MOLECULAR CHARACTERIZATION, GENOME-WIDE ASSOCIATION STUDIES AND GENOMIC SELECTION OF ADVANCED POTATO CLONES FROM THE TEXAS A&M POTATO BREEDING PROGRAM

A Dissertation

by

JEEWAN PANDEY

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	M. Isabel Vales
Committee Members,	Patricia E. Klein
	Oscar Riera-Lizarazu
	William L. Rooney
Interim Head of Department,	Patricia E. Klein

December 2021

Major Subject: Horticulture

ABSTRACT

Potato (*Solanum tuberosum* L.) is the world's fourth most important crop after maize, rice, and wheat. The Texas A&M University (TAMU) Potato Breeding Program is working to produce early maturing russet, specialty, chipping, and red varieties suited to Texas growing conditions. The increasing urgency to develop superior performing varieties to meet the needs of the fresh and processing potato market necessitates the validation and application of new approaches in potato breeding, such as genome-wide association studies (GWAS) and genomic selection (GS). Recent advances in the development of high-throughput genotyping platforms and whole-genome coverage and affordability have turned single nucleotide polymorphisms (SNPs) into one of the most promising tools for the investigation of genetic diversity and application of GWAS and GS in potato. Two hundred fourteen advanced clones selected over forty years were studied to assess the genetic diversity, population structure, linkage disequilibrium (LD), detect signatures of selection, and find marker trait associations for tuber morphology using the Illumina Infinium 22 K V3 Potato Array. Likewise, 384 unique chipping potato clones were used to evaluate GS for chipping quality in tetraploid potatoes. Results showed that most of the potato clones had high levels of heterozygosity, ranging from 0.22 to 0.80 with a mean of 0.59. Three groups of tetraploid clones, primarily based on potato market classes, were detected using the STRUCTURE software. The highest coefficient of differentiation observed between the groups was 0.14. Signatures of selection were uncovered in genes controlling potato flesh and skin color, length of

plant cycle and tuberization, and carbohydrate metabolism. The GWAS experiment yielded putative and novel markers/genes involved with tuber morphology traits. Genomic prediction using fry color, chip quality, specific gravity and yield gave reliabilities of 0.52, 0.17, 0.40 and 0.11, respectively. Even for limited reference populations and traits with low heritability, the accuracies found were encouraging. The comprehensive molecular characterization will help to better understand the genetic diversity of existing potato resources. Results from GWAS and GS will be helpful in increasing breeding efficiency.

DEDICATION

I dedicate this work:

In memory of Professor J. Creighton Miller, Jr (March 6, 1940 – November 3, 2019) who founded the Texas A&M Potato Breeding Program.

To my parents Tulsi Bhakta Pandey and Sharda Pandey, my wife Indu Dhakal and my sisters Durga Pandey and Sabita Pandey for their endless love, support, and encouragement.

ACKNOWLEDGEMENTS

I am grateful to my committee chair, Dr. Isabel Vales, for the unwavering support and excellent guidance over the years. I would like to thank her for always encouraging me to strive for excellence and to pay attention to details. I would like to express gratitude to my committee member Dr. Patricia Klein for providing me the constant guidance, opportunities, and counsel I need to succeed in the program. I am also indebted to my advisory committee members Dr. Oscar Riera-Lizarazu and Dr. William Rooney for helping me to improve the quality of my research by providing thorough feedback.

My sincere thanks go to research scientists Douglas and Jeff, research associates Angel and Mythreyi, my lab mates Sanjeev and Sam, research technician Mike and program aide Ao for their gracious help in the lab, in the field, and in the greenhouse. I also would like to thank student workers Ruth, Preston, Nina, Megan, Ganga, Hannah, and Brianna for their help with the tissue culture, DNA extractions, and phenotype data collections. The field evaluations were possible thanks to the generous contribution of research sites: Barrett Farms at Springlake and CSS Farms at Dalhart, Texas.

I thank Dr. David Douches and Joseph Coombs at Michigan State University, Dr. Jeffrey Endelman at University of Wisconsin-Madison, and other collaborators for providing software, tools, equipment and for hosting me for several useful training sessions. In addition, I would like to offer my special thanks to Dr. David Reed. It was an extremely rewarding experience working with him as a teaching assistant for HORT

201 and HORT 202. I learned a lot about visual teaching style, student engagement, and other valuable teaching strategies under his guidance. Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to my mother, father, and sisters for their encouragement and to my wife for her patience and love.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Associate Professor M. Isabel Vales, Professor Patricia E. Klein, Associate Professor Oscar Riera-Lizarazu of the Department of Horticultural Sciences and Professor William L. Rooney of the Department of Soil and Crop Sciences. USDA-ARS Centers (Aberdeen, Idaho; Beltsville, Maryland; and Prosser, Washington) and State Universities (Colorado State University, University of Wisconsin, University of Minnesota, North Dakota State University, Michigan State University, Oregon State University, North Carolina State University, Cornell University, University of Maine) potato breeding programs contributed selected potato clones and unselected families to the TAMU Potato Breeding Program. All the work conducted for the dissertation was completed by the student with the technical assistance from Texas A&M potato breeding team under the supervision of the advisor and committee members.

Funding Sources

Graduate study was supported by an Excellence Fellowship from Texas A&M University College of Agriculture and Life Sciences, teaching assistantship from the Department of Horticultural Sciences and an academic scholarship from the Potato Leadership, Education and Advancement Foundation. This work was also made possible in part by USDA-NIFA under Grant Number 2019-03814 and Potatoes USA under National Chip Program (NCP). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the USDA-NIFA.

NOMENCLATURE

AIC	Akaike Information Criterion
BLUE	Best Linear Unbiased Estimates
BLUP	Best Linear Unbiased Prediction
CIP	International Potato Center
CIS	Cold-induced Sweetening
DAL	Dalhart
DAPC	Discriminant Analysis of Principal Components
EHH	Extended Haplotype Homozygosity
gBLUP	Genomic best linear unbiased prediction
GBS	Genotyping by Sequencing
GCA	General Combining Ability
GEBV	Genomic Estimated Breeding Value
GS	Genomic Selection
GWAS	Genome Wide Association Study
IHN	Internal Heat Necrosis
LD	Linkage Disequilibrium
MAF	Minor allele frequency
MAS	Marker-assisted Selection
MCMC	Markov Chain Monte Carlo
MME	Mixed Model Equation

MLM	Mixed Linear Model
NCPT	National Chip Processing Trial
PGSC	Potato Genome Sequencing Consortium
PIC	Polymorphic Information Content
QTL	Quantitative Trait Locus
SCA	Specific Combining Ability
SFA	Snack Food Association
SNP	Single Nucleotide Polymorphism
SPR	Springlake
TAMU	Texas A&M University
USDA	United States Department of Agriculture

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1. Potato: A major food source

Potato (*Solanum tuberosum* L.) is the world's fourth most produced food crop after maize, rice, and wheat. According to FAOSTAT (2021), 370.4 million tonnes of potatoes were grown in 17.3 million ha in 2019. This represents a significant improvement (11%) over 2010's production of 333.6 million tons. China is the world's top potato producer, followed by India, Russia, and Ukraine. The United States is the fifth largest producer of potatoes in the world (NPC, 2019).

Potatoes, the top vegetable crop in the United States, are grown commercially in 30 states. Total potato production in the United States amounted to 19.2 million tonnes in 2019 (FAOSTAT, 2021). Idaho grows more potatoes than any other state, followed by Washington, North Dakota, Wisconsin, and Colorado (USDA NASS, 2019). Even though the area devoted to potato production in Texas is not large (approximately 8000 ha.), growers can harvest and provide fresh potatoes to the market earlier than other states, and often receive two to three times higher prices than growers from Northern States (USDA NASS, 2019).

Potato is an essential crop and is recommended as a food security crop by the United Nations (FAO, 2009; Devaux et al., 2014). Potatoes together with rice, wheat, and corn account for half of the world's food energy requirements (FAO, 2021). The potato's significance may be derived from its esteemed place in the global diet as well as its economic significance for growers, processors, packers, and retailers (Ortiz & Mares, 2017; Wijesinha-Bettoni & Mouillé, 2019). About 63 percent of potato sales are destined

for the processing market: French fries, chips, dehydrated potatoes, and other potato products. The remainder goes to the fresh market, as feed for farm animals, or used as seed tubers to grow the next season's crop (USDA NASS, 2019). In the past, people consumed much of their potatoes fresh. Processed potatoes, such as French fries and hash browns, have increased in popularity since the 1950s, as the technology to freeze vegetables has improved (Keijbets, 2008). Likewise, potato chips are becoming the most consumed snack universally due to changes in consumer's lifestyles and food consumption preferences (Liyanage et al., 2021).

Potatoes is an ingredient in many dishes. Potatoes are nutrient-rich, contributing macro and micronutrients to the diet. They are especially high in vitamin C, potassium, and dietary fiber (Beals, 2019). Some animal studies and some human research studies suggest that potato consumption can have a positive impact on chronic disease risk factors by reducing blood pressure, blood lipids, and inflammation (McGill et al., 2013). Potato is also recognized as a functional food for athletes because it is rich in carbohydrates and nutrients of high quality (Kanter & Elkin, 2019). Thus, potato is one of the most important food crops and is becoming increasingly popular in the world due to its incomparable nutritional value, rising consumption, and economic significance.

1.2. Potato species and origin of cultivated potatoes

The Solanaceae family comprises approximately 90 genera including 3,000-4,000 species. Several members of this family have been important in human civilizations as food sources (potato, tomato, pepper, eggplant, tamarillo), ornamentals (petunia, *Datura*), and sources for drugs (tobacco, *Atropa*) (Gebhardt, 2016). Potato

belongs to the genus *Solanum*. Tuber-bearing species of *Solanum* are grouped in the *Petota* section which contains over 100 species (Spooner, 2009), about 70 percent of which are diploid, and the rest are triploids, tetraploids, pentaploid, and hexaploids. *S. tuberosum* L. includes thousands of varieties that vary by size, shape, color, and sensory characteristics. Commonly cultivated potato varieties are mostly tetraploid (2n = 4x = 48), with a basic chromosome number of x = 12. The potato was domesticated about 8,000 years ago around Lake Titicaca, which borders modern-day Bolivia and Peru, but potatoes have a wide-ranging center of diversity including Venezuela, Colombia, Ecuador, Argentina, Chile, across the Pampa and Chaco regions of Argentina, Uruguay, Paraguay, and southern Brazil and northward into Central America, Mexico, and the southwestern United States (Hijmans & Spooner, 2001). The wild relatives of potatoes provide a rich, diverse and unique source of genetic variation for potato breeding. They can be utilized in breeding programs as sources of disease/pest resistance, environmental stress tolerance, agronomic traits, and fresh and processing tuber quality.

1.3. Potato breeding

Potato (*Solanum tuberosum* L.) is an herbaceous, dicotyledonous plant usually vegetatively propagated at the commercial level using tuber (modified stem) seed pieces. The life cycle of vegetatively propagated potato plants is divided into five stages: sprout growth, plant establishment, tuber initiation, tuber bulking, and tuber maturation. Potatoes can also be reproduced sexually. Potato fruits (berries) contain botanical seeds, called true seeds (Sarkar, 2008). The potato growing season is typically 90-180 days but

can be as short as 75 days in subtropical lowlands and up to 180 days in the Andean highlands (Bradshaw & Ramsay, 2009).

Hybridization is used by breeding programs to generate genetic variation and to combine, in progenies, traits of interest present in parental clones. During new cultivar development, approximately 40 traits are evaluated (Gebhardt, 2013) that are generally divided into yield and tuber quality characteristics, as well as tolerances to biotic and abiotic stresses. External quality traits required for fresh market and processed potatoes include tuber size and shape, eye depth, skin color, skin texture, and lack of blemishes due to bruising and diseases. Internal quality traits include dry matter content, nutritional quality, flavor, starch quantity and quality, and lack of defects such as vascular discoloration, hollow heart, and internal necrosis (Jansky, 2009). Cultivar development in potatoes is a multistep process, including first, identifying parents with desirable characteristics and crossing them to produce true potato seed (TPS). TPS is then planted typically under greenhouse conditions to produce segregating families of seedling tubers (50,000-100,000) for field planting. Evaluation and selection during the first and second field years are based on tuber morphological traits. Additional evaluation for important agronomic, quality, disease and pest resistance, and abiotic stresses take place in replicated trials over multiple years and locations. After the development of a cultivar profile, potato clones are cleaned *in-vitro*, increased for commercial testing, released as varieties, and protected (plant variety protection or registration). During the cleaning process, promising advanced clones are tested for the presence of virus and bacteria and cleaned if necessary, using thermotherapy and/or chemotherapy to eliminate the

pathogens (Wang et al., 2018). Multiplication includes tissue culture micropropagation in the lab, followed by greenhouse production and subsequent seed increase (production of certified seed) in the field for several years. Finally, market development and promotion of new varieties are necessary to explore grower, consumer, industry, and market retail for the intended use. All these components are essential for the successful development of new potato varieties and market acceptance.

1.4. The Texas A&M Potato Breeding Program

The Texas A&M Potato Breeding and Variety Development Program was established by Dr. J. Creighton Miller Jr. in 1972. Potato production responds to a varied set of environmental conditions, pest problems, and market niches, thus requiring the development of cultivars that are either widely adapted or suited for specific production areas. The goal of the Texas A&M Potato Breeding program is to provide improved cultivars with high yield, quality, disease/pest resistance, and tolerance to high temperatures that will enable producers in the U.S. Southwest region to remain viable and competitive and supply high quality, healthy and nutritious products to consumers. The Texas A&M Potato Breeding program has developed/co-developed and released 17 new cultivars/clonal selections. Some of the new cultivars/clonal selections make up a substantial and increasing share of the regional/national potato acreage and have become important contributors to the economies of several states. Of all the cultivars released over the past 15 years by the 12 US potato breeding programs, those developed by the Texas program have ranked in the top four to five nationally in total acreage approved for seed certification over the past several years (NPC, 2019). This has been due, in large

measure, to the popularity of the four Texas Russet Norkotah selections (112, 223, 278, and 296). This is especially significant since Texas does not have a seed-potato industry.

1.5. Limitations of cultivated tetraploid breeding

Most commercial grown potatoes are autotetraploid (2n = 4x = 48), but diploid (2n = 2x = 24), triploid (2n = 3x = 36), and pentaploid (2n = 5x = 60) potatoes are also grown by farmers in South America (Watanabe, 2015). Improving yield, processing, storage traits, and disease resistance in potatoes through conventional breeding is more difficult than in other major food crops. Phenotypic recurrent selection is the main breeding strategy used in potatoes. Recurrent selection involves obtaining the best clones from the progeny, their assessment, and cross with established varieties and elite clones in a cyclic process. It can take 10-15 years to develop a new potato variety (Slater et al., 2014). This is primarily due to the low rate of seed increase and the need to evaluate materials over several years and locations. Additionally, each multiplication cycle carries the risk of infection with tuber-borne diseases, especially viruses. Another factor influencing to the time to develop a potato variety is the growing environment which can affect many of the characteristics that contribute to a potato plant's phenotype, thus requiring multi-year, multi-location evaluations (Jansky, 2009). Finally, the search for useful genetic variability in wild relatives may be laborious and its introgression in cultivated varieties may be another challenge due to different ploidy levels and self and cross incompatibility issues (Slater et al., 2014). The cultivated potato has a complex genetic structure. Its autotetraploidy and vegetative propagation have led to the accumulation of mutations and a highly heterozygous genome (Watanabe, 2015).

Solanum tuberosum L. is a tetraploid with four homologues, each with 12 distinct chromosomes, resulting tetrasomic inheritance (Bradshaw, 2007). The selfing of a tetraploid with two alleles at a locus (AAaa) gives five different genotypic classes in its offspring: AAAA (quadruplex), AAAa (triplex), AAaa (duplex), Aaaa (simplex), and aaaa (nulliplex) based on the allelic dosage. If an allele has an additive effect, then one locus can result in considerable quantitative segregation. In potatoes, most diploid species are self-incompatible (Kirch et al., 1989) whereas most cultivated tetraploid species are self-compatible. Self-incompatibility in potatoes is controlled by the highly polymorphic *S*-locus (Kao & McCubbin, 1996). Though the tetraploids are compatible, the homozygosity approach is too slow for the development of an inbred line and suffers inbreeding depression upon selfing to be used as a breeding strategy. Major advances in genetics, molecular biology, and statistical genomics are providing breeders with molecular tools and software to increase genetic gain in the conventional potato breeding programs.

1.6. Potato genome, genetic markers, and genetic maps

Most cultivated potatoes are autotetraploids (2n = 4x = 48). Since cultivated potato has a large complex genome, the development of genetic and genomic data has lagged behind other crop species. The whole-genome sequence of potatoes has led to the development of genetic tools and resources for potatoes. A homozygous doubled-monoploid potato clone (DM1-3 516 R44) was sequenced via a whole-genome shotgun approach, providing a reference genome spanning 86% of the 844 Mbp genome and containing 39,031 predicted protein-coding genes (Xu et al., 2011). The sequencing also

identified 755 genes encoding polypeptides representing most of the currently known plant resistance (R) proteins (Bakker et al., 2011; Jupe et al., 2013). Likewise, sequencing and assembly of the genome of a diploid, self-compatible, inbred clone (M6) of *S. chacoense* generated an assembly of 826 Mbp and genome annotation yielded high-confidence gene models representing 37,740 genes (Leisner et al., 2018). Genome sequences from cultivated potato covering traditional landraces including modern cultivars would be a powerful resource to understand the genome structure of potato (Ortiz, 2020). However, the sequencing difficulties arise due to the existence of four genome sequences resulting from autopolyploidization and repeated introgression of wild chromosome segments (Rodríguez et al., 2010). Recently, six cultivated varieties have been sequenced (Potato pan-genome consortium) and genome sequences were successfully assembled from modern potato cultivars (NRGene at http:/www.nrgene.com). Pan-genome analysis and NRGene's genomes will help to map traits on the level of haplotypes. Many molecular methods have been developed in the past to detect natural DNA variations, beginning with a hybridization-based analysis of restriction fragment length polymorphisms (RFLPs) (Botstein et al., 1980). Thereafter, research advanced to PCR-based marker systems such as microsatellite (SSR) markers, RAPD markers (Welsh & McClelland, 1990), AFLP markers (P. Vos et al., 1995), and cleaved amplified polymorphic sequences (CAPS) markers (Konieczny & Ausubel, 1993). Sequencing techniques, such as Sanger amplicon sequencing (Sanger et al., 1977), then permitted the identification of single nucleotide polymorphisms (SNPs). SSRs and RAPDs have been used in potatoes for varietal identification and genetic

diversity assessments (Jamali et al., 2019). SNPs are increasingly being used due to their biallelic nature, high throughput, and abundance in most crop plants. Methods for minimizing genome complexity, such as genotyping by sequencing (GBS), restriction site-associated DNA sequencing (RADseq), and diversity array technology (DArTseq) have been developed to help discover new SNPs; however, the detection of SNP markers in polyploids is often difficult due to the repetitive segments of the genome and multiple ploidy levels (Bertioli et al., 2014). The development of potato markers and genetic map construction is mainly performed at the diploid level, using traditional quantitative trait loci (QTL) mapping. Relatively few software tools exist that can perform linkage mapping for polysomic species. TetraploidMap is perhaps the most popular and widely used software tool for Windows (Hackett et al., 2017). Other polyploid mapping software recently in use are the PERGOLA package in R (Grandke et al., 2017), polymapR (Bourke et al., 2018), and MAPpoly (Mollinari & Garcia, 2019). These advancements have aided in the identification of QTL that accounts for significant amounts of phenotypic variance within a polyploid population.

The SolCAP Infinium Potato Array has been used to build linkage maps (Felcher et al., 2012), genotype populations for QTL analysis (Douches et al., 2014), and to determine variability in the biosynthesis of glycoalkaloids (Manrique-Carpintero et al., 2014). To date, SNPs have been successfully used to detect novel QTLs and candidate SNPs associated with late blight resistance (Lindqvist-Kreuze et al., 2014; Massa et al., 2015), internal heat necrosis (Schumann et al., 2017), Potato virus Y (da Silva et al.,

2017), agronomic traits (Manrique-Carpintero et al., 2018), and processing quality (Massa et al., 2018).

1.7. Potato SolCAP SNP array

A single nucleotide polymorphism (SNP) is a variation at a single position in the DNA sequence. SNPs have rapidly gained the center stage of molecular genetics in recent years due to their abundance in genomes, biallelic nature, and their amenability for high-throughput detection formats and platforms. A consortium led by SolCAP (http://solcap.msu.edu/) identified 8,303 SNP markers from the commercial potato cultivars Atlantic, Premier Russet and Snowden (Hamilton et al., 2011). To validate these SNPs, Felcher et al. (2012) developed linkage maps for two diploid mapping populations (DRH and D84). Both populations used the doubled monoploid reference genotype DM1-3 516 R44 as the female parent but had different heterozygous diploid male parents (RH89-039-16 and 84SD22). Of the mapped markers, 3,018 SNPs (36%) are present in candidate genes; 536 SNPs (6%) are previously identified genetic markers; the remaining 4,749 SNPs (57%) are new SNPs spread across the chromosomes. Two additional public SNP arrays Infinium V2 12K (Hamilton et al., 2011) and V3 20K (Vos et al., 2015) have been developed and are helping potato breeders to incorporate genomic tools to improve the efficiency in potato improvement.

1.8. Tetraploid SNP calling

To determine the allele dosage in tetraploid individuals, one of five possible genotypes must be recognized for each SNP locus. The five different genotypic classes are: AAAA (quadruplex), AAAa (triplex), AAaa (duplex), Aaaa (simplex), and aaaa

(nulliplex). During SNP calling, Genome Studio software (Illumina, 2005) generates SNP theta scores, which are used to determine the allele dosage. Several different methods can be used to transform the continuous signal scores (theta values) to discrete genotype groups, including 1) pre-determined dose cluster calling boundaries (http:/solcap.msu.edu/potato infinium.shtml); 2) mixed models, FitTetra (Voorrips et al., 2011; Hackett et al., 2013) 3) NbClust cluster analysis (Charrad et al., 2014) and 4) ClusterCall (Schmitz Carley et al., 2017). NbClust provides 30 indices to determine the number of clusters while ClusterCall uses hierarchical clustering and multiple F_1 populations to calibrate the relationship between signal intensity and allele dosage.

1.9. Genetic diversity

Among DNA markers, SSR and SNP have been successfully used in polyploid species for the evaluation of genetic diversity. SSR markers were initially preferred because of their random genome distribution, co-dominant nature, high polymorphism, ease of use, high clarity and reproducibility, low operating costs, automation convenience, ease of multiplexing, and low-quantity DNA use (Vieira et al., 2016). Bali et al. (2018) quantified the genetic variability among the Russet potato clones and studied the genetic relationships based on their pedigrees by using SSR markers. Recent advances in the development of high-throughput genotyping platforms have turned SNPs into one of the most promising tools for the investigation of genetic diversity. Assessment of genetic diversity could provide valuable information for the genetic improvement of potatoes. The greater the genetic variability in the base population, the greater the opportunity for breeding. Several studies have been published implementing

the SolCAP Potato SNP Array for genetic diversity studies. Hirsch et al. (2013) retrospectively explored the effects of potato breeding at the genome level using the 8K SNP array. Kolech et al. (2016) used the SNP array to evaluate the genetic diversity of Ethiopian potato cultivars. Vos et al. (2015) developed a 20K SNP array and used it to genotype a total of 569 potato genotypes to identify footprints of the breeding history in contemporary breeding material. Similarly, other diversity studies have been done with the Colombian Central Collection (Berdugo-Cely et al., 2017), International Potato Center (CIP) collection (Ellis et al., 2018), and Japanese potatoes (Igarashi et al., 2019).

1.10. Genome-Wide Association Study (GWAS)

Genome-wide association study (GWAS) is a method for detecting marker-trait associations. GWAS provides the benefit of detecting QTLs within a range of individuals and is likely to achieve high mapping resolution for identifying candidate genes (Zhu et al., 2008). GWAS can be used for the study of complex traits when the development of mapping populations from crosses is not feasible (Gebhardt et al., 2004). An unstructured population in a GWAS contains many more cases of recombination and often traces back to shared ancestors. Therefore, it gives the potential of analyzing a widely genetically diverse population to have a greater genetic resolution (Rafalski, 2010). However, the predictive strength of natural population-based mapping depends heavily on the degree of linkage disequilibrium (LD) and population structure, as well as the sample size and the frequency of minor alleles (MAF). LD is defined as the nonrandom association of two alleles in a population (Flint-Garcia et al., 2003). The faster LD declines, the denser the collection of genome-wide markers must be used to

detect causative loci markers in LD. The ability to detect significant marker-trait associations in GWAS is influenced by the trait's genetic architecture, heritability, the nature of the phenotypic data, and sample size (Cortes et al., 2021). GWAS aims to find SNPs where genotype variation is correlated with phenotype variation. The naive method has a high incidence of false positives, which arise when a result is declared significant when it is not. Two popular methods of multiple testing correction are limiting the false discovery rate (FDR) (Storey & Tibshirani, 2003), or using the Bonferroni correction. When compared to traditional approaches, the implementation of the mixed linear model for GWAS has significantly reduced the number of false positives (Yu et al., 2006). The MLM accounts for relatedness at two levels: population structure (Q) and kinship (K).

GWAS has evolved to be an effective and widely used technique for studying complex traits in major crops. Some of the major highlights include GWAS conducted in maize to study leaf architecture (Tian et al., 2011), husk traits (Cui et al., 2016), and kernel test weight (X. Zhang et al., 2020). In rice, GWAS identified a gene comprehensively controlling rice architecture (Yano et al., 2019), mesocotyl elongation (Wu et al., 2015), and salt tolerance (Yuan et al., 2020). Likewise, Beyer et al. (2019) identified loci and candidate genes controlling root traits in wheat seedlings using GWAS.

Improvements in both marker development technologies and analytical statistical methods have advanced efforts to map quantitative traits in polyploid species. The ability of GWAS to explore more diverse germplasm facilitates the detection of variants

with a clear effect across the discovery population. One challenge in applying GWAS to polyploid species is how to define relatedness (K) between polyploid individuals. Software tools like the GWASpoly R package (Rosyara et al., 2016) and the SHEsisPlus package (Shen et al., 2016) are available for polyploid GWAS. Of these, GWASpoly looks critically at the form of the kinship matrix K. The relationship matrix best mitigated the inflation of significance values and is the default K provided in GWASpoly. Alternatively, netgwas (Behrouzi et al., 2019) provides GWAS mapping for polyploids which uses graphical models to go beyond the mapping of single-marker associations to investigate genotype-phenotype interactions using all markers simultaneously in a graph structure. In potato, marker-trait associations were detected using high-throughput genome-wide markers such as for late blight resistance (Lindqvist-Kreuze et al., 2014), tuber shape and eye depth (Rosyara et al., 2016), protein content (Klaassen et al., 2019), morpho-agronomic traits (Zia et al., 2020), common scab resistance (Kaiser et al., 2020), and potato stolon traits and root traits (Yousaf et al., 2021).

1.11. Population structure analysis

To minimize false-positive and false-negative correlations due to population stratification, GWAS must account for population structure (Teo, 2008). Although a marker may not be linked to a QTL, there is a significant risk that a significant association may be identified only based on the genetic relationship between individuals (Pritchard et al., 2000). There are several methods for evaluating population structure based on genetic markers. The factor analysis approach is principal component analysis

(PCA) (D'hoop et al., 2010), where genotypic information from molecular marker data is processed. Another approach is Bayesian clustering, implemented in the software Structure (Pritchard et al., 2000). Analysis of molecular variance (AMOVA) and hierarchical clustering (D'hoop et al., 2010) are similar approaches for population structure assessment. Likewise, the kinship matrix (K) account for hidden relationships between individuals (Yu et al., 2006).

1.12. The Breeder's Equation and Genetic Gain

The breeder's equation ($R = irσ_A/t$) has long been used to show how genetic response (R) changes in response to selection intensity (i), the square root of the additive genetic variance ($σ_A$), and selection accuracy (r). In genomic selection (GS), r is the correlation between true breeding values (TBVs) and genomic-estimated breeding values (GEBVs), while in the phenotypic selection, r is equal to the square root of the narrow-sense heritability (h) (Heffner et al., 2010). In the denominator, the number of years per cycle (t) measures efficiency by expressing the response to selection as a change over time.

Plant breeders must increase at least one of the three components of the breeder's equation within a given period (t) in a breeding cycle, or reduce the length of the breeding cycle, if they want to increase the genetic gain of their breeding program (Cobb et al., 2019). It can be difficult to maximize genetic gains through phenotypic selection because trait heritability, environmental interactions, and time for trait expression all play a role (Yang Xu et al., 2017). Alternatively, genetic gain can be improved by using

various forms of marker-assisted selection (MAS), being genomic selection (GS) one of the most promising options.

1.13. Marker-assisted selection (MAS) and Genomic Selection (GS)

Marker-assisted selection (MAS) is an indirect selection process where a trait of interest is selected based on a marker linked to that trait (Collard & Mackill, 2008). Traditional MAS can only be used to select for qualitative traits or quantitative traits dominated by a small number of loci and can only be used to select one or a few QTL or gene loci since it uses a limited number of markers (Bernardo, 2008). However, GS jointly estimates all marker effects without significance testing to capture small-effect QTL that is excluded by conventional MAS (Meuwissen et al., 2001; Xu et al., 2020). Many simulations and observational trials have shown the superiority of GS over conventional MAS for the enhancement of complex traits (Krishnappa et al., 2021).

Genomic selection (GS) uses genome-wide single nucleotide polymorphism (SNP) marker data to predict and perform selection based on genomic estimated breeding values (GEBVs) of individuals (Meuwissen et al., 2001). Lines with the highest GEBVs are then chosen as parents or for advancement in a GS breeding program, while lines with the lowest GEBVs are eliminated. Traditional breeding strategies supplemented by GS reduce the need for large-scale phenotyping and accelerate genetic gain by shortening breeding cycles (Heffner et al., 2010). Another significant benefit of GS is that it allows for the selection of low-heritability traits (Heffner et al., 2011; J. Zhang et al., 2016; Klápště et al., 2020) and novel traits that are difficult or costly to test in practice but can be tested on a small reference population (Calus et al., 2013). GS

accuracy depends on genetic composition and size of the training population (TP), level of LD, marker density, trait heritability, the effective number of loci, model performance, and relationship between TP and the validation population (VP) (Rutkoski et al., 2015).

The animal breeding sector pioneered GS and tested it (Meuwissen et al., 2001). Recently, it has been extended to crops, including wheat (Poland et al., 2012; Rutkoski et al., 2015), maize (Zhao et al., 2012; Beyene et al., 2015), rice (Spindel & Iwata, 2018), sorghum (Fernandes et al., 2018), barley (Lorenz et al., 2012), tomato (Duangjit et al., 2016), and potato (Habyarimana et al., 2017; Sverrisdóttir et al., 2017, 2018; Endelman et al., 2018) for identification of superior lines as well as predicting the performance potential. The polyBreedR package (https://github.com/jendelman/polyBreedR) has been recently designed to facilitate the use of genome-wide markers for GS in autotetraploid (4x) species.

1.14. Tuber morphology traits in potato

In the context of potato, morphological traits like tuber shape, eye depth, degree of russeting, tuber number, tuber weight, skin color, and flesh color are essential. Tuber shape is extremely important, and its expectation is based on historical regional preferences coupled with local culinary practices (Stark et al., 2020). Chips are made with round tubers, whereas French fries are made with long tubers. Likewise, the consumer expects table cultivars to be round or oval (van Eck et al., 1994; Chen et al., 2018). Van Eck et al. (1994) identified a single locus, *Ro*, that explains the inheritance of qualitative tuber form, with round being dominant over long on chromosome 10, using

restriction fragment length polymorphisms (RFLPs). Other reports using populations with different genetic backgrounds mentioned QTLs on chromosomes 2, 5, and 11 (Bradshaw et al., 2008), and 2 and 11 (ŚLiwka et al., 2008). Using a highly heterozygous diploid potato population and QTL analysis the major loci segregating for tuber shape was mapped to loci on chromosomes 2 and 10, with smaller effects mapping to three other chromosomes (Prashar et al., 2014). Thus, the findings of tuber shape inheritance whether it is monogenic or polygenic have been inconsistent. The irregular shape and deep eyes in potatoes lead to higher costs due to significant peeling losses (Li et al., 2005; Prashar et al., 2014). Eye depth tends to be dominated by a single locus, which is 4 cM away from the *Ro* locus on chromosome 10 and is closely linked to it (Li et al., 2005; ŚLiwka et al., 2008). Another trait, russeting is an inherited trait in some potato cultivars such as Russet Burbank and Russet Norkotah (De Jong, 1981) but the development of a russeted skin in red-skinned cultivars like 'Magen russeting' is a physiological disorder (Lulai, 2007; Ginzberg et al., 2012). Tuber skin color and flesh color traits are important to consumers and industry contingent the end use of the potato. The presence of carotenoids causes the yellow to the orange coloration of potato tuber flesh (Nesterenko & Sink, 2003). A dominant allele at the *Y* (yellow) locus on chromosome 3 is responsible for the yellow flesh color (Bonierbale et al., 1988). A beta-carotene hydroxylase gene (*BCH*) was found in the same position as the *Y* locus, suggesting that this is the most significant candidate gene for yellow flesh (Thorup et al., 2000; Kloosterman et al., 2010). Anthocyanin pigments accumulate in the tuber flesh, giving it a red or purple color (Eichhorn & Winterhalter, 2005). The *Pf* locus tightly

linked with the *I* locus (encodes a MYB transcription factor) maps to chromosome 10 and confers pigmented tuber flesh (De Jong, 1987). Also, the *R* locus which encodes dihydroflavonol 4-reductase (*dfr*) maps to chromosome 2 (De Jong et al., 2003) and it is essential for red anthocyanin production. For purple pigment formation in the potato skin, the *P* locus which encodes flavonoid 3′,5′-hydroxylase (*f3′5′h*) is required and it maps to chromosome 11 (Jung et al., 2005). The *f3′5′h* gene is expressed in the tuber skin only in the presence of *I* (Jung et al., 2005). Thus, breeding for morphological traits in potatoes is crucial as these traits affect consumption.

1.15. Chipping quality traits in potato

In terms of tuber appearance, the industry prefers shallow eyes, and a round-oval shape for processing chips, whereas long-oval-shaped tubers are favored for processing French fries. Greening, cracking, hollow tubers, secondary damage, rusty spots, and other issues should be avoided. The color of the peel, as well as the color of the flesh, should appeal to consumers' preferences (Nacheva & Pevicharova, 2008; Wayumba et al., 2019). In the case of processing potatoes (chippers and French fries), most flesh is white. The preferred texture of chippers is smooth and of potatoes used for French fries is russet. The starch, dry matter, reducing sugars, proteins, and vitamin C content of tubers determine their post-harvest quality and nutritional value. For processing, high starch and high dry matter-rich cultivars are favored because they cook faster, have a better texture, high net weight of the final product, and use less oil while frying (Habyarimana et al., 2017). Starch content of 13% or more, solids or dry matter content of 20% or more, and/or specific gravity of 1.080 or more are desired for most processing

potatoes (Stark et al., 2020). Another, important trait in potato that determines the quality of processed potato is the reducing sugar content. Potatoes are best stored in low temperature (just above freezing) and high relative humidity (98%) conditions to ensure stable storage. Cold storage conditions eliminate issues related to diseases, sprouting, and shrinking (Kleinkopf et al., 2003). However, cold temperatures promote the breakdown of starch and the accumulation of reducing sugar in potato tubers. This is known as cold-induced sweetening (CIS). Accumulated reducing sugars react with free amino acids when processed at high temperatures, resulting in undesirable brown, bitter-tasting products. Internal defects like internal heat necrosis (IHN), vascular browning, Fusarium wilt, stem end browning, hollow heart with discoloration, brown center, and internal black spot also contribute to browning and thus reduce chip quality. The potato processing industry is, therefore, searching for strategies to reduce browning. These processes are complex and controlled by several enzymes (Leonel et al., 2017). Selection for processing quality after long storage is costly and time-consuming and identification of clones using markers is desirable. Breeding programs have made efforts to develop cultivars that generate less sugar in cold storage, resulting in a lighter color potato chip. Much of the success has been attributed to the Lenape variety, a common parental stock in chip crosses since its introduction (Love et al., 1998). Lenape is a parent of chipping varieties including Atlantic, Trent, Belchip, and Snowden. Lenape had the highest dry matter content and was particularly responsible for a trend of increased dry matter content in newer varieties (Love et al., 1998). QTL mapping in populations of diploid potatoes reported QTL for chip color and reduced sugar accumulation associated with

candidate genes (Li et al., 2008). Recently, Byrne et al. (2020) identified a major QTL on chromosome 10 for fry color and predicted fry color with moderate accuracy using genome-wide markers. They also identified a small subset of SNPs for processing characteristics which can give moderate predictive ability. Likewise, genotyping-by-sequencing developed genomic prediction models were used for starch content and chipping quality (Sverrisdóttir et al., 2017) and dry matter content and chipping quality (Sverrisdóttir et al., 2018).

1.16. The approach of the current study

Reports on the use of advanced clones selected over multiple years for diversity studies, association mapping, and genomic selection remain limited. The current study hypothesized that the Texas A&M University Potato Breeding Program's clone bank collection is diverse and has variation in tuber morphology traits and chipping quality traits that can be used to identify genomic regions associated with these traits. Recent marker-dense platforms are expected to effectively account for relatedness in GWAS models, which will aid in the identification of genomic regions associated with important breeding traits. Finally, successful implementation of GS strategies for the potato breeding program is anticipated. The specific objectives are as follows:

Objective 1: Investigate potato varieties and advanced clones of the TAMU Potato Breeding Program (entered in the clone bank over 40 years of breeding) at the molecular level to assess genetic diversity for further genetic enhancement of important economic traits.

Objective 2: Identify genomic regions and superior individuals for tuber morphology traits in tetraploid advanced potato clones.

Objective 3: Evaluate how well GS can predict chipping quality in tetraploid potatoes.

1.17. References

Bakker, E., Borm, T., Prins, P., van der Vossen, E., Uenk, G., Arens, M., de Boer, J., van Eck, H., Muskens, M., Vossen, J., van der Linden, G., van Ham, R., Klein-Lankhorst, R., Visser, R., Smant, G., Bakker, J., & Goverse, A. (2011). A genome-wide genetic map of NB-LRR disease resistance loci in potato. *Theoretical and Applied Genetics*, *123*(3), 493–508. https://doi.org/10.1007/s00122-011-1602-z

Bali, S., Patel, G., Novy, R., Vining, K., Brown, C., Holm, D., Porter, G., Endelman, J., Thompson, A., & Sathuvalli, V. (2018). Evaluation of genetic diversity among Russet potato clones and varieties from breeding programs across the United States. *PLOS ONE*, *13*(8), e0201415. https://doi.org/10.1371/journal.pone.0201415

Beals, K. A. (2019). Potatoes, nutrition and health. *American Journal of Potato Research*, *96*(2), 102–110. https://doi.org/10.1007/s12230-018-09705-4

Behrouzi, P., Arends, D., & Wit, E. C. (2019). netgwas: An R package for network-based genome-wide association studies. *ArXiv:1710.01236 [q-Bio, Stat]*. http://arxiv.org/abs/1710.01236

Berdugo-Cely, J., Valbuena, R. I., Sánchez-Betancourt, E., Barrero, L. S., & Yockteng, R. (2017). Genetic diversity and association mapping in the Colombian Central Collection of *Solanum tuberosum* L. Andigenum group using SNPs markers. *PLoS ONE*, *12*(3). https://doi.org/10.1371/journal.pone.0173039

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: Learning from the last 20 years. *Crop Science*, *48*(5), 1649–1664. https://doi.org/10.2135/cropsci2008.03.0131

Bertioli, D. J., Ozias-Akins, P., Chu, Y., Dantas, K. M., Santos, S. P., Gouvea, E., Guimarães, P. M., Leal-Bertioli, S. C. M., Knapp, S. J., & Moretzsohn, M. C. (2014). The use of SNP markers for linkage mapping in diploid and tetraploid peanuts. *G3: Genes, Genomes, Genetics*, *4*(1), 89–96. https://doi.org/10.1534/g3.113.007617

Beyene, Y., Semagn, K., Mugo, S., Tarekegne, A., Babu, R., Meisel, B., Sehabiague, P., Makumbi, D., Magorokosho, C., Oikeh, S., Gakunga, J., Vargas, M., Olsen, M., Prasanna, B. M., Banziger, M., & Crossa, J. (2015). Genetic gains in grain yield through genomic selection in eight bi-parental maize populations under drought stress. *Crop Science*, *55*(1), 154–163. https://doi.org/10.2135/cropsci2014.07.0460

Beyer, S., Daba, S., Tyagi, P., Bockelman, H., Brown-Guedira, G., & Mohammadi, M. (2019). Loci and candidate genes controlling root traits in wheat seedlings—A wheat root GWAS. *Functional & Integrative Genomics*, *19*(1), 91–107. https://doi.org/10.1007/s10142-018-0630-z

Bonierbale, M. W., Plaisted, R. L., & Tanksley, S. D. (1988). RFLP Maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics*, *120*(4), 1095–1103.

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, *32*(3), 314–331.

Bourke, P. M., Voorrips, R. E., Visser, R. G. F., & Maliepaard, C. (2018). Tools for genetic studies in experimental populations of polyploids. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.00513

Bradshaw, J. E., Hackett, C. A., Pande, B., Waugh, R., & Bryan, G. J. (2008). QTL mapping of yield, agronomic and quality traits in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*). *Theoretical and Applied Genetics*, *116*(2), 193–211. https://doi.org/10.1007/s00122-007-0659-1

Bradshaw, J. E., & Ramsay, G. (2009). Chapter 1—Potato origin and production. In J. Singh & L. Kaur (Eds.), *Advances in Potato Chemistry and Technology* (pp. 1–26). Academic Press. https://doi.org/10.1016/B978-0-12-374349-7.00001-5

Bradshaw, John. E. (2007). The canon of potato science: 4. Tetrasomic inheritance. *Potato Research*, *50*(3), 219–222. https://doi.org/10.1007/s11540-008-9041-1

Byrne, S., Meade, F., Mesiti, F., Griffin, D., Kennedy, C., & Milbourne, D. (2020). Genome-wide association and genomic prediction for fry color in potato. *Agronomy*, *10*(1), 90. https://doi.org/10.3390/agronomy10010090

Calus, M. P. L., de Haas, Y., Pszczola, M., & Veerkamp, R. F. (2013). Predicted accuracy of and response to genomic selection for new traits in dairy cattle. *Animal*, *7*(2), 183–191. https://doi.org/10.1017/S1751731112001450

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, *61*(1), 1–36. https://doi.org/10.18637/jss.v061.i06

Chen, N., Zhu, W., Xu, J., Duan, S., Bian, C., Hu, J., Wang, W., Li, G., & Jin, L. (2018). Molecular marker development and primary physical map construction for the tuber shape *Ro* gene locus in diploid potato (*Solanum tuberosum* L.). *Molecular Breeding*, *39*(1), 6. https://doi.org/10.1007/s11032-018-0913-z

Cobb, J. N., Juma, R. U., Biswas, P. S., Arbelaez, J. D., Rutkoski, J., Atlin, G., Hagen, T., Quinn, M., & Ng, E. H. (2019). Enhancing the rate of genetic gain in public-sector plant breeding programs: Lessons from the breeder's equation. *Theoretical and Applied Genetics*, *132*(3), 627–645. https://doi.org/10.1007/s00122-019-03317-0

Collard, B. C. Y., & Mackill, D. J. (2008). Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1491), 557–572. https://doi.org/10.1098/rstb.2007.2170

Cortes, L. T., Zhang, Z., & Yu, J. (2021). Status and prospects of genome-wide association studies in plants. *The Plant Genome*, *14*(1), e20077. https://doi.org/10.1002/tpg2.20077

Cui, Z., Luo, J., Qi, C., Ruan, Y., Li, J., Zhang, A., Yang, X., & He, Y. (2016). Genome-wide association study (GWAS) reveals the genetic architecture of four husk

traits in maize. *BMC Genomics, 17*(1), 946. https://doi.org/10.1186/s12864-016-3229-6

da Silva, W. L., Ingram, J., Hackett, C. A., Coombs, J. J., Douches, D., Bryan, G. J., De Jong, W., & Gray, S. (2017). Mapping loci that control tuber and foliar symptoms caused by PVY in autotetraploid potato (*Solanum tuberosum* L.). *G3: Genes, Genomes, Genetics, 7*(11), 3587–3595. https://doi.org/10.1534/g3.117.300264

De Jong, H. (1981). Inheritance of russeting in cultivated diploid potatoes. *Potato Research, 24*(3), 309–313. https://doi.org/10.1007/BF02360368

Devaux, A., Kromann, P., & Ortiz, O. (2014). Potatoes for sustainable global food security. *Potato Research, 57*(3), 185–199. https://doi.org/10.1007/s11540-014-9265-1

D'hoop, B. B., Paulo, M. J., Kowitwanich, K., Sengers, M., Visser, R. G. F., van Eck, H. J., & van Eeuwijk, F. A. (2010). Population structure and linkage disequilibrium unravelled in tetraploid potato. *Theoretical and Applied Genetics, 121*(6), 1151–1170. https://doi.org/10.1007/s00122-010-1379-5

Douches, D., Hirsch, C., Manrique-Carpintero, N., Massa, A., Coombs, J., Hardigan, M., Bisognin, D., De Jong, W., & Buell, C. (2014). The contribution of the Solanaceae coordinated agricultural project to potato breeding. *Potato Research, 57*, 215–224. https://doi.org/10.1007/s11540-014-9267-z

Duangjit, J., Causse, M., & Sauvage, C. (2016). Efficiency of genomic selection for tomato fruit quality. *Molecular Breeding*, *36*(3), 29. https://doi.org/10.1007/s11032-016-0453-3

Ellis, D., Chavez, O., Coombs, J., Soto, J., Gomez, R., Douches, D., Panta, A., Silvestre, R., & Anglin, N. L. (2018). Genetic identity in genebanks: Application of the SolCAP 12K SNP array in fingerprinting and diversity analysis in the global in trust potato collection. *Genome*, *61*(7), 523–537. https://doi.org/10.1139/gen-2017-0201

Endelman, J. B., Carley, C. A. S., Bethke, P. C., Coombs, J. J., Clough, M. E., Silva, W. L. da, Jong, W. S. D., Douches, D. S., Frederick, C. M., Haynes, K. G., Holm, D. G., Miller, J. C., Muñoz, P. R., Navarro, F. M., Novy, R. G., Palta, J. P., Porter, G. A., Rak, K. T., Sathuvalli, V. R., … Yencho, G. C. (2018). Genetic variance partitioning and genome-wide prediction with allele dosage information in autotetraploid potato. *Genetics*, *209*(1), 77–87. https://doi.org/10.1534/genetics.118.300685

FAO. (2009). *Why potato? - International year of the potato 2008, Rome*. FAO. http://www.fao.org/potato-2008/en/aboutiyp/index.html

FAO. (2021). *Dimensions of need—Staple foods: What do people eat?, Rome*. FAO. http://www.fao.org/3/u8480e/u8480e07.htm

FAOSTAT. (2021). *Statistical data. Food and Agriculture Organization of the United Nations, Rome*. www.fao.org/faostat

Felcher, K. J., Coombs, J. J., Massa, A. N., Hansey, C. N., Hamilton, J. P., Veilleux, R. E., Buell, C. R., & Douches, D. S. (2012). Integration of two diploid potato linkage maps with the potato genome sequence. *PLOS ONE*, *7*(4), e36347. https://doi.org/10.1371/journal.pone.0036347

Fernandes, S. B., Dias, K. O. G., Ferreira, D. F., & Brown, P. J. (2018). Efficiency of multi-trait, indirect, and trait-assisted genomic selection for improvement of biomass sorghum. *Theoretical and Applied Genetics*, *131*(3), 747–755. https://doi.org/10.1007/s00122-017-3033-y

Flint-Garcia, S. A., Thornsberry, J. M., & Buckler, E. S. (2003). Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology*, *54*, 357–374. https://doi.org/10.1146/annurev.arplant.54.031902.134907

Gebhardt, C. (2013). Bridging the gap between genome analysis and precision breeding in potato. *Trends in Genetics*, *29*(4), 248–256. https://doi.org/10.1016/j.tig.2012.11.006

Gebhardt, C. (2016). The historical role of species from the Solanaceae plant family in genetic research. *Theoretical and Applied Genetics*, *129*(12), 2281–2294. https://doi.org/10.1007/s00122-016-2804-1

Gebhardt, C., Ballvora, A., Walkemeier, B., Oberhagemann, P., & Schüler, K. (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. *Molecular Breeding*, *13*(1), 93–102. https://doi.org/10.1023/B:MOLB.0000012878.89855.df

Ginzberg, I., Minz, D., Faingold, I., Soriano, S., Mints, M., Fogelman, E., Warshavsky, S., Zig, U., & Yermiyahu, U. (2012). Calcium mitigated potato skin physiological disorder. *American Journal of Potato Research*, *89*(5), 351–362. https://doi.org/10.1007/s12230-012-9249-0

Grandke, F., Ranganathan, S., van Bers, N., de Haan, J. R., & Metzler, D. (2017). PERGOLA: Fast and deterministic linkage mapping of polyploids. *BMC Bioinformatics*, *18*(1), 12. https://doi.org/10.1186/s12859-016-1416-8

Habyarimana, E., Parisi, B., & Mandolino, G. (2017). Genomic prediction for yields, processing and nutritional quality traits in cultivated potato (*Solanum tuberosum* L.). *Plant Breeding*, *136*(2), 245–252. https://doi.org/10.1111/pbr.12461

Hackett, C. A., Boskamp, B., Vogogias, A., Preedy, K. F., & Milne, I. (2017). TetraploidSNPMap: Software for linkage analysis and QTL mapping in autotetraploid populations using SNP dosage data. *Journal of Heredity*, *108*(4), 438–442. https://doi.org/10.1093/jhered/esx022

Hackett, C. A., McLean, K., & Bryan, G. J. (2013). Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. *PLOS ONE*, *8*(5), e63939. https://doi.org/10.1371/journal.pone.0063939

Hamilton, J. P., Hansey, C. N., Whitty, B. R., Stoffel, K., Massa, A. N., Van Deynze, A., De Jong, W. S., Douches, D. S., & Buell, C. R. (2011). Single nucleotide polymorphism discovery in elite North American potato germplasm. *BMC Genomics*, *12*(1), 302. https://doi.org/10.1186/1471-2164-12-302

Heffner, E. L., Jannink, J.-L., & Sorrells, M. E. (2011). Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *The Plant Genome*, *4*(1). https://doi.org/10.3835/plantgenome2010.12.0029

Heffner, E. L., Lorenz, A. J., Jannink, J.-L., & Sorrells, M. E. (2010). Plant breeding with genomic selection: Gain per unit time and cost. *Crop Science*, *50*(5), 1681–1690. https://doi.org/10.2135/cropsci2009.11.0662

Hijmans, R. J., & Spooner, D. M. (2001). Geographic distribution of wild potato species. *American Journal of Botany*, *88*(11), 2101–2112. https://doi.org/10.2307/3558435

Hirsch, C. N., Hirsch, C. D., Felcher, K., Coombs, J., Zarka, D., Deynze, A. V., Jong, W. D., Veilleux, R. E., Jansky, S., Bethke, P., Douches, D. S., & Buell, C. R. (2013). Retrospective view of North American potato (*Solanum tuberosum* L.) breeding in the 20th and 21st centuries. *G3: Genes, Genomes, Genetics*, *3*(6), 1003–1013. https://doi.org/10.1534/g3.113.005595

Igarashi, T., Tsuyama, M., Ogawa, K., Koizumi, E., Sanetomo, R., & Hosaka, K. (2019). Evaluation of Japanese potatoes using single nucleotide polymorphisms (SNPs). *Molecular Breeding*, *39*(1). https://doi.org/10.1007/s11032-018-0917-8

Jamali, S. H., Cockram, J., & Hickey, L. T. (2019). Insights into deployment of DNA markers in plant variety protection and registration. *Theoretical and Applied Genetics*, *132*(7), 1911–1929. https://doi.org/10.1007/s00122-019-03348-7

Jansky, S. (2009). Chapter 2—Breeding, genetics, and cultivar development. In J. Singh & L. Kaur (Eds.), *Advances in Potato Chemistry and Technology* (pp. 27–62). Academic Press. https://doi.org/10.1016/B978-0-12-374349-7.00002-7

Jupe, F., Witek, K., Verweij, W., Śliwka, J., Pritchard, L., Etherington, G. J., Maclean, D., Cock, P. J., Leggett, R. M., Bryan, G. J., Cardle, L., Hein, I., & Jones, J. D. G. (2013). Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *The Plant Journal*, *76*(3), 530–544. https://doi.org/10.1111/tpj.12307

Kaiser, N. R., Coombs, J. J., Felcher, K. J., Hammerschmidt, R., Zuehlke, M. L., Buell, C. R., & Douches, D. S. (2020). Genome-wide association analysis of common scab resistance and expression profiling of tubers in response to thaxtomin a treatment underscore the complexity of common scab resistance in tetraploid potato. *American Journal of Potato Research*, *97*(5), 513–522. https://doi.org/10.1007/s12230-020-09800-5

Kanter, M., & Elkin, C. (2019). Potato as a source of nutrition for physical performance. *American Journal of Potato Research*, *96*(2), 201–205. https://doi.org/10.1007/s12230-018-09701-8

Kao, T. H., & McCubbin, A. G. (1996). How flowering plants discriminate between self and non-self pollen to prevent inbreeding. *Proceedings of the National Academy of Sciences*, *93*(22), 12059–12065. https://doi.org/10.1073/pnas.93.22.12059

Keijbets, M. J. H. (2008). Potato processing for the consumer: Developments and future challenges. *Potato Research*, *51*(3–4), 271–281. https://doi.org/10.1007/s11540-008-9104-3

Kirch, H. H., Uhrig, H., Lottspeich, F., Salamini, F., & Thompson, R. D. (1989). Characterization of proteins associated with self-incompatibility in *Solanum tuberosum*. *Theoretical and Applied Genetics*, *78*(4), 581–588. https://doi.org/10.1007/BF00290845

Klaassen, M. T., Willemsen, J. H., Vos, P. G., Visser, R. G. F., van Eck, H. J., Maliepaard, C., & Trindade, L. M. (2019). Genome-wide association analysis in tetraploid potato reveals four QTLs for protein content. *Molecular Breeding*, *39*(11), 151. https://doi.org/10.1007/s11032-019-1070-8

Klápště, J., Dungey, H. S., Telfer, E. J., Suontama, M., Graham, N. J., Li, Y., & McKinley, R. (2020). Marker selection in multivariate genomic prediction improves accuracy of low heritability traits. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.499094

Kleinkopf, G. E., Oberg, N. A., & Olsen, N. L. (2003). Sprout inhibition in storage: Current status, new chemistries and natural compounds. *American Journal of Potato Research*, *80*(5), 317–327. https://doi.org/10.1007/BF02854316

Kloosterman, B., Oortwijn, M., uitdeWilligen, J., America, T., de Vos, R., Visser, R. G., & Bachem, C. W. (2010). From QTL to candidate gene: Genetical genomics of simple and complex traits in potato using a pooling strategy. *BMC Genomics*, *11*(1), 158. https://doi.org/10.1186/1471-2164-11-158

Kolech, S. A., Halseth, D., Perry, K., Wolfe, D., Douches, D. S., Coombs, J., & De Jong, W. (2016). Genetic diversity and relationship of Ethiopian potato varieties to germplasm from North America, Europe and the International Potato Center. *American Journal of Potato Research*, *93*(6), 609–619. https://doi.org/10.1007/s12230-016-9543-3

Konieczny, A., & Ausubel, F. M. (1993). A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *The Plant Journal*, *4*(2), 403–410. https://doi.org/10.1046/j.1365-313X.1993.04020403.x

Krishnappa, G., Savadi, S., Tyagi, B. S., Singh, S. K., Mamrutha, H. M., Kumar, S., Mishra, C. N., Khan, H., Gangadhara, K., Uday, G., Singh, G., & Singh, G. P. (2021). Integrated genomic selection for rapid improvement of crops. *Genomics*, *113*(3), 1070–1086. https://doi.org/10.1016/j.ygeno.2021.02.007

Leonel, M., do Carmo, E. L., Fernandes, A. M., Soratto, R. P., Ebúrneo, J. A. M., Garcia, É. L., & dos Santos, T. P. R. (2017). Chemical composition of potato tubers: The effect of cultivars and growth conditions. *Journal of Food Science and Technology*, *54*(8), 2372–2378. https://doi.org/10.1007/s13197-017-2677-6

Leisner, C. P., Hamilton, J. P., Crisovan, E., Manrique-Carpintero, N. C., Marand, A. P., Newton, L., Pham, G. M., Jiang, J., Douches, D. S., Jansky, S. H., & Buell, C. R. (2018). Genome sequence of M6, a diploid inbred clone of the high-glycoalkaloid-producing tuber-bearing potato species *Solanum chacoense*, reveals residual heterozygosity. *The Plant Journal*, 94(3), 562–570. https://doi.org/10.1111/tpj.13857

Li, L., Paulo, M.-J., Strahwald, J., Lübeck, J., Hofferbert, H.-R., Tacke, E., Junghans, H., Wunder, J., Draffehn, A., van Eeuwijk, F., & Gebhardt, C. (2008). Natural DNA variation at candidate loci is associated with potato chip color, tuber starch content, yield and starch yield. *Theoretical and Applied Genetics*, *116*(8), 1167–1181. https://doi.org/10.1007/s00122-008-0746-y

Li, X.-Q., De Jong, H., De Jong, D. M., & De Jong, W. S. (2005). Inheritance and genetic mapping of tuber eye depth in cultivated diploid potatoes. *Theoretical and Applied Genetics*, *110*(6), 1068–1073. https://doi.org/10.1007/s00122-005-1927-6

Lindqvist-Kreuze, H., Gastelo, M., Perez, W., Forbes, G. A., de Koeyer, D., & Bonierbale, M. (2014). Phenotypic stability and genome-wide association study of late blight resistance in potato genotypes adapted to the tropical highlands. *Phytopathology*, *104*(6), 624–633. https://doi.org/10.1094/PHYTO-10-13-0270-R

Liyanage, D. W. K., Yevtushenko, D. P., Konschuh, M., Bizimungu, B., & Lu, Z.-X. (2021). Processing strategies to decrease acrylamide formation, reducing sugars and free asparagine content in potato chips from three commercial cultivars. *Food Control*, *119*, 107452. https://doi.org/10.1016/j.foodcont.2020.107452

Lorenz, A. J., Smith, K. P., & Jannink, J.-L. (2012). Potential and optimization of genomic selection for fusarium head blight resistance in six-row barley. *Crop Science*, *52*(4), 1609–1621. https://doi.org/10.2135/cropsci2011.09.0503

Love, S. L., Pavek, J. J., Thompson-Johns, A., & Bohl, W. (1998). Breeding progress for potato chip quality in North American cultivars. *American Journal of Potato Research*, *75*(1), 27–36. https://doi.org/10.1007/BF02883514

Lulai, E. C. (2007). Chapter 22—Skin-set, wound healing, and related defects. In D. Vreugdenhil, J. Bradshaw, C. Gebhardt, F. Govers, D. K. L. Mackerron, M. A. Taylor, & H. A. Ross (Eds.), *Potato Biology and Biotechnology* (pp. 471–500). Elsevier Science B.V. https://doi.org/10.1016/B978-044451018-1/50064-6

Manrique-Carpintero, Norma C., Coombs, J. J., Pham, G. M., Laimbeer, F. P. E., Braz, G. T., Jiang, J., Veilleux, R. E., Buell, C. R., & Douches, D. S. (2018). Genome reduction in tetraploid potato reveals genetic load, haplotype variation, and loci associated with agronomic traits. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.00944

Manrique-Carpintero, Norma Constanza, Tokuhisa, J. G., Ginzberg, I., & Veilleux, R. E. (2014). Allelic variation in genes contributing to glycoalkaloid biosynthesis in a diploid interspecific population of potato. *Theoretical and Applied Genetics*, *127*(2), 391–405. https://doi.org/10.1007/s00122-013-2226-2

Massa, A. N., Manrique-Carpintero, N. C., Coombs, J., Haynes, K. G., Bethke, P. C., Brandt, T. L., Gupta, S. K., Yencho, G. C., Novy, R. G., & Douches, D. S. (2018). Linkage analysis and QTL mapping in a tetraploid russet mapping population of potato. *BMC Genetics*, *19*(1). https://doi.org/10.1186/s12863-018-0672-1

Massa, A. N., Manrique-Carpintero, N. C., Coombs, J. J., Zarka, D. G., Boone, A. E., Kirk, W. W., Hackett, C. A., Bryan, G. J., & Douches, D. S. (2015). Genetic Linkage mapping of economically important traits in cultivated tetraploid potato (*Solanum tuberosum* L.). *G3: Genes, Genomes, Genetics*, *5*(11), 2357–2364. https://doi.org/10.1534/g3.115.019646

McGill, C. R., Kurilich, A. C., & Davignon, J. (2013). The role of potatoes and potato components in cardiometabolic health: A review. *Annals of Medicine*, *45*(7), 467–473. https://doi.org/10.3109/07853890.2013.813633

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Mollinari, M., & Garcia, A. A. F. (2019). Linkage analysis and haplotype phasing in experimental autopolyploid populations with high ploidy level using hidden Markov models. *G3: Genes, Genomes, Genetics*, *9*(10), 3297–3314. https://doi.org/10.1534/g3.119.400378

Nacheva, E., & Pevicharova, G. (2008). Potato breeding lines for processing. *Genetics and Breeding*, *37*, 77–84.

Nesterenko, S., & Sink, K. C. (2003). Carotenoid profiles of potato breeding lines and selected cultivars. *HortScience*, *38*(6), 1173–1177. https://doi.org/10.21273/HORTSCI.38.6.1173

NPC National Potato Council. (2019). *Potato statistical yearbook 2019, Washington, D.C.*

https://www.nationalpotatocouncil.org/files/5015/6380/8213/2019_Stat_Book__fnl.pdf

Ortiz, O., & Mares, V. (2017). The historical, social, and economic importance of the potato crop. In S. Kumar Chakrabarti, C. Xie, & J. Kumar Tiwari (Eds.), *The Potato Genome* (pp. 1–10). Springer International Publishing. https://doi.org/10.1007/978-3-319-66135-3_1

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., Dreisigacker, S., Crossa, J., Sánchez-Villeda, H., Sorrells, M., & Jannink, J.-L. (2012). Genomic selection in wheat breeding using genotyping-by-sequencing. *The Plant Genome*, *5*(3). https://doi.org/10.3835/plantgenome2012.06.0006

Prashar, A., Hornyik, C., Young, V., McLean, K., Sharma, S. K., Dale, M. F. B., & Bryan, G. J. (2014). Construction of a dense SNP map of a highly heterozygous diploid potato population and QTL analysis of tuber shape and eye depth. *Theoretical and Applied Genetics*, *127*(10), 2159–2171. https://doi.org/10.1007/s00122-014-2369-9

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Rafalski, J. A. (2010). Association genetics in crop improvement. *Current Opinion in Plant Biology*, *13*(2), 174–180. https://doi.org/10.1016/j.pbi.2009.12.004

Rodríguez, F., Ghislain, M., Clausen, A. M., Jansky, S. H., & Spooner, D. M. (2010). Hybrid origins of cultivated potatoes. *Theoretical and Applied Genetics*, *121*(6), 1187–1198. https://doi.org/10.1007/s00122-010-1422-6

Rosyara, U. R., De Jong, W. S., Douches, D. S., & Endelman, J. B. (2016). Software for genome-wide association studies in autopolyploids and its application to potato. *The Plant Genome*, *9*(2). https://doi.org/10.3835/plantgenome2015.08.0073

Rutkoski, J., Singh, R. P., Huerta-Espino, J., Bhavani, S., Poland, J., Jannink, J. L., & Sorrells, M. E. (2015). Efficient use of historical data for genomic selection: A case study of stem rust resistance in wheat. *The Plant Genome*, *8*(1), plantgenome2014.09.0046. https://doi.org/10.3835/plantgenome2014.09.0046

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, *74*(12), 5463–5467. https://doi.org/10.1073/pnas.74.12.5463

Sarkar, D. (2008). The signal transduction pathways controlling in planta tuberization in potato: An emerging synthesis. *Plant Cell Reports*, *27*(1), 1–8. https://doi.org/10.1007/s00299-007-0457-x

Schmitz Carley, C. A., Coombs, J. J., Douches, D. S., Bethke, P. C., Palta, J. P., Novy, R. G., & Endelman, J. B. (2017). Automated tetraploid genotype calling by hierarchical clustering. *Theoretical and Applied Genetics*, *130*(4), 717–726. https://doi.org/10.1007/s00122-016-2845-5

Schumann, M. J., Zeng, Z.-B., Clough, M. E., & Yencho, G. C. (2017). Linkage map construction and QTL analysis for internal heat necrosis in autotetraploid potato. *Theoretical and Applied Genetics*, *130*(10), 2045–2056. https://doi.org/10.1007/s00122-017-2941-1

Shen, J., Li, Z., Chen, J., Song, Z., Zhou, Z., & Shi, Y. (2016). SHEsisPlus, a toolset for genetic studies on polyploid species. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep24095

Slater, A. T., Cogan, N. O. I., Hayes, B. J., Schultz, L., Dale, M. F. B., Bryan, G. J., & Forster, J. W. (2014). Improving breeding efficiency in potato using molecular and quantitative genetics. *Theoretical and Applied Genetics*, *127*(11), 2279–2292. https://doi.org/10.1007/s00122-014-2386-8

ŚLiwka, J., Wasilewicz-Flis, I., Jakuczun, H., & Gebhardt, C. (2008). Tagging quantitative trait loci for dormancy, tuber shape, regularity of tuber shape, eye depth and flesh colour in diploid potato originated from six Solanum species. *Plant Breeding*, *127*(1), 49–55. https://doi.org/10.1111/j.1439-0523.2008.01420.x

Spindel, J., & Iwata, H. (2018). Genomic selection in rice breeding. In T. Sasaki & M. Ashikari (Eds.), *Rice Genomics, Genetics and Breeding* (pp. 473–496). Springer. https://doi.org/10.1007/978-981-10-7461-5_24

Spooner, D. M. (2009). DNA barcoding will frequently fail in complicated groups: An example in wild potatoes. *American Journal of Botany*, *96*(6), 1177–1189. https://doi.org/10.3732/ajb.0800246

Stark, J. C., Love, S. L., & Knowles, N. R. (2020). Tuber quality. In J. C. Stark, M. Thornton, & P. Nolte (Eds.), *Potato Production Systems* (pp. 479–497). Springer International Publishing. https://doi.org/10.1007/978-3-030-39157-7_15

Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences, 100*(16), 9440–9445. https://doi.org/10.1073/pnas.1530509100

Sverrisdóttir, E., Byrne, S., Sundmark, E. H. R., Johnsen, H. Ø., Kirk, H. G., Asp, T., Janss, L., & Nielsen, K. L. (2017). Genomic prediction of starch content and chipping quality in tetraploid potato using genotyping-by-sequencing. *Theoretical and Applied Genetics*, *130*(10), 2091–2108. https://doi.org/10.1007/s00122-017-2944-y

Sverrisdóttir, E., Sundmark, E. H. R., Johnsen, H. Ø., Kirk, H. G., Asp, T., Janss, L., Bryan, G., & Nielsen, K. L. (2018). The value of expanding the training population to improve genomic selection models in tetraploid potato. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.01118

Teo, Y. Y. (2008). Common statistical issues in genome-wide association studies: A review on power, data quality control, genotype calling and population structure. *Current Opinion in Lipidology*, *19*(2), 133–143. https://doi.org/10.1097/MOL.0b013e3282f5dd77

Thorup, T. A., Tanyolac, B., Livingstone, K. D., Popovsky, S., Paran, I., & Jahn, M. (2000). Candidate gene analysis of organ pigmentation loci in the Solanaceae. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(21), 11192–11197. https://doi.org/10.1073/pnas.97.21.11192

Tian, F., Bradbury, P. J., Brown, P. J., Hung, H., Sun, Q., Flint-Garcia, S., Rocheford, T. R., McMullen, M. D., Holland, J. B., & Buckler, E. S. (2011). Genome-wide

association study of leaf architecture in the maize nested association mapping population. *Nature Genetics*, *43*(2), 159–162. https://doi.org/10.1038/ng.746

US Department of Agriculture National Agricultural Statistics Service. (2019). *Potatoes 2018 summary, Washington, D.C.* https://www.nass.usda.gov/Publications/Todays_Reports/reports/pots0919.pdf

van Eck, H. J., Jacobs, J. E., Ton, J., Stiekemat, W. J., & Jacobsen, E. (1994). Multiple alleles for tuber shape in diploid potato detected by qualitative and quantitative genetic analysis using RFLPs. *Genetics*, *137*(1), 303–309.

Vieira, M. L. C., Santini, L., Diniz, A. L., & Munhoz, C. de F. (2016). Microsatellite markers: What they mean and why they are so useful. *Genetics and Molecular Biology*, *39*(3), 312–328. https://doi.org/10.1590/1678-4685-GMB-2016-0027

Voorrips, R. E., Gort, G., & Vosman, B. (2011). Genotype calling in tetraploid species from bi-allelic marker data using mixture models. *BMC Bioinformatics*, *12*, 172. https://doi.org/10.1186/1471-2105-12-172

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., & Kuiper, M. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*, *23*(21), 4407–4414.

Vos, Peter. G., Uitdewilligen, J. G. A. M. L., Voorrips, R. E., Visser, R. G. F., & van Eck, H. J. (2015). Development and analysis of a 20K SNP array for potato (*Solanum tuberosum*): An insight into the breeding history. *Theoretical and Applied Genetics*, *128*(12), 2387–2401. https://doi.org/10.1007/s00122-015-2593-y

Wang, M.-R., Cui, Z.-H., Li, J.-W., Hao, X.-Y., Zhao, L., & Wang, Q.-C. (2018). *In vitro* thermotherapy-based methods for plant virus eradication. *Plant Methods*, 14, 87. https://doi.org/10.1186/s13007-018-0355-y

Watanabe, K. (2015). Potato genetics, genomics, and applications. *Breeding Science*, *65*(1), 53–68. https://doi.org/10.1270/jsbbs.65.53

Wayumba, B. O., Choi, H. S., & Seok, L. Y. (2019). Selection and evaluation of 21 potato (*Solanum tuberosum*) breeding clones for cold chip processing. *Foods*, *8*(3), 98. https://doi.org/10.3390/foods8030098

Welsh, J., & McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research*, *18*(24), 7213–7218.

Wijesinha-Bettoni, R., & Mouillé, B. (2019). The contribution of potatoes to global food security, nutrition and healthy diets. *American Journal of Potato Research*, *96*(2), 139–149. https://doi.org/10.1007/s12230-018-09697-1

Wu, J., Feng, F., Lian, X., Teng, X., Wei, H., Yu, H., Xie, W., Yan, M., Fan, P., Li, Y., Ma, X., Liu, H., Yu, S., Wang, G., Zhou, F., Luo, L., & Mei, H. (2015). Genome-wide association study (GWAS) of mesocotyl elongation based on re-sequencing approach in rice. *BMC Plant Biology*, *15*(1), 218. https://doi.org/10.1186/s12870-015-0608-0

Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., Wang, J., Orjeda, G., Guzman, F., Torres, M., Lozano, R., Ponce, O., Martinez, D., De la Cruz, G., Chakrabarti, S. K., Patil, V. U. (2011). Genome sequence and

analysis of the tuber crop potato. *Nature*, *475*(7355), 189–195. https://doi.org/10.1038/nature10158

Xu, Yang, Li, P., Yang, Z., & Xu, C. (2017). Genetic mapping of quantitative trait loci in crops. *The Crop Journal*, *5*(2), 175–184. https://doi.org/10.1016/j.cj.2016.06.003

Xu, Yunbi, Liu, X., Fu, J., Wang, H., Wang, J., Huang, C., Prasanna, B. M., Olsen, M. S., Wang, G., & Zhang, A. (2020). Enhancing genetic gain through genomic selection: From livestock to plants. *Plant Communications*, *1*(1), 100005. https://doi.org/10.1016/j.xplc.2019.100005

Yano, K., Morinaka, Y., Wang, F., Huang, P., Takehara, S., Hirai, T., Ito, A., Koketsu, E., Kawamura, M., Kotake, K., Yoshida, S., Endo, M., Tamiya, G., Kitano, H., Ueguchi-Tanaka, M., Hirano, K., & Matsuoka, M. (2019). GWAS with principal component analysis identifies a gene comprehensively controlling rice architecture. *Proceedings of the National Academy of Sciences*, *116*(42), 21262–21267. https://doi.org/10.1073/pnas.1904964116

Yousaf, M. F., Demirel, U., Naeem, M., & Çalışkan, M. E. (2021). Association mapping reveals novel genomic regions controlling some root and stolon traits in tetraploid potato (*Solanum tuberosum* L.). *3 Biotech*, *11*(4), 174. https://doi.org/10.1007/s13205-021-02727-6

Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., & Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts

for multiple levels of relatedness. *Nature Genetics*, *38*(2), 203–208. https://doi.org/10.1038/ng1702

Yuan, J., Wang, X., Zhao, Y., Khan, N. U., Zhao, Z., Zhang, Y., Wen, X., Tang, F., Wang, F., & Li, Z. (2020). Genetic basis and identification of candidate genes for salt tolerance in rice by GWAS. *Scientific Reports*, *10*(1), 9958. https://doi.org/10.1038/s41598-020-66604-7

Zhang, J., Song, Q., Cregan, P. B., & Jiang, G.-L. (2016). Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). *Theoretical and Applied Genetics*, *129*(1), 117–130. https://doi.org/10.1007/s00122-015-2614-x

Zhang, X., Guan, Z., Wang, L., Fu, J., Zhang, Y., Li, Z., Ma, L., Liu, P., Zhang, Y., Liu, M., Li, P., Zou, C., He, Y., Lin, H., Yuan, G., Gao, S., Pan, G., & Shen, Y. (2020). Combined GWAS and QTL analysis for dissecting the genetic architecture of kernel test weight in maize. *Molecular Genetics and Genomics*, *295*(2), 409–420. https://doi.org/10.1007/s00438-019-01631-2

Zhao, Y., Gowda, M., Liu, W., Würschum, T., Maurer, H. P., Longin, F. H., Ranc, N., & Reif, J. C. (2012). Accuracy of genomic selection in European maize elite breeding populations. *Theoretical and Applied Genetics*, *124*(4), 769–776. https://doi.org/10.1007/s00122-011-1745-y

Zhu, C., Gore, M., Buckler, E. S., & Yu, J. (2008). Status and prospects of association mapping in plants. *The Plant Genome*, *1*(1), 5–20. https://doi.org/10.3835/plantgenome2008.02.0089

Zia, M. A. B., Demirel, U., Nadeem, M. A., & Çaliskan, M. E. (2020). Genome-wide association study identifies various loci underlying agronomic and morphological traits in diversified potato panel. *Physiology and Molecular Biology of Plants*, *26*(5), 1003–1020. https://doi.org/10.1007/s12298-020-00785-3

2. GENETIC DIVERSITY AND POPULATION STRUCTURE OF ADVANCED CLONES SELECTED OVER FORTY YEARS BY A POTATO BREEDING PROGRAM IN THE USA*

2.1. Introduction

Potato (*Solanum tuberosum* L.) is the world's fourth most important crop after maize, rice, and wheat (FAOSTAT, 2018). Worldwide, over one billion people consume potatoes as a staple food (USD NASS, 2019). Potatoes, the leading vegetable crop in the United States, are grown commercially in 30 states. Idaho grows more potatoes than any other state, followed by Washington, North Dakota, Wisconsin, and Colorado (USD NASS, 2019). Even though potato production in Texas is comparatively lower (approximately 8,000 ha), growers can harvest and provide fresh potatoes to the market earlier in the growing season than other states, and often receive two to three times higher prices than growers from Northern States (USD NASS, 2019).

The Texas A&M University (TAMU) Potato Breeding Program was established by the late J. Creighton Miller Jr., Ph.D. in 1972 aiming to provide improved early maturing cultivars with high yield and quality that would enable Texas potato producers to remain viable and competitive, and to supply superior products to consumers. In recent years, the emphasis has been placed on increasing yield and quality in addition to

---

*Reprinted with permission from "Genetic diversity and population structure of advanced clones selected over forty years by a potato breeding program in the USA" by Jeewan Pandey, Douglas C. Scheuring, Jeffrey W. Koym, Joseph Coombs, Richard G. Novy, Asunta L. Thompson, David G. Holm, David S. Douches, J. Creighton Miller Jr. & M. Isabel Vales, 2021. *Scientific Reports*, 11, 8344, https://doi.org/10.1038/s41598-021-87284-x, Copyright © 2021, The Author(s)

disease/pest resistances like Potato Virus Y (PVY), nematodes (*Globodera rostochiensis* and *Globodera pallida*), late blight (*Phytophthora infestans*), potato psyllid (*Bactericera cockerelli* carrying *Candidatus* Liberibacter, the causal agent of the zebra chip disease), high-temperature tolerance, cold sweetening resistance, health and nutritional properties, and broad adaptability. The TAMU Potato Breeding program has developed/co-developed and released 17 cultivars, including clonal selections. Some of them make up a substantial and increasing share of the regional/national potato production and have become important contributors to the economies of several states. Of all the cultivars released over the past 15 years by the 12 US potato breeding programs, those developed by the Texas program have ranked in the top four to five nationally in the total area approved for seed certification over the past several years (NPC, 2019). This has been due, in large measure, to the popularity of the four Texas Russet Norkotah strains (Russet Norkotah 112, Russet Norkotah 223, Russet Norkotah 278, and Russet Norkotah 296) with improved plant type to withstand environmental stresses. The Texas Russet Norkotah strains with increased vine vigor and some resistance to early dying (*Verticillium* wilt) are an outstanding early market alternative to the standard Russet Norkotah variety (Miller et al., 1999).

Despite many available potato cultivars, there is a need for new cultivars. New cultivars must produce high yields under low inputs, have disease and pest resistance, and environmental stress tolerance such as high or low temperature, drought, and salinity. If possible, they should also have improved nutritional and health properties (Bradshaw, 2007a). Exploration of potato genetic diversity has been proposed to create

new varieties well adapted to these challenges, and also to better manage these collections. The development of new, improved varieties is done through breeding, which involves identifying superior and complementary parents from the available germplasm and crossing them to generate variability and permit selection of clones combining trait of interest. Breeders maintain valuable germplasm in tissue culture for long-term conservation of genetic resources, and also to initiate limited generation seed production of potato varieties from disease-free stocks. Thus, breeders have to think strategically to capture allelic diversity from a smaller set of parent combinations. For this, a breeder can use genetic distance based on molecular markers to complement co-ancestry/pedigree analysis to avoid crossing closely related parents and hence prevent inbreeding depression and to ensure genetic variation for continued selection progress. Genetic distance-based criteria have also been strongly recommended for evaluation and creation of core sets (Odong et al., 2013).

Further, the genetic characterization of clone bank collections is essential to assess their diversity and population structure. The identification of suitable genotypes from the study could serve as a source of new alleles in potato breeding programs. Molecular markers have been used to test the genetic diversity of potatoes. Recent advances in the development of high-throughput genotyping platforms together with whole-genome coverage and affordability have turned single nucleotide polymorphisms (SNPs) into one of the most promising tools for the investigation of genetic diversity. Several studies have implemented the Infinium Potato Array (Illumina Inc., San Diego, CA, USA) for genetic diversity studies. The 8K SNP array distinguished diverse North

American varieties based mainly on market classes (Hirsch et al., 2013). Kolech et al. (2016) used the same set of 8K SNPs to evaluate the genetic diversity of Ethiopian potato cultivars. Vos et al. (2015) developed a 20K SNP array and used it to genotype a total of 569 potato genotypes and found fingerprints of the breeding history in recent breeding materials such as identification of introgression segments, selection, and founder signatures. Genetic diversity in the Colombian Central Collection of *Solanum tuberosum* L. using SNP markers found that the Andigena (autotetraploid) population was more genetically diverse, but less genetically sub structured than the Phureja (diploid) population (Berdugo-Cely et al., 2017). Ellis et al. (2018) used the 12K SNP array for fingerprinting and diversity analysis of the cultivated potato collection from the International Potato Center (CIP) in Peru and reported some genetic redundancies among individual accessions with some putative misclassified accessions. Recently, Igarashi et al. (2019) used the 12K SNP array to characterize and compare 164 Japanese potatoes, including 70 breeding clones for chip processing with North American and European potatoes. Thus, the success of potato breeding depends on the understanding and use of the available gene pool of varieties and breeding clones. The Potato SNP array has been very useful for performing a robust and direct comparison of genetic diversity among different gene pools but has never been applied to the advanced clones selected over multiple years.

Further, with the availability of high-density genotype data, it is possible to identify regions of the genome that provide evidence of selective pressure commonly known as “signatures of selection” (Cadzow et al., 2014). Different statistical

approaches have been developed to identify selection footprints. According to Vitti et al. (2013), they are of three main types: (a) measures based on the allelic frequencies (e.g., Tajima's D, PCAdapt), (b) measures based on the differentiation between and within species/groups (e.g., XP-EHH, Fst), and (c) within population/groups measures based on extended haplotype homozygosity (e.g., iHS). These methods have been applied to several crops, including wheat (Pont et al., 2019), oat (Bekele et al., 2018), maize (Li et al., 2021), rice (He et al., 2015), tomato (Sauvage et al., 2017), and potato (Hardigan et al., 2017). The PCAdapt method tests how much each variant is associated with population structure, assuming that outlier variants are indicative of local adaptation. It does not need grouping of individuals into populations and can handle admixed individuals (Luu et al., 2017). The iHS approach measures the amount of extended haplotype homozygosity (EHH) for a given SNP within-population whereas XP-EHH compares the extended haplotype homozygosity between two populations (Sabeti et al., 2007). Recent selection events in which haplotypes have almost or fully risen to fixation are detected by iHS and XP-EHH statistics (Sabeti et al., 2007). Thus, methods for detecting evidence of selection provides a mechanism for highlighting genomic regions which are often associated with functional traits.

The goal of this research was to investigate potato varieties and advanced clones of the TAMU Potato Breeding Program (entered in the clone bank over 40 years of breeding) at the molecular level to assess genetic diversity for further genetic enhancement of important economic traits. In this study 214 TAMU potato clones were genotyped using 22K SNP markers to (a) examine the genetic diversity and the

population structure in the TAMU Potato Breeding clone bank collection, (b) to identify candidate loci under selection, (c) identify a "core set" to better manage the clone bank collection, and (d) check the accuracy of pedigree records of the clones.

2.2. Methods

2.2.1. Plant material

Two hundred fourteen potato clones were included in this study. The clones represent fresh and processing market classes with a variation for skin type (russet and smooth), flesh and skin color, shape, agronomic, biotic, abiotic, and quality traits. The collection comprised 31 chipping, 62 russet, 32 yellow-skinned, 68 red-skinned, and 21 purple-skinned clones. The collection was initiated during the 1980s and consisted mainly of early generation and advanced clones selected by the TAMU Potato Breeding Program. The introduction into tissue culture and virus eradication of early and advanced potato selections is a regular practice in the TAMU Potato Breeding Program since disease-free stocks from Texas selections are typically transferred to Colorado State University to produce clean seed for regional trials (SW and W). Some commercial cultivars developed by the TAMU Program were also preserved in the clone bank, including Russet Norkotah clonal selections (Russet Norkotah 112, Russet Norkotah 223, Russet Norkotah 278, and Russet Norkotah 296), Sierra Rose™ (ATTX961014-1R/Y), Sierra Gold™ (TX1523-1Ru/Y), Rio Rojo, COTX09022-3RuRE/Y (released under experimental ID), Reveille Russet (ATX91137-1Ru), Vanguard Russet (TX08352-5Ru) and Stampede Russet (TXAV657-27Ru). Commercially popular varieties, including Russet Norkotah (standard), Atlantic, and Russet Burbank were included as

reference genotypes. White LaSoda (a white skinned mutant of Red LaSoda selected by the TAMU Program) and Yukon Gold Strain (TXYG79) were also included in the study. All of the clones are now maintained in the TAMU breeding clone bank. To micropropagate the clones, tissue culture media consisting of Murashige and Skoog (MS) (4.8 g/L), sucrose (30 g/L), and agar (8 g/L) was used. Clean (disease-free status confirmed by ELISA assays) plant materials were multiplied and moved to the greenhouse in Fall 2018 to produce mini tubers. In the greenhouse, 12 plantlets of each clone were grown for 110 days in a standard flat insert (TO Plastics, MN, USA) (26.82 cm x 53.49 cm) with 32 cells (each cell size: 4.04 cm x 2.92 cm x 5.72 cm) filled with Sunshine Mix #1 (Sungro, Agawam, MA) with starter fertilizer Osmocote (Scotts Miracle-Gro, Marysville, OH). The photoperiod (light: dark) was 16:8 until flowering and 12:12 afterward to enhance tuberization. The greenhouse temperature averaged 20°C with a minimum of 14°C and a maximum of 31°C. Minitubers were harvested in Spring 2019 and stored at room temperature for approximately a week to confirm the skin and flesh color of the clones.

2.2.2. DNA extractions

Genomic DNA was extracted from 50-80 mg of fresh young potato leaves from tissue culture plantlets using the DNeasy Plant Pro Kit® (Qiagen, Valencia, CA, USA). DNA quality was examined using 1% agarose gel 1X TBE (Tris-Borate and ethylenediaminetetraacetic acid) and staining with GelRed (Biotium Inc., CA, USA) using a V.U.V. transilluminator (Benchtop VUV Transilluminators, UVP). Quantification of DNA was performed in a spectrophotometer (Nano Drop™, Thermo

Scientific, Waltham, Massachusetts, USA). DNA concentration was verified using the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, SanDiego, CA) and samples with uniform DNA concentration (50 ng μL^{-1}) were prepared.

2.2.3. SNP genotyping

Samples were assayed using the Infinium 22K V3 Potato Array on the Illumina iScan (Illumina Inc., San Diego, CA, USA) at Michigan State University. V3 Potato array includes the SNP from the Infinium 8303 Potato Array with additional markers from the Infinium high-confidence SNPs (69K) (Hamilton et al., 2011) and the SolSTW 20K array (Peter. G. Vos et al., 2015). Samples were SNP genotyped using the Illumina GenomeStudio 2.0.4 software (Illumina, San Diego, CA) for five-cluster (nulliplex = AAAA, simplex = AAAB, duplex = AABB, triplex = ABBB, and quadruplex = BBBB) marker calling using a custom tetraploid cluster file based on the PolyGentrain polyploid module calling of reference tetraploid samples (Illumina, San Diego, CA). The SNP genotype data were filtered to exclude low-quality, monomorphic SNPs, and loci with ≥10% missing data. Also, the alleles-design option was displayed in GenomeStudio to get genotypes in nucleotide format for STRUCTURE input. The genotyping data were transformed into diploid form as AAAA=AA, BBBB=BB, and AAAB, AABB, ABBB=AB to use in analysis packages which do not support polyploid data.

2.2.4. Genetic diversity

SNP genotypic data were used to study genetic diversity and to understand the genetic relationship among clones. Allele frequencies, polymorphic information content

(PIC), heterozygosity, and inbreeding coefficient were calculated in *snpReady* (Granato et al., 2018) package in R using the diploid genotypic calls.

The average pairwise divergence among genotypes, which represents the nucleotide diversity per bp, π (pi), and the expected number of polymorphic sites per nucleotide, θ (theta), were estimated in TASSEL v5.2.39 (Bradbury et al., 2007) using the default settings for the diploid genotypic calls. The normalized measure of the difference between the observed (π) and expected (θ) nucleotide diversity, known as Tajima's D, was also computed in TASSEL.

2.2.5. Population structure

Population structure was determined using STRUCTURE software version 2.3.4 (Pritchard et al., 2000) using an admixture model of the diploid genotypic calls. STRUCTURE places clones in subpopulations based on similar patterns of variation. For each dataset, three replicates were performed for each value of K from one to ten with a 50,000 burn-in time, and the number of Markov Chain Monte Carlo replicates also set to 50,000. For each K, we checked whether the run parameters (likelihood, posterior probability of data and alpha) reach convergence within the burn-in period. The most probable K-value was determined by STRUCTURE Harvester (Earl & vonHoldt, 2012), using the log probability of the data [LnP(D)] and delta K (ΔK) based on the rate of change in [LnP(D)] between successive K-values.

2.2.6. Discriminant analysis of principal components (DAPC)

DAPC was done using the *adegenet* package (Jombart & Ahmed, 2011) in R to identify and describe clusters based on genetic relationships using a diploid form of

genotyping data. The feature *find.clusters* was used to identify the number of clusters within the population. The K-means clustering decomposes the variable's total variance into between-group and within-group components. The lowest associated BIC had defined the best number of subpopulations. The correct number of principal components (PCs) to be maintained was verified using a cross-validation feature (*Xval.dapc*). In this analysis, the data is divided into two sets: a training set (90 percent of the data) and a validation set (10 percent of the data). The members of each group are chosen by stratified random sampling, ensuring that at least one member of each group or population is reflected in the original data in both training and validation sets. DAPC is performed on the training set with a variable number of retained PCs, and the degree to which the analysis can accurately predict group membership of excluded individuals (those in the validation set) is used to determine the optimum number of retained PC. The sampling and DAPC procedures are repeated many times at every PC retention level. The best number of PCs that should be taken is associated with the lowest root mean square error. *SNPZIP* analysis was used to identify alleles with the largest contributions to form the linear discriminants and allocate the genotypes to the clusters. The coefficient of genetic differentiation among groups (F_{st}) was calculated using *stamppFst* in StAMPP package (Pembleton et al., 2013) in R.

2.2.7. Hierarchal clustering

Pairwise Nei genetic distance (Nei, 1972) was calculated, and a distance matrix was obtained with the StAMPP package (Pembleton et al., 2013) of R software using the tetraploid SNP genotype calls. The resulting matrix was used to build a dendrogram

using the hierarchical clustering (method= “Ward D”) implementing in the *Ape* package (Paradis et al., 2004) in R. Duplications, mislabeling, and errors with the naming were identified from the dendrogram based on clustering. After removing duplicates and mislabeled clones, a core set of clone bank collection was developed for long-term *in-vitro* maintenance.

2.2.8. Identification of selection signatures

Signatures of selection analyses were performed using 10,106 SNPs applying three complementary statistical methods. The outlier test PCAdapt (Luu et al., 2017) was based on allele frequency differentiation whereas, the iHS (Voight et al., 2006) and the XP-EHH (Sabeti et al., 2007) were based on linkage disequilibrium (LD) patterns.

PCAdapt Version 4 (Privé et al., 2020) was used to identify loci related to diversification in R. The option for performing LD clumping was applied, this removes variants in LD and ensures that more PCs capture population structure instead of LD structure (Privé et al., 2020). The initial number of PCs was set as K = 20, and the scree plot was used to pick the K that explains much of the variance. The choice of K was also verified by projecting individuals on the principal components (called PCAdapt's score plot) to see if the clustering level was consistent with the value selected for K. The Mahalanobis distances were then used to search for outlier SNPs and transformed into p-values to perform hypothesis testing. A Q-Q plot of the predicted p-values vs. observed p-values was used to visualize the distribution of the p-values. The cut-off for identifying selections was then based on the q-value method using the qvalue R package, using 5% as false discovery rate threshold.

SHAPEIT2 (Delaneau et al., 2013) with a window size of 1 Mb and 500 iterations, including 200 burn-in and pruning iterations, was used to derive haplotypes for iHS and XP-EHH analyses. The iHS and XP-EHH analysis was done using the rehh package (Gautier & Vitalis, 2012) in R v. 3.4.4. To allow better visualization and analysis of regions under selection, the iHS and XP-EHH scores were standardized to a distribution with zero mean and unit variance. In addition, p-values were calculated with the threshold set at 1 percent, as defined in Gautier and Naves (Gautier & Naves, 2011) and FDR performed following Storey and Tibshirani (Storey & Tibshirani, 2003). Candidate selection sweep regions were classified as SNP regions identified as being under selection by at least two of the statistics applied. Genes spanning 250 kb upstream and downstream of the candidate selection regions were retrieved from the representative gene annotation for the pseudomolecules from the Potato Genome Sequencing Consortium (PGSC) public data (Pham et al., 2020; Sharma et al., 2013; X. Xu et al., 2011) retrieved from http://solanaceae.plantbiology.msu.edu/pgsc_download.shtml.

2.2.9. Core Set Identification

A core set of most diverse clones was identified using Core Hunter 3 (De Beukelaer et al., 2018). This software generates subsets based on multiple genetic measures, including both distance measures and allelic diversity indices (http://www.corehunter.org). The function *sampleCore* in the R package of Core Hunter 3 was run on a precomputed Nei's distance matrix of 214 clones.

2.2.10. Pedigree information

A curated dataset was used to check the accuracy of pedigree records using the methodology of Endelman et al. (2017). Pedigree information was assembled from variety release publications, published potato pedigree database, and TAMU potato program breeding records. If both parents were genotyped, the pedigree conflict rate was used to identify pedigree errors. For each of the parent-offspring trios (two parents and one offspring) in the dataset, a pedigree conflict metric was calculated as the percentage of monomorphic (i.e., non-segregating) markers in the cross at which the genotype of the offspring was different. When only one parent was genotyped, the marker vs. pedigree plot was used to confirm (or not) the known parent.

2.3. Results

Two hundred fourteen clones, including commercial and reference varieties maintained by the TAMU Potato Breeding Program, were genotyped with the Infinium 22K V3 Potato Array. Stringent screening of the SNP markers using MAF removed 10,669 SNP (50.7%) markers and additional filtering for more than 10% “No call rate” removed 252 (1.2%) SNP markers. After filtering, a total of 10,106 polymorphic SNP markers were selected for analysis.

2.3.1. Genome-wide distribution of SNPs

The SNPs were distributed across the 12 chromosomes. 10,106 SNP markers (after filtering) were mapped to 12 chromosomes represented as the 12 pseudomolecules of the potato genome DMv4.03 (Xu et al., 2011). Each chromosome had an average of 842 markers ranging from 1,389 markers on Chr. 1 to 617 on Chr. 10. The average

distance between SNPs was 71 kb, but the SNP-to-SNP distribution was skewed: 39% of the marker to marker distances were less than 1 kb, and 18% were less than 10 kb. SNPs were enriched toward chromosome ends (Figure 2.1).

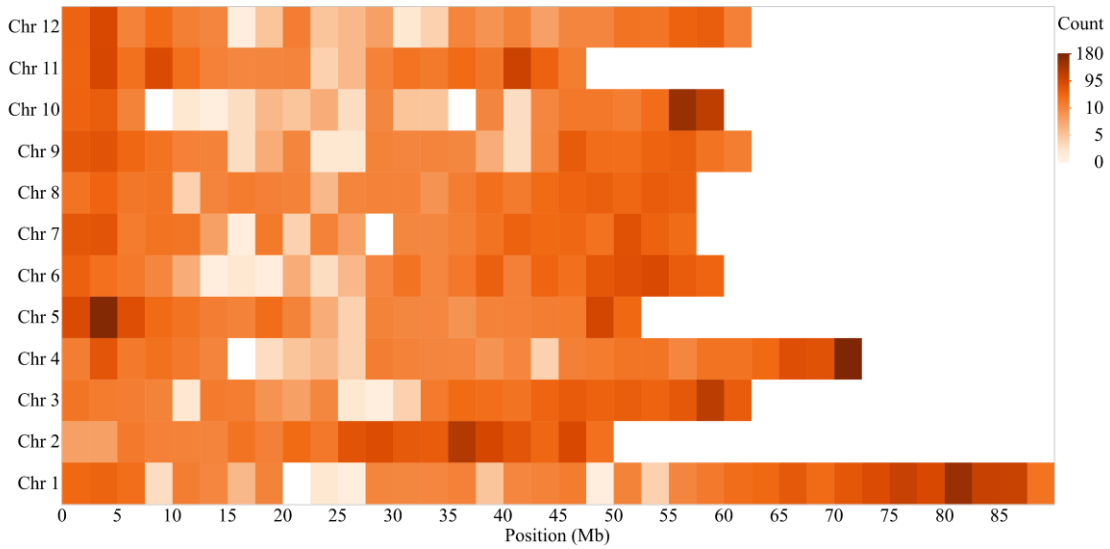


Figure 2.1 Heatmap of 10,106 SNPs across the twelve potato chromosomes. The color intensity indicates the density of markers in that segment of the chromosome (white, low density; maroon, high density). SNP density is shown to increase toward the ends of the chromosomes where gene density is higher.

2.3.2. Evaluation of SNP characteristics

The mean expected heterozygosity value of the SNP markers was 0.39, ranging from 0.10 to 0.50. Minor allele frequency (MAF) ranged from 0.05 to 0.50, with a mean of 0.31 (Figure 2.2). The polymorphic information content (PIC), which denotes the relative informativeness of each marker, ranged from 0.09 to 0.38 with a mean of 0.31 (Figure 2.3). Most of the clones had high levels of heterozygosity, ranging from 0.22 to 0.80 with a mean of 0.59 (Appendix A12). The mean heterozygosity values for different

market classes were 0.62 (Chipping), 0.59 (Russet), and 0.58 (Red/Specialties). A clone (ATX91322-2Y/Y) with very low frequencies of simplex and triplex was found. Those two allelic classes are absent in diploids. ATX91322-2Y/Y produces very small potatoes, yellow skin, and very intense yellow flesh. Thus, we are declaring this clone as a diploid potato. The inbreeding coefficient was negative for many highly heterozygous clones ranging from -1.00 to 0.44, with a mean of -0.51 (Appendix A12).

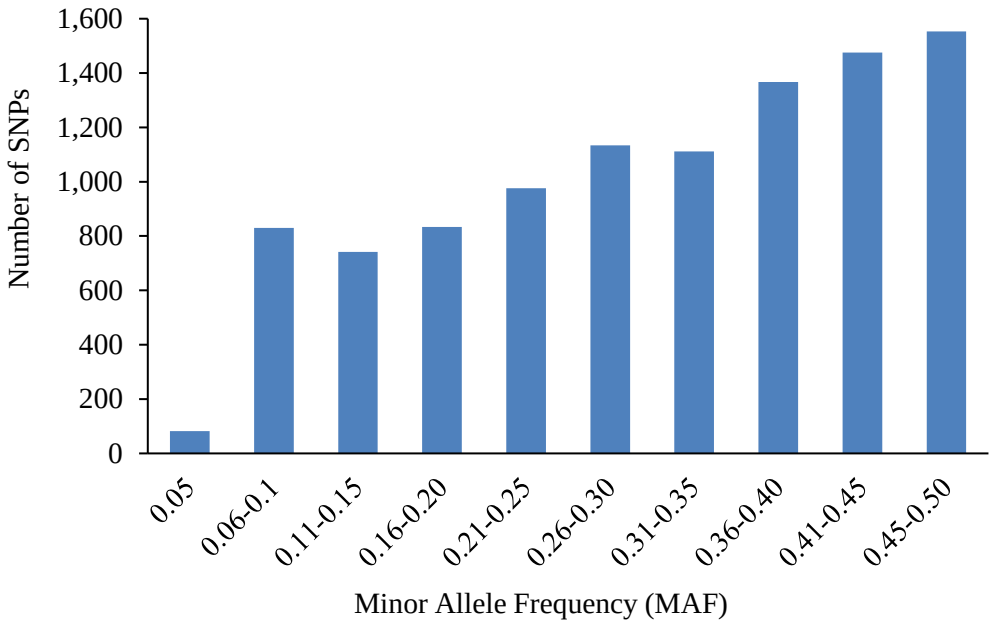


Figure 2.2 Distribution of minor allele frequency (MAF) of 10,106 SNPs in 214 tetraploid clones.

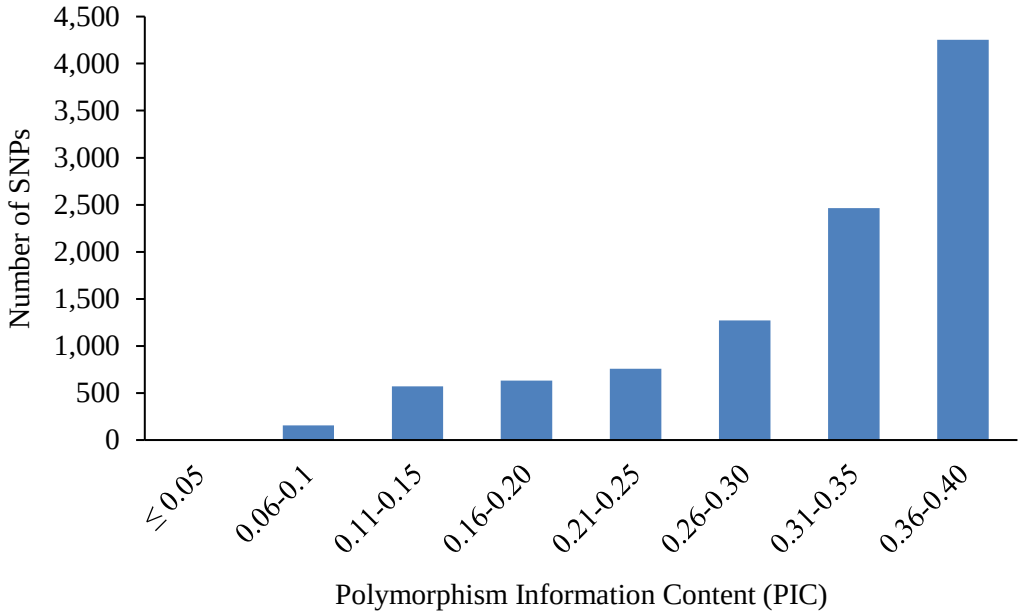


Figure 2.3 Distribution of polymorphic information content (PIC) values of 10,106 SNPs calculated for 214 tetraploid clones.

2.3.3. Genetic diversity

Based on the diploid genotypic calls analysis using TASSEL, the average pairwise divergence among genotypes (π), at SNP locations, was 0.39. This represents the nucleotide diversity per assayed SNP in the clones. The expected number of polymorphic sites per nucleotide (θ), which estimates the mutation rate in the collection, was 0.169 with 10,106 segregating sites. Tajima's D, which estimates the normalized measure of the difference between the observed (π) and expected (θ) nucleotide diversity was 4.29.

2.3.4. Population structure analysis

STRUCTURE analysis showed that the number of subpopulations (K) ranged from zero to ten when using the diploid genotyping model (AA, AB, BB; 10,106 SNP

markers). The K value with the maximum likelihood was K = 3 (Figure 2.4, Appendix A1). Clones were assigned to a subpopulation if they had at least 50% membership within that group. Most of the reds, purples, and yellows (46.6% of total clones) grouped in subpopulation 1 (Red) (Figure 2.4).

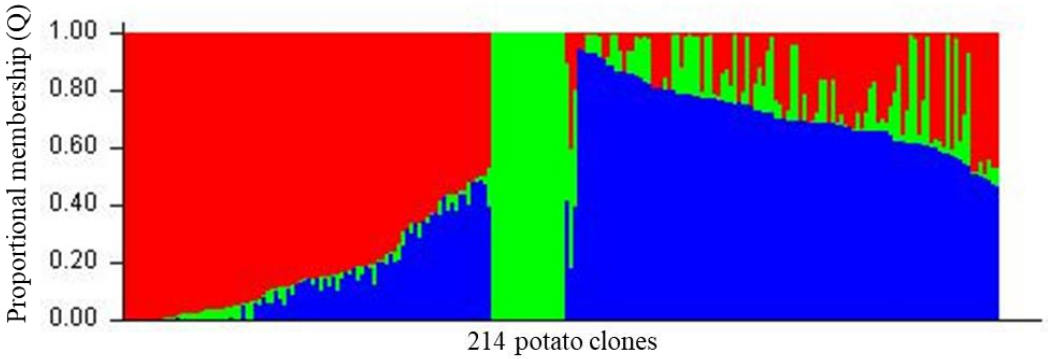


Figure 2.4 Proportional membership (Q) of each clone in the genetic clusters inferred by STRUCTURE (K= 3). This figure represents each individual as a vertical bar and its membership probability in each subpopulation. Individuals with the highest proportion of membership to subpopulation 1 (red color) corresponded to clones with red, purple, and yellow skin; Individuals belonging mainly to subpopulation 2 (green) include Russet Norkotah strain selections, and; Individuals with predominate membership to subpopulation 3 (blue) were russet and chipping clones.

For instance, the red skinned yellow flesh clone Sierra Rose, the purple skinned yellow flesh clone ATTX88654-2P/Y, and the yellow skin yellow flesh clone ATX91322-2Y/Y had complete membership in subpopulation 1. Russet Norkotah and its strain selections (8.4% of total clones) were grouped in a separate subpopulation 2 (Green) (Figure 2.4). Most of the russet and chipping clones (40.6% of total clones) were grouped to subpopulation 3 (Blue) (Figure 2.4). For instance, the chipping clone Atlantic and the russet clone Reveille Russet had complete membership to subpopulation 3.

STRUCTURE analysis revealed significant admixtures in 4.67% of the total clones. e.g., White LaSoda, TX11454-9Ru/Y, and COTX87601-2Ru (Figure 2.4; Appendix A2).

2.3.5. Discriminant analysis of principal components (DAPC) analysis

The lowest Bayesian information criterion (BIC) value obtained using *find.clusters* function was three (Appendix A3), which was in concordance with the delta K obtained in STRUCTURE. These three clusters were used to analyze the DAPC (Figure 2.5). Twenty principal components capturing 34.3% variance and two discriminant eigenvalues were retained. These values were confirmed by a cross-validation analysis (Appendix A4). Genotypes had membership coefficients to each group ranging from 0.5 to 1, thus confirming low admixture and high structured population. Exceptions to these values were clone NDTX059775-1W (chipper with white flesh), COTX10118-4Wpe/Y (specialty with white skin purple eyes, and yellow flesh), COTX03079-1W (chipper with white flesh), and COTX94216-1R (red skin white flesh) whose values were 0.36, 0.40, 0.43 and 0.44, respectively. In Figure 2.5, linear discriminant 1 separated Russet Norkotah and Red groups from the Chip & Russet group and linear discriminant 2 separated Red and Chip & Russet groups from the Russet Norkotah group. *SNPZIP* analysis detected 18 SNPs with the largest contribution to cluster identification. Two of them corresponded to linear discriminant 1, and the remaining 16 to the linear discriminant 2. Most of them annotated with known gene functions. The coefficient of genetic differentiation among groups was highest (0.14) between Red/Specialties and Russet Norkotah/strains, followed by Chip & Russet and

Russet Norkotah/strains (0.10). The lowest value (0.02) was found between the Chip & Russet and Red/Specialties groups suggesting low genetic differentiation among them.

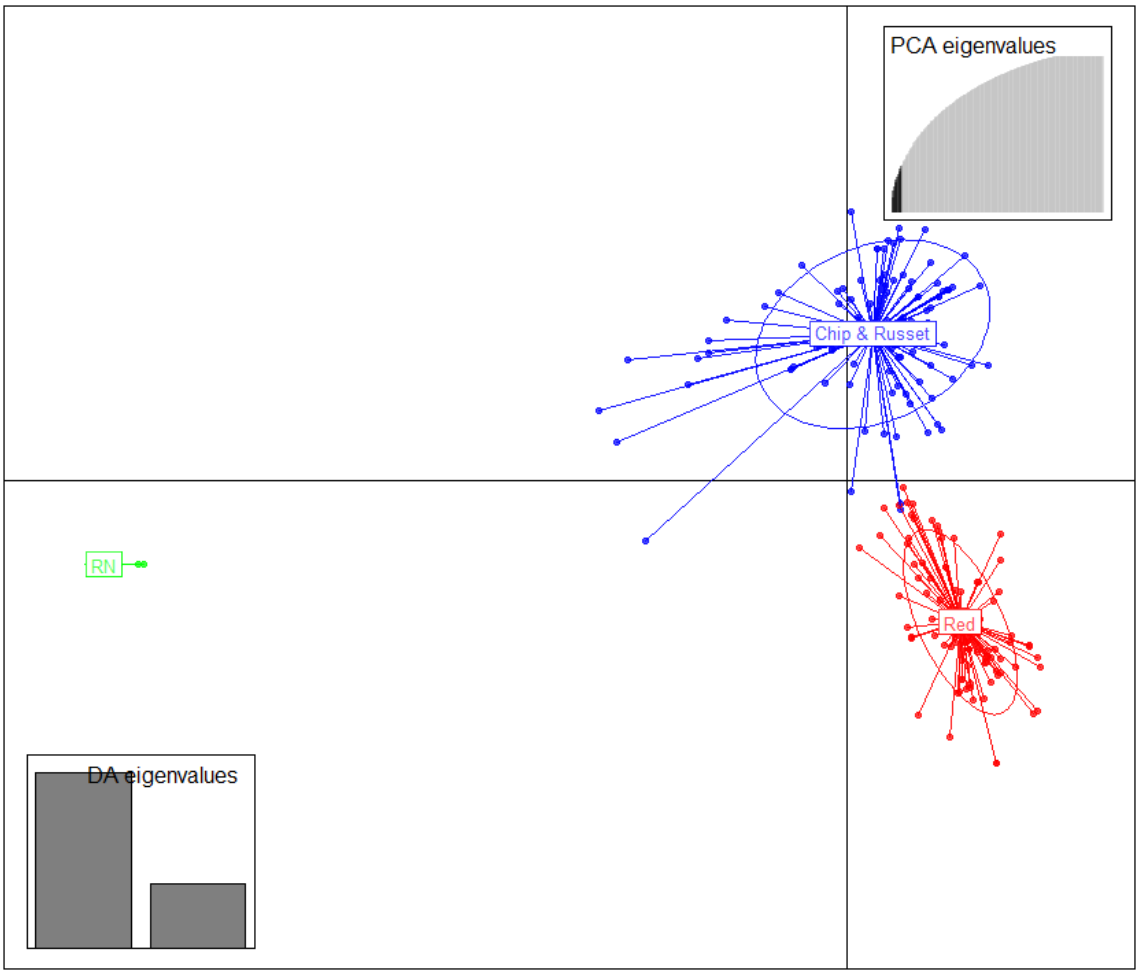


Figure 2.5 Discriminant analysis of principal components (DAPC) for 214 clones. The axes represent the first two linear discriminants. Circumferences surround each group, and small solid dots represent individual clones. Labels inside circles indicate the different subpopulations identified by DAPC analysis (Chip & Russet= chipping and russet clones, Red= red/specialties, and RN= Russet Norkotah and its strains).

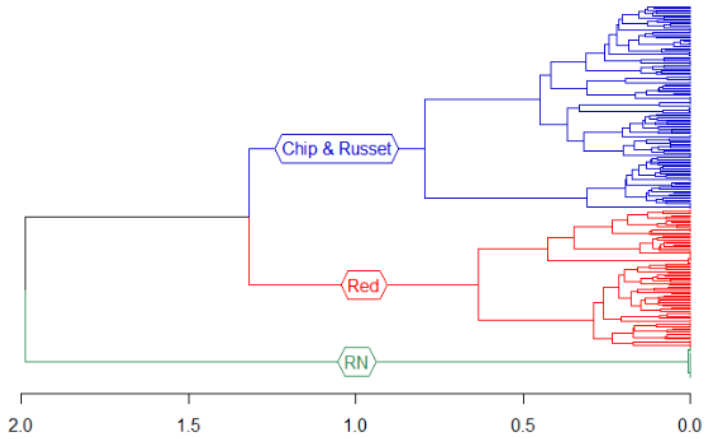
2.3.6. Phylogenetic cluster analysis

The dendrogram generated using Nei genetic distance and hierarchical clustering also revealed the presence of three clusters in the population (Figure 2.6). The assignment of the clones to the groups in the dendrogram corresponded to 92% and 93% with the allocation made by the STRUCTURE and DAPC analysis, respectively. Groupings of the clones were observed based on their lineage/pedigree. As a result, selections with one or both common parents clustered together along with their parental clones in the same group. However, the clones were not separated based on cross location. Cluster 1 (18 clones) comprised mainly of Russet Norkotah, its eight strains (TXNS 106, TXNS 118, TXNS 249, Russet Norkotah 102, Russet Norkotah 112, Russet Norkotah 223, Russet Norkotah 278, and Russet Norkotah 296), and nine other russet clones (Figure 2.6a). This is equivalent to 27% of the total russet clones used in this study. The dendrogram shows very low/no genetic distances between them. The origin of the clones in this cluster traces to crosses made by four breeding programs (Idaho, North Dakota, Colorado, and Texas). Similarly, cluster 2 (94 clones) was comprised mostly of reds, yellows, and purple clones (Figure 2.6b). This is equivalent to 96%, 40%, and 76% of the total red, total yellows, and total purple clones, respectively. In the cluster, Sierra Rose^{TM} and four additional red clones were distinct from the remaining clones in the group. Four chipping clones (AOTX95309−2W, ATTX95490−2W, TX12484−4W, and NDTX059828−2W) appeared as exceptions in this cluster consisting predominantly of red clones. The most prominent varieties in this cluster include White LaSoda, Sierra Rose^{TM,} and Rio Rojo. The cross-location of the clones in this cluster

traces to eight potato breeding programs in the United States. Lastly, Cluster 3 (102 clones) was comprised of chipping clones, russets, yellows, and purple. This is equivalent to 90%, 73%, 59%, and 23% of the total chipping clones, total russets, total yellows, and total purple clones, respectively (Figure 2.6c). Seven red clones appeared as exceptions in this cluster. COTX03187−1W grouped with russets rather than grouping with chippers. The most prominent varieties in cluster 3 include Atlantic, Tacna, Tokio, Sierra Gold^{TM}, Krantz, and a recently released Texas variety COTX09022-3RuRE/Y (russet skin red eyes and yellow flesh, released under the experimental name). The origin of the clones in this cluster traces to crosses originally made by eight potato breeding programs in the United States.

Under the current naming convention, the Texas Potato Breeding program typically uses a clone code that provides information about the place where the cross was made, where it was selected, year the cross was made, family number, selection number, type of skin, and sometimes type of flesh and other characteristics. For example, in ATX91137-1Ru, ‘A’ indicates the cross was made in Aberdeen, Idaho, TX= selected in Texas, 91 = year cross was made, 137 = family number, -1= selection number, and Ru = russet skin. After inspecting the dendrogram, we observed mislabeling in 5 clones (2.3 % of the total clones). For instance, a russet clone AOTX98096−1Ru was mislabeled as a red clone AOTX98096−1R. Likewise, COTX04303-3Ru/Y was mislabeled as COTX04303-3R/Y. The SNPs grouped both of them with russets and Russet Norkotah strains in the dendrogram. Inspection of parentage gave a hint about the error and the minitubers produced in the greenhouse further confirmed that these should

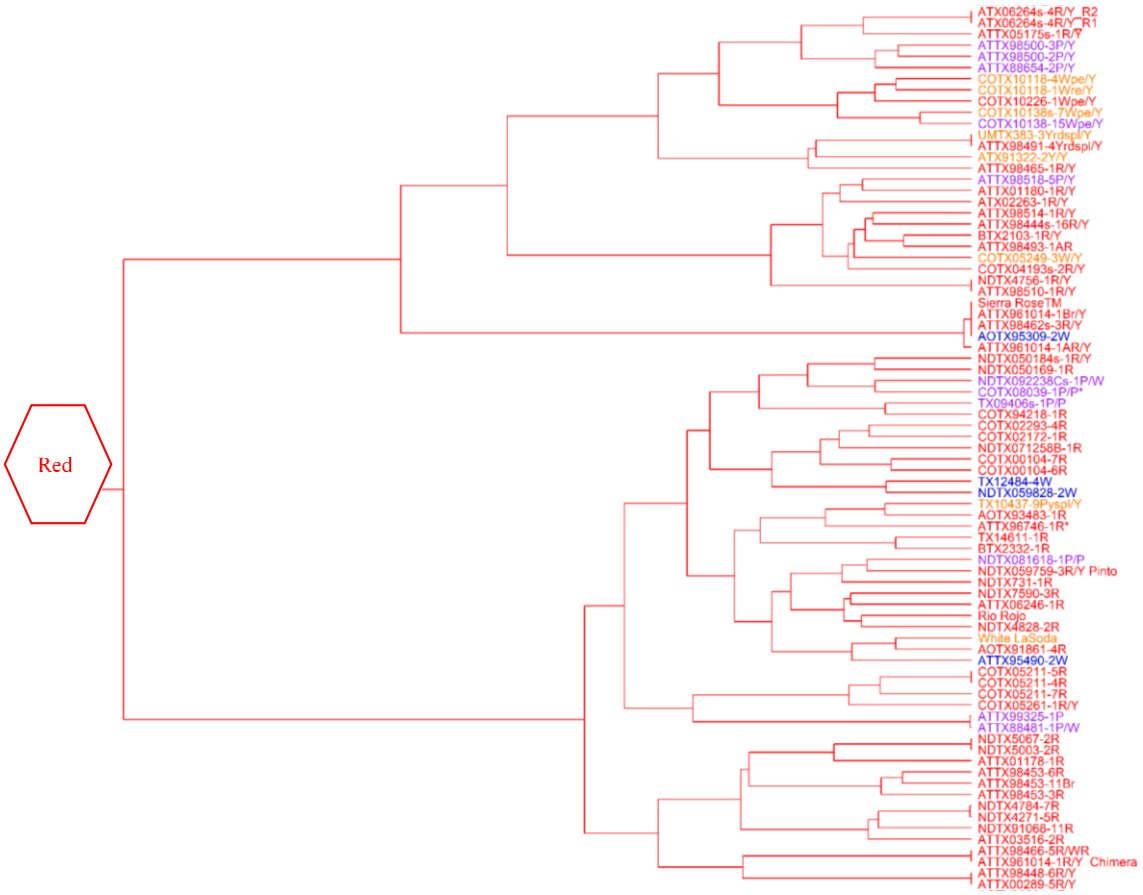
be russet clones (Appendix A5). All the corrected names are reflected in the dendrogram with an asterisk sign (*) at the end of the name. Atlantic was repeated (TAMU and MSU versions) as a quality control to detect duplications and they had almost zero Nei's distance between them. After SNP comparisons, we found that some clones were identical. For instance, sister lines TX09403-15W and TX09403-21W cluster together and had almost zero Nei's distance between them. Similarly, Russet Norkotah and Russet Norkotah strains could not be distinguished by the SNPs used in the current study. In another instance, clone AOTX95309-2W did not group with sister line AOTX95309-1W. AOTX95309-2W clustered together with Reds and tubers were red. Based on the dendrogram, parentage, and tuber color, the clone AOTX95309-2W could be considered an unintended mix and should be removed from the program. The use of SNP genotyping aided the discovery of typographic errors that occur during handling clonal material in the breeding program and/or tissue culture operations. Further, SNPs can also be used to define unique molecular fingerprints of released varieties and advanced clones and to calculate similarities (or distances) between new varieties and reference varieties and other released varieties.



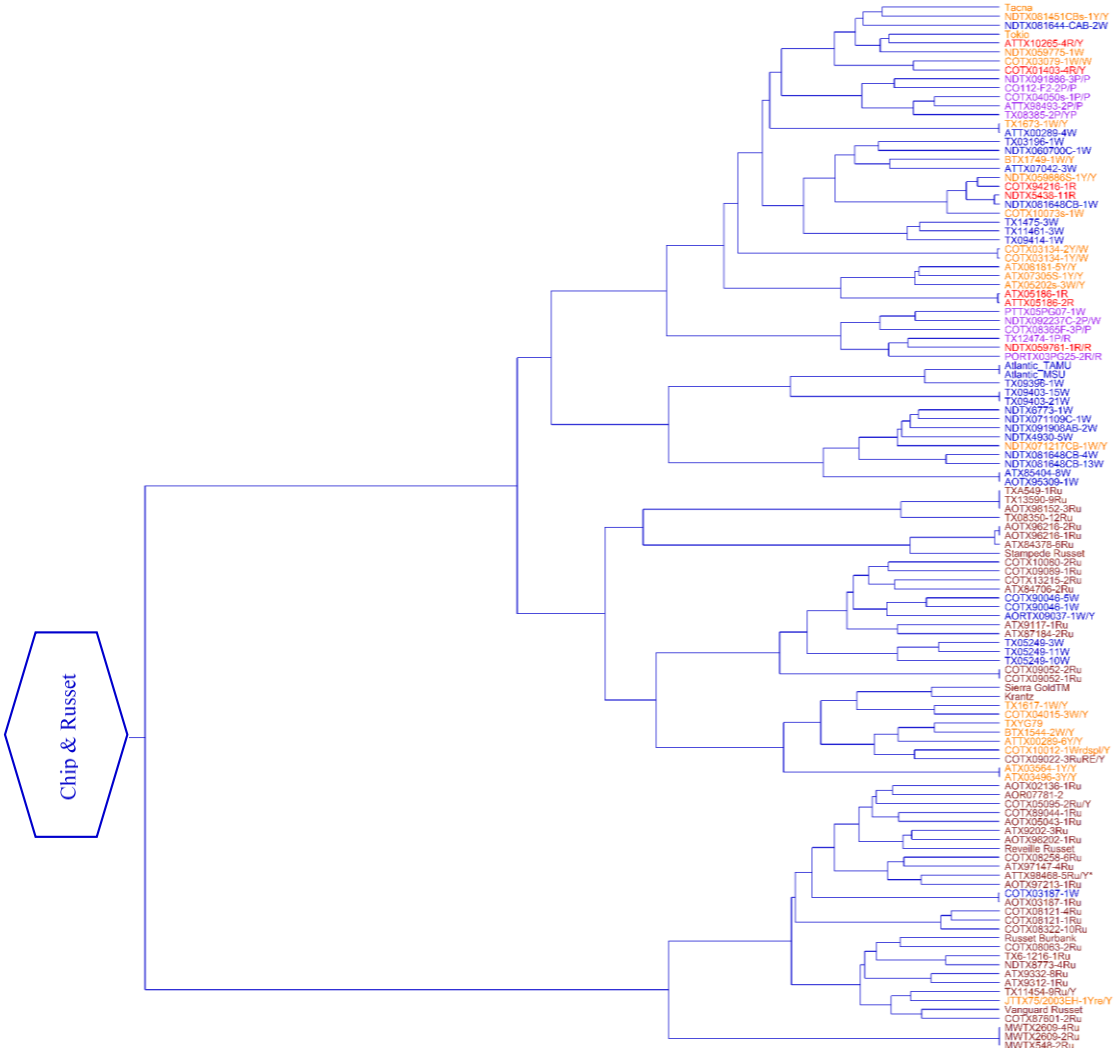
a.

Figure 2.6 A Ward Dendrogram of the 214 clones using hierarchical clustering (method = "ward D"), the lower part a. representing Russet Norkotah and its strains, the middle part b. representing red and specialties and top part c. representing chipping and russet clones are shown separately. In the X-axis are represented the Nei's genetic distances between clones. The color of the clones represents the market class (Red = red clones, Purple = purple clones, and Yellow = yellow clones; Green = Russet Clones, and; Blue = chipping clones). Corrected names are indicated by an asterisk sign (*) at the end of the clone name.

b.


Red
ATX06264s-4R/Y_R2
ATX06264s-4R/Y_R1
ATTX05175s-1R/Y
ATTX98500-3P/Y
ATTX98500-2P/Y
ATTX88654-2P/Y
COTX10118-4Wpe/Y
COTX10118-1Wre/Y
COTX10226-1Wpe/Y
COTX10138s-7Wpe/Y
COTX10138-15Wpe/Y
UMTX383-3Yrdspl/Y
ATTX98491-4Yrdspl/Y
ATX91322-2Y/Y
ATTX98465-1R/Y
ATTX98518-5P/Y
ATTX01180-1R/Y
ATX02263-1R/Y
ATTX98514-1R/Y
ATTX98444s-16R/Y
BTX2103-1R/Y
ATTX98493-1AR
COTX05249-3W/Y
COTX04193s-2R/Y
NDTX4756-1R/Y
ATTX98510-1R/Y
Sierra RoseTM
ATTX961014-1Br/Y
ATTX98462s-3R/Y
AOTX95309-2W
ATTX961014-1AR/Y
NDTX050184s-1R/Y
NDTX050169-1R
NDTX092238Cs-1P/W
COTX08039-1P/P*
TX09406s-1P/P
COTX94218-1R
COTX02293-4R
COTX02172-1R
NDTX071258B-1R
COTX00104-7R
COTX00104-6R
TX12484-4W
NDTX059828-2W
TX10437-9Pyspl/Y
AOTX93483-1R
ATTX96746-1R*
TX14611-1R
BTX2332-1R
NDTX081618-1P/P
NDTX059759-3R/Y Pinto
NDTX731-1R
NDTX7590-3R
ATTX06246-1R
Rio Rojo
NDTX4828-2R
White LaSoda
AOTX91861-4R
ATTX95490-2W
COTX05211-5R
COTX05211-4R
COTX05211-7R
COTX05261-1R/Y
ATTX99325-1P
ATTX88481-1P/W
NDTX5067-2R
NDTX5003-2R
ATTX01178-1R
ATTX98453-6R
ATTX98453-11Br
ATTX98453-3R
NDTX4784-7R
NDTX4271-5R
NDTX91068-11R
ATTX03516-2R
ATTX98466-5R/WR
ATTX961014-1R/Y Chimera
ATTX98448-6R/Y
ATTX00289-5R/Y

c.



2.3.7. Identification of candidate loci under selection

Using the proportion of explained variance displayed, and projecting individuals on the principal components as a score plot (Appendix A6), we estimated the optimal number of PCs from the SNP matrix to be three. At α=0.05 corrected for the genomic inflation factor (λGC=1.20), 26 SNPs were found under selection on chromosomes 1, 2, 3, 4, 5, 7, 8, and 10

using the PCA-based method (Appendix A13). Some of the selected SNPs had known functions. For example, a SNP (PotVar0120627) was selected at 48.6 Mb on chromosome 3. It had been reported that the *Y*-locus controlling the white-to-yellow flesh color in potato mapped to chromosome 3 (Bonierbale et al., 1988) and is believed to be regulated primarily by the b-carotene hydroxylase (*BCH*) gene (Kloosterman et al., 2010). Likewise, after adopting the false discovery rate of 0.01, 127 SNPs and 100 SNPs were found under selection using the iHS and XP-EHH tests, respectively. Figure 2.7 to Figure 2.9 shows the Manhattan plots illustrating the SNPs identified as being under selection pressure on all potato chromosomes according to the three tests assayed.

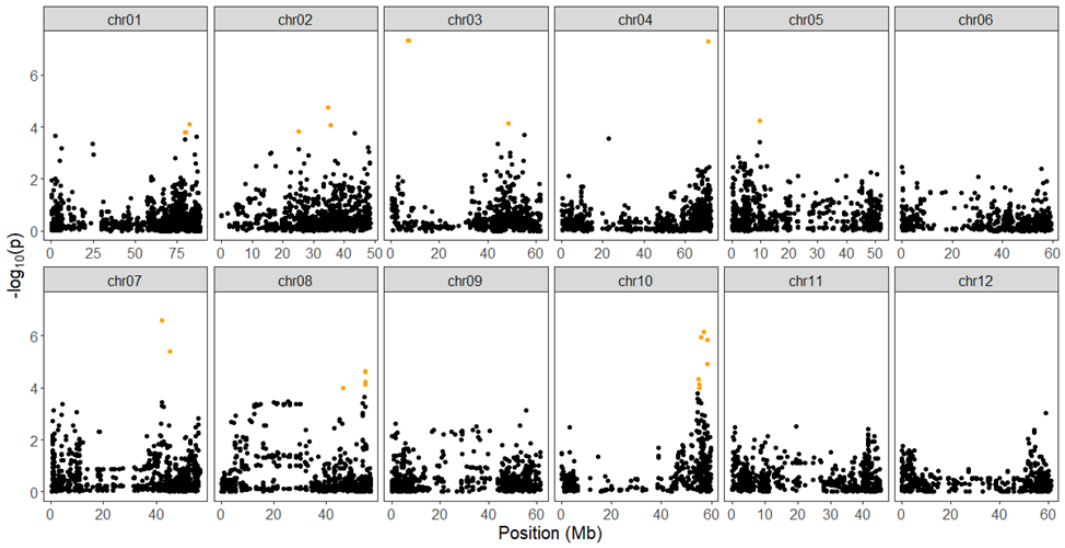


Figure 2.7 Manhattan plot showing the distribution of candidate outlier SNPs from PCAdapt where Y-axes represent P values and significant SNPs above α=0.05 are in orange color.

Eighteen regions were identified under selection by at least two of the statistics applied and were defined as candidate selection sweep regions. These regions occur on all chromosomes except 3 and 5 (Table 2.1). These SNPs were related to diversification and some of them were found associated with a specific phenotype. Some of the candidate genes had known functions and are discussed in the next section.

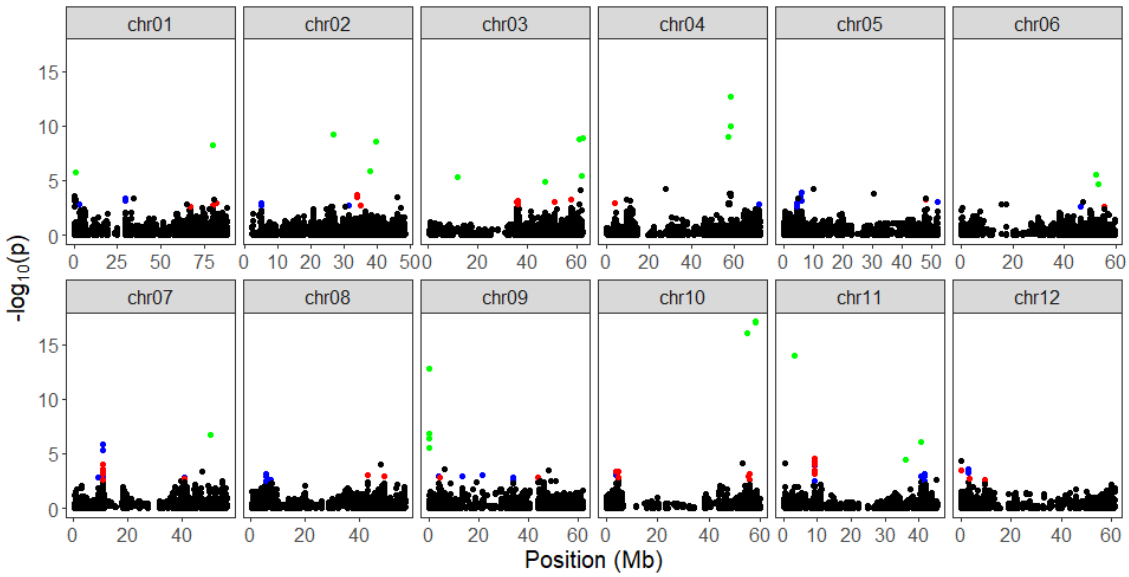


Figure 2.8 Distribution of standardized iHS scores in three groups of potatoes. Significant SNPs above 1% false discovery rate (FDR) threshold are colored according to groups (iHS_chipru = blue; iHS_red = red; iHS_rn = green)

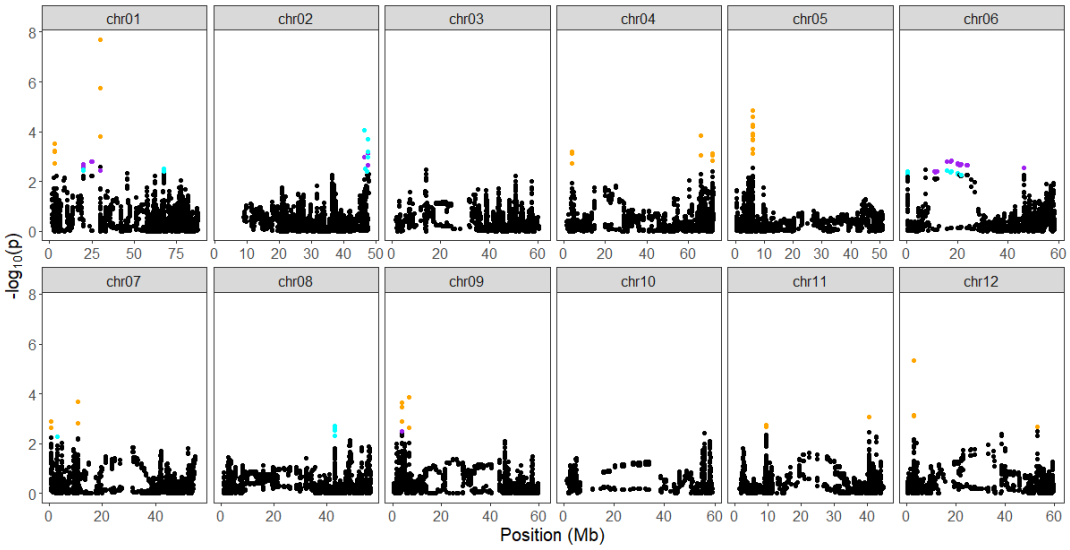


Figure 2.9 Distribution of standardized XP-EHH scores between three groups of potatoes. Significant SNPs above 1% false discovery rate (FDR) threshold are colored according to groups (XP-EHH chirpru vs red = orange; XP-EHH chipru vs rn = purple; XP-EHH red vs rn = cyan)

Table 2.1 Description of the candidate selective sweep regions detected using PCAdapt, iHS and XP-EHH analyses in potato.

Chr	Selective sweep region (Mb)	No. of candidate genes	Top significant SNP	Functional Annotation	Max iHS/XPEHH/PCAdapt statistic	P-val
1	2.60-3.10	47	PotVar0044963	Flesh Color	3.1, 3.4	2.8, 3.2
1	19.41-19.91	11	solcap_snp_c2_49758		3.1, 2.9	2.7, 2.5
1	29.43-29.93	15	PotVar0122478		-3.5, -5.6	3.4, 7.7
1	66.98-67.48	44	PotVar0098497		-3.0, -2.9	2.6, 2.5
2	46.15-46.65	68	solcap_snp_c2_24864	Length of plant cycle and tuberization	3.3, 3.9	3, 4
4	3.66-4.16	37	PotVar0106879		3.2, -3.1	2.9, 2.7
6	15.52-16.02	12	solcap_snp_c2_18787		3.2, 2.9	2.8, 2.4
6	17.14-17.64	8	solcap_snp_c1_9601		3.2, 2.9	2.8, 2.4
6	20.06-20.56	14	solcap_snp_c2_56793		3.1, 2.8	2.7, 2.3
6	21.00-21.50	16	solcap_snp_c2_33233		3.1, 2.8	2.7, 2.2
6	21.48-21.98	8	PotVar0083629		3.1, 2.8	2.7, 2.2
7	10.58-11.08	20	solcap_snp_c1_2404	Stolon attachment	3.2, -3.4	2.9, 3.2
7	40.36-40.86	25	solcap_snp_c2_9380		3.2, 3.1	2.8, 2.7
8	42.87-43.37	32	PotVar0086811		3.3, -2.8	3, 2.3
9	3.48-3.98	43	PotVar0012376	Flesh Color	-3.6, -3.0	3.5, 2.5
10	55.60-56.10	57	PotVar0005291	Cytokinin metabolism, Pelargonidin	30.5, 3.4	6.0, 3.2
11	9.12-9.62	28	solcap_snp_c2_53683		-4.1, -3.5	4.4, 3.3
12	2.83-3.33	54	PotVar0031150		3.7, 3.4	3.6, 3.2

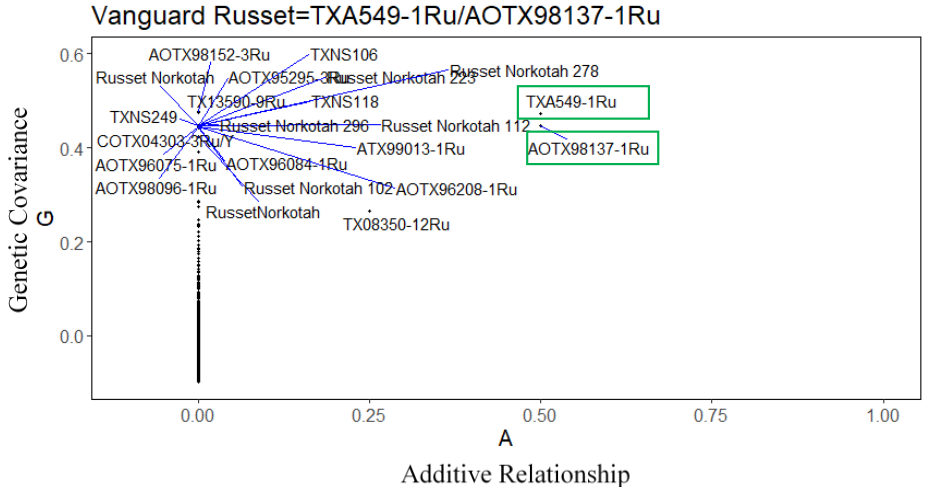


Figure 2.10 Population-wide comparison of genetic covariance calculated from markers with the additive relationship calculated from pedigree records, for clone Vanguard Russet.

2.3.8. Core set identification

The analysis of genetic diversity and population structure of 214 clones identified sub-populations in the clone bank and some of the genotypes were quite similar. A core set of 43 clones (Appendix A14) was selected to maximize diversity and minimize redundancy using Core Hunter 3 software. Among the core set, 14 clones were from the Chipping market class, 11 Russet, and seven, five, and six from the Red, Purple, and Yellow market classes, respectively. The genetic diversity of the core set was estimated to represent the extent of diversity captured from the total collection. Comparisons of all genetic parameters indicated that the values for the core set were almost equal to those for the total collection (Table 2.2). The mean genetic distance of the whole collection

was 0.09, but this value increased to 0.10 in the core set. Similarly, the mean PIC and the mean MAF of the whole collection were both 0.31, while those of the core set were 0.31 and 0.30, respectively. DAPC analysis and hierarchical clustering showed the presence of four clusters in the core set (Appendix A7 and Appendix A8).

Table 2.2 Comparison of the genetic diversity of the whole collection (214 clones) versus a core set (43 clones) based on genetic distance, polymorphism information content (PIC) and minor allele frequency (MAF).

Nei's genetic distance				PIC				MAF			
Whole collection (214 clones)		Core set (43 clones)		Whole collection (214 clones)		Core set (43 clones)		Whole collection (214 clones)		Core set (43 clones)	
Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
0.09	0-0.17	0.10	0.07-0.17	0.31	0.09-0.38	0.31	0.10-0.38	0.31	0.05-0.50	0.30	0.06-0.50

2.3.9. Pedigree information

For 12 of the parent-offspring trios having genotyping data for both parents, pedigree was found accurate for ten trios with no pedigree conflict. Figure 2.10 shows the marker vs. pedigree plot for Vanguard Russet without pedigree errors. The parents are in the top right corner of the figure, with both an additive relationship (A) and genetic covariance (G) of 0.5 and unrelated individuals with additive relationship A=0 to the parents. One of the grandparents of Vanguard Russet, TX08350-12Ru is plotted at A=0.25 and has less genetic covariance than the parents based on markers. Whereas, for clone NDTX4930-5W and TX11461-3W the conflict rate was 24.5% and 19%, respectively (Table 2.3). The male parent of NDTX4930-5W and female parent of

TX11461-3W is found erroneous from the marker vs. pedigree plot. In the case of 12 clones with one parent genotyped, five parents were found correct and seven parents were found erroneous (Appendix A9-A11)

Table 2.3 Parent-offspring trios for which both recorded parents were genotyped. Pedigree conflict rates are shown in percentages.

Clone Name	Market Class	Female Parent	Male Parent	Pedigree Conflict %
Atlantic	Chipping	Wauseon	Lenape	0.0
NDTX4930-5W	Chipping	ND860-2	**Western Russet**†	24.5
TX09396-1W	Chipping	Atlantic	Lamoka	0.0
TX09403-21W	Chipping	Waneta	Ivory Crisp	0.0
TX11461-3W	Chipping	**Waneta**†	Chipeta	19.0
TX09403-15W	Chipping	Waneta	Ivory Crisp	0.0
COTX08121-1Ru	Russet	AC96052-1Ru	Blazer	0.0
COTX08121-4Ru	Russet	AC96052-1Ru	Blazer	0.0
TX08350-12Ru	Russet	TXA549-1Ru	AC96052-1Ru	0.0
Vanguard Russet	Russet	TXA549-1Ru	AOTX98137-1Ru	0.0
COTX08322-10Ru	Russet	Blazer	AC96052-1Ru	0.0

† The erroneous parent is shown in bold.

2.4. Discussion

Tissue culture clone banks often contain potato varieties and breeding lines from several different regions and programs. Characterizing breeding collection germplasm is crucial in plant breeding, as the genetic advancement of economically valuable traits relies on the genetic diversity available within the breeding gene pool. Knowledge about

genetic diversity also assists in minimizing the use of closely related clones as parents in breeding programs, which might lead to a high risk of inbreeding depression and reduced genetic variation. Genetic diversity, population structure, and molecular marker knowledge may accelerate the selection of desirable traits in potato. In the present study, population structure and genetic diversity were evaluated in a tissue culture clone bank collection composed of 214 diverse clones. This collection contains advanced selections from the TAMU Potato Breeding Program entered in tissue culture over 40 years of breeding efforts.

The availability of SNP arrays has enabled genotyping the germplasm of crops like potatoes. SNP distribution across the genome assessed by analyzing filtered SNP density shows the typical pattern of distinctly reduced recombination in pericentromeric regions and increased varying recombination rates in euchromatic regions for all chromosomes. Larger regions with no SNP coverage are usually found in large pericentromeric regions, where repetitive DNA makes it difficult to distinguish unique flanking regions around SNPs (Otyama et al., 2019; Stupar et al., 2002). A similar distribution has been observed in sorghum (Evans et al., 2013), wild tomatoes (Bhardwaj et al., 2016), and *Prunus* (Guajardo et al., 2020). SNPs offer high-resolution markers to breeding programs far beyond traditionally used approaches depending solely on pedigrees (Brown, 1989) or phenotypic data (Bretting & Widrlechner, 2010). In the present study, the average PIC value was 0.31. Most SNPs (69%) had PIC values ranging from 0.30 to 0.38, while for the remaining percentage, it was < 0.30. The SNPs having PIC values ranging from 0.25 to 0.5 are considered moderately informative

(Botstein et al., 1980). This may support the idea that, despite breeding efforts, genetic diversity in our set of clones has not been narrowed. Similar PIC values were previously reported in collections of potatoes tested for genetic diversity using the SolCap SNP array (Bali et al., 2018; Berdugo-Cely et al., 2017).

Heterozygosity is an indicator of genetic variability in a population, and it is related to the polymorphic nature at each locus. In this study, a high level of heterozygosity was observed; in the TAMU potato breeding collection, this could be due to the high levels of genetic variation at loci with vital significance for adaptive response to environmental changes. Loss of heterozygosity was related to lower fitness (Manrique-Carpintero et al., 2018). Potato is an outcrossing species; thus, heterozygosity is usually higher than expected. Selection, migration, mutation, hybridization, polyploidization, and introgression elucidate the high diversity of potatoes (Berdugo-Cely et al., 2017). The average percent heterozygosity (59%) observed in the TAMU potato collection was similar to Hirsch et al. (2013) and Igarashi et al. (2019), who reported the average value of 56% and 60%, respectively. However, the lowest heterozygosity value we observed was 22% in clone ATX91322-2Y/Y, which was found to be a diploid selected by the program based upon five cluster genotypes calling. Igarashi et al. (2019) reported that the average percent heterozygosity of the 2x varieties (23%) was much lower than that of the 4x genotypes. Five cluster genotypes calling has successfully been used to predict ploidy determinations of diploid, triploid, and tetraploid samples (Alsahlany et al., 2019; Ellis et al., 2018). When the simplex and triplex frequency was close to zero, the sample was considered to be diploid, and when

the frequency was over 0.20, then the progeny samples were considered to be tetraploid. This finding is similar to Hirsch et al. (2013), indicating that greater ploidy can be correlated with greater heterozygosity, and vice versa. Almost all the highly heterozygous clones had negative inbreeding coefficients, which happens when observed heterozygous clones are larger than expected due to an excess outcrossing. Increasing the heterozygosity of clones and widening their genetic base are important aspects of breeding programs to have desired combinations of abiotic and biotic stress tolerances and high yield. It is clear from the heterozygosity analyses and SNP evaluation, that the Texas A&M potato breeding program harbors considerable genetic diversity.

Tajima's D (Tajima, 1989) provides a distinction between randomly altering loci and non-randomly evolving loci arising from directional selection, introgression, genetic bottleneck, and/or drift. Generally, a positive value of Tajima's D arises from an excess of intermediate frequency alleles and can result from population bottlenecks, structure, and/or balancing selection (Ching et al., 2002). These factors are likely present in potato breeding clones. The observed Tajima's D value of 4.29 in this study is higher than in sorghum (0.30) (Hamblin et al., 2004) and soybean (1.08) (Zhu et al., 2003), both of which show a significant bottleneck in their population history. To elucidate the possibility that the elevated D value occurred due to population subdivision, we assessed population structure as suggested by Pritchard et al. (2000). When SNPs from all chromosomes were included in the analysis, a significant subdivision was observed (Figure 2.4) indicating a relatively heterogeneous population.

Understanding population structure helps allow the successful use of genotypes for breeding purposes. Alleles that are divergent among clusters are a guide to detecting the principal differences due to breeding strategies and different origins among subpopulations. The STRUCTURE analysis provided further insight into the admixture and the number of populations in this collection. Structure analysis identified subgroups, as in other studies (Hirsch et al., 2013; Peter. G. Vos et al., 2015). Our results support hybridization or outcrossing among the individuals and a five percent admixture.

DAPC analysis divided the population into well-defined clusters according to their genetic structure and market classes. The DAPC approach offers an alternative to STRUCTURE software as it does not require the populations to be in Hardy-Weinberg equilibrium, and it can support large sets of data (Campoy et al., 2016). Our results identified good consistency between STRUCTURE and DAPC analysis when admixed clones were not considered.

The clustering of individuals gives interesting clues for increasing diversity in breeding programs and germplasm collections. Clear knowledge of the germplasm structure and clusters assists in parental choice in breeding programs, improving genetic diversity, and enhancing the potential gain from the selection. Both help to increase breeding program efficiency to face new demands from consumers and the industry, as well as new ecological issues like adaptation to climate change and pest resistances. In this study, clustering corresponded with the similarity in the genetic background of the clones. However, clustering was not dependent on the place where the original cross was done suggesting a huge gene flow across breeding programs due to the reciprocal

exchange of true potato seed of unselected families and use in crossing blocks of parental germplasm from potato breeding programs throughout the United States. Bali et al. (2018) were unable to separate Russet potato clones according to the breeding programs they originated from, which was an indicator of the free flow of germplasm among the potato breeding programs. Several quantitative differences (e.g., vine size, maturity, average tuber size, yield, etc.) existed among the strains and Russet Norkotah (Miller et al., 2004). However, clonal selections (strains) were not differentiated genetically despite using more than 11,000 genome-wide SNP markers. Previous studies using Amplified Fragment Length Polymorphisms (AFLPs) and microsatellite markers were also unable to detect differences between intraclonal variants of the potato cultivar Russet Norkotah (Hale et al., 2005). The differences between the strains could be due to epigenetic variation, and most of them may not be observed by SNPs (Xu et al., 2019).

Maintaining consistent and unique clone names in the clone bank is important for future cultivar identification, research, and breeding. There are many instances in the program where naming errors could be introduced. The longer a clone is handled in the program, the greater the potential of mixing or mislabeling. The Infinium 22K V3 Potato SNP Array generated unique genetic fingerprints to identify accessions where errors had occurred. The majority (97.6%) of the clones evaluated had no errors in genetic identity. It is common in most gene banks to have some errors in the collection. Ellis et al. (2018) at the International Potato Center (CIP) found 4.4% of accessions were genetically mismatched, and in some cases, the SNP results identified the mixed accession. Barcoding, automatic data collection, curation, and other quality control strategies will

help to minimize errors. Studies such as this can help identify and correct errors in the breeding program. In addition, SNP fingerprints and genetic distance comparisons can be useful for plant variety protection (PVP), as well as for the verification of the identity of clones in the foundation, certification, and breeding programs.

Our main goal for selection signature analysis was to detect regions that show preferential selection in the genome of potatoes. To accomplish this, we used three different, but complementary, statistical methods: PCAdapt, iHS, and XP-EHH. The use of a combination of methods for selection sweep detection allows different emerging patterns of selection to be identified, and it also improves the reliability and accuracy of the analyses (Pavlidis & Alachiotis, 2017). Potato breeding efforts currently center on improving different market classes such as chip and French fry processing, pigmented, table russets, and yellows (Hamilton et al., 2011). Most (but not all) hybridizations are made between clones within a market class. Over time one might expect these market classes to diverge, not only in terms of the few traits that define each class, but also in terms of unlinked, selectively neutral DNA markers (Hamilton et al., 2011). Several of the identified SNPs and sweep regions in this study are associated with functions of interest and warrant further investigation.

A SNP (PotVar0120627) was selected by PCAdapt at 48.6 Mb on chromosome 3 controlling the white-to-yellow flesh color. Sharma et al. (2018) using genome-wide association mapping found a strong association for flesh color at 49.4 Mb on the same chromosome. Similarly, SNPs belonging to the sweep region detected by this study on chromosomes 1 and 9 (Table 2.1) had been previously identified as significant SNPs for

flesh color (Sharma et al., 2018). The SNPs at the sweep region (10.58-11.08 Mb) on chromosome 7 were reported to have a significant association for the stolon attachment trait in potato (Sharma et al., 2018). In a previous QTL study, Manrique-Carpintero et al. (2018) identified candidate genes (*Dof Zinc Finger Protein-StCDF3, CONSTANS-CO*) in the photoperiod regulatory pathway associated with length of plant cycle and tuberization in the QTL region on chromosome 2 around 46 Mb. The XPEHH test from our study has also detected a sweep at the same location (46.15-46.65 Mb) on the same chromosome. Likewise, Manrique-Carpintero et al. (2018) had reported a *cytokinin riboside 5'-monophosphate phosphoribohydrolase LOG3* gene located in the QTL region at 56 Mb on chromosome 10. Differential expression and pleiotropic effects of *LOG* genes show their major role in cytokinin metabolism to modulate plant growth and development in *Arabidopsis* (Kuroha et al., 2009) and tomato (Mu et al., 2017). The sweep region at chromosome 10 (55.60-56.10 Mb) detected by both iHS and PCAdapt tests in our study matched with this finding. The *Sucrose transporter 4 (SUT4)* gene involving an accumulation of sucrose and starch in the terminal sink, is located at 65.8 Mb on chromosome 4 (Chincinska et al., 2013; Manrique-Carpintero et al., 2018). Our XPEHH test between the ChipRu and Red groups selected three SNPs (solcap_snp_c2_55781, solcap_snp_c2_55780 and solcap_snp_c2_55779) at 65.9 Mb on the same chromosome. It is well known that Red potatoes have less starch and more sugars than Chip and Russet potatoes (Stark et al., 2020). Parra-Galindo et al. (2019) reported the QTL AnthoX_Adeny, colocalized on chromosome 10 at 57.3 Mb (PGSC0003DMT400060833/*Adenylyl-sulfate kinase* gene), explaining 41.1% of the

phenotypic variance of pelargonidin. As illustrated in a colored potato study, the red cultivars contained predominantly pelargonidin derivatives, while the purple/blue varieties had peonidin, petunidin, and malvidin as the main aglycones (Oertel et al., 2017). These results will allow a better understanding of the genetic architecture and will open avenues for studying candidates for biochemical and functional studies of admixed advanced potato selections.

Most plant breeders want to make better use of plant genetic resources in their breeding programs but have trouble maintaining many clones and prioritizing clones for parental selection. Some breeders define a subset of clones that reflect the greater collection. The core subset can also be maintained as a backup collection to preserve important genes. In this study, we proposed a core set of 43 potato clones, accounting for 20% of the total collection using CoreHunter software. A sampling percentage of 20~30% was suggested by Hintum et al. (2000). Nevertheless, all core germplasm sets do not have a fixed size, as different crops and targets require different sampling percentages. In earlier studies, a core set of 48 was defined for capturing the genetic diversity of a collection of 350 tetraploid cultivated potato varieties using SSR data (Esnault et al., 2016). A core set of 27 genotypes was developed from 138 accessions of potato cultivars from the Western Highlands region of Cameroon using SSR markers (Anoumaa et al., 2017). Core Hunter software has been used for core set selection in earlier studies of different crops, including wheat (Ambati et al., 2020), Brazilian grapevine germplasm (Oliveira et al., 2020), banana (Bawin et al., 2019), sweetpotato (Su et al., 2017), and common bean (Mahajan et al., 2017). In many reports, genetic

diversity and cluster analysis were used to evaluate the efficiency of the development of the core germplasm set. In the present study, the genetic distance increased as expected after the removal of genetically similar clones during core germplasm set development. Having a core collection as a backup of important genes and source of parents is a good idea. However, in reality, core collections may not meet the needs of modern breeding approaches, such as genomic selection and genome-wide association studies where more individuals are desired to increase the statistical detection power.

2.5. Conclusions

Our analysis of the genetic diversity and population structure of the advanced clones in the TAMU Potato Breeding Program found a significant subdivision among clones, indicating a heterogeneous collection. Further, the SNP markers used in the study allowed the differentiation among breeding clones and the development of a core germplasm set of 43 clones, accounting for 20% of the total collection. Additionally, we used the SNP array to validate pedigree information. The genome-wide SNP characterization of these 214 clones, development of the core set, and reporting of the correct pedigree in this study will be useful for future genomic studies, parental selection, and germplasm management in potato breeding program.

2.6. References

Alsahlany, M., Zarka, D., Coombs, J., & Douches, D. S. (2019). Comparison of methods to distinguish diploid and tetraploid potato in applied diploid breeding. *American Journal of Potato Research, 96*(3), 244–254. https://doi.org/10.1007/s12230-018-09710-7

Ambati, D., Phuke, R. M., Vani, V., Sai Prasad, S. V., Singh, J. B., Patidar, C. P., Malviya, P., Gautam, A., & Dubey, V. G. (2020). Assessment of genetic diversity and development of core germplasm in durum wheat using agronomic and grain quality traits. *Cereal Research Communications*. https://doi.org/10.1007/s42976-020-00050-z

Anoumaa, M., Yao, N. K., Kouam, E. B., Kanmegne, G., Machuka, E., Osama, S. K., Nzuki, I., Kamga, Y. B., Fonkou, T., & Omokolo, D. N. (2017). Genetic diversity and core collection for potato (*Solanum tuberosum* L.) cultivars from Cameroon as revealed by SSR markers. *American Journal of Potato Research*, *94*(4), 449–463. https://doi.org/10.1007/s12230-017-9584-2

Bakker, E., Borm, T., Prins, P., van der Vossen, E., Uenk, G., Arens, M., de Boer, J., van Eck, H., Muskens, M., Vossen, J., van der Linden, G., van Ham, R., Klein-Lankhorst, R., Visser, R., Smant, G., Bakker, J., & Goverse, A. (2011). A genome-wide genetic map of NB-LRR disease resistance loci in potato. *Theoretical and Applied Genetics*, *123*(3), 493–508. https://doi.org/10.1007/s00122-011-1602-z

Bali, S., Patel, G., Novy, R., Vining, K., Brown, C., Holm, D., Porter, G., Endelman, J., Thompson, A., & Sathuvalli, V. (2018). Evaluation of genetic diversity among Russet potato clones and varieties from breeding programs across the United States. *PLOS ONE*, *13*(8), e0201415. https://doi.org/10.1371/journal.pone.0201415

Bawin, Y., Panis, B., Vanden Abeele, S., Li, Z., Sardos, J., Paofa, J., Ge, X.-J., Mertens, A., Honnay, O., & Janssens, S. B. (2019). Genetic diversity and core subset selection in *ex situ* seed collections of the banana crop wild relative *Musa balbisiana*. *Plant Genetic Resources: Characterization and Utilization*, *17*(6), 536–544. https://doi.org/10.1017/S1479262119000376

Beals, K. A. (2019). Potatoes, nutrition and health. *American Journal of Potato Research*, *96*(2), 102–110. https://doi.org/10.1007/s12230-018-09705-4

Bekele, W. A., Wight, C. P., Chao, S., Howarth, C. J., & Tinker, N. A. (2018). Haplotype-based genotyping-by-sequencing in oat genome research. *Plant Biotechnology Journal*, *16*(8), 1452–1463. https://doi.org/10.1111/pbi.12888

Berdugo-Cely, J., Valbuena, R. I., Sánchez-Betancourt, E., Barrero, L. S., & Yockteng, R. (2017). Genetic diversity and association mapping in the Colombian Central Collection of *Solanum tuberosum* L. Andigenum group using SNPs markers. *PLoS ONE*, *12*(3). https://doi.org/10.1371/journal.pone.0173039

Bertioli, D. J., Ozias-Akins, P., Chu, Y., Dantas, K. M., Santos, S. P., Gouvea, E., Guimarães, P. M., Leal-Bertioli, S. C. M., Knapp, S. J., & Moretzsohn, M. C. (2014). The use of SNP markers for linkage mapping in diploid and tetraploid

peanuts. *G3: Genes, Genomes, Genetics*, *4*(1), 89–96. https://doi.org/10.1534/g3.113.007617

Bhardwaj, A., Dhar, Y. V., Asif, M. H., & Bag, S. K. (2016). In Silico identification of SNP diversity in cultivated and wild tomato species: Insight from molecular simulations. *Scientific Reports*, *6*(1), 38715. https://doi.org/10.1038/srep38715

Bonierbale, M. W., Plaisted, R. L., & Tanksley, S. D. (1988). RFLP Maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics*, *120*(4), 1095–1103.

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, *32*(3), 314–331.

Bourke, P. M., Voorrips, R. E., Visser, R. G. F., & Maliepaard, C. (2018). Tools for genetic studies in experimental populations of polyploids. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.00513

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Bradshaw, J. E. (2007a). Chapter 8—Potato-breeding strategy. In D. Vreugdenhil, J. Bradshaw, C. Gebhardt, F. Govers, D. K. L. Mackerron, M. A. Taylor, & H. A. Ross (Eds.), *Potato Biology and Biotechnology* (pp. 157–177). Elsevier Science B.V. https://doi.org/10.1016/B978-044451018-1/50050-6

Bradshaw, J. E., & Ramsay, G. (2009). Chapter 1—Potato origin and production. In J. Singh & L. Kaur (Eds.), *Advances in Potato Chemistry and Technology* (pp. 1–26). Academic Press. https://doi.org/10.1016/B978-0-12-374349-7.00001-5

Bradshaw, John. E. (2007b). The canon of potato science: 4. Tetrasomic inheritance. *Potato Research*, *50*(3), 219–222. https://doi.org/10.1007/s11540-008-9041-1

Bretting, P. K., & Widrlechner, M. P. (2010). Genetic markers and plant genetic resource management. In J. Janick (Ed.), *Plant Breeding Reviews* (pp. 11–86). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470650059.ch2

Brown, A. H. D. (1989). Core collections: A practical approach to genetic resources management. *Genome*, *31*(2), 818–824. https://doi.org/10.1139/g89-144

Cadzow, M., Boocock, J., Nguyen, H. T., Wilcox, P., Merriman, T. R., & Black, M. A. (2014). A bioinformatics workflow for detecting signatures of selection in genomic data. *Frontiers in Genetics*, *5*. https://doi.org/10.3389/fgene.2014.00293

Campoy, J. A., Lerigoleur-Balsemin, E., Christmann, H., Beauvieux, R., Girollet, N., Quero-García, J., Dirlewanger, E., & Barreneche, T. (2016). Genetic diversity, linkage disequilibrium, population structure and construction of a core collection of Prunus avium L. landraces and bred cultivars. *BMC Plant Biology*, *16*(1), 49. https://doi.org/10.1186/s12870-016-0712-9

Chincinska, I., Gier, K., Krügel, U., Liesche, J., He, H., Grimm, B., Harren, F. J. M., Cristescu, S., & Kuehn, C. (2013). Photoperiodic regulation of the sucrose transporter StSUT4 affects the expression of circadian-regulated genes and

ethylene production. *Frontiers in Plant Science*, *4*. https://doi.org/10.3389/fpls.2013.00026

Ching, A., Caldwell, K. S., Jung, M., Dolan, M., Smith, O. S. (Howie), Tingey, S., Morgante, M., & Rafalski, A. J. (2002). SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genetics*, *3*, 19. https://doi.org/10.1186/1471-2156-3-19

da Silva, W. L., Ingram, J., Hackett, C. A., Coombs, J. J., Douches, D., Bryan, G. J., De Jong, W., & Gray, S. (2017). Mapping loci that control tuber and foliar symptoms caused by PVY in autotetraploid potato (*Solanum tuberosum*L.). *G3: Genes, Genomes, Genetics*, *7*(11), 3587–3595. https://doi.org/10.1534/g3.117.300264

De Beukelaer, H., Davenport, G. F., & Fack, V. (2018). Core Hunter 3: Flexible core subset selection. *BMC Bioinformatics*, *19*(1), 203. https://doi.org/10.1186/s12859-018-2209-z

Delaneau, O., Zagury, J.-F., & Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nature Methods*, *10*(1), 5–6. https://doi.org/10.1038/nmeth.2307

Devaux, A., Kromann, P., & Ortiz, O. (2014). Potatoes for sustainable global food security. *Potato Research*, *57*(3), 185–199. https://doi.org/10.1007/s11540-014-9265-1

Douches, D., Hirsch, C., Manrique-Carpintero, N., Massa, A., Coombs, J., Hardigan, M., Bisognin, D., De Jong, W., & Buell, C. (2014). The contribution of the

Solanaceae coordinated agricultural project to potato breeding. *Potato Research*, *57*, 215–224. https://doi.org/10.1007/s11540-014-9267-z

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Ellis, D., Chavez, O., Coombs, J., Soto, J., Gomez, R., Douches, D., Panta, A., Silvestre, R., & Anglin, N. L. (2018). Genetic identity in genebanks: Application of the SolCAP 12K SNP array in fingerprinting and diversity analysis in the global in trust potato collection. *Genome*, *61*(7), 523–537. https://doi.org/10.1139/gen-2017-0201

Endelman, J. B., Schmitz Carley, C. A., Douches, D. S., Coombs, J. J., Bizimungu, B., De Jong, W. S., Haynes, K. G., Holm, D. G., Miller, J. C., Novy, R. G., Palta, J. P., Parish, D. L., Porter, G. A., Sathuvalli, V. R., Thompson, A. L., & Yencho, G. C. (2017). Pedigree reconstruction with genome-wide markers in potato. *American Journal of Potato Research*, *94*(2), 184–190. https://doi.org/10.1007/s12230-016-9556-y

Esnault, F., Pellé, R., Dantec, J.-P., Bérard, A., Le Paslier, M.-C., & Chauvin, J.-E. (2016). Development of a potato cultivar (*Solanum tuberosum* L.) core collection, a valuable tool to prospect genetic variation for novel traits. *Potato Research*, *59*(4), 329–343. https://doi.org/10.1007/s11540-016-9332-x

Evans, J., McCormick, R. F., Morishige, D., Olson, S. N., Weers, B., Hilley, J., Klein, P., Rooney, W., & Mullet, J. (2013). Extensive variation in the density and distribution of DNA polymorphism in sorghum genomes. *PLOS ONE*, *8*(11), e79192. https://doi.org/10.1371/journal.pone.0079192

FAO. (2009). *Why potato? - International year of the potato 2008, Rome*. FAO. http://www.fao.org/potato-2008/en/aboutiyp/index.html

FAO. (2021). *Dimensions of need—Staple foods: What do people eat?, Rome*. FAO. http://www.fao.org/3/u8480e/u8480e07.htm

FAOSTAT. (2018). *Statistical data. Food and Agriculture Organization of the United Nations, Rome*. www.fao.org/faostat

FAOSTAT. (2021). *Statistical data. Food and Agriculture Organization of the United Nations, Rome*. www.fao.org/faostat

Felcher, K. J., Coombs, J. J., Massa, A. N., Hansey, C. N., Hamilton, J. P., Veilleux, R. E., Buell, C. R., & Douches, D. S. (2012). Integration of two diploid potato linkage maps with the potato genome sequence. *PLOS ONE*, *7*(4), e36347. https://doi.org/10.1371/journal.pone.0036347

Gautier, M., & Naves, M. (2011). Footprints of selection in the ancestral admixture of a New World Creole cattle breed. *Molecular Ecology*, *20*(15), 3128–3143. https://doi.org/10.1111/j.1365-294X.2011.05163.x

Gautier, M., & Vitalis, R. (2012). rehh: An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics (Oxford, England)*, *28*(8), 1176–1177. https://doi.org/10.1093/bioinformatics/bts115

Gebhardt, C. (2013). Bridging the gap between genome analysis and precision breeding in potato. *Trends in Genetics*, *29*(4), 248–256. https://doi.org/10.1016/j.tig.2012.11.006

Gebhardt, C. (2016). The historical role of species from the Solanaceae plant family in genetic research. *Theoretical and Applied Genetics*, *129*(12), 2281–2294. https://doi.org/10.1007/s00122-016-2804-1

Granato, I. S. C., Galli, G., de Oliveira Couto, E. G., e Souza, M. B., Mendonça, L. F., & Fritsche-Neto, R. (2018). snpReady: A tool to assist breeders in genomic analysis. *Molecular Breeding*, *38*(8), 102. https://doi.org/10.1007/s11032-018-0844-8

Grandke, F., Ranganathan, S., van Bers, N., de Haan, J. R., & Metzler, D. (2017). PERGOLA: Fast and deterministic linkage mapping of polyploids. *BMC Bioinformatics*, *18*(1), 12. https://doi.org/10.1186/s12859-016-1416-8

Guajardo, V., Solís, S., Almada, R., Saski, C., Gasic, K., & Moreno, M. Á. (2020). Genome-wide SNP identification in Prunus rootstocks germplasm collections using genotyping-by-sequencing: Phylogenetic analysis, distribution of SNPs and prediction of their effect on gene function. *Scientific Reports*, *10*(1), 1467. https://doi.org/10.1038/s41598-020-58271-5

Hackett, C. A., Boskamp, B., Vogogias, A., Preedy, K. F., & Milne, I. (2017). TetraploidSNPMap: Software for linkage analysis and QTL mapping in autotetraploid populations using SNP dosage data. *Journal of Heredity*, *108*(4), 438–442. https://doi.org/10.1093/jhered/esx022

Hale, A. L., Miller, J. C., Renganayaki, K., Fritz, A. K., Coombs, J. J., Frank, L. M., & Douches, D. S. (2005). Suitability of AFLP and microsatellite marker analysis for discriminating intraclonal variants of the potato cultivar Russet Norkotah. *Journal of the American Society for Horticultural Science*, *130*(4), 624–630. https://doi.org/10.21273/JASHS.130.4.624

Hamblin, M. T., Mitchell, S. E., White, G. M., Gallego, J., Kukatla, R., Wing, R. A., Paterson, A. H., & Kresovich, S. (2004). Comparative population genetics of the panicoid grasses: Sequence polymorphism, linkage disequilibrium and selection in a diverse sample of *Sorghum bicolor*. *Genetics*, *167*(1), 471–483. https://doi.org/10.1534/genetics.167.1.471

Hamilton, J. P., Hansey, C. N., Whitty, B. R., Stoffel, K., Massa, A. N., Van Deynze, A., De Jong, W. S., Douches, D. S., & Buell, C. R. (2011). Single nucleotide polymorphism discovery in elite North American potato germplasm. *BMC Genomics*, *12*(1), 302. https://doi.org/10.1186/1471-2164-12-302

Hardigan, M. A., Laimbeer, F. P. E., Newton, L., Crisovan, E., Hamilton, J. P., Vaillancourt, B., Wiegert-Rininger, K., Wood, J. C., Douches, D. S., Farré, E. M., Veilleux, R. E., & Buell, C. R. (2017). Genome diversity of tuber-bearing *Solanum* uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proceedings of the National Academy of Sciences*, *114*(46), E9999–E10008. https://doi.org/10.1073/pnas.1714380114

He, Q., Yu, J., Kim, T.-S., Cho, Y.-H., Lee, Y.-S., & Park, Y.-J. (2015). Resequencing reveals different domestication rate for *BADH1* and *BADH2* in rice (*Oryza*

sativa). *PLOS ONE*, *10*(8), e0134801. https://doi.org/10.1371/journal.pone.0134801

Hijmans, R. J., & Spooner, D. M. (2001). Geographic distribution of wild potato species. *American Journal of Botany*, *88*(11), 2101–2112. https://doi.org/10.2307/3558435

Hintum, T. J. L. van, Brown, A. H. D., Spillane, C., & Hodgkin, T. (2000). *Core collections of plant genetic resources*. International Plant Genetic Resources Institute.

Hirsch, C. N., Hirsch, C. D., Felcher, K., Coombs, J., Zarka, D., Deynze, A. V., Jong, W. D., Veilleux, R. E., Jansky, S., Bethke, P., Douches, D. S., & Buell, C. R. (2013). Retrospective view of North American potato (*Solanum tuberosum* L.) breeding in the 20th and 21st centuries. *G3: Genes, Genomes, Genetics*, *3*(6), 1003–1013. https://doi.org/10.1534/g3.113.005595

Igarashi, T., Tsuyama, M., Ogawa, K., Koizumi, E., Sanetomo, R., & Hosaka, K. (2019). Evaluation of Japanese potatoes using single nucleotide polymorphisms (SNPs). *Molecular Breeding*, *39*(1). https://doi.org/10.1007/s11032-018-0917-8

Jamali, S. H., Cockram, J., & Hickey, L. T. (2019). Insights into deployment of DNA markers in plant variety protection and registration. *Theoretical and Applied Genetics*, *132*(7), 1911–1929. https://doi.org/10.1007/s00122-019-03348-7

Jansky, S. (2009). Chapter 2—Breeding, genetics, and cultivar development. In J. Singh & L. Kaur (Eds.), *Advances in Potato Chemistry and Technology* (pp. 27–62). Academic Press. https://doi.org/10.1016/B978-0-12-374349-7.00002-7

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jupe, F., Witek, K., Verweij, W., Śliwka, J., Pritchard, L., Etherington, G. J., Maclean, D., Cock, P. J., Leggett, R. M., Bryan, G. J., Cardle, L., Hein, I., & Jones, J. D. G. (2013). Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *The Plant Journal*, *76*(3), 530–544. https://doi.org/10.1111/tpj.12307

Kanter, M., & Elkin, C. (2019). Potato as a source of nutrition for physical performance. *American Journal of Potato Research*, *96*(2), 201–205. https://doi.org/10.1007/s12230-018-09701-8

Kao, T. H., & McCubbin, A. G. (1996). How flowering plants discriminate between self and non-self pollen to prevent inbreeding. *Proceedings of the National Academy of Sciences*, *93*(22), 12059–12065. https://doi.org/10.1073/pnas.93.22.12059

Keijbets, M. J. H. (2008). Potato processing for the consumer: Developments and future challenges. *Potato Research*, *51*(3–4), 271–281. https://doi.org/10.1007/s11540-008-9104-3

Kirch, H. H., Uhrig, H., Lottspeich, F., Salamini, F., & Thompson, R. D. (1989). Characterization of proteins associated with self-incompatibility in *Solanum tuberosum*. *Theoretical and Applied Genetics*, *78*(4), 581–588. https://doi.org/10.1007/BF00290845

Kloosterman, B., Oortwijn, M., uitdeWilligen, J., America, T., de Vos, R., Visser, R. G., & Bachem, C. W. (2010). From QTL to candidate gene: Genetical genomics of simple and complex traits in potato using a pooling strategy. *BMC Genomics*, *11*(1), 158. https://doi.org/10.1186/1471-2164-11-158

Kolech, S. A., Halseth, D., Perry, K., Wolfe, D., Douches, D. S., Coombs, J., & De Jong, W. (2016). Genetic diversity and relationship of Ethiopian potato varieties to germplasm from North America, Europe and the International Potato Center. *American Journal of Potato Research*, *93*(6), 609–619. https://doi.org/10.1007/s12230-016-9543-3

Konieczny, A., & Ausubel, F. M. (1993). A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *The Plant Journal*, *4*(2), 403–410. https://doi.org/10.1046/j.1365-313X.1993.04020403.x

Kuroha, T., Tokunaga, H., Kojima, M., Ueda, N., Ishida, T., Nagawa, S., Fukuda, H., Sugimoto, K., & Sakakibara, H. (2009). Functional analyses of *LONELY GUY* cytokinin-activating enzymes reveal the importance of the direct activation pathway in Arabidopsis. *The Plant Cell*, *21*(10), 3152–3169. https://doi.org/10.1105/tpc.109.068676

Li, J., Li, D., Espinosa, C. Z., Pastor, V. T., Rasheed, A., Rojas, N. P., Wang, J., Varela, A. S., Carolina de Almeida Silva, N., Schnable, P. S., Costich, D. E., & Li, H. (2021). Genome-wide analyses reveal footprints of divergent selection and popping-related traits in CIMMYT's maize inbred lines. *Journal of Experimental Botany*, *72*(4), 1307–1320. https://doi.org/10.1093/jxb/eraa480

Lindqvist-Kreuze, H., Gastelo, M., Perez, W., Forbes, G. A., de Koeyer, D., & Bonierbale, M. (2014). Phenotypic stability and genome-wide association study of late blight resistance in potato genotypes adapted to the tropical highlands. *Phytopathology*, *104*(6), 624–633. https://doi.org/10.1094/PHYTO-10-13-0270-R

Liyanage, D. W. K., Yevtushenko, D. P., Konschuh, M., Bizimungu, B., & Lu, Z.-X. (2021). Processing strategies to decrease acrylamide formation, reducing sugars and free asparagine content in potato chips from three commercial cultivars. *Food Control*, *119*, 107452. https://doi.org/10.1016/j.foodcont.2020.107452

Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, *17*(1), 67–77. https://doi.org/10.1111/1755-0998.12592

Mahajan, R., Zargar, S. M., Singh, R., Salgotra, R. K., Farhat, S., & Sonah, H. (2017). Population structure analysis and selection of core set among common bean genotypes from Jammu and Kashmir, India. *Applied Biochemistry and Biotechnology*, *182*(1), 16–28. https://doi.org/10.1007/s12010-016-2307-1

Manrique-Carpintero, Norma C., Coombs, J. J., Pham, G. M., Laimbeer, F. P. E., Braz, G. T., Jiang, J., Veilleux, R. E., Buell, C. R., & Douches, D. S. (2018). Genome reduction in tetraploid potato reveals genetic load, haplotype variation, and loci associated with agronomic traits. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.00944

Manrique-Carpintero, Norma Constanza, Tokuhisa, J. G., Ginzberg, I., & Veilleux, R. E. (2014). Allelic variation in genes contributing to glycoalkaloid biosynthesis in a diploid interspecific population of potato. *Theoretical and Applied Genetics*, *127*(2), 391–405. https://doi.org/10.1007/s00122-013-2226-2

Massa, A. N., Manrique-Carpintero, N. C., Coombs, J., Haynes, K. G., Bethke, P. C., Brandt, T. L., Gupta, S. K., Yencho, G. C., Novy, R. G., & Douches, D. S. (2018). Linkage analysis and QTL mapping in a tetraploid russet mapping population of potato. *BMC Genetics*, *19*(1). https://doi.org/10.1186/s12863-018-0672-1

Massa, A. N., Manrique-Carpintero, N. C., Coombs, J. J., Zarka, D. G., Boone, A. E., Kirk, W. W., Hackett, C. A., Bryan, G. J., & Douches, D. S. (2015). Genetic Linkage mapping of economically important traits in cultivated tetraploid potato (*Solanum tuberosum* L.). *G3: Genes, Genomes, Genetics*, *5*(11), 2357–2364. https://doi.org/10.1534/g3.115.019646

McGill, C. R., Kurilich, A. C., & Davignon, J. (2013). The role of potatoes and potato components in cardiometabolic health: A review. *Annals of Medicine*, *45*(7), 467–473. https://doi.org/10.3109/07853890.2013.813633

Miller, J. C., Scheuring, D. C., Miller, J. P., & Fernandez, G. C. J. (1999). Selection, evaluation, and identification of improved Russet Norkotah strains. *American Journal of Potato Research*, *76*(3), 161–167. https://doi.org/10.1007/BF02853581

Miller, J. C., Tai, G. C. C., Ouellette, B., & Miller, J. P. (2004). Discriminating Russet Norkotah intraclonal selections using canonical and cluster analysis. *American Journal of Potato Research*, *81*(3), 203–207. https://doi.org/10.1007/BF02871750

Mollinari, M., & Garcia, A. A. F. (2019). Linkage analysis and haplotype phasing in experimental autopolyploid populations with high ploidy level using hidden Markov models. *G3: Genes, Genomes, Genetics*, *9*(10), 3297–3314. https://doi.org/10.1534/g3.119.400378

Mu, Q., Huang, Z., Chakrabarti, M., Illa-Berenguer, E., Liu, X., Wang, Y., Ramos, A., & Knaap, E. van der. (2017). Fruit weight is controlled by cell size regulator encoding a novel protein that is expressed in maturing tomato fruits. *PLOS Genetics*, *13*(8), e1006930. https://doi.org/10.1371/journal.pgen.1006930

Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, *106*(949), 283–292. https://doi.org/10.1086/282771

NPC National Potato Council. (2019). *Potato statistical yearbook 2019, Washington, D.C.* https://www.nationalpotatocouncil.org/files/5015/6380/8213/2019_Stat_Book__fnl.pdf

Odong, T. L., Jansen, J., van Eeuwijk, F. A., & van Hintum, T. J. L. (2013). Quality of core collections for effective utilisation of genetic resources review, discussion and interpretation. *Theoretical and Applied Genetics*, *126*(2), 289–305. https://doi.org/10.1007/s00122-012-1971-y

Oertel, A., Matros, A., Hartmann, A., Arapitsas, P., Dehmer, K. J., Martens, S., & Mock, H.-P. (2017). Metabolite profiling of red and blue potatoes revealed cultivar and tissue specific patterns for anthocyanins and other polyphenols. *Planta*, *246*(2), 281–297. https://doi.org/10.1007/s00425-017-2718-4

Oliveira, G. L. de, Souza, A. P. de, Oliveira, F. A. de, Zucchi, M. I., Souza, L. M. de, & Moura, M. F. (2020). Genetic structure and molecular diversity of Brazilian grapevine germplasm: Management and use in breeding programs. *BioRxiv*, 2020.05.05.078865. https://doi.org/10.1101/2020.05.05.078865

Ortiz, O., & Mares, V. (2017). The historical, social, and economic importance of the potato crop. In S. Kumar Chakrabarti, C. Xie, & J. Kumar Tiwari (Eds.), *The Potato Genome* (pp. 1–10). Springer International Publishing. https://doi.org/10.1007/978-3-319-66135-3_1

Otyama, P. I., Wilkey, A., Kulkarni, R., Assefa, T., Chu, Y., Clevenger, J., O'Connor, D. J., Wright, G. C., Dezern, S. W., MacDonald, G. E., Anglin, N. L., Cannon, E. K. S., Ozias-Akins, P., & Cannon, S. B. (2019). Evaluation of linkage disequilibrium, population structure, and genetic diversity in the U.S. peanut mini core collection. *BMC Genomics*, *20*(1), 481. https://doi.org/10.1186/s12864-019-5824-9

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20*(2), 289–290. https://doi.org/10.1093/bioinformatics/btg412

Parra-Galindo, M.-A., Piñeros-Niño, C., Soto-Sedano, J. C., & Mosquera-Vasquez, T. (2019). Chromosomes I and X harbor consistent genetic factors associated with the anthocyanin variation in potato. *Agronomy*, *9*(7), 366. https://doi.org/10.3390/agronomy9070366

Pavlidis, P., & Alachiotis, N. (2017). A survey of methods and tools to detect recent and strong positive selection. *Journal of Biological Research*, *24*. https://doi.org/10.1186/s40709-017-0064-0

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, *13*(5), 946–952. https://doi.org/10.1111/1755-0998.12129

Pham, G. M., Hamilton, J. P., Wood, J. C., Burke, J. T., Zhao, H., Vaillancourt, B., Ou, S., Jiang, J., & Buell, C. R. (2020). Construction of a chromosome-scale long-read reference genome assembly for potato. *GigaScience*, *9*(giaa100). https://doi.org/10.1093/gigascience/giaa100

Pont, C., Leroy, T., Seidel, M., Tondelli, A., Duchemin, W., Armisen, D., Lang, D., Bustos-Korts, D., Goué, N., Balfourier, F., Molnár-Láng, M., Lage, J., Kilian, B., Özkan, H., Waite, D., Dyer, S., Letellier, T., Alaux, M., Russell, J., … Salse, J. (2019). Tracing the ancestry of modern bread wheats. *Nature Genetics*, *51*(5), 905–911. https://doi.org/10.1038/s41588-019-0393-z

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Privé, F., Luu, K., Vilhjálmsson, B. J., & Blum, M. G. B. (2020). Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. *Molecular Biology and Evolution*, *37*(7), 2153–2154. https://doi.org/10.1093/molbev/msaa053

Rodríguez, F., Ghislain, M., Clausen, A. M., Jansky, S. H., & Spooner, D. M. (2010). Hybrid origins of cultivated potatoes. *Theoretical and Applied Genetics*, *121*(6), 1187–1198. https://doi.org/10.1007/s00122-010-1422-6

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E. H., McCarroll, S. A., Gaudet, R., Schaffner, S. F., & Lander, E. S. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature*, *449*(7164), 913–918. https://doi.org/10.1038/nature06250

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, *74*(12), 5463–5467. https://doi.org/10.1073/pnas.74.12.5463

Sarkar, D. (2008). The signal transduction pathways controlling in planta tuberization in potato: An emerging synthesis. *Plant Cell Reports*, *27*(1), 1–8. https://doi.org/10.1007/s00299-007-0457-x

Sauvage, C., Rau, A., Aichholz, C., Chadoeuf, J., Sarah, G., Ruiz, M., Santoni, S., Causse, M., David, J., & Glémin, S. (2017). Domestication rewired gene expression and nucleotide diversity patterns in tomato. *The Plant Journal*, *91*(4), 631–645. https://doi.org/10.1111/tpj.13592

Schumann, M. J., Zeng, Z.-B., Clough, M. E., & Yencho, G. C. (2017). Linkage map construction and QTL analysis for internal heat necrosis in autotetraploid potato. *Theoretical and Applied Genetics*, *130*(10), 2045–2056. https://doi.org/10.1007/s00122-017-2941-1

Sharma, S. K., Bolser, D., Boer, J. de, Sønderkær, M., Amoros, W., Carboni, M. F., D'Ambrosio, J. M., Cruz, G. de la, Genova, A. D., Douches, D. S., Eguiluz, M., Guo, X., Guzman, F., Hackett, C. A., Hamilton, J. P., Li, G., Li, Y., Lozano, R., Maass, A., … Bryan, G. J. (2013). Construction of reference chromosome-scale pseudomolecules for potato: Integrating the potato genome with genetic and physical maps. *G3: Genes, Genomes, Genetics*, *3*(11), 2031–2047. https://doi.org/10.1534/g3.113.007153

Sharma, S. K., MacKenzie, K., McLean, K., Dale, F., Daniels, S., & Bryan, G. J. (2018). Linkage disequilibrium and evaluation of genome-wide association mapping models in tetraploid potato. *G3: Genes*, *Genomes*, *Genetics*, *8*(10), 3185–3202. https://doi.org/10.1534/g3.118.200377

Slater, A. T., Cogan, N. O. I., Hayes, B. J., Schultz, L., Dale, M. F. B., Bryan, G. J., & Forster, J. W. (2014). Improving breeding efficiency in potato using molecular and quantitative genetics. *Theoretical and Applied Genetics*, *127*(11), 2279–2292. https://doi.org/10.1007/s00122-014-2386-8

Spooner, D. M. (2009). DNA barcoding will frequently fail in complicated groups: An example in wild potatoes. *American Journal of Botany*, *96*(6), 1177–1189. https://doi.org/10.3732/ajb.0800246

Stark, J. C., Love, S. L., & Knowles, N. R. (2020). Tuber quality. In J. C. Stark, M. Thornton, & P. Nolte (Eds.), *Potato Production Systems* (pp. 479–497). Springer International Publishing. https://doi.org/10.1007/978-3-030-39157-7_15

Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences*, *100*(16), 9440–9445. https://doi.org/10.1073/pnas.1530509100

Stupar, R. M., Song, J., Tek, A. L., Cheng, Z., Dong, F., & Jiang, J. (2002). Highly condensed potato pericentromeric heterochromatin contains rDNA-related tandem repeats. *Genetics*, *162*(3), 1435–1444.

Su, W., Wang, L., Lei, J., Chai, S., Liu, Y., Yang, Y., Yang, X., & Jiao, C. (2017). Genome-wide assessment of population structure and genetic diversity and development of a core germplasm set for sweet potato based on specific length amplified fragment (SLAF) sequencing. *PLOS ONE*, *12*(2), e0172066. https://doi.org/10.1371/journal.pone.0172066

Tajima, F. (1989). DNA Polymorphism in a subdivided population: The expected number of segregating sites in the two-subpopulation model. *Genetics*, *123*(1), 229–240.

US Department of Agriculture National Agricultural Statistics Service. (2019). *Potatoes 2018 summary, Washington, D.C.* https://www.nass.usda.gov/Publications/Todays_Reports/reports/pots0919.pdf

Vitti, J. J., Grossman, S. R., & Sabeti, P. C. (2013). Detecting natural selection in genomic data. *Annual Review of Genetics*, *47*, 97–120. https://doi.org/10.1146/annurev-genet-111212-133526

Voight, B. F., Kudaravalli, S., Wen, X., & Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLOS Biology*, *4*(3), e72. https://doi.org/10.1371/journal.pbio.0040072

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., & Kuiper, M. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*, *23*(21), 4407–4414.

Vos, Peter. G., Uitdewilligen, J. G. A. M. L., Voorrips, R. E., Visser, R. G. F., & van Eck, H. J. (2015). Development and analysis of a 20K SNP array for potato (*Solanum tuberosum*): An insight into the breeding history. *Theoretical and Applied Genetics*, *128*(12), 2387–2401. https://doi.org/10.1007/s00122-015-2593-y

Watanabe, K. (2015). Potato genetics, genomics, and applications. *Breeding Science*, *65*(1), 53–68. https://doi.org/10.1270/jsbbs.65.53

Welsh, J., & McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research*, *18*(24), 7213–7218.

Wijesinha-Bettoni, R., & Mouillé, B. (2019). The contribution of potatoes to global food security, nutrition and healthy diets. *American Journal of Potato Research*, *96*(2), 139–149. https://doi.org/10.1007/s12230-018-09697-1

Xu, J., Chen, G., Hermanson, P. J., Xu, Q., Sun, C., Chen, W., Kan, Q., Li, M., Crisp, P. A., Yan, J., Li, L., Springer, N. M., & Li, Q. (2019). Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize. *Genome Biology*, *20*(1), 243. https://doi.org/10.1186/s13059-019-1859-0

Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., Wang, J., Orjeda, G., Guzman, F., Torres, M., Lozano, R., Ponce, O., Martinez, D., De la Cruz, G., Chakrabarti, S. K., Patil, V. U. (2011). Genome sequence and analysis of the tuber crop potato. *Nature*, *475*(7355), 189–195. https://doi.org/10.1038/nature10158

Zhu, Y. L., Song, Q. J., Hyten, D. L., Tassell, C. P. V., Matukumalli, L. K., Grimm, D. R., Hyatt, S. M., Fickus, E. W., Young, N. D., & Cregan, P. B. (2003). Single nucleotide polymorphisms in soybean. *Genetics*, *163*(3), 1123–1134.

3. IDENTIFICATION OF GENOMIC REGIONS AND SUPERIOR INDIVIDUALS FOR TUBER MORPHOLOGY TRAITS IN TETRAPLOID POTATO BREEDING CLONES

3.1. Introduction

Domestication and selective breeding have resulted in significant morphological variation in organ shape in cultivated plant species (Sun et al., 2017). Examples of studies on phenotypic diversity and their genetic basis include the wheat grain size and shape (Gegas et al., 2010), tomato and melon fruit shape (Monforte et al., 2014), and morphotype diversification in *Brassica rapa* and *Brassica oleracea* (Cheng et al., 2016). Potato domestication took place around 8,000 years ago from a wild diploid *Solanum* species in an area located on the border of present-day Peru and Bolivia (Spooner et al., 2005). Modern cultivars are the result of extensive crossbreeding between cultivar groups as well as wild species. Andean tetraploid varieties likely resulted from repeated sexual polyploidization of early landrace diploids (Spooner et al., 2014). The domestication of the potato involved selection of shorter stolons, larger tubers, diverse tuber shape, and reduction of bitter tuber glycoalkaloids (Spooner et al., 2005).

Currently, potato breeding activities in the United States are concentrated on developing six distinct market groups (chip processing, French fry processing, pigmented, table russet, round white table, and yellow) (Hamilton et al., 2011). The market-class-specific morphological traits include tuber shape, eye depth, degree of russeting, tuber number, tuber weight, skin color, and flesh color. The tuber shape is extremely important and is based on historical regional preferences coupled with local

culinary practices (Stark et al., 2020). Chips are made with round tubers, whereas French fries are made with long tubers. Likewise, the consumer expects table cultivars to be round or oval (van Eck, Jacobs, van den Berg, et al., 1994; Chen et al., 2018). The irregular shape and deep eyes in potatoes lead to higher costs due to significant peeling losses (Li et al., 2005; ŚLiwka et al., 2008). Another trait, skin russeting is an inherited trait in some potato cultivars such as Russet Burbank and Russet Norkotah (De Jong, 1981), but the development of a russeted skin in red-skinned cultivars like 'Magen russeting' is a physiological disorder (Lulai, 2007; Ginzberg et al., 2012). Tuber sizes, skin color, and flesh color traits are also preferred by consumers and industry depending on the use of potatoes. Thus, breeding for morphological traits in potatoes is crucial as these traits affect consumption.

The cultivated potato has a complex genetic structure. It is a vegetatively propagated, highly heterozygous, autotetraploid (2n = 4x = 48) crop with a basic chromosome number of 12 and genome size of 844 Mbp (Xu et al., 2011). Due to high heterozygosity, broad segregation for many traits, not just target traits, will occur when a breeder makes a cross (Bonierbale et al., 2020). Some of these traits are poorly characterized or the heritability of traits is low (Thiele et al., 2021). Due to this complexity in autopolyploids, most molecular methods and techniques have been limited to diploids (Bourke, Voorrips, et al., 2018).

Utilizing bi-parental populations in diploid potatoes, quantitative trait loci (QTL) mapping studies have been performed to map several quality traits. The findings of tuber shape inheritance, whether it is monogenic or polygenic have been inconsistent. Van Eck

et al. (1994a), using restriction fragment length polymorphisms (RFLPs), identified a single locus, the *Ro*, that explains the inheritance of qualitative tuber form, with round being dominant over long on chromosome 10. Other reports using populations with different genetic backgrounds mention QTLs on chromosomes 2, 5, and 11 (Bradshaw et al., 2008), and 2 and 11 (ŚLiwka et al., 2008). Eye depth was linked to the *Ro* locus on chromosome 10 (Li et al., 2005; ŚLiwka et al., 2008). Similarly, using QTL analysis, a dominant allele at the *Y* (yellow) locus on potato chromosome 3 is responsible for the yellow flesh color (Bonierbale et al., 1988). A beta-carotene hydroxylase gene (*BCH*) was found in the same position as the *Y* locus, suggesting that this is the most likely candidate gene for yellow flesh (Thorup et al., 2000; Kloosterman et al., 2010). Anthocyanin pigments accumulate in the tuber flesh, giving it a red or purple color (Eichhorn & Winterhalter, 2005). The *Pf* locus tightly linked with the *I* locus (encode a MYB transcription factor) maps to chromosome 10 and confers pigmented tuber flesh (De Jong, 1987). Also, the *R* locus which encodes dihydroflavonol 4-reductase (*dfr*) maps to chromosome 2 (De Jong et al., 2003) and it is essential for red anthocyanin production. For purple pigment formation in the potato skin, the *P* locus which encodes flavonoid 3′,5′-hydroxylase (*f3′5′h*) is required and it maps to chromosome 11 (Jung et al., 2005). The *f3′5′h* gene co-segregates with *P,* and it is expressed in the tuber skin only in the presence of *I* (Jung et al., 2005).

Recent molecular technologies and analytical platforms hold the promise of unlocking previously unattainable levels of understanding of polyploid genomes. TetraploidSNPMap (Hackett et al., 2017), allows extensive use of allelic dosage data to

carry out linkage analysis and QTL mapping in autotetraploid species. Other polyploid mapping software recently in use are the PERGOLA package in R (Grandke et al., 2017), polymapR (Bourke et al., 2018), and MAPpoly (Mollinari & Garcia, 2019). These advancements have aided in the identification of QTL that account for significant amounts of phenotypic variance within a polyploid population. However, the number, effect, phenotypic variance explained, and resolution of individual QTL in a bi-parental population often obstruct causal gene identification (Zhu et al., 2008).

Another powerful strategy to discover markers linked to complex traits is to perform a Genome-Wide Association Study (GWAS) (Cortes et al., 2021). GWAS is an effective method for circumventing many of the limitations of bi-parental linkage mapping. It takes advantage of genetic variation that exists among natural or developed population utilizing historical recombination and knowledge of population structure (Ortiz, 2020). GWAS pinpoints QTLs by analyzing the marker-trait associations that can be attributed to the strength of linkage disequilibrium (LD) between markers and functional polymorphisms across a set of diverse germplasm (Zhu et al., 2008).

One challenge in applying GWAS to polyploid species is how to define relatedness between polyploid individuals (i.e., how to generate the kinship matrix, K). GWAS can now account for the kinship matrix (K) in potatoes using GWASpoly (Rosyara et al., 2016). GWASpoly also considers the allele dosage (AAAA, AAAa, AAaa, Aaaa, aaaa). For computational efficiency, Yang et al. (2014) proposed the leave-one-chromosome-out (LOCO) method. In LOCO method a different covariance matrix

is calculated for each chromosome based on the markers from all other chromosomes (Yang et al., 2014) and is now implemented in version 2 of GWASPoly.

Sharma et al. (2018) examined various GWAS models in cultivated potato genotypes using the Infinium 8K Potato SNP Array and found that kinship, not population structure, is the most important factor in determining the extent of false associations. Klaassen et al. (2019) evaluated a panel of 277 varieties using SolSTW 20K Infinium SNP marker array (Vos et al., 2015) and revealed four QTLs for protein content in tetraploid potato. Similarly, Zia et al. (2020) performed a GWAS using SolCAP 12K array for various morpho-agronomic traits in a panel of 237 tetraploid potato genotypes. Kaiser et al. (2020) used GWAS to identify genetic features associated with common scab resistance in the tetraploid population using the Illumina Infinium 8303 Potato Array. Recently, Yousaf et al. (2021) reported SNPs associated with potato stolon traits and root traits.

Traditional breeding strategies in potato like progeny testing and phenotypic recurrent selection are made more efficient by genomic estimated breeding values (GEBVs) particularly for low heritability traits. Sood et al. (2020) implemented best linear unbiased prediction (BLUP) for the rapid identification of superior individuals for tuber yield and late blight resistance in potato. Slater et al. (2014) reported that BLUP increased genetic gains for low heritability traits in autotetraploid potato.

The current study hypothesized that since the Texas A&M University Potato Breeding Program's collection of 214 advanced clones selected over 40 years and maintained *in vitro* (Pandey et al., 2021) has variation in tuber morphology traits, it can

be used to identify genomic regions associated with tuber traits and find superior individuals to use as parents. Recent marker-dense platforms including the 20K Infinium SNP marker array (Vos et al., 2015) and the Illumina 22K V3 Potato Array are expected to provide more precise correction for different levels of relatedness in GWAS models, as well as help in the identification of genomic regions associated with traits. In this study, we aimed to identify QTLs and genes and superior individuals associated with tuber shape, eye depth, degree of russeting, tuber number, tuber weight, skin color, and flesh color using the Illumina Infinium 22 K V3 Potato Array.

3.2. Materials and methods

3.2.1. Plant materials

The association panel (N = 214) consisted of tetraploid (2n = 4x =48) advanced clones and varieties. The collection comprised 31 chipping, 62 russet, 32 yellow-skinned, 68 red-skinned, and 21 purple-skinned clones. Analysis of population structure and discriminant analysis of principal components in the panel displayed three sub-populations, as reported earlier (Pandey et al., 2021).

3.2.2. Genotyping

Genomic DNA was extracted from 50 to 80 mg of fresh young potato leaves from tissue culture plantlets using the DNeasy Plant Pro Kit (Qiagen, Valencia, CA, USA). Samples were assayed using the Infinium 22 K V3 Potato Array on the Illumina iScan (Illumina Inc., San Diego, CA, USA) at Michigan State University. The marker dataset was filtered for polymorphism and minor allele frequency as described in Pandey et al. (2021).

3.2.3. Field experiment and phenotyping

Three field experiments were conducted in Texas to derive phenotypic data at commercial potato grower fields. In 2019, traits were evaluated at Dalhart (35°58'15.31"N, 102°44'36.33"W) and in 2020 traits were evaluated in Dalhart and Springlake (34°6'43.43"N, 102°19'40.02"W). In both all environments, entries were planted in a 12-hill plot with 2 replications. Seed tubers were planted with 30 cm spacing between hills and 70 cm spacing between rows. In Springlake, trials were planted in late March to early April and harvested in early July, whereas in Dalhart, trials were planted in early May and harvested in early September, with vine desiccation 2–3 weeks before harvest. Standard potato production practices were followed during the growing period in all years of the Texas A&M potato breeding program (https://potato.tamu.edu/reports/). Reference varieties for various market classes included in this study were: Russet Norkotah (standard, fresh market russet), Atlantic (chipper), Russet Burbank (processing russet), White LaSoda (a white-skinned mutant of Red LaSoda selected by the TAMU Program, fresh market white flesh), and Yukon Gold Strain (TXYG79, fresh market yellow flesh). The environmental conditions during the potato growing season (Figure 3.1) were representative of the subregion (High Plains) which is characterized by cold semi-arid climate with a range of temperatures. Rainfall was supplemented with with center pivot irrigation for a total of around 475 mm.

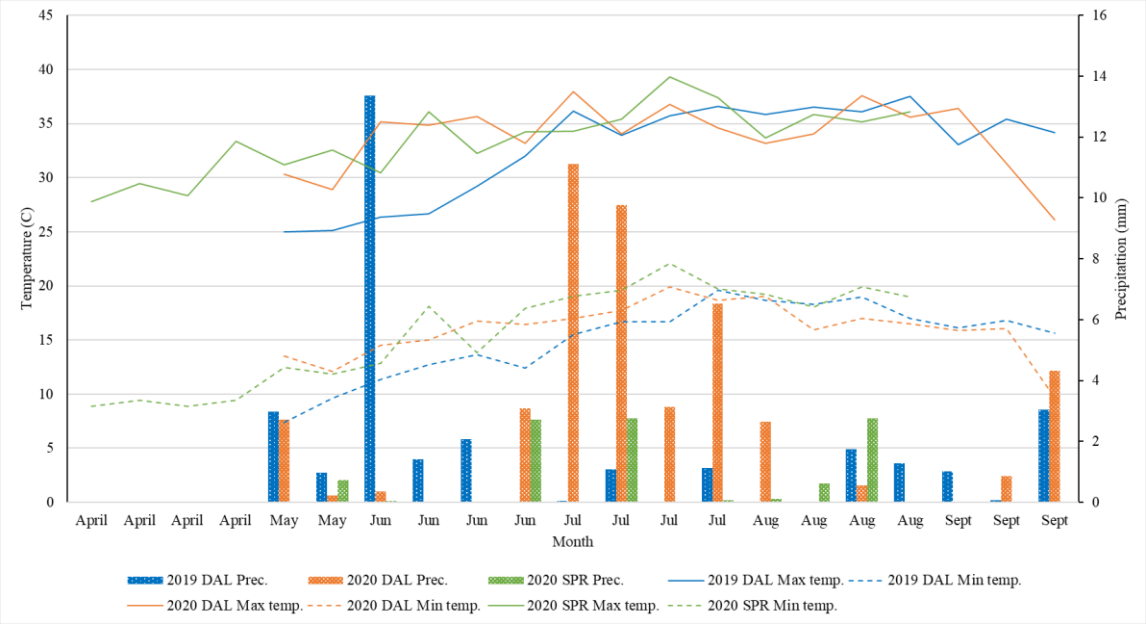


Figure 3.1 Precipitation, maximum and minimum temperatures during the experiment in 2019 and 2020 at Dalhart and Springlake, TX. Supplemental water was provided to the crop via center pivot irrigation.

For the traits assessed visually, the scores were given by replication and were representative of all tubers harvested. Tuber shape was determined using a 1 to 5 scale, where 1 = round, to 5 = long. Tuber shape was also measured as the ratio of tuber length to width (roundness score) and tuber length to thickness (flatness score). Length (mm) of the tuber was the distance from the stolon end to the bud end of the tuber, width (mm) was measured as the highest dimension at the equatorial area of the tuber, and thickness (mm) was measured perpendicularly to width. For each clone, length, width, and thickness data were collected from five tubers (113.4 – 170.1 g grading group) per replication. These numbers were used to calculate the length-width ratio and width-thickness ratio. Eye depth was evaluated on a 1–5 scale, where 1 = eyes deeper than 5 mm, to 5 = eyes not perceivable by touch. The degree of russeting was evaluated on a 1–

5 scale, where 1=none to 5=heavy russet. Mean tuber number per plant was calculated as a total number of tubers harvested per plot per replication divided by plant stand at 60 DAP (days after planting), mean tuber weight in g was calculated by dividing the total tuber yield of the plot divided by the number of tubers.

Tuber skin and flesh color intensity were evaluated visually on 10 tubers per replication according to a 1–5 scale, where 1=light to 5=dark. Objective color measurements of the skin and flesh color were obtained from 3 tubers per plot per replication with a colorimeter, Konica–Minolta Chroma Meter (CR-400 with 8-mm aperture and 0° viewing angle, Konica-Minolta, Inc., Tokyo, Japan) (Konica-Minolta 2013), using the CIE 1976 L*a*b* color spaces (HunterLab 2008). The instrument was calibrated against a standard white reference tile provided by the instrument manufacturer. CIE 1976 L*a*b* is a three-dimensional color space where L* (lightness) represents the white to black axis, a* represents the red to green axis, and b* represents the yellow to blue axis. The colorimeter collected and averaged three readings per tuber, and each reading measured 50.3 mm^2 (8-mm diameter area) of the surface. Data were recorded using Color Data Software CM-S100w SpectraMagic NX (Version 2.8) (Konica-Minolta 2014) and later transferred to Microsoft® Excel® (Microsoft Office Standard 2013, Microsoft Corp., Redmond, WA) spreadsheets. Also, the color was expressed in LCH color space, in which L* = lightness, C is chroma (saturation) and H is hue angle. Purple skin tuber pigmentation was scored as a binary trait (Purple/Non-purple).

3.2.4. Statistical analyses

Phenotypic data of tuber morphology traits were analyzed separately for each of the three environments (location-year) and also combined across environments using software package META-R (Alvarado et al., 2020). Summary statistics (mean, SE, range, LSD, CV) were also generated using standard procedures implemented in META-R. The phenotypic correlations between pairs of traits were calculated as simple Pearson correlations. Broad-sense heritability of a given trait in an individual environment was calculated on a genotype mean basis as

$$H^2 = \frac{{\sigma^2}_G}{{\sigma^2}_G + \frac{{\sigma^2}_e}{r}}$$

where ${\sigma^2}_G$ is the genotypic variance, ${\sigma^2}_e$ is the error variance, and r is the number of replications.

Broad sense heritability on a genotype mean basis for the combination of all environments was calculated as:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + (\sigma_{GE}^2/n) + (\sigma_e^2/rn)}$$

where σ_G^2 is the genotypic variance, σ_{GE}^2 is the variance of the interaction of genotype and environment, σ_e^2 is the variance of the experimental error, r is the number of replications, and n is the number of environments.

Best linear unbiased estimates (BLUEs) were calculated for each genotype, considering the effects of genotypes as fixed. BLUEs used for GWAS for a single environment were estimated as:

$$Y_{ik} = \mu + Rep_i + Gen_k + \varepsilon_{ik}$$

where Y_{ik} is the trait of interest, μ is the mean effect, Rep_i is the effect of the i^{th} replicate, Gen_k is the effect of the k^{th} genotype, ε_{ik} is the error associated with the i^{th} replication and the k^{th} genotype.

Narrow sense heritability was calculated using polyBreedR (https://github.com/jendelman/polyBreedR) utilizing a two-stage approach described by Damesa et al. (2017), using ASReml-R (VSN International, UK) for variance component estimation. To predict the breeding values, the polyBreedR function named predict_MME was used to implement mixed model equations (MME). Likewise, BLUEs used for GWAS for the combined environment considering genotypes fixed was estimated as:

$$Y_{ijr} = m + G_i + E_j + R_r(E) + GE_{ij} + e_{ijr}$$

where Y_{ijr} = trait value for the genotype i in the environment j and rep r, m = grand mean of trait, G = genotype, E = environment, R(E) = replication within environment, e = random error

3.2.5. Assessment of linkage disequilibrium

Linkage disequilibrium was estimated for all SNPs with pairwise correlation coefficient (r^2) based on allele frequencies (Hill & Robertson, 1968) using LD.plot function in GWASpoly (Rosyara et al., 2016). A monotone decreasing, convex spline was fit using the R package scam. The appropriate window size to filter the most significant markers was determined based on the extent of LD in the panel, visualized using the function LD.plot.

3.2.6. Genome wide association studies (GWAS)

Association analysis was performed for tuber traits with 10,116 SNPs using the GWASpoly Version 2 package in R (Rosyara et al., 2016). To control population structure, the leave-one-chromosome-out (LOCO) method (Yang et al., 2014) was used. Additive and dominant genetic models were tested for each trait. A Bonferroni test was run for each trait and year to establish a LOD threshold corresponding to a genome-wide false-positive rate of 5%. Manhattan plots were produced using GWASpoly. The proportion of phenotypic variance explained by significant SNPs was estimated by the function fit.QTL in GWASpoly. Contextual sequences of candidate SNPs were used in a BLAST search of DM1–3 pseudomolecules (Version 4.03) in the SpudDB database (http://solanaceae.plantbiology.msu.edu/) to identify putative candidate genes.

3.3. Results

3.3.1. Phenotypic data analysis

Analysis of variance and other descriptive statistics for the tuber morphology traits showed significant variability between the genotypes of the association mapping panel (Table 3.1). Frequency distribution of phenotypic data showed that most are not normally distributed (Figure 3.2).

Pearson correlation (r) analysis of combined analysis detected correlations between traits (Figure 3.3). Positive correlation was found between visual tuber shape and L/W ratio ($r = 0.89$; $P < 0.001$), visual tuber shape and russeting ($r = 0.78$; $P < 0.001$) (russet potatoes tend to be oblong to long), L/W and russeting ($r = 0.70$; $P < 0.001$) (similar interpretation but using objective values L/W), visual tuber shape and

average tuber weight ($r = 0.66$; $P < 0.001$), visual tuber shape and grading at table ($r = 0.27$; $P < 0.001$).

Negative correlation was found between tuber shape and average tubers per plant ($r = -0.70$; $P < 0.001$) (plants with more tubers tend to be rounder), L/W and average tubers per plant ($r = -0.59$; $P < 0.001$), tuber shape and flesh color chroma value ($r = -0.44$; $P < 0.001$), L/W and flesh color chroma value ($r = -0.40$; $P < 0.001$), average tuber per plant and average tuber weight, average tuber weight and chroma flesh, grading at table and chroma flesh color.

High broad-sense heritability ($H^2 > 0.95$) was observed for tuber shape, L/W, degree of russeting, and chroma value for flesh color (Table 3.1). In our experiment, higher H^2 values indicated that genetics contributed to the observed variability. The highest narrow sense heritability (h^2) of 0.91 was for flesh chroma value and lowest (0.33) was for average tuber weight per plant (Table 3.1).

Table 3.1 Estimates of best linear unbiased estimators, variance components, and broad-sense heritability for tuber morphology traits based on 214 clones evaluated in three environments (Dalhart 2019, Dalhart 2020, Springlake 2020).

		Tuber Shape	L/W	Russeting	Eye Depth	Av. Tuber Weight	Av. Tubers per Plant	Av. tuber wt per plant	Grading at Table	Flesh Chroma
		(1-5)	(ratio)	(1-5)	(1-5)	(g)	(no.)	(g)	(1-5)	(value)
Grand Mean	DAL 2019	2.7	1.4	1.9	4.0	113.1	8	804.2	3.5	19.6
	DAL 2020	2.7	1.4	2.1	4.0	126.2	8	931.8	3.7	17.7
	SPR 2020	2.4	1.4	1.9	4.0	87.8	9	650.8	3.6	17.1
	Overall	2.6	1.4	2.0	4.0	108.9	8	792.2	3.6	18.1
Range	DAL 2019	1-5	0.8-2.7	1-4.5	3.2-4.8	6-713	1-31	13-5664	1.0-4.6	11-52
	DAL 2020	1-5	0.8-2.8	1-4.6	3.0-4.6	9-405	2-22	108-3102	2.5-4.6	8-46
	SPR 2020	1-5	0.9-2.8	1-4.6	3.5-5.0	9-240	2-30	42-1302	2.0-4.5	7-36
	Overall	1-5	0.8-2.8	1-4.6	3.0-5.0	6-713	1-31	13-56664	1.0-4.6	7-52
LSD	DAL 2019	0.5	0.2	0.3	0.2	59.6	5.9	754.5	0.5	4.4
	DAL 2020	0.5	0.2	0.3	0.1	46.1	3.9	517.0	0.4	3.1
	SPR 2020	0.5	0.3	0.2	0.1	27.3	4.6	309.3	0.4	3.7
	Overall	0.6	0.2	0.3	0.2	38.4	3.4	314.5	0.4	2.4
CV	DAL 2019	9.0	8.3	6.5	2.8	28.6	36.0	45.7	7.6	10.9
	DAL 2020	9.3	8.8	6.2	1.8	19.7	24.1	27.8	4.9	8.7
	SPR 2020	11.3	12.1	5.1	1.3	15.7	27.3	23.9	5.4	11.0
	Overall	9.8	9.4	5.9	2.0	22.8	29.3	34.6	6.0	10.3
H^2	DAL 2019	0.98	0.95	0.99	0.90	0.83	0.72	0.58	0.82	0.96
	DAL 2020	0.98	0.94	0.99	0.93	0.87	0.79	0.57	0.82	0.98
	SPR 2020	0.98	0.90	0.99	0.98	0.93	0.79	0.72	0.82	0.96
	Overall	0.97	0.98	0.99	0.70	0.89	0.78	0.59	0.64	0.98
h^2	Overall	0.78	0.75	0.82	0.46	0.58	0.45	0.33	0.43	0.91
Genotypic Variance	DAL 2019	1.8	0.1	1.7	0.1	2570.4	10.7	92747	0.2	51.5
	DAL 2020	1.5	0.1	2.1	0.0	2139.2	7.3	44468	0.1	48.6
	SPR 2020	1.7	0.1	1.9	0.1	1445.2	10.0	31584	0.1	41.4
	Overall	1.6	0.1	1.9	0.0	1780.3	6.5	30433	0.1	47.7
Residual Variance	DAL 2019	0.1	0.0	0.0	0.0	1049.9	8.3	134935	0.1	4.6
	DAL 2020	0.1	0.0	0.0	0.0	617.4	3.9	67303.9	0.0	2.4
	SPR 2020	0.1	0.0	0.0	0.0	189.5	5.3	24258.9	0.0	3.6
	Overall	0.1	0.0	0.0	0.0	616.4	5.8	75241.9	0.0	3.5
GenxLoc Variance		0.1	0.0	0.0	0.0	313.0	2.9	26359.5	0.1	0.5

Table 3.2 Mean values (BLUEs) of potatoes belonging to various market classes for tuber morphological traits.

Market Class	Tuber Shape	L/W	Russeting	Eye Depth	Av. Tuber Weight	Av. Tubers Per Plant	Av. Tuber Weight Per Plant	Grading at table	Flesh Chroma
Market Class	(1-5)	(ratio)	(1-5)	(1-5)	(g)	(no.)	(g)	(1-5)	(value)
Russet	4.0[a]	1.7[a]	3.8[a]	3.9[c]	151[a]	5[d]	816[b]	3.7[a]	13.2[e]
Purple	2.4[b]	1.4[b]	1.0[c]	4.0[a]	76[c]	9[b]	653[c]	3.5[bc]	21.0[b]
Chip	1.8[c]	1.1[c]	1.1[b]	4.0[bc]	117[b]	7[c]	892[a]	3.6[b]	15.3[d]
Yellow	1.7[c]	1.1[c]	1.0[c]	4.0[ab]	78[c]	10[a]	768[b]	3.4[c]	27.2[a]
Red	1.7[c]	1.1[c]	1.0[c]	3.9[c]	83[c]	10[a]	801[b]	3.5[b]	19.6[c]

Values followed by the same letter (within each column) were not significantly different based on Students t test ($P < 0.05$)

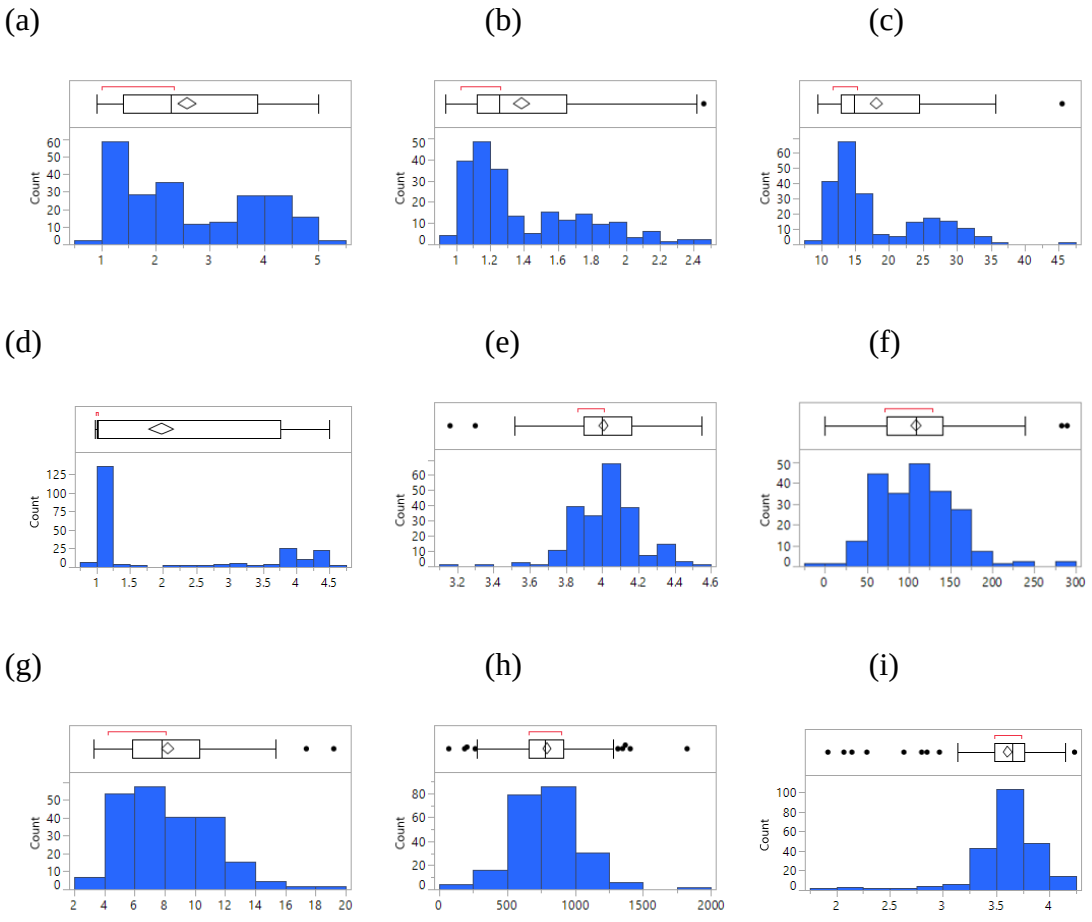


Figure 3.2 Distribution of (a) tuber shape, (b) L/W, (c) flesh chroma value, (d) russeting, (e) eye depth, (f) average weight of a tuber, (g) average tuber number per plant, (h) average tuber weight per plant, and (i) grading at table in combined environment.

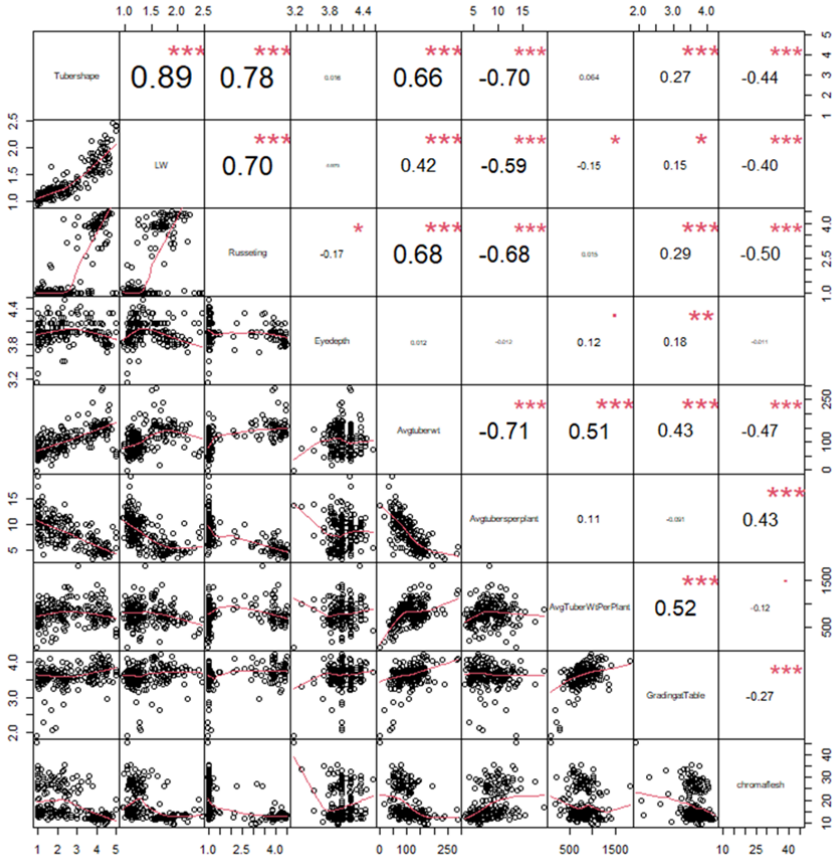


Figure 3.3 The bivariate scatter plots with a fitted line for the traits combined across all environments are displayed on the bottom of the diagonal. On the top of the diagonal are the value of the correlation plus the significance level as stars. Each significance level is associated to a symbol : p < 0.001 = *; p < 0.01 = ** and p < 0.05 = *.**

3.3.2. Linkage disequilibrium (LD)

Genome-wide LD is quite low (Figure 3.4). From the shape of the curve, a 5-10 Mb window seems appropriate to filter the most significant marker in the output.

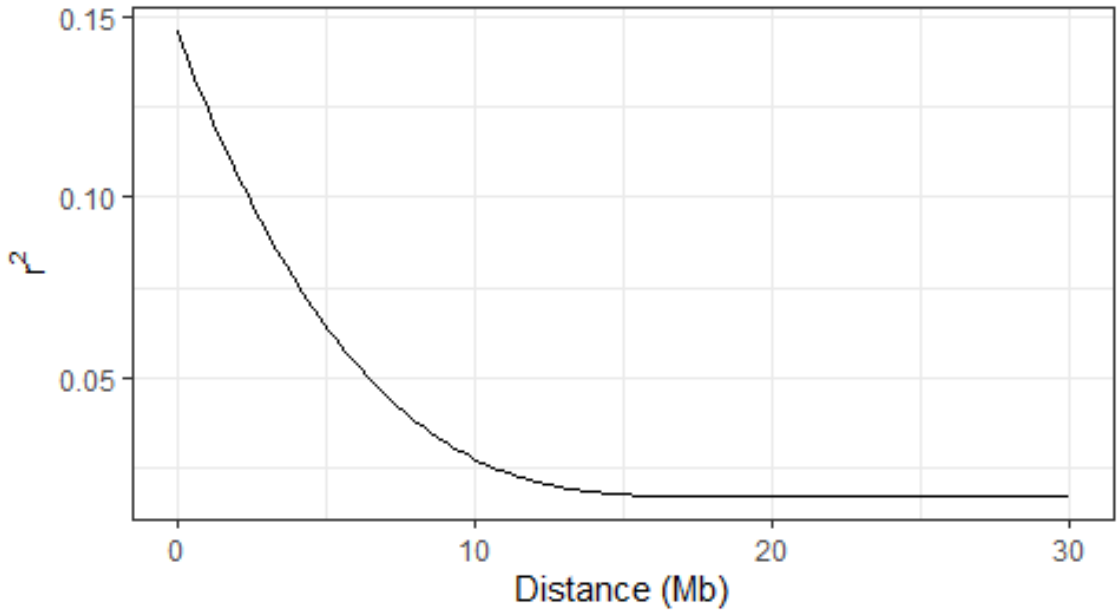


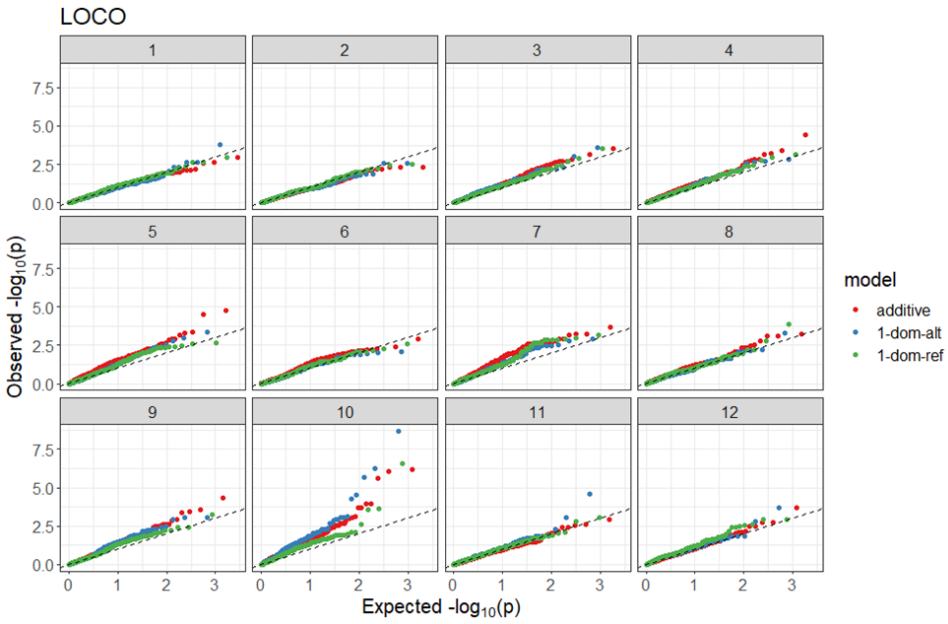
Figure 3.4 Genome-wide LD decay (in Mb) in potato using LD.plot function in GWASpoly (Rosyara et al., 2016). A monotone decreasing, convex spline was fit using R package scam.

3.3.3. GWAS Analysis

The inflation of the -log10(p) was examined using a quantile-quantile plot (Q–Q plots) of the observed vs. expected values under the null hypothesis, which follows a uniform distribution. Manhattan plots showed the significance threshold for each locus along with the location of SNPs corresponding to tuber shape (Figure 3.5), L/W (Figure 3.6), eye depth (Figure 3.7), degree of russeting (Figure 3.8), grading at table (Figure 3.9), intensity (chroma) of flesh color (Figure 3.10) and purple skin color (Figure 3.11). A QTL defined by three significant SNPs on chromosome 10 at 48 Mb was identified for tuber shape in the additive and dominant models (Table 3.4). A significant SNP (solcap_snp_c2_25485) at position of 48.73 Mbp explained 6% and solcap_snp_c2_25522 at a position of 48.61Mbp explained 4% of the tuber shape

variation. QTLs maps to the genome super scaffold PGSC0003DMB000000385 which annotates as a ribosomal protein S6 kinase. The QTLs for L/W on chromosome 10 explained 6% of the variation whereas each QTL on chromosomes 1 and 5 explained 4% and on chromosome 9 explained 3%. QTLs for eye depth were found on chromosomes 3, 5 and 10. QTLs on chromosomes 1 and 12 for degree of russeting explained 10% of the trait variance. A QTL for grading at table score was identified on chromosome 4. QTLs for intensity (chroma) of flesh color were detected on chromosome 1 and 3. The SNP (PotVar0120627) explained 26% of the trait variance. This marker was mapped in super scaffold PGSC0003DMG400010169 where beta-carotene hydroxylase is located. Likewise, QTLs for purple skin color were detected on chromosomes 1, 3 and 11 and explained 11%, 14%, and 25% of the trait variance.

(a)



(b)

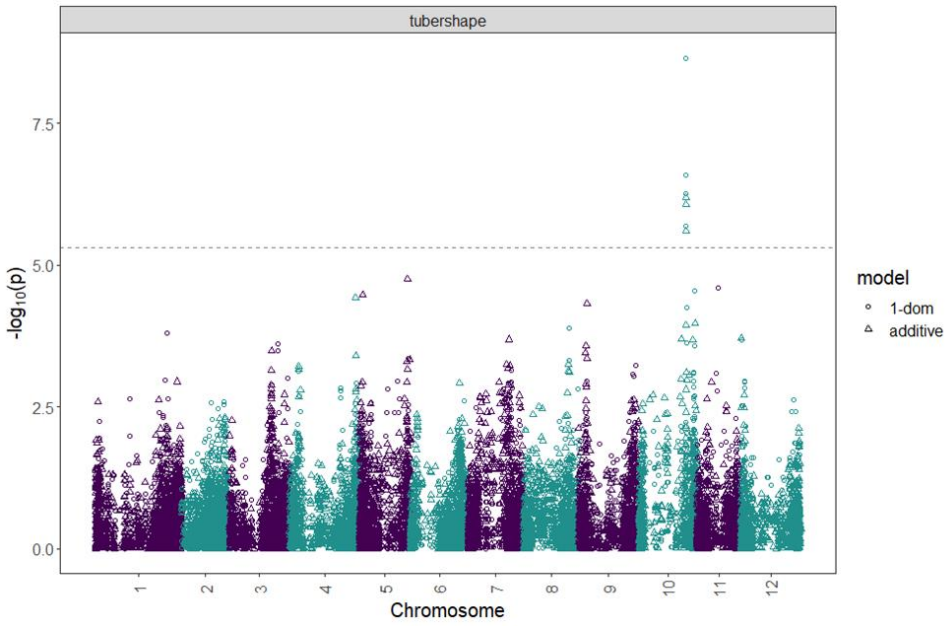
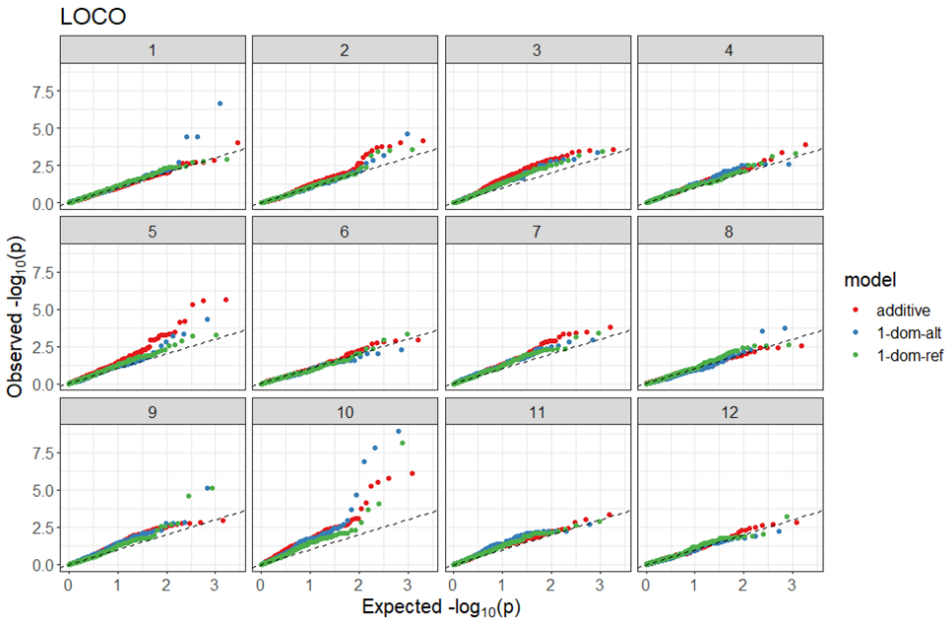


Figure 3.5 Q–Q plots (a) and corresponding Manhattan plot (b) of observed versus expected -log10 (P values) for tuber shape using the additive and dominant model in three combined environments. The Bonferroni threshold is at 5.31 for the additive, 4.96 for 1-dom-alt and 5.09 for the 1-dom-ref model.

(a)



(b)

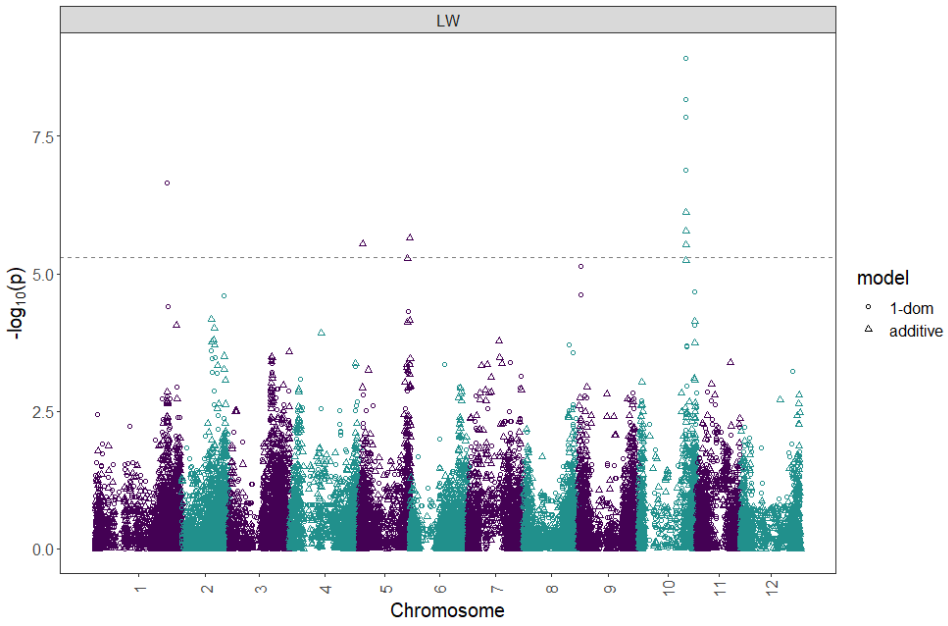
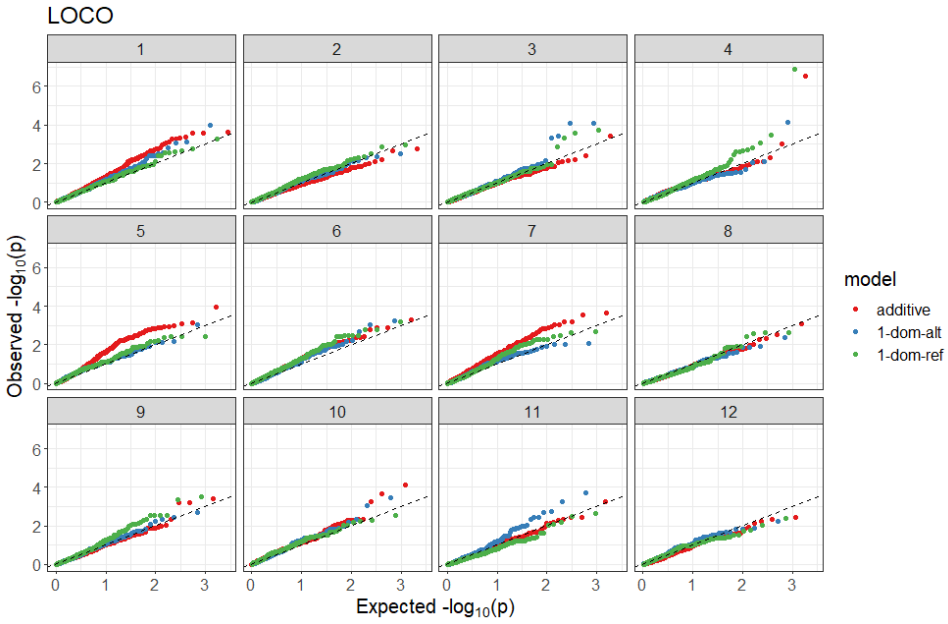


Figure 3.6 Q–Q plots (a) and corresponding Manhattan plot (b) of observed versus expected -log10 (P values) for length width ration (L/W) using the additive and dominant model in three combined environments. The Bonferroni threshold is at 5.31 for the additive, 4.96 for 1-dom-alt and 5.09 for the 1-dom-ref model.

(a)



(b)

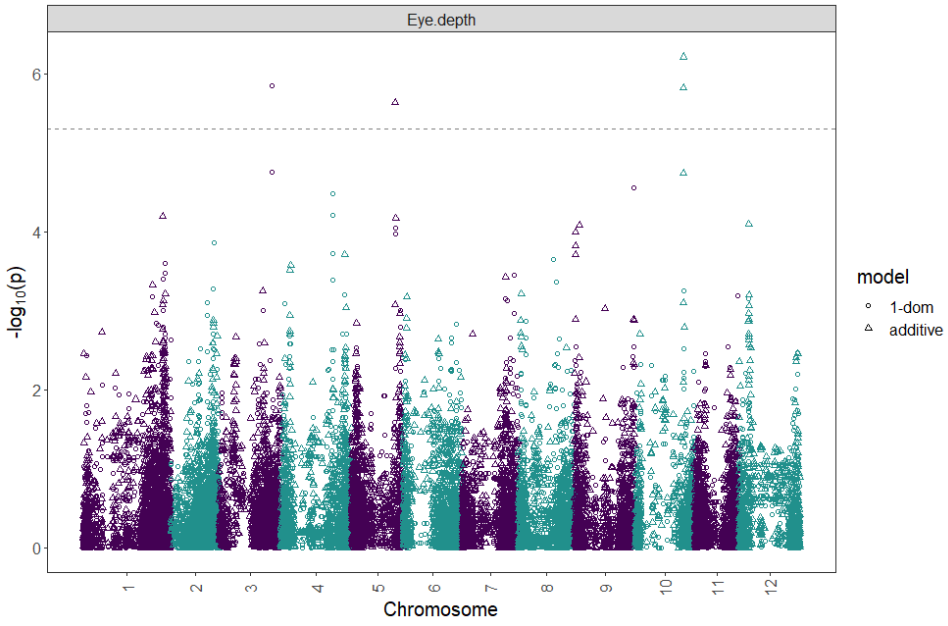
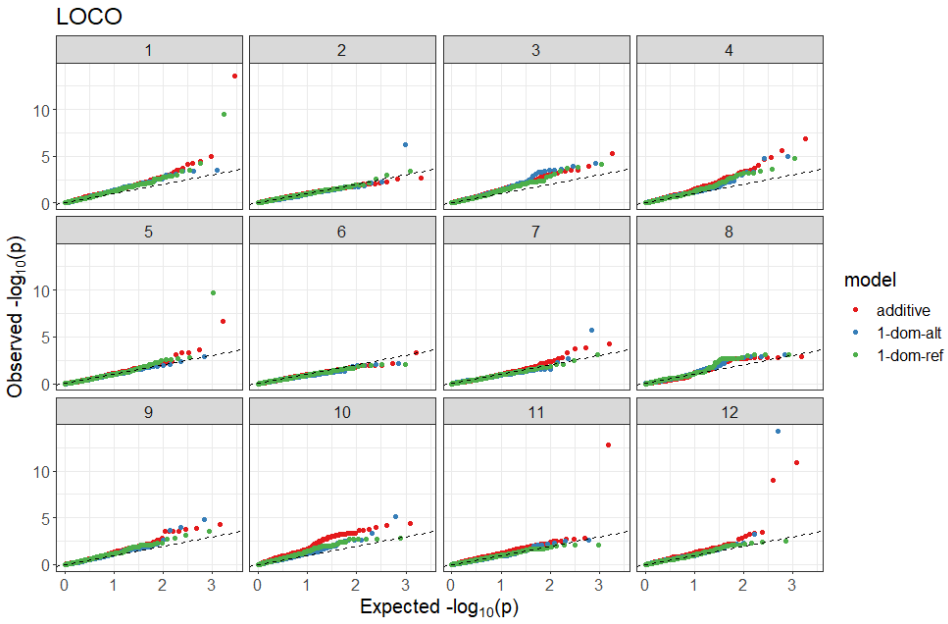


Figure 3.7 Q–Q plots (a) and corresponding Manhattan plot (b) of observed versus expected -log10 (P values) for eye depth using the additive and dominant model in three combined environments. The Bonferroni threshold is at 5.31 for the additive, 4.96 for 1-dom-alt and 5.09 for the 1-dom-ref model.

(a)



(b)

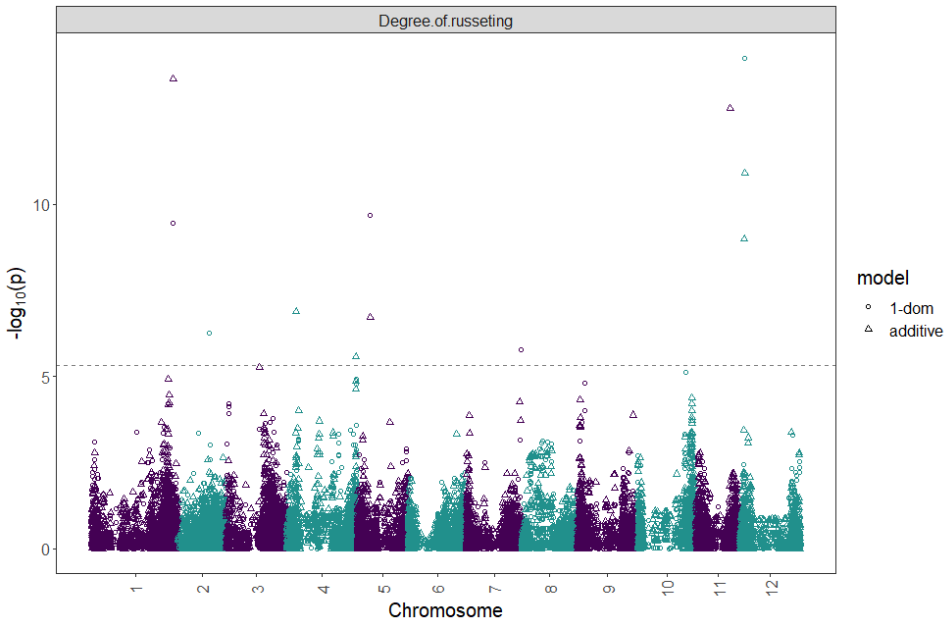
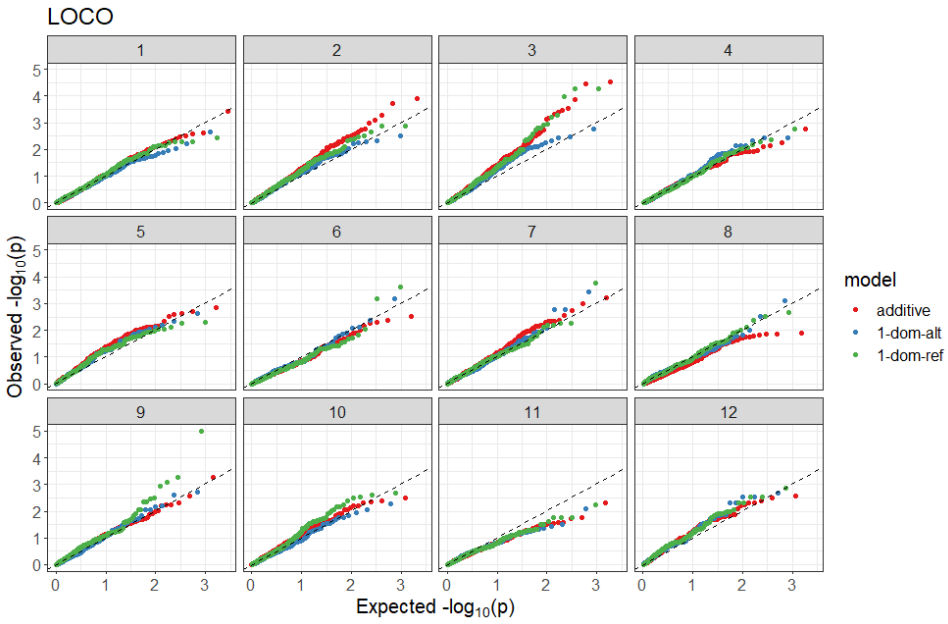


Figure 3.8 Q–Q plots (a) and corresponding Manhattan plot (b) of observed versus expected -log10 (P values) for degree of russeting using the additive and dominant model in three combined environments. The Bonferroni threshold is at 5.31 for the additive, 4.96 for 1-dom-alt and 5.09 for the 1-dom-ref model.

(a)



(b)

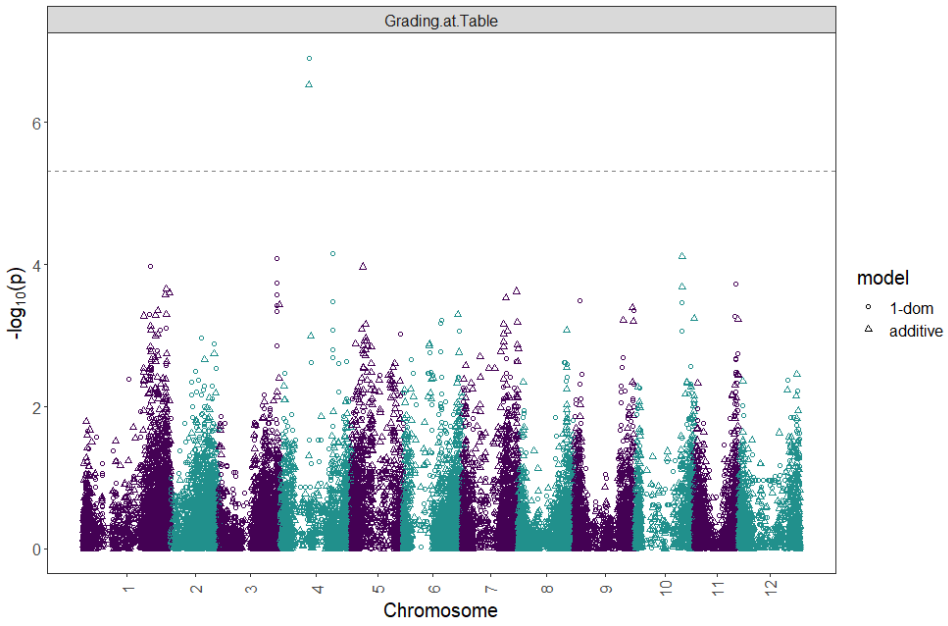
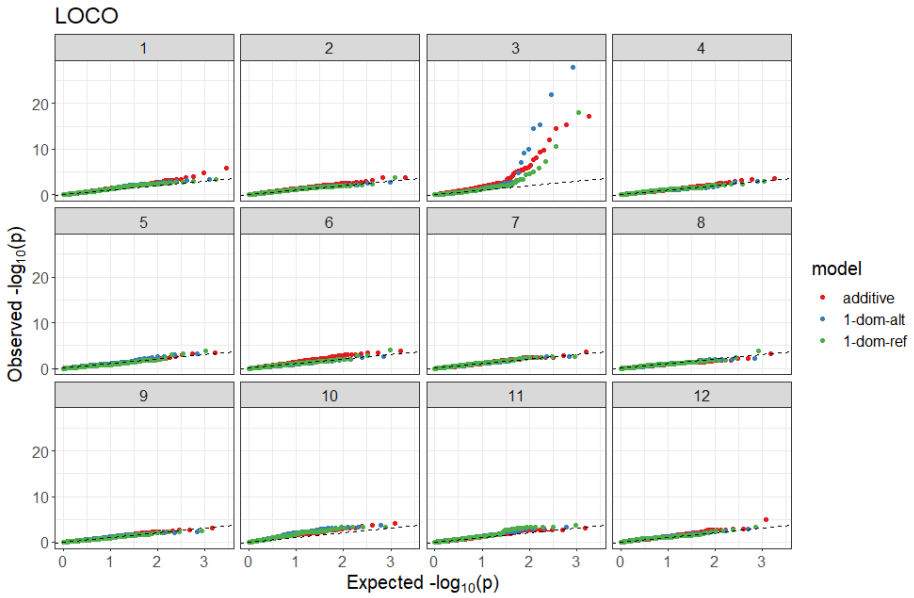


Figure 3.9 Q–Q plots (a) and corresponding Manhattan plot (b) of observed versus expected -log10 (P values) for grading at table using the additive and dominant model in three combined environments. The Bonferroni threshold is at 5.31 for the additive, 4.96 for 1 dom alt and 5.09 for the 1-dom-ref model.

(a)



(b)

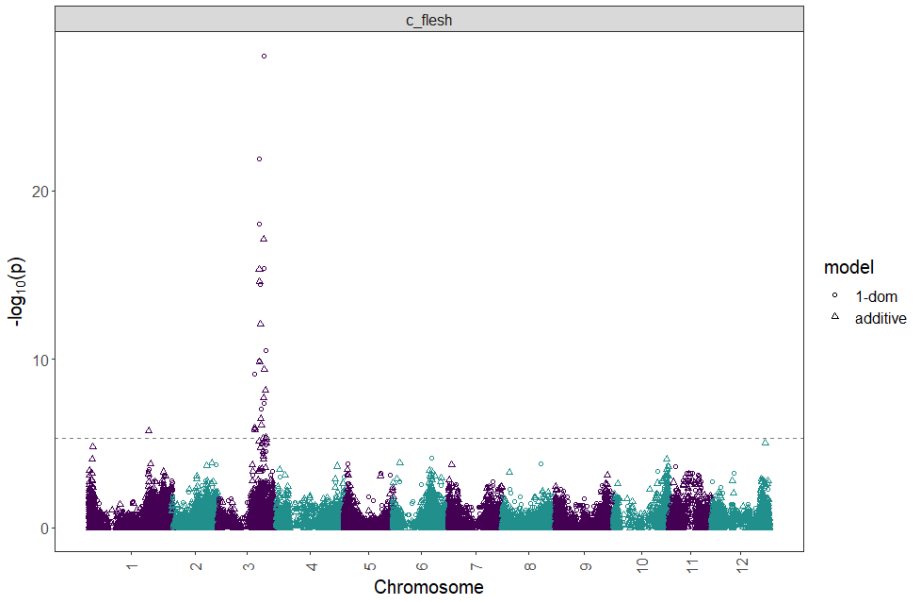
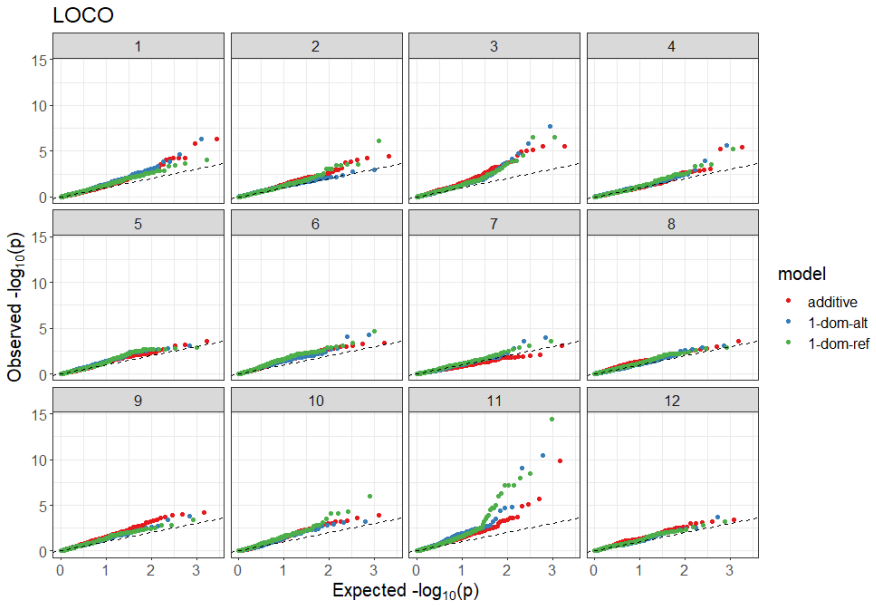


Figure 3.10 Q–Q plots (a) and corresponding Manhattan plot (b) of observed versus expected -log10 (P values) for intensity (chroma) of flesh color (c_flesh) using the additive and dominant model in three combined environments. The Bonferroni threshold is at 5.31 for the additive, 4.96 for 1-dom-alt and 5.09 for the 1-dom-ref model.

(a)



(b)

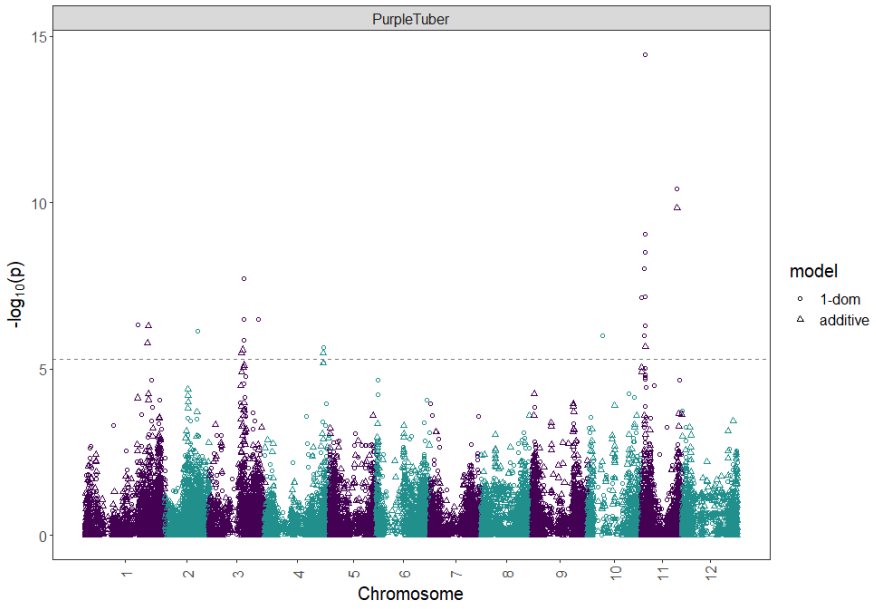


Figure 3.11 Q–Q plots (a) and corresponding Manhattan plot (b) of observed versus expected -log10 (P values) for purple skin color using the additive and dominant model in three combined environments. The Bonferroni threshold is at 5.31 for the additive, 4.96 for 1-dom-alt and 5.09 for the 1-dom-ref model.

Table 3.3 Significant SNPs associated with tuber morphology traits in tetraploid advanced potato clones.

Trait	Model	Threshold	Marker	Chrom	Position	Score	Effect	R2
Tuber shape	additive	5.31	solcap_snp_c2_25485	10	48737840	6.2	-0.38	0.04
Tuber shape	1-dom-alt	4.96	solcap_snp_c2_25485	10	48737840	8.7	-1.14	0.06
Tuber shape	1-dom-ref	5.09	solcap_snp_c2_25522	10	48617457	6.6	0.98	0.04
LW ratio	additive	5.31	PotVar0024712	5	2963647	5.6	0.10	0.03
LW ratio	additive	5.31	solcap_snp_c1_1124	5	51698153	5.7	0.09	0.04
LW ratio	additive	5.31	solcap_snp_c2_25522	10	48617457	6.1	0.10	0.04
LW ratio	1-dom-alt	4.96	solcap_snp_c1_14093	1	73743526	6.7	-0.34	0.04
LW ratio	1-dom-alt	4.96	PotVar0011392	9	2113876	5.1	-0.16	0.03
LW ratio	1-dom-alt	4.96	solcap_snp_c2_25485	10	48737840	8.9	-0.31	0.06
LW ratio	1-dom-ref	5.09	solcap_snp_c1_3597	9	2081425	5.1	0.16	0.03
LW ratio	1-dom-ref	5.09	solcap_snp_c2_25522	10	48617457	8.2	0.29	0.05
Eye depth	additive	5.31	PotVar0033968	5	43920348	5.6	0.12	0.03
Eye depth	additive	5.31	solcap_snp_c2_25510	10	48679881	6.2	-0.07	0.04
Eye depth	1-dom-ref	5.09	solcap_snp_c1_4577	3	53162686	5.9	-0.19	0.04
Degree of russeting	additive	5.31	solcap_snp_c2_4840	1	83042096	13.6	-0.63	0.10
Degree of russeting	additive	5.31	solcap_snp_c1_8346	4	9523734	6.9	-0.32	0.04
Degree of russeting	additive	5.31	PotVar0015716	4	70761963	5.6	-0.59	0.04
Degree of russeting	additive	5.31	PotVar0109849	5	13342810	6.7	0.76	0.06
Degree of russeting	additive	5.31	solcap_snp_c1_9499	11	34592587	12.8	0.41	0.09
Degree of russeting	additive	5.31	solcap_snp_c2_31247	12	4322254	10.9	0.38	0.08
Degree of russeting	1-dom-alt	4.96	PotVar0094383	2	31449005	6.3	-0.51	0.05
Degree of russeting	1-dom-alt	4.96	PotVar0037013	7	55736406	5.8	0.37	0.04
Degree of russeting	1-dom-alt	4.96	solcap_snp_c2_25485	10	48737840	5.1	-0.70	0.03
Degree of russeting	1-dom-alt	4.96	solcap_snp_c2_31301	12	4189099	14.2	-0.90	0.10
Degree of russeting	1-dom-ref	5.09	solcap_snp_c2_4840	1	83042096	9.4	-0.69	0.07
Degree of russeting	1-dom-ref	5.09	PotVar0109849	5	13342810	9.7	0.93	0.08
c_flesh	additive	5.31	PotVar0126039	1	63841740	5.8	-1.68	0.04
c_flesh	additive	5.31	PotVar0120627	3	48550473	17.1	5.66	0.13
c_flesh	1-dom-alt	4.96	PotVar0120627	3	48550473	28.0	11.52	0.26
c_flesh	1-dom-ref	5.09	solcap_snp_c2_17552	3	44178120	18.0	-7.39	0.14
c_flesh	1-dom-ref	5.09	PotVar0056507	3	50884766	10.5	-6.15	0.07
Grading at table	additive	5.31	PotVar0073716	4	28848091	6.5	0.23	0.04
Grading at table	1-dom-ref	5.09	PotVar0073716	4	28848091	6.9	0.27	0.04
Purple tuber	additive	5.31	PotVar0043559	1	69871607	6.3	0.13	0.11
Purple tuber	additive	5.31	solcap_snp_c2_52494	3	37495967	5.6	-0.09	0.10
Purple tuber	additive	5.31	PotVar0000760	4	64804249	5.5	-0.18	0.10
Purple tuber	additive	5.31	solcap_snp_c1_14082	11	4327085	5.7	-0.10	0.10
Purple tuber	additive	5.31	PotVar0047229	11	39417891	9.8	-0.29	0.17
Purple tuber	1-dom-alt	4.97	solcap_snp_c2_49619	1	58072346	6.3	0.22	0.11
Purple tuber	1-dom-alt	4.97	solcap_snp_c2_38058	3	38756280	7.7	0.23	0.14
Purple tuber	1-dom-alt	4.97	PotVar0000760	4	64804249	5.6	-0.23	0.10
Purple tuber	1-dom-alt	4.97	solcap_snp_c2_6185	11	3572445	9.1	0.29	0.16
Purple tuber	1-dom-alt	4.97	PotVar0047229	11	39417891	10.4	-0.32	0.18
Purple tuber	1-dom-ref	5.09	PotVar0128389	2	35666972	6.1	-0.38	0.11
Purple tuber	1-dom-ref	5.09	solcap_snp_c1_11350	3	38435099	6.5	-0.23	0.11
Purple tuber	1-dom-ref	5.09	solcap_snp_c2_22512	3	54971939	6.5	-0.27	0.11
Purple tuber	1-dom-ref	5.09	PotVar0000767	4	64803299	5.2	0.21	0.09
Purple tuber	1-dom-ref	5.09	solcap_snp_c2_55	10	15913000	6.0	-0.37	0.11
Purple tuber	1-dom-ref	5.09	solcap_snp_c1_14082	11	4327085	14.4	-0.48	0.25

3.3.4. Breeding values of tuber traits

In a potato breeding program, estimated breeding values utilizing best linear unbiased prediction (BLUP) through pedigree relationships can improve selection efficiency and save time and resources required for tuber multiplication and phenotyping. Based on the genomic estimated breeding values BLUP/EBVs (Appendix B1), clones PTTX05PG07-1W, COTX08365F-3P/P and ATX84378-6Ru were the best for long tuber shape and clones COTX10138-15Wpe/Y, NDTX5003-2R and ATX91322-2Y/Y were the best for round tuber shape. According to BLUP, clone TX08385-2P/YP clone was the best to eliminate russeting and clone AOTX95265-4Ru was the best for high degree of russeting. Clone ATX91322-2Y/Y was predicted to have the shallowest eye depth. Likewise, clones ATX84706-2Ru, NDTX092238Cs-1P/W and Sierra Gold were the best for average tuber weight, average tubers per plant and average tuber weight per plant respectively. Clone TX13590-9Ru was predicted to be the best while grading at table.

3.4. Discussion

Breeding for quality traits in potato requires knowledge of the genetics of tuber morphological characteristics. Molecular markers could provide an efficient and economical means to identify potatoes with desirable traits earlier in the breeding process. Identifying the current allelic variation of the primary genes and using the allele that delivers the desired phenotype in breeding programs is crucial. In this study, we present the identified SNPs associated with tuber shape, eye depth, degree of russeting, tuber number, tuber weight, skin color, and flesh color.

There was a significant variation for most evaluated traits. However, the frequency distribution of phenotypic data was not normally distributed. A common approach for these types of analysis is to carry out a transformation on the data to have a normal distribution (Feng et al., 2014). Beside transformations method, generalized linear mixed models provides a more flexible approach to evaluating non-normal data with random effects (Piepho et al., 2003).

Strong correlation existed between the traits that display phenotypic variability among the genotypes. Prashar et al. (2014) used a 1–4 ‘breeder’s scale’ and more ‘quantitative LW’ method achieved a correlation coefficient of 0.91 between the two scoring methods. Our study found a correlation of 0.83 between ‘assessed’ tuber shape and ‘measured’ (i.e., L/W). The high correlation of both implies that the visual evaluation approach may be rapid and useful for shape analysis in potato breeding programs. However, to achieve higher accuracies and to perform high-throughput analysis digital calipers and digital image-based methods are being explored (Si et al., 2017; Przybylak et al., 2020). Most of the russet potatoes are long and bulky (Table 3.2), so the correlation coefficient of russeting appears to be high with tuber shape and L/W. A negative correlation (r= -0.50) was observed between tuber shape and tubers per plant which means that if tubers are round there will be many tubers per plant and if tubers are elongated there will be few tubers per plant. Thus, the understanding of defined characteristics will make a visual and indirect selection in potato breeding programs more efficient.

High broad-sense heritability achieved for tuber shape, L/W, degree of russeting, and chroma value for flesh color traits indicated that primarily genetic factors control tuber morphology traits. Love et al. (1997) had observed broad sense heritability above 0.80 for russeting; our study found it to be 0.99. Moderate broad-sense heritability was observed in eye depth, and average tuber weight per plant. These moderate to high broad-sense heritability estimates suggested that either traditional selection or genomic selection would be able to achieve the genetic gains.

Likewise, narrow-sense heritability also referred to as "SNP-based heritability" for russeting was highest (h^2= 0.82) compared to other traits. For all the traits studied, SNP-based heritability estimates were lower than broad-sense heritability estimates. The lower heritability estimates can be due to the underrepresentation of significant causal loci by markers (Speed et al., 2012). In accordance with Sood et al. (2020), narrow-sense heritability was moderate for eye depth, average tuber weight, and average tubers per plant in our study. A trait with a high narrow-sense heritability means that most of the phenotypic variation is genetically controlled, and that selection will be successful (Ozturk & Yildirim, 2014; Ozimati et al., 2019). Hence, narrow-sense heritability, the measure of the additive genetic effect, would be extremely useful in a breeding program aimed at exploiting continuous genetic gain.

In this study, the GWAS experiment yielded putative and novel markers/genes involved with tuber morphology traits on a representative diversity panel. High-density SNP markers were applied to the panel of 214 clones considering dosages and allele interactions. Some of the associations explained larger proportions of phenotypic

variation than others. This study provides useful genomic resources in potatoes as well as empowers GWAS analyses in the polyploid species.

We report a significant SNP, solcap_snp_c2_25485 at 48.7 Mb on chromosome 10 to be associated with tuber shape explaining 6% of the trait variance in the dominant model. Solcap_snp_c2_25522 on the same chromosome, explained around 5% of the phenotypic variance in the dominant model. Zia et al. (2020) also detected the SNP solcap_snp_c2_25485 associated with tuber shape, but the SNP solcap_snp_c2_25522 was not. Previous biparental mapping studies (ŚLiwka et al., 2008; Prashar et al., 2014; Hara-Skrzypiec et al., 2018) have also mapped major QTL for tuber shape to the same locus on chromosome 10. *Ro* locus on chromosome 10 has been identified as the major locus controlling tuber shape (Van Eck et al., 1994a). GWAS by Sharma et al. (2018) reported 'solcap_snp_c1_8019' as the most significant association located at 48.9 Mb. In addition to the QTLs identified for the tuber shape, we report additional QTLs on chromosome 1, 5 and 9 for L/W.

Li et al. (2005) had identified locus *Ro* closely linked with a major locus for eye depth. In our study also a QTL on chromosome 10 for eye depth co-localized with tuber shape. The QTL peak for eye depth identified on chromosome 5 is located near the CDF1 gene, an important regulator of maturity in potato (Kloosterman et al., 2013). Clark (1933) had discovered that the action of complementary factors resulted in the russet skin in tubers, but the number of genes involved were not indicated. Our study found that multiple QTLs on chromosomes 1, 2, 4, 5, 7, 10, 11 and 12 govern russeting in potato.

In most tetraploid potato cultivars, the flesh color varies from white to dark yellow due to varying amounts of carotenoids which convey the yellow color. The dominant *Y*-locus is largely responsible for the yellow flesh color of potatoes, and it maps to chromosome 3 (Bonierbale et al., 1988; Thorup et al., 2000). Beta-carotene hydroxylase (BCH) is the most likely candidate gene responsible for the yellow flesh color (Brown et al., 2006; Kloosterman et al., 2010). One isoform (PGSC0003DMG400009501) of the *BCH* gene is located at 44.1 Mbp (Xu et al., 2011). In our study, a strong association for flesh color was observed on chromosome 3 at 48.5 Mb (PotVar0120627) and another significant association at 43.9 Mb (PotVar0070260) approximately 0.2 Mb away from the causal gene. The most significant SNPs (PotVar0120627 and solcap_snp_c2_17552) explained 26% and 14% of the trait variance, respectively.

The presence of anthocyanins in a limited number of cultivars result in red or blue/purple flesh (Bonar et al., 2018). The detection of genomic regions and desirable alleles linked to anthocyanin content could be used in potato breeding programs to increase bioactive compounds. In our study, QTLs for purple skin color were detected on chromosome 1 at 69.8 Mbp and chromosome 11 at 4.3 Mbp. Parra-Galindo et al. (2019) detected QTLs associated with cyanidin and delphinidin concentrations, on chromosome 1 at 50.7 Mbp and chromosome 11 (40 Mbp), respectively. Previous studies showed that the potato *P* locus is required to produce blue/purple anthocyanin pigments and it is located on chromosome 11 (De Jong et al., 2004; Jung et al., 2005).

GWAS methodology is capable of dissecting some complex genetic traits. However, no QTL were detected for yield related traits like average tubers per plant, and tubers weight per plant. The reason for this failure was likely. To acquire appropriate statistical power, GWAS require much larger sample sizes (Hong & Park, 2012). Also, these traits are highly quantitative, GWAS may fall short of uncovering the causative loci we seek. Korte & Farlow (2013) indicated that GWAS has difficulties in detecting more subtle or multiple QTL. One potential approach to deal with this issue could be refining the phenotype of interest to reduce the number of loci involved in the trait and thus increase the detection ability (Benjamin et al., 2012).

Plant breeding is a continuous genetic improvement process that involves the selection and recombination of superior lines. The majority of important traits in potato breeding are complex and controlled by several genes. Because estimated breeding values (EBVs) only assess the additive genetic effect, BLUP is an appropriate technique for complex traits with low heritability (Slater, Cogan, et al., 2014). This study displayed the potential advantage of EBVs using BLUP for effective selection for tuber traits in potato breeding. For low heritability variables like tuber yield, BLUP calculated breeding values outperform phenotypic selection (Sood et al., 2020). Parent selection based on BLUPs may ultimately enable new and improved potatoes.

3.5. Conclusions

The main purpose of the current study was to find SNPs associated with tuber shape, eye depth, degree of russeting, tuber number, tuber weight, skin color, and flesh color in potatoes. Besides the rediscovery of a QTL for tuber morphology traits in

potatoes, we identified new loci affecting the variation in studied traits. Following validation with genotyping technologies such as Kompetitive allele-specific PCR (KASP) markers, these markers could be potentially utilized in future potato-breeding programs for marker-assisted selection (MAS). The BLUPs for clones could be used to develop a selection index for selecting superior individuals across all traits.

3.6. References

Achenbach, U., Paulo, J., Ilarionova, E., Lübeck, J., Strahwald, J., Tacke, E., Hofferbert, H.-R., & Gebhardt, C. (2009). Using SNP markers to dissect linkage disequilibrium at a major quantitative trait locus for resistance to the potato cyst nematode *Globodera pallida* on potato chromosome V. *Theoretical and Applied Genetics*, *118*(3), 619–629. https://doi.org/10.1007/s00122-008-0925-x

Alvarado, G., Rodríguez, F. M., Pacheco, A., Burgueño, J., Crossa, J., Vargas, M., Pérez-Rodríguez, P., & Lopez-Cruz, M. A. (2020). META-R: A software to analyze data from multi-environment plant breeding trials. *The Crop Journal*, *8*(5), 745–756. https://doi.org/10.1016/j.cj.2020.03.010

Benjamin, D. J., Cesarini, D., Chabris, C. F., Glaeser, E. L., Laibson, D. I., Guðnason, V., Harris, T. B., Launer, L. J., Purcell, S., Smith, A. V., Johannesson, M., Magnusson, P. K. E., Beauchamp, J. P., Christakis, N. A., Atwood, C. S., Hebert, B., Freese, J., Hauser, R. M., Hauser, T. S., Lichtenstein, P. (2012). The promises and pitfalls of genoeconomics*. *Annual Review of Economics*, *4*, 627–662. https://doi.org/10.1146/annurev-economics-080511-110939

Bonar, N., Liney, M., Zhang, R., Austin, C., Dessoly, J., Davidson, D., Stephens, J., McDougall, G., Taylor, M., Bryan, G. J., & Hornyik, C. (2018). Potato miR828 is associated with purple tuber skin and flesh color. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.01742

Bonierbale, M. W., Amoros, W. R., Salas, E., & de Jong, W. (2020). Potato breeding. In H. Campos & O. Ortiz (Eds.), *The Potato Crop: Its Agricultural, Nutritional and*

Social Contribution to Humankind (pp. 163–217). Springer International Publishing. https://doi.org/10.1007/978-3-030-28683-5_6

Bonierbale, M. W., Plaisted, R. L., & Tanksley, S. D. (1988). RFLP Maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics*, *120*(4), 1095–1103.

Bourke, P. M., van Geest, G., Voorrips, R. E., Jansen, J., Kranenburg, T., Shahin, A., Visser, R. G. F., Arens, P., Smulders, M. J. M., & Maliepaard, C. (2018). PolymapR—linkage analysis and genetic map construction from F1 populations of outcrossing polyploids. *Bioinformatics*, *34*(20), 3496–3502. https://doi.org/10.1093/bioinformatics/bty371

Bourke, P. M., Voorrips, R. E., Visser, R. G. F., & Maliepaard, C. (2018). Tools for genetic studies in experimental populations of polyploids. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.00513

Bradshaw, J. E., Hackett, C. A., Pande, B., Waugh, R., & Bryan, G. J. (2008). QTL mapping of yield, agronomic and quality traits in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*). *Theoretical and Applied Genetics*, *116*(2), 193–211. https://doi.org/10.1007/s00122-007-0659-1

Brown, C. R., Kim, T. S., Ganga, Z., Haynes, K., De Jong, D., Jahn, M., Paran, I., & De Jong, W. (2006). Segregation of total carotenoid in high level potato germplasm and its relationship to beta-carotene hydroxylase polymorphism. *American Journal of Potato Research*, *83*(5), 365–372. https://doi.org/10.1007/BF02872013

Chen, N., Zhu, W., Xu, J., Duan, S., Bian, C., Hu, J., Wang, W., Li, G., & Jin, L. (2018). Molecular marker development and primary physical map construction for the tuber shape Ro gene locus in diploid potato (*Solanum tuberosum* L.). *Molecular Breeding*, *39*(1), 6. https://doi.org/10.1007/s11032-018-0913-z

Cheng, F., Sun, R., Hou, X., Zheng, H., Zhang, F., Zhang, Y., Liu, B., Liang, J., Zhuang, M., Liu, Y., Liu, D., Wang, X., Li, P., Liu, Y., Lin, K., Bucher, J., Zhang, N., Wang, Y., Wang, H., … Wang, X. (2016). Subgenome parallel selection is associated with morphotype diversification and convergent crop domestication in *Brassica rapa* and *Brassica oleracea*. *Nature Genetics*, *48*(10), 1218–1224. https://doi.org/10.1038/ng.3634

Clark, C. F. (1933). Further studies of the origin of russeting in the potato. *American Potato Journal*, *10*(5), 88–91. https://doi.org/10.1007/BF02880771

Cortes, L. T., Zhang, Z., & Yu, J. (2021). Status and prospects of genome-wide association studies in plants. *The Plant Genome*, *14*(1), e20077. https://doi.org/10.1002/tpg2.20077

Damesa, T. M., Möhring, J., Worku, M., & Piepho, H.-P. (2017). One step at a time: Stage-wise analysis of a series of experiments. *Agronomy Journal*, *109*(3), 845–857. https://doi.org/10.2134/agronj2016.07.0395

De Jong, H. (1981). Inheritance of russeting in cultivated diploid potatoes. *Potato Research*, *24*(3), 309–313. https://doi.org/10.1007/BF02360368

De Jong, W. S., De Jong, D. M., De Jong, H., Kalazich, J., & Bodis, M. (2003). An allele of dihydroflavonol 4-reductase associated with the ability to produce red

anthocyanin pigments in potato (*Solanum tuberosum* L.). *Theoretical and Applied Genetics*, *107*(8), 1375–1383. https://doi.org/10.1007/s00122-003-1395-9

De Jong, W. S., Eannetta, N. T., De Jong, D. M., & Bodis, M. (2004). Candidate gene analysis of anthocyanin pigmentation loci in the Solanaceae. *Theoretical and Applied Genetics*, *108*(3), 423–432. https://doi.org/10.1007/s00122-003-1455-1

Eichhorn, S., & Winterhalter, P. (2005). Anthocyanins from pigmented potato (*Solanum tuberosum* L.) varieties. *Food Research International*, *38*(8), 943–948. https://doi.org/10.1016/j.foodres.2005.03.011

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., & Tu, X. M. (2014). Log-transformation and its implications for data analysis. *Shanghai Archives of Psychiatry*, 26(2), 105–109. https://doi.org/10.3969/j.issn.1002-0829.2014.02.009

Gegas, V. C., Nazari, A., Griffiths, S., Simmonds, J., Fish, L., Orford, S., Sayers, L., Doonan, J. H., & Snape, J. W. (2010). A genetic framework for grain size and shape variation in wheat. *The Plant Cell*, *22*(4), 1046–1056. https://doi.org/10.1105/tpc.110.074153

Ginzberg, I., Minz, D., Faingold, I., Soriano, S., Mints, M., Fogelman, E., Warshavsky, S., Zig, U., & Yermiyahu, U. (2012). Calcium mitigated potato skin physiological disorder. *American Journal of Potato Research*, *89*(5), 351–362. https://doi.org/10.1007/s12230-012-9249-0

Grandke, F., Ranganathan, S., van Bers, N., de Haan, J. R., & Metzler, D. (2017). PERGOLA: Fast and deterministic linkage mapping of polyploids. *BMC Bioinformatics*, *18*(1), 12. https://doi.org/10.1186/s12859-016-1416-8

Hackett, C. A., Boskamp, B., Vogogias, A., Preedy, K. F., & Milne, I. (2017). TetraploidSNPMap: Software for linkage analysis and QTL mapping in autotetraploid populations using SNP dosage data. *Journal of Heredity*, *108*(4), 438–442. https://doi.org/10.1093/jhered/esx022

Hamilton, J. P., Hansey, C. N., Whitty, B. R., Stoffel, K., Massa, A. N., Van Deynze, A., De Jong, W. S., Douches, D. S., & Buell, C. R. (2011). Single nucleotide polymorphism discovery in elite North American potato germplasm. *BMC Genomics*, *12*(1), 302. https://doi.org/10.1186/1471-2164-12-302

Hara-Skrzypiec, A., Śliwka, J., Jakuczun, H., & Zimnoch-Guzowska, E. (2018). QTL for tuber morphology traits in diploid potato. *Journal of Applied Genetics*, *59*(2), 123–132. https://doi.org/10.1007/s13353-018-0433-x

Hill, W. G., & Robertson, A. (1968). Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics*, *38*(6), 226–231. https://doi.org/10.1007/BF01245622

Hong, E. P., & Park, J. W. (2012). Sample size and statistical power calculation in genetic association studies. *Genomics & Informatics*, *10*(2), 117–122. https://doi.org/10.5808/GI.2012.10.2.117

Jung, C. S., Griffiths, H. M., De Jong, D. M., Cheng, S., Bodis, M., & De Jong, W. S. (2005). The potato *P* locus codes for flavonoid 3′,5′-hydroxylase. *Theoretical*

and Applied Genetics, *110*(2), 269–275. https://doi.org/10.1007/s00122-004-1829-z

Kaiser, N. R., Coombs, J. J., Felcher, K. J., Hammerschmidt, R., Zuehlke, M. L., Buell, C. R., & Douches, D. S. (2020). Genome-wide association analysis of common scab resistance and expression profiling of tubers in response to thaxtomin a treatment underscore the complexity of common scab resistance in tetraploid potato. *American Journal of Potato Research*, *97*(5), 513–522. https://doi.org/10.1007/s12230-020-09800-5

Klaassen, M. T., Willemsen, J. H., Vos, P. G., Visser, R. G. F., van Eck, H. J., Maliepaard, C., & Trindade, L. M. (2019). Genome-wide association analysis in tetraploid potato reveals four QTLs for protein content. *Molecular Breeding*, *39*(11), 151. https://doi.org/10.1007/s11032-019-1070-8

Kloosterman, B., Abelenda, J. A., Gomez, M. del M. C., Oortwijn, M., de Boer, J. M., Kowitwanich, K., Horvath, B. M., van Eck, H. J., Smaczniak, C., Prat, S., Visser, R. G. F., & Bachem, C. W. B. (2013). Naturally occurring allele diversity allows potato cultivation in northern latitudes. *Nature*, *495*(7440), 246–250. https://doi.org/10.1038/nature11912

Kloosterman, B., Oortwijn, M., uitdeWilligen, J., America, T., de Vos, R., Visser, R. G., & Bachem, C. W. (2010). From QTL to candidate gene: Genetical genomics of simple and complex traits in potato using a pooling strategy. *BMC Genomics*, *11*(1), 158. https://doi.org/10.1186/1471-2164-11-158

Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods*, *9*, 29. https://doi.org/10.1186/1746-4811-9-29

Li, X.-Q., De Jong, H., De Jong, D. M., & De Jong, W. S. (2005). Inheritance and genetic mapping of tuber eye depth in cultivated diploid potatoes. *Theoretical and Applied Genetics*, *110*(6), 1068–1073. https://doi.org/10.1007/s00122-005-1927-6

Love, S. L., Werner, B. K., & Pavek, J. J. (1997). Selection for individual traits in the early generations of a potato breeding program dedicated to producing cultivars with tubers having long shape and russet skin. *American Potato Journal*, *74*(3), 199–213. https://doi.org/10.1007/BF02851598

Lulai, E. C. (2007). Chapter 22—Skin-set, wound healing, and related defects. In D. Vreugdenhil, J. Bradshaw, C. Gebhardt, F. Govers, D. K. L. Mackerron, M. A. Taylor, & H. A. Ross (Eds.), *Potato Biology and Biotechnology* (pp. 471–500). Elsevier Science B.V. https://doi.org/10.1016/B978-044451018-1/50064-6

Mollinari, M., & Garcia, A. A. F. (2019). Linkage analysis and haplotype phasing in experimental autopolyploid populations with high ploidy level using hidden Markov models. *G3: Genes, Genomes, Genetics*, *9*(10), 3297–3314. https://doi.org/10.1534/g3.119.400378

Monforte, A. J., Diaz, A., Caño-Delgado, A., & van der Knaap, E. (2014). The genetic basis of fruit morphology in horticultural crops: Lessons from tomato and melon. *Journal of Experimental Botany*, *65*(16), 4625–4637. https://doi.org/10.1093/jxb/eru017

Ortiz, R. (2020). Genomic-led potato breeding for increasing genetic gains: Achievements and outlook. *Crop Breeding, Genetics and Genomics*, *2*(2). https://doi.org/10.20900/cbgg20200010

Ozimati, A., Kawuki, R., Esuma, W., Kayondo, S. I., Pariyo, A., Wolfe, M., & Jannink, J.-L. (2019). Genetic variation and trait correlations in an East African cassava breeding population for genomic selection. *Crop Science*, *59*(2), 460–473. https://doi.org/10.2135/cropsci2018.01.0060

Ozturk, G., & Yildirim, Z. (2014). Heritability estimates of some quantitative traits in potatoes. *Turkish Journal of Field Crops*, *19*(2), 262. https://doi.org/10.17557/tjfc.66538

Pandey, J., Scheuring, D. C., Koym, J. W., Coombs, J., Novy, R. G., Thompson, A. L., Holm, D. G., Douches, D. S., Miller, J. C., & Vales, M. I. (2021). Genetic diversity and population structure of advanced clones selected over forty years by a potato breeding program in the USA. *Scientific Reports*, *11*(1), 8344. https://doi.org/10.1038/s41598-021-87284-x

Parra-Galindo, M.-A., Piñeros-Niño, C., Soto-Sedano, J. C., & Mosquera-Vasquez, T. (2019). Chromosomes I and X harbor consistent genetic factors associated with the anthocyanin variation in potato. *Agronomy*, *9*(7), 366. https://doi.org/10.3390/agronomy9070366

Piepho, H. P., Büchse, A., & Emrich, K. (2003). A Hitchhiker's guide to mixed models for randomized experiments. *Journal of Agronomy and Crop Science*, 189(5), 310–322. https://doi.org/10.1046/j.1439-037X.2003.00049.x

Prashar, A., Hornyik, C., Young, V., McLean, K., Sharma, S. K., Dale, M. F. B., & Bryan, G. J. (2014). Construction of a dense SNP map of a highly heterozygous diploid potato population and QTL analysis of tuber shape and eye depth. *Theoretical and Applied Genetics*, *127*(10), 2159–2171. https://doi.org/10.1007/s00122-014-2369-9

Przybylak, A., Kozłowski, R., Osuch, E., Osuch, A., Rybacki, P., & Przygodziński, P. (2020). Quality evaluation of potato tubers using neural image analysis method. *Agriculture*, *10*(4), 112. https://doi.org/10.3390/agriculture10040112

Rosyara, U. R., De Jong, W. S., Douches, D. S., & Endelman, J. B. (2016). Software for genome-wide association studies in autopolyploids and its application to potato. *The Plant Genome*, *9*(2). https://doi.org/10.3835/plantgenome2015.08.0073

Sharma, S. K., MacKenzie, K., McLean, K., Dale, F., Daniels, S., & Bryan, G. J. (2018). Linkage disequilibrium and evaluation of genome-wide association mapping models in tetraploid potato. *G3: Genes*, *Genomes*, *Genetics*, *8*(10), 3185–3202. https://doi.org/10.1534/g3.118.200377

Si, Y., Sankaran, S., Knowles, N. R., & Pavek, M. J. (2017). Potato tuber length-width ratio assessment using image analysis. *American Journal of Potato Research*, *94*(1), 88–93. https://doi.org/10.1007/s12230-016-9545-1

Slater, A. T., Cogan, N. O. I., Hayes, B. J., Schultz, L., Dale, M. F. B., Bryan, G. J., & Forster, J. W. (2014). Improving breeding efficiency in potato using molecular and quantitative genetics. *Theoretical and Applied Genetics*, *127*(11), 2279–2292. https://doi.org/10.1007/s00122-014-2386-8

Slater, A. T., Wilson, G. M., Cogan, N. O. I., Forster, J. W., & Hayes, B. J. (2014). Improving the analysis of low heritability complex traits for enhanced genetic gain in potato. *Theoretical and Applied Genetics*, *127*(4), 809–820. https://doi.org/10.1007/s00122-013-2258-7

ŚLiwka, J., Wasilewicz-Flis, I., Jakuczun, H., & Gebhardt, C. (2008). Tagging quantitative trait loci for dormancy, tuber shape, regularity of tuber shape, eye depth and flesh colour in diploid potato originated from six Solanum species. *Plant Breeding*, *127*(1), 49–55. https://doi.org/10.1111/j.1439-0523.2008.01420.x

Sood, S., Bhardwaj, V., Kaushik, S. K., & Sharma, S. (2020). Prediction based on estimated breeding values using genealogy for tuber yield and late blight resistance in auto-tetraploid potato (*Solanum tuberosum* L.). *Heliyon*, *6*(11). https://doi.org/10.1016/j.heliyon.2020.e05624

Sood, S., Lin, Z., Caruana, B., Slater, A. T., & Daetwyler, H. D. (2020). Making the most of all data: Combining non-genotyped and genotyped potato individuals with HBLUP. *The Plant Genome*, *13*(3), e20056. https://doi.org/10.1002/tpg2.20056

Speed, D., Hemani, G., Johnson, M. R., & Balding, D. J. (2012). Improved heritability estimation from genome-wide SNPs. *The American Journal of Human Genetics*, *91*(6), 1011–1021. https://doi.org/10.1016/j.ajhg.2012.10.010

Spooner, D. M., Ghislain, M., Simon, R., Jansky, S. H., & Gavrilenko, T. (2014). Systematics, Diversity, Genetics, and Evolution of Wild and Cultivated Potatoes.

The Botanical Review, *80*(4), 283–383. https://doi.org/10.1007/s12229-014-9146-y

Spooner, D. M., McLean, K., Ramsay, G., Waugh, R., & Bryan, G. J. (2005). A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences*, *102*(41), 14694–14699. https://doi.org/10.1073/pnas.0507400102

Stark, J. C., Love, S. L., & Knowles, N. R. (2020). Tuber quality. In J. C. Stark, M. Thornton, & P. Nolte (Eds.), *Potato Production Systems* (pp. 479–497). Springer International Publishing. https://doi.org/10.1007/978-3-030-39157-7_15

Stich, B., Urbany, C., Hoffmann, P., & Gebhardt, C. (2013). Population structure and linkage disequilibrium in diploid and tetraploid potato revealed by genome-wide high-density genotyping using the SolCAP SNP array. *Plant Breeding*, *132*(6), 718–724. https://doi.org/10.1111/pbr.12102

Sun, L., Chen, J., Xiao, K., & Yang, W. (2017). Origin of the domesticated horticultural species and molecular bases of fruit shape and size changes during the domestication, taking tomato as an example. *Horticultural Plant Journal*, *3*(3), 125–132. https://doi.org/10.1016/j.hpj.2017.07.007

Thiele, G., Dufour, D., Vernier, P., Mwanga, R. O. M., Parker, M. L., Geldermann, E. S., Teeken, B., Wossen, T., Gotor, E., Kikulwe, E., Tufan, H., Sinelle, S., Kouakou, A. M., Friedmann, M., Polar, V., & Hershey, C. (2021). A review of varietal change in roots, tubers and bananas: Consumer preferences and other

drivers of adoption and implications for breeding. *International Journal of Food Science & Technology*, *56*(3), 1076–1092. https://doi.org/10.1111/ijfs.14684

Thorup, T. A., Tanyolac, B., Livingstone, K. D., Popovsky, S., Paran, I., & Jahn, M. (2000). Candidate gene analysis of organ pigmentation loci in the Solanaceae. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(21), 11192–11197. https://doi.org/10.1073/pnas.97.21.11192

van Eck, H. J., Jacobs, J. E., Ton, J., Stiekemat, W. J., & Jacobsen, E. (1994). Multiple alleles for tuber shape in diploid potato detected by qualitative and quantitative genetic analysis using RFLPs. *Genetics*, *137*(1), 303–309.

van Eck, H. J., Jacobs, J. M. E., van den Berg, P. M. M. M., Stiekema, W. J., & Jacobsen, E. (1994). The inheritance of anthocyanin pigmentation in potato (*Solanum tuberosum* L.) and mapping of tuber skin colour loci using RFLPs. *Heredity*, *73*(4), 410–421. https://doi.org/10.1038/hdy.1994.189

Vos, Peter. G., Uitdewilligen, J. G. A. M. L., Voorrips, R. E., Visser, R. G. F., & van Eck, H. J. (2015). Development and analysis of a 20K SNP array for potato (*Solanum tuberosum*): An insight into the breeding history. *Theoretical and Applied Genetics*, *128*(12), 2387–2401. https://doi.org/10.1007/s00122-015-2593-y

Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., Wang, J., Orjeda, G., Guzman, F., Torres, M., Lozano, R., Ponce, O., Martinez, D., De la Cruz, G., Chakrabarti, S. K., Patil, V. U. (2011). Genome sequence and

analysis of the tuber crop potato. *Nature*, *475*(7355), 189–195. https://doi.org/10.1038/nature10158

Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., & Price, A. L. (2014). Advantages and pitfalls in the application of mixed-model association methods. *Nature Genetics*, *46*(2), 100–106. https://doi.org/10.1038/ng.2876

Yousaf, M. F., Demirel, U., Naeem, M., & Çalışkan, M. E. (2021). Association mapping reveals novel genomic regions controlling some root and stolon traits in tetraploid potato (*Solanum tuberosum* L.). *3 Biotech*, *11*(4), 174. https://doi.org/10.1007/s13205-021-02727-6

Zhu, C., Gore, M., Buckler, E. S., & Yu, J. (2008). Status and prospects of association mapping in plants. *The Plant Genome*, *1*(1), 5–20. https://doi.org/10.3835/plantgenome2008.02.0089

Zia, M. A. B., Demirel, U., Nadeem, M. A., & Çaliskan, M. E. (2020). Genome-wide association study identifies various loci underlying agronomic and morphological traits in diversified potato panel. *Physiology and Molecular Biology of Plants*, *26*(5), 1003–1020. https://doi.org/10.1007/s12298-020-00785-3

4. GENOMIC PREDICTION OF CHIPPING QUALITY IN TETRAPLOID POTATO

4.1. Introduction

Potato (*Solanum tuberosum* L.) is the most essential non-grain food on the planet and a highly recommended crop in ensuring food security for future generations (FAO, 2009; Devaux et al., 2014). Potatoes together with rice, wheat, and corn account for half of the world's food energy requirements (FAO, 2021). As such, the potato is economically importance for growers, processors, packers, and retailers (Ortiz & Mares, 2017; Wijesinha-Bettoni & Mouillé, 2019). In the U.S., potato consumption was around 22.4 kg per capita in 2019 (USDA ERS, 2021). Potato outnumbers the next most common vegetable in the U.S., tomato, by nearly a factor of two in terms of consumption.

Potatoes are used for fresh consumption, processing, and as seed. Among these, processed potato products have evolved as the most important. As such, most potatoes grown in the United States (60%) are intended for the processing industry. Frozen, dehydrated, chips and other packaged foods and snacks made from processed potatoes are among the best-selling items worldwide (Scott & Kleinwechter, 2017; Scott et al., 2019; Tul'cheev et al., 2020). Potato chips are the most consumed snack universally due to changes in consumer's lifestyles and food consumption (Liyanage et al., 2021). The global potato chip market reached a value of US $31.2 billion in 2020, and the market value is expected to rise even more in the coming years, reaching US$43.2 billion by 2026 (Statista, 2021).

Potato breeding programs focus on developing varieties well-suited to specific market classes including fresh market, specialty, and processing varieties. Important quality traits for processing potatoes include appearance, post-harvest quality, nutritional value, flavor properties, and productivity (Nacheva & Pevicharova, 2008). In terms of tuber appearance, the industry prefers shallow eyes and a round-oval shape for processing chips, whereas long-oval-shaped tubers are favored for processing French fries. Greening, cracking, hollow tubers, secondary damage, and rusty spots are undesirable traits. Most processing potatoes have a white flesh. Smooth skin is sought for making chips, whereas russet skin is preferred for making French fries.

The starch, dry matter, reducing sugars, proteins, and vitamin C contents of tubers determine their post-harvest quality and nutritional value. Potato tubers contain, on average, 80% water and 20% dry matter, with starch accounting for most of the dry matter (Stark et al., 2020). However, these composition numbers vary due to both genotype and environment (Zhou et al., 2017). Starch is the most important component of dry matter and determines tuber density, also referred to as tuber-specific gravity. For processing, high starch and dry matter-rich cultivars are favored because they cook faster, have a better texture, higher net weight of the final product, and use less oil while frying (Habyarimana et al., 2017). Starch content of 13% or more, solids or dry matter content of 20% or more, and specific gravity of 1.080 or more are preferred for most processing potatoes (Stark et al., 2020).

Another important component in potato that determines the quality of processed potato is the reducing sugar content (Wiberley-Bradford et al., 2014). During frying,

reducing sugars (glucose and fructose) in potato tubers react with free amino acids in a nonenzymatic Maillard reaction, resulting in undesirable brown, bitter-tasting products (Kumar et al., 2004; Wen et al., 2016). Although heat treatment contributes to the aroma, flavor, and color of the fried potatoes, the acrylamide products (when the amino acid asparagine is involved) of the Maillard reaction have raised health concerns due to their potential toxicity and carcinogenicity (Hogervorst et al., 2010). For these reasons, reducing sugar content in processing potatoes should not exceed 0.2-3.0% of the fresh weight (Fiselier & Grob, 2005).

Internal defects like internal heat necrosis (IHN), vascular browning, Fusarium wilt, stem end browning, hollow heart with discoloration, brown center, and internal black spot also contribute to browning and reduce chip quality. Consequently, the potato processing industry continues the search for strategies to reduce browning. These processes are complex, and the reducing sugar accumulation in tubers varies with variety, storage temperature, physiological maturity, plant stresses, and tuber age (Leonel et al., 2017). At the industrial level, cold storage conditions facilitate long-term storage and eliminate issues related to diseases, sprouting, and shrinking (Kleinkopf et al., 2003). However, cold temperatures promote the breakdown of starch and the accumulation of reducing sugar in potato tubers, a process known as cold-induced sweetening (CIS). As a result, developing potato cultivars that can be stored at low temperatures without undergoing CIS would be extremely beneficial.

Breeding programs have made efforts to develop cultivars that generate less sugar in cold storage, resulting in a lighter color potato chip. Improvements to date are

associated with traits from the Lenape variety which has been a common parent in chip crosses since its introduction (Love et al., 1998). Lenape is a parent of popular chipping varieties including Atlantic, Trent, Belchip and Snowden. Atlantic (Webb et al., 1978), has high yield and high specific gravity (1.085-1.100) but it is sensitive to environmental stress. Its major processing weakness is the accumulation of reducing sugars in long-term storage. Likewise, Snowden, selected in the late 1970s in Wisconsin, is like Atlantic except that it chips out of cold (10 °C) storage without reconditioning. However, it produces undersized tubers which results in lower yields. It is also susceptible to early blight, late blight, and common scab. Lamoka, selected at Cornell University, is a cold induced sweetening resistant variety with good chip color and resistance to the Golden Cyst Nematode (De Jong et al., 2017). However, Lamoka is susceptible to internal tuber defects, and moderately susceptible to early blight, late blight, powdery scab, potato virus X and Y. Therefore, it is necessary to breed a new potato cultivars with high chip quality, low acrylamide precursor contents, CIS resistance, high specific gravity, and high dry matter, and free of diseases and tuber defects.

Reasons that have hindered current potato breeding progress include the complex nature of the potato genome (polyploid with polysomic inheritance and highly heterozygous), quantitative nature of the most important traits, rapid inbreeding depression, and low intensity of selection in early generations (Ghislain & Douches, 2020). The current selective breeding process takes a long time (10–15 years) to produce a new marketable potato cultivar following the initial cross (Halterman et al., 2016).

Low multiplication rate delays multilocation field evaluations, and the contamination of seed tubers with pathogens and the need to clean them further delays evaluations. Another essential aspect that hinders progress in breeding is the slow introduction of advanced clones as parents to make new crosses. Speeding up the use of advanced parents with high breeding value for the traits of interest would shorten the breeding cycle and thus make recurrent selection breeding more effective.

The recent availability of dense single-nucleotide polymorphism (SNP) array and software packages, which can perform linkage and association mapping in autotetraploid species, presents opportunities to apply these optimized approaches to cultivated potato. An important step in the process of developing new chip varieties that produce light chip color is to identify the underlying genetic loci. Quantitative trait loci (QTL) mapping for chip color and sugar content have identified several regions in the potato genome which contain genes influencing the trait (Li et al., 2008). Several of these QTLs were co-localized with functional genes for carbohydrate metabolism and transport, including vacuolar invertase (*Pain-1/VInv*) on chromosome 3, apoplastic invertases (*Invap-a* and *Invap-b*) on chromosomes 10 and 9, and sucrose synthases (*Sus3* and *Sus4*) on chromosomes 7 and 12 (Gebhardt et al., 2014). Recently, Byrne et al. (2020) identified a major QTL on chromosome 10 for fry color and predicted fry color with moderate accuracy using genome-wide markers. Frederick & Bethke (2019) identified nine quantitative trait loci (QTL) for stem-end chip defect (SECD) seven of which overlapped with QTLs for chip color traits. Other efforts to practically apply the SNP panel have focused on Genome-Wide Association Studies (GWAS) to identify marker-trait

associations. The SolCAP SNP array was used for Genome-Wide Association Studies (GWAS) to identify marker-trait associations for tuber yield and starch content (Schönhals et al., 2016), tuber quality and fry color traits (D'hoop et al., 2014), and potato chip color and tuber traits (Rak et al., 2017).

Furthermore, important progress has been made in recent years for rapid genetic gain in crops using pedigree or genomic information-based methods such as genomic selection (GS) (Crossa et al., 2017). GS uses genome-wide molecular markers and historical phenotype data from a breeding program to predict the performance of related individuals. The genomic estimated breeding values (GEBVs) allow selections of superior parents or better individuals for next-generation advancement to be made more quickly than using phenotypes alone (Meuwissen et al., 2001; Crossa et al., 2017). Thus, GS enhances genetic gains by shortening the breeding cycle and/or enhancing testing efficiency. Habyarimana et al. (2017) demonstrated three GS models using SilicoDArT markers and indicated that GS can predict early clonal generations and can select traits with low heritability. Endelman et al. (2018), combining genotype and pedigree information with phenotype data for economically important traits, concluded that genome-wide prediction is feasible in autotetraploid potato. Likewise, genotyping-by-sequencing developed genomic prediction models were used for starch content and chipping quality (Sverrisdóttir et al., 2017), and for dry matter content and chipping quality (Sverrisdóttir et al., 2018). The polyBreedR package (https://github.com/jendelman/polyBreedR) has been recently designed to facilitate the use of genome-wide markers for autotetraploid (4x) species. It follows a two-stage

procedure (Damesa et al., 2017) where information on mean estimates and the associated variance-covariance matrix is forwarded from the first stage to the second stage to predict the breeding values.

Further, the usage of a selection index can aid in the selection of multiple traits at the same time. Using a multiple-trait selection index, it is possible to identify superior and inferior genotypes by integrating several attributes (Bernardo, 2020). The best index to use depends on the crop and the traits that are important for selection. The sum of standardized variables (∑Z) or Z-index (Mendes et al., 2009) is an alternative to other indices and has been used in common bean (Lima et al., 2015), upland rice (Ribeiro et al., 2016), and quince cultivars (Coutinho et al., 2019) to select superior lines.

The aim of this study was i) to evaluate how well GS could predict chipping quality in tetraploid potatoes, and ii) to generate a selection index for selection of superior individuals for important chipping traits together.

4.2. Materials and Methods

4.2.1. Plant materials and phenotype data

Five hundred and forty-nine unique chipping potato clones were evaluated between 2017 and 2020 near Dalhart Texas (35°58'15.31"N, 102°44'36.33"W) (Figure 4.1). These included entries from the National Chip Processing Trial (NCPT) and advanced chip selections from the Western, Southwestern Regional, and Texas trials contributed by public potato breeding programs in the United States. The NCPT uses a two-tier evaluation system, with one plot per location for tier 1 clones and two plots (replications) per location for tier 2 clones. Advanced chip selections from the Western,

Southwestern Regional, and Texas trials were planted in one to four replications. Plots in the NCPT trial contained 15 seed pieces, and advanced chip selections from the Western, Southwestern Regional and Texas trials contained 24 seed pieces. Seeds were planted with 30 cm spacing between hills and 70 cm spacing between rows. Trials were planted in early May and harvested in early September, with vine desiccation 2–3 weeks before harvest. Standard potato production practices were followed during the growing period in all years (https://potato.tamu.edu/reports/).

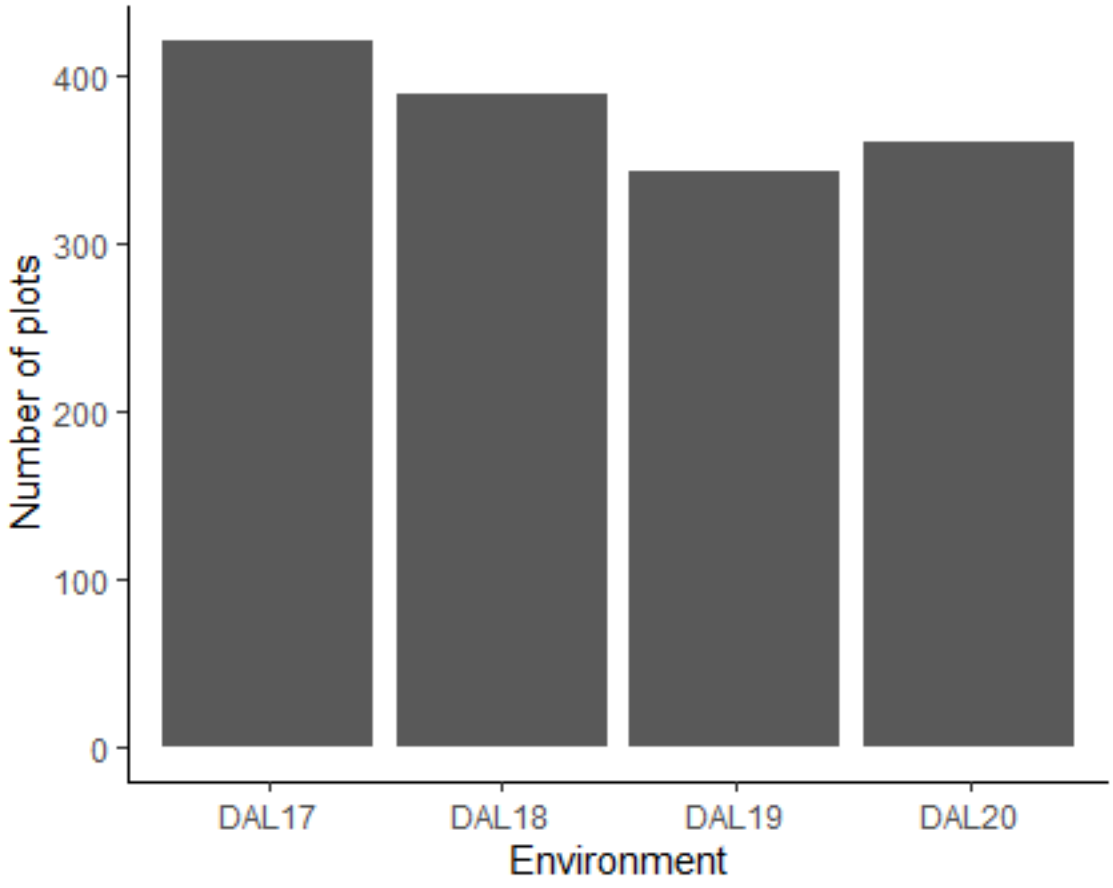


Figure 4.1 Number of chipping clone plots tested between 2017 and 2020 in Dalhart, TX.

Phenotype data for total yield, specific gravity, chip color, and chip quality were included in this study. Total yield was based on the weight of all harvested tubers (Mg ha-^{1}). Specific gravity was determined by water displacement, using 2–3 kg of tubers per plot (Wang et al., 2017). To evaluate fry color, ten tubers from each plot were cut into two halves (stem to bud), and a 1.3 mm-thick chip slice was prepared from each tuber. Chip slices were then rinsed in water and blot-dried on a paper towel to remove extra starch. Slices were fried in vegetable oil at 182 °C for 80 seconds or until bubbling ceased. The color of the chip (chips without internal defects) was scored using the Snack Food Association (SFA) scale (1 to 5; 1 = light, 5 = dark). Chip quality was scored using the SFA scale but focusing on the overall level of browning (including browning due to chip defects). The phenotypic dataset was analyzed using polyBreedR (https://github.com/jendelman/polyBreedR), considering partially replicated designs with incomplete blocks. The phenotypic correlations between pairs of traits were calculated as simple Pearson correlations.

4.2.2. Genotyping

Most of the clones in this dataset were genotyped with the Infinium 22 K V3 Potato Array on the Illumina iScan (Illumina Inc., San Diego, CA, USA), but some were genotyped with V2. V3 of the array contained 126 clones genotyped with 6,966 markers, and the V2 array consisted of 832 clones genotyped with 9,975 markers. The merge_impute function in the polyBreedR was used to merge the two datasets, and missing values were imputed by best linear unbiased prediction (BLUP). The imputation

process generated fractional values of dosage compatible with an estimation of the G matrix for predicting additive values using polyBreedR function G_mat.

4.2.3. Two-stage analysis of multi-environment trials

The two-stage procedure, as described by Damesa et al. (2017), was implemented in the polyBreedR package using ASReml-R (R 3.6.3 software), where information on mean estimates and the associated variance-covariance matrix were forwarded from the first stage to the second stage to predict the breeding values. In stage 1, the data for each environment x trait combination was analyzed independently with a linear mixed model. In the model, blocking was a fixed effect and factorial covariate. The plot-based broad-sense heritability for each environment x trait combination and the best linear unbiased estimates (BLUEs) for individual x trait combination were obtained. In stage 2, the function Stage2_prep removed the clones without marker data from stage 1 and returned a variance-covariance matrix called Omega, which must be saved as a variable in the global environment for ASReml-R. When used in conjunction with a random effect with covariance proportional to G, the stage 2 function provided an automatic random genetic effect with an identity (I) variance-covariance matrix, which describes non-additive effects. Genomic narrow-sense heritability, defined as the proportion of variance for the additive effects were estimated.

4.2.4. Genomic prediction by BLUP (gBLUP)

To predict the breeding values, the polyBreedR function predict_MME was used to implement mixed model equations (MME). Inversion of the MME coefficient matrix enabled the computation of the reliability (r^2) of the BLUPs, which is the expected

squared correlation between the true and predicted values. Reliability is an alternative to cross-validation (CV) for assessing genomic prediction accuracy. To estimate accuracy by cross-validation, the correlation between the predicted and observed values for the cross-validation group was divided by the square root of the narrow-sense h^2.

4.2.5. Standardized multiple selection index

The standardized multiple selection index was calculated based primarily on the standardization of the BLUPs for all clones, using the following estimator:

$z_i = (\bar{y}_i - \bar{y})/\sigma$

where z_i is the standardized average, $\bar{y}_i$ is the average of line i, $\bar{y}$ is the overall average, and σ is the standard deviation.

The index was calculated considering chip quality, specific gravity, yield, and chip color according to the following expression:

MIS= (-3) * $Z_{Chip\ quality}$ + (2) * $Z_{Specific\ gravity}$ + (2) * Z_{Yield} + (-1) * $Z_{Chip\ color}$

where MIS is the multiple index selection, $Z_{Chip\ quality}$ is the standardized average for chip quality, $Z_{Specific\ gravity}$ is the standardized average for specific gravity, Z_{Yield} is the standardized average for yield and $Z_{Chip\ color}$ is the standardized average for chip color.

The subjective economic weightage for each trait Z score was provided in the parenthesis. The highest weightage of value 3 was provided to chip quality as it is the most relevant trait to the industry. Chip quality trait includes both chip color and defects score. The weightage of value 2 was provided for specific gravity as it is directly correlated to starch content and amount of oil used during frying. The weightage of

value 2 was assigned for yield because it is the measure of profitability for growers. Chips made from some of the yellow-fleshed cultivars may be of equal preference to those prepared from white-fleshed potatoes. Therefore, lowest weightage of value 1 was given for chip color. Negative signs were used to flip the scales so high z values indicate better score. For example, negative weight was used for chip quality because the original chip quality score was on a 1-5 scale were 1 was the best. To be able to interpret the MIS in the context of z score distribution, MIS was also converted into a z value (Z_{MIS}). The top 5% clones of the z distribution were identified as potential parents or candidates to be advanced because they had the best combined breeding values.

4.3. Results

4.3.1. Phenotype and genotype data

Typically, in breeding trials, many clones are tested in one year and then dropped. However, the repetitions across years should be enough to estimate genotype x year interactions. We found that 81 clones tested in DAL17 were also tested in other years. Likewise, 114 clones in DAL18, 106 clones in DAL19, and 84 clones in DAL20 were common across environments. After merging and imputing V2, V3 and TAMU genotyped chipping clones, 14,401 markers were retained. The 384 clones which had both genotypic and phenotypic data were used for further analysis. Fractional values of dosage were obtained by the imputation process (Figure 4.2). These values are compatible with an estimation of the G matrix for predicting additive values, using polyBreedR function G_mat. The typical, unimodal, zero-mean distribution of G matrix coefficients for a homogeneous population is shown in Figure 4.3. The broad-sense

heritability for each env x trait combination is shown in Figure 4.4. For chip color, the highest broad-sense heritability of 0.77 was obtained in DAL19, and the lowest broad-sense heritability of 0.32 was obtained in DAL18. Likewise, broad-sense heritability of 0.59 was achieved for both chip quality and specific gravity in DAL19 and DAL18, respectively. Similarly, for yield, the broad sense heritability ranged from 0.43 in DAL17 to 0.61 in DAL20. Pearson correlation (r) analysis was done for all the traits to identify the relationship between them. A significant correlation was found between chip color and chip quality ($r = 0.39$; $P < 0.001$), and between specific gravity and yield ($r = 0.29$; $P < 0.01$) (Figure 4.5).

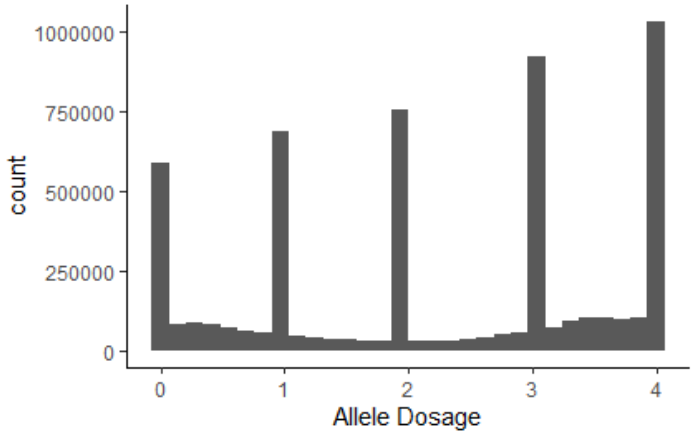


Figure 4.2 Fractional values of dosage obtained from 384 clones with the imputation of 14,401 markers compatible with an estimation of the G matrix.

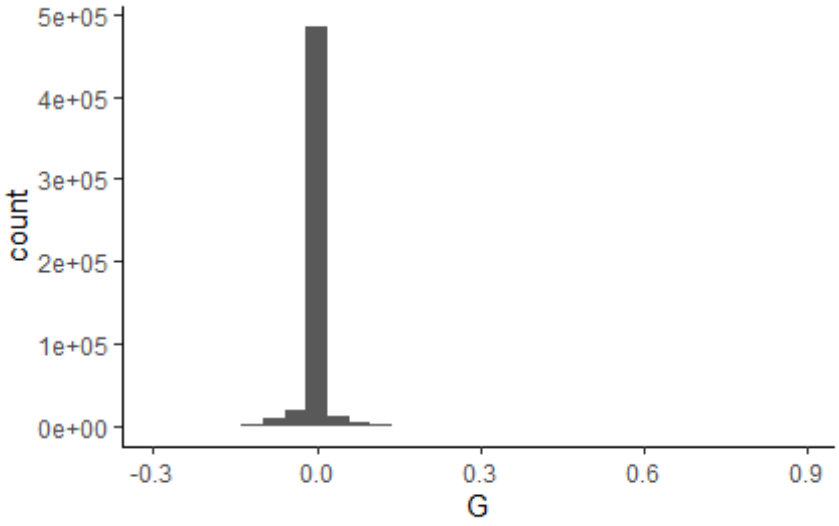


Figure 4.3 Distribution of G matrix coefficients obtained using 384 clones and 14,401 markers.

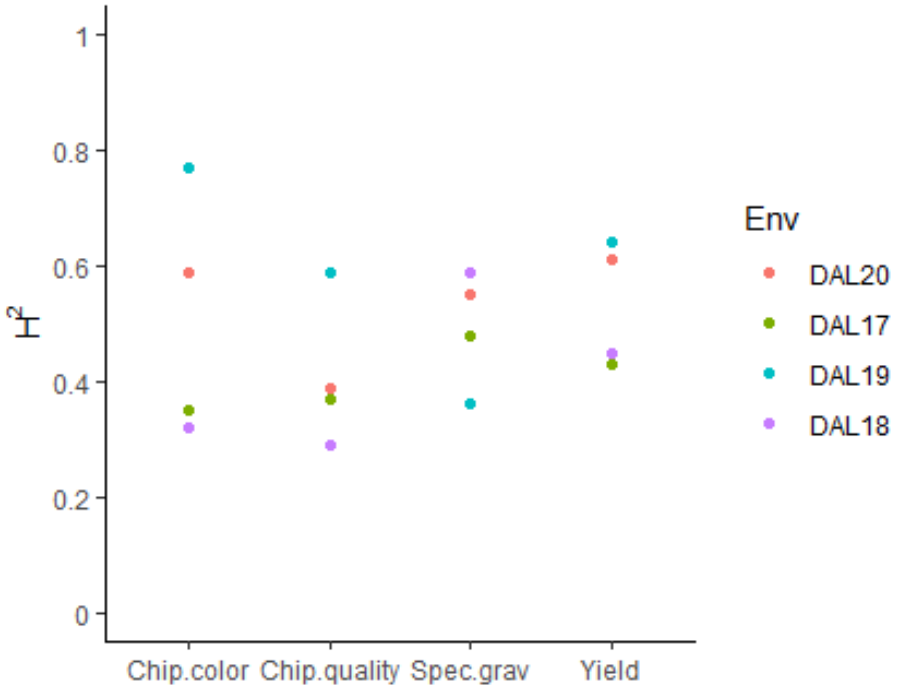


Figure 4.4 Plot based H2 using 384 unique clones for the Dalhart environment from 2017 to 2020

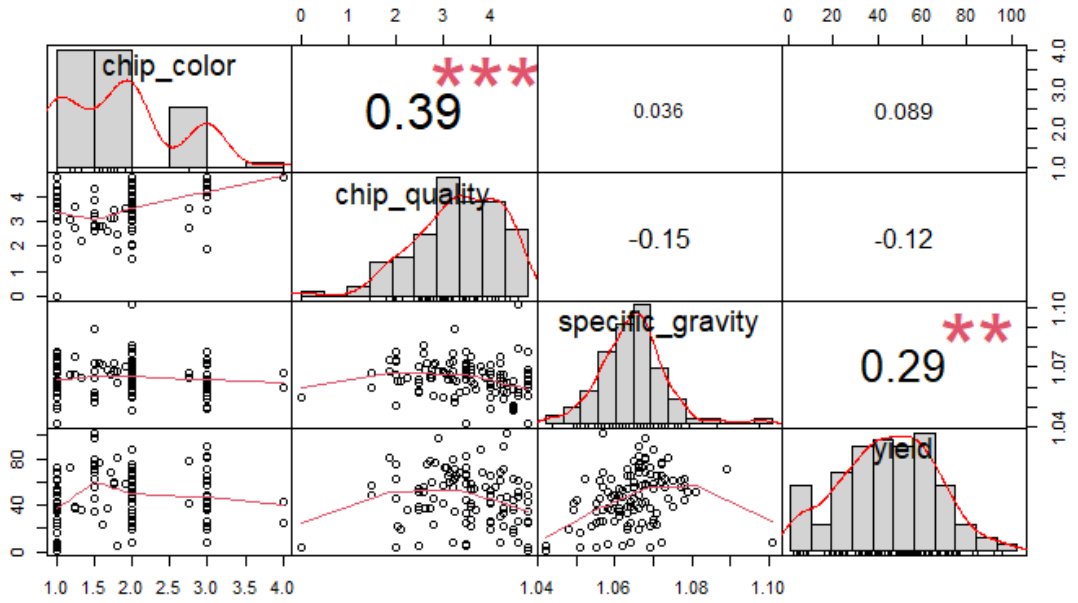


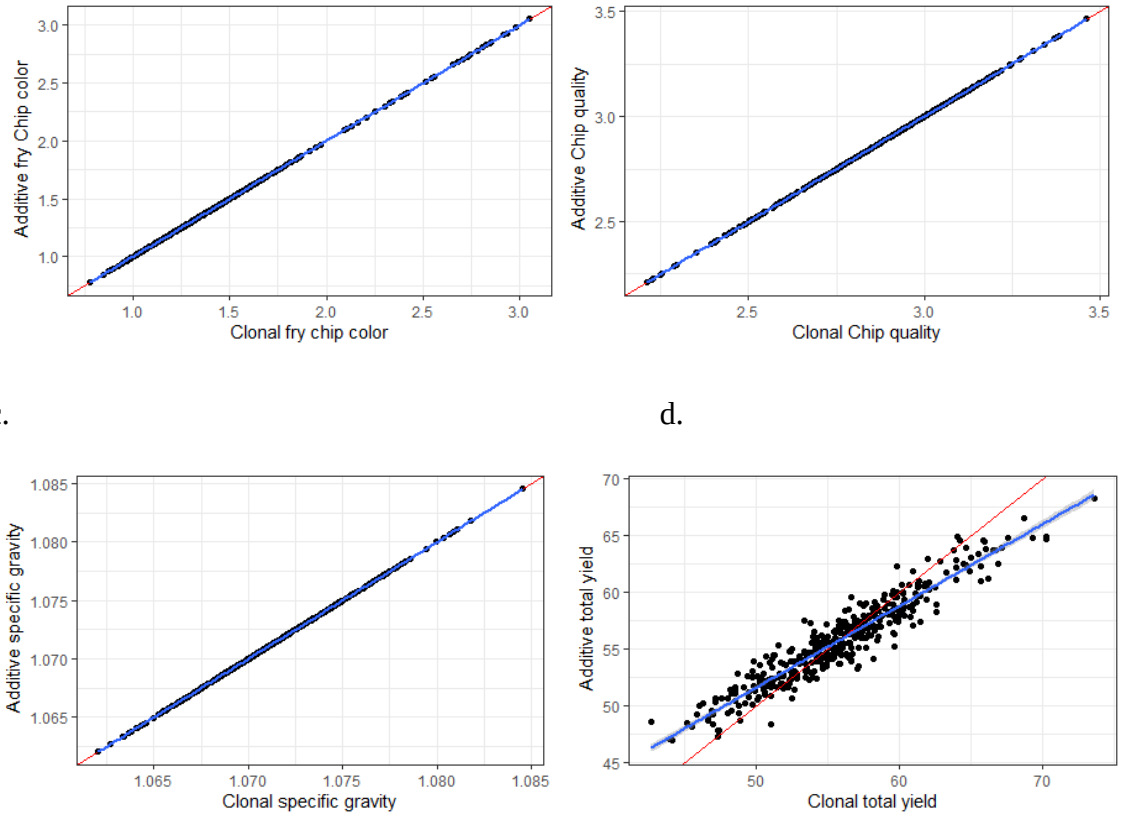
Figure 4.5 Bivariate scatter plots with a fitted line for the Dalhart chipping traits are displayed on the bottom of the diagonal. On the top of the diagonal are the value of the correlation plus the significance level as stars. Each significance level is associated to a symbol: p < 0.001 = * and p < 0.01 = **. The distribution of each chipping trait is shown on the diagonal.**

4.3.2. Genomic prediction by BLUP (gBLUP)

Breeding values for fried chip color, chip quality, specific gravity, and total yield were predicted using the mixed model equations (MME) (Appendix C1). The top-performing clone for fry chip color was W14NYQ9-2 with BLUP 0.80 and a reliability score of 0.69. COTX12235-2W was predicted to have the best chip quality with BLUP 2.13 and a reliability score of 0.63. ATTX10333-1W/Y was predicted to be superior for specific gravity with BLUP 1.083 and a reliability score of 0.70. NYQ29-1 was predicted to be superior for total yield with BLUP 68.5 and a reliability score of 0.4. The reliability range plotted against the number of close relationships for each clone,

quantified by the 95th percentile of its G coefficients (G95), is shown in Figure 4.7. The individuals that had a lot of close relatives were predicted with higher reliability. The number of phenotype measurements (i.e., plots) for each clone also contributed to the accuracy (Figure 4.7)

a. b.



.

c. d.

Figure 4.6 Comparison of the predicted additive (y-axis) and clonal values (x-axis) for fry chip color (a) chip quality (b), specific gravity (c), and total yield (d). The slope of the regression line (blue) represented the additive proportion of the clonal value in comparison with the line with a slope equal to 1 (red).

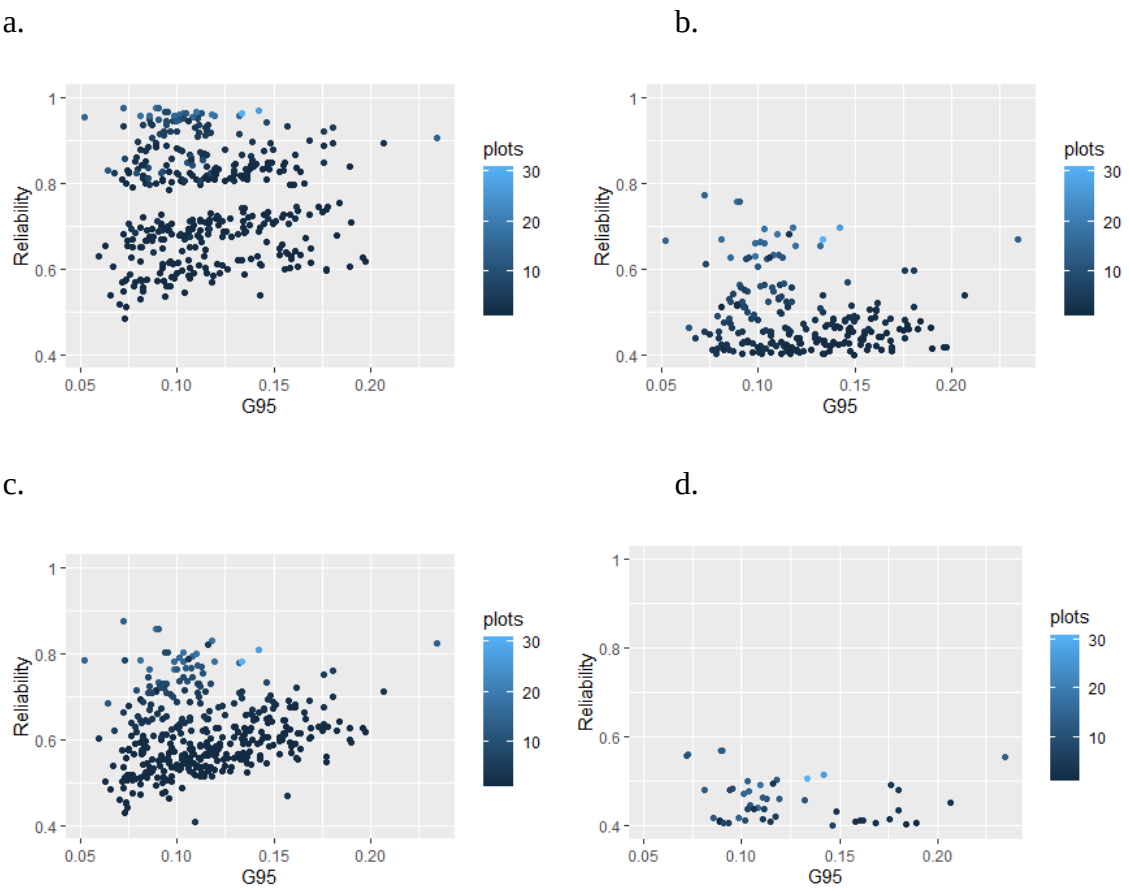


Figure 4.7 Expected squared correlation between the true and predicted values known as reliability (y-axis) plotted against the number of close relationships for each clone, quantified by the 95th percentile of its G coefficients (G95) (x-axis) for fry chip color (a), chip quality (b), specific gravity (c), and total yield (d).

4.3.3. Cross-validation

To illustrate cross-validation, we masked phenotypes for the clones in the 2020 Texas preliminary yield trial. Figure 4.8 compares the reliability of the predicted additive values with (y-axis) and without (x-axis) phenotypes of the DAL20 group (black line = slope of 1). The reliability was substantially higher for the DAL20 clones when their phenotype data are included. For the rest of the population, there was a slight

advantage to including the DAL20 phenotypes. The accuracy by cross-validation was estimated to be 0.62.

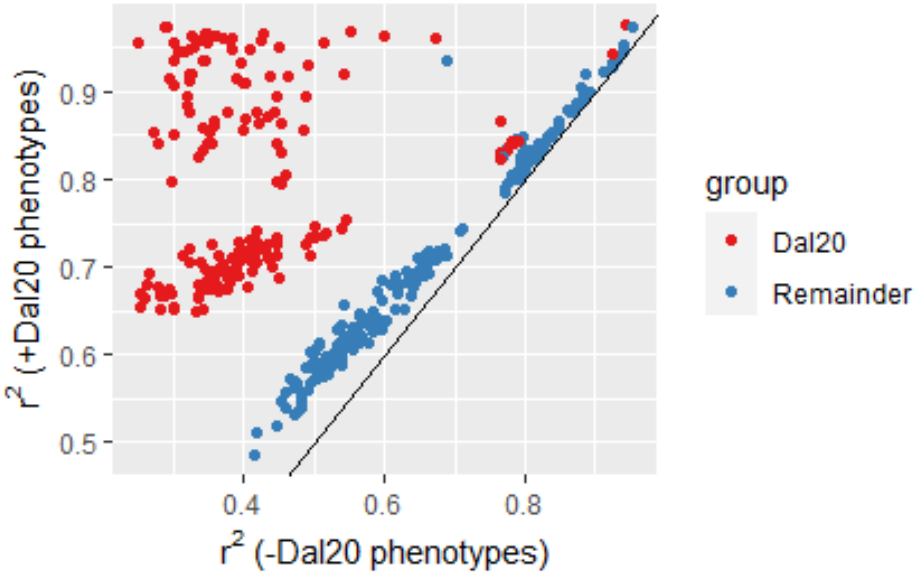


Figure 4.8 Comparison of the reliability of the predicted additive values with phenotypes for the DAL20 group (y-axis) and without (x-axis) phenotypes for the DAL20 group (black line = slope of 1)

4.3.4. Two-stage analysis of multi-environment trials

Genomic narrow-sense heritability, defined as the proportion of variance for the additive effects, was estimated at 0.52, 0.17, 0.40, and 0.11 for chip color, chip quality, specific gravity, and yield, respectively. Non-additive genetic effects and genotype x year interaction effects were negligible for chip color. However, the estimated genotype x year interaction effect for chip quality was 0.32. Likewise, for specific gravity the proportion of non-additive genetic effects and genotype x year interaction effects were 0.09 and 0.08, respectively. Similarly, for yield, the proportion of non-additive genetic effects and genotype x year interaction effects were 0.08 and 0.41 respectively. The

benefit of including Omega, which accounts for the micro and macro-environmental variation in each of the estimates in the analysis, was reflected in a lower Akaike information criterion (AIC) compared to a model without it ((https://github.com/jendelman/polyBreedR). The AIC was lowered by 24, 25, 56, and 19 for chip color, chip quality, specific gravity, and yield, respectively. Comparison of the predicted additive and clonal values showed that the slope of the regression line (blue) is less than 1 (shown in red) for specific gravity and yield (Figure 4.6c and 4.6d). However, the regression line for chip color and chip quality is almost equal to one, which implied that the proportion of non-additive genetic effects for these traits were negligible (Figure 4.6a and 4.6b).

4.3.5. Standardized multiple selection index

Selection indexes are important in the processes of breeding and selecting superior chipping cultivars. The results for the standardized multiple selection index (Appendix C2) allowed the identification of chipping clones with high chip quality, yield, specific gravity, and light-colored chips. It can be observed that clones with standardized multiple selection index (z) > 2 are in the top 5%. Most of the clones within the top 5% (NYR102-3, NYR102-7, NY169, NYN24-2, NYR102-8, NYP116-6, NYP108-6) were bred at Cornell University. Three clones (NDTX1246-3W, NDTX113030C-3W, NDTX1246-5W/Y) selected by the Texas potato program were also found to be among the best chipping clones.

4.4. Discussion

Genomic selection can accelerate breeding for a range of traits, including those with complex genetic regulation. To increase GS accuracy, the breeding program's historical data can be effectively used (Atanda et al., 2021). Proper biometric procedures are needed to combine all this historical information while accounting for micro- and macro-environmental variation. The polyBreedR package uses a fully efficient, two-stage procedure (Damesa et al., 2017) that requires some entries to be replicated in each environment. In this study, using the chipping quality data of tetraploid potato, the genotype estimate for each clone and environment combination was determined in stage one by assuming independent effects, and genomic covariance matrices were used in stage two.

GS can be incorporated into potato breeding, particularly in early clonal generations, to predict and select for all traits, but traits difficult or expensive to measure and those with low heritability could benefit the most from GS. GS could overcome the need for evaluating many plants per clone over several locations and years. The heritability indicates that the environment plays a large role for fry chip color. It ranged from 0.32 to 0.77 in different env x trait combinations. Trait heritability is affected by changes in allelic frequencies, introduction of new alleles (Latta, 2010), or changes in genetic influence caused by altered genetic backgrounds or environmental factors (Chandler et al., 2017) or error. Broad-sense heritability of 0.59 for specific gravity was observed in our study. However, Slater et al. (2014) estimated the heritability of specific

gravity to be 0.74. Temperature fluctuations between years affects specific gravity and the Dalhart location is prone to year-to-year environmental variation.

The partial replication over years allowed for explicit modeling of a residual genetic effect in addition to the additive, dominance, and epistatic random effects. We found that more of the genetic variances were additive for fry chip color (52%), chip quality (17%), and specific gravity (40%). Specific gravity is an important component of chipping quality. Endelman et al. (2018) reported add itive genetic variance was 20% for specific gravity. Based on diallel studies, Tai (1976) reported significant specific combining ability (SCA) /general combining ability (GCA) of 0.60 for specific gravity. Lynch et al. (1992) found the high GCA effects and non-significant SCA effects for specific gravity and suggested that mainly additive genetic factors account for the genetic variation. Wang et al. (2017) also showed relatively little genotype × environment interaction for the specific gravity. Thus, specific gravity as a proxy for dry matter content can be used for selecting chip processing potatoes.

Breeding values can now be predicted by regressing phenotypic values on all available markers, thanks to the advent of higher-density single-nucleotide polymorphisms (SNPs) spanning the entire genome of many plants (Crossa et al., 2017). The option of molecular marker platform has a big impact on GS applicability. Our study is based on the SNP arrays. SNP chips provide high-quality SNPs, but the cost per sample is significantly higher. It can be costly for small breeding programs since genomic selection often necessitates extensive population genotyping. On the other hand, genotyping by sequencing (GBS) offers many markers, but there can be a lot of

missing data (Elbasyoni et al., 2018). In potato, Endelman et al. (2018) used the SNP array, whereas Sverrisdóttir et al. (2017, 2018) and Byrne et al. (2020) utilized GBS to identify SNPs for genomic prediction models. Marker density, a major concern for using GS, has been studied in some crops like wheat (Crossa et al., 2014; Elbasyoni et al., 2018) and pea (Tayeh et al., 2015) but has yielded mixed results. Since not all SNP sites influence traits, selecting the most effective markers for genomic prediction will likely improve prediction accuracy.

The reliability of genomic predictions is the expected squared correlation between the true and predicted values and indicates the proportion of the genetic variance that is explained (de Roos et al., 2009). Reliable prediction of complex quality traits will be essential to attain the objective of successful chip potato breeding. Reliability (r^2) of the BLUPs is an alternative to cross-validation for assessing genomic prediction accuracy. In this study, we aimed to predict chipping quality with relatively high accuracy. Predictions were made for chipping quality traits like fry color, chip quality, and specific gravity using the mixed model equations. The mean reliability of the BLUPs obtained were 0.75, 0.43, 0.61 and 0.33 for chip color, chip quality, specific gravity, and total yield, respectively. For chip color, the reliability ranged from 0.49-0.97. Sverrisdóttir et al. (2018) had obtained cross-validated prediction correlations of 0.39–0.79 for chipping quality. Similarly, the reliability for the chip quality, specific gravity and total yield ranged from 0.25-0.77, 0.38-0.87 and 0.13-0.58, respectively. Since these predictions utilized phenotypic data for the selected candidates, the number of phenotype measurements (i.e., plots) for each clone also contributed to accuracy

(Figure 4.7). The GS strategy's superiority is determined not only by the ability to predict an individual's genetic merit with greater precision but also by the ability to shorten the breeding period and maximize genetic gains (Heffner et al., 2011). Breeders hope to pass useful traits from parental lines to their progeny by choosing parental lines with the highest GEBVs. Taken together, these factors make genomic selection in potatoes thoroughly feasible and attractive.

The use of an index is a useful approach for considering more than one trait in the selection process. This strategy has been applied for selection of early soybean inbred lines by Gesteira et al. (2018). In potato, Terres et al. (2015) estimated genetic gains using different selection index in potato populations. However, no reports were found on the use of this tool in the selection of chipping cultivars for chipping quality, yield, specific gravity, and chip color. The index was efficient in selecting cultivars for multiple traits simultaneously. The New York clones stood out in the evaluation. The index can be used to select superior clones to use as parents and ultimately to enhance the efficiency of selection in breeding programs.

4.5. Conclusions

Genomic selection (GS) leverages historical phenotype data from a breeding program to predict the genetic value of related individuals more accurately. Data from a subset of highly maintained trials can be used to predict success in an untested environment for chipping quality, which is labor-intensive and costly to phenotype. Even for limited reference populations and traits with low heritability, the accuracies found were motivating. For phenotyped individuals, GS can be used to improve the accuracy of

selection. The accuracy of the predictions will allow for a gradual transition to a full-scale genomic selection program. The selection index obtained will make it possible to select superior clones, considering multiple traits together.

4.6. References

Atanda, S. A., Olsen, M., Burgueño, J., Crossa, J., Dzidzienyo, D., Beyene, Y., Gowda, M., Dreher, K., Zhang, X., Prasanna, B. M., Tongoona, P., Danquah, E. Y., Olaoye, G., & Robbins, K. R. (2021). Maximizing efficiency of genomic selection in CIMMYT's tropical maize breeding program. *Theoretical and Applied Genetics*, *134*(1), 279–294. https://doi.org/10.1007/s00122-020-03696-9

Bernardo, R. N. (2020). *Breeding for quantitative traits in plants* (Third edition). Stemma Press.

Byrne, S., Meade, F., Mesiti, F., Griffin, D., Kennedy, C., & Milbourne, D. (2020). Genome-wide association and genomic prediction for fry color in potato. *Agronomy*, *10*(1), 90. https://doi.org/10.3390/agronomy10010090

Campos, H., & Ortiz, O. (Eds.). (2020). *The Potato Crop: Its Agricultural, Nutritional and Social Contribution to Humankind*. Springer International Publishing. https://doi.org/10.1007/978-3-030-28683-5

Chandler, C. H., Chari, S., Kowalski, A., Choi, L., Tack, D., DeNieu, M., Pitchers, W., Sonnenschein, A., Marvin, L., Hummel, K., Marier, C., Victory, A., Porter, C., Mammel, A., Holms, J., Sivaratnam, G., & Dworkin, I. (2017). How well do you know your mutation? Complex effects of genetic background on expressivity, complementation, and ordering of allelic effects. *PLOS Genetics*, *13*(11), e1007075. https://doi.org/10.1371/journal.pgen.1007075

Coutinho, G., Pio, R., Souza, F. B. M. de, Farias, D. da H., Bruzi, A. T., & Guimarães, P. H. S. (2019). Multivariate analysis and selection indices to identify superior

quince cultivars for cultivation in the tropics. *HortScience*, *54*(8), 1324–1329. https://doi.org/10.21273/HORTSCI14004-19

Crossa, J., Pérez, P., Hickey, J., Burgueño, J., Ornella, L., Cerón-Rojas, J., Zhang, X., Dreisigacker, S., Babu, R., Li, Y., Bonnett, D., & Mathews, K. (2014). Genomic prediction in CIMMYT maize and wheat breeding programs. *Heredity*, *112*(1), 48–60. https://doi.org/10.1038/hdy.2013.16

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., & Varshney, R. K. (2017). Genomic selection in plant breeding: Methods, models, and perspectives. *Trends in Plant Science*, *22*(11), 961–975. https://doi.org/10.1016/j.tplants.2017.08.011

Damesa, T. M., Möhring, J., Worku, M., & Piepho, H.-P. (2017). One step at a time: Stage-wise analysis of a series of experiments. *Agronomy Journal*, *109*(3), 845–857. https://doi.org/10.2134/agronj2016.07.0395

De Jong, W. S., Halseth, D. E., Plaisted, R. L., Wang, X., Perry, K. L., Qu, X., Paddock, K. M., Falise, M., Christ, B. J., & Porter, G. A. (2017). Lamoka, a Variety with Excellent Chip Color Out of Cold Storage and Resistance to the Golden Cyst Nematode. *American Journal of Potato Research*, *94*(2), 148–152. https://doi.org/10.1007/s12230-016-9557-x

de Roos, A. P. W., Hayes, B. J., & Goddard, M. E. (2009). Reliability of genomic predictions across multiple populations. *Genetics*, *183*(4), 1545–1553. https://doi.org/10.1534/genetics.109.104935

Devaux, A., Kromann, P., & Ortiz, O. (2014). Potatoes for sustainable global food security. *Potato Research*, *57*(3), 185–199. https://doi.org/10.1007/s11540-014-9265-1

D'hoop, B. B., Keizer, P. L. C., Paulo, M. J., Visser, R. G. F., van Eeuwijk, F. A., & van Eck, H. J. (2014). Identification of agronomically important QTL in tetraploid potato cultivars using a marker–trait association analysis. *Theoretical and Applied Genetics*, *127*(3), 731–748. https://doi.org/10.1007/s00122-013-2254-y

Elbasyoni, I. S., Lorenz, A. J., Guttieri, M., Frels, K., Baenziger, P. S., Poland, J., & Akhunov, E. (2018). A comparison between genotyping-by-sequencing and array-based scoring of SNPs for genomic prediction accuracy in winter wheat. *Plant Science*, *270*, 123–130. https://doi.org/10.1016/j.plantsci.2018.02.019

Endelman, J. B., Carley, C. A. S., Bethke, P. C., Coombs, J. J., Clough, M. E., Silva, W. L. da, Jong, W. S. D., Douches, D. S., Frederick, C. M., Haynes, K. G., Holm, D. G., Miller, J. C., Muñoz, P. R., Navarro, F. M., Novy, R. G., Palta, J. P., Porter, G. A., Rak, K. T., Sathuvalli, V. R., … Yencho, G. C. (2018). Genetic variance partitioning and genome-wide prediction with allele dosage information in autotetraploid potato. *Genetics*, *209*(1), 77–87. https://doi.org/10.1534/genetics.118.300685

FAO. (2009). *Why potato? - International year of the potato 2008, Rome*. FAO. http://www.fao.org/potato-2008/en/aboutiyp/index.html

FAO. (2021). *Dimensions of need—Staple foods: What do people eat?, Rome*. FAO. http://www.fao.org/3/u8480e/u8480e07.htm

Fiselier, K., & Grob, K. (2005). Legal limit for reducing sugars in prefabricates targeting 50 μg/kg acrylamide in French fries. *European Food Research and Technology*, *220*(5), 451–458. https://doi.org/10.1007/s00217-004-1081-4

Frederick, C. M., & Bethke, P. C. (2019). Identification of quantitative trait loci for stem-end chip defect and potato chip color traits in a 'Lenape'-derived full-sib population. *American Journal of Potato Research*, *96*(6), 564–577. https://doi.org/10.1007/s12230-019-09746-3

Gebhardt, C., Urbany, C., & Stich, B. (2014). Dissection of potato complex traits by linkage and association genetics as basis for developing molecular diagnostics in breeding programs. In R. Tuberosa, A. Graner, & E. Frison (Eds.), *Genomics of Plant Genetic Resources: Volume 2. Crop productivity, food security and nutritional quality* (pp. 47–85). Springer Netherlands. https://doi.org/10.1007/978-94-007-7575-6_3

Ghislain, M., & Douches, D. S. (2020). The genes and genomes of the potato. In H. Campos & O. Ortiz (Eds.), *The Potato Crop: Its Agricultural, Nutritional and Social Contribution to Humankind* (pp. 139–162). Springer International Publishing. https://doi.org/10.1007/978-3-030-28683-5_5

Habyarimana, E., Parisi, B., & Mandolino, G. (2017). Genomic prediction for yields, processing and nutritional quality traits in cultivated potato (*Solanum tuberosum* L.). *Plant Breeding*, *136*(2), 245–252. https://doi.org/10.1111/pbr.12461

Halterman, D., Guenthner, J., Collinge, S., Butler, N., & Douches, D. (2016). Biotech potatoes in the 21st century: 20 years since the first biotech potato. *American Journal of Potato Research*, *93*(1), 1–20. https://doi.org/10.1007/s12230-015-9485-1

Heffner, E. L., Jannink, J.-L., & Sorrells, M. E. (2011). Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *The Plant Genome*, *4*(1). https://doi.org/10.3835/plantgenome2010.12.0029

Hogervorst, J. G. F., Baars, B.-J., Schouten, L. J., Konings, E. J. M., Goldbohm, R. A., & Brandt, P. A. van den. (2010). The carcinogenicity of dietary acrylamide intake: A comparative discussion of epidemiological and experimental animal research. *Critical Reviews in Toxicology*, *40*(6), 485–512. https://doi.org/10.3109/10408440903524254

Kleinkopf, G. E., Oberg, N. A., & Olsen, N. L. (2003). Sprout inhibition in storage: Current status, new chemistries and natural compounds. *American Journal of Potato Research*, *80*(5), 317–327. https://doi.org/10.1007/BF02854316

Kumar, D., Singh, B. P., & Kumar, P. (2004). An overview of the factors affecting sugar content of potatoes. *Annals of Applied Biology*, *145*(3), 247–256. https://doi.org/10.1111/j.1744-7348.2004.tb00380.x

Latta, R. G. (Ed.). (2010). Natural selection, variation, adaptation, and evolution: A primer of interrelated concepts. *International Journal of Plant Sciences*, *171*(9), 930–944. https://doi.org/10.1086/656220

Leonel, M., do Carmo, E. L., Fernandes, A. M., Soratto, R. P., Ebúrneo, J. A. M., Garcia, É. L., & dos Santos, T. P. R. (2017). Chemical composition of potato tubers: The effect of cultivars and growth conditions. *Journal of Food Science and Technology*, *54*(8), 2372–2378. https://doi.org/10.1007/s13197-017-2677-6

Li, L., Paulo, M.-J., Strahwald, J., Lübeck, J., Hofferbert, H.-R., Tacke, E., Junghans, H., Wunder, J., Draffehn, A., van Eeuwijk, F., & Gebhardt, C. (2008). Natural DNA variation at candidate loci is associated with potato chip color, tuber starch content, yield and starch yield. *Theoretical and Applied Genetics*, *116*(8), 1167–1181. https://doi.org/10.1007/s00122-008-0746-y

Lima, D. C., Abreu, Â. de F. B., Ferreira, R. A. D. C., & Ramalho, M. A. P. (2015). Breeding common bean populations for traits using selection index. *Scientia Agricola*, *72*, 132–137. https://doi.org/10.1590/0103-9016-2014-0130

Liyanage, D. W. K., Yevtushenko, D. P., Konschuh, M., Bizimungu, B., & Lu, Z.-X. (2021). Processing strategies to decrease acrylamide formation, reducing sugars and free asparagine content in potato chips from three commercial cultivars. *Food Control*, *119*, 107452. https://doi.org/10.1016/j.foodcont.2020.107452

Love, S. L., Pavek, J. J., Thompson-Johns, A., & Bohl, W. (1998). Breeding progress for potato chip quality in North American cultivars. *American Journal of Potato Research*, *75*(1), 27–36. https://doi.org/10.1007/BF02883514

Lynch, D. R., Tai, G. C. C., & Coffin, R. H. (1992). Genetic components of potato chip quality evaluated in three environments and under various storage regimes. *Canadian Journal of Plant Science*, *72*(2), 535–543. https://doi.org/10.4141/cjps92-067

Mendes, F. F., Ramalho, M. A. P., & Abreu, Â. de F. B. (2009). Selection index for choosing segregating populations of common bean. *Pesquisa Agropecuária Brasileira*, *44*, 1312–1318. https://doi.org/10.1590/S0100-204X2009001000015

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Nacheva, E., & Pevicharova, G. (2008). Potato breeding lines for processing. *Genetics and Breeding*, *37*, 77–84.

Ortiz, O., & Mares, V. (2017). The historical, social, and economic importance of the potato crop. In S. Kumar Chakrabarti, C. Xie, & J. Kumar Tiwari (Eds.), *The Potato Genome* (pp. 1–10). Springer International Publishing. https://doi.org/10.1007/978-3-319-66135-3_1

Rak, K., Bethke, P. C., & Palta, J. P. (2017). QTL mapping of potato chip color and tuber traits within an autotetraploid family. *Molecular Breeding*, *37*(2), 15. https://doi.org/10.1007/s11032-017-0619-7

Ribeiro, D. C., de Lima, I. P., Mendes, M. P., & Oliveira, H. (2016). Selection for multiple traits in upland rice progenies using the z index. *International Journal of Current Research*, *08*, 5.

Schönhals, E. M., Ortega, F., Barandalla, L., Aragones, A., Ruiz de Galarreta, J. I., Liao, J.-C., Sanetomo, R., Walkemeier, B., Tacke, E., Ritter, E., & Gebhardt, C. (2016). Identification and reproducibility of diagnostic DNA markers for tuber starch and yield optimization in a novel association mapping population of potato (*Solanum tuberosum* L.). *Theoretical and Applied Genetics*, *129*, 767–785. https://doi.org/10.1007/s00122-016-2665-7

Scott, G. J., & Kleinwechter, U. (2017). Future scenarios for potato demand, supply and trade in South America to 2030. *Potato Research*, *60*(1), 23–45. https://doi.org/10.1007/s11540-017-9338-z

Scott, G. J., Petsakos, A., & Suarez, V. (2019). Not by bread alone: Estimating potato demand in india in 2030. *Potato Research*, *62*(3), 281–304. https://doi.org/10.1007/s11540-019-9411-x

Slater, A. T., Cogan, N. O. I., Hayes, B. J., Schultz, L., Dale, M. F. B., Bryan, G. J., & Forster, J. W. (2014). Improving breeding efficiency in potato using molecular and quantitative genetics. *Theoretical and Applied Genetics*, *127*(11), 2279–2292. https://doi.org/10.1007/s00122-014-2386-8

Stark, J. C., Love, S. L., & Knowles, N. R. (2020). Tuber quality. In J. C. Stark, M. Thornton, & P. Nolte (Eds.), *Potato Production Systems* (pp. 479–497). Springer International Publishing. https://doi.org/10.1007/978-3-030-39157-7_15

Statista. (2021). *Global market size: Packaged & unpackaged potato chips 2026*. Statista. https://www.statista.com/statistics/1199637/global-packaged-and-unpackaged-potato-chips-market-size/

Sverrisdóttir, E., Byrne, S., Sundmark, E. H. R., Johnsen, H. Ø., Kirk, H. G., Asp, T., Janss, L., & Nielsen, K. L. (2017). Genomic prediction of starch content and chipping quality in tetraploid potato using genotyping-by-sequencing. *Theoretical and Applied Genetics*, *130*(10), 2091–2108. https://doi.org/10.1007/s00122-017-2944-y

Sverrisdóttir, E., Sundmark, E. H. R., Johnsen, H. Ø., Kirk, H. G., Asp, T., Janss, L., Bryan, G., & Nielsen, K. L. (2018). The value of expanding the training population to improve genomic selection models in tetraploid potato. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.01118

Tayeh, N., Klein, A., Le Paslier, M.-C., Jacquin, F., Houtin, H., Rond, C., Chabert-Martinello, M., Magnin-Robert, J.-B., Marget, P., Aubert, G., & Burstin, J. (2015). Genomic prediction in pea: Effect of marker density and training population size and composition on prediction accuracy. *Frontiers in Plant Science*, *6*. https://doi.org/10.3389/fpls.2015.00941

Terres, L. R., Lenz, E., Castro, C. M., & Pereira, A. S. (2015). Genetic gain estimates using different selection index methods in potato hybrid populations. *Horticultura Brasileira*, *33*, 305–310. https://doi.org/10.1590/S0102-053620150000300005

Tul'cheev, V. V., Zhevora, S. V., Borisov, M. Yu., & Gordienko, N. N. (2020). Perspectives for development of Russian and global potato markets. *Studies on Russian Economic Development*, *31*(1), 85–89. https://doi.org/10.1134/S1075700720010177

USDA ERS. (2021). *USDA ERS - Food Availability and Consumption*. https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/food-availability-and-consumption/

Wang, Y., Snodgrass, L. B., Bethke, P. C., Bussan, A. J., Holm, D. G., Novy, R. G., Pavek, M. J., Porter, G. A., Rosen, C. J., Sathuvalli, V., Thompson, A. L., Thornton, M. T., & Endelman, J. B. (2017). Reliability of measurement and genotype × environment interaction for potato specific gravity. *Crop Science*, *57*(4), 1966–1972. https://doi.org/10.2135/cropsci2016.12.0976

Wayumba, B. O., Choi, H. S., & Seok, L. Y. (2019). Selection and evaluation of 21 potato (*Solanum tuberosum*) breeding clones for cold chip processing. *Foods*, *8*(3), 98. https://doi.org/10.3390/foods8030098

Webb, R. E., Wilson, D. R., Shumaker, J. R., Graves, B., Henninger, M. R., Watts, J., Frank, J. A., & Murphy, H. J. (1978). Atlantic: A new potato variety with high solids, good processing quality, and resistance to pests. *American Potato Journal*, *55*(3), 141–145. https://doi.org/10.1007/BF02852087

Wen, C., Shi, X., Wang, Z., Gao, W., Jiang, L., Xiao, Q., Liu, X., & Deng, F. (2016). Effects of metal ions on formation of acrylamide and 5-hydroxymethylfurfural in asparagine–glucose model system. *International Journal of Food Science & Technology*, *51*(2), 279–285. https://doi.org/10.1111/ijfs.12966

Wiberley-Bradford, A. E., Busse, J. S., Jiang, J., & Bethke, P. C. (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 Research Notes*, *7*(1), 801. https://doi.org/10.1186/1756-0500-7-801

Wijesinha-Bettoni, R., & Mouillé, B. (2019). The contribution of potatoes to global food security, nutrition and healthy diets. *American Journal of Potato Research*, *96*(2), 139–149. https://doi.org/10.1007/s12230-018-09697-1

Zhou, Z., Plauborg, F., Kristensen, K., & Andersen, M. N. (2017). Dry matter production, radiation interception and radiation use efficiency of potato in response to temperature and nitrogen application regimes. *Agricultural and Forest Meteorology*, *232*, 595–605. https://doi.org/10.1016/j.agrformet.2016.10.017

5. CONCLUSIONS

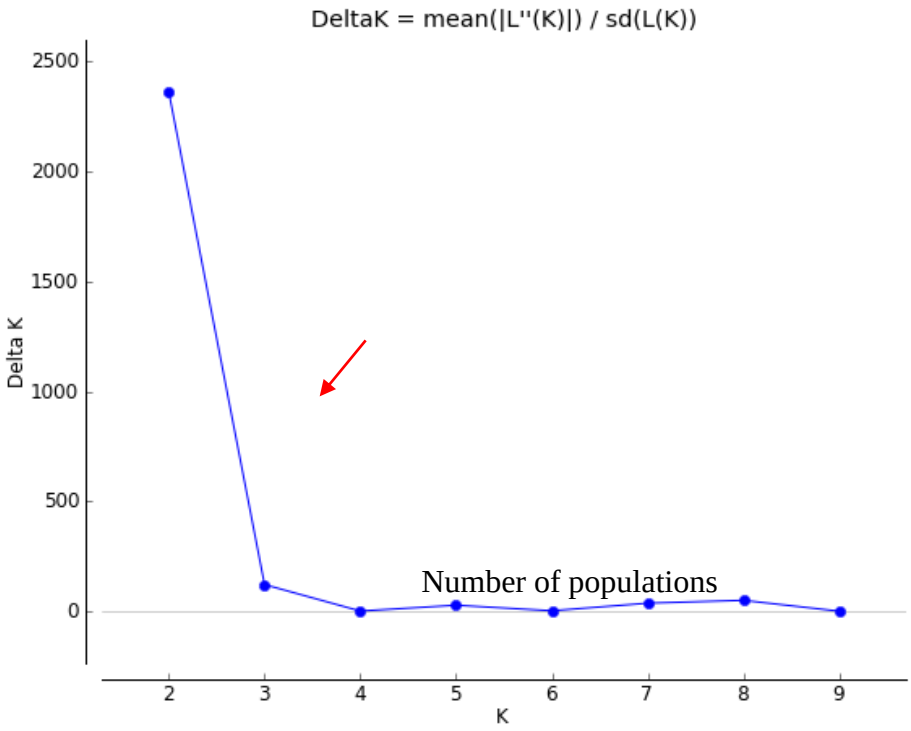
This research aimed to assess the genetic diversity, population structure, identify genetic variation, and predict the breeding values associated with important phenotypic and economic traits. Based on the assessment of genetic diversity in the advanced tetraploid clones selected over forty years by the Texas A&M potato program, it can be concluded that this collection is highly diverse and provides opportunity to develop new and improved varieties with desirable traits. The results indicate moderate to high levels of heterozygosity, low inbreeding, and significant population structure. By obtaining a core set, this research has shown how a small-scale breeding program can maintain a sub-collection that retains similar genetic diversity as the whole population for long-term conservation of genetic resources. This study was able to make the discovery of typographic errors and pedigree errors that occurred during handling clonal material in the breeding program and/or tissue culture operations. Studies such as this will help identify and correct errors in the breeding program. Also, SNP fingerprints and genetic distance comparisons from this study can be useful for plant variety protection (PVP), as well as for the verification of the identity of clones.

In most complicated quantitative traits, it appears that many small-effect alleles are responsible for controlling the observed variation. The results presented here constitute the effort to identify genetic variation associated with tuber shape, eye depth, degree of russeting, tuber number, tuber weight, skin color, and flesh color in advanced potato clones using genome wide association studies. Besides the rediscovery of a QTL for tuber morphology traits in potatoes, we introduce nonredundant markers that are

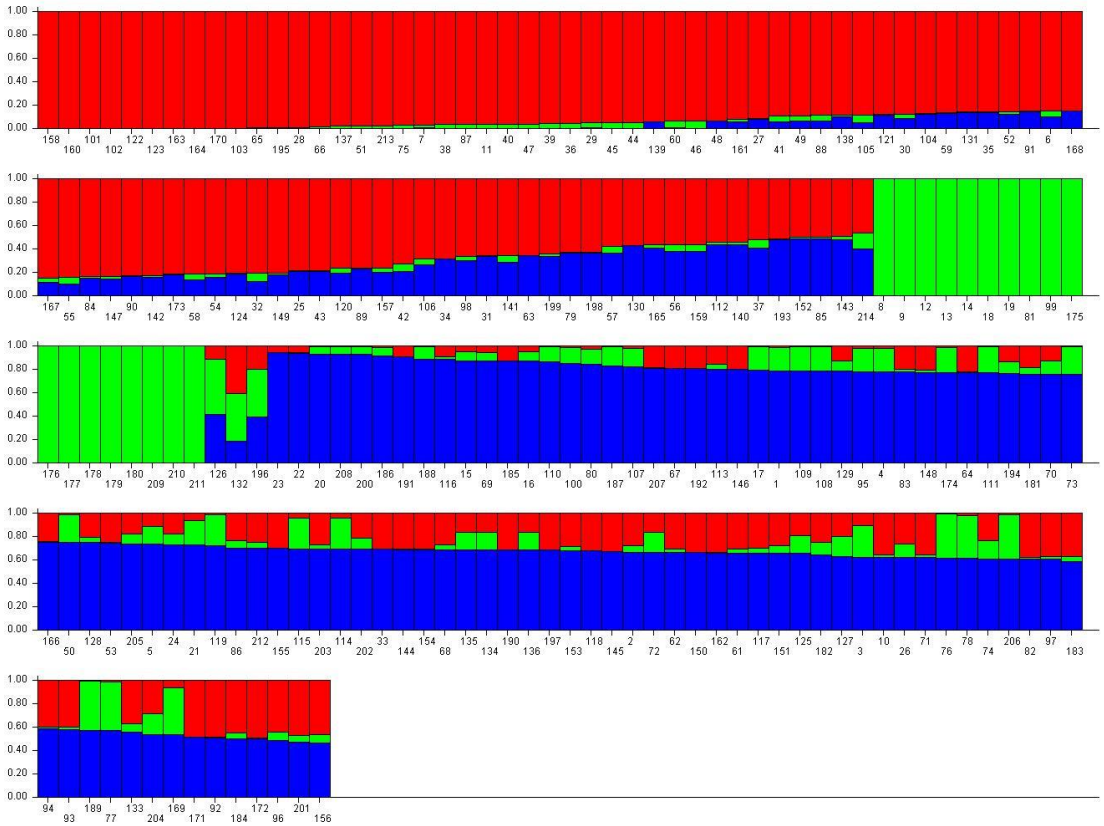
playing important roles in the control of these interesting traits. Since, GWAS require much larger sample sizes to acquire appropriate statistical power, we were not able to identify QTLs for some of the traits. Instead, in the future, phenotype of interest could be refined, or sample size and marker density could be increased. The findings of this dissertation research are likely to be valuable in the future design and implementation of genomic studies for marker-assisted selection.

This study shows that GS could be successfully implemented for chipping quality, which is labor-intensive and costly to phenotype. The accuracy is encouraging even for small reference populations and traits with low heritability. This study demonstrates that the index is a useful approach for considering more than one trait in the selection process. More research is needed to see how genetic prediction accuracy can be improved (expanding the training population, appropriate prediction model, higher marker density). The results of genomic diversity, GWAS and GS will be useful in improving breeding efficiency in the potato program.

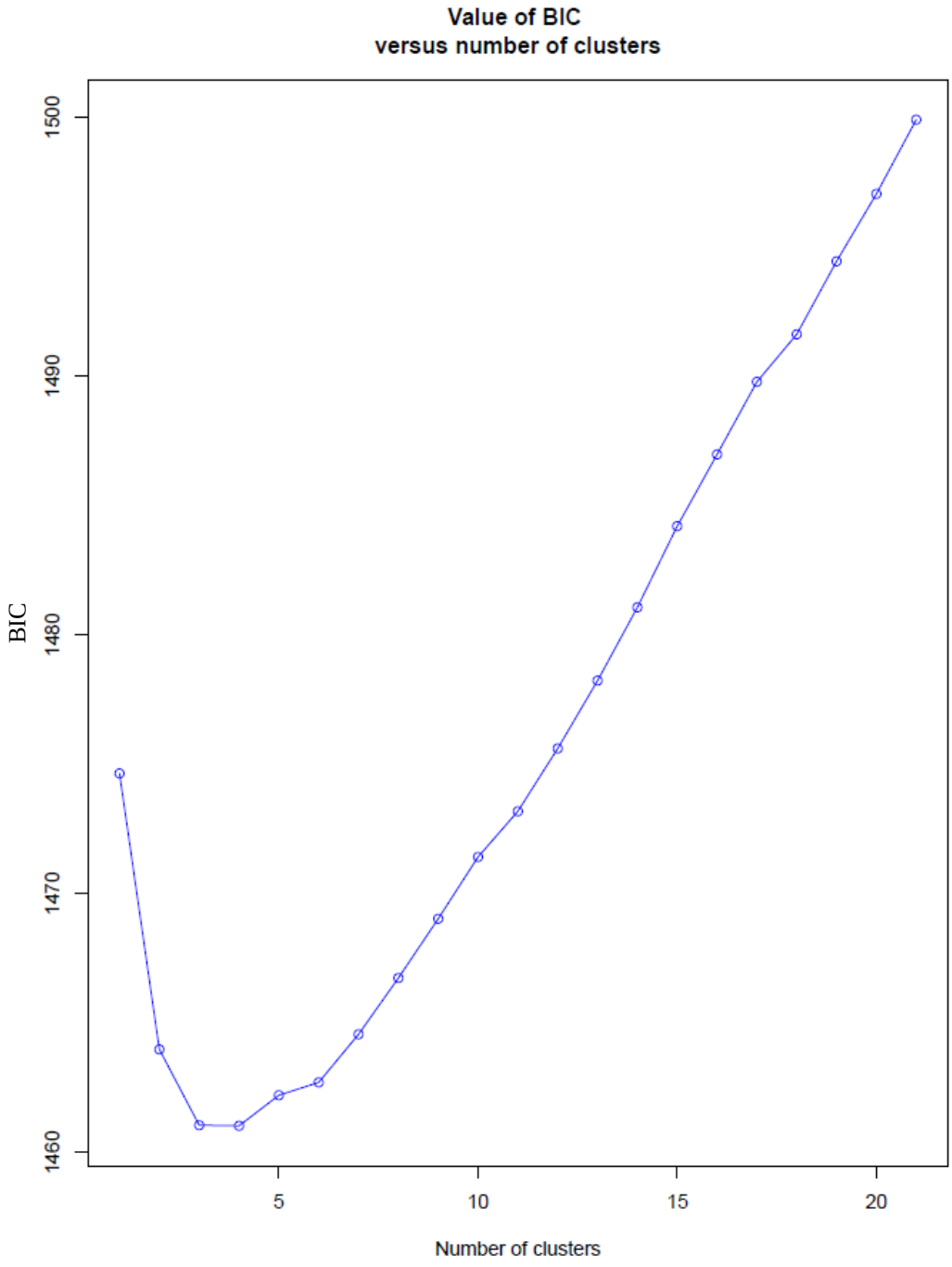
APPENDIX A



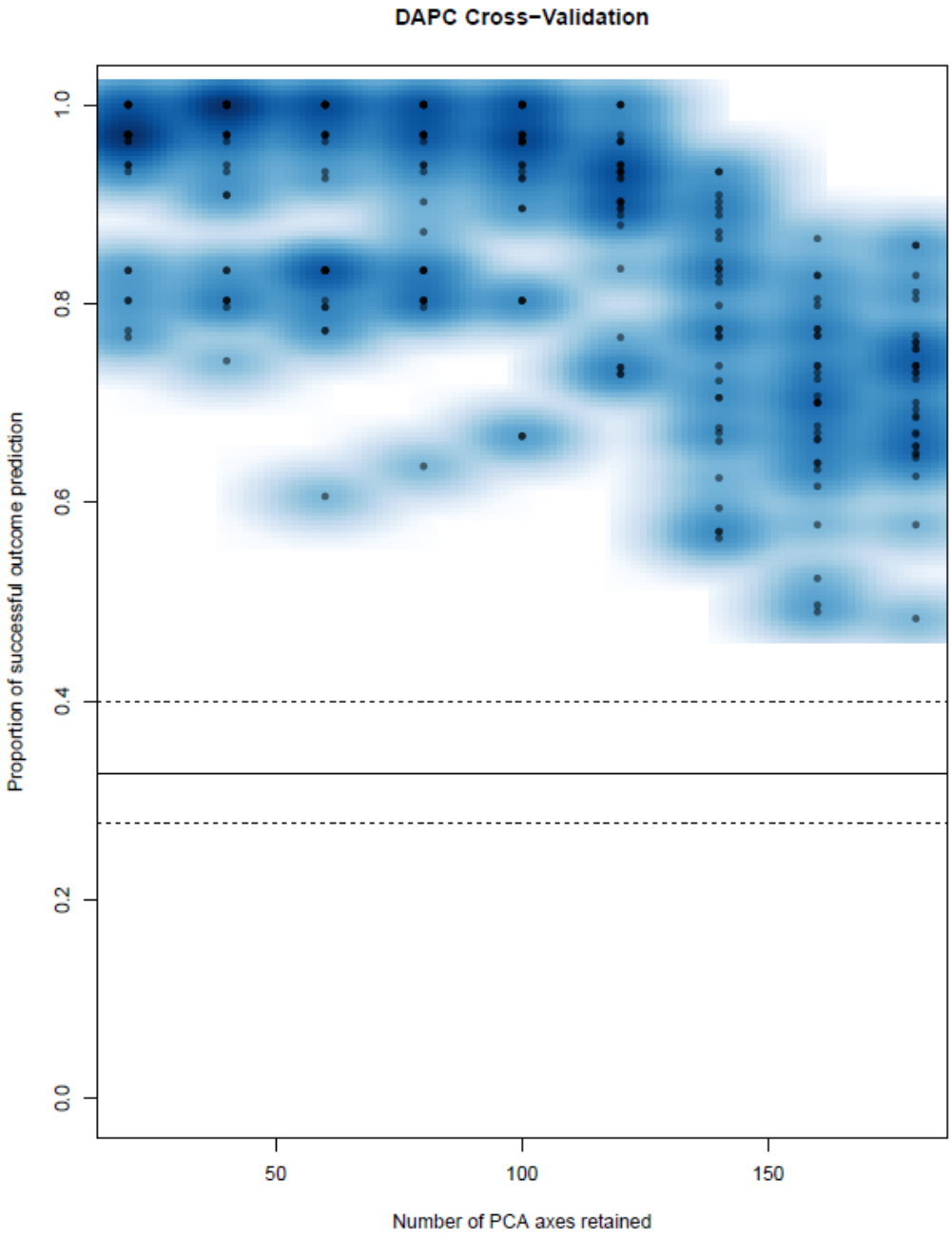
A1. Estimation of the number of populations using LnP(D) derived Δk for K from 1 to 10 using 10,106 SNPs. The maximum of adhoc measure ΔK determined by structure harvester was found to be K = 3. The red arrow in the graph indicates the clear inflection point.



A2. Model-based clustering using software STRUCTURE showing individual genotypes (x-axis) in the population structure of 214 potato clones (K = 3, red, green and blue) based on 10,106 SNPs. The y-axis indicates the subpopulation membership (relative scale 0-1). Individual potato genotypes are shown in the x-axis.



A3. The lowest Bayesian information criterion (BIC) value was obtained using *find.clusters* function using *adegenet* R package[67] for discriminant analysis of principal components (DAPC) analysis using 10K SNPs in the population of 214 potato clones.



A4. Cross-validation plot to guide the selection of the number of principal component axes (PCA) to retain a Discriminant Analysis of Principal Components Analysis using *adegenet* R package[67]. The PCA value that maximizes the proportion of successful outcomes and minimizes the mean square error (MSE) is 20.

a.

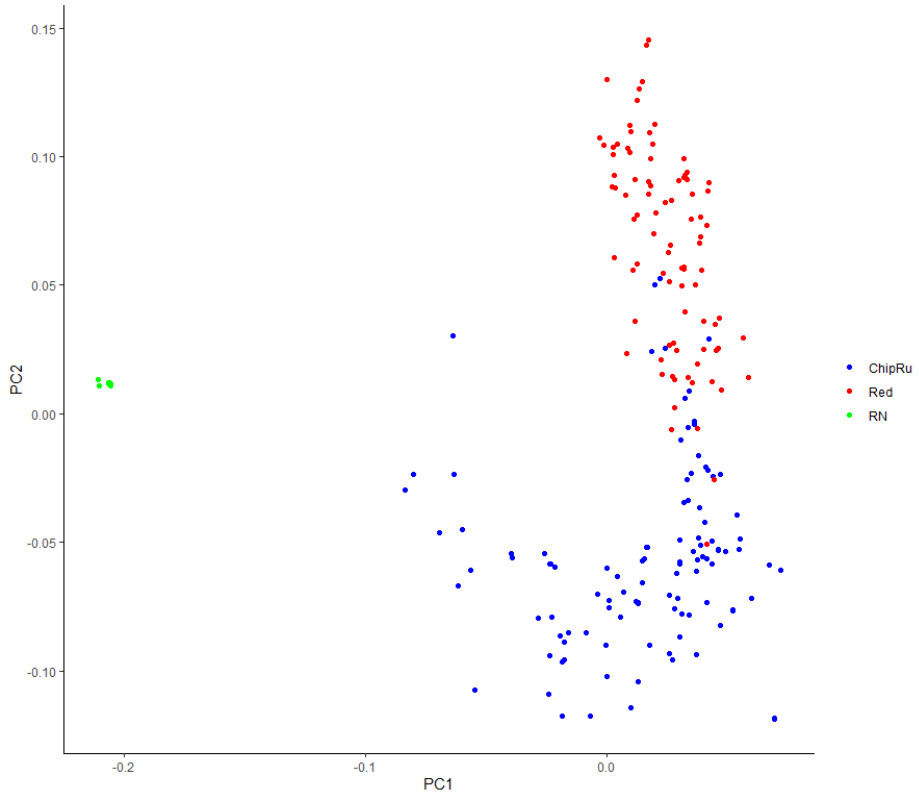
b.

c.

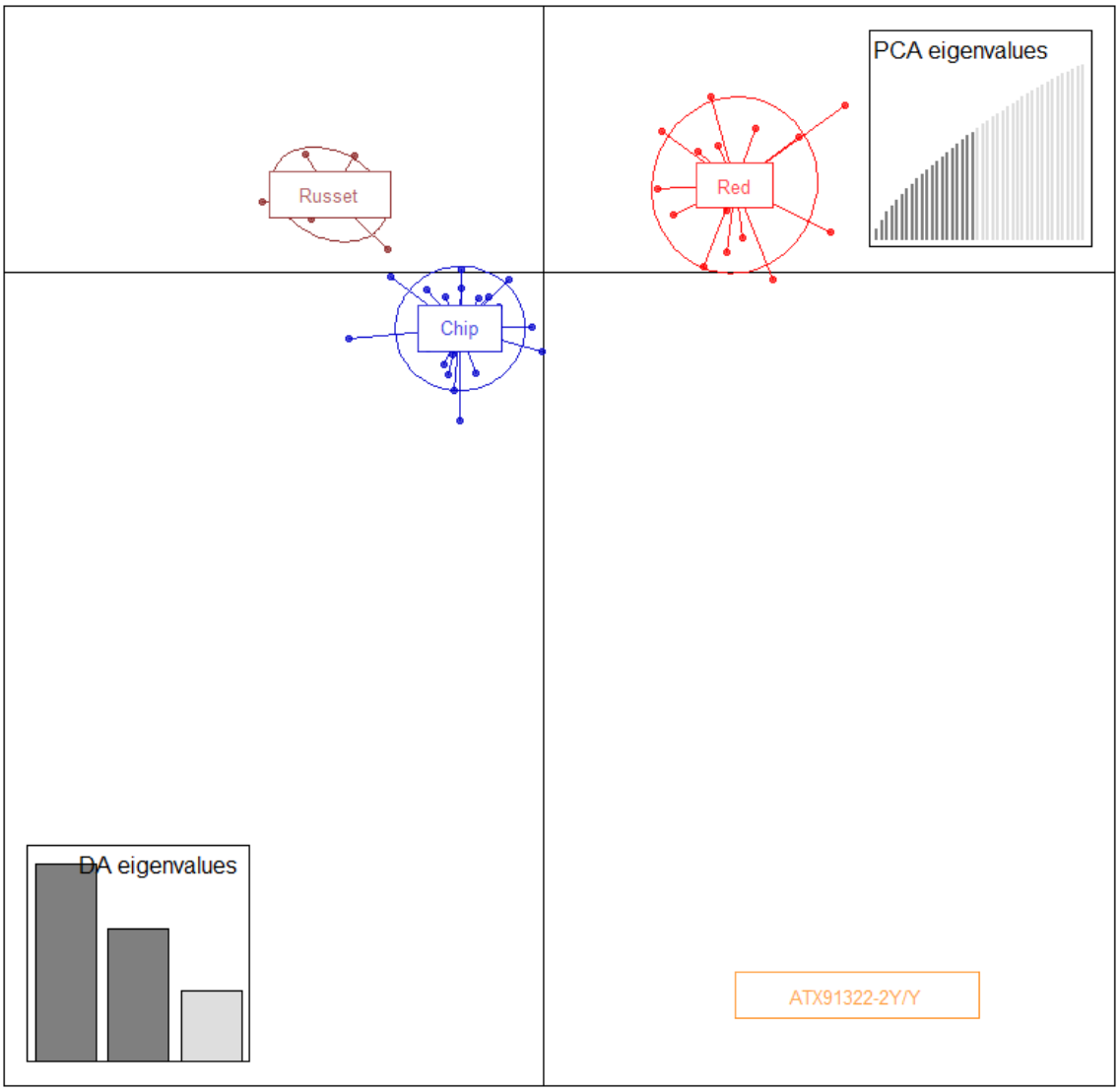
d.

e.

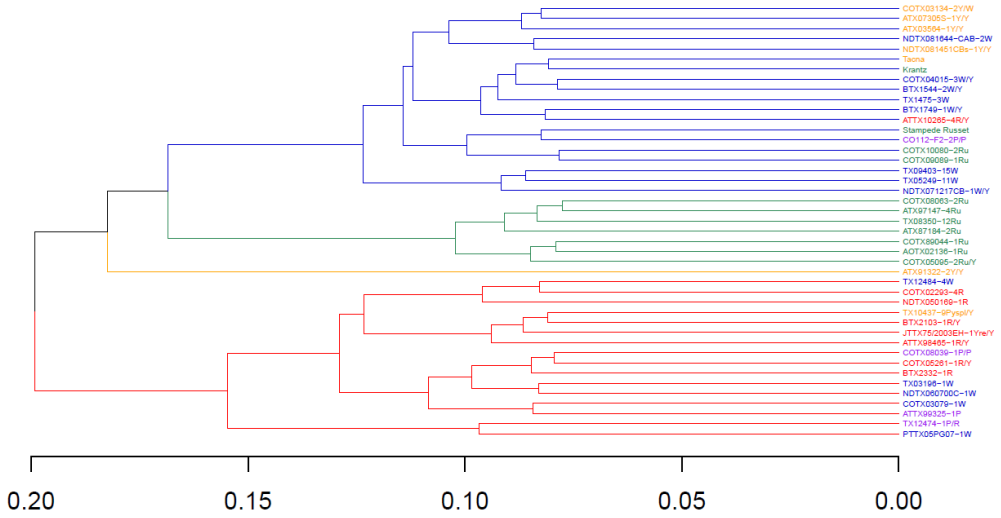
A5. Illustration of skin and Flesh color verification from minitubers when the naming error was found from phylogenetic analysis. a. AOTX98096-1R correct name AOTX98096-1Ru b. ATTX98468-5R/Y correct name ATTX98468-5Ru/Y c. ATX96746-1Ru correct name ATTX96746-1R d. COTX04303-3R/Y correct name COTX04303-3Ru/Y e. COTX08039-1R/R correct name COTX08039-1P/P



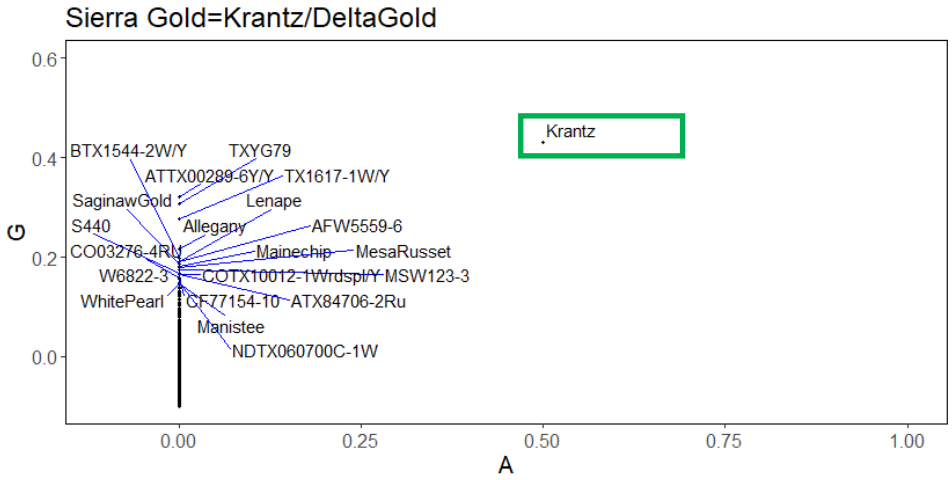
A6. Score plot showing the population structure (first two principal components PC1 and PC2) for the observed data set produced with PCAdapt where dots correspond to individuals and color indicates sub-populations (ChipRu = Blue, Red = Red and RN = Green).



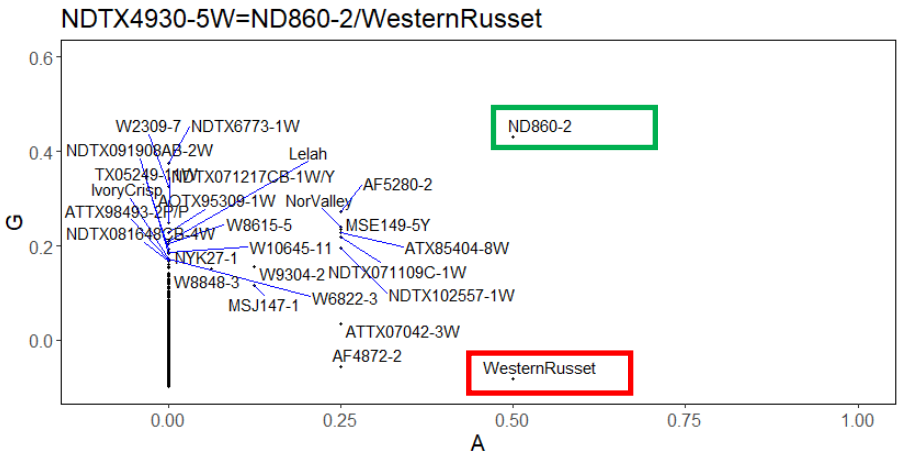
A7. Discriminant analysis of principal components (DAPC) for 43 clones in the core set using *adegenet* R package[67]. The axes represent the first two linear discriminants. Circles represent groups and dots represent individual clones. Numbers represent the different groups identified by DAPC analysis (Russet, Red and Purple and Chip and ATX91322-2Y/Y –-diploid-).



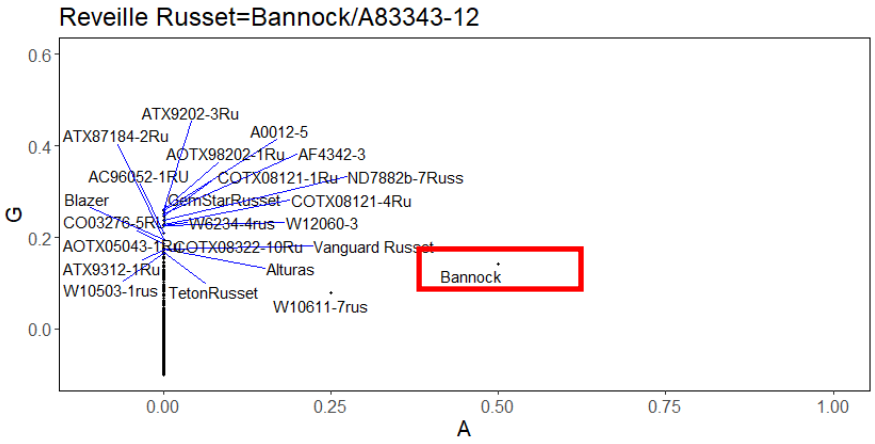
A8. Dendrogram of the core set (43 genotypes) divided into clusters using Ape R package[70]. In the X-axis represents the Nei's genetic distances between clones. In the Y-axis are represented the subpopulations by color in the core set. The color of the clones represents the market class (Red = red clones, Purple = purple clones, and Yellow = yellow clones; Green = Russet Clones, and; Blue = chipping clones)



A9. Population-wide comparison of genetic covariance calculated from markers with the additive relationship calculated from pedigree records, for clone Sierra Gold using R software[82] with one parent (Krantz) genotyped.



A10. Population-wide comparison of genetic covariance calculated from markers with the additive relationship calculated from pedigree records, for clone NDTX4930-5W using R software[82] in which one of the parent Western Russet seems to be erroneous.



A11. Population-wide comparison of genetic covariance calculated from markers with the additive relationship calculated from pedigree records, for clone Reveille Russet using R software[82] in which one of the genotyped parent Bannock seems to be erroneous.

A12. Heterozygosity and inbreeding coefficient of two hundred fourteen potato clones

Clone Name	Heterozygosity	Inbreeding coefficient (Fi)
AOR07781-2	0.60	-0.54
AORTX09037-1W/Y	0.60	-0.54
AOTX02136-1Ru	0.59	-0.50
AOTX03187-1Ru	0.56	-0.44
AOTX05043-1Ru	0.57	-0.46
AOTX91861-4R	0.58	-0.47
AOTX93483-1R	0.59	-0.51
AOTX95265-2Ru	0.57	-0.46
AOTX95295-3Ru	0.59	-0.51
AOTX95309-1W	0.60	-0.53
AOTX95309-2W	0.59	-0.49
AOTX96075-1Ru	0.59	-0.50
AOTX96084-1Ru	0.57	-0.47
AOTX96208-1Ru	0.58	-0.48
AOTX96216-1Ru	0.58	-0.49
AOTX96216-2Ru	0.58	-0.49
AOTX97213-1Ru	0.57	-0.46
AOTX98096-1Ru*	0.59	-0.50
AOTX98137-1Ru	0.59	-0.51
AOTX98152-3Ru	0.63	-0.59
AOTX98202-1Ru	0.60	-0.52
Atlantic_MSU	0.73	-0.85
Atlantic_TAMU	0.73	-0.85
ATTX00289-4W	0.61	-0.56
ATTX00289-5R/Y	0.59	-0.50
ATTX00289-6Y/Y	0.60	-0.51
ATTX01178-1R	0.59	-0.51
ATTX01180-1R/Y	0.58	-0.46
ATTX03516-2R	0.57	-0.45
ATTX05175s-1R/Y	0.56	-0.42
ATTX05186-2R	0.59	-0.52
ATTX06246-1R	0.58	-0.47
ATTX07042-3W	0.63	-0.59
ATTX10265-4R/Y	0.58	-0.47
ATTX88481-1P/W	0.57	-0.45
ATTX88654-2P/Y	0.57	-0.44
ATTX95490-2W	0.59	-0.51
ATTX961014-1AR/Y	0.57	-0.44
ATTX961014-1Br/Y	0.59	-0.49

Clone Name	Heterozygosity	Inbreeding coefficient (Fi)
Sierra RoseTM	0.59	-0.49
ATTX961014-1R/Y Chimera	0.58	-0.49
ATTX98444s-16R/Y	0.59	-0.50
ATTX98448-6R/Y	0.59	-0.50
ATTX98453-11Br	0.57	-0.44
ATTX98453-3R	0.57	-0.46
ATTX98453-6R	0.60	-0.52
ATTX98462s-3R/Y	0.59	-0.49
ATTX98465-1R/Y	0.52	-0.32
ATTX98466-5R/WR	0.59	-0.49
ATTX98468-5Ru/Y*	0.57	-0.46
ATTX98491-4Yrdspl/Y	0.55	-0.40
ATTX98493-1AR	0.57	-0.46
ATTX98493-2P/P	0.61	-0.56
ATTX98500-2P/Y	0.57	-0.46
ATTX98500-3P/Y	0.56	-0.43
ATTX98510-1R/Y	0.61	-0.55
ATTX98514-1R/Y	0.60	-0.54
ATTX98518-5P/Y	0.60	-0.54
ATTX99325-1P	0.56	-0.43
ATX02263-1R/Y	0.59	-0.51
ATX03496-3Y/Y	0.57	-0.46
ATX03564-1Y/Y	0.58	-0.46
ATX05186-1R	0.62	-0.56
ATX05202s-3W/Y	0.61	-0.54
ATX06264s-4R/Y_R1	0.59	-0.49
ATX06264s-4R/Y_R2	0.59	-0.50
ATX07305S-1Y/Y	0.61	-0.54
ATX08181-5Y/Y	0.58	-0.48
ATX84378-6Ru	0.61	-0.55
ATX84706-2Ru	0.61	-0.56
ATX85404-8W	0.60	-0.53
ATX87184-2Ru	0.56	-0.43
Reveille Russet	0.62	-0.58
ATX9117-1Ru	0.58	-0.48
ATX91322-2Y/Y	0.22	0.45
ATX9202-3Ru	0.60	-0.52
ATX9312-1Ru	0.59	-0.49
ATX9332-8Ru	0.59	-0.51
ATTX96746-1R*	0.60	-0.52
ATX97147-4Ru	0.62	-0.58

Clone Name	Heterozygosity	Inbreeding coefficient (Fi)
ATX99013-1Ru	0.59	-0.51
BTX1544-2W/Y	0.62	-0.58
BTX1749-1W/Y	0.62	-0.57
BTX2103-1R/Y	0.57	-0.45
BTX2332-1R	0.64	-0.64
CO112-F2-2P/P	0.62	-0.58
COTX00104-6R	0.57	-0.45
COTX00104-7R	0.59	-0.49
COTX01403-4R/Y	0.60	-0.53
COTX02172-1R	0.58	-0.48
COTX02293-4R	0.57	-0.46
COTX03079-1W/W	0.60	-0.52
COTX03134-1Y/W	0.61	-0.55
COTX03134-2Y/W	0.60	-0.52
COTX03187-1W	0.57	-0.46
COTX04015-3W/Y	0.60	-0.53
COTX04050s-1P/P	0.61	-0.55
COTX04193s-2R/Y	0.58	-0.47
COTX04303-3Ru/Y*	0.59	-0.51
COTX05095-2Ru/Y	0.57	-0.44
COTX05211-4R	0.55	-0.41
COTX05211-5R	0.55	-0.41
COTX05211-7R	0.58	-0.48
COTX05249-3W/Y	0.57	-0.44
COTX05261-1R/Y	0.56	-0.41
COTX08039-1P/P*	0.56	-0.42
COTX08063-2Ru	0.59	-0.51
COTX08121-1Ru	0.60	-0.53
COTX08121-4Ru	0.58	-0.48
COTX08258-6Ru	0.60	-0.51
COTX08322-10Ru	0.57	-0.46
COTX08365F-3P/P	0.58	-0.46
COTX09022-3RuRE/Y	0.61	-0.55
COTX09052-1Ru	0.60	-0.53
COTX09052-2Ru	0.60	-0.53
COTX09089-1Ru	0.63	-0.61
COTX10012-1Wrdspl/Y	0.59	-0.51
COTX10073s-1W	0.73	-0.87
COTX10080-2Ru	0.59	-0.50
COTX10118-1Wre/Y	0.58	-0.48
COTX10118-4Wpe/Y	0.54	-0.38

Clone Name	Heterozygosity	Inbreeding coefficient (Fi)
COTX10138-15Wpe/Y	0.54	-0.38
COTX10138s-7Wpe/Y	0.54	-0.38
COTX10226-1Wpe/Y	0.54	-0.38
COTX13215-2Ru	0.62	-0.57
COTX87601-2Ru	0.61	-0.56
COTX89044-1Ru	0.61	-0.55
COTX90046-1W	0.64	-0.64
COTX90046-5W	0.61	-0.55
COTX94216-1R	0.75	-0.89
COTX94218-1R	0.57	-0.45
JTTX75/2003EH-1Yre/Y	0.58	-0.47
Krantz	0.63	-0.61
MWTX2609-2Ru	0.61	-0.56
MWTX2609-4Ru	0.61	-0.56
MWTX548-2Ru	0.61	-0.56
NDTX050169-1R	0.49	-0.25
NDTX050184s-1R/Y	0.57	-0.44
NDTX059759-3R/Y Pinto	0.58	-0.47
NDTX059761-1R/R	0.61	-0.54
NDTX059775-1W	0.58	-0.48
NDTX059828-2W	0.56	-0.42
NDTX059886S-1Y/Y	0.77	-0.94
NDTX060700C-1W	0.58	-0.48
NDTX071109C-1W	0.59	-0.49
NDTX071217CB-1W/Y	0.59	-0.49
NDTX071258B-1R	0.55	-0.40
NDTX081451CBs-1Y/Y	0.57	-0.46
NDTX081618-1P/P	0.62	-0.58
NDTX081644-CAB-2W	0.57	-0.44
NDTX081648CB-13W	0.63	-0.59
NDTX081648CB-1W	0.80	-1.00
NDTX081648CB-4W	0.62	-0.58
NDTX091886-3P/P	0.63	-0.60
NDTX091908AB-2W	0.61	-0.56
NDTX092237C-2P/W	0.58	-0.48
NDTX092238Cs-1P/W	0.57	-0.45
NDTX4271-5R	0.58	-0.48
NDTX4756-1R/Y	0.61	-0.55
NDTX4784-7R	0.58	-0.47
NDTX4828-2R	0.58	-0.48
NDTX4930-5W	0.61	-0.55

Clone Name	Heterozygosity	Inbreeding coefficient (Fi)
NDTX5003-2R	0.59	-0.49
NDTX5067-2R	0.58	-0.49
NDTX5438-11R	0.79	-0.97
NDTX6773-1W	0.60	-0.52
NDTX731-1R	0.58	-0.46
NDTX7590-3R	0.58	-0.48
NDTX8773-4Ru	0.59	-0.49
NDTX91068-11R	0.59	-0.49
PORTX03PG25-2R/R	0.61	-0.55
PTTX05PG07-1W	0.54	-0.38
Rio Rojo	0.56	-0.42
Russet Burbank	0.63	-0.60
Russet Norkotah 102	0.59	-0.50
Russet Norkotah 112	0.59	-0.50
Russet Norkotah 223	0.59	-0.51
Russet Norkotah 278	0.59	-0.51
Russet Norkotah 296	0.59	-0.51
Russet Norkotah	0.59	-0.51
Stampede Russet	0.62	-0.58
Tacna	0.62	-0.56
Tokio	0.56	-0.43
TX03196-1W	0.59	-0.51
TX05249-10W	0.62	-0.57
TX05249-11W	0.62	-0.57
TX05249-3W	0.64	-0.63
TX08350-12Ru	0.60	-0.53
Vanguard Russet	0.63	-0.59
TX08385-2P/YP	0.63	-0.60
TX09396-1W	0.66	-0.67
TX09403-21W	0.63	-0.59
TX09406s-1P/P	0.61	-0.55
TX09414-1W	0.61	-0.55
TX10437-9Pyspl/Y	0.54	-0.37
TX11454-9Ru/Y	0.60	-0.52
TX11461-3W	0.60	-0.53
TX12474-1P/R	0.48	-0.22
TX12484-4W	0.59	-0.50
TX13590-9Ru	0.63	-0.59
TX14611-1R	0.60	-0.53
TX1475-3W	0.58	-0.49
Sierra GoldTM	0.62	-0.58

Clone Name	Heterozygosity	Inbreeding coefficient (Fi)
TX1617-1W/Y	0.64	-0.62
TX1673-1W/Y	0.62	-0.57
TX6-1216-1Ru	0.62	-0.59
TX09403-15W	0.63	-0.59
TXA549-1Ru	0.63	-0.59
TXNS106	0.59	-0.51
TXNS118	0.59	-0.51
TXNS249	0.59	-0.51
TXYG79	0.64	-0.62
UMTX383-3Yrdspl/Y	0.55	-0.40
White LaSoda	0.63	-0.61

A13. SNPs found under selection on chromosomes 1, 2, 3, 4, 5, 7, 8, and 10 using the PCAdapt method at α=0.05 corrected for the genomic inflation factor (λGC=1.20)

snp	chr	pos	pval	maf	stat	nlog10pval
PotVar0096714	chr01	79.47	0.00	0.06	20.22	3.82
PotVar0049884	chr01	79.82	0.00	0.06	20.22	3.82
PotVar0060730	chr01	82.03	0.00	0.06	21.56	4.09
solcap_snp_c2_21752	chr02	25.05	0.00	0.21	20.23	3.82
PotVar0038601	chr02	34.81	0.00	0.07	24.82	4.77
PotVar0128499	chr02	35.67	0.00	0.08	21.48	4.08
PotVar0055453	chr03	6.78	0.00	0.05	36.96	7.33
PotVar0021465	chr03	7.33	0.00	0.05	36.96	7.33
PotVar0021405	chr03	7.34	0.00	0.05	36.96	7.33
PotVar0120627	chr03	48.55	0.00	0.15	21.79	4.14
PotVar0015935	chr04	70.87	0.00	0.05	36.86	7.31
PotVar0091295	chr05	9.84	0.00	0.17	22.21	4.23
PotVar0092752	chr07	42.06	0.00	0.06	33.56	6.61
solcap_snp_c2_33489	chr07	45.15	0.00	0.09	27.83	5.40
solcap_snp_c2_51053	chr08	46.36	0.00	0.29	21.02	3.98
PotVar0023940	chr08	54.75	0.00	0.23	24.21	4.65
PotVar0023779	chr08	54.76	0.00	0.22	22.24	4.24
PotVar0023676	chr08	54.78	0.00	0.22	23.95	4.59
solcap_snp_c2_16996	chr08	54.84	0.00	0.22	21.65	4.11
PotVar0004789	chr10	54.88	0.00	0.41	22.75	4.34
solcap_snp_c1_5001	chr10	55.16	0.00	0.41	21.65	4.11
solcap_snp_c2_15533	chr10	55.29	0.00	0.39	20.98	3.97
PotVar0005291	chr10	55.86	0.00	0.24	30.46	5.96

snp	chr	pos	pval	maf	stat	nlog10pval
PotVar0005683	chr10	56.86	0.00	0.25	31.40	6.15
PotVar0057721	chr10	58.19	0.00	0.28	25.45	4.91
PotVar0057635	chr10	58.30	0.00	0.25	29.92	5.84

A14. Grouping of core set of potatoes based on Ward.D distance

Clone Name	**Group**
AOTX02136-1Ru	1
ATX87184-2Ru	1
ATX97147-4Ru	1
COTX05095-2Ru/Y	1
COTX08063-2Ru	1
COTX89044-1Ru	1
TX08350-12Ru	1
ATTX10265-4R/Y	2
ATTX98465-1R/Y	2
ATTX99325-1P	2
BTX2103-1R/Y	2
BTX2332-1R	2
COTX02293-4R	2
COTX05261-1R/Y	2
COTX08039-1P/P	2
JTTX75/2003EH-1Yre/Y	2
NDTX050169-1R	2
PTTX05PG07-1W	2
TX03196-1W	2
TX10437-9Pyspl/Y	2
TX12474-1P/R	2
TX12484-4W	2
ATX03564-1Y/Y	3
ATX07305S-1Y/Y	3
BTX1544-2W/Y	3
BTX1749-1W/Y	3
CO112-F2-2P/P	3
COTX03079-1W/W	3
COTX03134-2Y/W	3
COTX04015-3W/Y	3
COTX09089-1Ru	3
COTX10080-2Ru	3

Clone Name	Group
Krantz	3
NDTX060700C-1W	3
NDTX071217CB-1W/Y	3
NDTX081451CBs-1Y/Y	3
NDTX081644-CAB-2W	3
Stampede Russet	3
Tacna	3
TX05249-11W	3
TX1475-3W	3
TX09403-15W	3
ATX91322-2Y/Y	4

APPENDIX B

B1. Best linear unbiased prediction (gBLUP) of tuber traits using 214 advanced clones from Texas potato program using polyBreedR

Clone id	Tubershape (1-5)	LW ratio	Russeting (1-5)	Eye depth (1-5)	Avg tuber wt (g)	Avg tubers per plant (no.)	Avg tuber wt per plant (g)	Grading at table (1-5)	C_Flesh (value)
AOR07781-2	3.7	1.6	3.4	4.1	144	7	1009	3.8	12
AORTX09037-1W/Y	1.5	1.2	1.0	4.2	114	9	923	3.7	23
AOTX02136-1Ru	3.9	1.7	3.8	4.2	151	6	835	3.7	12
AOTX03187-1Ru	4.1	2.2	1.3	4.3	86	9	728	3.5	12
AOTX05043-1Ru	3.6	1.6	3.9	4.0	121	5	656	3.6	12
AOTX91861-4R	1.1	1.1	1.1	3.8	111	9	940	3.7	14
AOTX93483-1R	2.6	1.2	1.1	4.1	100	9	825	3.6	14
AOTX95265-2Ru	4.4	1.9	4.2	3.9	145	5	749	3.7	14
AOTX95295-3Ru	4.5	2.0	4.3	3.8	148	5	759	3.8	13
AOTX95309-1W	3.3	1.3	2.3	4.0	145	6	812	3.5	15
AOTX95309-2W*	2.7	1.3	1.0	4.2	105	9	916	3.7	24
AOTX96075-1Ru	4.4	2.0	4.3	3.8	148	5	758	3.8	13
AOTX96084-1Ru	4.4	1.9	4.2	3.8	141	5	748	3.7	14
AOTX96208-1Ru	4.3	1.9	4.2	3.8	144	5	752	3.8	13
AOTX96216-1Ru	4.4	1.6	4.3	3.9	193	4	847	3.8	13
AOTX96216-2Ru	4.5	1.6	4.3	4.0	195	5	844	3.8	13
AOTX97213-1Ru	4.0	1.8	3.8	4.0	132	6	861	3.7	13
AOTX98096-1Ru*	4.4	2.0	4.3	3.8	147	5	759	3.8	13
AOTX98137-1Ru	4.5	2.0	4.3	3.8	148	5	759	3.8	13
AOTX98152-3Ru	4.0	1.5	3.8	4.0	182	7	1115	4.0	12
AOTX98202-1Ru	4.1	1.7	3.7	4.1	139	7	875	3.8	13
Atlantic	1.7	1.1	2.1	3.8	126	8	957	3.8	15
ATTX00289-4W	2.1	1.2	1.1	4.1	124	8	964	3.7	17
ATTX00289-5R/Y	2.5	1.2	1.0	3.9	109	11	1031	3.5	23
ATTX00289-6Y/Y	2.0	1.2	1.1	4.2	112	9	894	3.5	31
ATTX01178-1R	1.1	1.1	1.0	4.0	106	9	888	3.7	15
ATTX01180-1R/Y	2.6	1.3	1.1	4.0	89	10	828	3.5	29
ATTX03516-2R	1.5	1.1	1.0	4.0	98	9	807	3.7	18
ATTX05175s-1R/Y	1.0	1.0	1.0	3.8	51	12	610	3.6	34
ATTX05186-2R	1.1	1.1	1.0	4.0	56	12	686	3.6	14
ATTX06246-1R	1.3	1.2	1.1	4.1	67	12	760	3.6	15
ATTX07042-3W	1.3	1.1	1.4	4.0	114	8	811	3.4	25
ATTX10265-4R/Y	1.7	1.2	1.0	4.1	77	10	768	3.6	30
ATTX88481-1P/W	3.5	1.4	1.1	4.2	128	8	921	3.8	15

Clone id	Tubershape (1-5)	LW ratio	Russeting (1-5)	Eye depth (1-5)	Avg tuber wt (g)	Avg tubers per plant (no.)	Avg tuber wt per plant (g)	Grading at table (1-5)	C_Flesh (value)
ATTX88654-2P/Y	1.2	1.0	1.0	3.8	85	9	685	3.4	31
ATTX95490-2W	2.3	1.2	1.0	3.9	120	10	1068	3.4	14
ATTX961014-1AR/Y	2.7	1.3	1.0	4.1	109	9	920	3.8	24
ATTX961014-1Br/Y	2.7	1.3	1.0	4.2	105	9	912	3.7	24
ATTX961014-1R/Y Chimera	1.4	1.2	1.0	4.0	86	10	823	3.6	16
ATTX96746-1R*	2.5	1.3	1.1	4.1	92	8	724	3.5	13
ATTX98444s-16R/Y	1.3	1.2	1.1	4.0	57	10	587	3.2	23
ATTX98448-6R/Y	2.5	1.2	1.0	3.9	110	11	1032	3.5	23
ATTX98453-11Br	1.9	1.1	1.1	3.9	79	11	798	3.6	14
ATTX98453-3R	1.6	1.1	1.0	3.9	85	11	875	3.6	12
ATTX98453-6R	1.8	1.2	1.2	4.1	108	8	820	3.6	13
ATTX98462s-3R/Y	2.7	1.3	1.0	4.2	106	9	916	3.8	24
ATTX98465-1R/Y	1.1	1.1	1.1	3.5	83	12	932	3.3	33
ATTX98466-5R/WR	1.4	1.1	1.0	4.0	86	10	822	3.6	15
ATTX98468-5Ru/Y*	3.9	1.8	4.1	4.1	132	5	648	3.6	12
ATTX98491-4Yrdspl/Y	1.7	1.1	1.0	3.8	71	10	704	3.4	34
ATTX98493-1AR	2.0	1.2	1.1	4.0	80	10	803	3.5	27
ATTX98493-2P/P	1.9	1.3	1.2	4.2	83	10	755	3.6	17
ATTX98500-2P/Y	1.6	1.2	1.0	4.0	68	9	653	3.4	32
ATTX98500-3P/Y	2.2	1.4	1.1	4.0	74	8	544	3.0	33
ATTX98510-1R/Y	1.4	1.1	1.0	3.9	84	10	782	3.5	26
ATTX98514-1R/Y	1.8	1.2	1.0	4.0	83	7	554	3.4	25
ATTX98518-5P/Y	3.2	1.7	1.0	4.1	127	8	834	3.5	28
ATTX99325-1P	3.5	1.4	1.1	4.2	126	8	920	3.8	15
ATX02263-1R/Y	2.0	1.3	1.0	4.2	70	10	660	3.7	24
ATX03496-3Y/Y	2.3	1.3	1.1	4.1	89	10	844	3.5	25
ATX03564-1Y/Y	2.3	1.3	1.0	4.1	88	10	846	3.5	24
ATX05186-1R	1.2	1.1	0.9	4.0	56	13	699	3.6	14
ATX05202s-3W/Y	1.0	1.1	1.1	4.2	71	11	748	3.7	27
ATX06264s-4R/Y	1.8	1.2	1.0	4.1	68	11	701	3.3	30
ATX07305S-1Y/Y	1.2	1.1	1.1	3.9	57	12	643	3.4	26
ATX08181-5Y/Y	1.0	1.1	1.1	4.1	47	15	633	3.5	28
ATX84378-6Ru	4.7	1.6	4.4	4.0	209	4	887	3.9	12
ATX84706-2Ru	4.2	1.5	3.6	4.2	240	5	1062	4.0	16
ATX85404-8W	3.3	1.3	2.4	4.0	146	6	814	3.5	15
ATX87184-2Ru	4.0	1.5	3.7	4.0	175	4	791	3.6	13
ATX9117-1Ru	3.8	1.6	3.6	4.0	151	5	704	3.4	13
ATX9130-1Ru	4.6	1.8	3.6	4.3	170	5	873	3.5	13

Clone id	Tubershape (1-5)	LW ratio	Russeting (1-5)	Eye depth (1-5)	Avg tuber wt (g)	Avg tubers per plant (no.)	Avg tuber wt per plant (g)	Grading at table (1-5)	C_Flesh (value)
ATX91322-2Y/Y	1.0	1.1	1.0	3.3	14	13	290	2.4	45
ATX9202-3Ru	3.5	1.8	3.8	4.1	111	6	664	3.6	12
ATX9312-1Ru	4.0	1.8	4.1	3.9	130	5	701	3.5	12
ATX9332-8Ru	4.0	1.8	3.8	3.9	146	4	701	3.5	11
ATX97147-4Ru	4.0	1.7	3.9	4.0	134	5	746	3.6	12
ATX99013-1Ru	4.5	2.0	4.3	3.8	148	5	759	3.8	13
BTX1544-2W/Y	2.2	1.3	1.3	4.0	105	8	831	3.5	26
BTX1749-1W/Y	1.7	1.2	1.1	4.2	87	10	857	3.7	29
BTX2103-1R/Y	1.5	1.1	1.0	3.9	77	10	751	3.6	27
BTX2332-1R	1.2	1.1	1.0	4.2	110	8	829	3.8	15
CO112-F2-2P/P	1.8	1.6	1.1	3.9	42	10	420	3.0	18
COTX00104-6R	2.0	1.2	1.0	4.1	106	8	724	3.4	14
COTX00104-7R	2.4	1.2	1.0	4.0	109	8	818	3.5	13
COTX01403-4R/Y	3.3	1.2	1.0	4.0	97	9	823	3.5	28
COTX02172-1R	2.1	1.2	1.0	4.1	88	9	721	3.7	14
COTX02293-4R	1.4	1.2	1.0	4.0	88	9	774	3.7	15
COTX03079-1W/W	2.4	1.2	1.0	4.0	88	11	903	3.6	28
COTX03134-1Y/W	1.9	1.3	1.1	4.0	55	9	573	2.8	21
COTX03187-1W	4.4	2.2	1.4	4.3	87	9	725	3.5	13
COTX04015-3W/Y	2.2	1.3	1.1	4.1	82	9	731	3.3	33
COTX04050s-1P/P	1.2	1.1	1.1	4.1	73	10	719	3.6	17
COTX04193s-2R/Y	1.2	1.1	1.0	4.1	69	10	701	3.6	29
COTX04303-3Ru/Y*	4.5	2.0	4.3	3.8	148	5	759	3.8	13
COTX05095-2Ru/Y	3.8	1.8	2.4	4.1	117	7	838	3.7	26
COTX05211-4R	2.0	1.4	1.0	4.0	75	10	690	3.5	15
COTX05211-5R	2.0	1.4	1.0	4.0	75	10	689	3.5	15
COTX05211-7R	1.2	1.1	1.0	4.0	78	11	847	3.5	14
COTX05249-3W/Y	1.1	1.2	1.0	4.0	64	11	688	3.7	26
COTX05261-1R/Y	3.4	1.5	1.0	4.0	83	10	801	3.5	30
COTX08039-1P/P*	2.4	1.5	1.0	4.2	72	9	649	3.4	12
COTX08063-2Ru	3.8	1.9	3.4	4.0	133	5	749	3.6	12
COTX08121-1Ru	3.9	1.7	3.9	4.1	142	6	806	3.7	12
COTX08121-4Ru	3.5	1.7	3.8	4.0	122	6	781	3.7	12
COTX08258-6Ru	3.5	1.6	2.8	4.2	126	8	1016	3.8	12
COTX08322-10Ru	4.1	1.6	3.9	4.0	168	6	1028	4.0	12
COTX08365F-3P/P	4.7	2.1	1.1	4.0	70	9	572	3.6	16
COTX09022-3RuRE/Y	2.4	1.3	3.0	4.0	120	7	839	3.9	26
COTX09052-1Ru	3.7	1.7	3.7	4.1	105	7	814	3.7	12

Clone id	Tubershape (1-5)	LW ratio	Russeting (1-5)	Eye depth (1-5)	Avg tuber wt (g)	Avg tubers per plant (no.)	Avg tuber wt per plant (g)	Grading at table (1-5)	C_Flesh (value)
COTX09052-2Ru	3.7	1.8	3.7	4.1	106	7	816	3.7	12
COTX09089-1Ru	2.4	1.3	3.9	4.0	116	7	824	3.4	15
COTX10012-1Wrdspl/Y	1.5	1.2	1.1	4.0	94	10	889	3.7	28
COTX10073s-1W	1.2	1.2	1.1	4.1	72	11	725	3.5	16
COTX10080-2Ru	3.8	1.7	4.0	4.0	122	7	755	3.8	13
COTX10118-1Wre/Y	1.1	1.1	1.0	3.9	71	12	774	3.6	25
COTX10118-4Wpe/Y	1.8	1.2	1.0	4.0	74	12	827	3.7	29
COTX10138-15Wpe/Y	1.0	1.1	0.9	4.1	67	12	764	3.8	28
COTX10138-19P/Y	1.9	1.1	1.0	4.0	113	9	811	3.5	28
COTX10138s-7Wpe/Y	1.3	1.1	0.9	4.0	68	13	823	3.7	29
COTX10226-1Wpe/Y	1.0	1.0	1.0	3.9	51	11	597	3.6	31
COTX13215-2Ru	3.0	1.5	2.9	4.1	137	6	821	3.6	14
COTX87601-2Ru	4.4	1.8	3.9	3.9	160	6	893	3.9	12
COTX89044-1Ru	3.7	1.8	3.6	4.0	136	5	748	3.5	12
COTX90046-1W	2.2	1.2	1.2	4.0	110	8	832	3.7	14
COTX90046-5W	3.1	1.2	1.7	4.0	131	6	853	3.6	15
COTX94216-1R	1.1	1.1	1.0	4.1	79	11	817	3.7	15
COTX94218-1R	1.0	1.1	0.9	4.0	72	11	731	3.6	15
JTTX75/2003EH-1Yre/Y	2.9	1.4	1.5	3.9	94	8	709	3.5	25
Krantz	3.3	1.3	3.7	4.1	168	6	1023	3.7	16
MWTX2609-2Ru	4.3	1.9	3.0	4.0	164	6	1007	3.7	10
MWTX2609-4Ru	4.3	1.9	3.0	4.0	164	6	1009	3.7	10
MWTX548-2Ru	4.3	1.9	3.0	4.0	163	6	1004	3.7	10
NDTX050169-1R	1.0	1.1	1.1	4.0	50	14	641	3.6	13
NDTX050184s-1R/Y	1.1	1.1	1.0	4.0	65	13	784	3.8	25
NDTX059759-3R/Y Pinto	2.2	1.3	0.9	4.3	74	7	570	3.6	29
NDTX059761-1R/R	2.5	1.4	1.1	4.0	61	11	620	3.5	16
NDTX059775-1W	2.1	1.3	1.0	3.9	70	11	756	3.6	25
NDTX059828-2W	1.1	1.1	0.9	3.9	74	10	741	3.7	15
NDTX059886S-1Y/Y	1.7	1.2	1.1	4.1	101	9	866	3.6	25
NDTX060700C-1W	1.3	1.1	1.2	4.2	76	9	710	3.6	16
NDTX071109C-1W	1.3	1.0	1.0	3.9	137	7	939	3.7	15
NDTX071217CB-1W/Y	1.6	1.1	1.0	4.0	97	8	754	3.6	27
NDTX071258B-1R	1.1	1.1	1.1	3.9	67	11	664	3.6	16
NDTX081451CBs-1Y/Y	1.8	1.3	1.2	4.1	86	11	879	3.7	29
NDTX081618-1P/P	2.3	1.3	1.0	4.3	87	10	797	3.8	15
NDTX081644-CAB-2W	1.0	1.1	1.0	4.1	57	9	567	3.4	21
NDTX081648CB-13W	1.7	1.1	1.2	4.0	107	9	935	3.8	16

Clone id	Tubershape (1-5)	LW ratio	Russeting (1-5)	Eye depth (1-5)	Avg tuber wt (g)	Avg tubers per plant (no.)	Avg tuber wt per plant (g)	Grading at table (1-5)	C_Flesh (value)
NDTX081648CB-1W	1.5	1.1	1.1	4.0	98	9	844	3.7	16
NDTX081648CB-4W	1.2	1.1	1.2	4.0	118	8	886	3.6	15
NDTX091886-3P/P	2.2	1.2	1.0	4.2	75	7	559	3.5	17
NDTX091908AB-2W	1.6	1.2	1.0	4.1	106	9	916	3.7	15
NDTX092237C-2P/W	2.4	1.4	1.0	3.7	61	12	685	3.4	23
NDTX092238Cs-1P/W	1.0	1.1	1.0	4.1	52	16	769	3.6	17
NDTX4271-5R	1.2	1.1	0.9	4.0	103	9	856	3.7	14
NDTX4756-1R/Y	1.4	1.1	0.9	3.9	84	10	782	3.5	26
NDTX4784-7R	1.1	1.1	1.0	4.0	106	9	866	3.7	14
NDTX4828-2R	1.2	1.1	1.1	4.0	77	10	711	3.6	14
NDTX4930-5W	3.4	1.2	1.1	4.0	173	6	862	3.6	13
NDTX5003-2R	1.0	1.0	1.1	3.9	83	11	843	3.7	17
NDTX5067-2R	1.0	1.0	1.2	3.9	84	11	845	3.7	17
NDTX5438-11R	2.1	1.2	1.1	4.0	99	9	854	3.6	16
NDTX6773-1W	1.1	1.1	0.9	3.9	117	8	845	3.7	15
NDTX731-1R	1.1	1.0	1.0	3.9	99	11	966	3.7	13
NDTX7590-3R	2.9	1.3	1.0	4.2	103	10	983	3.5	12
NDTX8773-4Ru	3.2	1.5	4.1	4.1	123	6	685	3.5	17
NDTX91068-11R	2.4	1.1	1.1	4.1	127	9	1016	3.6	13
PORTX03PG25-2R/R	4.6	2.1	1.0	4.1	56	9	538	3.7	20
PTTX05PG07-1W	4.8	2.2	1.0	4.1	60	7	458	3.7	16
Reveille Russet	3.8	1.6	3.8	4.0	138	6	760	3.7	13
Rio Rojo	1.5	1.2	1.0	3.9	59	12	620	3.5	16
Russet Burbank	4.4	2.0	3.2	3.8	112	6	708	3.2	11
Russet Norkotah	4.5	2.0	4.3	3.8	148	5	758	3.8	13
Russet Norkotah 102	4.4	2.0	4.3	3.8	147	5	758	3.8	13
Russet Norkotah 112	4.4	2.0	4.3	3.8	147	5	756	3.8	13
Russet Norkotah 223	4.4	2.0	4.3	3.8	148	5	758	3.8	13
Russet Norkotah 278	4.5	2.0	4.3	3.8	148	5	758	3.8	13
Russet Norkotah 296	4.5	2.0	4.3	3.8	148	5	759	3.8	13
Sierra GoldTM	2.9	1.3	3.1	4.2	172	8	1205	3.8	25
Sierra RoseTM	2.7	1.3	1.0	4.2	106	9	914	3.8	24
Stampede Russet	4.0	1.6	3.8	3.9	156	6	889	4.0	12
Tacna	1.9	1.3	1.1	4.0	72	7	603	2.8	14
Tokio	1.1	1.2	1.1	4.0	63	11	626	3.5	31
TX03196-1W	1.7	1.1	1.0	4.0	107	9	860	3.6	14
TX05249-10W	1.7	1.1	2.0	4.0	137	7	881	3.7	12
TX05249-11W	1.6	1.2	1.4	4.3	113	6	729	3.8	16

Clone id	Tubershape (1-5)	LW ratio	Russeting (1-5)	Eye depth (1-5)	Avg tuber wt (g)	Avg tubers per plant (no.)	Avg tuber wt per plant (g)	Grading at table (1-5)	C_Flesh (value)
TX05249-3W	1.4	1.1	1.4	4.1	131	5	750	3.6	15
TX08350-12Ru	4.3	1.6	3.1	4.0	162	6	924	3.9	12
TX08385-2P/YP	2.2	1.1	0.8	4.1	79	9	671	3.5	24
TX09396-1W	1.5	1.1	1.3	3.9	131	7	818	3.5	16
TX09403-15W	2.3	1.1	1.2	4.0	145	8	1049	3.5	15
TX09403-21W	2.3	1.1	1.2	4.0	147	8	1057	3.5	16
TX09414-1W	2.2	1.2	1.3	4.0	124	6	770	3.5	14
TX10437-9Pyspl/Y	1.5	1.3	1.0	4.0	69	8	571	2.9	29
TX11454-9Ru/Y	2.9	1.3	3.7	4.0	143	6	817	3.6	19
TX11461-2W	1.3	1.1	1.4	4.0	114	9	969	3.6	15
TX11461-3W	1.9	1.1	1.2	4.0	106	9	941	3.6	15
TX12474-1P/R	1.1	1.2	1.1	4.1	46	10	481	3.4	21
TX12484-4W	2.0	1.2	1.2	4.0	96	9	793	3.7	13
TX13590-9Ru	4.0	1.5	3.9	4.0	181	7	1112	4.0	12
TX14611-1R	1.0	1.1	1.0	4.1	77	9	646	3.6	16
TX1475-3W	2.2	1.1	1.3	3.9	149	6	923	3.7	16
TX1617-1W/Y	3.0	1.4	1.2	4.3	116	7	885	3.6	30
TX1673-1W/Y	2.3	1.2	1.1	4.1	124	8	966	3.7	18
TX6-1216-1Ru	4.1	1.6	3.8	3.9	149	5	767	3.7	12
TXA549-1Ru	4.0	1.6	3.9	4.0	181	7	1112	4.0	12
TXNS106	4.4	2.0	4.3	3.8	148	5	758	3.8	13
TXNS118	4.5	2.0	4.3	3.8	148	5	758	3.8	13
TXNS249	4.4	2.0	4.3	3.8	147	5	758	3.8	13
TXYG79	2.2	1.2	1.2	4.1	126	7	884	3.6	28
UMTX383-3Yrdspl/Y	1.8	1.1	1.0	3.8	72	10	704	3.4	34
Vanguard Russet	4.0	1.7	4.1	4.0	159	7	1069	4.0	12
White LaSoda	2.3	1.2	1.2	3.8	143	8	1040	3.6	15

APPENDIX C

C1. Breeding values and reliability score for fried chip color, chip quality, specific gravity, and total yield

Clone id	Chip Color	r^2	Chip quality	r^2	Specific Gravity	r^2	Yield Ton/Ha	r^2
NYR102-3	1.1	0.7	2.3	0.4	1.1	0.6	60.7	0.3
NYR102-7	1.2	0.7	2.4	0.4	1.1	0.6	61.5	0.3
NY169	1.2	0.8	2.3	0.5	1.1	0.7	64.1	0.4
NDTX1246-3W	1.5	0.9	2.4	0.6	1.1	0.8	62.5	0.4
W13NYP19-2	1.6	0.6	2.5	0.4	1.1	0.6	58.9	0.3
NYN24-2	1.5	0.6	2.5	0.4	1.1	0.6	57.4	0.3
NYR102-8	1.1	0.7	2.4	0.4	1.1	0.6	59.6	0.3
NDTX113030C-3W	1.3	1.0	2.4	0.7	1.1	0.8	56.6	0.5
NYP116-6	1.1	0.6	2.4	0.4	1.1	0.6	59.3	0.3
NYP108-6	1.3	0.6	2.5	0.4	1.1	0.6	56.9	0.3
NDTX1246-5W/Y	2.8	1.0	2.2	0.7	1.1	0.8	63.1	0.5
MSZ246-1	1.2	0.7	2.6	0.5	1.1	0.7	60.1	0.4
NYM15-3	1.1	0.7	2.4	0.5	1.1	0.6	57.2	0.4
NYR105-11	0.9	0.7	2.4	0.5	1.1	0.6	57.9	0.4
NYQ29-1	1.4	0.9	2.7	0.5	1.1	0.7	68.5	0.4
NYR101-7	1.0	0.7	2.5	0.4	1.1	0.6	63.4	0.3
NYQ37-1	1.2	0.8	2.6	0.4	1.1	0.6	60.2	0.3
NYP119-3	1.4	0.5	2.4	0.3	1.1	0.6	54.0	0.3
NYP19-2	1.0	0.6	2.5	0.4	1.1	0.6	56.5	0.3
NDTX14362AB-1W	1.2	1.0	2.1	0.8	1.1	0.9	57.0	0.6
B3174-4	1.7	0.6	2.7	0.4	1.1	0.5	59.5	0.3
NDTX12203AB-1W	1.2	1.0	2.1	0.8	1.1	0.9	57.0	0.6
NDTX1244-3W/Y	2.5	1.0	2.3	0.7	1.1	0.8	63.6	0.5
ATTX10333-1W/Y	2.3	0.9	2.5	0.5	1.1	0.7	53.9	0.4
NDTX14475-1W/Y	2.5	0.9	2.4	0.7	1.1	0.8	63.6	0.5
NYP103-1	1.3	0.6	2.7	0.4	1.1	0.6	57.1	0.3
WAF13066-2	1.4	0.6	2.5	0.4	1.1	0.6	59.3	0.3
NDTX1444-5W	1.1	0.9	2.4	0.5	1.1	0.7	60.6	0.4
AORTX09033-11W	1.2	1.0	2.4	0.7	1.1	0.8	54.2	0.5
AOR11484-2	1.3	0.7	2.7	0.4	1.1	0.6	59.5	0.3
NDOR13317CB-2	1.3	0.7	2.8	0.4	1.1	0.6	64.8	0.3
AORTX09037-5W/Y	1.6	1.0	2.3	0.7	1.1	0.8	59.1	0.5
W14NYQ29-5	1.0	0.7	2.7	0.4	1.1	0.6	65.1	0.4
COTX12235-2W	1.2	1.0	2.1	0.6	1.1	0.8	54.6	0.4

Clone id	Chip Color	r^2	Chip quality	r^2	Specific Gravity	r^2	Yield Ton/Ha	r^2
TX09403-21W	1.3	1.0	2.2	0.6	1.1	0.8	58.1	0.5
NYP103-22	1.0	0.6	2.6	0.4	1.1	0.6	57.9	0.3
TX09403-15W	1.3	1.0	2.2	0.6	1.1	0.8	58.1	0.5
AORTX09037-4W/Y	1.7	1.0	2.5	0.7	1.1	0.8	59.7	0.5
NDTX1482YB-1W	1.2	0.9	2.4	0.5	1.1	0.7	57.4	0.4
W15NYR11-8	1.2	0.7	2.6	0.4	1.1	0.6	59.2	0.3
NYR107-4	1.4	0.7	2.7	0.4	1.1	0.6	58.1	0.4
Snowden	1.3	1.0	2.6	0.7	1.1	0.8	62.6	0.5
AF6550-2	1.2	0.7	2.7	0.3	1.1	0.5	58.7	0.3
W13058-19	1.3	0.6	2.6	0.4	1.1	0.6	56.4	0.3
AORTX09032-3W	1.2	0.8	2.4	0.5	1.1	0.7	58.6	0.4
TX12484-3W	1.2	1.0	2.3	0.7	1.1	0.8	55.8	0.3
NYN25-1	1.4	0.6	2.7	0.4	1.1	0.6	59.2	0.3
MSW075-2	1.2	0.9	2.4	0.4	1.1	0.4	55.9	0.2
MSX540-4	1.4	0.9	2.6	0.6	1.1	0.7	55.1	0.4
NYQ29-2	1.2	0.8	2.7	0.4	1.1	0.6	66.6	0.4
NYP118-6	1.1	0.6	2.7	0.4	1.1	0.6	58.3	0.3
CO10073-7W	1.2	1.0	2.3	0.8	1.1	0.9	54.4	0.6
COTX12428-1W	1.2	1.0	2.3	0.8	1.1	0.9	54.3	0.6
NYQ38-4	1.2	0.9	2.8	0.4	1.1	0.6	62.1	0.3
WAF15184-4	1.2	0.7	2.6	0.4	1.1	0.6	58.7	0.3
NCB3171-7	1.3	0.7	2.7	0.5	1.1	0.7	57.5	0.3
NYM18-2	1.3	0.6	2.6	0.4	1.1	0.6	54.7	0.3
NYN16-10	1.3	0.6	2.7	0.4	1.1	0.6	56.8	0.3
ND124C-1	1.2	0.8	2.8	0.4	1.1	0.6	59.9	0.4
B2904-2	1.3	0.9	2.6	0.6	1.1	0.7	57.6	0.4
NDA081453CAB-2C	1.5	0.8	2.6	0.5	1.1	0.7	58.8	0.4
NY166	1.1	0.9	2.7	0.5	1.1	0.7	56.5	0.4
Mackinaw	1.2	0.9	2.7	0.5	1.1	0.7	55.2	0.4
NDOR13320CAB-2	1.5	0.7	2.5	0.4	1.1	0.5	57.0	0.3
NYN6-2	1.1	0.6	2.7	0.4	1.1	0.6	57.9	0.3
B3012-1	1.2	0.9	2.7	0.5	1.1	0.7	57.7	0.4
NYN11-4	1.2	0.7	2.8	0.4	1.1	0.6	61.9	0.4
NYQ106-13	1.6	0.8	2.8	0.4	1.1	0.6	57.5	0.4
ATX13134-3W/Y	2.9	0.9	2.5	0.6	1.1	0.7	58.1	0.4
W12078-77	1.2	0.6	2.5	0.3	1.1	0.5	52.8	0.3
AORTX09037-1W/Y	2.9	1.0	2.6	0.7	1.1	0.8	65.1	0.5
W13057-3	1.3	0.5	2.7	0.3	1.1	0.6	56.2	0.3
MSY041-1	1.2	0.7	2.8	0.4	1.1	0.7	57.6	0.3
NDTX13280CB-3W	1.2	0.9	2.5	0.6	1.1	0.7	57.3	0.4
W10659-16	1.1	0.5	2.6	0.3	1.1	0.5	55.1	0.2

Clone id	Chip Color	r^2	Chip quality	r^2	Specific Gravity	r^2	Yield Ton/Ha	r^2
MSV030-4	1.2	1.0	2.8	0.6	1.1	0.8	55.0	0.4
W15NYR5-2	1.2	0.8	2.9	0.5	1.1	0.7	64.6	0.4
ND1336-5	1.6	0.7	2.6	0.3	1.1	0.5	59.6	0.3
BNC544-2	1.5	0.7	2.8	0.4	1.1	0.7	52.7	0.3
AF6253-1	1.1	0.7	2.7	0.3	1.1	0.5	58.1	0.2
NYR5-6	1.2	0.8	2.9	0.5	1.1	0.7	64.6	0.4
CO12235-3W	1.3	0.9	2.7	0.5	1.1	0.7	56.3	0.4
Petoskey	1.1	0.9	2.8	0.6	1.1	0.8	55.0	0.4
TX12484-2W	1.3	1.0	2.5	0.7	1.1	0.8	58.4	0.3
AF4157-6	1.4	0.9	2.5	0.4	1.1	0.4	55.9	0.2
MSAFB635-15	1.3	0.9	2.9	0.5	1.1	0.7	64.1	0.4
W15NYR11-13	1.0	0.7	2.5	0.4	1.1	0.6	60.9	0.3
NDTX12135-1W	1.3	1.0	2.6	0.6	1.1	0.8	57.0	0.4
Lamoka	1.2	0.9	2.8	0.7	1.1	0.8	53.4	0.6
COOR13428-1	1.4	0.6	2.7	0.4	1.1	0.6	58.3	0.3
W12078-76	1.0	0.8	2.7	0.5	1.1	0.7	51.7	0.3
W14NYQ9-2	0.8	0.7	3.0	0.4	1.1	0.6	63.1	0.3
NYR2-2	1.2	0.7	2.7	0.4	1.1	0.6	59.2	0.3
TX12483-6W	1.6	0.8	2.6	0.5	1.1	0.7	58.7	0.3
AF6522-1	1.2	0.7	2.8	0.4	1.1	0.6	57.9	0.3
MSCC009-1	1.7	0.7	2.8	0.4	1.1	0.6	62.2	0.3
NDAF14477C-2	1.3	0.8	2.9	0.3	1.1	0.5	60.2	0.2
NYR10-4	1.2	0.7	2.9	0.4	1.1	0.5	63.8	0.3
AF5960-4	1.5	0.8	2.8	0.3	1.1	0.5	60.3	0.2
MSAA072-05	1.3	0.6	2.7	0.4	1.1	0.6	51.7	0.3
B3378-3	2.4	0.7	2.8	0.3	1.1	0.5	60.9	0.3
AOR09034-3	1.0	0.8	2.8	0.5	1.1	0.7	58.6	0.3
B3194-1	1.3	0.6	2.5	0.3	1.1	0.5	52.8	0.2
COOR13270-2	1.3	0.9	2.8	0.6	1.1	0.7	61.6	0.4
AORTX09144-2W	1.8	0.9	2.6	0.6	1.1	0.8	55.9	0.4
NYOR14Q9-5	1.1	0.9	2.9	0.5	1.1	0.7	61.0	0.4
NYR3-5	1.0	0.7	2.9	0.4	1.1	0.6	59.3	0.3
NYP114-1	1.3	0.6	2.7	0.4	1.1	0.6	53.0	0.3
BNC742-2	1.4	0.8	2.8	0.4	1.1	0.6	59.4	0.3
BNC541-5	1.6	0.6	2.9	0.4	1.1	0.6	57.5	0.3
NYR107-6	1.7	0.7	2.8	0.5	1.1	0.6	56.8	0.4
NYN44-7	0.8	0.6	2.7	0.4	1.1	0.6	57.3	0.3
AF6192-3	1.1	0.9	2.7	0.5	1.1	0.7	53.1	0.4
NYR101-2	1.2	0.7	2.8	0.4	1.1	0.6	60.2	0.3
BNC726-5	1.3	0.8	3.1	0.4	1.1	0.6	59.2	0.3
WAF15133-3	1.3	0.9	2.7	0.5	1.1	0.7	53.4	0.4

Clone id	Chip Color	r^2	Chip quality	r^2	Specific Gravity	r^2	Yield Ton/Ha	r^2
NC733-7	1.3	0.7	2.8	0.3	1.1	0.5	62.0	0.2
NYOR14Q9-9	1.3	0.8	2.9	0.4	1.1	0.6	61.4	0.3
CO11023-9W	1.2	1.0	2.5	0.6	1.1	0.8	55.2	0.5
CO11087-1W	0.9	0.7	2.8	0.5	1.1	0.6	58.8	0.4
MSX225-2	1.4	0.7	2.8	0.4	1.1	0.6	55.9	0.3
NYQ23-4	1.1	0.8	2.8	0.4	1.1	0.6	57.0	0.3
BNC472-3	1.3	0.6	2.8	0.3	1.1	0.5	58.1	0.2
BNC549-1	1.2	0.9	3.0	0.6	1.1	0.7	52.7	0.4
AF5563-2	1.4	0.7	2.8	0.5	1.1	0.7	57.1	0.4
W12082-5	1.5	0.5	2.8	0.3	1.1	0.4	56.0	0.2
CO02321-4W	1.6	0.9	2.7	0.5	1.1	0.6	57.1	0.3
WIA14067-1C	0.9	0.6	2.7	0.4	1.1	0.6	53.3	0.3
CO10076-4W	1.3	1.0	2.6	0.7	1.1	0.8	57.1	0.5
AC11453-7W	1.5	0.9	2.8	0.5	1.1	0.7	58.6	0.4
B3305-1	1.6	0.6	2.7	0.3	1.1	0.6	52.3	0.3
NDTX13224CAB-1W	1.4	0.6	2.9	0.4	1.1	0.6	59.0	0.3
TX13580-1W	1.4	0.8	2.8	0.5	1.1	0.7	58.7	0.3
AF5563-5	1.3	0.7	2.8	0.4	1.1	0.6	59.0	0.3
NDTX13315CB-1W	1.3	0.9	2.8	0.5	1.1	0.7	58.5	0.4
WAF15204-4	1.1	0.7	2.8	0.4	1.1	0.6	50.9	0.3
W15200-3	1.0	0.7	2.7	0.4	1.1	0.6	51.1	0.3
Atlantic	1.3	1.0	3.0	0.7	1.1	0.8	55.6	0.5
MSCC168-1	1.4	0.7	2.9	0.4	1.1	0.5	60.5	0.3
NYQ36-6	1.2	0.8	2.8	0.3	1.1	0.5	56.9	0.2
AF5846-3	1.4	0.6	2.8	0.3	1.1	0.5	59.0	0.2
AF6037-2	1.2	0.9	3.0	0.4	1.1	0.6	60.1	0.3
ND13221C-3	1.6	0.8	2.7	0.4	1.1	0.5	57.6	0.3
NC727-6	1.2	0.7	2.8	0.3	1.1	0.5	57.4	0.2
AF5429-3	1.0	0.7	3.0	0.4	1.1	0.6	57.0	0.3
NYORN6-8	1.2	0.8	2.8	0.4	1.1	0.5	56.1	0.3
NY168	1.3	0.9	3.0	0.4	1.1	0.6	56.6	0.3
AF6206-5	1.7	0.7	2.8	0.4	1.1	0.6	55.4	0.3
CO11023-2W	1.4	1.0	2.8	0.6	1.1	0.8	53.8	0.5
AF5973-3	1.2	0.8	3.0	0.3	1.1	0.5	56.8	0.3
B3388-3	2.2	0.7	2.8	0.3	1.1	0.5	57.9	0.2
MSV507-001	1.3	0.6	2.9	0.4	1.1	0.6	55.2	0.3
AF5933-4	1.3	0.8	3.0	0.4	1.1	0.6	58.4	0.3
MSAA324-04	1.2	0.9	2.9	0.5	1.1	0.7	53.4	0.4
ND113307C-3	1.2	0.8	2.8	0.4	1.1	0.6	58.0	0.3
ATTX07042-3W	2.7	0.9	2.6	0.5	1.1	0.7	57.6	0.3
MSBB633-18	1.4	0.9	3.0	0.4	1.1	0.6	64.4	0.4

Clone id	Chip Color	r^2	Chip quality	r^2	Specific Gravity	r^2	Yield Ton/Ha	r^2
AF5975-1	1.4	0.8	2.9	0.3	1.1	0.5	56.3	0.2
ND13217C-1	1.6	0.8	2.9	0.4	1.1	0.6	57.5	0.3
ATX13127-1Ru	1.3	0.6	2.8	0.4	1.1	0.6	53.9	0.3
NDOR1480Y-3	1.0	0.7	2.9	0.4	1.1	0.6	58.6	0.3
W14187-2	1.1	0.7	3.0	0.4	1.1	0.6	56.5	0.3
MSZ248-02	1.4	0.7	3.1	0.3	1.1	0.5	62.5	0.3
AF5938-6	1.2	0.6	2.8	0.4	1.1	0.6	55.5	0.3
AF5825-3	1.4	0.7	2.9	0.5	1.1	0.7	54.9	0.4
AF6551-1	1.1	0.7	2.8	0.4	1.1	0.6	53.4	0.4
CO12293-1W	1.2	0.9	2.7	0.5	1.1	0.7	56.4	0.4
NYP111-16	1.1	0.6	3.0	0.4	1.1	0.6	55.6	0.3
ATTX10389-1W	1.3	0.9	2.6	0.4	1.1	0.6	55.0	0.3
MSX542-2	1.3	0.5	2.9	0.3	1.1	0.5	55.8	0.2
ND13220C-3	1.3	0.9	2.9	0.5	1.1	0.6	59.9	0.4
AF6036-1	1.2	0.8	3.0	0.4	1.1	0.6	54.1	0.3
MSBB038-1	1.2	0.7	2.8	0.4	1.1	0.6	51.6	0.3
MSBB636-11	1.3	0.8	3.1	0.4	1.1	0.6	64.2	0.3
COTX13231-5W	1.3	0.9	2.4	0.5	1.1	0.7	51.5	0.4
MSAFB635-3	1.3	0.9	3.1	0.6	1.1	0.7	63.4	0.4
TX12484-1W	1.1	0.9	2.7	0.6	1.1	0.8	56.7	0.3
MSV507-003	1.4	0.7	2.9	0.5	1.1	0.7	51.1	0.4
WAF15221-2	1.5	0.7	2.9	0.3	1.1	0.5	56.3	0.2
MSAA309-15	1.2	0.8	3.0	0.5	1.1	0.7	53.7	0.4
TX14668-3W	1.5	0.9	2.7	0.6	1.1	0.8	57.2	0.3
AF6520-5	1.3	0.7	2.9	0.4	1.1	0.6	55.7	0.3
MSBB634-8	1.8	0.7	2.9	0.4	1.1	0.6	56.1	0.3
CO13232-25W	1.0	0.7	2.8	0.4	1.1	0.6	57.2	0.3
COTX13230-1W	1.3	1.0	2.5	0.5	1.1	0.7	51.4	0.4
W13069-5	1.3	0.7	2.8	0.4	1.1	0.6	50.6	0.4
AF6602-3	1.3	0.7	3.1	0.3	1.1	0.5	55.3	0.3
CO13232-5W	1.0	0.7	2.8	0.4	1.1	0.6	56.6	0.3
AF6188-9	1.2	0.9	3.0	0.5	1.1	0.7	54.3	0.4
AC01144-1W	1.2	0.8	2.6	0.5	1.1	0.7	55.4	0.4
MSBB018-1	1.3	0.7	3.0	0.3	1.1	0.5	54.6	0.3
MSZ154-1	1.2	0.6	2.8	0.3	1.1	0.5	54.7	0.2
MSZ022-16	1.4	0.7	2.9	0.5	1.1	0.7	51.4	0.4
TX14695-2W	1.2	0.8	2.6	0.5	1.1	0.7	48.4	0.2
AF6037-3	1.1	0.6	3.2	0.4	1.1	0.6	60.8	0.3
MSAA513-01	1.2	0.9	2.9	0.5	1.1	0.7	51.0	0.4
MSAA678-01	1.2	0.9	2.9	0.4	1.1	0.6	54.9	0.3
MSW501-2	0.9	0.6	2.9	0.3	1.1	0.5	56.1	0.3

Clone id	Chip Color	r^2	Chip quality	r^2	Specific Gravity	r^2	Yield Ton/Ha	r^2
W15125-4	1.3	0.7	2.9	0.4	1.1	0.5	55.4	0.3
B3317-1	1.1	0.9	3.0	0.4	1.1	0.6	56.0	0.3
NC744-4	1.3	0.7	2.9	0.3	1.1	0.5	54.9	0.2
AF6542-1	1.4	0.7	3.0	0.4	1.1	0.5	51.9	0.3
W14184-5	1.2	0.7	3.0	0.4	1.1	0.6	56.7	0.3
ATX11482-3W	1.6	0.7	2.8	0.5	1.1	0.7	56.0	0.3
MSZ063-2	1.7	0.8	3.1	0.4	1.1	0.7	57.9	0.3
AF5846-4	1.3	0.6	2.8	0.3	1.1	0.5	57.0	0.3
MSAA100-1	1.3	0.8	3.0	0.5	1.1	0.6	56.6	0.4
MSZ013-3	1.8	0.5	3.0	0.3	1.1	0.5	58.6	0.2
BNC811-9	2.2	0.7	3.1	0.4	1.1	0.5	56.4	0.3
AF6165-9	1.2	0.7	2.9	0.4	1.1	0.5	54.2	0.3
AF5677-4	1.4	0.6	2.9	0.4	1.1	0.6	56.5	0.3
MSBB058-01	1.4	0.8	3.0	0.5	1.1	0.6	51.7	0.4
MSAA100-01	1.3	0.8	3.0	0.5	1.1	0.6	56.5	0.4
NDAF14477C-7	1.3	0.8	3.0	0.3	1.1	0.5	54.3	0.3
MSW044-1	1.7	0.7	2.9	0.4	1.1	0.7	51.9	0.3
MSV507-073	1.3	0.6	3.0	0.4	1.1	0.6	51.1	0.4
COTX13231-4W	1.3	0.9	2.5	0.5	1.1	0.7	50.8	0.4
MSBB058-1	1.4	0.8	2.9	0.5	1.1	0.7	51.6	0.4
MSCC256-02	1.3	0.8	3.1	0.4	1.1	0.6	54.0	0.3
MSZ242-09	1.6	0.6	2.9	0.4	1.1	0.6	51.6	0.3
AF5819-2	1.5	0.6	2.7	0.4	1.1	0.6	57.7	0.3
AF6531-3	1.3	0.6	2.9	0.3	1.1	0.5	54.4	0.3
CO13232-11W	1.0	0.7	2.9	0.4	1.1	0.6	56.4	0.3
COTX13231-1W	1.2	1.0	2.6	0.6	1.1	0.7	52.3	0.4
NYR1-7	1.0	0.7	2.9	0.4	1.1	0.6	53.5	0.4
MSX111-3	1.5	0.6	2.8	0.4	1.1	0.6	54.2	0.3
MSCC376-1	1.2	0.8	2.9	0.5	1.1	0.6	52.2	0.4
MSBB230-01	1.9	0.9	3.1	0.4	1.1	0.6	58.1	0.3
MSZ269-18	2.0	0.6	2.9	0.4	1.1	0.5	54.2	0.3
MSAA217-3	1.3	0.9	3.0	0.5	1.1	0.6	54.3	0.4
MSX042-3	1.3	0.6	2.9	0.3	1.1	0.5	53.1	0.2
MSAA076-04	1.4	0.8	2.9	0.5	1.1	0.7	54.9	0.4
CO11048-8W	1.4	0.6	2.8	0.4	1.1	0.5	54.4	0.3
MSX472-2	1.3	0.7	2.9	0.4	1.1	0.6	54.3	0.3
MSAFB619-2	1.5	0.6	2.8	0.4	1.1	0.6	51.8	0.3
NYORQ6-6	1.0	0.8	3.1	0.4	1.1	0.6	58.0	0.3
BNC469-7	1.5	0.7	3.0	0.4	1.1	0.6	56.8	0.3
MSCC248-3	1.5	0.8	3.0	0.4	1.1	0.6	49.7	0.4
AF6027-2	1.0	0.6	2.8	0.3	1.1	0.6	55.0	0.3

Clone id	Chip Color	r^2	Chip quality	r^2	Specific Gravity	r^2	Yield Ton/Ha	r^2
MSCC248-03	1.5	0.8	3.0	0.4	1.1	0.6	49.6	0.4
AF6241-4	1.5	0.8	2.9	0.5	1.1	0.6	47.8	0.4
MSAFB614-4	1.6	0.7	3.0	0.5	1.1	0.7	52.7	0.4
MSCC129-04	1.6	0.8	3.0	0.4	1.1	0.5	48.8	0.3
MSBB166-01	1.8	0.8	3.0	0.3	1.1	0.5	54.3	0.3
CO12428-2W	1.3	0.8	2.9	0.4	1.1	0.6	52.8	0.3
AF6024-1	2.8	0.8	3.0	0.3	1.1	0.5	63.7	0.3
MSAA328-04	1.3	0.8	2.9	0.4	1.1	0.6	53.6	0.3
TX09396-1W	1.7	0.8	2.9	0.5	1.1	0.7	51.5	0.4
MSAFB609-5	1.4	0.9	2.9	0.6	1.1	0.8	51.6	0.5
AF6031-2	1.6	0.8	3.1	0.3	1.1	0.5	56.9	0.3
B3306-2	2.9	0.9	3.0	0.4	1.1	0.6	57.2	0.3
MSW324-1	1.4	0.6	2.9	0.4	1.1	0.5	56.7	0.3
WAF13076-2	2.2	0.6	3.0	0.4	1.1	0.5	59.8	0.3
NC475-3	1.3	0.8	3.2	0.3	1.1	0.5	57.9	0.3
NDTX059828-2W	1.4	0.7	2.6	0.5	1.1	0.7	56.8	0.3
MSV505-2	1.6	0.6	2.9	0.4	1.1	0.5	50.0	0.3
MSBB617-2	1.3	0.8	3.0	0.5	1.1	0.6	52.6	0.4
ND13219C-4	1.3	0.8	3.0	0.4	1.1	0.6	54.0	0.3
TX05249-10W	1.3	0.8	2.8	0.5	1.1	0.7	55.4	0.3
MSCC058-1	1.7	0.7	2.9	0.4	1.1	0.6	54.5	0.3
AF5801-1	1.7	0.5	2.9	0.3	1.1	0.5	55.9	0.2
MSBB060-01	1.5	0.8	3.0	0.4	1.1	0.6	59.3	0.3
MSAA076-6	1.4	0.7	3.0	0.5	1.1	0.6	52.4	0.4
NDTX081648CB-13W	1.5	0.8	2.9	0.5	1.1	0.7	54.3	0.3
MSBB617-02	1.4	0.8	3.0	0.5	1.1	0.6	52.5	0.4
NDTX1488-1W/Y	3.1	0.9	2.6	0.3	1.1	0.5	52.4	0.3
WAF14067-6	1.1	0.8	2.9	0.4	1.1	0.6	53.9	0.3
NC473-2	1.0	0.7	3.0	0.4	1.1	0.6	58.6	0.3
AF5583-3	1.3	0.6	3.1	0.3	1.1	0.5	54.5	0.2
MSBB625-2	1.5	0.8	3.1	0.4	1.1	0.6	55.1	0.3
AOR13125-9	1.5	0.6	3.0	0.4	1.1	0.6	50.4	0.3
MSBB625-02	1.5	0.8	3.1	0.4	1.1	0.6	55.1	0.3
MSZ251-1	1.5	0.6	2.8	0.3	1.1	0.6	51.4	0.3
MSX277-1	1.2	0.6	2.7	0.4	1.1	0.5	48.9	0.3
MSCC081-1	1.1	0.7	2.9	0.4	1.1	0.6	49.2	0.3
AF6237-3	1.0	0.7	2.9	0.4	1.1	0.6	54.9	0.3
NC470-3	1.3	0.7	3.2	0.4	1.1	0.6	56.6	0.3
MSAA373-3	2.6	0.8	3.0	0.4	1.1	0.5	56.3	0.3
MSAA091-01	1.5	0.8	3.1	0.4	1.1	0.6	52.8	0.3
AF5648-3	1.5	0.6	3.0	0.4	1.1	0.6	52.1	0.3

Clone id	Chip Color	r^2	Chip quality	r^2	Specific Gravity	r^2	Yield Ton/Ha	r^2
MSAA072-04	1.5	0.8	3.0	0.3	1.1	0.6	53.3	0.3
MSY156-2	1.5	0.7	3.0	0.5	1.1	0.7	54.8	0.4
A11516-1C	1.8	0.9	3.1	0.5	1.1	0.7	57.1	0.3
AF6200-4	1.2	0.9	3.1	0.4	1.1	0.6	54.1	0.4
AOR13137-2	1.4	0.7	3.0	0.4	1.1	0.5	54.4	0.3
B3181-7	2.7	0.5	3.1	0.3	1.1	0.5	56.2	0.2
CO11037-5W	1.3	1.0	2.9	0.6	1.1	0.8	52.3	0.5
MSCC246-07	1.8	0.8	3.0	0.4	1.1	0.6	50.5	0.3
MSBB631-04	1.2	0.8	3.1	0.3	1.1	0.5	51.6	0.3
MSX526-1	1.5	0.6	3.0	0.4	1.1	0.5	52.8	0.3
MSCC248-02	1.9	0.8	3.1	0.4	1.1	0.5	51.0	0.3
COTX15425-1Y/Y	2.9	0.8	2.7	0.3	1.1	0.4	55.8	0.2
MSAA271-05	1.2	0.6	3.1	0.4	1.1	0.6	54.8	0.3
AC11494-6W	1.4	0.9	3.0	0.5	1.1	0.7	54.4	0.4
NYL8-12	1.5	0.6	3.2	0.4	1.1	0.6	57.1	0.3
A13125-3C	1.3	0.7	3.0	0.4	1.1	0.6	52.6	0.2
MSAA240-06	1.4	0.8	2.9	0.3	1.1	0.5	53.6	0.3
NC472-1	1.3	0.7	3.2	0.4	1.1	0.6	61.0	0.3
MSCC266-1	1.2	0.7	3.2	0.4	1.1	0.6	55.1	0.4
MSAFB618-2	1.6	0.6	3.0	0.4	1.1	0.5	52.3	0.3
MSBB008-3	1.4	0.8	3.0	0.4	1.1	0.6	53.6	0.4
MSBB008-03	1.4	0.8	3.0	0.4	1.1	0.6	53.5	0.4
AF6197-8	0.9	0.7	2.9	0.4	1.1	0.5	50.7	0.3
MSZ219-14	1.3	0.9	2.9	0.5	1.1	0.5	47.2	0.3
AF6034-1	1.2	0.8	3.1	0.3	1.1	0.5	55.9	0.3
AF6236-7	1.5	0.9	3.1	0.5	1.1	0.7	62.1	0.4
AF6232-1	1.5	0.9	3.1	0.5	1.1	0.7	62.1	0.4
MSBB618-9	1.3	0.8	3.0	0.4	1.1	0.6	53.2	0.4
MSBB635-14	1.5	0.7	3.0	0.4	1.1	0.6	60.0	0.3
ATX13126-5Ru	1.6	0.6	3.0	0.3	1.1	0.6	55.4	0.2
MSBB618-09	1.3	0.8	3.0	0.4	1.1	0.6	53.1	0.4
MSAFB609-12	1.1	0.8	3.1	0.5	1.1	0.7	52.7	0.4
B3183-6	1.9	0.7	3.1	0.5	1.1	0.7	55.4	0.4
MSBB613-04	1.5	0.8	3.0	0.4	1.1	0.5	50.4	0.3
AF6603-5	1.5	0.7	3.2	0.5	1.1	0.6	54.1	0.4
MSZ157-3	1.6	0.6	3.1	0.4	1.1	0.6	57.5	0.3
MSBB193-1	2.1	0.8	2.9	0.5	1.1	0.6	52.1	0.4
MSAA725-03	1.1	0.6	3.1	0.3	1.1	0.6	53.7	0.3
MSBB047-1	1.2	0.7	3.2	0.4	1.1	0.5	50.2	0.3
AF5819-6	1.7	0.6	2.9	0.4	1.1	0.6	55.1	0.3
MSBB193-01	2.1	0.8	2.9	0.5	1.1	0.6	52.0	0.4

Clone id	Chip Color	r^2	Chip quality	r^2	Specific Gravity	r^2	Yield Ton/Ha	r^2
MSAFB626-2	1.7	0.6	3.1	0.4	1.1	0.6	53.5	0.3
MSCC129-02	1.5	0.8	3.0	0.4	1.1	0.5	48.3	0.3
NYM7-4	1.5	0.7	3.0	0.4	1.1	0.6	54.7	0.3
MSAA311-1	1.4	0.8	3.2	0.4	1.1	0.5	52.5	0.3
AF6200-7	1.3	0.7	3.0	0.4	1.1	0.6	47.0	0.3
MSX194-3	1.3	0.7	2.9	0.4	1.1	0.6	50.7	0.3
MSAA038-09Y	1.6	0.9	3.0	0.4	1.1	0.6	53.1	0.3
MSAFB614-6	1.4	0.7	3.0	0.5	1.1	0.7	48.1	0.4
MSAFB618-3	1.5	0.6	3.1	0.3	1.1	0.5	50.8	0.3
MSAFB605-4	1.2	0.9	3.3	0.6	1.1	0.8	54.0	0.5
MST437-1	1.7	0.5	3.1	0.3	1.1	0.5	55.0	0.2
BNC537-5	1.8	0.6	3.1	0.3	1.1	0.5	53.9	0.2
MSAA313-01	1.6	0.8	3.1	0.4	1.1	0.6	51.2	0.3
MSBB623-12	1.5	0.8	3.1	0.4	1.1	0.6	49.7	0.3
MSX345-6Y	1.9	0.6	2.9	0.3	1.1	0.5	52.9	0.2
MSDD085-13	1.6	0.7	3.1	0.5	1.1	0.6	52.2	0.4
MSV507-128	1.6	0.7	3.2	0.5	1.1	0.7	52.6	0.4
MSBB067-02	1.2	0.9	3.0	0.5	1.1	0.7	51.3	0.5
AFC5687-2W	1.4	0.6	3.0	0.4	1.1	0.6	56.2	0.3
NYORQ2-2	1.2	0.8	3.2	0.4	1.1	0.6	55.2	0.3
MSAA260-03	1.8	0.9	2.9	0.4	1.1	0.7	48.9	0.3
COTX14284-2W/Y	2.8	0.7	3.1	0.4	1.1	0.6	62.3	0.3
MSBB079-2	1.4	0.7	3.0	0.4	1.1	0.5	54.4	0.3
W14NYQ4-1	1.1	0.7	3.1	0.4	1.1	0.6	50.2	0.3
CO11074-1W	1.9	0.6	3.1	0.3	1.1	0.5	53.9	0.2
B3379-6	2.6	0.7	3.0	0.3	1.1	0.5	55.3	0.2
WAF15195-3	1.4	0.8	3.2	0.4	1.1	0.5	56.5	0.3
AF6616-1	1.4	0.7	3.1	0.4	1.1	0.6	55.8	0.3
B3182-1	2.8	0.6	3.2	0.3	1.1	0.5	58.9	0.3
B3381-2	2.5	0.7	3.1	0.3	1.1	0.5	54.0	0.3
NYORQ6-3	1.2	0.8	3.2	0.4	1.1	0.6	54.2	0.3
MSBB610-13	1.3	0.8	3.2	0.4	1.1	0.6	54.5	0.4
MSBB618-02	1.5	0.8	3.2	0.4	1.1	0.6	51.6	0.3
MSBB614-10	1.5	0.8	3.2	0.4	1.1	0.6	49.1	0.4
AF6614-4	2.1	0.7	3.2	0.4	1.1	0.6	53.8	0.3
MSAFB611-5	1.5	0.6	3.3	0.4	1.1	0.6	52.0	0.3
AOR12197-4	2.4	0.9	3.4	0.5	1.1	0.7	57.6	0.4
MSAA498-18	1.6	0.9	3.2	0.4	1.1	0.6	48.6	0.3
MSBB032-1	1.5	0.7	3.1	0.4	1.1	0.6	53.5	0.3
AOR13124-6	1.5	0.7	3.3	0.3	1.1	0.5	52.5	0.3
MSBB626-11	1.5	0.9	3.2	0.5	1.1	0.7	48.8	0.4

Clone id	Chip Color	r^2	Chip quality	r^2	Specific Gravity	r^2	Yield Ton/Ha	r^2
MSAA252-07	1.5	0.9	3.1	0.4	1.1	0.6	54.0	0.3
MSAA085-1	2.4	0.7	3.1	0.3	1.1	0.5	53.7	0.3
MSBB611-03	2.0	0.9	3.2	0.5	1.1	0.7	56.0	0.4
MSBB131-01	2.8	0.9	3.0	0.4	1.1	0.6	55.4	0.3
B3381-4	2.9	0.7	3.0	0.4	1.1	0.6	51.7	0.3
AF6616-4	2.3	0.7	3.2	0.4	1.1	0.6	54.4	0.3
MSAFB626-8	1.4	0.7	3.4	0.4	1.1	0.6	52.9	0.4
MSAA036-09	1.4	0.9	3.1	0.5	1.1	0.7	53.4	0.4
MSBB190-1	1.8	0.7	3.1	0.4	1.1	0.6	51.5	0.3
MSBB190-2	1.2	0.7	3.3	0.4	1.1	0.6	53.5	0.3
MSAA556-03Y	2.4	0.5	3.1	0.3	1.1	0.5	54.6	0.3
MSAA342-11Y	2.7	0.8	3.0	0.4	1.1	0.6	54.7	0.3
MSV507-012	1.6	0.7	3.3	0.5	1.1	0.7	50.5	0.4
MSAA241-01	2.0	0.9	3.1	0.5	1.1	0.7	48.0	0.4
ATX13126-3W	1.7	0.6	3.1	0.4	1.1	0.6	54.5	0.3
MSAFB610-2	1.6	0.7	3.5	0.5	1.1	0.6	53.7	0.4
MSAA290-02	2.3	0.8	3.1	0.4	1.1	0.6	50.4	0.3
NDTX1432Y-2Y/Y	3.0	0.9	3.0	0.5	1.1	0.7	55.7	0.4
MSAA498-07	1.6	0.6	3.3	0.4	1.1	0.6	49.5	0.3
MSAFB626-5	1.4	0.6	3.3	0.4	1.1	0.6	51.2	0.3
TX12474-1P/R	3.7	0.8	2.8	0.5	1.1	0.7	51.3	0.1
MSAFB610-4	1.6	0.7	3.4	0.5	1.1	0.7	47.0	0.4

C2. Standardized multiple selection index from Z score

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
NYR102-3	-1.0	-2.4	2.1	1.3	15.0	3.1
NYR102-7	-0.6	-2.0	2.5	1.5	14.7	3.0
NY169	-0.7	-2.2	1.0	2.2	14.0	2.9
NDTX1246-3W	0.1	-2.0	1.2	1.8	12.1	2.5
W13NYP19-2	0.2	-1.4	2.9	0.8	11.2	2.3
NYN24-2	0.1	-1.4	3.0	0.4	10.9	2.2
NYR102-8	-0.9	-1.7	1.5	1.0	10.9	2.2
NDTX113030C-3W	-0.4	-2.0	1.9	0.2	10.6	2.2
NYP116-6	-0.8	-1.9	1.1	0.9	10.5	2.1
NYP108-6	-0.5	-1.6	2.1	0.3	10.1	2.1
NDTX1246-5W/Y	3.1	-2.7	0.6	2.0	10.1	2.1
MSZ246-1	-0.6	-1.2	1.7	1.2	10.0	2.0

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
NYM15-3	-0.7	-1.8	1.3	0.4	9.7	2.0
NYR105-11	-1.4	-1.9	0.7	0.6	9.5	1.9
NYQ29-1	-0.2	-0.8	0.0	3.4	9.3	1.9
NYR101-7	-1.1	-1.5	-0.4	2.0	8.9	1.8
NYQ37-1	-0.6	-1.1	1.2	1.2	8.7	1.8
NYP119-3	-0.2	-1.8	1.9	-0.4	8.7	1.8
NYP19-2	-1.1	-1.3	1.6	0.2	8.7	1.8
NDTX14362AB-1W	-0.5	-2.9	-0.7	0.3	8.6	1.8
B3174-4	0.6	-0.8	2.4	1.0	8.6	1.8
NDTX12203AB-1W	-0.5	-3.0	-0.7	0.3	8.6	1.7
NDTX1244-3W/Y	2.4	-2.1	0.3	2.1	8.6	1.7
ATTX10333-1W/Y	2.1	-1.5	3.4	-0.5	8.4	1.7
NDTX14475-1W/Y	2.4	-2.0	0.2	2.1	8.4	1.7
NYP103-1	-0.3	-0.8	2.4	0.4	8.1	1.6
WAF13066-2	-0.1	-1.4	0.9	0.9	7.9	1.6
NDTX1444-5W	-0.8	-1.8	-0.4	1.3	7.9	1.6
AORTX09033-11W	-0.5	-2.0	1.1	-0.4	7.9	1.6
AOR11484-2	-0.4	-0.5	1.9	1.0	7.7	1.6
NDOR13317CB-2	-0.4	-0.2	0.8	2.4	7.6	1.5
AORTX09037-5W/Y	0.3	-2.3	-0.5	0.9	7.4	1.5
W14NYQ29-5	-1.0	-0.8	-0.5	2.5	7.4	1.5
COTX12235-2W	-0.7	-3.0	-0.9	-0.3	7.3	1.5
TX09403-21W	-0.4	-2.7	-1.3	0.6	7.2	1.5
NYP103-22	-1.2	-0.9	1.0	0.6	7.2	1.5
TX09403-15W	-0.4	-2.7	-1.4	0.6	7.1	1.5
AORTX09037-4W/Y	0.5	-1.6	0.4	1.1	7.1	1.4
NDTX1482YB-1W	-0.6	-2.0	-0.2	0.4	7.0	1.4
W15NYR11-8	-0.5	-1.2	0.5	0.9	7.0	1.4
NYR107-4	-0.1	-0.6	1.9	0.6	6.9	1.4
Snowden	-0.4	-1.1	-0.2	1.8	6.8	1.4
AF6550-2	-0.7	-0.8	1.0	0.8	6.6	1.3
W13058-19	-0.5	-1.2	1.0	0.2	6.5	1.3
AORTX09032-3W	-0.5	-1.7	-0.4	0.8	6.5	1.3
TX12484-3W	-0.7	-2.3	-0.6	0.0	6.4	1.3
NYN25-1	-0.1	-0.7	1.2	0.9	6.3	1.3
MSW075-2	-0.7	-1.9	-0.2	0.1	6.3	1.3
MSX540-4	-0.2	-1.0	1.6	-0.2	6.2	1.3
NYQ29-2	-0.6	-0.5	-1.0	2.9	6.0	1.2
NYP118-6	-0.8	-0.7	0.7	0.7	5.8	1.2
CO10073-7W	-0.5	-2.2	-0.4	-0.4	5.6	1.1

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
COTX12428-1W	-0.6	-2.2	-0.4	-0.4	5.6	1.1
NYQ38-4	-0.7	-0.1	0.5	1.7	5.5	1.1
WAF15184-4	-0.6	-0.9	0.3	0.8	5.4	1.1
NCB3171-7	-0.4	-0.7	1.1	0.5	5.4	1.1
NYM18-2	-0.5	-1.2	1.0	-0.3	5.4	1.1
NYN16-10	-0.4	-0.5	1.5	0.3	5.3	1.1
ND124C-1	-0.7	-0.2	0.9	1.1	5.3	1.1
B2904-2	-0.5	-1.2	0.1	0.5	5.2	1.1
NDA081453CAB-2C	0.2	-0.9	0.5	0.8	5.1	1.0
NY166	-0.8	-0.5	1.1	0.2	5.0	1.0
Mackinaw	-0.5	-0.5	1.5	-0.1	4.9	1.0
NDOR13320CAB-2	0.1	-1.4	0.1	0.3	4.9	1.0
NYN6-2	-0.9	-0.5	0.7	0.6	4.8	1.0
B3012-1	-0.5	-0.6	0.6	0.5	4.6	0.9
NYN11-4	-0.7	-0.3	-0.2	1.6	4.5	0.9
NYQ106-13	0.2	-0.1	1.8	0.5	4.5	0.9
ATX13134-3W/Y	3.5	-1.3	1.3	0.6	4.3	0.9
W12078-77	-0.7	-1.3	0.7	-0.8	4.3	0.9
AORTX09037-1W/Y	3.3	-1.2	-0.5	2.5	4.3	0.9
W13057-3	-0.3	-0.8	0.7	0.1	4.2	0.9
MSY041-1	-0.5	-0.3	1.0	0.5	4.2	0.9
NDTX13280CB-3W	-0.5	-1.4	-0.8	0.4	4.0	0.8
W10659-16	-0.8	-1.2	0.0	-0.2	4.0	0.8
MSV030-4	-0.7	-0.4	1.2	-0.2	3.9	0.8
W15NYR5-2	-0.7	0.3	-0.4	2.4	3.9	0.8
ND1336-5	0.4	-1.0	-0.3	1.0	3.9	0.8
BNC544-2	0.1	-0.2	2.6	-0.8	3.9	0.8
AF6253-1	-0.8	-0.7	-0.1	0.6	3.8	0.8
NYR5-6	-0.7	0.3	-0.3	2.4	3.8	0.8
CO12235-3W	-0.5	-0.5	0.7	0.1	3.8	0.8
Petoskey	-0.8	-0.3	1.2	-0.2	3.8	0.8
TX12484-2W	-0.4	-1.3	-1.0	0.7	3.8	0.8
AF4157-6	-0.2	-1.4	-0.4	0.0	3.7	0.8
MSAFB635-15	-0.3	0.2	-0.3	2.2	3.7	0.8
W15NYR11-13	-1.0	-1.3	-2.0	1.4	3.7	0.8
NDTX12135-1W	-0.4	-1.0	-0.2	0.3	3.7	0.8
Lamoka	-0.6	-0.4	1.5	-0.6	3.6	0.7
COOR13428-1	-0.1	-0.8	-0.2	0.7	3.6	0.7
W12078-76	-1.0	-0.8	1.1	-1.1	3.6	0.7
W14NYQ9-2	-1.5	0.5	-0.3	2.0	3.6	0.7

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
NYR2-2	-0.7	-0.6	-0.4	0.9	3.5	0.7
TX12483-6W	0.3	-0.9	-0.3	0.8	3.5	0.7
AF6522-1	-0.7	-0.2	0.5	0.6	3.5	0.7
MSCC009-1	0.5	-0.3	-0.3	1.7	3.4	0.7
NDAF14477C-2	-0.5	0.1	0.4	1.2	3.4	0.7
NYR10-4	-0.6	0.0	-0.9	2.2	3.2	0.6
AF5960-4	0.1	-0.1	0.3	1.2	3.2	0.6
MSAA072-05	-0.4	-0.5	1.7	-1.1	3.1	0.6
B3378-3	2.2	-0.4	0.6	1.4	3.0	0.6
AOR09034-3	-1.0	-0.1	0.1	0.8	3.0	0.6
B3194-1	-0.5	-1.5	-0.1	-0.8	3.0	0.6
COOR13270-2	-0.3	-0.2	-0.6	1.6	2.9	0.6
AORTX09144-2W	0.7	-1.1	0.1	0.0	2.8	0.6
NYOR14Q9-5	-0.8	0.0	-0.3	1.4	2.8	0.6
NYR3-5	-1.0	0.0	-0.1	1.0	2.7	0.6
NYP114-1	-0.3	-0.5	1.1	-0.7	2.7	0.6
BNC742-2	-0.1	-0.1	0.2	1.0	2.6	0.5
BNC541-5	0.4	0.2	1.4	0.5	2.6	0.5
NYR107-6	0.6	-0.3	0.8	0.3	2.6	0.5
NYN44-7	-1.5	-0.5	-0.6	0.4	2.5	0.5
AF6192-3	-0.9	-0.7	0.5	-0.7	2.5	0.5
NYR101-2	-0.6	-0.1	-0.5	1.2	2.5	0.5
BNC726-5	-0.3	1.0	1.7	0.9	2.4	0.5
WAF15133-3	-0.3	-0.6	0.7	-0.6	2.4	0.5
NC733-7	-0.3	-0.3	-1.0	1.7	2.4	0.5
NYOR14Q9-9	-0.3	0.4	0.1	1.5	2.4	0.5
CO11023-9W	-0.6	-1.3	-0.9	-0.2	2.4	0.5
CO11087-1W	-1.3	-0.2	-0.5	0.8	2.4	0.5
MSX225-2	-0.1	-0.4	0.5	0.0	2.4	0.5
NYQ23-4	-0.8	-0.3	0.1	0.3	2.4	0.5
BNC472-3	-0.3	-0.1	0.3	0.6	2.3	0.5
BNC549-1	-0.6	0.4	2.2	-0.8	2.3	0.5
AF5563-2	-0.2	-0.3	0.2	0.4	2.3	0.5
W12082-5	0.2	-0.2	0.9	0.1	2.3	0.5
CO02321-4W	0.3	-0.8	-0.3	0.4	2.3	0.5
WIA14067-1C	-1.4	-0.6	0.1	-0.6	2.0	0.4
CO10076-4W	-0.4	-1.1	-1.2	0.4	2.0	0.4
AC11453-7W	0.1	-0.4	-0.2	0.8	2.0	0.4
B3305-1	0.3	-0.6	1.1	-0.9	1.9	0.4
NDTX13224CAB-1W	-0.2	0.0	0.0	0.9	1.8	0.4

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
TX13580-1W	-0.2	-0.2	-0.2	0.8	1.8	0.4
AF5563-5	-0.4	-0.1	-0.3	0.9	1.7	0.4
NDTX13315CB-1W	-0.4	-0.2	-0.3	0.7	1.7	0.4
WAF15204-4	-0.9	-0.4	0.9	-1.3	1.6	0.3
W15200-3	-1.0	-0.5	0.7	-1.2	1.5	0.3
Atlantic	-0.4	0.5	1.3	0.0	1.4	0.3
MSCC168-1	0.0	0.0	-0.6	1.3	1.4	0.3
NYQ36-6	-0.7	-0.1	-0.1	0.3	1.4	0.3
AF5846-3	0.0	-0.4	-0.8	0.9	1.3	0.3
AF6037-2	-0.5	0.5	0.0	1.2	1.3	0.3
ND13221C-3	0.2	-0.5	-0.5	0.5	1.3	0.3
NC727-6	-0.7	-0.1	-0.4	0.4	1.2	0.2
AF5429-3	-1.0	0.6	0.6	0.3	1.2	0.2
NYORN6-8	-0.6	-0.2	-0.1	0.1	1.2	0.2
NY168	-0.4	0.4	0.8	0.2	1.1	0.2
AF6206-5	0.6	-0.2	0.6	-0.1	1.0	0.2
CO11023-2W	-0.1	-0.1	0.8	-0.5	1.0	0.2
AF5973-3	-0.5	0.6	0.9	0.3	0.9	0.2
B3388-3	1.8	-0.4	0.1	0.6	0.8	0.2
MSV507-001	-0.4	0.3	0.7	-0.2	0.8	0.2
AF5933-4	-0.4	0.4	0.1	0.7	0.7	0.2
MSAA324-04	-0.6	0.0	0.7	-0.6	0.7	0.1
ND113307C-3	-0.7	-0.1	-0.8	0.6	0.7	0.1
ATTX07042-3W	3.0	-1.1	-0.4	0.5	0.6	0.1
MSBB633-18	0.0	0.4	-1.4	2.3	0.6	0.1
AF5975-1	-0.2	0.1	0.3	0.1	0.6	0.1
ND13217C-1	0.2	0.0	0.0	0.5	0.6	0.1
ATX13127-1Ru	-0.5	-0.3	0.1	-0.5	0.6	0.1
NDOR1480Y-3	-1.0	0.3	-0.6	0.8	0.5	0.1
W14187-2	-0.8	0.5	0.4	0.2	0.5	0.1
MSZ248-02	-0.2	1.1	0.0	1.8	0.5	0.1
AF5938-6	-0.5	-0.1	-0.1	-0.1	0.5	0.1
AF5825-3	-0.2	0.2	0.7	-0.2	0.5	0.1
AF6551-1	-0.7	-0.4	-0.1	-0.6	0.5	0.1
CO12293-1W	-0.6	-0.5	-1.0	0.2	0.5	0.1
NYP111-16	-0.9	0.5	0.5	0.0	0.5	0.1
ATTX10389-1W	-0.3	-0.9	-1.0	-0.2	0.5	0.1
MSX542-2	-0.3	0.0	0.1	0.0	0.4	0.1
ND13220C-3	-0.4	0.1	-1.0	1.1	0.4	0.1
AF6036-1	-0.5	0.6	1.2	-0.4	0.3	0.1

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
MSBB038-1	-0.7	-0.1	0.8	-1.1	0.3	0.1
MSBB636-11	-0.4	1.1	-0.7	2.3	0.3	0.1
COTX13231-5W	-0.3	-1.8	-1.7	-1.1	0.3	0.1
MSAFB635-3	-0.4	1.1	-0.5	2.0	0.2	0.0
TX12484-1W	-0.8	-0.7	-1.6	0.3	0.2	0.0
MSV507-003	-0.1	0.1	1.3	-1.2	0.1	0.0
WAF15221-2	0.0	0.1	0.1	0.1	0.0	0.0
MSAA309-15	-0.7	0.6	1.1	-0.5	0.0	0.0
TX14668-3W	0.0	-0.7	-1.4	0.4	-0.1	0.0
AF6520-5	-0.5	0.1	-0.1	0.0	-0.1	0.0
MSBB634-8	0.8	0.3	0.6	0.1	-0.2	0.0
CO13232-25W	-1.2	-0.3	-1.5	0.4	-0.2	0.0
COTX13230-1W	-0.3	-1.5	-1.3	-1.2	-0.2	0.0
W13069-5	-0.3	-0.1	0.9	-1.4	-0.2	0.0
AF6602-3	-0.3	1.0	1.3	-0.1	-0.3	-0.1
CO13232-5W	-1.0	-0.3	-1.3	0.2	-0.3	-0.1
AF6188-9	-0.6	0.5	0.6	-0.4	-0.3	-0.1
AC01144-1W	-0.5	-1.2	-2.1	-0.1	-0.4	-0.1
MSBB018-1	-0.4	0.5	0.7	-0.3	-0.5	-0.1
MSZ154-1	-0.7	-0.4	-0.9	-0.3	-0.5	-0.1
MSZ022-16	-0.2	0.0	0.8	-1.2	-0.5	-0.1
TX14695-2W	-0.7	-1.0	-0.3	-2.0	-0.6	-0.1
AF6037-3	-0.7	1.2	-0.3	1.3	-0.7	-0.1
MSAA513-01	-0.6	0.1	0.8	-1.3	-0.7	-0.1
MSAA678-01	-0.5	0.0	-0.4	-0.2	-0.7	-0.1
MSW501-2	-1.3	0.3	-0.6	0.1	-0.7	-0.1
W15125-4	-0.3	0.3	0.1	-0.1	-0.7	-0.1
B3317-1	-0.7	0.4	-0.2	0.1	-0.8	-0.2
NC744-4	-0.3	0.1	-0.3	-0.2	-0.9	-0.2
AF6542-1	-0.1	0.7	1.6	-1.0	-0.9	-0.2
W14184-5	-0.7	0.5	-0.4	0.2	-0.9	-0.2
ATX11482-3W	0.4	-0.3	-0.8	0.1	-0.9	-0.2
MSZ063-2	0.5	0.9	0.6	0.6	-1.0	-0.2
AF5846-4	-0.3	-0.4	-1.6	0.3	-1.0	-0.2
MSAA100-1	-0.4	0.7	0.1	0.2	-1.0	-0.2
MSZ013-3	0.7	0.8	0.2	0.8	-1.1	-0.2
BNC811-9	1.7	1.0	1.6	0.2	-1.2	-0.2
AF6165-9	-0.6	0.3	0.0	-0.4	-1.2	-0.2
AF5677-4	0.0	0.0	-0.8	0.2	-1.2	-0.2
MSBB058-01	-0.2	0.4	0.9	-1.1	-1.2	-0.2

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
MSAA100-01	-0.4	0.7	0.1	0.2	-1.2	-0.2
NDAF14477C-7	-0.5	0.5	0.3	-0.4	-1.2	-0.2
MSW044-1	0.5	0.0	0.7	-1.0	-1.2	-0.3
MSV507-073	-0.4	0.4	1.1	-1.2	-1.3	-0.3
COTX13231-4W	-0.3	-1.4	-1.6	-1.3	-1.3	-0.3
MSBB058-1	-0.2	0.4	0.9	-1.1	-1.3	-0.3
MSCC256-02	-0.3	0.9	0.9	-0.5	-1.4	-0.3
MSZ242-09	0.3	0.2	0.8	-1.1	-1.5	-0.3
AF5819-2	0.1	-0.7	-2.3	0.5	-1.5	-0.3
AF6531-3	-0.4	0.2	-0.2	-0.4	-1.5	-0.3
CO13232-11W	-1.0	0.2	-1.2	0.2	-1.5	-0.3
COTX13231-1W	-0.5	-1.1	-1.7	-0.9	-1.6	-0.3
NYR1-7	-1.2	0.1	-0.7	-0.6	-1.6	-0.3
MSX111-3	0.0	-0.3	-0.9	-0.4	-1.6	-0.3
MSCC376-1	-0.6	0.1	0.0	-0.9	-1.7	-0.3
MSBB230-01	1.1	1.1	0.8	0.6	-1.7	-0.3
MSZ269-18	1.2	0.3	0.7	-0.4	-1.7	-0.3
MSAA217-3	-0.3	0.7	0.5	-0.4	-1.7	-0.4
MSX042-3	-0.3	0.1	-0.2	-0.7	-1.8	-0.4
MSAA076-04	0.0	0.2	-0.4	-0.2	-1.8	-0.4
CO11048-8W	-0.1	-0.3	-1.0	-0.4	-1.9	-0.4
MSX472-2	-0.3	0.1	-0.5	-0.4	-1.9	-0.4
MSAFB619-2	0.2	0.0	0.1	-1.1	-1.9	-0.4
NYORQ6-6	-1.0	0.9	-0.7	0.6	-1.9	-0.4
BNC469-7	0.1	0.5	-0.5	0.3	-1.9	-0.4
MSCC248-3	0.2	0.7	1.9	-1.6	-1.9	-0.4
AF6027-2	-1.0	-0.1	-1.5	-0.2	-1.9	-0.4
MSCC248-03	0.2	0.7	1.8	-1.6	-1.9	-0.4
AF6241-4	0.1	0.3	1.7	-2.1	-1.9	-0.4
MSAFB614-4	0.3	0.5	0.7	-0.8	-1.9	-0.4
MSCC129-04	0.2	0.4	1.6	-1.9	-2.0	-0.4
MSBB166-01	0.8	0.7	0.8	-0.4	-2.0	-0.4
CO12428-2W	-0.4	0.2	-0.1	-0.8	-2.0	-0.4
AF6024-1	3.1	0.6	-0.7	2.1	-2.1	-0.4
MSAA328-04	-0.3	0.4	-0.1	-0.6	-2.2	-0.4
TX09396-1W	0.5	0.3	0.7	-1.1	-2.2	-0.5
MSAFB609-5	-0.2	0.4	0.4	-1.1	-2.3	-0.5
AF6031-2	0.5	1.0	0.2	0.3	-2.3	-0.5
B3306-2	3.5	0.5	0.9	0.4	-2.3	-0.5
MSW324-1	-0.2	0.3	-1.0	0.3	-2.4	-0.5

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
WAF13076-2	1.8	0.6	-0.5	1.1	-2.4	-0.5
NC475-3	-0.4	1.2	-0.2	0.6	-2.4	-0.5
NDTX059828-2W	-0.1	-0.9	-2.9	0.3	-2.4	-0.5
MSV505-2	0.3	0.1	0.6	-1.5	-2.4	-0.5
MSBB617-2	-0.3	0.6	0.3	-0.8	-2.5	-0.5
ND13219C-4	-0.4	0.6	-0.1	-0.5	-2.5	-0.5
TX05249-10W	-0.4	-0.1	-1.5	-0.1	-2.5	-0.5
MSCC058-1	0.5	0.2	-0.3	-0.3	-2.5	-0.5
AF5801-1	0.5	0.2	-0.8	0.0	-2.6	-0.5
MSBB060-01	0.1	0.7	-1.2	1.0	-2.6	-0.5
MSAA076-6	-0.2	0.8	0.6	-0.9	-2.6	-0.5
NDTX081648CB-13W	0.2	0.2	-0.6	-0.4	-2.6	-0.5
MSBB617-02	-0.3	0.6	0.4	-0.9	-2.6	-0.5
NDTX1488-1W/Y	3.8	-1.2	-0.4	-0.9	-2.7	-0.6
WAF14067-6	-0.7	0.3	-0.9	-0.5	-2.8	-0.6
NC473-2	-1.1	0.6	-1.8	0.8	-2.8	-0.6
AF5583-3	-0.5	0.8	-0.1	-0.3	-2.9	-0.6
MSBB625-2	0.1	1.1	0.4	-0.2	-2.9	-0.6
AOR13125-9	0.2	0.6	0.9	-1.4	-3.0	-0.6
MSBB625-02	0.0	1.1	0.4	-0.2	-3.0	-0.6
MSZ251-1	0.0	-0.1	-0.6	-1.2	-3.0	-0.6
MSX277-1	-0.5	-0.6	-0.8	-1.8	-3.0	-0.6
MSCC081-1	-0.9	0.0	-0.3	-1.7	-3.0	-0.6
AF6237-3	-1.1	0.3	-1.3	-0.2	-3.1	-0.6
NC470-3	-0.3	1.2	-0.1	0.2	-3.1	-0.6
MSAA373-3	2.6	0.5	0.3	0.1	-3.1	-0.6
MSAA091-01	0.1	1.1	0.9	-0.8	-3.2	-0.6
AF5648-3	0.2	0.7	0.4	-1.0	-3.3	-0.7
MSAA072-04	0.1	0.7	0.1	-0.6	-3.3	-0.7
MSY156-2	0.2	0.7	-0.3	-0.2	-3.3	-0.7
A11516-1C	0.9	0.9	-0.3	0.4	-3.3	-0.7
AF6200-4	-0.7	1.1	0.0	-0.4	-3.4	-0.7
AOR13137-2	-0.1	0.7	-0.4	-0.4	-3.5	-0.7
B3181-7	2.9	0.8	0.8	0.1	-3.5	-0.7
CO11037-5W	-0.3	0.1	-0.8	-0.9	-3.5	-0.7
MSCC246-07	0.9	0.7	1.2	-1.4	-3.5	-0.7
MSBB631-04	-0.7	0.9	0.3	-1.1	-3.5	-0.7
MSX526-1	0.1	0.6	-0.1	-0.8	-3.5	-0.7
MSCC248-02	1.0	1.1	1.6	-1.3	-3.6	-0.7
COTX15425-1Y/Y	3.4	-0.8	-1.3	0.0	-3.6	-0.7

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
MSAA271-05	-0.6	0.9	-0.5	-0.2	-3.7	-0.7
AC11494-6W	-0.2	0.7	-0.6	-0.4	-3.7	-0.8
NYL8-12	0.2	1.3	-0.2	0.4	-3.7	-0.8
A13125-3C	-0.3	0.7	-0.1	-0.8	-3.7	-0.8
MSAA240-06	-0.1	0.2	-1.1	-0.6	-3.8	-0.8
NC472-1	-0.4	1.2	-1.6	1.4	-3.8	-0.8
MSCC266-1	-0.7	1.5	0.2	-0.2	-3.8	-0.8
MSAFB618-2	0.2	0.8	0.3	-0.9	-3.8	-0.8
MSBB008-3	-0.2	0.6	-0.5	-0.6	-3.8	-0.8
MSBB008-03	-0.2	0.6	-0.5	-0.6	-3.8	-0.8
AF6197-8	-1.2	0.0	-1.1	-1.3	-3.8	-0.8
MSZ219-14	-0.4	0.0	0.2	-2.3	-3.8	-0.8
AF6034-1	-0.6	1.1	-0.6	0.0	-3.8	-0.8
AF6236-7	0.1	1.1	-2.0	1.7	-3.9	-0.8
AF6232-1	0.0	1.1	-2.0	1.7	-3.9	-0.8
MSBB618-9	-0.4	0.7	-0.4	-0.7	-4.0	-0.8
MSBB635-14	0.2	0.6	-2.1	1.1	-4.0	-0.8
ATX13126-5Ru	0.4	0.6	-0.8	-0.1	-4.0	-0.8
MSBB618-09	-0.3	0.7	-0.4	-0.7	-4.1	-0.8
MSAFB609-12	-0.8	0.8	-0.5	-0.8	-4.2	-0.9
B3183-6	1.2	0.9	0.0	-0.1	-4.2	-0.9
MSBB613-04	0.1	0.5	0.1	-1.4	-4.2	-0.9
AF6603-5	0.1	1.3	0.2	-0.4	-4.3	-0.9
MSZ157-3	0.5	0.8	-1.2	0.5	-4.3	-0.9
MSBB193-1	1.5	0.2	-0.3	-1.0	-4.5	-0.9
MSAA725-03	-0.9	0.9	-0.9	-0.5	-4.5	-0.9
MSBB047-1	-0.6	1.2	0.7	-1.5	-4.6	-0.9
AF5819-6	0.6	0.0	-1.8	-0.2	-4.6	-0.9
MSBB193-01	1.5	0.2	-0.3	-1.0	-4.6	-0.9
MSAFB626-2	0.5	1.1	0.2	-0.6	-4.7	-1.0
MSCC129-02	0.0	0.6	0.5	-2.0	-4.7	-1.0
NYM7-4	0.0	0.7	-1.1	-0.3	-4.7	-1.0
MSAA311-1	0.0	1.3	0.4	-0.9	-4.7	-1.0
AF6200-7	-0.3	0.5	0.5	-2.3	-4.8	-1.0
MSX194-3	-0.3	0.3	-0.8	-1.4	-4.8	-1.0
MSAA038-09Y	0.3	0.7	-0.6	-0.7	-4.9	-1.0
MSAFB614-6	-0.2	0.5	0.2	-2.0	-4.9	-1.0
MSAFB618-3	0.1	1.0	0.4	-1.3	-4.9	-1.0
MSAFB605-4	-0.6	1.7	0.2	-0.5	-5.0	-1.0
MST437-1	0.6	1.1	-0.3	-0.2	-5.1	-1.0

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
BNC537-5	0.7	0.9	-0.4	-0.5	-5.1	-1.0
MSAA313-01	0.3	1.1	0.4	-1.2	-5.2	-1.1
MSBB623-12	0.1	0.8	0.2	-1.6	-5.2	-1.1
MSX345-6Y	0.9	0.2	-1.1	-0.8	-5.2	-1.1
MSDD085-13	0.2	1.1	0.1	-0.9	-5.3	-1.1
MSV507-128	0.4	1.4	0.4	-0.8	-5.4	-1.1
MSBB067-02	-0.5	0.6	-0.8	-1.2	-5.4	-1.1
AFC5687-2W	0.0	0.6	-1.9	0.1	-5.4	-1.1
NYORQ2-2	-0.6	1.2	-1.1	-0.1	-5.4	-1.1
MSAA260-03	0.9	0.2	-0.1	-1.8	-5.5	-1.1
COTX14284-2W/Y	3.2	1.2	-1.1	1.7	-5.5	-1.1
MSBB079-2	-0.2	0.7	-1.4	-0.3	-5.5	-1.1
W14NYQ4-1	-1.0	1.0	-0.4	-1.5	-5.7	-1.2
CO11074-1W	1.0	0.8	-0.7	-0.5	-5.7	-1.2
B3379-6	2.6	0.7	-0.5	-0.1	-5.9	-1.2
WAF15195-3	0.0	1.3	-1.2	0.2	-5.9	-1.2
AF6616-1	-0.1	1.1	-1.5	0.0	-6.0	-1.2
B3182-1	3.2	1.2	-0.5	0.9	-6.0	-1.2
B3381-2	2.5	1.2	0.4	-0.5	-6.1	-1.2
NYORQ6-3	-0.5	1.4	-0.9	-0.4	-6.1	-1.2
MSBB610-13	-0.3	1.4	-0.8	-0.3	-6.1	-1.3
MSBB618-02	0.1	1.2	-0.2	-1.1	-6.3	-1.3
MSBB614-10	0.0	1.4	0.5	-1.8	-6.6	-1.3
AF6614-4	1.5	1.5	0.1	-0.5	-6.7	-1.4
MSAFB611-5	0.2	1.6	0.1	-1.0	-6.8	-1.4
AOR12197-4	2.3	2.0	0.3	0.5	-6.8	-1.4
MSAA498-18	0.3	1.4	0.7	-1.9	-6.9	-1.4
MSBB032-1	0.1	1.0	-1.3	-0.6	-6.9	-1.4
AOR13124-6	0.0	1.8	0.1	-0.9	-6.9	-1.4
MSBB626-11	0.2	1.3	0.4	-1.8	-7.0	-1.4
MSAA252-07	0.0	1.1	-1.4	-0.5	-7.0	-1.4
MSAA085-1	2.2	1.0	-0.5	-0.5	-7.2	-1.5
MSBB611-03	1.2	1.6	-0.7	0.1	-7.3	-1.5
MSBB131-01	3.2	0.7	-0.9	-0.1	-7.3	-1.5
B3381-4	3.3	0.6	-0.2	-1.1	-7.8	-1.6
AF6616-4	2.0	1.4	-0.5	-0.4	-8.0	-1.6
MSAFB626-8	-0.2	2.1	-0.2	-0.8	-8.1	-1.6
MSAA036-09	-0.2	1.1	-1.9	-0.6	-8.2	-1.7
MSBB190-1	0.8	1.1	-1.0	-1.1	-8.3	-1.7
MSBB190-2	-0.6	1.9	-1.0	-0.6	-8.4	-1.7

Clone id	Chip Color (z)	Chip Quality (z)	Specific gravity (z)	Yield (z)	Multiple index selection (MIS)	Z_{MIS}
MSAA556-03Y	2.1	1.1	-1.2	-0.3	-8.4	-1.7
MSAA342-11Y	2.9	0.4	-2.1	-0.3	-8.7	-1.8
MSV507-012	0.2	1.7	-0.4	-1.4	-8.8	-1.8
MSAA241-01	1.2	0.9	-0.5	-2.1	-9.0	-1.8
ATX13126-3W	0.5	0.9	-2.6	-0.3	-9.0	-1.8
MSAFB610-2	0.3	2.5	-0.3	-0.5	-9.6	-2.0
MSAA290-02	1.9	1.1	-0.8	-1.4	-9.7	-2.0
NDTX1432Y-2Y/Y	3.7	0.5	-2.5	0.0	-10.1	-2.1
MSAA498-07	0.4	2.0	-0.4	-1.7	-10.5	-2.1
MSAFB626-5	-0.1	1.8	-1.9	-1.2	-11.5	-2.4
TX12474-1P/R	5.3	0.0	-2.4	-1.2	-12.2	-2.5
MSAFB610-4	0.3	2.3	-1.8	-2.3	-15.5	-3.2