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, et al. 1992. High density molecular linkage maps of the tomato and potato genomes. Genetics. 132:1141–1160.

Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, Beck-Bunn T. 1996. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theor Appl Genet. 92:213–224.

Tareke E, Karlsson P, Eriksson S, Rnqvist MT. 2002. Analysis of acrylamide, a carcinogen formed in heated foodstuffs. :4998–5006.

Tavaud M, Zanetto A, David JL, Laigret F, Dirlewanger E. 2004. Genetic relationships between diploid and allotetraploid cherry species (*Prunus avium*, *Prunus* x *gondouinii* and *Prunus cerasus*). Heredity (Edinb). 93:631–638.

Thompson-Johns A. 1998. Inheritance of the sugar end disorder in potato (*Solanum tuberosum L.*). Moscow, ID: University of Idaho.

Thompson AL, Bamberg J, Bizimungu B, editors. 2015. NCCC215: Potato breeding and genetics technical committee. In: North Cent Coord Comm 215. Chicago, IL.

Thompson AL, Love SL, Sowokinos JR, Thornton MK, Shock CC. 2008. Review of the sugar end disorder in potato (*Solanum tuberosum L.*). Am J Potato Res. 85:375–386.

Tobias RD. 1995. An introduction to partial least squares regression. Proc Ann SAS Users Gr Int Conf, 20th, Orlando, FL.:2–5.

USDA-NASS. 2016. Potatoes 2015 summary. Washington, D.C.

Voorrips RE, Gort G, Vosman B. 2011. Genotype calling in tetraploid species from bi-allelic marker data using mixture models. BMC Bioinformatics. 12:172.

Vos PG, Uitdewilligen JGAML, Voorrips RE, Visser RGF, van Eck HJ. 2015. Development and analysis of a 20K SNP array for potato (*Solanum tuberosum*): an insight into the breeding history. Theor Appl Genet. 128:2387–2401.

Wakeling IN, Morris JJ. 1993. A test of significance for partial least squares regression. J Chemom.

7:291–304.

Wang Y, Bethke PC. 2013. Effects of infection on stem-end chip defect development in potatoes. Crop Sci. 53:595.

Wang Y, Bethke PC, Bussan AJ, Glynn MT, Holm DG, Navarro FM, Novy RG, Palta JP, Pavek MJ, Porter GA, et al. 2016. Acrylamide-forming potential and agronomic properties of elite US potato germplasm from the national fry processing trial. Crop Sci. 56:30–39.

Wang Y, Bethke PC, Drilias MJ, Schmitt WG, Bussan AJ. 2015. A multi-year survey of stem-end chip defect in chipping potatoes (*Solanum tuberosum L.*). Am J Potato Res. 92:79–90.

Wang Y, Bussan AJ, Bethke PC. 2012. Stem-end defect in chipping potatoes (*Solanum tuberosum L.*) as influenced by mild environmental stresses. Am J Potato Res. 89:392–399.

Wells HF. 2012. USDA ERS - Vegetables & pulses : Potatoes [Internet]. [cited 2016 May 31]:1–4. Available from: http://www.ers.usda.gov/topics/crops/vegetables-pulses/potatoes.aspx

Wharton P, Thornton M, Olsen N, Miller J, Whitworth J. 2015. Potato progress. Moscow, ID.

Wiberley-Bradford AE, Busse JS, Jiang J, Bethke PC. 2014. Sugar metabolism, chip color, invertase activity, and gene expression during long-term cold storage of potato (*Solanum tuberosum*) tubers from wild-type and vacuolar invertase silencing lines of Katahdin. BMC Res Notes. 7:801.

Wilson AM, Work TM, Bushway AA, Bushway RJ. 1981. HPLC determination of fructose, glucose, and sucrose in potatoes. J Food Sci. 46:300–301.

Wisconsin Historical Society. 2003. Chip chat: Red Dot and the potato chip. Wisconsin Hist Soc [Internet]. [cited 2012 Jan 1]. Available from: http://www.wisconsinhistory.org/museum/exhibits/chip/index.asp

Wold S, Ruhe A, Wold H, Dunn WJ. 1984. The collinearity problem in linear regression: The partial least squares (PLS) approach to generalized inverses. Soc Industiral Appl Math. 5:735–743.

Wold S, Sjöström M, Eriksson L. 2001. PLS-regression: A basic tool of chemometrics. Chemom Intell Lab Syst. 58:109–130.

Wright AJ, Mowers RP. 1994. Multiple regression for molecular-marker, quantitative trait data from large F2 populations. Theor Appl Genet. 89:305–312.

Wu R, Gallo-Meagher M, Littell RC, Zeng ZB. 2001. A general polyploid model for analyzing gene segregation in outcrossing tetraploid species. Genetics. 159:869–882.

Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J, et al. 2011. Genome sequence and analysis of the tuber crop potato. Nature. 475:189–95.

Ye J, Shakya R, Shrestha P, Rommens CM. 2010. Tuber-specific silencing of the acid invertase gene substantially lowers the acrylamide-forming potential of potato. J Agric Food Chem. 58:12162–12167.

Zeng ZB. 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc Natl Acad Sci U S A. 90:10972–6.

Zheng C, Voorrips RE, Jansen J, Hackett CA, Ho J, Bink MCAM. 2016. Probabilistic multilocus haplotype reconstruction in outcrossing tetraploids. Genetics. 203:119–131.

Zhu X, Richael C, Chamberlain P, Busse JS, Bussan AJ, Jiang J, Bethke PC. 2014. Vacuolar invertase gene silencing in potato (*Solanum tuberosum L.*) improves processing quality by decreasing the frequency

of sugar-end defects. PLoS One. 9.

Zorrilla C, Navarro F, Vega S, Bamberg J, Palta J. 2014. Identification and selection for tuber calcium, internal quality and pitted scab in segregating “Atlantic” x “Superior” reciprocal tetraploid populations. Am J Potato Res. 91:673–687.

Zrenner R, Schüler K, Sonnewald U. 1996. Soluble acid invertase determines the hexose-to-sucrose ratio in cold-stored potato tubers. Planta. 198:246–52.

Zrenner R, Willmitzer L, Sonnewald U. 1993. Analysis of the expression of potato uridinediphosphate-glucose pyrophosphorylase and its inhibition by antisense RNA. Planta. 190:247–252.

Table 1.1. Autotetraploid genotype class ratio for a single allele in progeny from crosses of the different parental dosage classes. Numbers designate proportion of AAAA:AAAB:AABB:ABBB:BBBB. These segregation patterns are free from double reduction

Parent B \ Parent A	AAAA	AAAB	AABB	ABBB	BBBB
AAAA	**1:0:0:0:0**	**1:1:0:0:0**	**1:4:1:0:0**	**0:1:1:0:0**	**0:0:1:0:0**
AAAB	**1:1:0:0:0**	**1:2:1:0:0**	**1:5:5:1:0**	**0:1:2:1:0**	**0:0:1:1:0**
AABB	**1:4:1:0:0**	**1:5:5:1:0**	**1:8:18:8:1**	**0:1:5:5:1**	**0:0:1:4:1**
ABBB	**0:1:1:0:0**	**0:1:2:1:0**	**0:1:5:5:1**	**0:0:1:2:1**	**0:0:0:1:1**
BBBB	**0:0:1:0:0**	**0:0:1:1:0**	**0:0:1:4:1**	**0:0:0:1:1**	**0:0:0:0:1**

Figure 1.1. Rating scale for stem-end chipping defect (SECD) as developed by (Wang et al., 2012)

Figure 1.2. An era study evaluating tuber solids in chip processing cultivars. Lenape and cultivars with it present in their pedigree perform favorably for dry matter accumulation (Love et al. 1998).

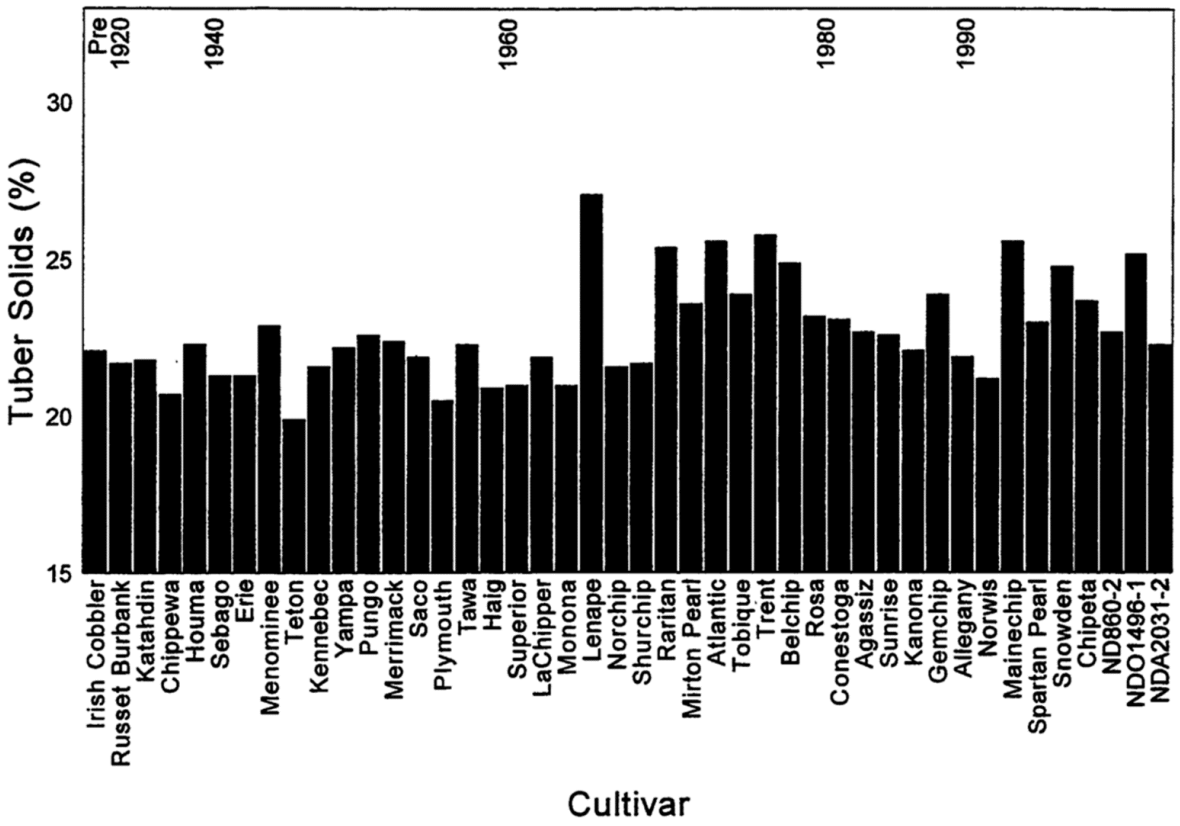
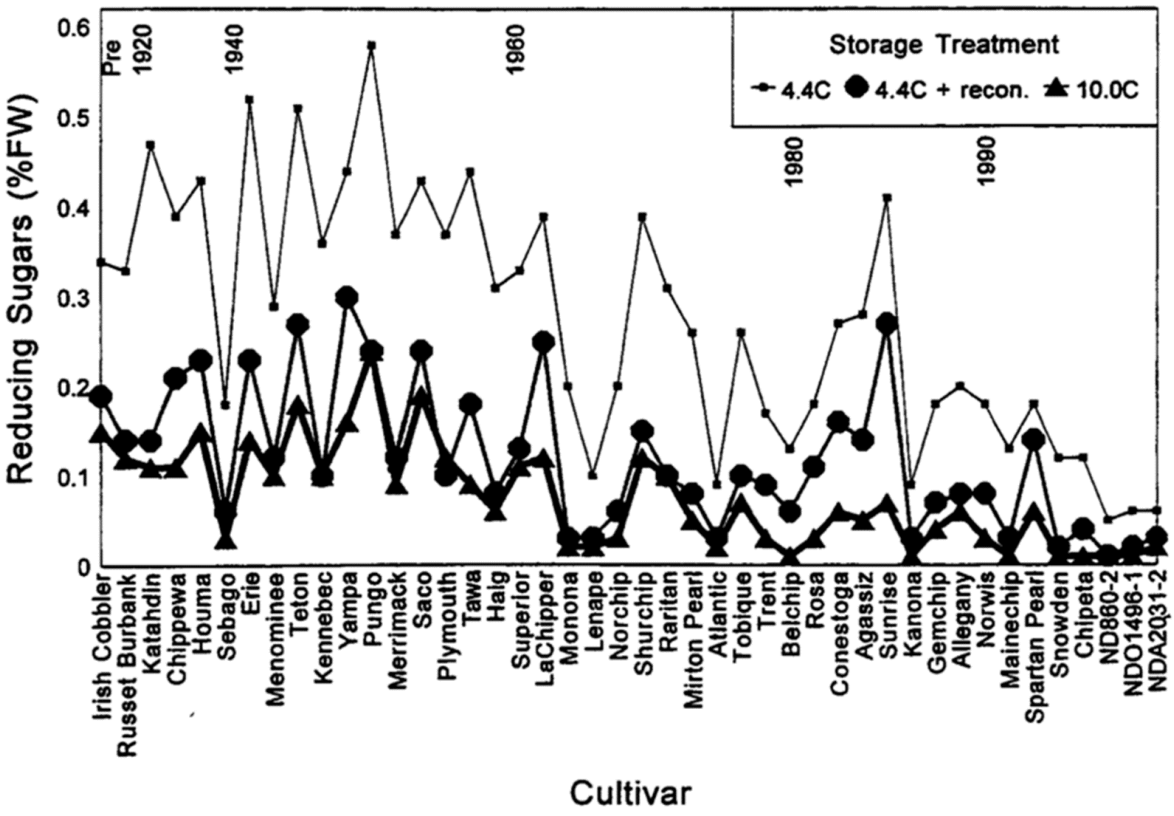


Figure 1.3. An era study evaluating reducing sugar accumulation in chip processing cultivars under different storage temperatures. Lenape and cultivars with it present in their pedigree perform favorably for cold storage sweetening resistance (Love et al. 1998).



Chapter 2: Metabolite and expression-based QTL study of loci contributing to cold-induced sweetening and construction of a linkage map in a prominent chip processing population

2.1 Abstract

The objective of this study was to identify QTL controlling the regulation and biochemistry of early cold induced sweetening (CIS) in chip processing potatoes. A bi-parental population was developed by crossing the chip processing varieties Wauseon and Lenape to develop 191 full siblings. The population was genotyped using the SolCAP 8300 Infinium Chip and a linkage map was constructed using 2355 of the single nucleotide polymorphism (SNP) markers and the process is described herein. Parents and progeny were evaluated in 2012, 2013, and 2014 for sucrose, glucose, fructose, malic and citric acids, and relative expression levels of vacuolar acid invertase (*VInv*) two weeks after an abrupt transition to 3^0C storage. QTL mapping was performed using the TetraploidMap 2.0 software and 33 significant QTL were detected: 4 for sugar concentration traits, 5 for organic acids, 15 for sugar relationship traits, and 9 for *VInv* relative expression traits. QTL were detected on all chromosomes except VIII, IX, and X and there were regions where multiple QTL from different traits overlapped. The largest effect QTL was for fructose on chromosome XII, co-located with a large effect QTL for glucose, and mapped to a position within 200 kb of phosphofructokinase genes. *VInv* relative expression had large genotype x year effects and low heritability estimates but one and two year data sets detected QTL on chromosomes III and VII. The detected QTL provide additional explanation of Lenape's resistance to CIS.

2.2 Introduction

A highly desirable trait for processing potatoes is resistance to cold-induced sweetening (CIS), in which a cold-mediated biological pathway accelerates the net conversion of starch to sugars, generally at temperatures less than 10° C (Müller-Thurgau, 1882)(Burton, 1969)(Coffin et al., 1987). CIS resistance

allows potatoes to be stored at temperatures that decrease rates of respiration, losses due to disease, and prolong dormancy (Brown et al., 1990a; b) without a loss of processing quality. CIS is a complex trait and many of the enzymes involved have been characterized at the molecular and biochemical levels (Kumar et al., 2004)(Sowokinos, 2001b)(Chen et al., 2001). CIS is primarily comprised of several metabolic processes; the breakdown of starch, the synthesis of sucrose, and the hydrolysis of sucrose into the six carbon reducing sugars glucose and fructose (Sowokinos, 2001b). During CIS, the synthesis of sucrose is accelerated (Isherwood, 1973) and much of this sucrose is hydrolyzed in the vacuole to form glucose and fructose (Isla et al., 1992)(Blenkinsop et al., 2004). Most of the variation in fry color for potatoes out of cold storage is highly associated with the concentration of these reducing sugars that accumulate in the vacuole (Pritchard and Adam, 1994).

One potato variety that had greater resistance to CIS when released than other varieties was the cultivar Lenape (Love et al., 1998). Lenape was released in 1968 and was derived from a cross between USDA seedlings B3672-3 and 47156. B3672-3 contained a *S. chacoense* accession in its background that presumably conferred CIS resistance (Akeley et al. 1968). While not widely adopted itself due to concerns over moderately high tuber glykoalkaloid content, Lenape is prominent in the pedigree of numerous successful chipping and fry processing cultivars including Snowden, Atlantic, Belchip, Gemchip, Pike, NorValley, Tundra, Accumulator, Lamoka, Highland Russet, Blazer Russet, and Alturas amongst others (Van Berloo et al., 2007). The alleles that contribute to its CIS resistance are not well understood.

Many enzymes influence the rate or extent of CIS. These include starch phosphorylase, UDP glucose pyrophosphorylase, sucrose phosphate synthase, β – amylase, acid invertase and others (Zrenner et al., 1993)(McKenzie et al., 2005)(Sowokinos, 2001a)(Hill et al., 1996)(Nielsen et al., 1997)(Lorberth et al., 1998)(Richardson et al., 1990)(Zrenner et al., 1996). The production of reducing sugars has been strongly associated with the activity of VInv (Pressey, 1969)(Richardson et al., 1990). Transgenic plants expressing anti-sense or RNA interference (RNAi) constructs of *VInv* have provided evidence for the key role this enzyme plays in CIS. For example, Katahdin lines transformed with an RNAi-silencing construct

resulted in tubers that had minimal Maillard reaction browning and low reducing sugar (<0.3 mg g^{-1} FW) concentration after storage at 4 degrees C. Severity of CIS decreased as the level of *VInv* expression decreased (Bhaskar et al. 2010). Expression profiles show that *VInv* is upregulated in the tuber tissue under cold temperatures (Matsuura-Endo et al., 2006)(Liu et al., 2011)(Wiberley-Bradford et al., 2014).

Understanding the processes that regulate the abundance of *VInv* mRNA and activity could open new ways for incorporating additional levels of control over CIS. Temperatures that promote CIS also promote the up-regulation of *VInv* transcription (Matsuura-Endo et al., 2006)(Zrenner et al., 1996)(Liu et al., 2011). An increase in *VInv* transcript abundance occurs within 7 days when tuber tissue is suddenly exposed to storage temperatures of 4 degrees C (Zrenner et al., 1996). *VInv* transcript amount then gradually decreased over the next several weeks (Wiberley-Bradford et al., 2014)(Zrenner et al., 1996). Interestingly, differences in *VInv* gene expression in long-term storage between wild type and transgenic RNAi chipping potato lines are more pronounced in the first half of storage. This may be attributed to the persistence in VInv activity through the storage period regardless of transcript levels (Wiberley-Bradford et al., 2014).

Invertase activity is also regulated post-translationally by endogenous inhibitors (Pressey, 1969). These small, 17 kDa proteins bind irreversibly to invertase in vivo and form an inactive complex (Rausch and Greiner, 2004). Inhibitors of invertase from tobacco have been overexpressed in transgenic potato and this resulted in short term CIS resistance (Greiner et al., 1999). It has been shown that cultivars that have a higher level of translation of the native inhibitor gene INH2 have both a lower level of reducing sugar accumulation and a suppression of *VInv* expression. (Liu et al., 2010)(Brummell et al., 2011b)(McKenzie et al., 2013). It is not known how CIS resistant, North American chipping germplasm regulates the expression of *VInv*. It would be valuable to understand more about the regulation of this gene under cold storage stress.

2.3 Materials and Methods

Plant Materials; Field and Storage Conditions

The population evaluated in this study was created through a cross between Wauseon and Lenape and consisted of 192 individuals. The population, parents, and check varieties were grown at the Hancock Agricultural Research Station in Hancock, WI on Plainfield sandy loam soil in 2012, 2013, and 2014. The 2012 field year was a randomized complete block and the 2013 and 2014 field years were grown in a modified augmented block design II (Lin and Poushinsky, 1985) with Atlantic as the primary check. Wauseon, Lenape, Snowden, and Megachip served as the secondary checks in this design. Plots were single row entries consisting of 6 plants in 2012 and 10 plants in 2013 and 2014. Management of the plots was based on the growing practices of the central sands potato growing region of Wisconsin. Emergence and stem counts were recorded for all plots in all years. Vine maturity was assessed at 105 DAP on a 1-5 scale (1- Vine 100% senescence, 2- Yellowing, 3- Green vine, 4 – Very Green vine, 5- Very late, full bloom). Vine kill occurred 10 days and 3 days before harvest each year with harvest occurring 120 days after planting. Each plot was dug using a single row potato digger with tubers manually placed into containers to minimize bruising. Four replicates of 7 representative tubers from each genotype or control were placed in mesh bags and stored at 10 ^{0}C for wound-healing, skin set and CIPC sprout inhibitor treatment. Six weeks after harvest, replicate samples from each genotype were split into four equal sets, with one replicate from each genotype, for staggered storage environment treatment. In each of 4 successive days, one set was moved to a 3^{0}C storage locker and kept there for exactly two weeks.

Tissue Collection

After two weeks at 3˚C, the set of samples was removed for tissue sample processing. This involved collecting 7 mm diameter cores of tuber tissue collected from the center of the tuber perpendicular to the longitudinal axis. This core was dissected; the outer periderm and outside tissue was discarded and the center tissue was divided into three equal size core samples approximately 1.5 cm in

length. All three cores were placed into separate 1.5 ml microcentrifuge tubes and flash frozen in liquid nitrogen. Each core was designated for either sugar extraction, RNA extraction, or as a spare. Sugar extraction samples had their mass recorded prior to freezing. After being flash frozen, samples were stored at -80 ^{0}C until further processing.

HPLC Sample Extraction

Sugar sample cores were lyophilized and weighed for dry weight. Each core was placed into a 15 ml conical tube with approximately 10, 3 mm glass beads and shaken for 4 minutes until fully pulverized the tuber tissue. Each powdered sample was suspended in 4 ml of 80% EtOH and placed on a shaking water bath at 60^{0}C for 24 hours. Samples were then centrifuged at 4,000 rpm for 10 minutes to pellet the starch and other solid material and the supernatant was decanted into a labeled, 15 ml conical collection tube. The pellet was then resuspended with 4 ml of 80% EtOH and incubated in a shaking water bath at 60^{0}C for 24 hours. The sample was centrifuged at 4,000 rpm for 10 minutes and the supernatant was decanted into the 15 ml conical collection tube. Each collection tube was brought to 10 ml in volume with 80% EtOH.

From each collection tube, 1ml of liquid was pipetted into a 1.5 ml microcentrifuge tube and vacuum centrifuged until all liquid was removed from the sample. The pellet was resuspended with 1 ml of dH_2O, filtered using a 0.2 µm regenerated cellulose syringe filter, and 10 µL was injected into an HPLC (Shimadzu Scientific Instruments, Columbia, MD). This machine ran in isocratic mode and had a flow rate of 0.44 ml/min using as the mobile phase 0.0041% HPLC grade formic acid pH 3.92±0.02 (Sigma-Aldrich, St. Louis, MO) The samples were separated on a 300 x 8.0 mm Rezex ROA-organic acid column (Phenomenex, Torrance, CA) cooled to 20^{0}C. Reagent grade glucose, fructose, sucrose, citric acid, and malic acid were used as standards, and three sets of calibration standards were run through the system every 17 samples to ensure that the system was running stably. All 6 standards were brought to a concentration of 1.0 mM and placed together in one standard. A 2.0 mM standard was created separately for the three sugars and the two organic acids. The blank standard was dH_2O. Experimental samples were

organized into sequence runs of 60-90 samples. Peaks were detected using the Shimadzu refractive index detector and identified using the EZStart software (Shimadzu). The concentration of each compound in each of the samples was calculated using the compound standards within the sequence. Calculations were performed using R and a custom script for compound quantification.

Reverse Transcription qPCR

Relative expression of genes was performed using data collected from reverse transcription quantitative polymerase chain reactions (RT-qPCR). After tissue collection, samples were stored at -80^{0}C. Tuber cores from a plot were aggregated into two, three-core subsample groups with cores 1, 2, and 3 assigned to subsample 1 and core 4, 5, and 6 assigned to subsample 2. The subsamples were then ground under liquid nitrogen to prevent RNA degradation (model 6770, SPEX SamplePrep, Metuchen, NJ). After grinding, frozen powder was returned to -80 ^{0}C storage. RNA was extracted using the Agilent Plant RNA Isolation Mini Kit (Agilent Technologies, Inc., Santa Clara, CA) according to the manufacturer's instructions. Once extracted, RNA was visually inspected for integrity with gel electrophoresis and tested for contamination and quantity on a NanoDrop 1000 (NanoDrop Products, Wilmington, DE). RNA was DNase-treated using a DNase Treatment Kit (Ambion, Austin, TX) according to the manufacturer's instructions. Samples of 400 ng of total RNA were reverse transcribed using the SuperScript III kit (Invitrogen Corporation, Carlsbad, CA) according to the manufacturer's instructions.

Nuclease free water was added to each cDNA sample to bring the volume to 32 µL. The template for each qPCR assay consisted of 25 ng of total cDNA carried in a volume of 2 µL. For each gene, only a single product was formed and calculated PCR efficiencies were between 90% and 105%. Quantitative PCR was performed using a BioRad iCycler with Maxima SYBR Green/Fluorescein qPCR Master Mix (Fermentas Inc., Glen Burnie, MD). The thermal profile was dependent on the annealing temperature of the gene being amplified (Table 2.1) but consisted of one cycle of 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, the gene specific annealing temperature for 30 s, and 72°C for 30 s. A melt curve was

collected with temperatures from 55 to 95°C. Fluorescence data were collected during the extension phase and the melt curve. There were three technical replicates performed for each quantified gene. The gene primer sequence information along with thermal profile and primer concentration are indicated in Table 2.1.

Data filtering of Ct values was a necessary procedure for data analysis. All threshold decisions were based on Ct and Ct range distribution patterns within a single gene. Maximum Ct thresholds in this study were: 27.5 for *actin,* 29.0 for *EF1-α,* and 27.0 for *60S*. For *VInv*, samples were retained across all reactions and if no data were collected, a Ct value of 35.0 was assigned to the sample. All reactions within a plot subsample and gene were also filtered by their range within their technical replicates. For all genes except actin, single technical replicates were removed if the Ct range of the three samples was greater than 4 and if that sample was the largest deviant from the median value. For actin, the Ct range was 2.0. These filtering steps removed <5% of the samples from the experiment.

Relative expression values were calculated for *VInv* using the cycle threshold (C_T) values of *actin* and *EF1-α* reference genes. This was performed using a modified $\Delta\Delta C_T$ method. In this method the amount of target gene (*VInv*) is calculated using the equation:

$$target\ amount = 2^{-\Delta\Delta C_T}$$

Where $\Delta\Delta C_T$ is calculated with the equation:

$$-\Delta\Delta C_T = -(\Delta C_{T,q} - \Delta C_{T,cb})$$

In this equation $\Delta C_{T,q}$ is the cycle threshold value for the target gene and $\Delta C_{T,cb}$ is the cycle threshold for the reference gene (Livak and Schmittgen, 2001). In the case where two reference genes were used, the geometric mean was calculated and used for this value.

Before the target gene amount was calculated, the mean of a gene's C_T values was calculated among the technical replicates within a subsample. $-\Delta\Delta C_T$ values were calculated within a subsample

using both single reference genes as $\Delta C_{T,cb}$ and their mathematical average. The value for the target amount was then averaged among subsamples and used as the plot mean for data analysis (Schmittgen and Livak, 2008).

Data Analysis

Phenotypic data were evaluated using R statistical software (R Core Team, 2014) and custom scripts. All traits were evaluated for autocorrelation using the corplot package in R. Tuber cores 1, 2, and 3 for HPLC data were aggregated and paired with *VInv* subsample 1 and cores 4, 5, and 6 were aggregated and paired with *VInv* subsample 2. Pearson's correlation was performed on all complete observations between the traits at the subsample level. Linear mixed models were used to estimate the genotypic variance component for each trait across all years by treating genotype as a random effect. This was achieved using the R package lme4 (Bates et al., 2014). Best linear unbiased predictions (BLUP) were generated using a mixed model to get a genotypic performance value for each trait (Robinson, 1991)(Henderson, 1974).

The model used in the mixed model was:

$$y = \mu + G_i + Y_k + (GY)_{ik} + \varepsilon_{ijk}$$

Where μ is the overall mean, G is the effect of the genotype *i*, Y is the effect of year *k*, GY is effect of genotype *i* by year *k* and ε is the residual error. The field plots had incomplete blocks but block was omitted from the model because it did not account for a significant amount of the variance, nor did it affect the residuals in any of the reported traits. The residuals for all models met assumptions of normality and heteroscedasticity.

Broad-sense heritability was estimated for each trait by using the estimated variance components from the lme4 package. The equation was:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{GY}^2}{y} + \frac{\sigma_e^2}{ry}}$$ (Holland et al., 2003)

The broad-sense heritability is denoted by H^2, σ_G^2 is the genotypic variance, σ_{GY}^2 is the genotype by year variance, σ_e^2 is the residual variance, y is the number of years and r is the number of replicates within year.

Linkage Map and QTL Mapping

Theta Score Filtering and Dosage Calling

Theta values from the Illumina genotyping platform are transformed and normalized values based on the reflectance values from alleles at a single marker. Theta values range from 0 (if only probe A is detected) to 1 (if only B is detected) and heterozygous classes have mean theta values that are proportionate with the dosage of the B probe (LaFramboise, 2009). Before any data analysis, theta values were first filtered to eliminate outliers and low quality SNPs. First, the range of theta scores for each marker was analyzed with the expectation of certain theta value ranges depending on marker genotype. Segregating SNPs were expected to range between 0-0.25 or 0.75-1 for simplex SNPs, 0-0.5 or 0.5-1 for duplex or double-simplex SNPs, 0-0.75 or 0.25-1 for simplex-duplex SNPs, and 0-1 for double-duplex SNPs. The smallest expected range for an informative SNP would be approximately 0.25. Theta value ranges were trimmed to a range between 2% and 98% quantiles to eliminate the effect of outliers. Markers were removed if they were monomorphic (range of less than 0.1). SNPs with NA calls were removed prior to further analysis.

The conversion of theta values to marker genotypes followed the procedure developed by Hackett et al (2013) of Biomathematics & Statistics Scotland (BioSS) which is based on expected segregation ratios in a biparental tetraploid family. Marker dosages were called according to a maximum likelihood approach of assigning marker classes based on the possible segregation ratios in a tetraploid cross. The model assumed 13 possible allelic configurations in a biparental cross, which assumes non preferential bivalent formation and no segregation distortion (Hackett et al., 2013), making the identification and elimination of distorted regions important for accurate QTL discovery.

Genotypes were then called using a normal mixture model (Voorrips et al., 2011) that was adapted for using SNP data in Tetraploid Map (Hackett et al., 2013). For each marker, the minimum Bayesian information criterion (BIC) was used to select the most likely configuration (Schwarz, 1978). Markers were eliminated from the dataset if the assigned configuration explained less than 95% of the theta value variance. Genetic structure was analyzed using principal component analysis (PCA) and relatedness of individuals was determined using Nei's genetic distance (Nei, 1972).

Chi-Square Marker Filtering

After dosage calling was complete, a χ^2 test was performed to test for segregation distortion. The expectation classes depended on the parental genotypes due to tetraploid segregation patterns. For this reason, the overall χ^2 test p-value was used to compare markers and how they met expectations in the progeny. Only markers that had a χ^2 p-value > 0.3 were retained for further analysis as described in the TetraploidMap protocol.

Marker Clustering and Ordering

The SNP markers were clustered into twelve linkage groups using the TetraploidMap 2.0 software (Hackett et al., 2013)(Hackett et al., 2014) and a chi-square test for independent segregation (Luo et al., 2001). Recombination and logarithm of the odds (LOD) scores were calculated according to a maximum likelihood procedure between pairs of SNP markers within a linkage group (Hackett et al., 2013)(Luo et al., 2001). Markers were then ordered in each linkage group using the recombination fractions in a weighted least squares procedure implemented in JoinMap 4.0 (Stam, 1993).

Marker Phase Determination

For QTL mapping to proceed, the allelic makeup of each homologous chromosome in each parent was determined. Haplotype reconstruction, or phasing, has been a computational issue for genetic studies in potato populations (Simko, 2004)(Zheng et al., 2016). In this study, phasing was completed manually using the pairwise comparisons between markers and the maximum likelihood of their linkage phase.

Simplex markers had their haplotype reconstructed in a method presented by Hackett et al. (2013). The simplex markers then provided a framework for phasing by providing anchors to parental homologues. Once all the simplex markers were assigned to a parental homologue, the pairwise comparisons to markers of other dosage classes were evaluated and their minor alleles could be deductively assigned to a parental homologue. Stronger emphasis was placed on markers with high LOD difference values, meaning that the presented linkage phase configuration was likely.

After haplotype reconstruction, the quality of the linkage map was assessed by QTL mapping the original theta values. The assumption was that the QTL peak would align with the expected map location of the marker and that most of the variance would be explained. Markers that had theta values which mapped more than 5 cM away or had <75% of the variance explained were dropped. Remaining markers were used in a second iteration or clustering, ordering, phasing, and theta value mapping to ensure the creation of an accurate map.

Interval Mapping on Genotype Probabilities

With complete phase information for each parental homologue available, the most probable genotype at each locus was estimated for all individual progeny in the population. At a single locus in one progeny, there were 36 possible inheritance patterns as dictated by the independent inheritance of chromosomal segments from each parent. The most likely homologue combination was determined using a Hidden Markov Model (Hackett et al., 2013)(Rabiner, 1989).

Each trait then underwent interval QTL mapping which was performed at 1 cM intervals across the genome using the model:

$$y_j = \mu + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5 + a_6X_6 + a_7X_7 + a_8X_8 + \varepsilon_j$$

where y_j is the phenotypic value for the j th individual, μ is the population mean, a_i is the additive contribution of allele i at the locus, X_i is a dummy variable (either 1 or 0) to indicate the presence of

absence of a particular allele in individual j, and ε_j is the residual error. There is an additional constraint that for any individual, $X_1 + X_2 + X_3 + X_4 = 2$ and $X_5 + X_6 + X_7 + X_8 = 2$ because two alleles are inherited from each parent. With that constraint, the variables X_4 and X_8 can be replaced. Therefore, the model can be re-written

$$y_j = \mu + a_1X_1 + a_2X_2 + a_3X_3 + a_5X_5 + a_6X_6 + a_7X_7 + \varepsilon_j$$

where $\mu = \mu_1 + 2(a_4 + a_8), a_1 = a_1 - a_4 \ \ etc.$ This model is referred to as the main effects (additive) genotype probabilities model (Bourke, 2014)(Hackett et al., 2001)

LOD thresholds were determined for each trait by implementing a sequential method of permutations in order to establish a 95% confidence interval (Nettleton and Doerge, 2000). In this study 200 permutation were used. A LOD peak was considered to be significant if it exceeded the upper limit of the confidence interval.

2.4 Results

SNP Marker Filtering

The initial dataset included 8,303 genotyped markers that were filtered through many data quality steps to ensure that an accurate linkage map was constructed (Table 2.2). Removing monomorphic SNPs and markers with a high proportion of missing values resulted in 7,565 markers. The BioSS method for dosage designation eliminated double reduction products, leaving 7,114 markers. The remaining markers were filtered by the amount of variance that a SNP dosage model explained across the population. Removal of markers with less than 95% variance explained left 6,308 markers. After this, filtering for segregation distortion at a threshold $\chi 2$ p-value <0.3 left 2,896 markers. The assigned genome position (Sharma et al., 2013) was then used to eliminate markers that had no assigned physical position, leaving 2,848 markers. Markers were then clustered into linkage groups and ordered by genetic distance. Markers that did not order properly were removed, leaving 2,804. The initial linkage map was developed and theta values were treated as phenotypes, mapped as QTL, and theta values that did not map near their

respective marker or explain greater than 75 percent of the variance were filtered out of the dataset. The remaining markers were clustered and ordered again and this left 2,355 markers in the linkage map.

Linkage Map

The total length of the linkage map in the Wauseon x Lenape population was 1,282 cM long. The marker coverage averaged 1.83 markers per cM with the number of markers for each chromosome occurring in the following distribution: 247 on chromosome I, 165 on chromosome II, 279 on chromosome III, 254 on chromosome IV, 188 on chromosome V, 313 on chromosome VI, 152 on chromosome VII, 175 on chromosome VIII, 126 on chromosome IX, 185 on chromosome X, 159 on chromosome XI, and 112 on chromosome XII (Figure 2.5). Markers in the population could be categorized by the number of dosage groups found across all individuals with 930 occurring in two groups, 704 occurring in three groups, 543 occurring in four groups, and 178 occurring in five dosage groups. The genetic linkage map was in similar relative order as the physical distance reported by the sequenced genome PSGC DM v4.03 with a Pearson's correlation of $R^2 = 0.83$ (Figure 2.1)(Sharma et al., 2013).

Phenotypic Analysis

Sugar and Organic Acids

Sucrose ranged in concentration from 1.34 to 12.17 mg/g FW and had a mean value of 5.11 mg/g FW, glucose ranged from 0.062 to 2.540 mg/g FW and had a mean value of 0.641 mg/g FW, and fructose ranged from 0.0513 to 3.271 mg/g FW and had a mean of 0.838 mg/g FW across all years. Sucrose was normally distributed while fructose and glucose had a right skewed distribution. There was a significant year effect for sugar content in tubers (Figure 2.2). Citric acid was normally distributed and ranged in values from 0.768 to 5.207 mg/g FW with a mean of 1.943 mg/g FW. Malic acid ranged in concentration from 0.2852 to 1.797 mg/g FW and had a mean of 0.851 mg/g FW and was also normally distributed.

The heritabilities for the sugar traits and organic acids were high (Table 2.3). The highest broad sense heritability estimate was for fructose at 0.84. Sucrose heritability was similar at 0.83. Glucose heritability was lower at 0.75 as this compound had a higher proportion of variance explained by year and genotype x year interactions. Citric and malic acid had heritability estimates of 0.66 and 0.70. Citric acid had a large amount of variance attributed to the year component of the model.

Vacuolar Invertase Relative Expression

Relative expression of *VInv* had a range of values and distribution that was dependent on the reference genes used in the $-\Delta\Delta C_T$ method for target gene quantification, as well as the year that the tuber material was grown. *VInv*:(*Actin,EF1-α*) relative expression had an overall range of 0.0004 to 0.9490 with a mean of 0.0960 across all years. *VInv*:(*Actin,EF1-α*) relative expression in 2012 and 2013 had similar distributions; 2012 ranged from 0.0025 to 0.813 with a mean of 0.161 and 2013 ranged from 0.0176 to 0.949 with a mean of 0.161. The 2014 expression distribution was different and ranged from 0.00041 to 0.318 with a mean of 0.028. This pattern occurred in each of the *VInv* relative expression measurement methods. In all, the distribution of expression values in 2012 and 2013 was less skewed than the right skewed distribution of 2014 data (Figure 2.3).

The broad sense heritability for *VInv* relative expression varied depending on the years used in the BLUP model and the reference genes used to calculate the amount of target. In general, heritability was low for multiple year models. The highest heritability was 0.12 for *VInv* expression relative to the mean of *actin* and *EF1-α* in 2013 and 2014. In many cases heritability was below 0.05, such as the overall *VInv* expression relative to the mean of *actin* and *EF1-α* across all years. The variance attributed to genotype x year interactions was very high in low heritability traits, accounting for a comparable amount of variation as the year component *(VInv*:(*Actin,EF1-α*) Full, *VInv*:*Actin* Full). In some cases, variance attributed to genotype by year interactions was much higher than any other component (*VInv*:*EF1-α* 2012,2014) (Table 2.3).

Correlations

There were strong correlations between fructose and glucose concentrations with R^2=0.87. There were also strong, significant correlations between glucose and fructose and all of the sucrose to reducing sugar ratios (all R^2>0.46). Glucose and fructose did not have strong relationships with the glucose to fructose ratio or to the amount of sucrose (Figure 2.4). High positive correlations occurred between *VInv*:(*Actin,EF1-α*), *VInv*:*Actin*, and *VInv*:(*Actin,EF1-α,60S*) with all R^2>0.58. This correlation was lower for the *60S* and *EF1-α* single gene standardized *VInv* expression. Citric acid had the largest correlation with *VInv* expression traits with R^2 ranging from -0.15 to -0.23. Expression trait correlations with sugar related traits were all < 0.15 or not significant.

QTL Analysis

The QTL detected by the additive dosage genotype probabilities model are presented in Table 2.4. In this study, 33 QTL were detected for traits related to sugars, organic acids, and *VInv* expression. The mean 1.5 LOD support interval was 22 cM with a maximum LOD interval of 37 cM and a minimum of 8 cM. Significance of a QTL peak was determined by the sequential permutation test and was different for each trait in this study. However, there was minimal range of LOD threshold values between traits in this study (Table 2.4). The mean LOD threshold was 3.56 with a mean 95% confidence interval of (3.32-3.98). The detected QTL had moderate effect sizes ranging from 5.49 to 14.45 % variance explained. The largest effect QTL was for fructose and it had a LOD score of 8.37. The smallest effect QTL was for *VInv*:*Actin* Full and had LOD score of 4.19.

QTL for Sugars and Organic Acids

Four QTL were detected for the sugars of sucrose, glucose and fructose across the genome. Fructose had one detected QTL (Chr12, 86 cM) with a LOD interval of 19 cM and was the largest effect QTL in this study. Glucose had one detected QTL (Chr12, 84 cM) in the same location with similar LOD intervals. Sucrose had two QTL (Chr02, 81 cM; Chr05, 38 cM) accounting for 7.18 and 7.23 % variance

and had LOD intervals of 22 and 28 cM. Five QTL were detected for the organic acids. Citric acid had one detected QTL (Chr02, 13 cM), LOD interval of 19 cM, and accounted for 6.21% of variance. Malic acid had four detected QTL (Chr02, 14 cM; Chr03, 77 cM; Chr05, 18 cM; Chr06, 26 cM) with effect sizes between 6.46 and 7.41 percent variance explained. The QTL on chromosome VI had a LOD interval of 8 cM while the LOD intervals were 13 cM on chromosome II, 30 cM on chromosome III, and 24 cM on chromosome V.

Fifteen QTL were detected for the sugar relationship traits. There were five QTL for the ratio of glucose to fructose (Chr01, 46 cM; Chr05, 38 cM; Chr07, 55 cM; Chr11, 12 cM; Chr12, 84 cM). The QTL on chromosome XII had the largest effect for sugar ratios and accounted for 13.09 % variance. The remaining QTL ranged in variance between 5.57-8.36 and all QTL ranged in LOD interval size from 16 – 23 cM (Table 2.4). Four QTL were detected for sucrose to fructose ratio (Chr01, 46 cM; Chr03, 76 cM; Chr11, 44 cM; Chr12, 84 cM) and had LOD intervals of 20, 27, 32, and 17 cM respectively. The largest effect size was for the QTL on chromosome XII with 11.15% variance explained. Two QTL were detected for the sucrose to glucose ratio (Chr03, 76 cM; Chr12, 84 cM) with respective LOD intervals of 32 and 18 cM and 6.92 and 11.1 variance explained. Four QTL were detected for the ratio between sucrose and total reducing sugars (Chr01, 46 cM; Chr03, 76 cM; Chr11, 43 cM; Chr12, 84 cM). The largest effect QTL accounted for 12.1% variance and had a LOD peak of 7.25. The remaining QTL had effect sizes ranging from 5.65-6.85 percent variance explained.

QTL for Vacuolar Invertase Relative Expression

Across all reference gene and year combinations for *VInv* relative expression traits, there were 9 detected QTL. Effect size of QTL ranged from 5.49 – 9.85 percent variance explained and LOD intervals ranged in size from 14-34 cM. There were five QTL detected for *VInv* expression relative to the mean of *actin* and *EF1-α*; two QTL of trait data from 2012 and 2014, two QTL of trait data from 2012, and one QTL of trait data from 2014. The QTL detected from the 2012,2014 data (Chr03, 98 cM; Chr07, 34 cM) had effect sizes of 7.02 and 7.86 respectively. The expression QTL from 2012 (Chr03, 85 cM; Chr07, 34

cM) had similar LOD profiles and slightly higher effect sizes at 8.12 and 8.77 percent variance explained. The QTL from the trait measured in 2014 (Chr03, 57 cM) had an effect size of 5.98 percent variance explained. There were four QTL detected using a single reference gene. One QTL was detected using all years of data with *actin* as the reference (Chr04, 13 cM) and had a small effect size of 5.49 percent variance explained. *Actin* was also used in the QTL detected for trait data collected in 2012 (Chr03, 85 cM; Chr07, 34 cM) and had effect sizes of 8.12 and 8.77 percent variance explained. There was one QTL detected for trait data collected in 2014 with *60S* as the reference gene (Chr03, 64 cM). This QTL explained the largest amount of variance (9.85 %) and had a LOD interval of 14 cM.

2.5 Discussion

The range of sugar concentrations in this population under conditions that induce cold sweetening indicate a genetic component to their ability to tolerate this stress. This ability has been noted in the parent of this population, Lenape, and has been used to confer CIS resistance to its progeny (Love et al., 1998)(Van Berloo et al., 2007). In this study, there were many progeny that had sucrose concentrations below 3 mg/g FW and 14 that had glucose and fructose concentrations below 0.3 mg/g FW reliably. This population and the evaluation of its sugars under a strong cold storage signal identified some QTL with sizable effect on variation in sugar accumulation. The results related to *Vinv* expression were less clear and indicate the complexity of how this gene is regulated within the tuber.

QTL Related to Sugars

Chromosome XII had multiple large effect QTL that co-located and all of these QTL involved glucose, fructose, or ratios of sugars. The QTL for fructose, glucose, and the ratios of sucrose:glucose, sucrose:fructose, glucose:fructose, and sucrose:total reducing sugars all map to chromosome XII at 84 cM. The closest marker to this position is c1_2724 and is within 200 kb of two phosphofructokinase genes and is also 3Mb away from an annotated Fructose-1,6-bisphosphatase gene (Sharma et al., 2013). This maps to the same position on chromosome XII as QTL for the slope in chip color through storage in chapter 4 and chip color in late storage (Rak, 2015). Single marker analysis reveals a significant

relationship between dosage of the marker and the amount of fructose and glucose in an additive manner. Sucrose levels increase relative to the fructose and glucose levels (Figure 2.6). The allele dosage linked to the marker appears to behave as a limiting step for the hydrolysis of sucrose and is likely linked to acid invertase activity, a regulator in the ratio of hexoses to sucrose (Zrenner et al., 1996). The closest invertase-related gene is an invertase inhibitor 2.5 Mb away from marker c1_2724. In Figure 2.6, there is not a clear relationship between a marker allele donation from Lenape and lower reducing sugars. However, this does occur for the sucrose to fructose ratio QTL on chromosome XI. However, this pattern is not conveniently elucidated using the mapping software used in this study. This is partly due to the limitations imposed by SNP marker technology where bi allelic markers can mask the true number of alleles and due to constraints imposed by experimental size. To fully determine parental contribution, haplotype block analysis should be conducted.

Other QTL for sugars have co-localized near some areas of the genome with potential candidate genes. Chromosome I had QTL detected at 86 cM for sucrose:total reducing sugars, sucrose:fructose, and glucose:fructose. The estimated position was close to the marker c2_20803 and is 1.6 Mb from 1,4-alpha-glucan-maltohydrolase genes, genes previously associated with total sugar yield (Schreiber et al., 2014)(Sharma et al., 2013) This gene has been shown to be upregulated in transcription under cold stress in cultivated potato (Oufir et al., 2008). The QTL detected on chromosome III for sucrose:total reducing sugars, sucrose:glucose, and sucrose:fructose all map to 84 cM near marker c2_29513. This is 3.7 Mb away from a several putative acid invertase genes not reported in Schreiber et al. (2014) as being associated with chipping quality.

A QTL for sucrose concentration maps to the same location as glucose:fructose ratio on chromosome V at 38 cM. The closest marker is c2_47607, which is 60 kb from an alpha-amylase gene that has been associated with chip quality (Schreiber et al., 2014). Alpha-amylase activity is known to increase during cold storage conditions (Cochrane et al., 1991) and would result in increased sugar content in the cell (Sowokinos, 2001b). Chromosome XI contains two QTL, one for sucrose:total

reducing sugars and one for sucrose:fructose that map close to each other at 43 and 44 cM respectively. The two closest markers, c2_53687 and c2_24318 span a region that is 300 kb. The marker c2_24318 is adjacent to three sucrose transport proteins, which were identified by Schreiber et al. (2014) as a candidate gene but were not found in that case to have an association with sugar accumulation. This region of the genome was identified previously in a diploid mapping population as being located close to the sucrose transporter 1 (*Sut1*) gene (Menendez et al., 2002). This difference may be due to differences in genetic background or the effects of different alleles.

The glucose:fructose ratio QTL maps to chromosome XI at 12 cM. The closest marker is c2_33661 and is located at a gene of unknown function. This gene mapped to the same location as L color in January in Chapter 4. QTL for sucrose maps to a position at 81 cM on chromosome II. The closest marker is c1_8490 which is directly annotated as a stress regulated protein and is 1.4 Mb from several Granule-bound starch synthase 2 (*GBSS2*) genes. Reductions in *GBSS2* gene expression in transgenic tubers resulted in lower amylose content (Kuipers et al., 1994) and expression decreases as storage temperature declines (Wiberley-Bradford et al., 2014). This gene has been associated with sugar accumulation in a diploid mapping population (Menendez et al., 2002).

QTL for Organic Acids

Citric acid and malic acid both had QTL that mapped to 13 and 14 cM on chromosome II. The nearest markers were c2_49068 and c2_16362 and they spanned a region of 9 Mb. Citric and malic acid accumulate in the vacuole in tuber tissue (Farre et al., 2014). These organic acids have been shown to be important in the decrease in tuber pH after movement to cold storage conditions (Hyde and Morrison, 1964). In storage, citric acid increases while malic decreases in cold storage over time (Sweeney et al., 1969)(Lisinska and Aniolowski, 1990) but in this short storage duration study, there was a positive correlation between the two compounds. In growing tubers under stress, organic acids are known to have a role in osmotic potential regulation (Bethke et al., 2009). The marker c2_16362 is adjacent to Potassium transporter 11, a potassium ion transmembrane transporter. This might be a valuable region in

understanding organic acid accumulation under cold stress. Malic acid had QTL mapped to chromosome V at 18 cM and chromosome VI at 26 cM. The QTL on chromosome V is located closest to marker c2_11604 and the QTL on chromosome VI is nearest to marker c2_32938. Both markers were located in or near genes of unknown function. Examination of these regions might provide answers to the role that citric and malic acids have in the CIS response pathway.

VInv relative expression and QTL

VInv expression levels were comparable to wild type levels observed by Wiberley-Bradford et al. (2014) during the early storage season. However, the heritability of trait severity in this population is not very high, largely because there was high genotypic x year variation. Interestingly, there were no strong correlations between *VInv* expression and sugar concentration at the time point sampled. It is well documented that *VInv* activity is associated with increased reducing sugars (McKenzie et al., 2005)(Bhaskar et al. 2010)(Liu et al., 2011). The relationship between relative expression and activity, however, is difficult to quantify due to mechanisms like post translational modification (McKenzie et al., 2013), interactions with inhibitor genes that limits activity (Brummell et al., 2011a), and allelic variants that affect the *VInv* protein at a functional level (Draffehn et al., 2012). The potential for multiple levels of cellular regulation for this gene offers an explanation for the high level of genotype by year interaction.

Most of the QTL detected for *VInv* expression were identified using traits measured in a single or two year period and no QTL mapped to the vacuolar acid invertase position on chromosome III. QTL for *VInv*:(*Actin, EFI-α*) in 2012 and *VInv:Actin* in 2012 both mapped to 85 cM on chromosome III. The closest marker to this position was c2_1717 which is annotated within a cluster of nodulation receptor kinases and within 400 kb of miRNA/protein/siRNA binding proteins and translation initiation factors (Sharma et al., 2013). QTL for *VInv*:(*Actin, EFI-α*) in 2012 and 2014 had the same LOD intervals but mapped to a position at 98 cM on chromosome III. Its closest marker is c1-111 but the QTL maps to a location between this marker and marker c2_26454. In the center of the 15 annotated genes in this 800 kb region are two signal recognition particle 54 kDa proteins that could potentially play a role in targeting

invertase protein to the ER (Sharma et al., 2013). There were 2 QTL with overlapping LOD intervals on another section of chromosome III. QTL for *VInv*:(*Actin, EFI-α*) in 2014 mapped to 57 cM and the closest marker was c1_6355. QTL for *VInv*:*60S* in 2014 mapped to 64 cM and the closest marker was c1_13842. These markers mapped to within 2 Mb of each other and are within 700 kb of Kunitz-type tuber invertase inhibitors and c1_13843 is close to a gene annotated as an ubiquitin-protein ligase (Sharma et al., 2013).

There were three QTL for invertase expression that co-located on chromosome VII at position 34 cM. The marker that maps closest to the QTL at this position is c2_46755 and is annotated as a 60S ribosomal protein and is within 1 Mb of 3 genes of unknown function. QTL for *VInv*:*Actin* measured under all years mapped to chromosome IV on a position at 13 cM and mapped to a broad region without many markers. The nearest markers, c2_26792 and c2_47320 spanned a region of 1.3 Mb with the most common annotations in that region being: Leucine-rich repeat containing protein, Conserved gene of unknown function, NBS-LRR disease resistance protein, NAD-dependent epimerase/dehydratase, Phosphatidylinositol-4-phosphate 5-kinase, and Glycosyl transferase, family 8. While there are no clearly understood roles for the genes in the regions of these QTL, it is an initial step towards understanding the regulation of this important class of protein under stress mediated response (Ruan et al., 2010).

2.6 Conclusion

The high quality linkage map constructed using this population has provided a framework for identifying genomic regions important for CIS resistance. Many of the QTL for sugar traits mapped close to loci previously discovered in work with diploid potato and this extends that work to cultivated autotetraploid potato. The population in this study is a good representative of the chip processing market segment and allows for an understanding of the beneficial alleles contained in Lenape. Large effect QTL on chromosome XII for fructose and glucose concentration are significantly linked to a single marker and this region should be investigated further to determine the candidate gene. Vacuolar Acid invertase expression had some small effect QTL that explained expression levels but much of the variation in this trait was accounted for by other factors.

2.7 References

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2014. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67(1): 51.

Van Berloo, R., R.C.B. Hutten, H.J. Van Eck, and R.G.F. Visser. 2007. An online potato pedigree database resource. Potato Res. 50(1): 45–57.

Bethke, P.C., R.P. Sabba, and A.J. Bussan. 2009. Tuber water and pressure potentials decrease and sucrose contents increase in response to moderate drought and heat stress. Am. J. Potato Res. 86(6): 519–532.

Bhaskar, P.B., L. Wu, J.S. Busse, B.R. Whitty, A.J. Hamernik, S.H. Jansky, C.R. Buell, P.C. Bethke, and J. Jiang. 2010a. Suppression of the vacuolar invertase gene prevents cold-induced sweetening in potato. Plant Physiol. 154(2): 939–948.

Blenkinsop, R.W., R.Y. Yada, and A.G. Marangoni. 2004. Metabolic control of low-temperature sweetening in potato tubers during postharvest storage. Hortic. Rev. (Am. Soc. Hortic. Sci). 30: 317–354.

Bourke, P. 2014. QTL analysis in polyploids: Model testing and power calculations. Wageningen, The Netherlands. Wageningen University.

Brown, J., G.R. Mackay, H. Bain, D.W. Griffith, and M.J. Allison. 1990a. The processing potential of tubers of the cultivated potato, *Solanum tuberosum L.*, after storage at low temperatures. 2. Sugar concentration. Potato Res. 33: 211–218.

Brown, J., G.R. Mackay, H. Bain, D.W. Griffith, and M.J. Allison. 1990b. The processing potential of tubers of the cultivated potato, *Solanum tuberosum L.*, after storage at low temperatures. 2. Sugar concentration. Potato Res. 33: 219–227.

Brummell, D.A., R.K.Y. Chen, J.C. Harris, H. Zhang, C. Hamiaux, A. V Kralicek, and M.J. Mckenzie. 2011a. Induction of vacuolar invertase inhibitor mRNA in potato tubers contributes to cold-induced sweetening resistance and includes spliced hybrid mRNA variants. 62(10): 3519–3534.

Brummell, D.A., R.K.Y. Chen, J.C. Harris, H. Zhang, C. Hamiaux, A. V. Kralicek, and M.J. McKenzie. 2011b. Induction of vacuolar invertase inhibitor mRNA in potato tubers contributes to cold-induced sweetening resistance and includes spliced hybrid mRNA variants. J. Exp. Bot. 62(10): 3519–3534.

Burton, W.G. 1969. The sugar balance in some British potato varieties during storage. II. The effects of tuber age, previous storage temperature, and intermittent refrigeration upon low-temperature sweetening. Eur. Potato J. 12(2): 81–95.

Chen, X., F. Salamini, and C. Gebhardt. 2001. A potato molecular-function map for carbohydrate metabolism and transport. Theor. Appl. Genet. 102(2–3): 284–295.

Cochrane, M.P., C.M. Duffus, M.J. Allison, and G.R. Mackay. 1991. Amylolytic activity in stored potato tubers. 1. Estimation using p-nitrophenyloligosaccharides. Potato Res. 34(3): 325–332.

Coffin, R.H., R.Y. Yada, K.L. Parkin, B. Grodzinski, and D.W. Stanley. 1987. Effect of low temperature storage on sugar concentrations and chip color of certain processing potato cultivars and selections. J. Food Sci. 52(3): 639–645.

Draffehn, A.M., P. Durek, A. Nunes-Nesi, B. Stich, A.R. Fernie, and C. Gebhardt. 2012. Tapping natural variation at functional level reveals allele specific molecular characteristics of potato invertase Pain-1. Plant, Cell Environ. 35(12): 2143–2154.

Farré EM, Tiessen A, Roessner U, Geigenberger P, Trethewey RN, Willmitzer L. 2001. Analysis of the compartmentation of glycolytic intermediates, nucleotides, sugars, organic acids, amino acids, and sugar alcohols in potato tubers using a nonaqueous fractionation method. Plant Physiol. 127:685–700.

Greiner, S., T. Rausch, U. Sonnewald, and K. Herbers. 1999. Ectopic expression of a tobacco invertase inhibitor homolog prevents cold-induced sweetening of potato tubers. Nat. Biotechnol. 17(7): 708–711.

Hackett, C., J.E. Bradshaw, and G.J. Bryan. 2014. QTL mapping in autotetraploids using SNP dosage information. Theor. Appl. Genet. 127(9): 1885–1904.

Hackett, C., J.E. Bradshaw, and J.W. McNicol. 2001. Interval mapping of quantitative trait loci in autotetraploid species. Genetics 159(4): 1819–1832.

Hackett, C., K. McLean, and G.J. Bryan. 2013. Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. PLoS One 8(5): 1-21.

Henderson, C.R. 1974. General flexibility of linear model techniques for sire evaluation. J. Dairy Sci. 57(8): 963–972.

Hill, L.M., R. Reimholz, R. Schröder, T.H. Nielsen, and M. Stitt. 1996. The onset of sucrose accumulation in cold-stored potato tubers is caused by an increased rate of sucrose synthesis and coincides with low levels of hexose-phosphates, an activation of sucrose phosphate synthase and the appearance of a new form of amylase. Plant Cell Env. 19(11): 1223–1237.

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant bBreeding: An update. Plant Breed. Rev. 22: 9–112.

Hyde, R.B., and J.W. Morrison. 1964. The effect of storage temperature on reducing sugars, pH, and phosphorylase enzyme activity in potato tubers. Am. Potato J. 41(6): 163–168.

Isherwood, F.A. 1973. Starch-sugar interconversion in *Solanum tuberosum*. Phytochemistry 12(1965): 2579–2591.

Isla, M.I., D.P. Leal, M.A. Vattuone, and A.R. Sampietro. 1992. Cellular localization of the invertase, proteinaceous inhibitor and lectin from potato tubers. Phytochemistry 31(4): 1115–1118.

Kuipers, A., E. Jacobsen, and R. Visser. 1994. Formation and deposition of amylose in the potato tuber starch granule are affected by the reduction of granule-bound starch synthase gene expression. Plant Cell 6(1): 43–52.

Kumar, D., B.P. Singh, and P. Kumar. 2004. An overview of the factors affecting sugar content of potatoes. Ann. Appl. Biol. 145(3): 247–256.

LaFramboise, T. 2009. Single nucleotide polymorphism arrays: A decade of biological, computational and technological advances. Nucleic Acids Res. 37(13): 4181–4193.

Lisinska, G., and K. Aniolowski. 1990. Organic Acids in Potato Tubers : Part 1 The effect of storage temperatures and time on citric and malic acid contents of potato tubers. Food Chem. 38: 255–261.

Liu, X., B. Song, H. Zhang, X.Q. Li, C. Xie, and J. Liu. 2010. Cloning and molecular characterization of putative invertase inhibitor genes and their possible contributions to cold-induced sweetening of potato tubers. Mol. Genet. Genomics 284(3): 147–159.

Liu, X., C. Zhang, Y. Ou, Y. Lin, B. Song, C. Xie, J. Liu, and X.-Q. Li. 2011. Systematic analysis of potato acid invertase genes reveals that a cold-responsive member, StvacINV1, regulates cold-

induced sweetening of tubers. Mol. Genet. Genomics 286(2): 109–18.

Livak, K.J., and T.D. Schmittgen. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25(4): 402–408.

Lorberth, R., G. Ritte, L. Willmitzer, and J. Kossmann. 1998. Inhibition of a starch-granule–Bound protein leads to modified starch and repression of cold sweetening. Nat. Biotechnol. 16(May): 473–477.

Love, S., J. Pavek, A. Thompson-Johns, and W. Bohl. 1998. Breeding progress for potato chip quality in North American cultivars. Am. J. Potato Res. 75: 27–36.

Luo, Z.W., C. Hackett, J.E. Bradshaw, J.W. McNicol, and D. Milbourne. 2001. Construction of a genetic linkage map in tetraploid species using molecular markers. Genetics 157(3): 1369–1385.

Matsuura-Endo, C., A. Ohara-Takada, Y. Chuda, H. Ono, H. Yada, M. Yoshida, A. Kobayashi, S. Tsuda, S. Takigawa, T. Noda, H. Yamauchi, and M. Mori. 2006. Effects of storage temperature on the contents of sugars and free amino acids in tubers from different potato cultivars and acrylamide in chips. Biosci. Biotechnol. Biochem. 70(5): 1173–1180.

McKenzie, M.J., R.K.Y. Chen, J.C. Harris, M.J. Ashworth, and D.A. Brummell. 2013. Post-translational regulation of acid invertase activity by vacuolar invertase inhibitor affects resistance to cold-induced sweetening of potato tubers. Plant, Cell Environ. 36(1): 176–185.

McKenzie, M.J., J.R. Sowokinos, I.M. Shea, S.K. Gupta, R.R. Lindlauf, and J.A.D. Anderson. 2005. Investigations on the role of acid invertase and UDP-glucose pyrophosphorylase in potato clones with varying resistance to cold-induced sweetening. Amer J Potato Res 82: 231–239.

Menendez, C.M., E. Ritter, R. Schafer-Pregl, B. Walkeneier, A. Kalde, F. Salamini, and C. Gebhardt. 2002. Cold-sweetening in diploid potato: Mapping QTL and candidate genes. Genetics 162(November): 1423–1434.

Müller-Thurgau, H. 1882. Ueber zuckeranhäufung in pflanzentheilen in folge niederer temperatu. Landwirtsch Jahrb 11: 751–828.

Nei, M. 1972. Genetic distance between populations - Part 3: Wahlund's principle as related to genetic distance and an application. Am. Nat. 106(949): 283–292.

Nettleton, D., and R.W. Doerge. 2000. Accounting for variability in the use of permutation testing to detect quantitative trait loci. Biometrics 56(1): 52–58.

Nielsen, T.H., U. Deiting, and M. Stitt. 1997. A beta-amylase in potato tubers is induced by storage at low temperature. Plant Physiol. 113(2): 503–510.

Oufir, M., S. Legay, N. Nicot, K. Van Moer, L. Hoffmann, J. Renaut, J.F. Hausman, and D. Evers. 2008. Gene expression in potato during cold exposure: Changes in carbohydrate and polyamine metabolisms. Plant Sci. 175(6): 839–852.

Pressey, R. 1969. Role of invertase in the accumulation of sugars in cold-stored potatoes. Am. Potato J. 46(8): 291–297.

Pritchard, M.K., and L.R. Adam. 1994. Relationships between fry color and sugar concentrations in stored Russet Burbank and Shepody potatoes. Am. J. Potato Res. 71(1): 59–68.

Rabiner, L.R. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE 77(2): 257–286.

R Core Team. 2014. R: A language and environment for statistical computing. R Found. Stat. Comput. Vienna, Austria (January): 1–3. Available at https://www.r-project.org/.

Rak, K. 2015. Breeding and molecular genetics for the improvement of cold storage potato chip quality. Madison, WI. University of Wisconsin-Madison.

Rausch, T., and S. Greiner. 2004. Plant protein inhibitors of invertases. Biochim. Biophys. Acta - Proteins Proteomics 1696(2): 253–261.

Richardson, D.L., H. V Davies, H.A. Ross, and G.R. Mackay. 1990. Invertase activity and its relation to hexose accumulation in potato tubers. J. Exp. Bot. 41(222): 95–99.

Robinson, G.K. 1991. That BLUP is a good thing: The estimation of random effects. Stat. Sci. 6: 15–32.

Ruan, Y.-L., Y. Jin, Y.-J. Yang, G.-J. Li, and J.S. Boyer. 2010. Sugar input, metabolism, and signaling mediated by invertase: roles in development, yield potential, and response to drought and heat. Mol. Plant 3(6): 942–955.

Schmittgen, T.D., and K.J. Livak. 2008. Analyzing real-time PCR data by the comparative CT method. Nat. Protoc. 3(6): 1101–1108.

Schreiber, L., a. C. Nader-Nieto, E.M. Schonhals, B. Walkemeier, and C. Gebhardt. 2014. SNPs in genes functional in starch-sugar interconversion associate with natural variation of tuber starch and sugar content of p otato (*Solanum tuberosum L.*). G3 Genes|Genomes|Genetics 4(10): 1797–1811.

Sharma, S.K., D. Bolser, J. de Boer, M. Sønderkær, W. Amoros, M.F. Carboni, J.M. D'Ambrosio, G. de la Cruz, A. Di Genova, D.S. Douches, M. Eguiluz, X. Guo, F. Guzman, C. Hackett, J.P. Hamilton, G. Li, Y. Li, R. Lozano, A. Maass, D. Marshall, D. Martinez, K. McLean, N. Mejía, L. Milne, S. Munive, I. Nagy, O. Ponce, M. Ramirez, R. Simon, S.J. Thomson, Y. Torres, R. Waugh, Z. Zhang, S. Huang, R.G.F. Visser, C.W.B. Bachem, B. Sagredo, S.E. Feingold, G. Orjeda, R.E. Veilleux, M. Bonierbale, J.M.E. Jacobs, D. Milbourne, D.M.A. Martin, and G.J. Bryan. 2013. Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps. G3 (Bethesda). 3(11): 2031–2047.

Simko, I. 2004. One potato, two potato: Haplotype association mapping in autotetraploids. Trends Plant Sci. 9(9): 441–448.

Sowokinos, J.R. 2001a. Allele and isozyme patterns of UDP-glucose pyrophosphorylase as a marker for cold-sweetening resistance in potatoes. Am. J. Potato Res. 78: 57–64.

Sowokinos, J.R. 2001b. Biochemical and molecular control of cold-induced sweetening in potatoes. Am. J. Potato Res. 78(3): 221–236.

Stam, P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: Join Map. Plant J. 3(1 993): 739–744.

Sweeney, J., P. Hepner, and S. Libeck. 1969. Organic acid, amino acid, and ascorbic acid content of potatoes as affected by storage conditions. Am. Potato J.: 463–469.

Voorrips, R.E., G. Gort, and B. Vosman. 2011. Genotype calling in tetraploid species from bi-allelic marker data using mixture models. BMC Bioinformatics 12(1): 172.

Watts, J. 1968. Lenape: A new potato variety high in solids and chipping quality. Am. Potato J. 45: 142–145.

Wiberley-Bradford, A.E., J.S. Busse, J. Jiang, and P.C. Bethke. 2014. Sugar metabolism, chip color, invertase activity, and gene expression during long-term cold storage of potato (*Solanum tuberosum*)

tubers from wild-type and vacuolar invertase silencing lines of Katahdin. BMC Res. Notes 7(1): 801-811.

Zheng, C., R.E. Voorrips, J. Jansen, C.A. Hackett, J. Ho, and M.C.A.M. Bink. 2016. Probabilistic multilocus haplotype reconstruction in outcrossing tetraploids. Genetics 203(1): 119–131.

Zrenner, R., K. Schüler, and U. Sonnewald. 1996. Soluble acid invertase determines the hexose-to-sucrose ratio in cold-stored potato tubers. Planta 198: 246–52.

Zrenner, R., L. Willmitzer, and U. Sonnewald. 1993. Analysis of the expression of potato uridinediphosphate-glucose pyrophosphorylase and its inhibition by antisense RNA. Planta 190(2): 247–252.

Table 2.1. Primers used for quantitative PCR with concentration. All sequences are given 5'-3', concentrations are µM, and accessions are from the Potato Genome Sequencing Consortium.

Gene	Forward Primer (Concentration)	Reverse Primer (Concentration)	T_{anneal} ^0C
Actin (PGSC0003DMG400027746)	ATGTTCCCGGGTATTGCTGACAGA (0.4)	CTGCCTTTGCAATCCACATCTGCT (0.4)	55
Elongation Factor 1α (ef1-α) (PGSC0003DMG400023270)	TTCCACTTCAGGATGTTTACAAGA (0.8)	CAGCAACATTCTTAACATTGAACC (0.4)	53
60S ribosomal protein (PGSC0003DMG400015795)	GCAAAGAAGAAGAGAGAGGAGATG (0.4)	TTCAAAGCCATAATTGTGTCAAGT (0.4)	57
Vacuolar Acid Intervase (Vinv) (PGSC0003DMG400013856)	AAACGGGTTGGACACATCAT (0.2)	AACCCAATTCCACAATCCAA (0.2)	55

Table 2.2. Filtering steps of SNP markers and genotypes in the construction of the Wauseon x Lenape linkage map developed using TetraploidMap 2.0

Mapping Step	Markers Retained	Genotypes Retained
Start with raw theta values	8303	192
Filter monomorphic markers	7578	
Filter SNPs with NA calls	7565	
Evaluate population structure and remove unrelated individuals		191
Call tetra-allelic dosage and eliminate double reduction marker products	7114	
Eliminate SNPs with a dosage model <95% of variance explained	6308	
Test for segregation distortion χ^2 p-value <0.3	2896	
Eliminate SNPSs with no assigned genome physical position	2848	
Cluster and order markers, eliminate problematic markers	2804	
Run QTL on raw Theta values, manually examine markers with <75% variance explained and a QTL peak >10cM away from genetic position, drop problematic SNPs	2355	
Final Map for QTL analysis	2355	191

Table 2.3. Variance Components for sugar traits, organic acids, and *VInv* relative expression in a Wauseon x Lenape population. Relative expression was calculated using different years and combinations of reference genes. σ_g^2, σ_y^2, σ_{gy}^2, σ_ε^2, and h^2 for genotypes, years, genotypes x years, residual variance, and broad sense heritability on an entry mean basis

Trait	Abbreviation	σ_g^2	σ_y^2	σ_{gy}^2	σ_ε^2	h^2
Oxalic Acid	oa	0.01275	0.00012	0.00000	0.02804	0.58
Citric Acid	ca	0.08463	0.17410	0.06627	0.06235	0.66
Sucrose	suc	1.947	0.1187	0.405	0.8124	0.83
Malic Acid	ma	0.01855	0.00509	0.01091	0.01288	0.70
Glucose	glu	0.09200	0.04853	0.08377	0.00605	0.75
Fructose	fru	0.20090	0.04304	0.10040	0.01310	0.84
Dry Matter	dm	0.00020	0.00004	0.00013	0.00029	0.59
Glucose:Fructose	glfrr	0.01523	0.02926	0.01471	0.00979	0.65
Sucrose:Total Reducing Sugars	sutrsr	20.06	2.79	7.25	9.94	0.78
Sucrose:Fructose	sufrr	98.69	7.71	30.69	43.10	0.80
Sucrose:Glucose	suglur	61.44	28.66	61.66	57.15	0.61
VInv :(*Actin,EF1-α*) Full	vfmae	0.00009	0.00497	0.00407	0.00596	0.03
VInv:Actin Full	vfa	0.00026	0.00378	0.00307	0.00798	0.07
VInv :(*Actin,EF1-α*) 2012,2014	v24mae	0.00000	0.00511	0.00191	0.00479	0.00
VInv :*Actin* 2012,2014	v24a	0.00000	0.00363	0.00215	0.00620	0.00
VInv :*EF1-α* 2012,2014	v24e	0.00004	0.00443	0.00888	0.00090	0.01
VInv :(*Actin,EF1-α*) 2012,2013	v23mae	0.00013	0.00016	0.01315	0.00535	0.02
VInv :*Actin* 2012,2013	v23a	0.00013	0.00016	0.01315	0.00535	0.02
VInv :(*Actin,EF1-α*) 2013,2014	v34mae	0.00048	0.00929	0.00401	0.00602	0.13
VInv :*EF1-α* 2013,2014	v34e	0.00000	0.00024	0.00004	0.00102	0.00

Table 2.4. Estimated QTL locations and effect sizes for BLUPs of sugars, sugar ratios, organic acids, and relative expression of *VInv* under different years and reference genes from genotype probabilities complete additive model in Wauseon x Lenape population

Trait code[a]	Chrom, Position (cM) and 1.5 LOD support interval	LOD	LOD Threshold 95% C.I.[b]	R^2	Closest Marker and Physical Position[c]	Flanking Markers and Physical Position[d,c]
ca	Chr02 13 (0-19)	4.35	(3.38-4.02)	6.21	c2_49068 (17379552)	c2_4372,c1_15971 (4194794,18501616)
fru	Chr12 84 (82-101)	8.37	(3.44-4.12)	14.45	c1_2724 (57460060)	c1_2718,c1_2690 (57349369,58047558)
glfrr	Chr01 46 (36-53)	4.68	(3.31-3.94)	6.50	c2_20803 (60987055)	c2_35518,c2_32112 (60512580,42197568)
glfrr	Chr05 38 (31-47)	4.32	(3.31-3.94)	5.77	c2_47607 (5972364)	c1_14802,c2_55905 (5051766,9640118)
glfrr	Chr07 55 (48-71)	4.25	(3.31-3.94)	5.57	c1_7520 (41720062)	c2_9326,c2_28174 (41265653,51518820)
glfrr	Chr11 12 (1-24)	5.16	(3.31-3.94)	8.36	c2_33661 (2262006)	c1_16456,c2_13350 (4184287,437208)
glfrr	Chr12 84 (81-102)	7.55	(3.31-3.94)	13.09	c1_2724 (57460060)	c1_11657,c1_2690 (54319182,58047558)
glu	Chr12 84 (82-101)	6.45	(3.44-4.2)	10.63	c1_2724 (57460060)	c1_2718,c1_2690 (57349369,58047558)
ma	Chr02 14 (6-19)	4.71	(3.36-3.96)	6.90	c2_16362 (15157259)	c2_48734,c1_15971 (5552574,18501616)
ma	Chr03 77 (55-85)	4.34	(3.36-3.96)	6.57	c1_6298 (48470944)	c2_1717,c2_45697 (51404699,43326283)
ma	Chr05 18 (0-24)	4.97	(3.36-3.96)	6.46	c2_11604 (1957584)	c2_33543,c2_11924 (1505540,3708760)
ma	Chr06 26 (22-30)	4.65	(3.36-3.96)	7.41	c2_32938 (31631446)	c2_51765,c1_15811 (38211470,7995004)
suc	Chr02 81 (70-92)	4.75	(3.37-4.01)	7.18	c1_8490 (44698431)	c2_50405,c1_5920 (35657856,47808850)
suc	Chr05 38 (18-46)	5.00	(3.37-4.01)	7.23	c2_47607 (5972364)	c2_11604,c2_51915 (1957584,10714562)
sufrr	Chr01 46 (36-56)	4.39	(3.35-3.99)	5.96	c2_20803 (60987055)	c2_35518,c2_32112 (60512580,42197568)
sufrr	Chr03 76 (62-89)	4.43	(3.35-3.99)	6.89	c2_29513 (47493308)	c2_47801,c1_6405 (54033139,45305706)

(Table 2.4 Cont.)

sufrr	Chr11 44 (30-62)	4.65	(3.35-3.99)	7.08	c2_24318 (9100472)	c2_30369,c2_44868 (33082201,6771527)
sufrr	Chr12 84 (81-98)	6.78	(3.35-3.99)	11.15	c1_2724 (57460060)	c1_11657,c1_2690 (54319182,58047558)
suglur	Chr03 76 (49-81)	4.55	(3.38-4.14)	6.92	c2_29513 (47493308)	c1_433,c1_10737 (51038035,16887518)
suglur	Chr12 84 (81-99)	6.66	(3.38-4.14)	11.10	c1_2724 (57460060)	c1_11657,c1_2690 (54319182,58047558)
sutrsr	Chr01 46 (36-56)	4.26	(3.37-4.07)	5.65	c2_20803 (60987055)	c2_35518,c2_32112 (60512580,42197568)
sutrsr	Chr03 76 (59-88)	4.43	(3.37-4.07)	6.85	c2_29513 (47493308)	c2_1724,c1_6390 (51404231,45198406)
sutrsr	Chr11 43 (30-67)	4.26	(3.37-4.07)	5.96	c2_53687 (9380556)	c1_4822,c2_44868 (38231783,6771527)
sutrsr	Chr12 84 (81-98)	7.25	(3.37-4.07)	12.10	c1_2724 (57460060)	c1_11657,c1_2690 (54319182,58047558)
v24mae	Chr03 98 (73-101)	4.64	(3.24-3.87)	7.02	c1_111 (57295688)	c1_111,c2_11211 (57295688,46528226)
v24mae	Chr07 34 (28-47)	4.87	(3.24-3.87)	7.86	c2_46755 (4880359)	c2_36882,c1_3159 (3314985,41078607)
v2a	Chr03 85 (73-100)	4.60	(3.2-3.93)	8.12	c2_1717 (51404699)	c1_111,c2_11211 (57295688,46528226)
v2a	Chr07 34 (28-47)	5.17	(3.2-3.93)	8.77	c2_46755 (4880359)	c2_36882,c1_3159 (3314985,41078607)
v2mae	Chr03 85 (73-100)	4.60	(3.2-3.93)	8.12	c2_1717 (51404699)	c1_111,c2_11211 (57295688,46528226)
v2mae	Chr07 34 (28-47)	5.17	(3.2-3.93)	8.77	c2_46755 (4880359)	c2_36882,c1_3159 (3314985,41078607)
v46	Chr03 64 (54-68)	5.97	(3.25-3.9)	9.85	c1_13843 (43469293)	c1_8059,c2_57349 (45970632,41780626)
v4mae	Chr03 57 (48-69)	4.22	(3.3-3.89)	5.98	c1_6355 (45036809)	c1_526,c1_10519 (46395963,42366082)
vfa	Chr04 13 (0-34)	4.19	(3.23-3.87)	5.49	c2_26792 (9523209)	c2_39800,c2_51639 (59829029,5488735)

[a] Trait abbreviations are listed in Table 4.1. Traits are from overall genotype BLUPs, [b] 95% confidence intervals calculated from 200 permutations (Nettlton and Doerge, 2000), [c] Physical positions are from the PGSC DM v4.03 sequenced genome (Sharma et al., 2013), [d] Flanking markers were at each end of a 1.5 LOD support interval

Figure 2.1. Graph of 12 potato chromosomes comparing the genetic location (cM) to the physical position (Mb) of SNP markers (open circles) as published by the Potato Genome Sequencing Consortium.

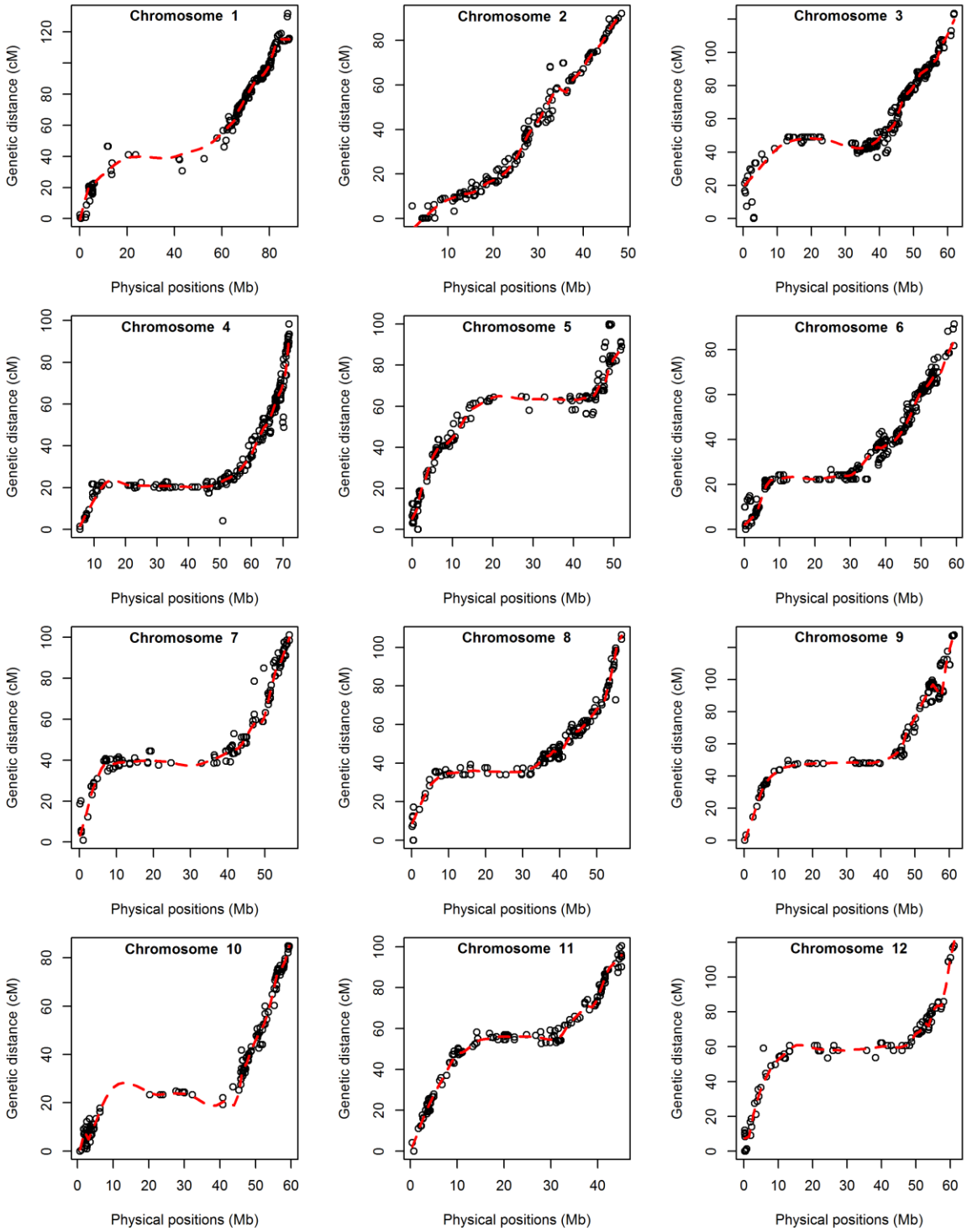


Figure 2.2. Trait distribution patterns for sugar concentrations in the Wauseon x Lenape population grown in the 2012, 2013, and 2014 field seasons.

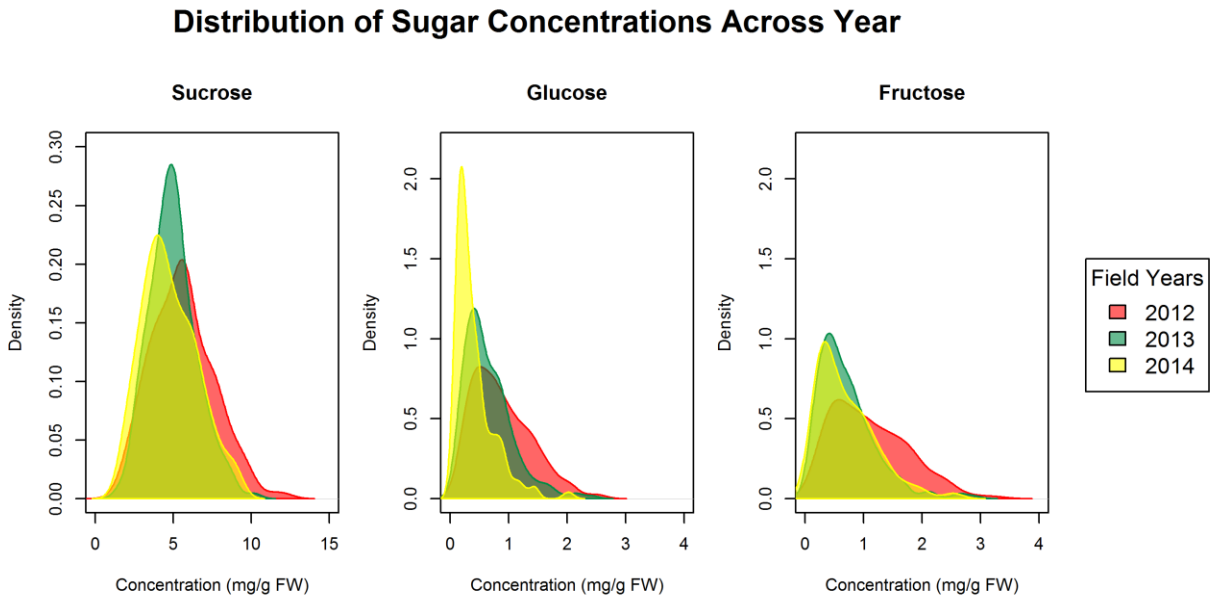


Figure 2.3. Trait distribution patterns for measures of *VInv* relative expression using mean of reference genes or single reference genes and the $-\Delta\Delta C_T$ method. Data were collected from the Wauseon x Lenape population grown in the 2012, 2013, and 2014 field seasons. 60S data was only available for the 2014 field season.

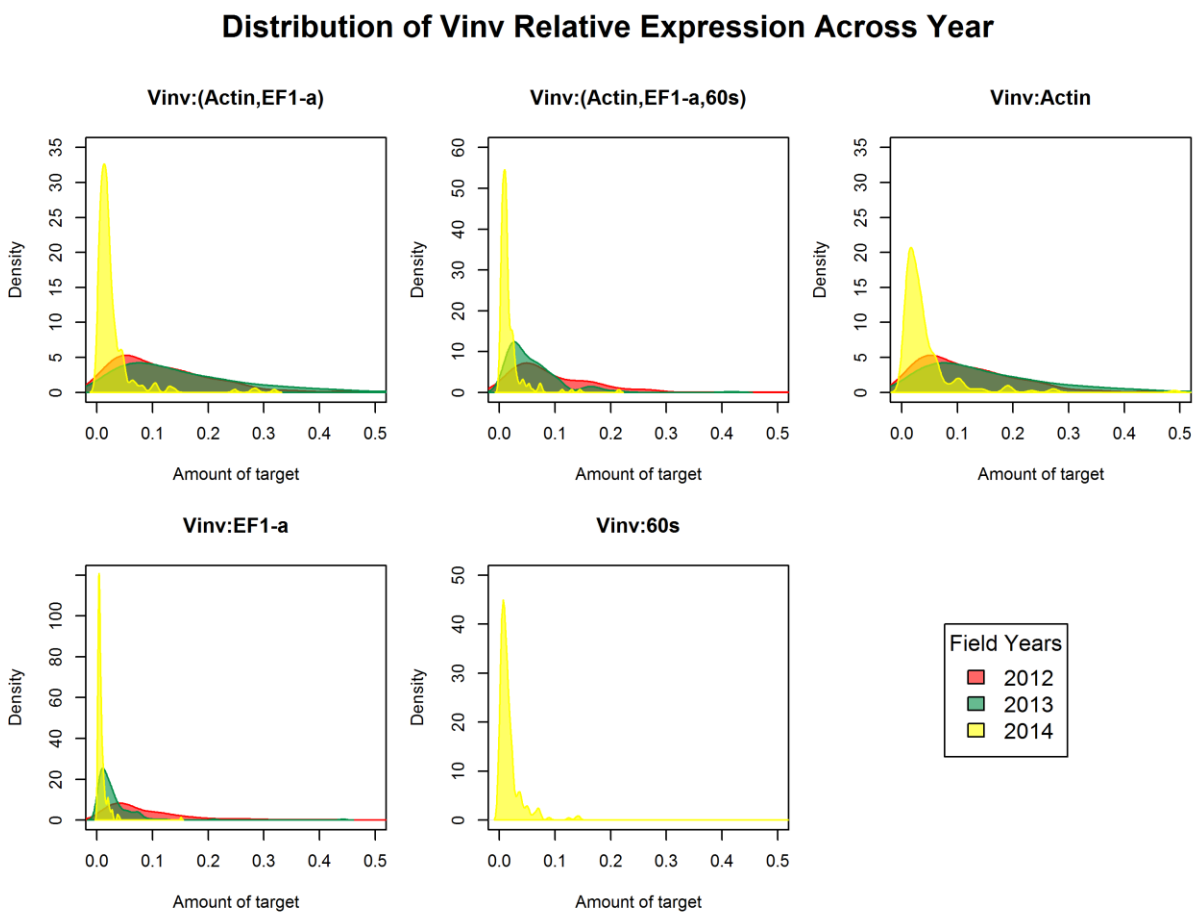


Figure 2.4. Correlation plot for sugar traits, organic acids, and *VInv* relative expression in a Wauseon x Lenape population of 192 individuals. Pearson's correlation coefficients are indicated numerically in one half of the plot and graphically in the other half.

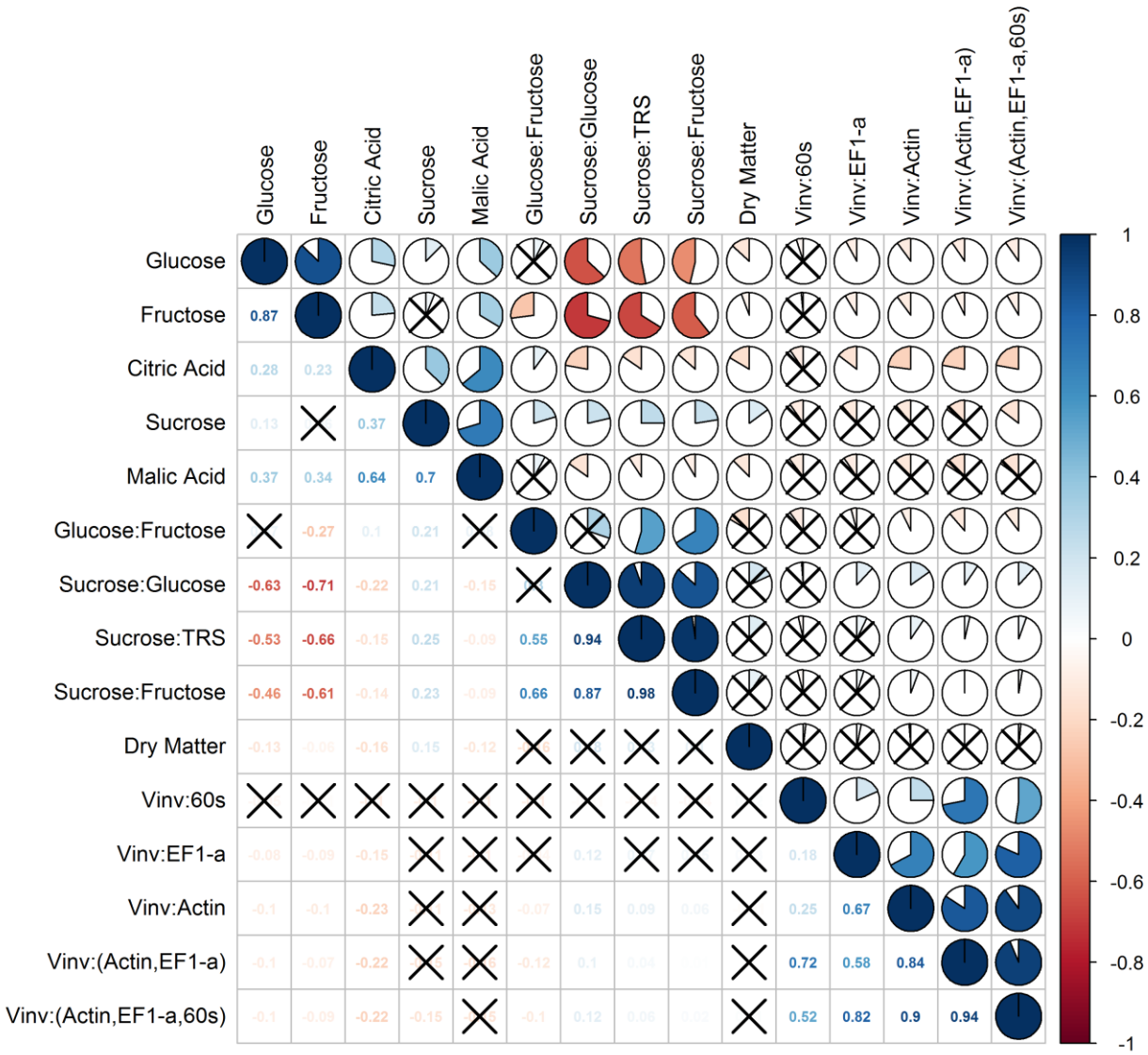
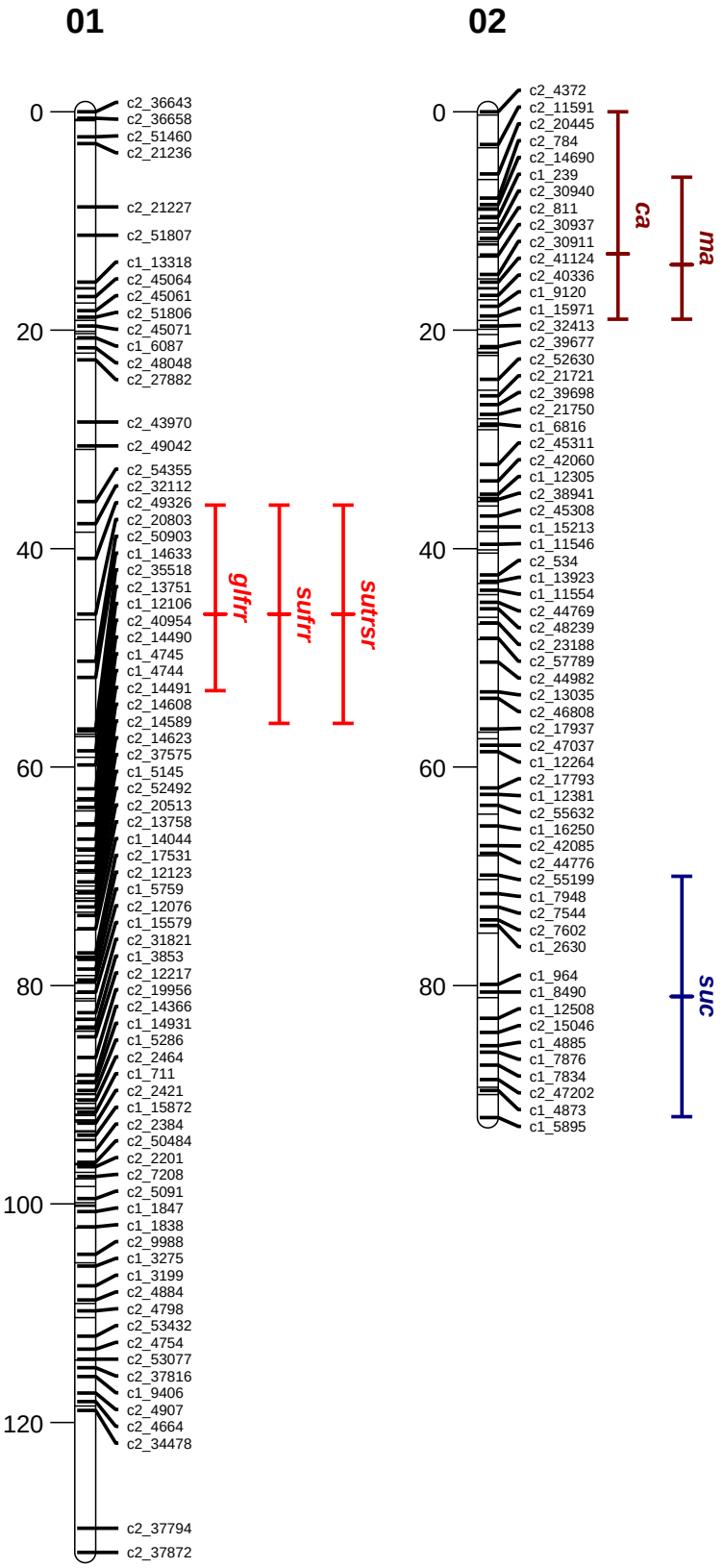
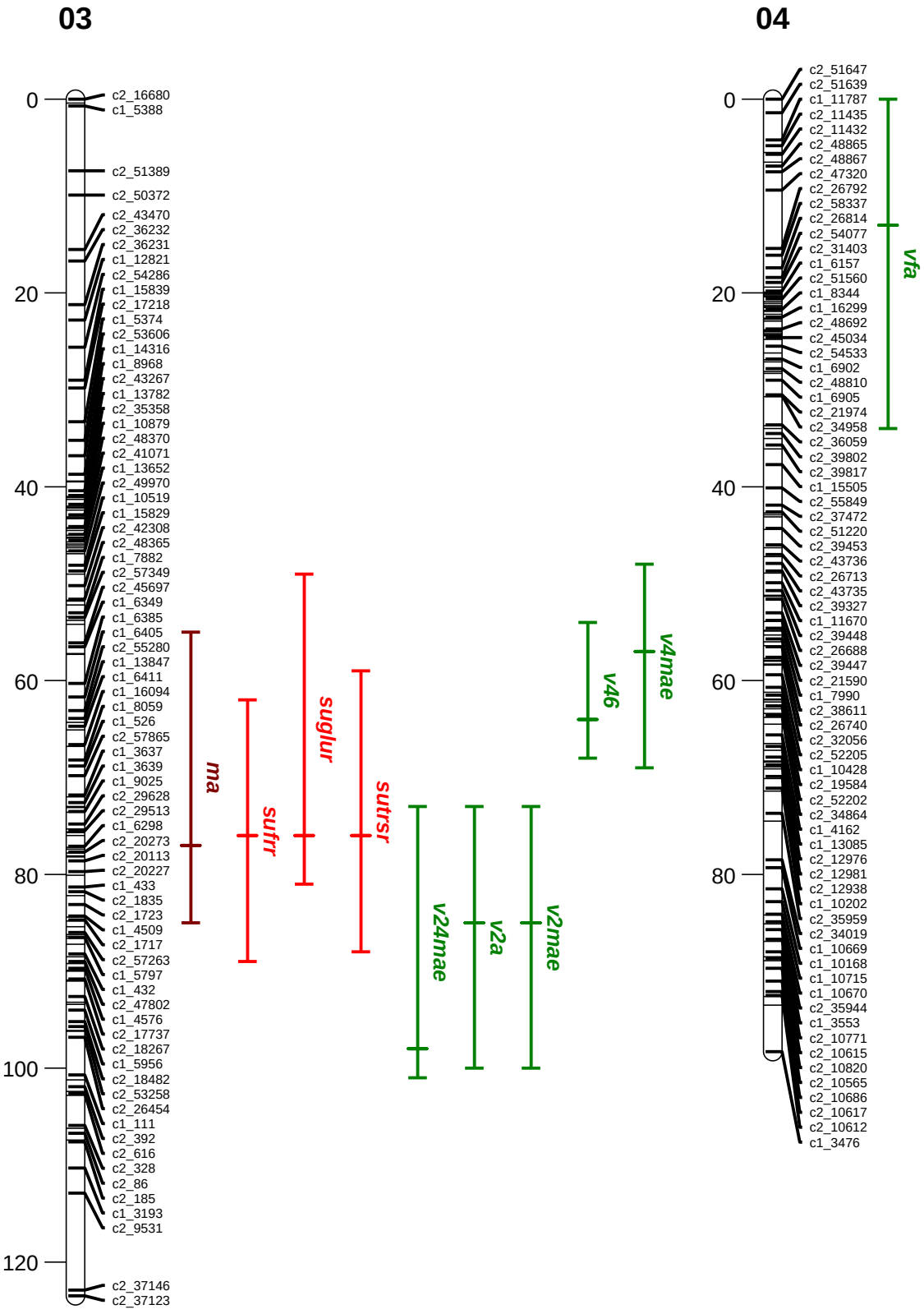


Figure 2.5. Genetic map of Wauseon x Lenape population with QTL for chip processing traits and their 1.5 LOD support intervals. 12 Linkage groups with markers binned every 1 cM and distance in cM. Trait abbreviations are the same as those in Table 2.1. Mapped QTL positions are indicated by vertical lines to the right of each chromosome. The length of the line indicates the 1.5 LOD support interval with the center tick indicating the mapped position. Colors of the QTL lines and text indicate the trait type: sugars are blue, sugar ratios are bright red, organic acids are maroon, and *VInv* relative expression values are green.

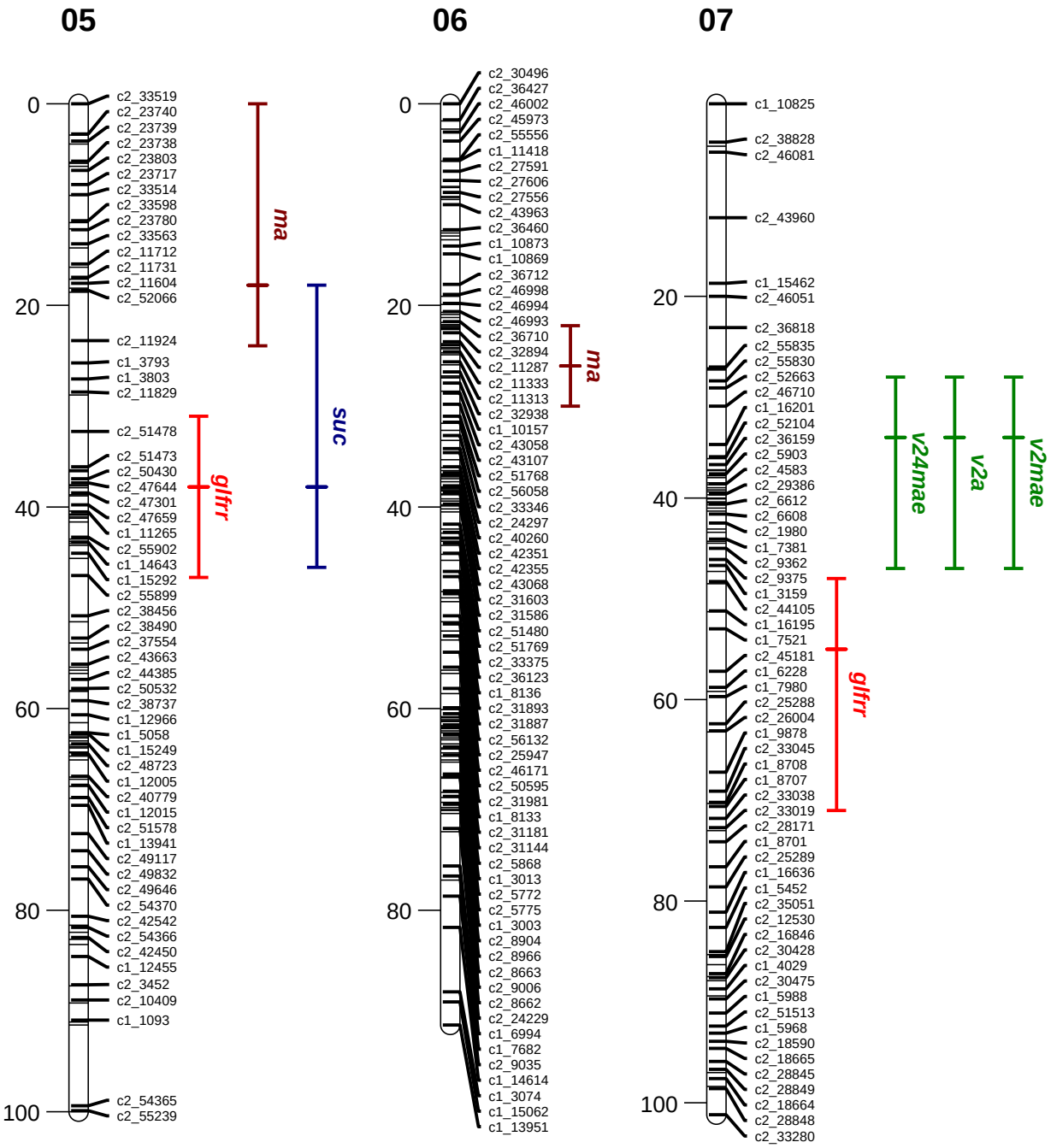
(Figure 2.5 Cont.)



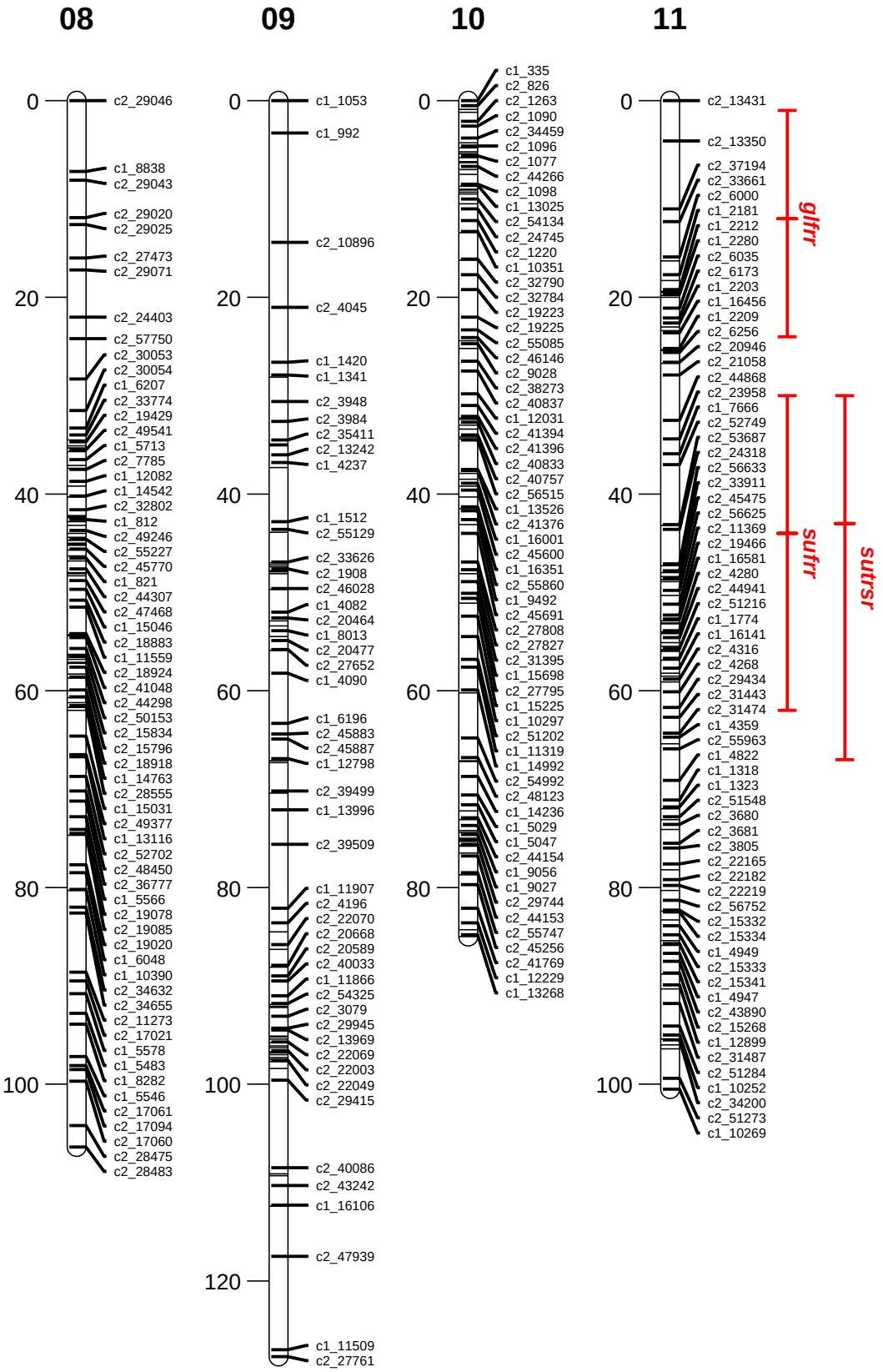
(Figure 2.5 Cont.)



(Figure 2.5 Cont.)



(Figure 2.5 Cont.)



(Figure 2.5 Cont.)

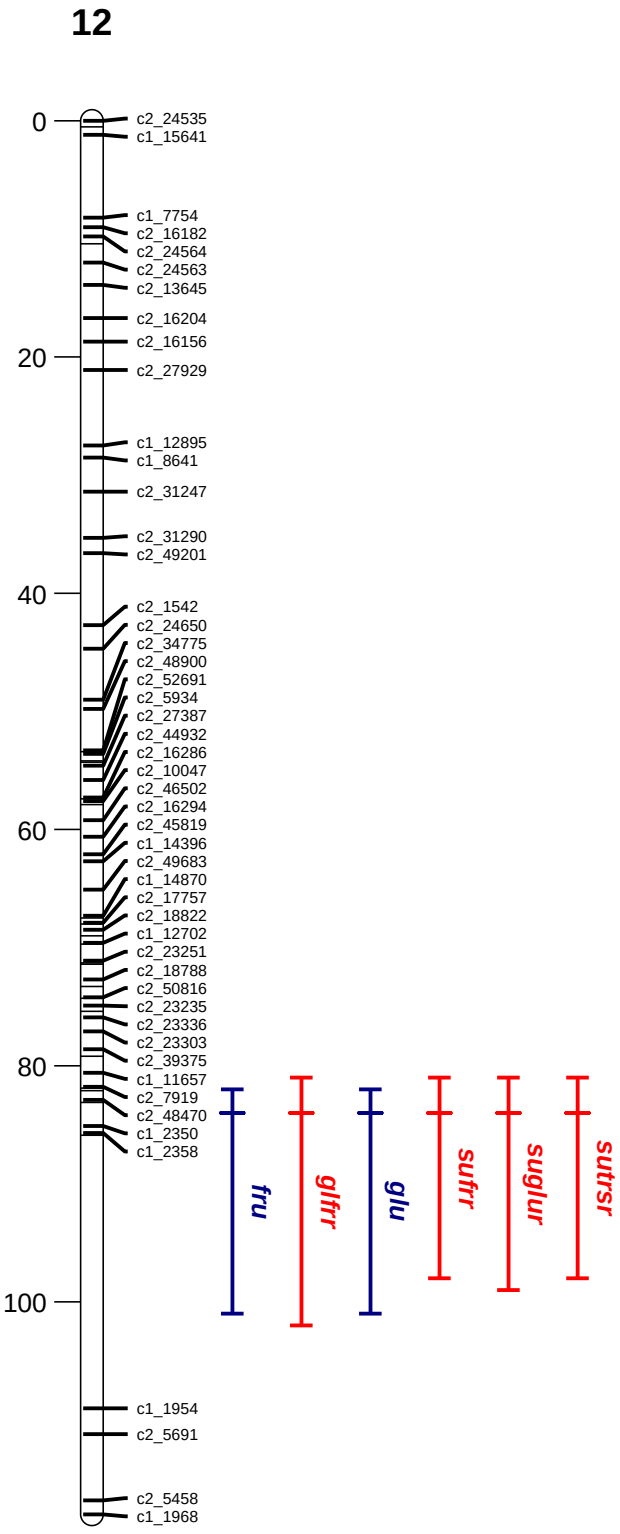
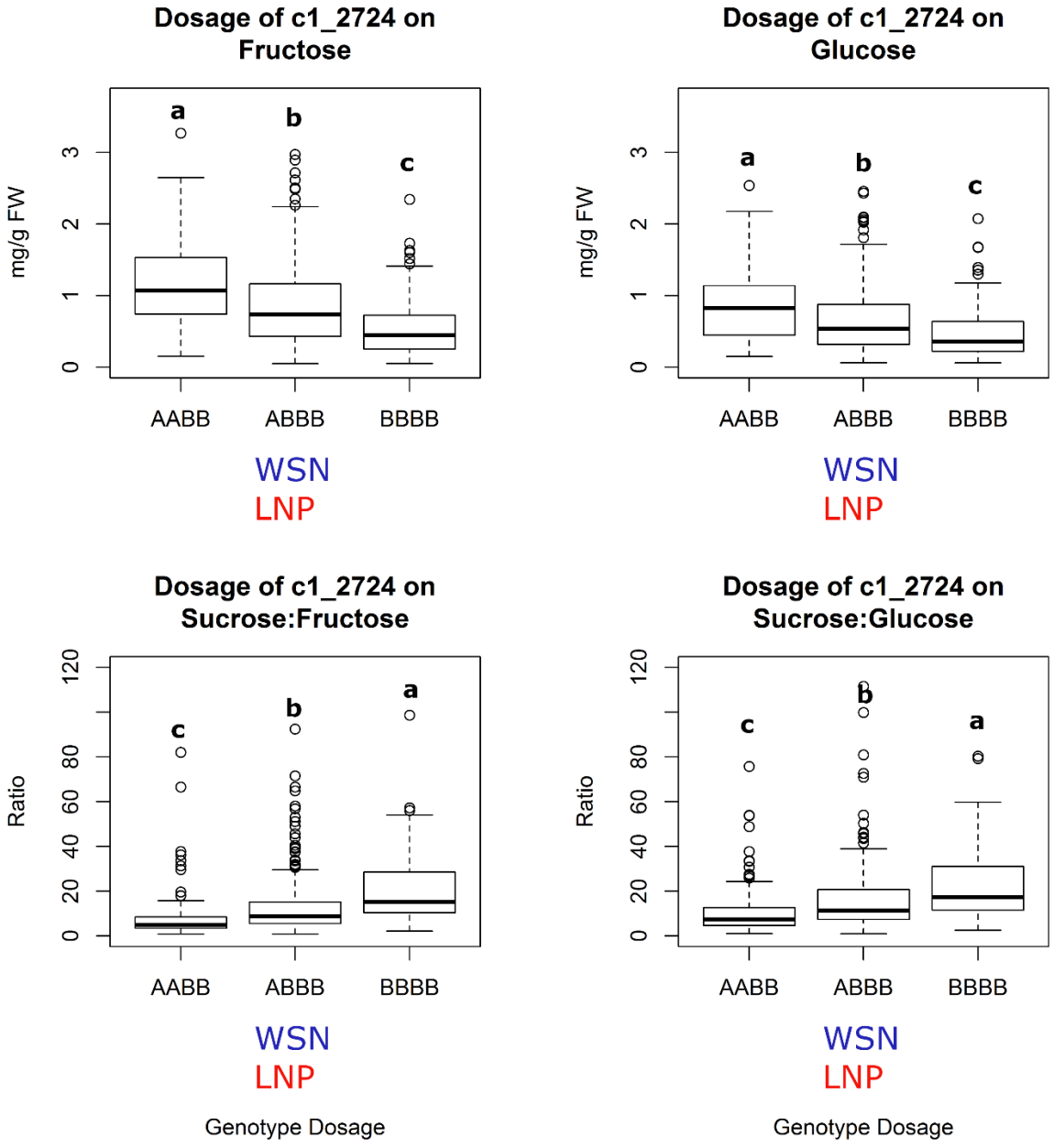


Figure 2.6. Single Marker Analysis of markers positioned closest to LOD peak for QTL mapping of fructose, glucose, sucrose:fructose ratio, and sucrose:glucose ratio on a population of 191 individuals across three seasons. The box plots measure the distribution of sugar (mg/g FW) on the top two plots and ratio of sugars in the bottom two plots. The bold line indicates median, box lines encompass the middle 50% quantile, and whiskers encompass the middle 99.3% quantile. Letters designate significance as determined by Tukey's HSD at a p value of 0.05. WSN and LNP designate the dosage of the two parents.



Chapter 3: Identification of QTL for Stem-End Chip Defect and Chip Color Traits in a Lenape-Derived Full-Sib Population

3.1 Abstract

The objective of this study was to identify QTL for economically important traits in chip processing potatoes. A bi-parental population was developed by crossing the chip processing varieties Wauseon and Lenape. Parents and 191 individual progeny were evaluated in 2012, 2013, 2014, and 2015 for chip color and stem-end chip defect (SECD) at January, March, and May; and for tuber specific gravity and vine maturity. The population was genotyped using the SolCAP 8303 Infinium array and a linkage map was constructed using 2,355 single nucleotide polymorphism (SNP) markers. QTL mapping was performed using TetraploidMap 2.0 software and 60 significant QTL were detected: 9 for SECD, 28 for chip color, 17 for chip color change in storage, 2 for maturity, and 4 for specific gravity. QTL were detected on every chromosome and there were regions where multiple QTL from different traits overlapped. The largest effect QTL for SECD was on chromosome III and mapped to a position within 2 Mb of vacuolar acid invertase and the largest effect QTL for chip color mapped to chromosome VI. There were several other QTL that provide insights into the genetics of chip processing quality traits.

3.2 Introduction

Potatoes grown for potato chip processing account for 21 percent of potatoes grown for food in the US (National Potato Council, 2015). Growing potatoes is an expensive endeavor that requires intensive management and large capital investment in inputs, equipment and facilities (Patterson, 2007). For the producer, the value of the crop is highly dependent on its quality and reductions in marketable yield can result in an unprofitable season (Patterson, 2007)(Arnott and Smart, 2016). Visual appearance of the potato chip product is important to consumers and is based on the color of the chip (Scanlon et al., 1994)(Sweetman, 1930). This chip aesthetic is caused by factors involving potato biochemistry,

physiology, and growing decisions. Processors adjust payment on a shipment of potatoes if chip color is not acceptable in order to incentivize better growing and storage practices (Curtis and McCluskey, 2003)(Bolotova and Patterson, 2009).

Maillard reaction products produced during cooking primarily impact the consumer's acceptance of a cooked potato chip (Scanlon et al., 1994). This is a non-enzymatic reaction between reducing sugars and alpha-amino groups of amino acids and proteins to form brown-colored pigments (Habib and Brown, 1957)(Shallenberger et al., 1959)(Scanlon et al., 1994). The reducing sugars glucose and fructose accumulate in potato tuber tissue as end products of a metabolic pathway that includes starch breakdown, sucrose synthesis, and hydrolysis of sucrose by vacuolar acid invertase to fructose and glucose. There are two major contributors to elevated amounts of reducing sugars in stored potatoes, cold-induced sweetening (CIS) and senescent sweetening. Cold induced sweetening, which generally occurs when potatoes are stored below 10°C (Burton, 1969), is managed carefully in order to balance crop longevity in storage with the ability to make quality processed products (Sowokinos, 2001). Senescent sweetening is an irreversible, developmental process that occurs near the end of the storage season when tubers begin accumulating sucrose and reducing sugars (Colgan and Rees, 2012). While many of the enzymes participating in these processes have been identified (Sowokinos, 2001)(Menendez et al., 2002)(Chen et al., 2001), there is sustained demand for cultivars that maintain low amounts of reducing sugars during low-temperature storage and late season storage.

There is growing interest in understanding how stress during the growing season results in chip quality defects (Bussan et al., 2009)(Wang et al., 2015). Stem-end chip defect (SECD) is a notable chip quality defect caused by abiotic stress that has affected the US potato chip industry. SECD is characterized by a dark-colored, post-fry blemish on chips at a position corresponding to the stem end of the tuber tissue. Reducing sugars accumulate adjacent to the vasculature and, when chips are fried, create a dark discoloration due to products from the Maillard reaction. SECD is weakly associated with early to mid-season heat and water stress (Wang et al., 2012) and its severity is correlated with high nighttime

temperatures (Dickman, 2016). Genotypic differences have been observed for both severity at harvest and change in SECD severity through storage (Dickman, 2016)(Wang et al., 2015). SECD in production systems occurs erratically, and results in raw product that is unsuitable for the chip industry (Dickman, 2016)(Wang et al., 2015). It is unclear how other processing traits might be related to the severity of SECD and its genetic basis has not been established.

SECD has similarities with sugar end defect, a dark-color defect observed in processing russets. Sugar end defects are usually found on the tuber stem end and are caused by increased reducing sugar accumulation (Iritani and Weller, 1973)(Kleinkopf, 1976)(Zhu, 2014). Environmental stress events including moderate heat and drought stress during early and middle bulking stage, are thought to cause sugar end defects (Eldredge et al., 1996)(Bethke et al., 2009)(Thompson, 2008). Genetic resistance to sugar end defects has been observed (Shock et al., 1993)(Wang et al., 2015)(Dickman, 2016)(Thompson-Johns, 1998).

In response to the need for better chip processing quality, there have been efforts to breed for improved chip processing varieties as well as to identify genetic regions that impart superior quality (Love et al., 1998)(Hirsch et al., 2013)(Douches et al., 1996). One individual cultivar that has been notable in the modern era of potato breeding is Lenape. At the time of its release, it had greater resistance to CIS than other varieties and tubers had very high dry matter content, another important chip processing trait (Love et al., 1998)(Lulai and Orr, 1979). The pedigree of this cultivar included a *Solanum chacoense* accession that is thought to have conferred cold sweetening resistance (Akeley, 1968) but also may have contributed to higher glycoalkoloid levels in Lenape tubers, which led to its withdrawal from the marketplace (Zitnak, 1970). Nevertheless, Lenape is in the pedigree of many important chip processing varieties including Atlantic, Snowden, Belchip, Gemchip, Pike, NorValley, Tundra, Accumulator, Lamoka, as well as fry processing varieties such as Highland Russet, Blazer Russet, and Alturas amongst others (Van Berloo et al., 2007). Understanding the allelic contributions from Lenape that confer preferential traits in progeny would assist development of new cultivars.

In classical breeding studies, QTL mapping methods have been used to uncover genomic regions related to a rare trait (Bernardo, 2010). QTL mapping in autotetraploid crops such as potato has been limited due to constraints imposed by non-preferentially pairing of homologous chromosomes (Milbourne et al., 2009)(Hackett et al., 2001). In addition, marker arrays based on single nucleotide polymorphism (SNP) markers were difficult to implement for autotetraploid linkage populations. However, Hackett et al. (Hackett et al., 2013) developed a computational method whereby SNP markers of all dosage categories could be used to map QTL on a single linkage map. This permitted QTL mapping studies to be performed on potato populations that have genetic makeups similar to those developed by potato breeding programs for commercial production (Hackett et al., 2014)(Rak, 2015)(Massa et al., 2015)(Zorilla Cisneros, 2013). This mapping software, TetraploidMap, was used in the evaluation of a bi-parental population developed from a cross between Wauseon and Lenape.

3.3 Materials and Methods

Population growth, phenotyping and Data Collection

The materials evaluated in this portion of the study were from a Wauseon x Lenape population consisting of 191 individuals in addition to the parents. These materials were grown in 2012, 2013, 2014, and 2015 at the Hancock Agricultural Research Station in Hancock Wisconsin. All plots were managed using standard commercial procedures for the central sands growing region of Wisconsin. Each year contained one replication of the set of full siblings. The 2013, 2014, and 2015 field years were grown in a modified augmented block design II (Lin and Poushinsky, 1985) with Atlantic as the primary check. Wauseon, Lenape, Snowden, and Megachip served as the secondary checks in this design. Planting occurred each year in the first week of May and harvest commenced 120 days after planting (DAP). Vine maturity was assessed at 105 DAP on a 1-5 scale (1- Vine 100% senescence, 2- Yellowing, 3- Green, prostrate vine, 4 –Green, upright vine, 5- Very late, full bloom). Vine killing was performed in two applications 10 days and 3 days prior to harvest.

The storage evaluations occurred in the winter following each of the growing seasons. Harvested tubers were placed in plastic crates and randomized in relation to their field order before being stored on pallets. Samples were stabilized at 12.7^{0}C and 95% relative humidity in storage lockers at the Storage Research Facility at the Hancock Agricultural Research Station. Samples were kept in this condition for approximately 8 weeks to allow for wound healing after which they were treated with the sprout inhibitor CIPC. At 10 weeks after harvest, storage temperature was gradually ramped down to a final temperature of 8.9^{0}C over the course of 25 days in order to avoid a stress response in the tubers that would result in elevated sugar accumulation.

Specific gravity was measured in January using 5 kg of tuber material (Welltech PW-2050, Welltech International Limited, Marshville, NC). Six representative tubers from each field plot were chip processed at three periods through the storage season: mid-January, mid-March, and mid-May. The exception to this was the 2012-13 storage season, when the amount of tuber material permitted evaluations only on the January and March processing dates. For chip processing, tubers were cut longitudinally from bud to stem end such that one tuber half contained both the stem and bud ends. Four chips were removed from this half using a custom mandolin that reliably created 1.1 mm thick chip slices. The first slice was discarded from the chosen tuber half and the other three slices were collected. Chip slices from all 6 tubers from a plot were aggregated (18 chips total), rinsed sequentially in two bowls of cold water, and blotted dry. Chips were fried flat in custom baskets so as to more easily score visual defects. All chip samples were fried at 182^{0}C for 2 minutes and 10 seconds. After cooling, chips were bagged for later visual scoring of SECD.

Chips were evaluated for SECD on a 0-5 scale as described by (Wang et al., 2012) (Figure 3.1) After this, chips were crushed to particle sizes smaller than 1 cm in diameter and read for reflectance color using a HunterLab D25LT colorimeter (Hunter Associates Laboratory Inc., Reston, VA) with measurements reported in the Lab color space. This color space is used because it correlates well with the visual range of the human eye (Tkalčič and Tasič, 2003). The Lab color space is comprised of 3

dimensions; L* which is relative lightness, a* is the color range between red and green, and b* is the color range between yellow and blue. There is strong correlation between visual scores and the L*, a*, and b* vectors in the context of fried potato chips (Rak, 2015).

Processing Traits Metrics

The processing traits of overall chip color (Lab) and SECD score were also analyzed to take into account the time series aspect of these measurements through the storage season. The slope of the L*, a*, b*, and SECD values were calculated between January and May.

Data Analysis

Phenotypic data were evaluated using R statistical software (R Core Team, 2014) and custom scripts. All traits were evaluated for autocorrelation using the corplot package in R. Pearson's correlation was performed on all complete observations between the data sets. Linear mixed models were used to estimate the genotypic variance component for each trait across all years by treating genotype as a random effect. This was achieved using the R package lme4 (Bates et al., 2014). Best linear unbiased predictions (BLUP) were generated using a mixed model to get a genotypic performance value for each trait (Robinson, 1991)(Henderson, 1974).

The model used in the mixed model was:

$$y = \mu + G_i + Y_k + (GY)_{ik} + \varepsilon_{ijk}$$

Where μ is the overall mean, G is the effect of the genotype *i*, Y is the effect of year *k*, GY is effect of genotype *i* by year *k* and ε is the residual error. The field plots had incomplete blocks but block was omitted from the model because it did not account for a significant amount of the variance, nor did it affect the residuals in any of the reported traits. The residuals for all models met assumptions of normality and heteroscedasticity.

Broad-sense heritability was estimated for each trait by using the estimated variance components from the lme4 package. The equation was:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{GY}^2}{y} + \frac{\sigma_e^2}{ry}}$$ (Holland et al., 2003)

The broad-sense heritability is denoted by H^2, σ_G^2 is the genotypic variance, σ_{GY}^2 is the genotype by year variance, σ_e^2 is the residual variance, y is the number of years and r is the number of replicates within year.

Genotyping and Mapping

The population in this experiment was genotyped using the SNP 8303 array (Hamilton et al., 2011). DNA was extracted using the Qiagen DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) and genotyped on an Illumina iScan system (Illumina, San Diego, CA) at Michigan State University. Called SNPs and theta value data sets were returned for each individual in the population. Marker data consisting of 8,303 SNP markers (Hamilton et al., 2011) were filtered and mapped as described in Chapter 2 with a final number of 2,355 SNP markers on 12 linkage groups. Mapping was performed using Tetraploid Map 2.0, developed by Christine Hackett at the Scottish Crops Research Institute. Significance thresholds for LOD scores was determined for each trait using 200 permutation tests to establish the 95% confidence interval (Nettleton and Doerge, 2000).

3.4 Results

Phenotypic analysis

Heritability estimates were high for many chip color traits (Tables 3.1 and 3.2). Variance components (Table 3.2) had differing proportions of total variance depending on the trait but most traits exhibited a strong genetic effect in this population and this resulted in identification of 60 QTL (Table 3.3).

Chip color traits

Overall L* values ranged from 23.89 to 59.21, a* values ranged from -0.31 to 11.30, and b* values ranged from 0.24 to 24.77 across all years. There was a significant year effect on all color scales at all processing periods (Figure 3.2). One notable result was for January L* value data where data from 2012 had a mean of 48.52 (range 35.54-57.49), 2013 had a mean of 52.07 (range 44.12-58.89), 2014 had a mean of 47.76 (range 27.35-56.25) and 2015 had a mean of 43.70 (range 28.54-55.06). The color values reported in the Lab color space had high heritability estimates. May fry processing date a* values reported the highest heritability of this trait group with 0.89. Heritability estimates for many of the March and May chip color traits were higher than those for the January processing period. January L* value heritability was 0.82 while March and May L* values were of 0.87 and 0.88. March processing time tended to have the highest variance attributed to genotype by year interactions for L*, a*, and b* values (Table 3.2). The most heritable slope traits were chip color slopes between January and May with L* (0.71), a* (0.66), and b* (0.74).

Stem-end chip defect

The SECD trait ranged from 0-5 in all years except 2012 when no 0 rankings were observed. Across the entire study, variance attributed to year was 0.333, 0.3657, and 0.505 for SECD ranking in January, March, and May, respectively. SECD incidence was highest in 2012 with an overall mean of 2.1. The overall means for 2013, 2014 and 2015 were 2.0, 1.9 and 2.1. There was a strong genetic component of the variance observed for all processing dates. The highest variance attributed to genotype was 0.717 for March SECD score. SECD scores for March and January were lower at 0.649 and 0.638 respectively (Table 3.2). Variance attributed to year was 0.011, 0.067, and 0.076 for SECD score in January, March, and May, respectively. Distribution of mean SECD score was normal for all years. In Figure 3.3, genotypes were binned based on whole plot score at the January processing period. Differences between parental phenotypes differed in severity between years but ranking remained consistent.

Specific gravity and Vine Maturity

Tuber specific gravity had a range of 1.056 to 1.108 and a mean of 1.082 across all years in this study. The majority of the variation in this trait was explained by the genetic component (0.000042) with variance attributed to year (0.000013), lower than residual variance (0.000016). The heritability of this trait was 0.88.

Vine maturity ratings of progeny covered the full breadth of the vine maturity scale in all years. The mean maturity ranking across all years was 2.05. The variance attributed to year (0.035) was smaller than that for residual (0.063), genotype by year (0.250), or genotype (0.386). This trait had a high proportion of genetic control and had a heritability of 0.79 (Table 3.2).

Trait Correlation

Several traits had strong correlations within this study. Many of the chip color traits were highly correlated (Figure 3.3). L* values for each processing period were positively correlated with L* values from other processing periods (all R^2>0.57). Similar correlations were also observed within the a* and b* metrics. There were strong negative correlation between a* values and L* values within and between processing dates, and strong positive correlations between b* and L* values. SECD score at the different processing periods had positive correlation coefficients of 0.47 to 0.55. SECD score had positive correlations with a* values across processing times, with correlation coefficients ranging from 0.30 to 0.46 and negative correlations with L* value measurements, with correlation coefficients ranging from -0.26 to -0.41 (Figure 3.3). Maturity had few significant correlations with other traits. The highest correlations were with the slope of L* (January – May) 0.27 and the slope of b* (January- May) 0.25. Maturity did have significant positive correlations with SECD score across the processing dates, and correlation coefficients ranged between 0.14 and 0.18. Maturity also had a positive correlations of 0.17 with L* value (May) and 0.16 with b* value (May).

QTL analysis

The QTL detected by the additive dosage genotype probabilities model are presented in Table 3.2 and Figure 3.5. There were 60 QTL detected in total for the traits of SECD in all months of processing, chip color as described as L* or a* values in all processing months, specific gravity, and vine maturity. There were also QTL detected for changes over time in both SECD and chip color metrics. Significance, determined by the sequential permutation test, was slightly different for each trait but with minimal variation. For all traits presented, the mean LOD threshold was 3.71 with a mean 95% confidence interval of (3.37-4.05). Most of the detected QTL were of moderate effect with 8-20% of the variation explained. QTL resolution, defined by a 1.5 LOD support interval, was on average 22.6 cM, an expected result given the limited recombination in a full-sib population. The best resolution was observed for vine maturity on chromosome V with 5 cM LOD support interval. Other high resolution QTL include Slope of L* (January – March) (Chr05, 5 cM), Slope of a* (January – March) (Chr09, 6 cM), and SECD score (January) (Chr02, 7 cM).

QTL for Chip color

Twenty-eight QTL for chip color descriptors at various measurement periods were identified. QTL in this category had both the highest variance explained, L* value (March) (Chr06, 69 cM) with 16.26% and one of the lowest, L* value (January) (Chr11, 11 cM) with 5.80%. There were twelve QTL amongst the L* and a* scores for January, March, and May that explained more than 10% of the variation in the trait. For example, the values for L*, a*, and b* at the May processing date accounted for 15.95%, 15.11%, and 14.99% variance, and all mapped between 34 and 59 cM on Chr06. Significant QTL for chip score clustered together in certain locations along the genome besides those found on chromosome VI. Chromosome I had mapped locations for L* value in January, March, and May; the b* value for January and May; and the a* value in January in the range of 44-51 cM. Chromosome III contained QTL for L* value in January and March; a* value for March and January; and b* value for January and March that

mapped between 30-63 cM. Additionally, chromosome IV had QTL of L* value for May and a* and b* values for March and May that mapped near 78 cM (Figure 3.3).

There were 17 significant QTL for metrics of chipping color change through the storage season. The slope traits had different effect sizes, LOD peaks, and LOD interval sizes. The slope of b* (January-May) had the highest LOD score of 7.03 in this trait segment and explained 11.61% variance (Chr06, 42 cM). There were traits in this segment with LOD intervals that overlapped with single processing time chip color traits. Chromosome IV contained QTL for slope of L* and b* from January-May and slope of b* from January-March with a LOD interval between 56 and 85 cM positions. The slope traits of L* from January-May and March-May as well as slope traits for b* from January-May and March-May were located on chromosome VI in the same interval as chip color traits in May for L*, a*, and b*. Slope traits located on chromosome X and XII were not collocated with chip color traits. Slope of a* (January-May) explained 8.32% variance and had a LOD score of 5.07 (Chr10, 34 cM). This was an area of the genome where a QTL for slope of a* (January-March) was also mapped. Chromosome XII also had QTL for slope of a* from January-May and January-March and slope of a* (January-May) had a LOD peak of 5.59, a LOD interval of 80-99 cM, and accounted for 9.25% variance.

QTL for SECD

Nine QTL for SECD were detected across the genome for all processing time points; 4 in January, 2 in March, and 3 in May (Figure 3.6). The strongest effect QTL was for SECD score (January) (Chr03, 47 cM) and accounted for 12.87% variance. The next strongest QTL was for SECD score in May, occupied the same area of the genome (Chr03, 41-54 cM), and accounted for 10.58% variance. QTL for SECD traits did not always collocate with QTL for chipping traits. Chromosome II contains a QTL for SECD score (January) that explained 7.25% variance and had a moderate LOD interval of 11-18 cM. Chromosome IV contains a QTL for SECD score (May) between 0-10 cM that explains 5.23% variance. Neither of these two QTL overlaps with a chip color QTL.

QTL for Specific Gravity and Maturity

Two QTL were detected for vine maturity and 4 for tuber specific gravity. The vine maturity QTL on chromosome V had the highest LOD peak of 25.46 in a LOD interval from 29-34 cM and explained 39.06% variance. The other maturity QTL was present on chromosome I and had a smaller effect size, accounting for 6.63% of trait variance, a 4.34 LOD peak, and a LOD interval from 42-66 cM. Specific gravity QTL were located on chromosomes I, III, VI, and VIII. The largest effect QTL for this trait was on chromosome VI, explaining 10.06% variance with a LOD interval of 56-85 cM. The QTL for specific gravity on chromosome I, 6.69% variance explained, and chromosome VIII, 6.29% variance explained, were not located within the LOD intervals of other chip processing traits.

Single Marker Analysis

Markers closest to the mapped QTL positions were used in single marker analysis (Figures 3.7 and 3.8). In this analysis, marker dosage was used to classify all individuals and the phenotype was summarized by marker dosage category. There were several instances in the study where the dosage of the marker was significantly related to the observed trait and each indicate a potential marker-trait association (Gebhardt et al., 2004)(Li et al., 2013). In Figure 3.7, 3 markers in January SECD QTL displayed significant relationships to SECD score. Marker c2_30911 showed evidence of Lenape contributing the beneficial allele. This is also the case for the QTL on chromosome 6 mapped to marker c2_51758. Figure 3.8 shows the single marker analysis for specific gravity and for chip color L* in January and March. Again, there are significant relationships between marker dosage and trait. The minor allele, contributed by Lenape, is associated with the desired phenotype in each example.

3.5 Discussion

This population and the phenotyping methods used were very favorable for the detection of QTL for chip processing traits. Several QTL identified occupy the same intervals along the genome and there are characteristics of these traits that suggest that the same underlying candidate genes are could be involved. This was expected for the chip color traits in the Lab color space because both L* and a* were

highly correlated (Figure 3.4) and L* and a* are impacted similarly by products of the Maillard reaction (Habib and Brown, 1957)(Rak, 2015)(Scanlon et al., 1994). However, there were similar QTL locations among traits such as SECD score at different sampling times with chip color traits that had lower phenotypic correlation.

Overlapping QTL for Short Storage Duration

QTL for chip color in January and March occupied overlapping regions on chromosomes I, III, and VI. The QTL for L* value in January, March and May, b* value in January and May, and a* value in January had LOD intervals that overlapped on chromosome I and all have estimated positions between 44 and 51 cM. Markers in genes associated with chipping quality traits have been mapped to this region and include genes such as Glucose_6_phosphate isomerase I and II, Glucose_6_phosphate/phosphate translocator, and a neutral invertase (Schreiber et al., 2014). Association panels have also found significant associations on this linkage group for chip color (D'hoop et al., 2014). On chromosome III, the location of QTL peaks for chip color is near that of soluble starch synthase and sucrose synthase. The LOD interval for this set of QTL is broad, observed between 24 and 74 cM positions. QTL for chip color observed on this chromosome have been observed previously on the same chromosome arm (Rak, 2015). In the Wauseon x Lenape population there was a separate QTL interval for a* color in March that maps to 30 cM, which is near several Kunitz-type tuber invertase inhibitors, proteins shown to inhibit soluble invertase *in vitro*. (Glaczinski et al., 2002)(Fischer et al., 2014) (Sharma et al., 2013). QTL for L*, a*, and b* color in January and March occupied the same region on chromosome VI. The LOD intervals for these 6 traits spanned from 56-78 cM and the LOD peaks were estimated to occur at 64, 69, and 71 cM positions. Specific gravity also mapped to this region and while its LOD interval was broader, its position was mapped to 71 cM. Candidate genes near the physical location of the nearest markers include fructokinase, fructokinase (StFrk2), and hexokinase (Schreiber et al., 2014). The QTL in this region had the highest variance explained for chip color related traits in this study.

Long Term Chip Color and Slope Trait QTL

There were distinct sets of clustered QTLs for May chip color traits that occupied the same regions with QTL for changes in chip color, i.e. slope measurements of chip color, on Chromosomes IV, VI, and VII that may be involved with tuber aging. Chromosome IV had QTL for L*, a*, and b* values at May processing co-located with QTL for a* and b* values at March processing. QTL for slope of b* value from January-May mapped to the same region of Chromosome IV as the b* value QTL at 78 cM with a similar LOD profile. QTL for slope of L* from January-May and slope of b* from January-March had LOD intervals similar to that of QTL for a* values in March and May. Previously, AFLP markers and a diversity panel were used to detect an association with April fry color on chromosome IV (D'hoop et al., 2014). In the region of the QTL in this study, candidate genes that might be associated to chip color traits include those for hexose transporters, hexokinases, starch debranching enzyme I, alpha-amylase (AmyZ), and Invertase (INV-4)(Schreiber et al., 2014). Chromosome VII had two QTL with similarly broad LOD intervals for b* value at May processing and slope of L* from March-May. Schreiber et al. 2014 found many genes associated with chip quality traits in this broad region, including several genes for starch synthase, beta-amylase, and sucrose phosphate synthase.

QTL for SECD Traits

There were strong phenotypic correlations between chip color and SECD trait values (Figure 3.4). This relationship could partially be attributed to the methods of trait measurement; the inclusion of chip material with dark SECD patterning with the chip material that was measured by the Hunter D25LT may have affected the Lab values. However, whether or not this made an appreciable contribution to chip color values is debatable given the proportionally low amount of chip material affected by SECD compared to overall chip area (Figure 3.1). It has been found that vacuolar invertase, an enzyme involved in cold-induced sweetening, was up regulated in tissue with high SECD scores (Wang et al., 2012) and here we found QTL for SECD and chip color traits co-localized on chromosomes III and VI. Chromosome VI has QTL for May chipping color, slope of chip color over different month intervals, and also 3 QTL for

SECD in January, March, and May (Figure 3.5). The chip color and SECD trait QTL map to positions 38, 42, 43, and 44 cM. Genes associated with this region include cell wall invertase and hexokinase (Schreiber et al., 2014). QTL on chromosome III for chip color have LOD intervals that overlap with SECD trait QTL although their positions appear to be offset from each other. Interestingly, the closest marker to the estimated SECD QTL at 47 cM is within 2 Mb of INV_Pain-1, a vacuolar invertase as well as InvInh-3/1, an invertase inhibitor. These QTL explain approximately 10% of SECD variance.

QTL for SECD also occurred in regions on chromosomes II and IV that did not contain QTL for other traits. The SECD QTL located on chromosome II maps to 15 cM. The closest marker, c2_30911, is annotated as a conserved gene of unknown function and is located near a high number of genes annotated for function in the electron transport chain and chloroplast (Sharma et al., 2013). This marker is 200 kb away from the nearest marker for overall chip color traits in another population (Rak, 2015) so it is plausible that the allele is involved with both traits. The populations phenotyped by Rak et al (2015) also contain Lenape in their pedigree. Chromosome IV contains a QTL for SECD in May that is positioned away from other chip processing QTL. Mapped to 3 cM, the closest marker is c1_11787 and this marker is located within 1 Mb of a set of cytosolic Fructose-1,6-bisphosphatase genes. This enzyme is thought to play a role in stress response and affects carbohydrate accumulation and photosynthetic activity in transgenic potato (Kossmann et al., 1994)(Sweetlove and Hill, 2000).

QTL for Vine Maturity and Specific Gravity

A vine maturity QTL was found in the well documented position on chromosome V and explained a large proportion of the observed variance for this trait (McCord et al., 2011b)(Bradshaw et al., 2004)(Bradshaw et al., 2004)(Massa et al., 2015). In this study, this vine maturity QTL was located near a QTL for SECD ranking in January which mapped to 38 cM. The closest marker to the LOD peak was within 53 kb of an alpha-amylase gene and 650 kb away from a sucrose transporter (Schreiber et al., 2014). Chip color associations were also found by others in this area of chromosome V (D'hoop et al., 2014). In this study, a set of slope color QTL were found nearby on chromosome V but clustered near 56

cM. This position is located within 200 kb of an invertase inhibitor (Schreiber et al., 2014) and invertase inhibition has been demonstrated to reduce sugar accumulation through storage time. A smaller effect QTL for vine maturity was found on chromosome I in the same region as the co-localized chip color QTL. Small effect QTL for maturity have previously been found on chromosome I in this region (D'Hoop et al., 2008).

Specific gravity was a trait of interest in part due to the sizable increase in tuber solids that Lenape exhibited in relation to other cultivars (Love et al., 1998). In this study, Lenape had greater mean specific gravity than 95% of the measured plots across all years. QTL were detected on chromosome I, III, IV, and VIII. The QTL on chromosome I was mapped to 92 cM but there were no obvious candidate genes in the region. Chromosome I has previously been associated with specific gravity related traits (D'hoop et al., 2014) and the QTL identified here is 200 kb from a QTL for specific gravity found in another tetraploid population (Zorilla Cisneros, 2013). The cultivar Atlantic was one of the parents of that population, and Lenape is a parent of Atlantic. Specific gravity on chromosome III maps within 500 kb of a soluble starch synthase gene (Schreiber et al., 2014) and that QTL might be the same as a QTL detected previously in studies using AFLPs (D'hoop et al., 2014)(McCord et al., 2011b). The largest effect QTL was present on chromosome VI and maps to the same location as chip color traits. This QTL location was not described in other mapping studies. Chromosome VIII contained a QTL that mapped to 68 cM and was 200 kb from an invertase inhibitor gene (Schreiber et al., 2014). There have been QTL for specific gravity and dry matter discovered on this chromosome in previous work that evaluated related material (Zorilla Cisneros, 2013)(McCord et al., 2011a). Interestingly, there is a possible link between higher specific gravity and the presence of a minor allele for the SNP markers closest to the QTLs found on both chromosome III and VI (Figure 3.8). It would be useful to validate this relationship for use in breeding.

Many of the QTL regions identified in this study are starting points for further research. The location of a QTL does not pinpoint the causal locus for trait variation. Schreiber et al. found relationships between traits and SNPs in genes of possible importance but that does not preclude a neighboring gene

from being the causal variant responsible for variation of the measured trait (Gebhardt et al., 2004). In this study, there were several QTL that mapped near these annotated genes, and this provides further evidence that these genes may influence chip quality. However, direct proof of a relationship between a candidate gene allele and a phenotype requires functional analysis. Single marker analysis, which showed significant differences for several traits in this study (Figures 3.7 and 3.8), could be performed in other populations to determine if there is a broad effect on a trait phenotype (Li et al., 2013). The single marker analysis for many of the QTL suggested an effect of the alleles carried by Lenape (Figures 3.7 and 3.8) but this requires further investigation. The haplotype block from Lenape that confers beneficial trait qualities must be identified to use the results presented in a practical breeding program (Simko, 2004). It must be proven that a minor allele from a mapped SNP marker co-segregates with the causal allele. This could be achieved through fine mapping regions of interest, such as the large effect QTL area of chromosome VI. Depending on the trait, the haplotype blocks controlling a trait could also be identified in these regions through bioinformatic approaches.

Improvements in phenotyping might also be made to better separate chip color from SECD traits. This could be achieved by dissecting SECD affected tissue from overall chip material used to measure Lab color with image analysis software routines. This would have the additional benefit of being able to objectively measure SECD (Pedreschi et al., 2006)(Mendoza et al., 2007). There was some correlation between overall chip color and SECD ranking. A computer vision approach could segment the image and grade the severity of defects independent of overall chip color. The specific gravity QTL could be further refined by measuring dry matter in raw tissue by employing the hyperspectral methods described in Chapter 4. This technology would permit rapid measurements of dry matter from many tubers. This could further identify genetic control of this trait.

3.6 Conclusion

Here we present significant QTL for important potato processing traits and the first QTL for SECD. The autotetraploid population was derived from Wauseon x Lenape, an important cross not only

in that it developed some notable varieties, but also that it can be used to explain the allelic contribution of Lenape, a source of significant improvement in chip processing varieties. TetraploidMap 2.0 was used to call dosage on markers, construct the linkage map, and map QTL based on a genotype probabilities model. While there are limitations to this mapping approach, there are many genomic regions that were identified as future areas of interest for cultivar improvement. This is promising for the QTL that were identified for SECD because it is difficult to screen due to its expression based on abiotic stresses. In order to move forward with this research, it would be valuable to determine if the underlying causative alleles for any of the QTL herein are present in other individuals in North American germplasm. GWAS approaches are now feasible in autotetraploid potato to identify alleles that are important on a broad diversity of genotypes. Other ways to focus on candidate genes would be to employ RNA sequencing analysis under different environmental conditions using potatoes with differing phenotypes. Once large effect alleles are identified for these processing traits, it is critical that closely linked markers are identified. Only then can breeding programs begin to use this information for rapid selection of chip processing varieties.

3.7 References

Arnott, R., and G. Smart. 2016. Guidelines for Estimating Potato Production Costs 2016 in Manitoba. Available at https://www.gov.mb.ca/agriculture/business-and-economics/financial-management/pubs/cop_crop_irrigatedpotato.pdf (verified 4 January 2017).

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2014. Fitting Linear Mixed-Effects Models using lme4. J. Stat. Softw. 67(1): 1-48.

Van Berloo, R., R.C.B. Hutten, H.J. Van Eck, and R.G.F. Visser. 2007. An online potato pedigree database resource. Potato Res. 50(1): 45–57.

Bernardo, R. 2010. Breeding for Quantitative Traits in Plants. Second. Stemma Press, Woodbury, Minnesota.

Bethke, P.C., R.P. Sabba, and A.J. Bussan. 2009. Tuber water and pressure potentials decrease and sucrose contents increase in response to moderate drought and heat stress. Am. J. Potato Res. 86(6): 519–532.

Bolotova, Y., and P.E. Patterson. 2009. An analysis of contracts in the idaho processing-potato industry. J. Food Distrib. Res. 40(1): 32–38.

Bradshaw, J.E., B. Pande, G.J. Bryan, C.A. Hackett, K. McLean, H.E. Stewart, and R. Waugh. 2004. Interval mapping of quantitative trait loci for resistance to late blight [*Phytophthora infestans* (Mont.) *de bary*], height and maturity in a tetraploid population of potato (*Solanum tuberosum* subsp. *tuberosum*). Genetics 168(2): 983–995.

Burton, W.G. 1969. The sugar balance in some British potato varieties during storage. II. The effects of tuber age, previous storage temperature, and intermittent refrigeration upon low-temperature sweetening. Eur. Potato J. 12(2): 81–95.

Bussan, A.J., R.P. Sabba, and M.J. Drilias. 2009. Tuber maturation and potato storability : Optimizing skin set, sugars, and solids.

Chen, X., F. Salamini, and C. Gebhardt. 2001. A potato molecular-function map for carbohydrate metabolism and transport. Theor. Appl. Genet. 102(2–3): 284–295.

Colgan, R., and D. Rees. 2012. Senescent Sweetening. UK Potato Counc. 2012(6).

Curtis, K.R., and J.J. McCluskey. 2003. Contract incentives in the processed potato industry.

D'hoop, B.B., P.L.C. Keizer, M.J. Paulo, R.G.F. Visser, F.A. van Eeuwijk, and H.J. van Eck. 2014. Identification of agronomically important QTL in tetraploid potato cultivars using a marker-trait association analysis. Theor. Appl. Genet. 127(3): 731–748.

D'Hoop, B.B., M.J. Paulo, R.A. Mank, H.J. Van Eck, and F.A. Van Eeuwijk. 2008. Association mapping of quality traits in potato (*Solanum tuberosum L.*). Euphytica 161(1–2): 47–60.

Dickman, L. V. 2016. Stem end chipping defect incidence and severity in potatoes (*Solanum tuberosum*). Madison, WI. University of Wisconsin-Madison.

Douches, D.S., K. Jastrzebski, D. Maas, and R.W. Chase. 1996. Assessment of potato breeding over the past century. Crop Sci. 36: 1544–1552.

Eldredge, E.P., Z. a. Holmes, a. R. Mosley, C.C. Shock, and T.D. Stieber. 1996. Effects of transitory

water stress on potato tuber stem-end reducing sugar and fry color. Am. Potato J. 73(11): 517–530.

Fischer, M., M. Kuckenberg, R. Kastilan, J. Muth, and C. Gebhardt. 2014. Novel in vitro inhibitory functions of potato tuber proteinaceous inhibitors. Mol. Genet. Genomics 290(1): 387–398.

Gebhardt, C., A. Ballvora, B. Walkemeier, P. Oberhagemann, and K. Schüler. 2004. Assessing genetic potential in germplasm collections of crop plants by marker-trait association: A case study for potatoes with quantitative variation of resistance to late blight and maturity type. Mol. Breed. 13(1): 93–102.

Glaczinski, H., A. Heibges, F. Salamini, and C. Gebhardt. 2002. Members of the Kunitz-type protease inhibitor gene family of potato inhibit soluble tuber invertase in vitro. Potato Res. 45(2): 163–176.

Habib, A.T., and H.D. Brown. 1957. Role of reducing sugars and amino acids in the browning of potato chips. Food Technol. 11(2): 85–89.

Hackett, C., J.E. Bradshaw, and G.J. Bryan. 2014. QTL mapping in autotetraploids using SNP dosage information. Theor. Appl. Genet. 127(9): 1885–1904.

Hackett, C., J.E. Bradshaw, and J.W. McNicol. 2001. Interval mapping of quantitative trait loci in autotetraploid species. Genetics 159(4): 1819–1832.

Hackett, C., K. McLean, and G.J. Bryan. 2013. Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. PLoS One 8(5): 1-21.

Hamilton, J.P., C.N. Hansey, B.R. Whitty, K. Stoffel, A.N. Massa, A. Van Deynze, W.S. De Jong, D.S. Douches, and C.R. Buell. 2011. Single nucleotide polymorphism discovery in elite North American potato germplasm. BMC Genomics 12(1): 302-313.

Henderson, C.R. 1974. General flexibility of linear model techniques for sire evaluation. J. Dairy Sci. 57(8): 963–972.

Hirsch, C.N., C.D. Hirsch, K. Felcher, J. Coombs, D. Zarka, A. Van Deynze, W. De Jong, R.E. Veilleux, S. Jansky, P. Bethke, D.S. Douches, and C.R. Buell. 2013. Retrospective view of North American potato (*Solanum tuberosum L.*) breeding in the 20th and 21st centuries. G3 (Bethesda). 3(6): 1003–1013.

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding: An Update. Plant Breed. Rev. 22: 9–112.

Iritani, W.M., and L. Weller. 1973. The development of translucent end tubers. Am. Potato J. 50(7): 223–233.

Kleinkopf, G.E. 1976. Translucent-End of Potatoes. Current Information Series No. 488. University of Idaho, College of Agriculture: Cooperative Extension Service.

Kossmann, J., U. Sonnewald, and L. Willmitzer. 1994. Reduction of the chloroplastic fructose-l,6-bisphosphatase in transgenic potato plants impairs photosynthesis and plant growth. Plant J. 6(5): 637–650.

Li, L., E. Tacke, H.-R. Hofferbert, J. Lübeck, J. Strahwald, A.M. Draffehn, B. Walkemeier, and C. Gebhardt. 2013. Validation of candidate gene markers for marker-assisted selection of potato cultivars with improved tuber quality. Theor. Appl. Genet. 126(4): 1039–1052.

Lin, C.S., and G. Poushinsky. 1985. A Modified Augmented Design (Type 2) for Rectangular Plots. Can. J. Plant Sci. 65(I): 743–749.

Love, S., J. Pavek, A. Thompson-Johns, and W. Bohl. 1998. Breeding progress for potato chip quality in North American cultivars. Am. J. Potato Res. 75: 27–36.

Lulai EC, Orr PH. 1979. Influence of potato specific gravity on yield and oil content of chips. Am Potato J. 56:379-390.

Massa, A.N., N.C. Manrique-Carpintero, J.J. Coombs, D.G. Zarka, A.E. Boone, W.W. Kirk, C.A. Hackett, G.J. Bryan, and D.S. Douches. 2015. Genetic Linkage Mapping of Economically Important Traits in Cultivated Tetraploid Potato (*Solanum tuberosum L.*). G3 Genes|Genomes|Genetics 5(11): 2357–2364.

McCord, P.H., B.R. Sosinski, K.G. Haynes, M.E. Clough, and G. Craig Yencho. 2011a. Linkage mapping and QTL analysis of agronomic traits in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*). Crop Sci. 51(2): 771–785.

McCord, P.H., B.R. Sosinski, K.G. Haynes, M.E. Clough, and G.C. Yencho. 2011b. QTL mapping of internal heat necrosis in tetraploid potato. Theor. Appl. Genet. 122(1): 129–142.

Mendoza, F., P. Dejmek, and J.M. Aguilera. 2007. Colour and image texture analysis in classification of commercial potato chips. Food Res. Int. 40(9): 1146–1154.

Menendez, C.M., E. Ritter, R. Schafer-Pregl, B. Walkeneier, A. Kalde, F. Salamini, and C. Gebhardt. 2002. Cold-sweetening in diploid potato : mapping QTL and candidate genes Cold Sweetening in Diploid Potato : Mapping Quantitative Trait Loci and Candidate Genes. Genetics 162(November): 1423–1434.

Milbourne, D., J.E. Bradshaw, and C. Hackett. 2009. Molecular mapping and breeding in polyploid crop plants. p. 355–394. *In* Principles and Practices of Plant Genomics 2.

National Potato Council. 2015. Potato Statistical Yearbook 2015.

Nettleton, D., and R.W. Doerge. 2000. Accounting for variability in the use of permutation testing to detect quantitative trait loci. Biometrics 56(1): 52–58.

Patterson, P.E. 2007. 2007 Cost of Potato Production Comparison for Idaho Commercial Potato Production. University of Idaho. Deapartment of Ag. Econ. And Rural Soc. Reports.

Pedreschi, F., J. León, D. Mery, and P. Moyano. 2006. Development of a computer vision system to measure the color of potato chips. Food Res. Int. 39(10): 1092–1098.

R Core Team. 2014. R: A language and environment for statistical computing. R Found. Stat. Comput. Vienna, Austria (January): 1–3Available at https://www.r-project.org/.

Rak, K. 2015. Breeding and Molecular Genetics for the Improvement of Cold Storage Potato Chip Quality. University of Wisconsin-Madison.

Robinson, G.K. 1991. That BLUP is a good thing: The estimation of random effects. Stat. Sci. 6: 15–32.

Scanlon, M.G., R. Roller, G. Mazza, and M.K. Pritchard. 1994. Computerized video image analysis to quantify color of potato chips. Am. Potato J. 71(11): 717–733.

Schreiber, L., a. C. Nader-Nieto, E.M. Schonhals, B. Walkemeier, and C. Gebhardt. 2014. SNPs in genes functional in starch-sugar interconversion associate with natural variation of tuber starch and sugar content of potato (*Solanum tuberosum L.*). G3 Genes|Genomes|Genetics 4(10): 1797–1811.

Shallenberger, R., O. Smith, and R. Treadway. 1959. Food Color Changes - Role of the sugars in the browning reaction in potato chips. J. Agric. Food Chem. 7(4): 274–277.

Sharma, S.K., D. Bolser, J. de Boer, M. Sønderkær, W. Amoros, M.F. Carboni, J.M. D'Ambrosio, G. de la Cruz, A. Di Genova, D.S. Douches, M. Eguiluz, X. Guo, F. Guzman, C. Hackett, J.P. Hamilton, G. Li, Y. Li, R. Lozano, A. Maass, D. Marshall, D. Martinez, K. McLean, N. Mejía, L. Milne, S. Munive, I. Nagy, O. Ponce, M. Ramirez, R. Simon, S.J. Thomson, Y. Torres, R. Waugh, Z. Zhang, S. Huang, R.G.F. Visser, C.W.B. Bachem, B. Sagredo, S.E. Feingold, G. Orjeda, R.E. Veilleux, M. Bonierbale, J.M.E. Jacobs, D. Milbourne, D.M.A. Martin, and G.J. Bryan. 2013. Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps. G3 (Bethesda). 3(11): 2031–47.

Shock, C.C., Z.A. Holmes, T.D. Stieber, E.E. Eldredge, and P. Zhang. 1993. The effect of timed water stress on quality, total solids, and reducing sugar content of potatoes. Am. Potato J. 70: 227–241.

Simko, I. 2004. One potato, two potato: Haplotype association mapping in autotetraploids. Trends Plant Sci. 9(9): 441–448.

Sowokinos, J.R. 2001. Biochemical and molecular control of cold-induced sweetening in potatoes. Am. J. Potato Res. 78(3): 221–236.

Sweetlove, L.J., and S.A. Hill. 2000. Source metabolism dominates the control of source to sink carbon flux in tuberizing potato plants throughout the diurnal cycle and under a range of environmental conditions. Plant, Cell Environ. 23(5): 523–529.

Sweetman, M.D. 1930. Color of potato chips as influenced by storage temperatures of the tubers and other factors. J. Agric. Res. 41: 479–490.

Thompson-Johns, A. 1998. Inheritance of the sugar end disorder in potato (*Solanum tuberosum L.*). University of Idaho.

Tkalčič, M., and J.F. Tasič. 2003. Colour spaces - Perceptual, historical and applicational background. IEEE Reg. 8 EUROCON 2003 Comput. as a Tool - Proc. A: 304–308.

Wang, Y., P.C. Bethke, M.J. Drilias, W.G. Schmitt, and A.J. Bussan. 2015. A multi-year survey of stem-end chip defect in chipping potatoes (*Solanum tuberosum L.*). Am. J. Potato Res. 92(1): 79–90.

Wang, Y., A.J. Bussan, and P.C. Bethke. 2012. Stem-end defect in chipping potatoes (*Solanum tuberosum L.*) as Influenced by Mild Environmental Stresses. Am. J. Potato Res. 89(5): 392–399.

Akeley RV, Mills WR, Cunningham CE, Watts J. 1968. Lenape: A new potato variety high in solids and chipping quality. Am Potato J. 45:142–145.

Zorilla Cisneros, C. 2013. Understanding the Genetics of Potato Tuber Calcium and its Implications in Breeding for Improved Quality. University of Wisconsin-Madion. Ann Arbor, MI: ProQuest LLC.

Table 3.1 Broad sense heritability estimates on an entry-mean basis for select processing traits. The years used in each analysis are listed under the Storage Years column. Storage years refers to the calendar year following the growing season. Abbreviation is the trait name in Table 3.2 and Figure 3.5.

Trait	Abbreviation	Heritability	Storage Years
Specific Gravity	spgrvf	0.88	2013, 2014, 2015, 2016
a* value (March)	a_marf	0.88	2013, 2014, 2015, 2016
L* value (May)	L_mayf	0.88	2014, 2015, 2016
L* value (March)	L_marf	0.87	2013, 2014, 2015, 2016
b* value (May)	b_mayf	0.86	2014, 2015, 2016
a* value (January)	a_janf	0.85	2013, 2014, 2015, 2016
a* value (May)	a_mayf	0.83	2014, 2015, 2016
L* value (January)	L_janf	0.82	2013, 2014, 2015, 2016
maturity	matf	0.79	2014, 2015, 2016
b* value (March)	b_marf	0.76	2013, 2014, 2015, 2016
Slope of b* (January - May)	sbjyf	0.74	2014, 2015, 2016
b* value (January)	b_janf	0.73	2013, 2014, 2015, 2016
SECD ranking (March)	sedrf	0.72	2013, 2014, 2015, 2016
Slope of L* (January - May)	sLjyf	0.71	2014, 2015, 2016
Slope of a* (January - May)	sajyf	0.66	2014, 2015, 2016
SECD ranking (May)	sedyf	0.65	2014, 2015, 2016
SECD ranking (January)	sedjf	0.64	2013, 2014, 2015, 2016
Slope of L* (March - May)	sLryf	0.52	2014, 2015, 2016
Slope of a* (March - May)	saryf	0.49	2014, 2015, 2016
Slope of b* (March - May)	sbryf	0.48	2014, 2015, 2016
Slope of L* (January - March)	sLjrf	0.44	2013, 2014, 2015, 2016
Slope of b* (January - March)	sbjrf	0.37	2013, 2014, 2015, 2016
Slope of a* (January - March)	sajrf	0.32	2013, 2014, 2015, 2016
Slope of SECD (January - May)	ssdjyf	0.13	2014, 2015, 2016

Table 3.2. Variance components of chip processing traits. σ_g^2, σ_{gy}^2, σ_y^2, σ_ε^2, and h^2 for genotypes, genotypes x years, residual variance, and broad sense heritability on an entry mean basis

Trait	Abbreviation	σ_g^2	σ_y^2	σ_{gy}^2	σ_ε^2	h^2
Specific Gravity	spgrvf	0.000042	0.000013	0.000007	0.000016	0.88
a* value (March)	a_marf	2.125	2.129	0.642	0.529	0.88
L* value (May)	L_mayf	22.870	6.557	3.045	6.627	0.88
L* value (March)	L_marf	15.230	12.820	5.395	3.402	0.87
b* value (May)	b_mayf	3.992	3.443	1.014	0.948	0.86
a* value (January)	a_janf	1.556	2.819	0.408	0.684	0.85
a* value (May)	a_mayf	2.074	1.149	0.182	1.101	0.83
L* value (January)	L_janf	9.565	11.940	2.301	6.039	0.82
Maturity	matf	0.386	0.036	0.250	0.063	0.79
b* value (March)	b_marf	1.752	1.952	1.777	0.435	0.76
Slope of b* (January - May)	sbjyf	0.239	0.025	0.077	0.178	0.74
b* value (January)	b_janf	0.857	1.866	0.454	0.800	0.73
SECD ranking (March)	sedrf	0.260	0.067	0.045	0.366	0.72
Slope of L* (January - May)	sLjyf	1.257	0.617	0.121	1.439	0.71
Slope of a* (January - May)	sajyf	0.133	0.085	0.035	0.175	0.66
SECD ranking (May)	sedyf	0.312	0.076	0.000	0.505	0.65
SECD ranking (January)	sedjf	0.217	0.011	0.158	0.333	0.64
Slope of L* (March - May)	sLryf	1.233	1.179	1.236	2.137	0.52
Slope of a* (March - May)	saryf	0.141	0.072	0.085	0.353	0.49
Slope of b* (March - May)	sbryf	0.259	0.093	0.516	0.338	0.48
Slope of L* (January - March)	sLjrf	0.520	0.392	0.150	2.490	0.44
Slope of b* (January - March)	sbjrf	0.072	0.111	0.152	0.333	0.37
Slope of a* (January - March)	sajrf	0.051	0.015	0.142	0.288	0.32
Slope of SECD (January - May)	ssdjyf	0.004	0.003	0.033	0.058	0.13

Table 3.3. Estimated QTL locations, LOD scores, and effect sizes for BLUPs of chip color, SECD, and associated processing traits from genotype probabilities complete additive model in Wauseon x Lenape population

Trait code[a]	Chrom, Position (cM) and 1.5 LOD support interval	LOD	LOD Threshold 95% C.I.[b]	R^2	Closest Marker and Physical Position[c]	Flanking Markers and Physical Position[d,c]
a_janf	Chr01 44 (35-62)	4.84	(3.35-4.01)	6.95	c2_20803 (60987055)	c2_14490,c2_54355 (65782289,13740819)
a_janf	Chr03 63 (24-75)	4.86	(3.35-4.01)	7.35	c2_55281 (45670613)	c2_29628,c1_15783 (47155647,1276863)
a_janf	Chr06 69 (63-78)	7.95	(3.35-4.01)	14.02	c2_8662 (53053305)	c2_9243,c2_8707 (57959589,52063192)
a_marf	Chr03 30 (1-36)	4.70	(3.39-4.03)	7.55	c2_17218 (1954563)	c1_8968,c1_5388 (5430430,3161621)
a_marf	Chr04 78 (56-81)	4.54	(3.39-4.03)	6.85	c2_34019 (70222194)	c1_10168,c2_26742 (70185119,66204783)
a_marf	Chr06 71 (63-73)	8.36	(3.39-4.03)	14.50	c2_24229 (54218082)	c1_6997,c2_8707 (53678569,52063192)
a_mayf	Chr04 78 (57-84)	5.33	(3.4-4.04)	8.33	c2_34019 (70222194)	c2_35944,c2_32042 (70878898,67130719)
a_mayf	Chr06 42 (39-56)	8.43	(3.4-4.04)	15.11	c2_51758 (38212118)	c2_16863,c2_39576 (47604700,43919744)
b_janf	Chr01 51 (35-78)	4.55	(3.33-3.97)	6.00	c1_14633 (57931888)	c2_17530,c2_54355 (69918804,13740819)
b_janf	Chr03 63 (24-73)	5.37	(3.33-3.97)	8.43	c2_55281 (45670613)	c2_11211,c1_15783 (46528226,1276863)
b_janf	Chr06 64 (56-71)	6.92	(3.33-3.97)	12.43	c2_41412 (51166697)	c2_24229,c2_16863 (54218082,47604700)
b_marf	Chr03 62 (49-73)	5.39	(3.32-3.97)	8.55	c1_6405 (45305706)	c2_11211,c1_10737 (46528226,16887518)
b_marf	Chr04 78 (68-82)	4.76	(3.32-3.97)	7.29	c2_34019 (70222194)	c1_10711,c1_10169 (70954731,70186406)
b_marf	Chr06 69 (62-72)	6.05	(3.32-3.97)	10.53	c2_8662 (53053305)	c1_6997,c1_2091 (53678569,50658278)
b_mayf	Chr01 45 (38-54)	5.24	(3.42-4.15)	7.64	c2_20803 (60987055)	c2_35518,c2_41438 (60512580,52487732)
b_mayf	Chr04 78 (69-81)	6.45	(3.42-4.15)	10.78	c2_34019 (70222194)	c1_10168,c2_12938 (70185119,69206404)
b_mayf	Chr06 56 (34-59)	8.63	(3.42-4.15)	14.99	c2_16863 (47604700)	c2_31144,c2_40236 (48879137,40611306)

(Table 3.3 cont.)

b_mayf	Chr07 65 (38-88)	4.39	(3.42-4.15)	6.15	c2_26013 (50152841)	c2_4578,c1_9215 (13637552,54277246)
L_janf	Chr01 51 (35-75)	4.86	(3.33-3.99)	6.48	c1_14633 (57931888)	c2_17531,c2_54355 (70016241,13740819)
L_janf	Chr03 63 (51-69)	6.69	(3.33-3.99)	11.28	c2_55281 (45670613)	c1_526,c2_48359 (46395963,39259070)
L_janf	Chr06 64 (62-74)	8.23	(3.33-3.99)	14.84	c2_41412 (51166697)	c1_6997,c1_2091 (53678569,50658278)
L_janf	Chr11 11 (1-20)	4.46	(3.33-3.99)	5.80	c2_37194 (1771819)	c1_2280,c2_13350 (3948374,437208)
L_marf	Chr01 44 (34-59)	5.04	(3.35-3.97)	7.04	c2_20803 (60987055)	c2_32367,c2_54355 (63183063,13740819)
L_marf	Chr03 62 (23-74)	4.83	(3.35-3.97)	7.49	c1_6405 (45305706)	c1_9025,c1_12821 (46923611,859369)
L_marf	Chr06 69 (66-72)	8.84	(3.35-3.97)	16.26	c2_8662 (53053305)	c1_6997,c2_8904 (53678569,52599348)
L_mayf	Chr01 45 (36-54)	5.11	(3.43-4.13)	7.36	c2_20803 (60987055)	c2_35518,c2_32112 (60512580,42197568)
L_mayf	Chr04 78 (68-81)	6.36	(3.43-4.13)	10.60	c2_34019 (70222194)	c1_10168,c1_10169 (70185119,70186406)
L_mayf	Chr06 43 (34-59)	9.08	(3.43-4.13)	15.95	c2_57989 (44972343)	c2_31144,c2_40236 (48879137,40611306)
matf	Chr01 53 (42-66)	4.34	(3.38-4.12)	6.63	c2_35518 (60512580)	c2_14608,c2_20803 (66349597,60987055)
matf	Chr05 32 (29-34)	25.46	(3.38-4.12)	39.06	c1_14802 (5051766)	c1_14802,c2_51478 (5051766,5573355)
sajrf	Chr10 47 (22-51)	5.07	(3.39-4.05)	8.32	c2_27827 (50697563)	c2_45686,c2_19225 (51360322,40828602)
sajyf	Chr10 34 (25-38)	5.93	(3.4-3.96)	8.89	c2_40757 (47300568)	c1_11991,c2_48917 (47398435,45320802)
sajyf	Chr12 84 (80-99)	5.59	(3.4-3.96)	9.25	c1_2724 (57460060)	c2_48477,c1_2690 (54723235,58047558)
saryf	Chr09 89 (85-91)	4.74	(3.36-4.05)	7.09	c2_20588 (56940969)	c2_26979,c1_11866 (53091476,57422065)
saryf	Chr12 84 (67-103)	4.24	(3.36-4.05)	6.19	c1_2724 (57460060)	c2_49683,c1_2690 (48605748,58047558)
sbjrf	Chr04 63 (56-85)	4.92	(3.25-3.87)	8.03	c1_10425 (67985135)	c2_10799,c2_26742 (71383530,66204783)
sbjyf	Chr04 78 (67-82)	5.69	(3.46-4.12)	9.31	c2_34019 (70222194)	c1_10711,c2_12976 (70954731,69050482)

(Table 3.3 cont.)

sbjyf	Chr05 56 (36-58)	4.70	(3.46-4.12)	6.26	c1_14913 (44713672)	c2_51478,c2_50532 (5573355,29043101)
sbjyf	Chr06 42 (34-58)	7.03	(3.46-4.12)	11.61	c2_51758 (38212118)	c1_15371,c2_40236 (48416539,40611306)
sbryf	Chr06 38 (34-59)	5.76	(3.35-4.06)	9.32	c2_56056 (38291760)	c2_31144,c2_40236 (48879137,40611306)
sedjf	Chr02 15 (11-18)	4.95	(3.4-4.17)	7.25	c2_30911 (16229386)	c2_32253,c1_9120 (13697423,18883765)
sedjf	Chr03 47 (42-47)	7.51	(3.4-4.17)	12.87	c2_49970 (37121012)	c2_49970,c2_35358 (37121012,35991675)
sedjf	Chr05 38 (32-46)	5.88	(3.4-4.17)	8.73	c2_47607 (5972364)	c1_14802,c2_51915 (5051766,10714562)
sedjf	Chr06 42 (30-47)	5.14	(3.4-4.17)	8.66	c2_51758 (38212118)	c2_31887,c2_51765 (46390332,38211470)
sedrf	Chr03 47 (42-65)	6.04	(3.38-4.08)	9.71	c2_49970 (37121012)	c1_6411,c2_35358 (45345505,35991675)
sedrf	Chr06 44 (31-61)	4.38	(3.38-4.08)	6.79	c2_36123 (45245848)	c1_3013,c1_15233 (51930201,38333228)
sedyf	Chr03 47 (41-54)	6.30	(3.43-4.2)	10.58	c2_49970 (37121012)	c2_57349,c2_56611 (41780626,34073937)
sedyf	Chr04 3 (0-10)	4.33	(3.43-4.2)	5.23	c1_11787 (50918691)	c2_26792,c2_51639 (9523209,5488735)
sedyf	Chr06 43 (28-50)	4.47	(3.43-4.2)	7.17	c2_57989 (44972343)	c2_25947,c1_9846 (46145916,32103689)
sLjrf	Chr05 56 (54-59)	5.15	(3.36-4.09)	7.49	c1_14913 (44713672)	c2_53307,c1_15314 (13386684,40466421)
sLjyf	Chr04 73 (56-85)	5.03	(3.41-4.02)	7.51	c1_3450 (71076408)	c2_10799,c2_26742 (71383530,66204783)
sLjyf	Chr05 56 (45-58)	5.83	(3.41-4.02)	8.33	c1_14913 (44713672)	c1_15292,c2_50532 (8795719,29043101)
sLjyf	Chr06 42 (34-58)	4.93	(3.41-4.02)	7.29	c2_51758 (38212118)	c1_15371,c2_40236 (48416539,40611306)
sLryf	Chr05 51 (44-58)	5.27	(3.36-4.1)	7.77	c2_38456 (12209043)	c2_38167,c2_50532 (6625290,29043101)
sLryf	Chr06 42 (34-59)	5.37	(3.36-4.1)	8.52	c2_51758 (38212118)	c2_31144,c2_40236 (48879137,40611306)
sLryf	Chr07 49 (44-87)	4.57	(3.36-4.1)	6.86	c1_12976 (44649580)	c2_23394,c2_30427 (41846151,54329782)
spgrvf	Chr01 92 (87-95)	4.42	(3.33-3.95)	6.69	c2_2571 (77596089)	c2_50485,c2_14366 (78756913,74688218)

(Table 3.3 cont.)

spgrvf	Chr03 62 (42-73)	4.38	(3.33-3.95)	6.70	c1_6405 (45305706)	c2_11211,c2_35358 (46528226,35991675)
spgrvf	Chr06 71 (56-85)	6.07	(3.33-3.95)	10.06	c2_24229 (54218082)	c2_9039,c2_16863 (57643386,47604700)
spgrvf	Chr08 68 (54-73)	4.83	(3.33-3.95)	6.29	c2_34710 (50368189)	c2_18872,c2_49600 (42460813,49814298)

[a] Trait abbreviations are listed in Table 3.1. Traits are from overall genotype BLUPs, [b] 95% confidence intervals calculated from 200 permutations (Nettlton and Doerge, 2000), [c] Physical positions are from the PGSC DM v4.03 sequenced genome (Sharma et al., 2013), [d] Flanking markers were at each end of a 1.5 LOD support interval

Figure 3.1. Rating scale for stem-end chip defect (SECD) as developed by (Wang et al., 2012)

Figure 3.2. Trait distribution patterns for chip color values in Lab color space for Wauseon x Lenape Population grown during 2012, 2013, 2014, and 2015 field seasons across 3 processing times. Values are absent for late storage processing times in 2012 field season material.

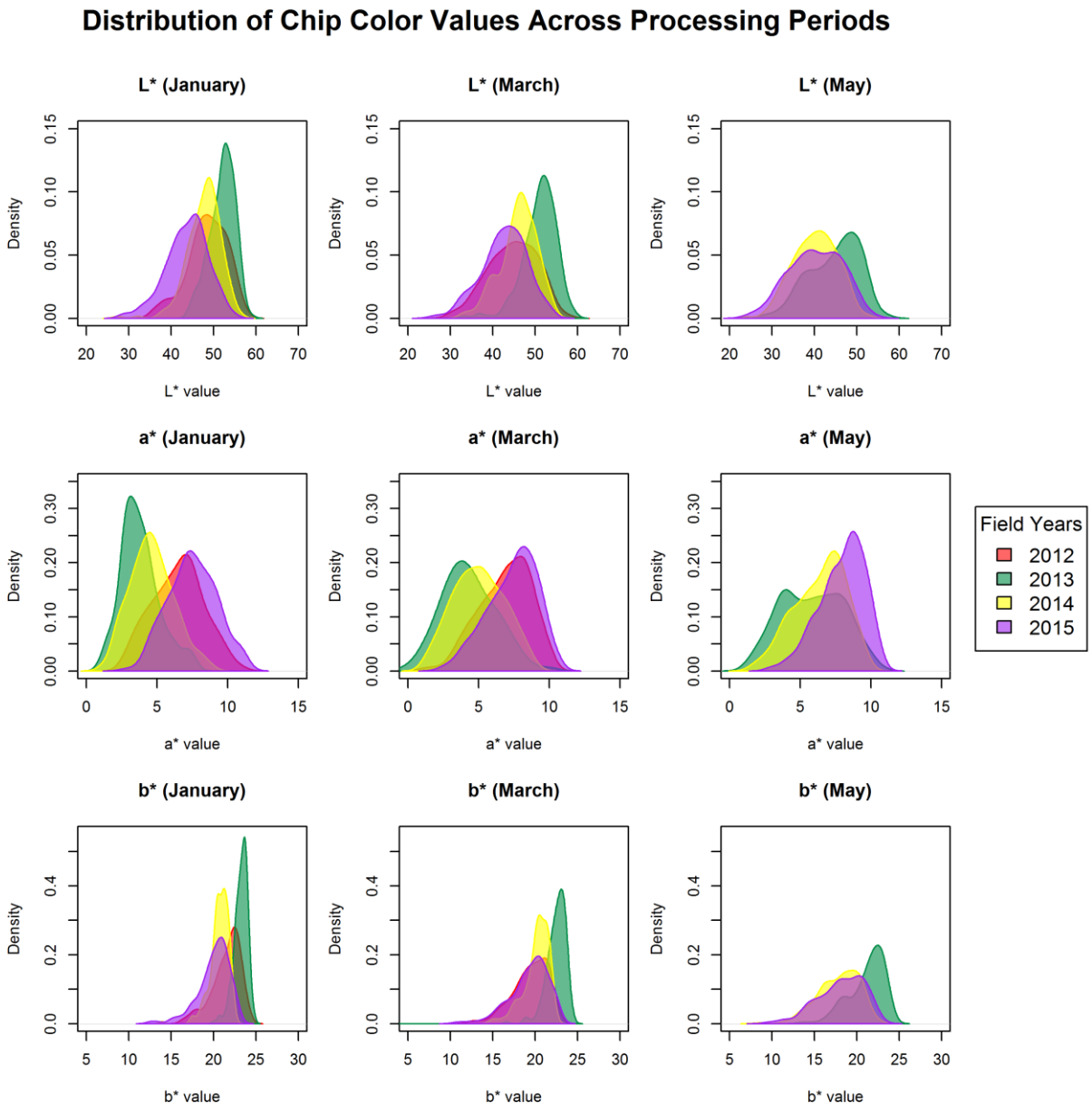


Figure 3.3. Trait distribution patterns for genotype mean January SECD score across 2012, 2013, 2014, and 2015 field years. Vertical lines represent mean SECD score for parents and full-sib population.

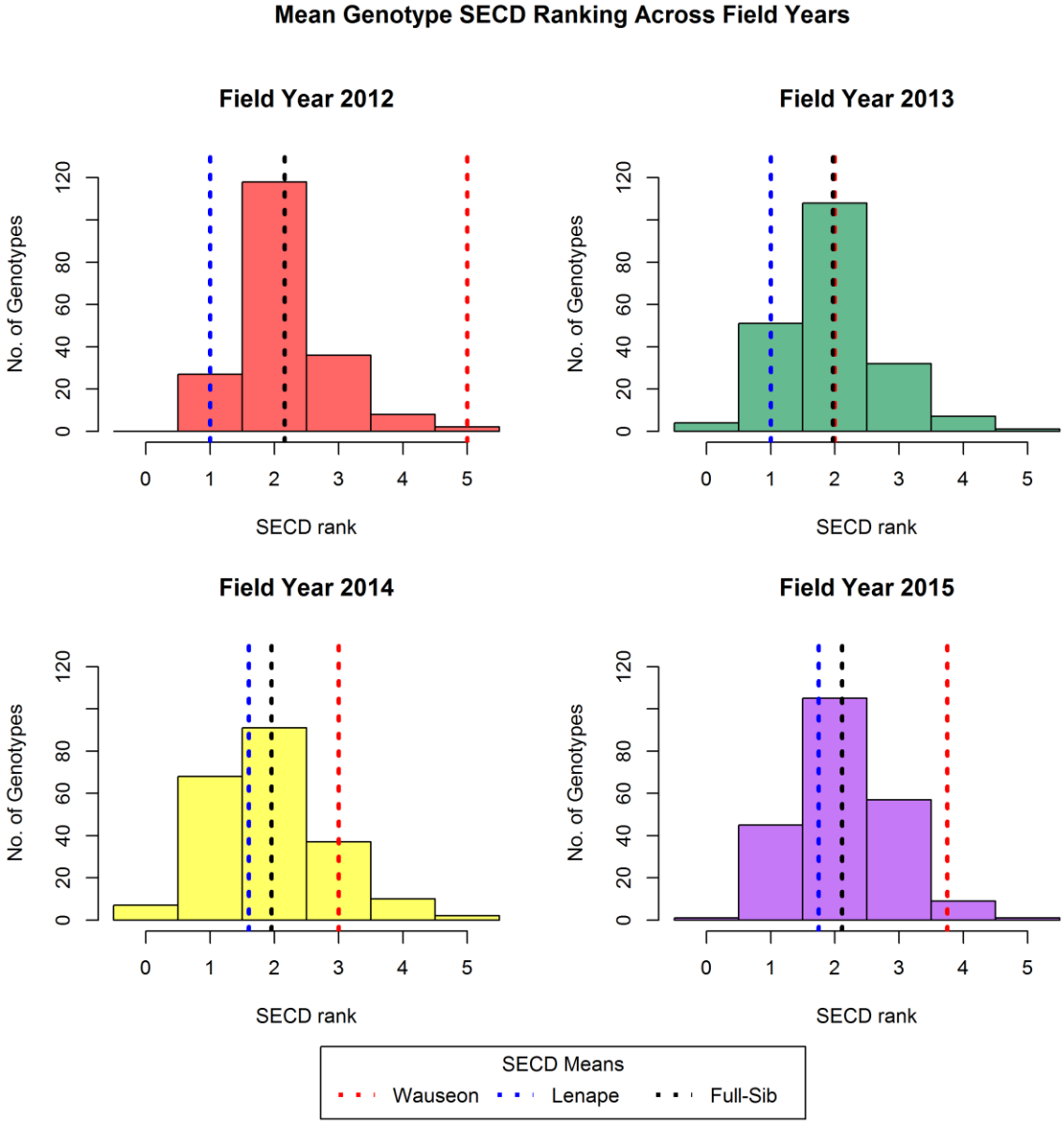
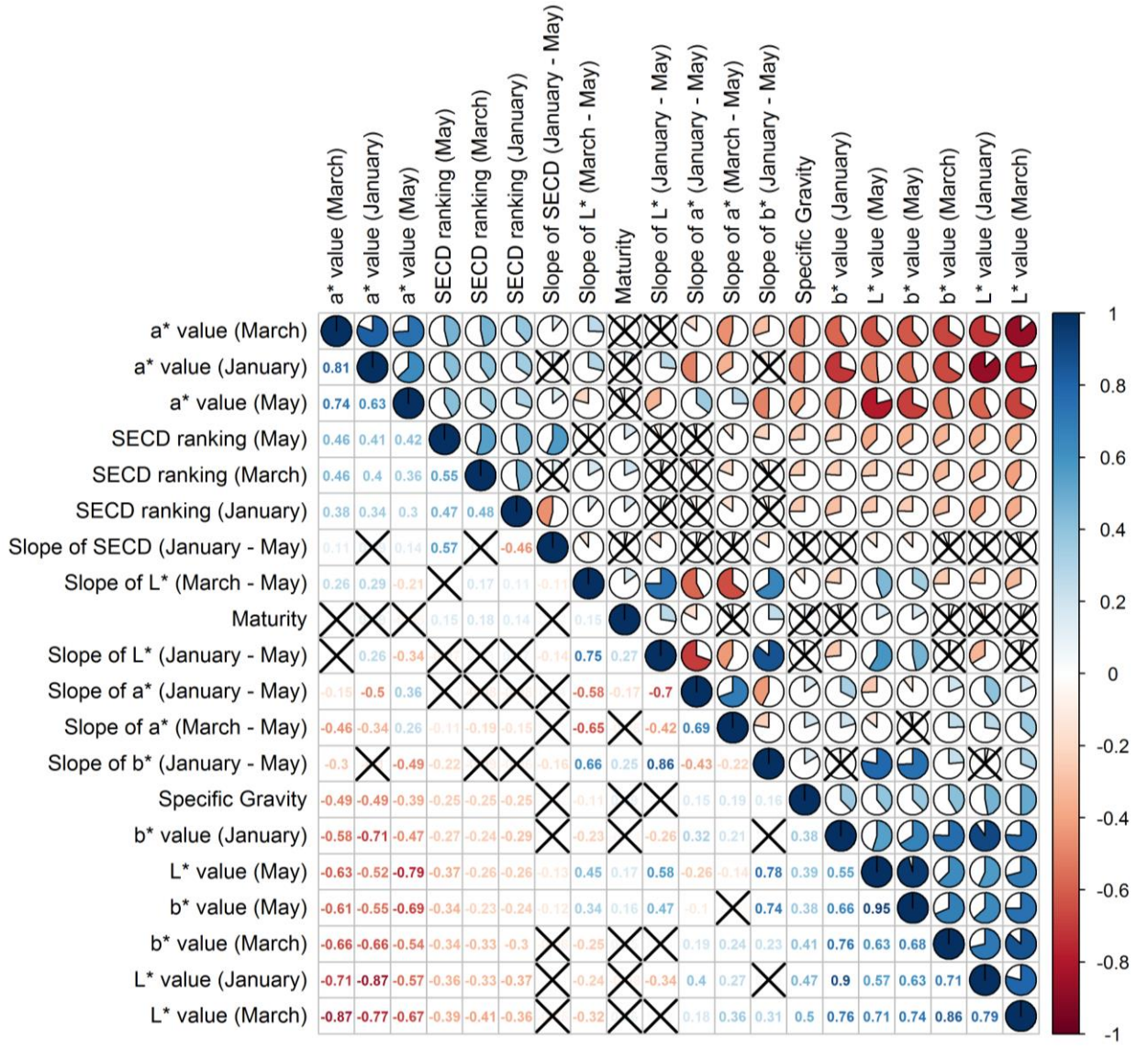


Figure 3.4. Correlation plot of chip processing traits from chip color analysis in the Lab color space, SECD rating, maturity rating, and specific gravity over 4 years of data within a processing potato population of 192 individuals. Pearson's correlation coefficients are indicated numerically in one half of the plot and graphically in the other half.



(An "X" denotes a non-significant correlation)

Figure 3.5. Genetic map of Wauseon x Lenape population with QTL for chip processing traits and their 1.5 LOD support intervals. 12 Linkage groups with markers binned every 1 cM and distance in cM. Trait abbreviations are the same as those in Table 3.1. Mapped QTL positions are indicated by vertical lines to the right of each chromosome. The length of the line indicates the 1.5 LOD support interval with the center tick indicating the mapped position. Colors of the QTL lines and text indicate the trait type: chip color is mustard, SECD ranking traits are blue, slope of chip color through storage is bright red, specific gravity is maroon, and maturity is green.

Figure 3.5 (cont.)

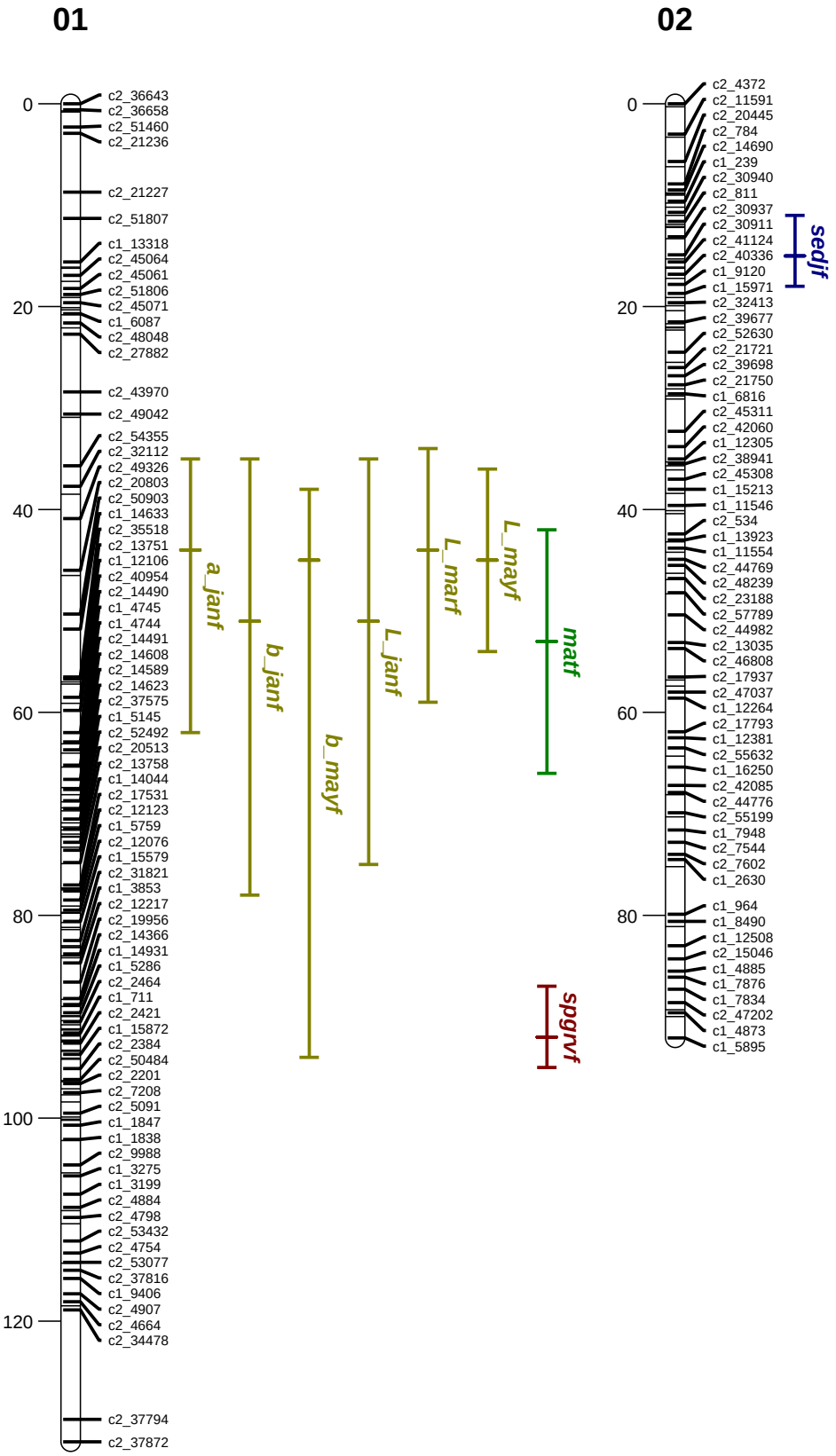


Figure 3.5 (cont.)

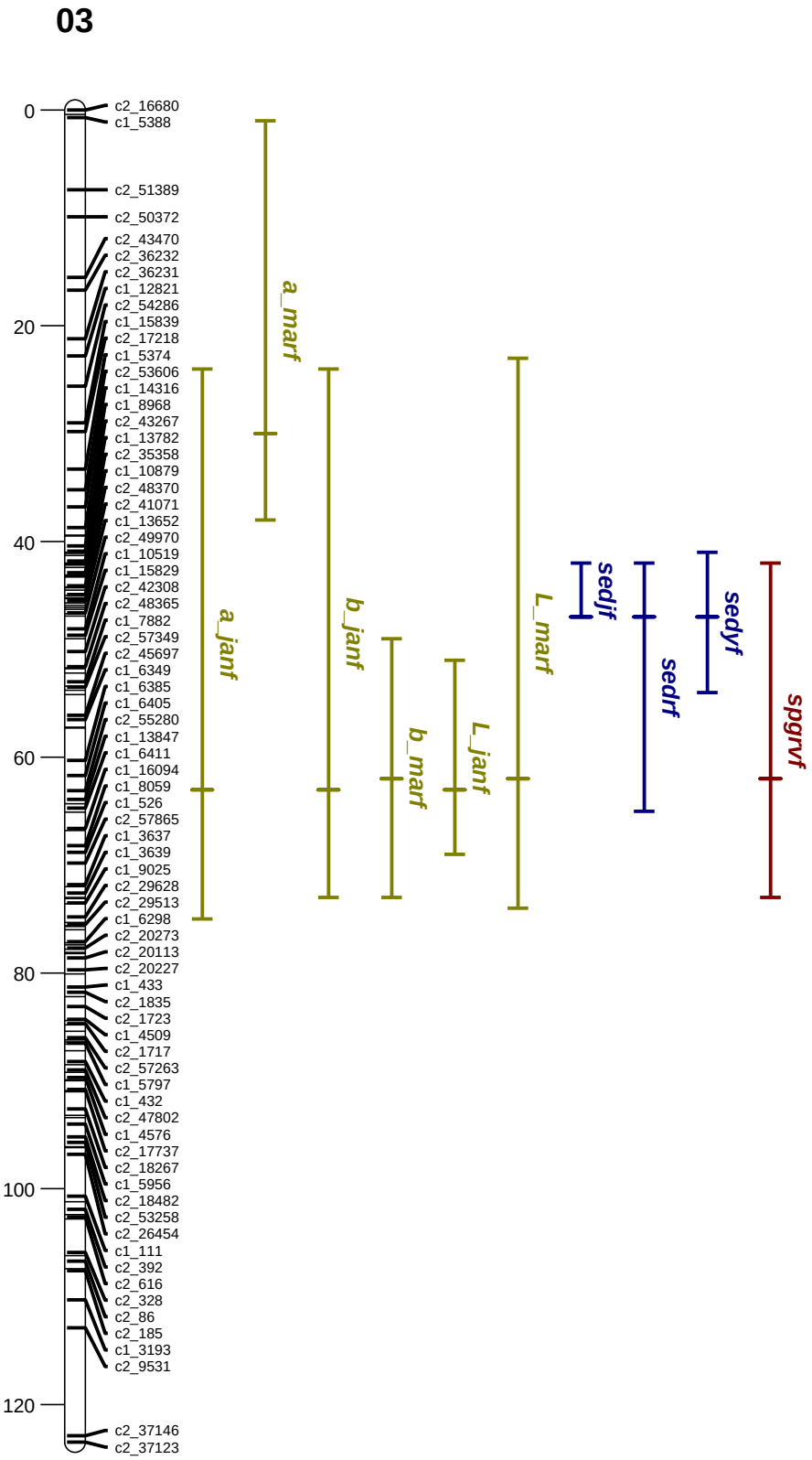


Figure 3.5 (cont.)

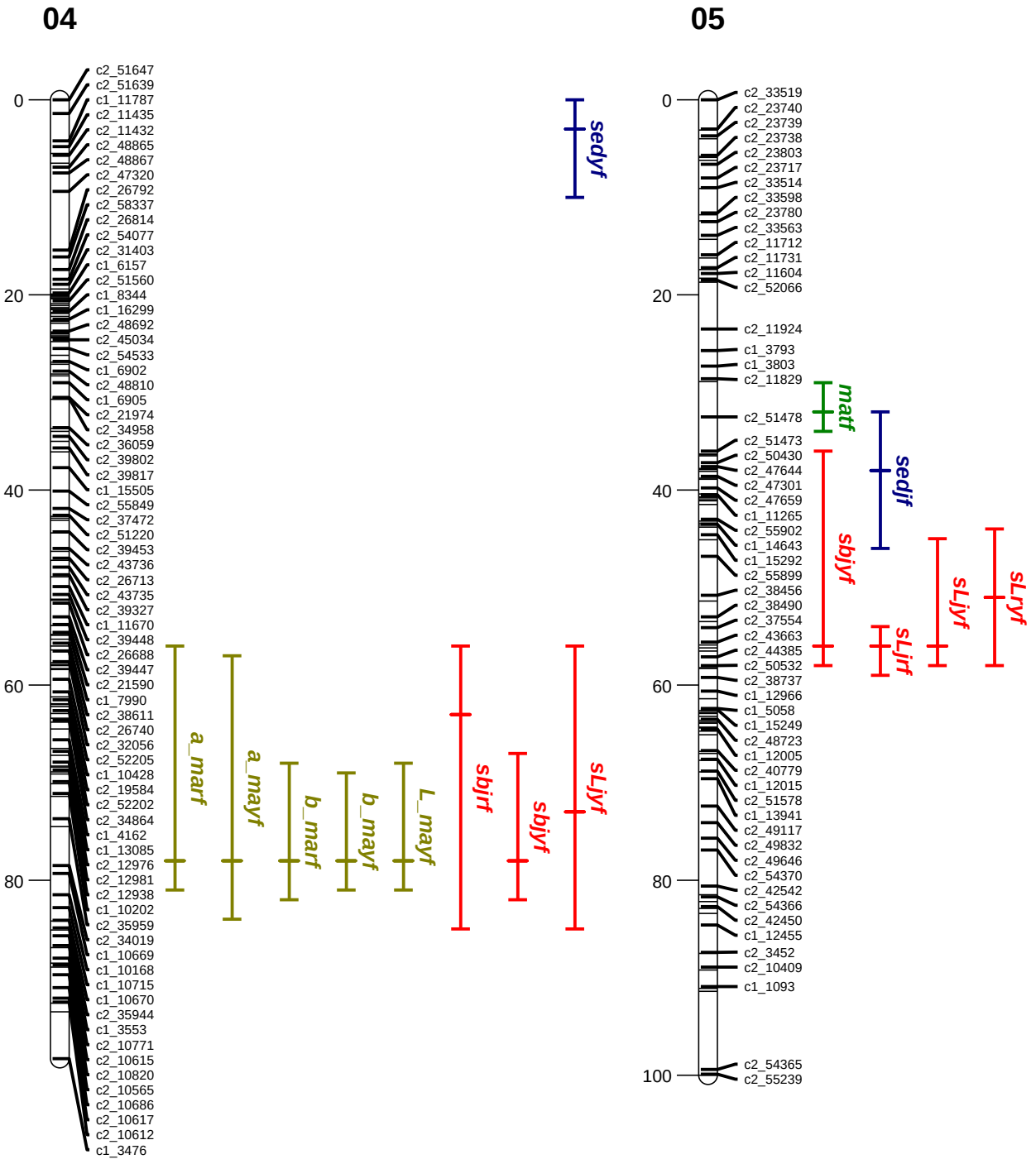


Figure 3.5 (cont.)

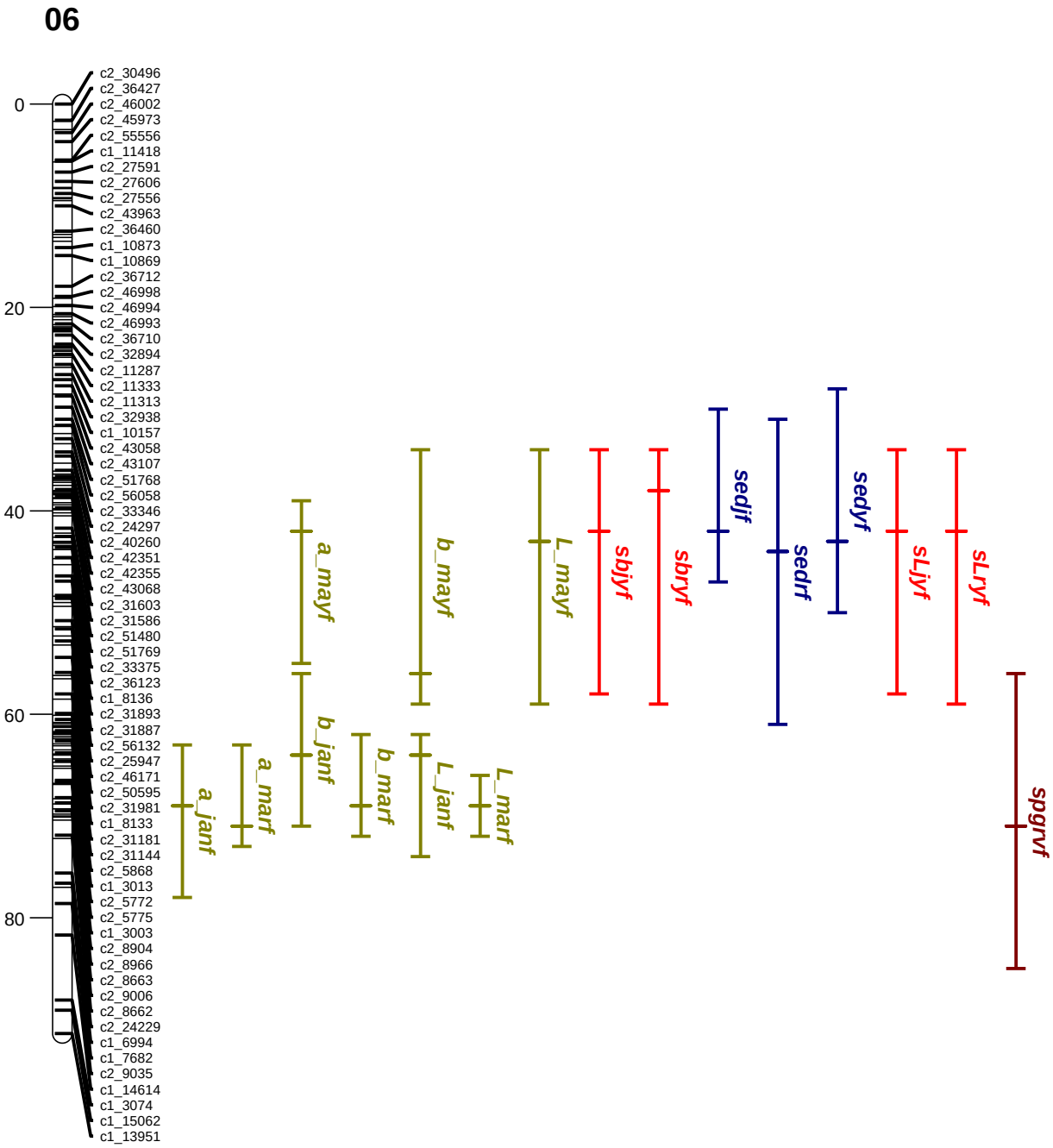


Figure 3.5 (cont.)

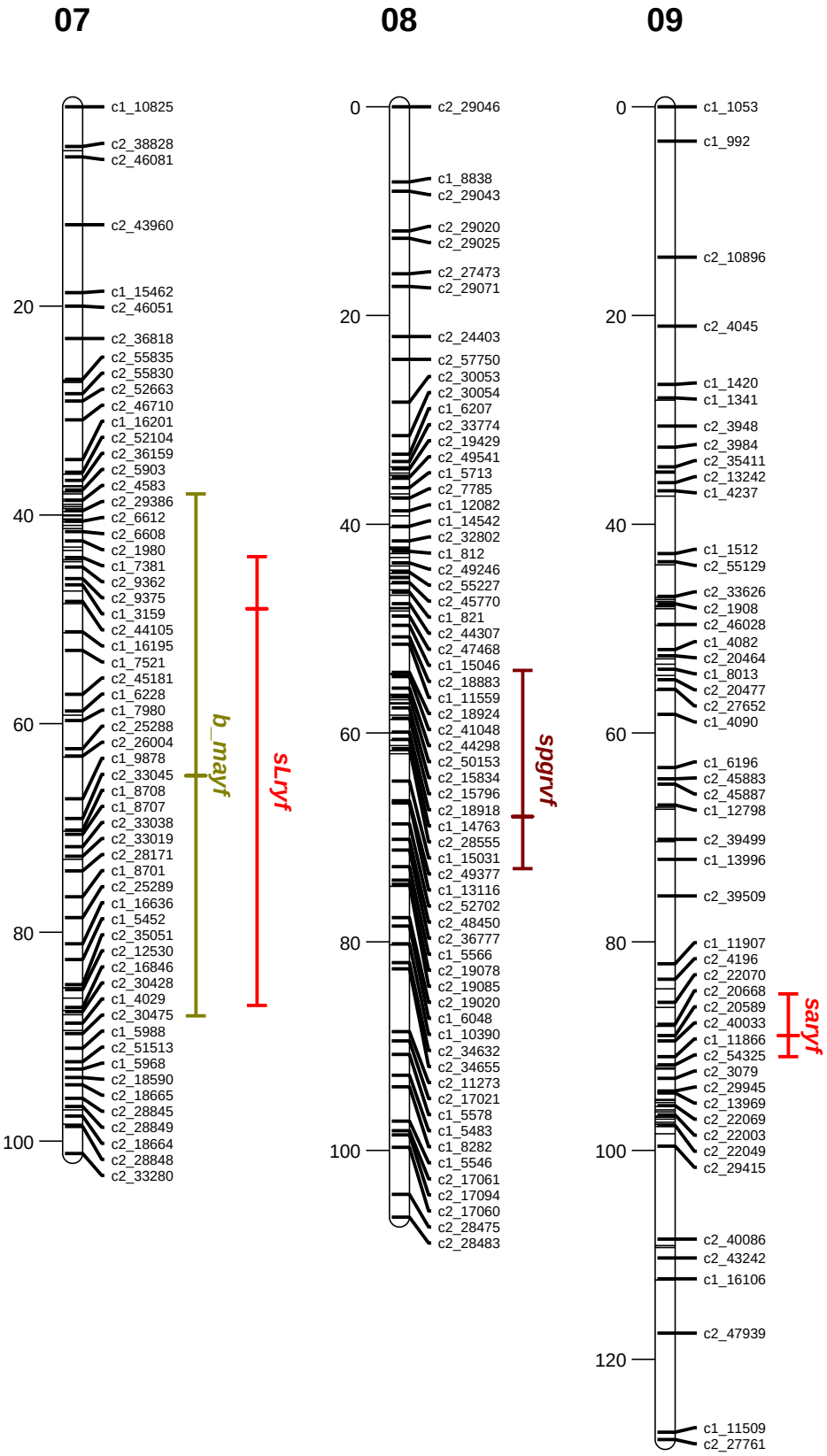


Figure 3.5 (cont.)

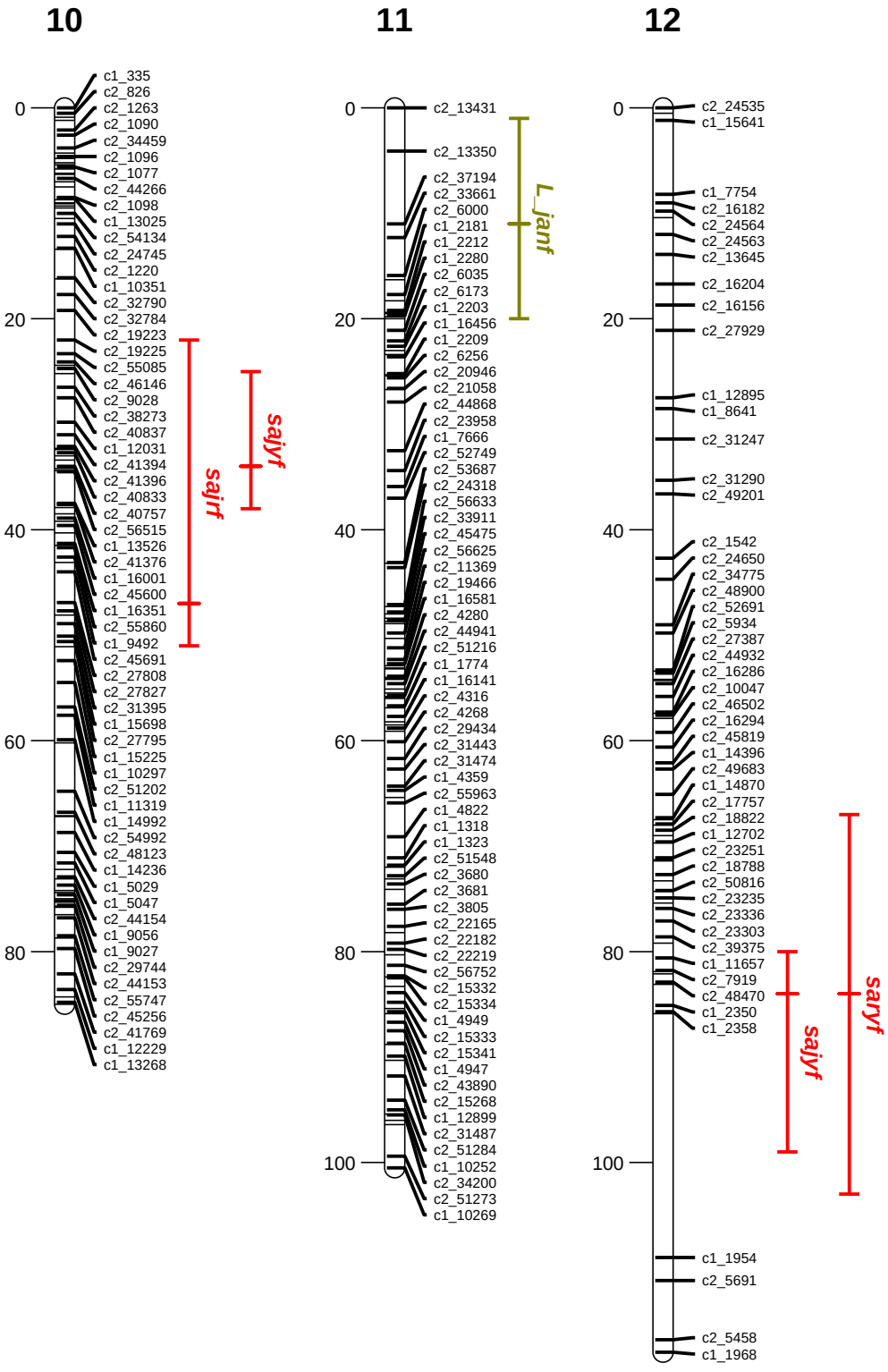


Figure 3.6. LOD profile across the genome for SECD rankings from potato chips processed in January, March, and May. LOD threshold is the mean of thresholds from all traits. Refer to table 3.2 for exact threshold values.

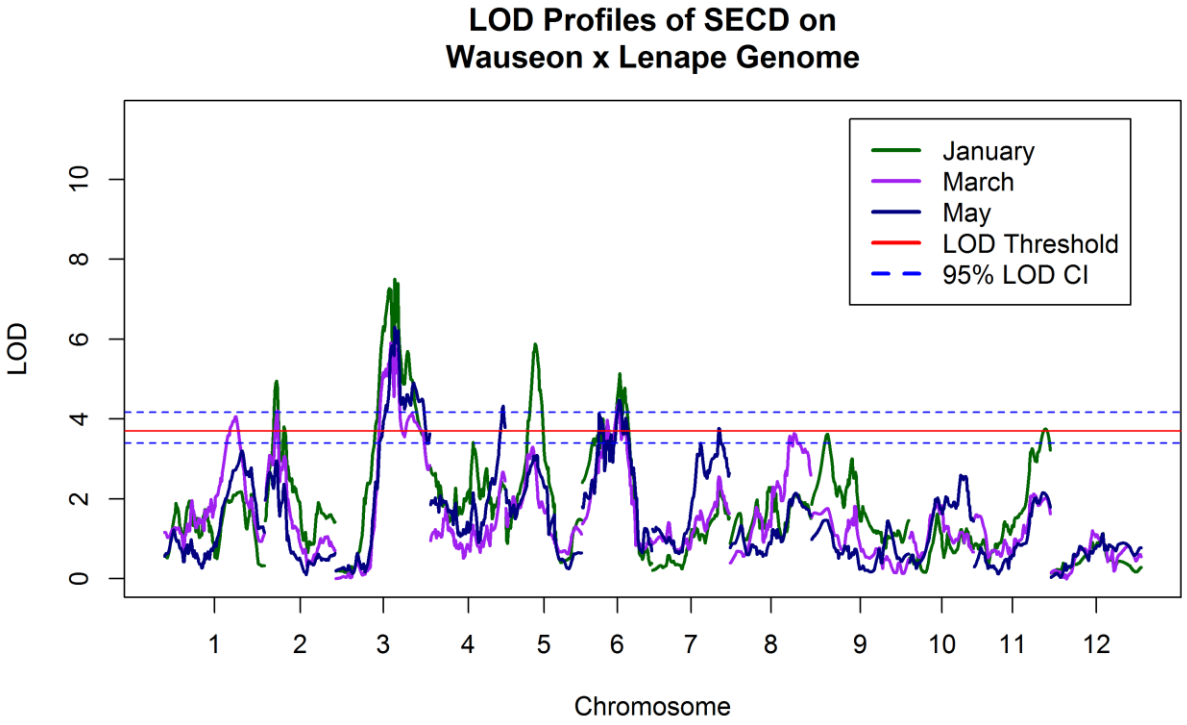


Figure 3.7. Single marker analysis of markers positioned closest to LOD peak for QTL mapping of SECD ranking in January. The box plots measure the distribution of SECD ranking within a genotypic class at a single marker with bold line indicating median, box lines encompass the middle 50% quantile, and whiskers encompassing the middle 99.3% quantile. Letters designate significance as determined by Tukey's HSD at a p value of 0.05. WSN and LNP designate the dosage of the two parents.

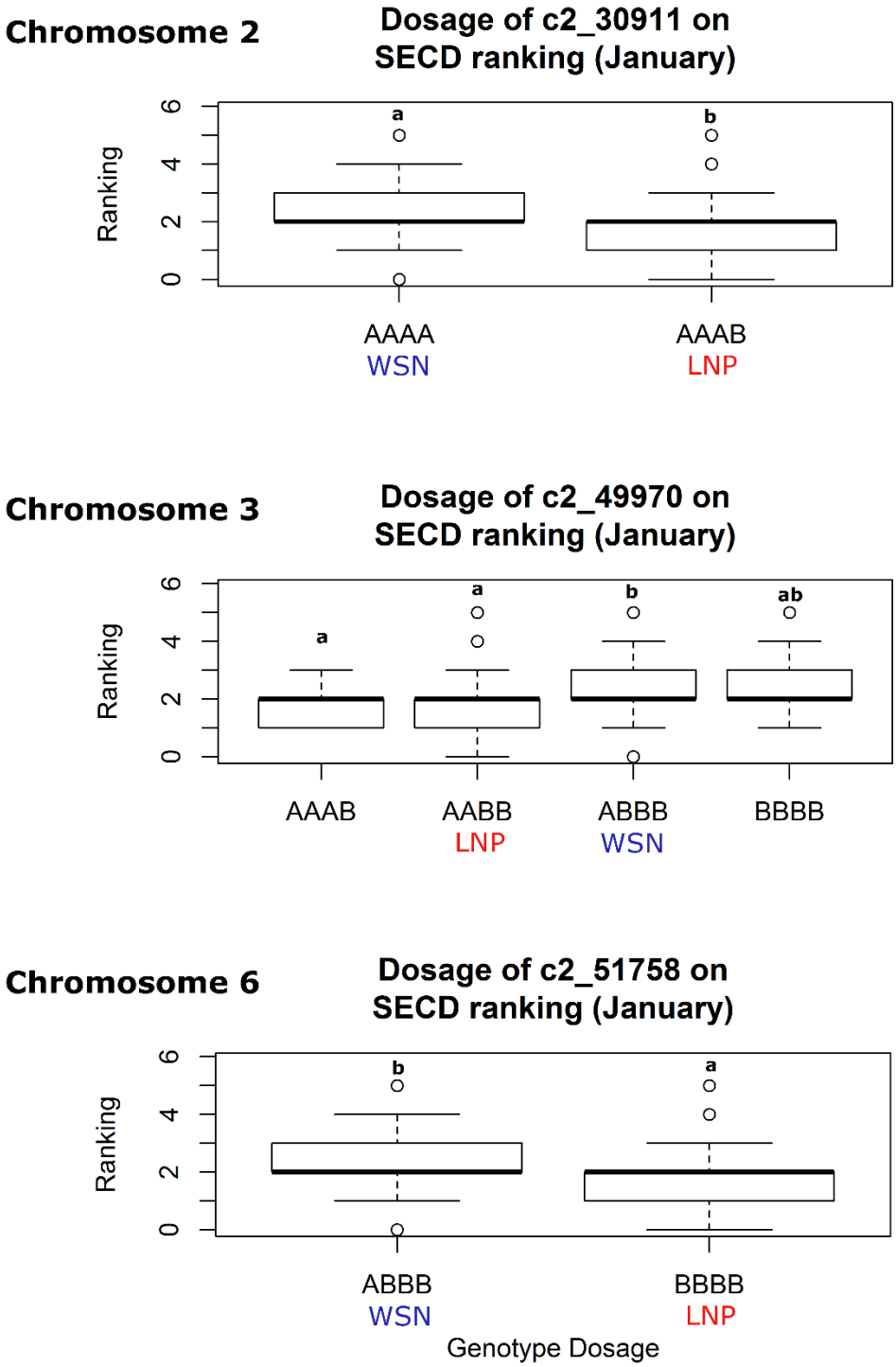
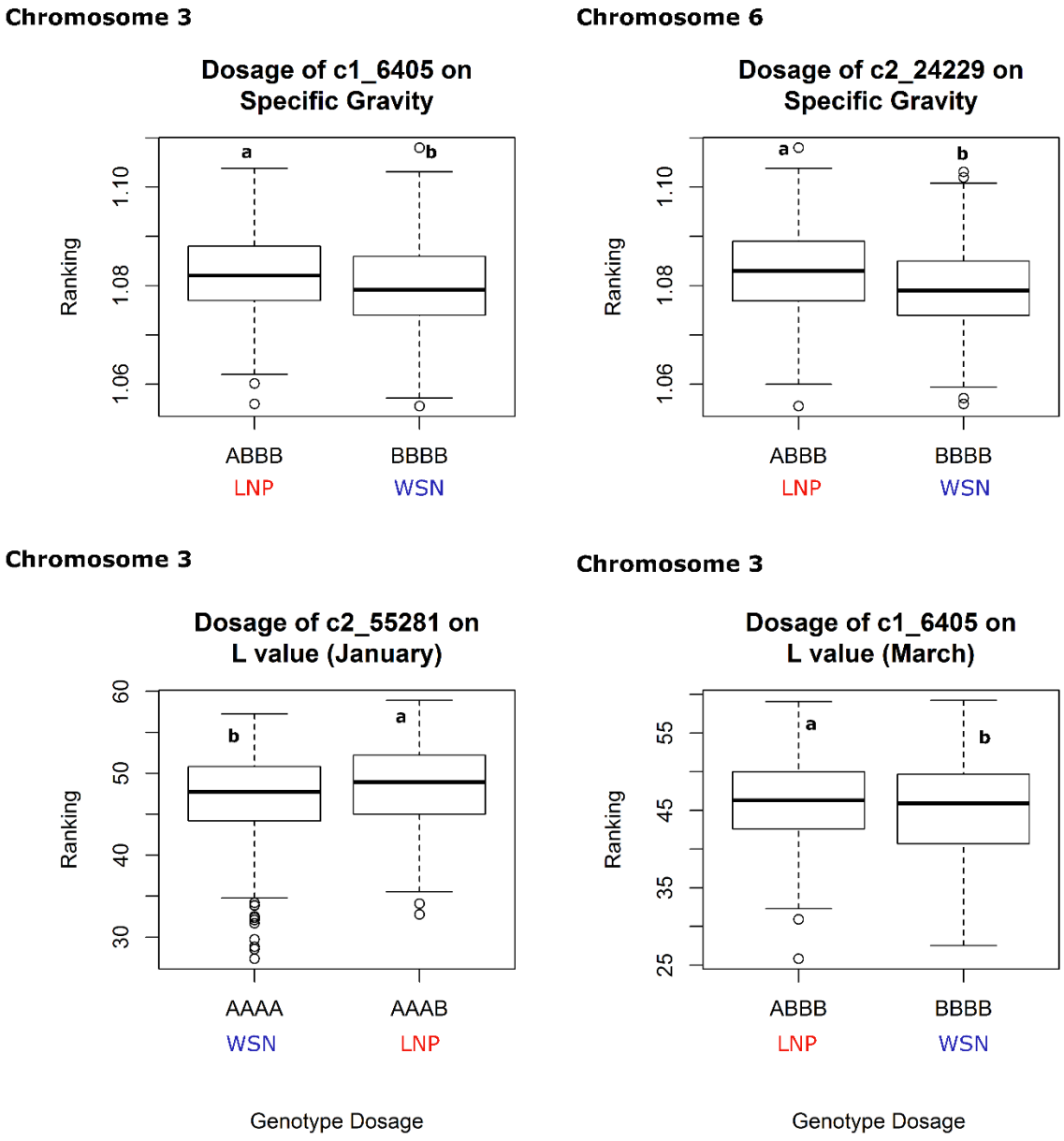


Figure 3.8. Single marker analysis of markers positioned closest to LOD peak for QTL mapping of specific gravity, L* value in January, and L* value in March. The box plots measure the distribution of traits within a genotypic class at a single marker with bold line indicating median, box lines encompass the middle 50% quantile, and whiskers encompassing the middle 99.3% quantile. Letters designate significance as determined by Tukey's HSD at a p value of 0.05. WSN and LNP designate the dosage class of the two parents.



Chapter 4: Evaluation of Hyperspectral Reflectance for Estimating Dry Matter and Sugar Concentration in a Processing Potato Population

4.1 Abstract

Rapid, accurate assessment of potato quality characteristics benefits potato breeders, growers and the potato processing industry. Hyperspectral reflectance is a proposed method to estimate tuber quality traits, such as dry matter and sugar contents, that are important to the potato processing industry. This technology can reduce costs and increase sample throughput for traits that require labor-intensive wet-chemistry analysis. Broad implementation across a diverse set of genotypes requires development of standardized tissue measurement protocols and derivation of chemometric equations to estimate the traits of interest from the hyperspectral measurements. We developed and analyzed computational models for estimating dry matter, sucrose, glucose and fructose contents in chipping potatoes using tubers of 86 individual genotypes from a bi-parental chipping potato population and 5 named varieties. Each individual tuber had spectral and physical samples taken from tissues including the face of a tuber cut in half, a 1.1 mm chip slice, a 10 mm slab, and trimmed tuber half. Models were constructed between spectral samples and physical trait values. Each data set combination was cross-validated 70:30 over 500 random sample permutations of a partial least squares regression. Our results indicate that dry matter is best estimated with 10 mm slab samples and the use of 700-1450 nm or 700-1800 nm spectra. The mean validation R^2 was 0.757 and mean RMSE was 0.0114 (w/w) dry matter. Sugar estimations were less accurate but obtained highest accuracy with tuber end tissue and 350-2500 nm spectra. External application of tissue and spectra parameters to create a new prediction model on a short wave infrared reflectance camera platform displayed accurate performance for dry matter prediction. The mean R^2 was 0.888 using 8 latent variables built on 359 tissue cores from 4 varieties. Coefficients from the PLSR model were applied to image pixels and displayed dry matter variation within individual tuber slabs.

4.2 Introduction

Processing potatoes account for the majority of potatoes grown in the United States (USDA-NASS 2016) and their value to industry is heavily impacted by a few biochemical traits. One of the most important is dry matter content of raw tubers, since dry matter content strongly influences recovery of product from raw tubers and finished product quality (Lulai & Orr 1979). Also important are tuber content of sucrose and the reducing sugars glucose and fructose. Elevated reducing sugars derived from sucrose contribute to the formation of undesirable, dark-colored pigments during frying as a result of the Maillard reaction (Habib & Brown 1957). Dark-colored products are undesirable to consumers, and the presence of potatoes with elevated reducing sugars in a shipment of potatoes can lead to a rejection at processing plants or deduction from the contracted price (Sowokinos 2001). For these reasons, the processing industry and researchers routinely measure these traits.

Measuring both tuber dry matter and sugars is a time and resource intensive activity. Dry matter content is typically estimated by measuring tuber specific gravity since the two are linearly related. In one method, specific gravity is calculated by dividing the weight of a potato sample in air by the weight of the same sample in water. This method does not permit the rapid collection of individual tuber values for this trait (Nissen 1955). An alternative method uses large tanks containing brine solutions that cover a range of specific gravities. Potatoes are passed from high to low specific gravity and the specific gravity of an individual tuber is taken to be halfway between that of the first solution in which it sinks and the previous solution. This method requires a large work area and lacks portability. Sucrose, glucose and fructose are the primary sugars quantified in potatoes. Currently, the methods employed include high performance liquid chromatography (HPLC) or an enzyme-based assay commonly implemented on instruments from YSI (Yellow Springs Instrument). Both methods require wet chemistry reagents, time to prepare samples, and intensive data analysis. Other methods include portable blood glucose monitors that are adapted for horticultural use. These have the disadvantages of not being very accurate due to their reliance on fresh potato juice and they are only able to directly measure glucose (Helgerud et al. 2016). Additionally, these

methods do not easily permit within tuber measurements of these traits. Within individual tubers there is heterogeneity for dry matter, sugars, and other culinary traits (Reeve et al. 1969). Thus, the ability to measure these traits at multiple locations within a tuber is highly desirable.

A proposed method to improve phenotyping of processing potatoes is to employ VIS-NIR-SWIR reflectance measurements from a portable, handheld spectrometer to estimate dry matter and sugar concentrations. Full-range hyperspectral reflectance spectroscopy captures the reflectance of light from a specimen at 350-2500 nm at narrow intervals (1-10 nm), yielding measurements of reflectance from hundreds of wavelengths or "bands". The spectral signature of an individual specimen is a consequence of its unique chemical composition, water content, and cellular structure. Differences in spectra among specimens occur in narrow wavelength regions because of their chemical or morphological differences. Thus, spectra can be used to estimate traits of interest through predictive models that relate the spectral signatures to measured physical or chemical properties. The quality of the spectral measurements and resulting models depend on measurement techniques, representative specimen samples, and the resolution and sensitivity of the spectrometer (Osborne 2000).

Partial least squares regression (PLSR) (Wold et al. 1984) is the chemometric method typically used to predict traits from spectra. PLSR is a generalization of multiple linear regression that creates a relationship between highly collinear predictor variables (spectral reflectance) and one or more response variables (traits of interest). PLSR has advantages over multiple linear regression by avoiding overfitting a model by reducing the predictor variables to orthogonal factors (latent vectors). PLSR is advantageous over principal components analysis (PCA) in the context of dimensional reduction for use in predictive modeling because it maximizes the covariance between predictor and response variables (Maitra & Yan 2008)(Wold et al. 2001)(Tobias 1995). Frequently, modifications of PLSR models in hyperspectral prediction include: number of selected latent vectors, spectral value transformation, removal of irrelevant spectral bands, data smoothing, and weighting of factor loadings.

Hyperspectral reflectance has been used previously to estimate dry matter and sugars in potato but only for a small number of varieties. Previous studies report correlations for model validation ranging from 0.90 - 0.98 and root mean squared error of prediction (RMSE) between 0.0130 – 0.0169 (w/w) for dry matter (Haase 2004)(Dull et al. 1989)(Subedi & Walsh 2009)(Singh et al. 2004). The prediction of glucose, fructose, and sucrose from reflectance spectroscopy has been less accurate, with calibration correlation coefficients between 0.38 – 0.95 and RMSE between 0.280 – 3.259 mg/g fresh weight (Rady et al. 2014)(Rady & Guyer 2015a) (Hartmann & Büning-Pfaue 1998) The aforementioned research evaluated a range of sample preparation techniques covering raw whole tubers, peeled, raw chip slice, slab, homogenized, and freeze dried (Dull et al. 1989)(Haase 2011)(Scanlon et al. 1999)(Hartmann & Büning-Pfaue 1998)(Haase 2004)(Singh et al. 2004). These and other hyperspectral estimation studies suggest that sample preparation method strongly influences prediction accuracy for multiple traits (López et al. 2013)(Rady & Guyer 2015b).

PLSR models in some crops have been developed based on a large number of varieties grown in multiple environments and years. These global equations have permitted the use of hyperspectral reflectance in breeding (De Leon 2006)(Nilsson 2011). In potato, one study evaluated hyperspectral reflectance for the estimation of sugars on a broad range of genotypes using freeze-dried samples. Sucrose, glucose, and fructose had prediction correlation coefficients between 0.95 – 0.97 and a standard error of prediction (RMSE) of 0.32 -1.15 mg/g fresh weight in a single hold out external validation method (Ayvaz et al. 2015). To our knowledge, there are no previous studies that have investigated the accuracy of a model applied across a genetic population to measure sugar or dry matter content from raw potato tissue.

Adoption of reflectance spectroscopy by potato researchers and industry requires high prediction accuracy on raw potato tissue from a broad diversity of genotypes. This would result in a global PLSR model for wide use, enabling reliable comparisons among investigators. A robust sampling protocol for raw potato tissue will minimize noise in spectral measurements and errors in estimation. Optimal sampling protocol for a single point spectroradiometer may then be used to develop accurate prediction models in a

hyperspectral reflectance imaging system. Here we report on models developed using a diverse range of potato genotypes and several easily implemented sampling methods for estimating the key processing traits of tuber dry matter, sucrose, glucose, and fructose. We identify the best tissue sampling procedures for dry matter prediction models under a single point spectroradiometer and then transfer this knowledge to a hyperspectral imaging system where we apply accurate models to map dry matter distribution in images of raw tuber tissue.

4.3 Materials and Methods

Tuber Material

The potatoes analyzed in this study included 86 tubers from a Wauseon x Lenape bi-parental population with each tuber having a different genotype, 4 tubers from var. Atlantic, 1 tuber from var. Lenape, 1 tuber from var. Megachip, 3 tubers from var. Snowden, and 1 tuber from var. Wauseon, amounting to 96 tubers in total. Tubers were from individual plots chosen randomly from a chip processing experiment. All tubers were grown at the Hancock Agricultural Research Station in the 2014 growing season in the same field plot. They were harvested 120 days after planting and stored at 10^0 C / 95% relative humidity until they were sampled for this experiment.

The tubers were sequentially cut into sample configurations that are typical for field analysis and fry processing (Figure 4.1). Potatoes were first cut longitudinally from apical to basal end with a knife to produce two tuber halves. One tuber half was discarded and two 1.1 mm thick slices were removed from the other tuber half with a custom-made chip mandolin. The first slice was discarded and the second slice was used for making spectral measurements. A 10 mm thick slab was then removed with a kitchen mandolin. The cut surface of the slab produced by the chip mandolin was much smoother than that produced by the kitchen mandolin. The remaining, trimmed tuber half (TTH) was the final sample.

Spectral Acquisition

Spectral readings were collected with an ASD Field Spec 3, a portable full range spectroradiometer (350-2500 nm) fitted with a flexible fiber optic cable, stable light source and a reflectance probe or leaf clamp attachment, depending on sample type (Analytical Spectral Devices, Boulder, CO). All cut surfaces were blotted dry with a Kimwipe (Kimberly-Clark, Irving, TX) before making a spectral reading to minimize reflectance from excess water. Readings were taken with the reflectance probe on the exposed tuber face after the first cut (t); with the leaf clamp on the chip slice (c); with leaf clamp on the side of the slab that was smoothed with the chip mandolin (s1); on the rougher side of the slab that was cut with the kitchen mandolin (s2), and with the reflectance probe on the trimmed tuber half (s3) (Figure 4.1). Spectral readings at each spectral sample type were collected in groups of 3 technical replicates and then averaged. A Spectralon white reference standard was measured between each of the 3 technical replicates to calculate relative reflectance from the spectral counts and ensure instrument stability (Labsphere, North Sutton, NH). Spectral data are interpolated by the instrument to 1 nm resolution.

Wet lab analysis

Tissue samples were collected from the same regions as spectral measurements. A 25 mm diameter disk was cut from the chip slice and a 16 mm diameter cylinder of tuber tissue was removed from the slab and trimmed tuber half using appropriately sized cork borers. These tissue samples were used with the c, s1, s2, and s3 spectral readings. All tissue samples were weighed on an analytical balance, placed in coin envelopes, flash frozen in liquid nitrogen, lyophilized and weighed to get sample dry weight. Dry matter content of each sample was calculated as dry weight divided by fresh weight. Tuber sugar contents of chip samples and ~0.19 g dry weight slab and trimmed tuber half subsamples were quantified using HPLC (Wilson et al., 1981)For the lab values associated with the tuber face (t) measurements, the chip, slab and trimmed tuber half laboratory values were combined using a weighted average.

Data Transformation and Outlier Detection/Subsetting

Prior to PLSR modeling, spectral data were analyzed within each tissue sample group to identify potential outliers. Principal components were applied to the spectral data and samples more than six standard deviations away from PC1 were considered unrepresentative of the population of tuber samples and removed from analysis since an erroneous sample with large leverage can exaggerate model prediction error (Shenk & Westerhaus 1991). Laboratory values were also removed if they were likely strong outliers. In both cases, errors could have been due to sample processing or user/instrument error during measurement. Six samples were identified as spectral outliers, leaving a final data set of 90 samples in all tissue types.

We tested several transformations of the 1 nm interval 350-2500 nm spectral data to smooth the noise in the spectra, including first-derivative of the reflectance, pseudo-absorbance (one divided by the log of reflectance), and a natural log transformation on the full spectral reflectance data set (Owen 1995). We also tested several spectral subsets including 700-1450 nm, 700-1800nm, and 700-2500 nm because discrete spectral regions can be important for prediction of particular traits (Curran 1989)(Ayvaz et al. 2015), while other regions (such as the visible) may be confounding. Although PLSR is designed to handle high dimensionality and collinearity in predictor data, inclusion of more spectral predictor variables than necessary still has the potential for model overfitting (Wold et al. 2001). In total, seven spectral data sets were tested for trait prediction.

Trait data for glucose and fructose were not normally distributed and were log normalized for the comparative analyses. The raw values of glucose and fructose were also combined into total reducing sugars (TRS) to evaluate whether prediction of TRS from spectra performed better than glucose or fructose alone (Hartmann & Büning-Pfaue 1998)(Haase 2011). TRS was also natural log transformed. Nine traits were evaluated in total: dry matter; concentrations of sucrose, glucose, fructose, and TRS, as well as the transformed concentration values of sucrose, glucose, fructose, and TRS. Each trait was modeled under each combination of spectral parameters for a total of 63 PLSR analyses.

Chemometrics

Partial least squares regression was used to predict traits from spectra (Wold et al. 1984)(Wold et al. 2001) using the pls package in R (Mevik & Wehrens 2007)(R Core Team 2014). To minimize prediction error in PLSR, the dimensions of the predictor data set were reduced to an optimal set of latent structures. The number of latent structures was determined through an iterative, leave one out, cross validation in which we fit PLSR models along a sequence of 1 through 40 latent variables and repeated each cross validation procedure 50 times on a single latent structure value. The number of latent vectors ultimately used for prediction was the set that minimized the median predicted residual sum of squares (PRESS) across the 50 permutations (Chen et al. 2004). This analysis was performed for each model combination of spectra and trait values.

Once the number of latent variables was selected, models were validated using a cross validation procedure where the data set of corresponding spectra and trait values was randomly split 500 times with 70% of the samples (63) used for model calibration and 30% (17) withheld for validation. We calculated the validation coefficient of determination (R^2), validation root mean squared error (RMSE), and validation bias by comparing the predicted trait values to the observed values. We report RMSE as the normalized root mean square error (NRMSE) where the RMSE was divided by the range of the predicted trait and is expressed as a percentage. NMRSE is a preferred metric as it enables comparison of model performance across data sets with different ranges and scales. These metrics convey the strength and robustness of the predictions by leveraging the distribution of each fit statistic across the 500 permutations of the data.

The outcome of PLSR is a chemometric equation weighting each wavelength for its contribution to prediction of the trait of interest. To have biological or chemical merit, a robust model should preferentially weight wavelengths known to be sensitive to the traits of interest. To assess our models, we evaluated the standardized coefficients of the PLSR models by wavelength as well as the variable importance of projection (VIP)(Mehmood et al. 2012) statistic for each wavelength. Together, the standardized coefficients and VIP indicate the importance of the variable for modeling both the predictor

variables and the trait, and those wavelengths that are identified as important in the PLSR can be compared to known or expected spectral features (Gosselin et al. 2010).

This analysis was performed for each combination of sample tissue, trait, and spectral dataset. This resulted in over 1,200 models constructed and cross validated. The data analysis was performed using the computing resources and assistance of the UW-Madison Center for High Throughput Computing (CHTC), which participates in the National Science Foundation and the U.S. Department of Energy's Office of Science Open Science Grid. Due to the number of simulations and models, only the most promising are reported.

Hyperspectral Imaging

The optimal tissue sample, spectral range, and data transformation was then validated using a different reflectance system to demonstrate the robustness of the chosen model parameters. The NIR hyperspectral imaging system used was a MSV-640 NIR Spectral camera with the ViaSpec II Scanner stage from Middleton Spectral Vision. Using this system, we collected images in the short wave infrared range of 950-1700 nm. Images were collected of a single 10 mm slab from 15 tubers per cultivar. For every slab image capture, a white reference image of a Spectralon standard and a dark current image from the closed aperture was captured as well. In this validation study, the cultivars Russet Burbank, Ranger Russet, Snowden, and Dark Red Norland were used due to their widespread cultivation and because they occupy different market segments. In each of the slabs, 6 tissue samples were collected along the transect perpendicular to the longitudinal axis with a 1 cm diameter cork borer. Slabs were scored with the cork borer prior to image collection so that location of a collected sample within a slab could be identified and cross references with the spectral signal from the sampling area. After image capture, cores were weighed and flash frozen in liquid nitrogen. All samples were lyophilized for 3 days before having their dry weight determined.

Images were converted from raw intensity to relative reflectance using spectral data from a white reference image and dark current image through a custom Python script. Systematic sensor noise was corrected in images using a script written by IDL+ENVI (Exelis Visual Information Solutions, Boulder, Colorado). During image analysis, the sample area identified by the score mark was used as the region of interest (ROI) boundary. The average spectral reflectance was determined for each core sample ROI. With spectral and trait data, we constructed PLSR models to evaluate this system's prediction accuracy. The number of latent variables was determined by minimizing the predicted residual sum of squares with importance placed on keeping the number of latent variables to a minimum (Chen et al. 2004). After the 500 permutations of the PLSR cross validation procedure, the mean raw coefficients for each wavelength predictor and the intercept was extracted. This equation was then used in a custom Python script to render color images in the scale of predicted dry matter for each slab image.

4.4 Results

Prediction of dry matter

Dry matter ranged from 0.12-0.25 (w/w) and was normally distributed across all samples. This trait was accurately predicted using PLSR models but certain factors sizably improved prediction accuracy. Mean R^2 and mean RMSE (Table 4.1) generally decreased within a tissue sample as the spectra wavelength range used for modeling was expanded.

The model based on the rough slab slice within the 700-1800 nm wavelengths had the highest accuracy with a mean R^2 of 0.757 (range 0.389 - 0.939) and a mean RMSE of 0.0114 (w/w) dry matter content. The slab tissue model was also highest in accuracy for the 700-1450 nm model with a mean R^2 of 0.750 (range 0.266-0.940) and a RMSE of 0.0116 (w/w) dry matter, a normalized error of 8.7%. The models constructed using samples collected from the chip slice tissue had comparable results using the same spectra wavelengths but then superseded the slab based models when using the larger 700-2500 nm and 350-2500 nm datasets. The smoother surface slab sample model had a lower mean R^2 for all of the models when

compared to the rougher surface slab sample. Models based on the TTH were lower in predication accuracy than chip slice, smooth slab, and rough slab when using all spectral data sets with the exception of the full 350-2500 nm spectra, where it outperformed the smooth slab tissue sampling. This particular trimmed tuber half model had a mean R^2 of 0.490 (range -0.009 - 0.809) and a RMSE of 0.0210 (w/w) dry matter. The tuber half model had the lowest prediction accuracy of all models with its best prediction occurring with the 700-1450 nm spectra with a R^2 of 0.400 (range -0.039 - 0.775), a NRMSD of 0.0203 (w/w) dry matter. Transformed spectra, natural log, first derivative, and absorbance, were not an improvement over the untransformed models with any of the data ranges (data not shown).

Figure 4.2 shows the R^2 distributions of the 12 top performing prediction models. This figure was limited to the 350-2500 nm, 700-1450 nm, and 700-1800 nm spectra combinations with the chip slice, smooth slab, rough slab, and TTH tissue/spectra samples. The rough slab model using 700-1800 nm spectra had the highest mean R^2 and the lowest mean RMSE and was significantly different from the rough slab model constructed using the 700-2500 nm spectra and all of the chip slice, smooth slab and trimmed tuber half models. For all tissue types illustrated in Figure 4.2, the validation R^2 of prediction models using 700-1450 nm was not significantly different from that using 700-1800 nm spectra.

Realistic predictive models for dry matter should compare favorably to previous investigations of spectral absorption features. Starch is known to have absorption properties at 970, 990, 1200, 1450, 1530, 1540, 1580, 1690, 1780, 1900, 1940, 2000, 2100, 2250, 2270, 2280, and 2320 nm and water has known absorption bands at 970, 1200, 1400, and 1940 nm (Curran 1989). Potato dry matter is approximately ¾ starch and the largest proportion of raw tuber tissue is water (Woolfe 1987). We compared the known absorption bands of starch and water with VIP from the best performing models (Figure 4.6). This model has clear trends where local maxima in VIP correspond to known absorption features of water and starch. For example, local maxima in VIP correspond to water absorption features at 970, 1200 and 1400 and starch absorption features at 970, 990, 1200, 1450 and 1580 (Figure 4.6).

Prediction of sugars

Tuber sucrose, glucose, and fructose contents ranged from 0.65-2.93, 0.03-1.61, and 0.03-2.09 mg/g fresh weight of tuber tissue, respectively. Sucrose followed a normal distribution in all samples whereas glucose and fructose did not and were right skewed.

Prediction models for sugars had prediction validation metrics that depended on tissue sampled, spectra data used, and trait transformation. The trait modeled also impacted the RMSE statistic due to the scale of the trait itself. Normalized Root Mean Squared Error (NRMSE) was calculated by dividing RMSE by the traits observed range. The highest prediction accuracies were found in glucose, fructose, and TRS. Models constructed using the TTH tissue for both spectral and trait samples had the highest prediction accuracies although all were significantly and sizably lower than predictions for dry matter.

The model constructed for predicting TRS using TTH tissue and 350 – 2500 nm spectra had a R^2 of 0.270 (range -0.047 – 0.645) and a mean NRMSE of 18.87%. Among the models using trimmed tuber half tissue and 350 -2500 nm spectra, glucose and fructose models (Table 4.2) were not significantly different in their prediction accuracy. Glucose had a mean R^2 of 0.258 (range -0.044 – 0.728) and mean NRMSE of 16.68 % and fructose had a mean R^2 of 0.244 (range -0.047 – 0.759) and mean NRMSE of 19.64%. Prediction models for sucrose were lower with a mean R^2 of 0.085 (range -0.047 – 0.608) and mean NRMSD of 20.73%. Transforming the traits did not increase the model accuracy and only in glucose did it not result in a significant decrease for both validation R^2 and NRMSE (Figure 4.4).

SWIR Camera Prediction Accuracy

This model was constructed using 359 samples from 60 tuber slabs that ranged in dry matter from 0.06 to 0.29 (w/w). Each variety had a different range of measured dry matter but data points overlapped considerably; Snowden ranged from 0.07 – 0.29, Russet Burbank ranged from 0.09 to 0.27, Ranger Russet ranged from 0.10 to 0.29, and Dark Red Norland ranged from 0.06 to 0.22. The PLSR model constructed for dry matter prediction had a mean R^2 of 0.888 (range 0.825 – 0.933) and mean RMSE of 0.0158 (w/w)

dry matter. The bias was centered on 0 across all model permutations (Figure 4.5). Images were created by scaling the pixel intensity values**.** The resulting images from the 4 cultivars are seen in Figure 4.7. The pixels are color scaled to the predicted dry matter value and their accuracy was described above and in Figure 4.5. It is visually suggestive that although the core samples had a wide trait range, there are areas in the slab that had dry matter contents that were higher and lower than those in the measured samples within a particular slab.

4.5 Discussion

Ninety potato genotypes were used for this study. Individual tubers varied widely in dry matter and sugar contents. Sample dry matter content ranged from less than 10% to more than 25%, a range comparable to that found across market classes of cultivated potato. Content of different sugars was equally variable, and observed data included values typical of chip processing, fry processing and fresh market potatoes. The data presented here demonstrate that potato dry matter can be accurately predicted from raw potato tissue using PLSR models of hyperspectral reflectance over a wide range of values and in heterogeneous genetic backgrounds (Figure 4.3). Model combinations that performed reliably were identified by examining a combination of spectra subsets, transformations, and tissue sampling methods. The robust testing methodologies used provide insights into the variability present within a cross validated prediction (Figure 4.2). These data allow others interested in this approach to determine whether implementing a similar methodology is appropriate for their application.

The most accurate dry matter model, the rough slab 700-1800 nm model, had a strong relationship between predicted and observed dry matter (Figure 4.3) and prediction errors were small compared to the range of the dry matter values in the population (NRMSE = 8.7%). Mean RMSE was 0.0114 (w/w) which compares favorably to values reported previously, 0.011-0.169 (w/w), for models developed for one potato variety or a particular market class (Dull et al. 1989)(Subedi & Walsh 2009)(Scanlon et al. 1999). Because of our permutation approach, the cross-validated results presented here, however, are likely to provide a

better estimation of what the true model statistics should be and are less prone to sampling error that can impact single set hold out model validations.

The tissue samples prepared as chips or slabs were demonstrably better for dry matter prediction than the other tissue types (smooth slabs and tuber halves, Table 4.1). The results were not always consistent among sample types. Within the full 350-2500 nm range, the best model was developed from the chip slice data. Conversely, the rough slab tissue data generated the best performance model with the 700 – 1800 nm spectra (R^2=0.757, RMSE = 0.0114, Table 4.1). Interactions between tissue sample and spectra were not observed when confined to the 700-1450 nm and 700-1800 nm spectra parameters. In all cases, the 700-1450 nm spectra had prediction accuracies comparable to those from models constructed using the larger 700-1800 nm spectra. Thus, the data suggest that dry matter can be predicted accurately using simpler and cheaper hyperspectral devices than the one use for this research.

Prediction models for all sugars, their combinations, and their transformants had lower mean R^2 values and higher mean NRMSE values than the dry matter prediction models. As reported, the tuber end models with the 350–2500 nm spectra had the highest accuracy with differences between the traits. The best sugar prediction model was for total reducing sugars and had a mean NRMSE of 18.87% of the trait range. In comparison, the best dry matter model using the rough slab and 700-1800 nm spectra had a NRMSE of 9.62 % trait range. When sugar models were reported as RMSE values in units of mg per g FW, sucrose, glucose, fructose, and TRS had values of 0.427, 0.272, 0.302, and 0.566.

Accuracy of prediction models might be improved using other spectral reflectance capture techniques. Comparable studies using interractance probes on raw tissue from smaller variety sets found higher R^2 values of 0.50-0.95 and RMSE of 0.439 – 3.410 for sucrose, R^2 values of 0.83-0.95 and RMSE of 0.515 – 0.870 for glucose, R^2 values of 0.71-0.95 and RMSE of 0.260 - 1.01 for fructose, and R^2 values of 0.81-0.93 and RMSE of 0.600 – 2.040 for TRS (Yaptenko et al. 2000)(Mehrubeoglu & Cote 1997)(Yu Chen et al. 2010)(Rady & Guyer 2014). Reflectance based spectral measurements in the literature tended to have poorer accuracy than other methods of hyperspectral capture.

Studies using homogenized or lyophilized tuber tissue in combination with diffuse reflectance spectroscopy have yielded more accurate predictions for sugars, but such methods require considerable sample processing. Our study measured raw tubers with a portable spectrometer, potentially offering a rapid method for a general characterization of tuber quality. Potentially these compounds are not abundant enough in raw potato tissue to predict accurately with hyperspectral reflectance methods (Hartmann & Büning-Pfaue 1998)(Haase 2011), especially fresh samples in which water absorption features can mask other compounds (Curran et al. 1992) (Dawson et al. 1999). We hypothesized that a full range spectrometer had the spectral resolution to overcome this masking effect, but the modest prediction accuracy achieved for all sugars suggests that either a finer resolution is required or that by using raw samples, signals related to sugar content are masked by the water signal.

The results of the dry matter models permitted further study evaluating the use of a hyperspectral camera for dry matter visualization (Figure 4.7). The dry matter prediction models identified reflectance spectroscopy on slab tissue samples in the wavelength range of 700-1800 nm to be suitable for accurate predictions. The SWIR camera only permitted the use of the 950-1700 nm but our investigation of the absorbance bands in Figure 4.6 (Curran 1989), the VIP, and the coefficients allowed for the assumption that it would still be an appropriate tool. This assumption was validated after evaluating the model fit statistics created from our PLSR routine (Figure 4.5). Even though the model was constructed using a hyperspectral camera, the mean R^2 of 0.888 was close to the previous R^2 value range of 0.93 - 0.97 from Dull and Subedi and sizably higher than the 0.77 reported by Scanlon. The mean RMSE of 0.0158 (w/w) is also comparable to other reflectance studies including the models in this study developed using a population of genotypes (Dull et al. 1989)(Subedi & Walsh 2009)(Scanlon et al. 1999)

Interestingly, there appears to be noticeable within tuber variation for dry matter. While cultivars are distinct, the level of variation within a cultivar is both expected (Reeve et al. 1969)(Peiris et al. 1999) and potentially valuable to the processing industry for maximizing fried product yield and quality (Sayre et al. 1975)(Lulai & Orr 1979). Hyperspectral imaging in the shortwave infrared spectral range has been

evaluated for the prediction of specific gravity in potato. In this case, wavelength selection and numerous mathematical procedures produced at best a prediction R^2 0.978 and a RMSEP of 0.009, (equates to approximately 0.0160 in dry matter). Generally, the vast majority of the fitted models had performance that fell within the range of this study's model, which is an interesting result considering the inclusion of sweet potato may have inflated the fit statistics (Su & Sun 2016). This lack of departure between models suggests that complex mathematical spectral processing may not be necessary for hyperspectral reflectance imaging of dry matter in potato.

4.6 Conclusions

This study provides a methodological framework for the development of a global dry matter prediction model in potato. The data show clearly that hyperspectral reflectance can be used to accurately predict dry matter content in heterogeneous genotypes of potato. We demonstrated that spectral samples from many genetically unique individuals can be used to quickly develop a functional model. The success across a broad set of samples in the creation of a calibration equation shows that a global calibration equation could potentially be used on a wider basis in the potato industry. The best performing PLSR model used a combination of a simply prepared slab tissue sample and spectra that did not extend to wavelengths beyond 1800 nm. These findings indicate that involved methods of sample preparation and full range hyperspectral devices up to 2500 nm are not required to accurately predict dry matter content of potato and this is further emphasized by its application to a SWIR imaging camera system. The results were less promising for predicting sugars, but spectral data could be used to provide coarse estimates. Additional work is needed to develop a method to measure sugars on raw samples for commercial practice using a portable spectrometer, perhaps due to noise or insufficient spectral resolution. Spectra on dried samples may perform better, as it is also possible that sugar concentrations are not sufficiently high to overcome confounding factors in spectra due to water content. Nevertheless, this study demonstrates the potential for handheld contact spectrometry to rapidly characterize dry matter content in potatoes.

4.7 References

Ayvaz H, Santos AM, Moyseenko J, Kleinhenz M, Rodriguez-Saona LE. 2015. Application of a portable infrared instrument for simultaneous analysis of sugars, asparagine and glutamine levels in raw potato tubers. Plant Foods Hum Nutr. 70:215–220.

Chen S, Hong X, Harris CJ, Sharkey PM. 2004. Sparse modeling using orthogonal forward regression with PRESS statistic and regularization. IEEE Trans Syst Man, Cybern Part B Cybern. 34:898–911.

Curran PJ. 1989. Remote sensing of foliar chemistry. Remote Sens Environ. 30:271–278.

Curran PJ, Dungan JL, Macler BA, Plummer SE, Peterson DL. 1992. Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. Remote Sens Environ. 39:153–166.

Dawson TP, Curran PJ, North PRJ, Plummer SE. 1999. The propagation of foliar biochemical absorbtion features in forest canopy reflectance: a theoretical analysis. Remote Sens Envirnonment. 67:147–159.

Dull G, Birth G, Leffler R. 1989. Use of near infrared analysis for the nondestructive measurement of dry matter in potatoes. Am Potato J. 66:215-225.

Gosselin R, Rodrigue D, Duchesne C. 2010. A Bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications. Chemom Intell Lab Syst. 100:12–21.

Haase NU. 2004. Estimation of dry matter and starch concentration in potatoes by determination of under-water weight and near infrared spectroscopy. 46:117–127.

Haase NU. 2011. Prediction of potato processing quality by near infrared reflectance spectroscopy of ground raw tubers. J Near Infrared Spectrosc. 19:37–45.

Habib AT, Brown HD. 1957. Role of reducing sugars and amino acids in the browning of potato chips. Food Technol. 11:85–89.

Hartmann R, Büning-Pfaue H. 1998. NIR determination of potato constituents. Potato Res. 41:327–334.

Helgerud T, Knutsen SH, Afseth NK, Stene KF, Rukke EO, Ballance S. 2016. Evaluation of hand-held instruments for representative determination of glucose in potatoes. Potato Res.:1–14.

De Leon N. 2006. NIRS Prediction Equations : Development of NIRS prediction equations [Internet]. [cited 2016 Nov 16]. Available from: http://cornbreeding.wisc.edu/analytics/nirs-prediction-equations/

López A, Arazuri S, García I, Mangado J, Jarén C. 2013. A review of the application of near-infrared spectroscopy for the analysis of potatoes. J Agric Food Chem. 61:5413–24.

Lulai EC, Orr PH. 1979. Influence of potato specific gravity on yield and oil content of chips. Am Potato J. 56.

Maitra S, Yan J. 2008. Principle component analysis and partial least squares : Two dimension reduction techniques for regression. Casualty Actuar Soc. Discusion Paper Program. p. 79–90.

Mehmood T, Liland KH, Snipen L, Sæbø S. 2012. A review of variable selection methods in partial least squares regression. Chemom Intell Lab Syst. 118:62–69.

Mehrubeoglu M, Cote GL. 1997. Determination of total reducing sugars in potato samples using near-infrared spectroscopy. Cereal Foods World. 42:409–413.

Mevik B-H, Wehrens R. 2007. The pls package: Principle component and partial least squares regression in R. J Stat Softw. 18:1–24.

Nilsson T. 2011. Comparison of the FOSS NIR global ANN calibration against reference methods : A five year pan-European study. Hilleroed, Denmark.

Nissen M. 1955. The weight of potatoes in water. Am Potato J. 32:332–339.

Osborne BG. 2000. Near-infrared spectroscopy in food analysis. In: Meyers R, editor. Encycl Anal Chem. Chichester: John Wiley & Sons Ltd; p. 1–14.

Owen AJ. 1995. Uses of derivative spectroscopy: application note. Waldbronn, Germany.

Peiris KHS, Dull G, Leffler RG, Kays SJ. 1999. Spatial variability of soluble solids or dry-matter content within individual fruits, bulbs, or tubers: Implications for the development and use of NIR spectrometric techniques. HortScience. 34:114–118.

R Core Team. 2014. R: A language and environment for statistical computing. R Found Stat Comput Vienna, Austria [Internet]. Available from: https://**www.r-project.org/**

Rady A, Guyer D. 2014. Utilization of visible/near-infrared spectroscopic and wavelength selection methods in sugar prediction and potatoes classification. J Food Meas Charact. 9:20–34.

Rady AM, Guyer DE. 2015a. Evaluation of sugar content in potatoes using NIR reflectance and wavelength selection techniques. Postharvest Biol Technol. 103:17–26.

Rady AM, Guyer DE. 2015b. Rapid and/or nondestructive quality evaluation methods for potatoes: A review. Comput Electron Agric. 117:31–48.

Rady AM, Guyer DE, Kirk W, Donis-González IR. 2014. The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict processing-related constituents of potatoes. J Food Eng. 135:11–25.

Reeve RM, Hautala E, Weaver ML. 1969. Anatomy and compositional variations within potatoes III. Gross compositional gradients. Am Potato J. 47:148–162.

Sayre RN, Nonaka M, Weaver ML. 1975. French fry quality related to specific gravity and solids content variation among potato strips within the same tuber. Am Potato J. 52:73–82.

Scanlon MG, Pritchard MK, Adam LR. 1999. Quality evaluation of processing potatoes by near infrared reflectance. 771:763–771.

Shenk JS, Westerhaus MO. 1991. Population definition, sample selection, and calibration procedures for near infrared reflectance spectroscopy. Crop Sci. 31:469.

Singh B, Wang N, Prasher S, Ngadi M. 2004. A spectroscopic technique for water content determination in potato. ASAE Annu Int Meet 2004.:1–11.

Sowokinos JR. 2001. Biochemical and molecular control of cold-induced sweetening in potatoes. Am J Potato Res. 78:221–236.

Su W-H, Sun D-W. 2016. Comparative assessment of feature-wavelength eligibility for measurement of water binding capacity and specific gravity of tuber using diverse spectral indices stemmed from hyperspectral images. Comput Electron Agric. 130:69–82.

Subedi PP, Walsh KB. 2009. Assessment of potato dry matter concentration using short-wave near-infrared spectroscopy. Potato Res. 52:67–77.

Tobias RD. 1995. An introduction to partial least squares regression. Proc Ann SAS Users Gr Int Conf, 20th, Orlando, FL.p. 2–5.

USDA-NASS. 2016. Potatoes 2015 summary. Washington, D.C.

Wilson AM, Work TM, Bushway AA, Bushway RJ. 1981. HPLC determination of fructose, glucose, and sucrose in potatoes. J Food Sci. 46:300–301.

Wold S, Ruhe A, Wold H, Dunn WJ. 1984. The collinearity problem in linear regression: The partial least squares (PLS) approach to generalized inverses. Soc Industiral Appl Math. 5:735–743.

Wold S, Sjöström M, Eriksson L. 2001. PLS-regression: A basic tool of chemometrics. Chemom Intell Lab Syst. 58:109–130.

Woolfe J. 1987. The potato in the human diet. New York, NY: Cambridge University Press.

Yaptenko K, Suzuki T, Kawakami S, Sato H, Takano K, Kozima T. 2000. Nondestructive determination of sugar content in “Danshaku” potato. J Agric Sceince. 44:284–294.

Yu Chen J, Zhang H, Miao Y, Asakura M. 2010. Nondestructive determination of sugar content in potato tubers using visible and near infrared spectroscopy. Japan J Food Eng. 11:59–64.

Table 4.1. Dry matter prediction model evaluation with 500 permutations of a randomly sampled cross validation procedure. Displayed here is the validation data set mean values for R^2 and root mean squared error (RMSE) of the randomly sampled cross validation procedure. Tissue represents the tuber tissue sample type used for both spectral and trait collection. This was combined with spectra subsets for partial least squares model calibration. The number of latent structures selected for each tissue/spectra combination is presented under Latent variables in the respective order of the four Spectra columns.

Tissue		**Latent variables**	**Statistic**	**Spectra**			
				700-1450 nm	*700-1800 nm*	*700-2500 nm*	*350-2500 nm*
Chip Slice	(c)	9, 9, 9, 13	valR2	0.723	0.713	0.67	0.661
			RMSE	0.0157	0.0161	0.0175	0.0177
Smooth Slab	(s1)	13, 13, 14, 8	valR2	0.63	0.612	0.541	0.312
			RMSE	0.0145	0.0149	0.0166	0.0202
Rough Slab	(s2)	12, 12, 12, 9	valR2	0.75	0.757	0.663	0.569
			RMSE	0.0117	0.0114	0.0136	0.0153
Trimmed	(s3)	6, 6, 6, 13	valR2	0.546	0.536	0.512	0.49
Tuber Half			RMSE	0.0186	0.0189	0.0191	0.0210
Tuber Half	(t)	10, 10, 10, 5	valR2	0.4	0.388	0.325	0.057
			RMSE	0.0204	0.0206	0.0224	0.0261

Table 4.2. Model prediction across 500 permutations of a randomly sampled cross validation procedure where 70 percent of the dataset was used for calibration and 30 percent was used for validation. Presented are the top models for sugar prediction using chip, rough slab, and trimmed tuber end tissues for spectra and trait collection using the 350 – 2500 nm spectra. L.V. designates the number of latent variables in the model, R^2 is the mean model coefficient of variation, NRMSE is the normalized root mean squared error (RMSE) reported as a percentage of the trait range. R^2 Range and NRMSE Range are the maximum and minimum values for both statistics in the 500 permutations.

Tissue	Trait	Latent variables	R^2	R^2 Range	NRMSE	NRMSE Range
Chip	Fructose	6	0.115	(-0.045 - 0.534)	16.45	(9.09 - 26.15)
	Glucose	5	0.157	(-0.045 - 0.578)	17.82	(8.80 - 28.98)
	Sucrose	3	0.114	(-0.048 - 0.581)	21.28	(11.17 - 29.58)
	Total Reducing Sugars	4	0.138	(-0.045 - 0.555)	17.78	(10.26 - 27.25)
Rough Slab	Fructose	5	0.071	(-0.048 - 0.49)	16.78	(8.65 - 27.57)
	Glucose	3	0.087	(-0.048 - 0.537)	16.39	(8.97 - 27.97)
	Sucrose	2	0.097	(-0.045 - 0.523)	21.03	(12.75 - 30.82)
	Total Reducing Sugars	10	0.129	(-0.048 - 0.455)	20.33	(12.56 - 35.29)
Tuber End	Fructose	6	0.244	(-0.048 - 0.76)	19.65	(12.43 - 29.65)
	Glucose	6	0.258	(-0.044 - 0.729)	16.68	(8.76 - 23.64)
	Sucrose	2	0.085	(-0.048 - 0.608)	20.74	(14.07 - 29.12)
	Total Reducing Sugars	6	0.270	(-0.048 - 0.645)	18.87	(11.29 - 25.73)

Figure 4.1. Diagram of the tuber dissection for spectral and trait sample collection. Tubers were bisected longitudinally from apical to basal end. Spectral readings were collected sequentially from the center of a tuber half (t), a chip (c), the smoother (s1) and rougher (s2) surfaces on a slab, and a trimmed tuber half (s3).

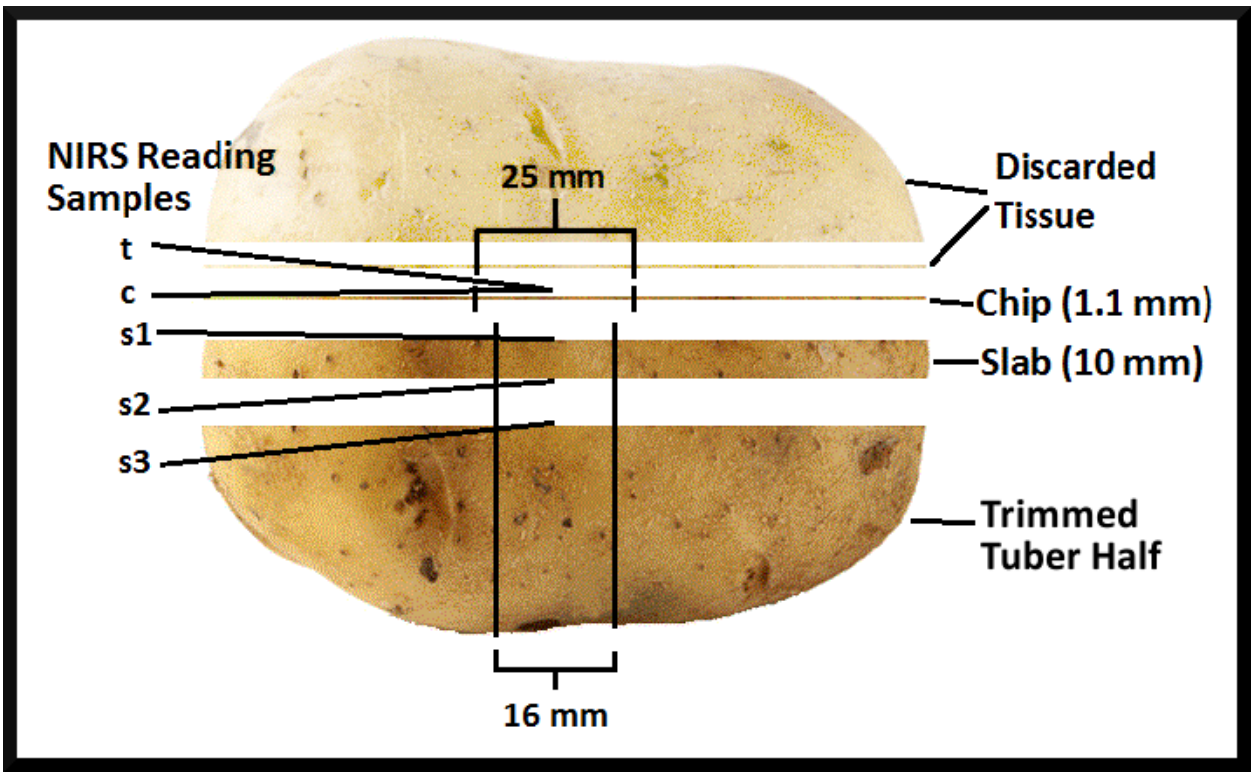


Figure 4.2. Validation R^2 distributions of dry matter prediction models from 500 permutations of a randomly sampled cross validation procedure. Each boxplot represents the combination of tissue sample used for spectral/trait collection and the spectral data wavelength range used for partial least squares model calibration. Letters represent Tukey honest significant difference grouping.

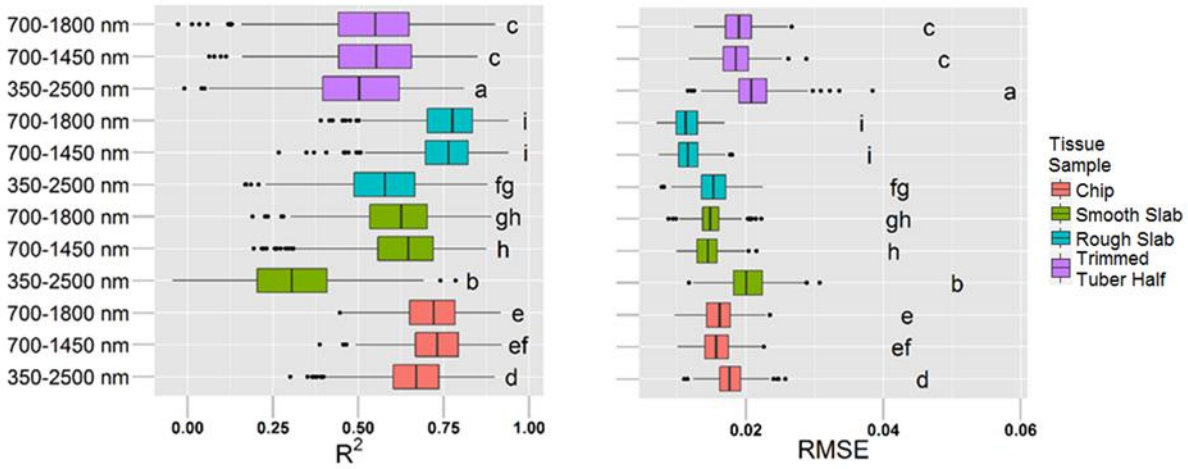


Figure 4.3. Plot of predicted vs. observed dry matter for the model constructed with rough slab spectra/trait sampling, 13 latent variables, and the spectral wavelengths of 700-1800 nm. The points displayed are from the validation of the 500 permutations of a randomly sampled cross validation procedure. Error bars are the standard deviation of the predicted value for that individual. The dashed line is the linear model of best fit between mean predicted and observed dry matter points.

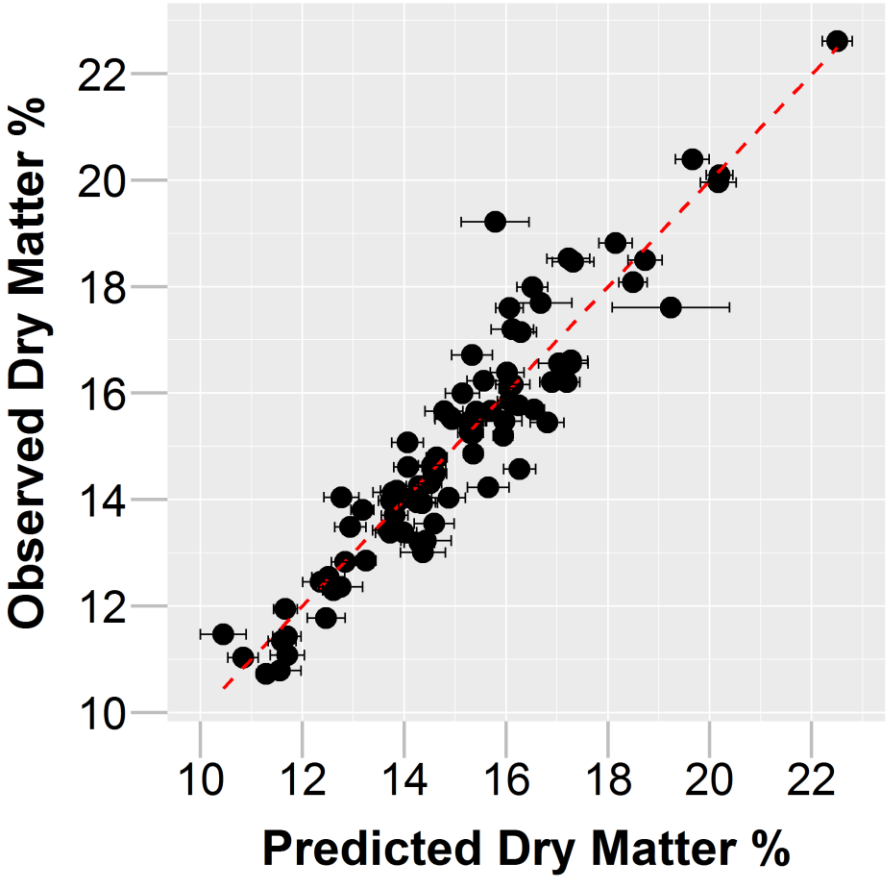


Figure 4.4. Validation R^2 and normalized root mean squared error (NRMSE) distributions of sugar prediction models based on tuber end tissue sampling from 500 permutations of a randomly sampled cross validation procedure. Each boxplot represents the model fit for the three directly measured traits of fructose, glucose, and sucrose, the combination of fructose and glucose into total reducing sugars (TRS) and the log transformation of glucose, fructose end TRS. Letters represent Tukey honest significant difference grouping.

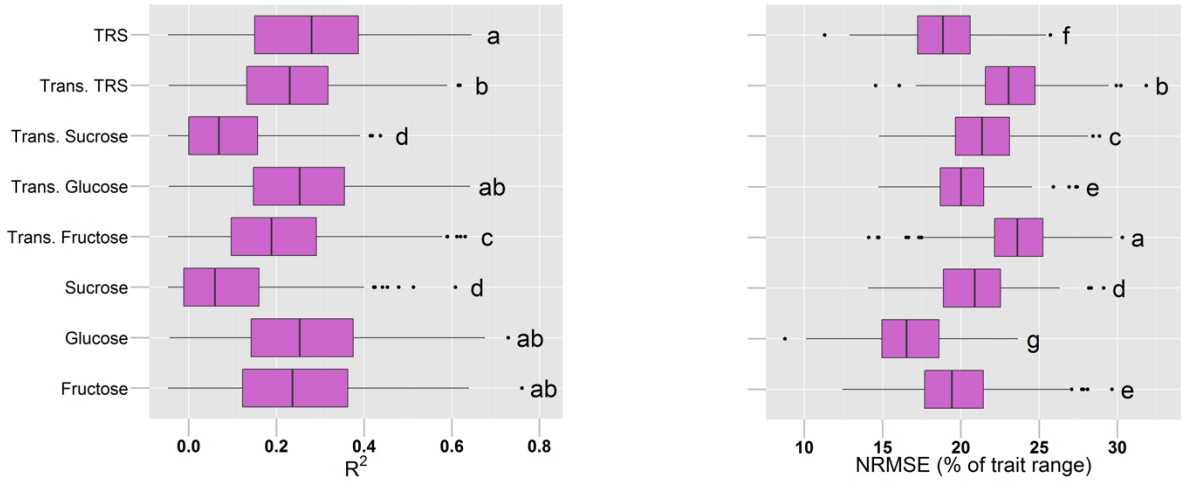


Figure 4.5. Short wave infrared camera suite of partial least square regression model fit statistics through 500 permutations of a randomly sampled cross validation procedure for dry matter prediction. Upper left panel contains the model fit for predicted vs observed dry matter on 359 samples. Upper right panel displays R^2 distributions, lower left panel displays root mean squared error (RMSE) distribution, and lower left panel contains bias in the scale of trait units. 8 latent variables were used to construct the model.

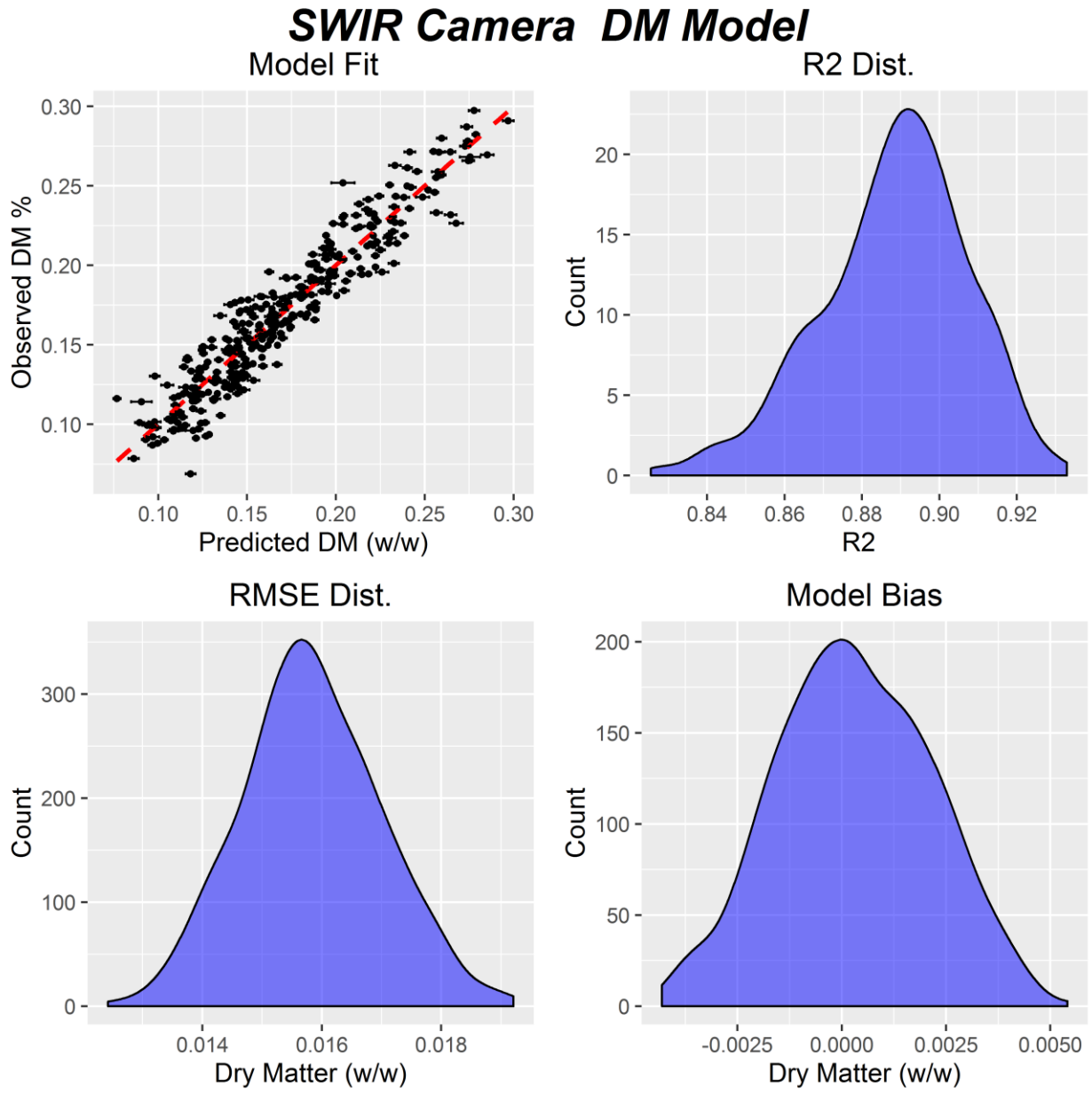


Figure 4.6. Mean variables important to the projection (VIP) and standardized coefficients (STD COEFF) for a dry matter prediction model constructed with rough slab tissue, 13 latent variables, and 700-1800 nm spectra across 500 permutations of a randomly sampled cross validation procedure. Vertical dotted lines represent absorption features as reported by Curran, 1989 for starch (red) and water (blue). VIP above VIP threshold are proposed to be more meaningful to prediction model. Standardized coefficients indicate the positive or negative relationship of reflectance value for trait predication.

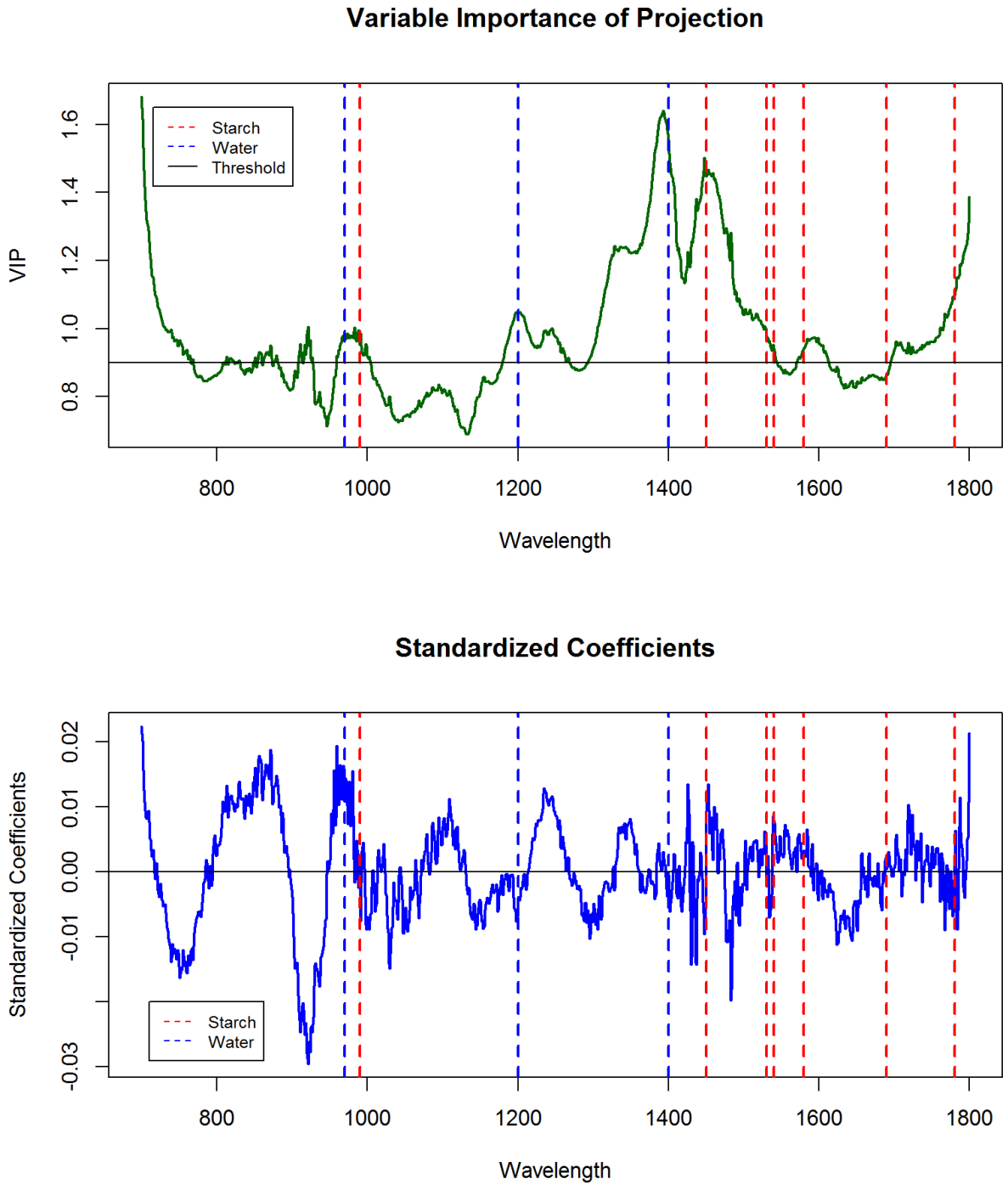
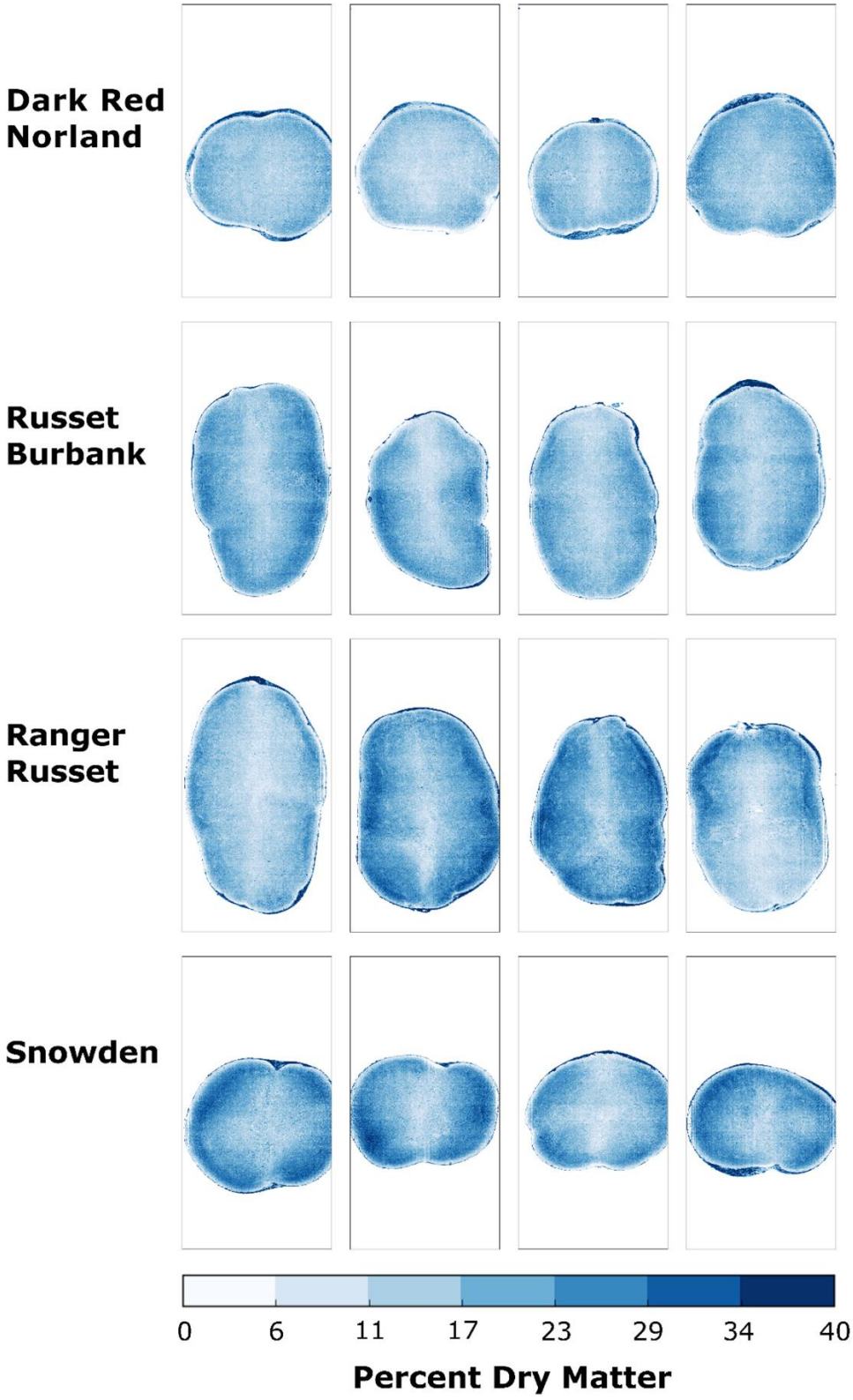


Figure 4.7. Images of 4 tuber slabs from 4 potato cultivars scaled to predicted dry matter as dictated by the average coefficients of the prediction equation based on 8 latent variables. This is a spatial visual representation of the prediction model developed by sampling tissue within the images.



Chapter 5: Summary and Perspectives

Breeding for chip processing traits in potato focuses on the end users; both the processor and the consumer. The aim of the research described herein is to improve our understanding of traits contributing to chip processing quality and provide potential molecular and phenotypic tools for use in breeding chipping potatoes. Chip processing cultivars must produce tubers with high dry matter contents, a desirable cooking attribute, and should be able to produce high quality products after long periods of storage to meet the continual needs of a chip processing plant. Cool storage temperatures, in the range of 4-7 ˚C, contribute to long-term storage because cool temperatures limit losses due to disease and weight loss due to respiration. However, cool temperature storage increases the net conversion of starch into reducing sugars in a process known as cold induced sweetening (CIS) (Burton 1969)(Sowokinos 2001). This accumulation of reducing sugars is undesirable for chip and fry processors because it promotes the formation of dark-colored pigments during cooking as a result of the Maillard reaction, a non-enzymatic reaction that occurs between glucose or fructose and free amino acids at high temperatures (Habib & Brown 1957). Because of the impact that CIS has on the quality of stored potatoes, this dissertation focused on increasing our understanding of this important biological process. Another trait that it related to the conversion of starch to sugar, and one that has become important to the chip processing industry, is stem-end chip defect (SECD). This abiotic-stress-induced trait causes an unacceptable finished chip product by creating a dark discoloration after frying at the location corresponding to the stem end of the tuber. The reliable expression of this trait in the population studied during this project allowed me to uncover some of the genetic loci contributing to this trait. My focus on chipping quality led me to evaluate novel tools for the quantification of tuber solids content and sugar concentration. Thus, a portion of this dissertation describes experiments to determine if hyperspectral reflectance data can be used to predict these traits in the context of raw tissue from a processing potato population.

One of the most notable varieties released in the modern breeding period is the cultivar Lenape (Akeley et al. 1968). In chapter one, I highlighted data showing that Lenape had markedly better performance for the traits of tuber solids and cold sweetening resistance than previous cultivars. Cultivars released after Lenape varied in performance, but varieties that were related to Lenape tended to be among the best in these regards (Figures 1.2 and 1.3)(Love et al. 1998). Here we evaluated a population developed from a cross between Wauseon and Lenape, a prominent cross used in the development of modern chipping germplasm, to begin to identify areas of the genome that confer favorable chip processing traits. We also wanted to assist others in the identification of specific regions on parental homologous chromosomes that contain genes that might be used by breeders in developing improved varieties.

Chapter two presents the results from an experiment in which tubers from this population were exposed to an extreme cold storage environment with the aim of understanding initial responses to cold stress. Sucrose, glucose, fructose, citric acid, and malic acid were measured in tuber tissue grown in three field years. There were high broad sense heritability estimates for these biochemical traits. The expression of *VInv* was also quantified using reverse transcription PCR cycle threshold values relative to different combinations of the reference genes. Heritability estimates for this trait were low in part because a large amount of variation was attributed to genotype by year interactions as well as variation between years.

Chapter 2 also describes how TetraploidMap v2.0 was used for marker dosage calling, linkage map construction, and QTL mapping (Hackett et al. 2014). Multiple QTL for chip-processing-related traits were detected in this experiment. Small effect QTL for relative *VInv* expression were identified, but most of those QTL were based on less than the full three years of data. The goal of this part of the project was to identify genomic locations that might contain direct or indirect regulators of *VInv* expression. Unfortunately, the data did not allow us to fully meet this goal, although the QTL identified may provide clues for further research in this area. My findings further reinforce the idea that control of *VInv* expression is complex.

In contrast to the invertase expression study, the sugar and organic acid data from the HPLC resulted in many detected QTL. The largest effect QTL were on chromosome XII where glucose, fructose, and sucrose to reducing sugar ratio QTL co-located. This region has been identified previously as an area of the potato genome associated with late chip color out of storage (Rak et al. 2017). This area maps within 2.5 Mb of an invertase inhibitor (Sharma et al. 2013) and, as described in chapter three of this project, changes in chip color a* value between January and March and January and May. Thus, it may be that this region of the potato genome plays an important role in sugar metabolism as it influences chip color stability.

Chapter three evaluated the same population used for the research described in chapter two, but in this case harvested tubers were stored using commercial storage practices. In this study, the entire population was chip processed at three time points that spanned the storage season in each of 4 years. Chip color was measured at all time points and specific gravity was measured once in January. Many QTL were identified for chip color traits at the different time points and for changes in chip color between time points. The study described in chapter three was originally devised to corroborate the QTL described in chapter two, and this occurred in places on chromosomes I, XI, and XII. The QTL located on chromosome XII are interesting because the short term cold stress experiment measured biochemical traits and the QTL from the chip processing experiment were for the slope of chip color a* value between January and May and between January and March. Additional QTL for chip color traits were identified that did not collocate with QTL associated with short term sugar accumulation during cold stress. This finding provides more evidence that chip color traits are quantitative and that storage environment plays a significant role in their expression.

Chapter three also highlights the importance of observation while evaluating a genetic population. In this chapter, I discussed stem-end chip defect (SECD) and its importance to growers and processors. SECD is a trait related to abiotic stress resistance, particularly heat stress during tuber bulking (Dickman 2016)(Wang et al. 2012). Tuber tissues that display strong SECD symptoms also have increased levels of

VInv activity (Wang et al. 2012). The original plan for phenotyping the Lenape x Wauseon population did not include the scoring of SECD, but I observed that the population displayed a reliable segregation pattern for this trait throughout the duration of this study. We identified several QTL for SECD; some of these collocated with chip color QTL but others mapped to locations on linkage groups that did not contain QTL related to chip processing. This portion of the research was notable as it was the first to detect QTL for SECD. The QTL found on chromosome III were located near an invertase gene and these QTL were of sizeable effect. There was a significant effect associated with the dosage of many of the SECD QTL markers. This study offers a starting point for future research investigating SECD and its regulation.

Chapter four occupies the field of study known as phenotyping. In plant breeding and genetic studies, phenotyping is an important aspect of research. In this chapter I discuss how I began to apply hyperspectral reflectance techniques to potato breeding and research. This research topic was influenced by my experience with measuring sugar content, as explained in chapter two. The extraction procedure for the measurement of sugars and organic acids was time consuming and required tedious supervision of HPLC processes to ensure high quality data. The measurement of dry matter, whether through the method presented in chapter two using fresh weight and dried weight of tuber tissue, or through the measurement of specific gravity used in chapter three, had many shortcomings. The dry matter measurement described in chapter two was time consuming and the specific gravity measurement in chapter three obtained mean density of many tubers from a plot. The HPLC method, and the methods for measuring dry matter, were unable to easily quantify both within plot variation among tubers and within tuber variation.

As an alternative to conventional methods, I evaluated hyperspectral reflectance for quantification of sugars and dry matter content of raw potato tissue. In the first portion of chapter four, a hyperspectral radiometer was used to collect reflectance spectra from a large set of genotypes taken from the Wauseon x Lenape population, as well as some related cultivars. A tuber from each individual genotype was dissected in different ways and these preparations of raw potato tissue were used to collect spectral data.

Corresponding tissue samples were used to measure sucrose, glucose, fructose and dry matter content. These data were used to create predictive models using partial least squares regression modeling. A randomly permuted cross validation approach tested model robustness using different parameters. The highest prediction accuracy was for the trait of dry matter. In this context, using a subset of the wavelength range increased the model accuracy for dry matter as it excluded spectral data with low relevance to the chemical properties of the trait. The best tissue type for prediction was a rough surface cut slab but good prediction accuracies were attained for several sampling types. The sugar predictions were not as accurate as the dry matter predictions, but could differentiate between high, medium, and low concentrations. This study provides evidence that this technology can be used on a broad set of genotypes to accurately predict dry matter.

In the second portion of chapter four, I applied the results from the first portion to a hyperspectral reflectance imaging system for dry matter prediction. This was possible because the wavelength subsets and tissue types that created accurate prediction models had been identified and because Middleton Spectral Vision donated use of their spectral imaging platform. In this study, 15 slabs from each of 4 varieties were imaged. After image capture, tissue samples were collected from slabs and the dry matter values were combined with spectral data to create a highly accurate prediction model for dry matter content. The coefficients from the model were then applied to each pixel in each of the slab images to create dry matter maps of each tuber slab. While this portion of the chapter included just four cultivars, there was visible variation within and between tubers for dry matter content in the images (Figure 4.7). The accuracy achieved by this system makes it a promising tool for industrial applications as well as research.

This dissertation and the results in each of the chapters are most relevant to a cross made between Wauseon and Lenape. This has advantages in understanding the contributions made by Lenape to modern chip processing varieties. In the single marker analyses performed in chapters two and three, there are cases where the superior trait value was associated with the genotype found in Lenape (Figures 2.6 and

3.7). However, the precise allele from a parental homologue that conferred beneficial attributes at each locus was not identified due to combinatoric complexities. TetraploidMap provided output that contained genotype effects consisting of allelic combinations for progeny at the estimated QTL position. This is a potential starting point for identifying alleles that confer beneficial attributes for each trait. However, it would be important to consider dominance and over-dominance effects in this analysis. Importantly, the effect of genotype classes may be poorly estimated due to low numbers of progeny in certain genotypic classes. While this is an important area to analyze to identify influential alleles, a more robust validation of allelic effects requires information from other related individuals derived from Lenape that contain varying copies of alleles at loci important to the processing traits.

Some of the QTL identified in this study are more valuable for future research aims than others. The SECD trait QTL could identify important genomic regions for physiologists to understand trait development and for breeders to use in new variety development. Identifying parental alleles at the mapped loci that contribute to the trait is an important goal. In the case of SECD, there are several ways to achieve this. One method is to utilize a dataset developed by the Bethke Potato Post Harvest Physiology Lab at University of Wisconsin – Madison. Multiple recently released cultivars and genetic lines were evaluated under controlled environment conditions to evaluate the effect of heat stress on SECD development across several years of greenhouse studies. From this evaluation, data are available for chip color, SECD ranking, tuber counts, and tuber weight for cultivars grown with a short period of heat stress and for unstressed controls. The majority of these lines have been genotyped using the same SOLCap Infinium genotyping platform that was used for this research and this information could be used to determine which markers or alleles located within the QTL identified in chapter 3 had the highest association with SECD in the context of these varieties.

RNA sequencing might also be used to observe relative expression profiles between stressed and unstressed plants within a genotype using the same controlled stress environment made possible by the UW Biotron facility. Suggested genotypes would include individuals from the Wauseon x Lenape

population segregating for SECD while taking care that different combinations of SECD QTL and marker dosages were included. This study might be more informative if other cultivars known to be both susceptible and resistant to SECD were included in the analysis. Another, more extensive method, would be to perform GWAS on a large set of tetraploid cultivars evaluated under field conditions shown to induce SECD symptoms (Dickman 2016)(Wang et al. 2015). This approach is problematic for small effect QTL or if alleles that confer SECD severity are rare in the population of individuals in the study. Finally, an improvement for all types of SECD research could be achieved if scoring was performed not on a subjective scale but rather by a computer vision system. There has been software development in this area for chip color and defects in potato chips (Mendoza et al. 2007)(Pedreschi et al. 2006) and if this were applied to a QTL mapping study, it could more clearly define traits, reduce noise in the data, and permit repeatable measurements between researchers.

The QTL detected in the cold storage experiment described in chapter two are also worth investigating further. The QTL detected on chromosome XII had large effects for reducing sugar accumulation and is worth investigating in other individuals. Interestingly, there was overlap between this QTL for sugar accumulation in cold storage and a QTL for change in chip color during storage described in chapter three. Many of the QTL relating to sugars and their accumulation in tubers are a valuable reference for further investigation into this process. This could be further augmented if markers from previous research studies performed using older marker technology were anchored on the reference genome. This would enable comparisons of discovered QTL for CIS related traits and provide insight into the most commonly discovered region.

A promising direction for future research involves further development of the hyperspectral systems evaluated in chapter four. As mentioned in the introduction of chapter four, global equations have been developed for hyperspectral systems to improve routine measurement of silage and grain quality traits. In these systems, dry matter, protein content, and starch type are often estimated using equations developed over a broad set of samples. Developing a global equation for potato would allow for both the

probe spectroradiometer and the imaging system to be turned into a tool for both genetic studies and selection. It is also worth investigating the accuracy of partial least square regression model predictions for other compounds commonly found in potato, such as protein or forms of starch, that affect nutrition and cooking properties of the crop.

An area to pursue in the dry matter prediction models is subset selection. There are potential improvements to be made by utilizing the VIP values and selecting wavelength subsets where the VIP exceeds the threshold (Gosselin et al. 2010). It could be worthwhile to spend considerable effort in defining a VIP threshold for these models. Using the VIP threshold might allow for a large reduction in the wavelengths included in the prediction models. With selected wavelengths, reductions in instrumentation and processing costs might permit the use of this technology on a processing line for the chip and French fry industry. This would also permit wider use in research on potato dry matter in fresh tuber tissue.

There will be a continual need for improved methods in breeding. The research presented in this dissertation has provided new information about processing potatoes for the chip market segment and the traits important to cultivars contained in this class. Much of the importance of this work is dependent on the parents of this population and the importance of Lenape in many other commercial cultivars. The presented research encompassed areas of genotypic and phenotypic advancement. The QTL detected for SECD was the first for this trait. There were sizable QTL detected for sugar development in cold storage, and multiple QTL detected for chip color out of storage. These findings provide information for additional research focused on identifying the alleles that confer trait expression. The research on hyperspectral prediction provided promising new avenues of research and potentially new tools for dry matter studies. Overall, the information obtained from this thesis project has helped to increase the level of understanding of the chip processing potato and, hopefully, it also assisted in the development of tools that aid future research in this crop.

5.1 References

Akeley RV, Mills WR, Cunningham CE, Watts J. 1968. Lenape: A new potato variety high in solids and chipping quality. Am Potato J. 45:142–145.

Bernardo R. 2010. Breeding for Quantitative Traits in Plants. Second. Woodbury, Minnesota: Stemma Press.

Bhaskar PB, Wu L, Busse JS, Whitty BR, Hamernik AJ, Jansky SH, Buell CR, Bethke PC, Jiang J. 2010. Suppression of the vacuolar invertase gene prevents cold-induced sweetening in potato. Plant Physiol. 154:939–948.

Burton WG. 1969. The sugar balance in some British potato varieties during storage. II. The effects of tuber age, previous storage temperature, and intermittent refrigeration upon low-temperature sweetening. Eur Potato J. 12:81–95.

Dickman L V. 2016. Stem end chipping defect incidence and severity in potatoes (*Solanum tuberosum*). Madison, Wisconsin: University of Wisconsin-Madison.

Falconer DS, Mackay T. 1996. Introduction to quantitative genetics. Fourth Edi. Essex, England: Pearson Education Limited.

Gosselin R, Rodrigue D, Duchesne C. 2010. A Bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications. Chemom Intell Lab Syst. 100:12–21.

Habib AT, Brown HD. 1957. Role of reducing sugars and amino acids in the browning of potato chips. Food Technol. 11:85–89.

Hackett C, Bradshaw JE, Bryan GJ. 2014. QTL mapping in autotetraploids using SNP dosage information. Theor Appl Genet. 127:1885–1904.

Hirsch CN, Hirsch CD, Felcher K, Coombs J, Zarka D, Van Deynze A, De Jong W, Veilleux RE, Jansky S, Bethke P, et al. 2013. Retrospective view of North American potato (*Solanum tuberosum L.*) breeding in the 20th and 21st centuries. G3 Genes|Genomes|Genetics. 3:1003–13.

Love S, Pavek J, Thompson-Johns A, Bohl W. 1998. Breeding progress for potato chip quality in North American cultivars. Am J Potato Res. 75:27–36.

Mendoza F, Dejmek P, Aguilera JM. 2007. Colour and image texture analysis in classification of commercial potato chips. Food Res Int. 40:1146–1154.

Pedreschi F, León J, Mery D, Moyano P. 2006. Development of a computer vision system to measure the color of potato chips. Food Res Int. 39:1092–1098.

Rak K, Bethke PC, Palta JP. 2017. QTL mapping of potato chip color and tuber traits within an autotetraploid family. Mol Breed. 37:15.

Rosyara UR, De Jong WS, Douches DS, Endelman JB. 2016. Software for genome-wide association studies in autopolyploids and its application to potato. Plant Genome. 9:1–10.

Schreiber L, Nader-Nieto a. C, Schonhals EM, Walkemeier B, Gebhardt C. 2014. SNPs in genes functional in starch-sugar interconversion associate with natural variation of tuber starch and sugar content of potato (*Solanum tuberosum L.*). G3 Genes|Genomes|Genetics. 4:1797–1811.

Sharma SK, Bolser D, de Boer J, Sønderkær M, Amoros W, Carboni MF, D'Ambrosio JM, de la Cruz G, Di Genova A, Douches DS, et al. 2013. Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps. G3 (Bethesda). 3:2031–47.

Sowokinos JR. 2001. Biochemical and molecular control of cold-induced sweetening in potatoes. Am J Potato Res. 78:221–236.

Wang Y, Bethke PC, Drilias MJ, Schmitt WG, Bussan AJ. 2015. A multi-year survey of stem-end chip defect in chipping potatoes (*Solanum tuberosum L.*). Am J Potato Res. 92:79–90.

Wang Y, Bussan AJ, Bethke PC. 2012. Stem-end defect in chipping potatoes (*Solanum tuberosum L.*) as influenced by mild environmental stresses. Am J Potato Res. 89:392–399.

Wiberley-Bradford AE, Busse JS, Jiang J, Bethke PC. 2014. Sugar metabolism, chip color, invertase activity, and gene expression during long-term cold storage of potato (*Solanum tuberosum*) tubers from wild-type and vacuolar invertase silencing lines of Katahdin. BMC Res Notes. 7:801.

Zhu X, Richael C, Chamberlain P, Busse JS, Bussan AJ, Jiang J, Bethke PC. 2014. Vacuolar invertase gene silencing in potato (*Solanum tuberosum L.*) improves processing quality by decreasing the frequency of sugar-end defects. PLoS One. 9.