HIGH DENSITY MAPPING AND QUANTITATIVE TRAIT LOCI ANALYSIS FOR FUNGAL DISEASES IN VITIS AESTIVALIS-DERIVED ‘NORTON’ AND BREEDING FOR LOW PALMITIC ACID IN GLYCINE MAX

A Dissertation

Presented to

The Faculty of the Graduate School

At the University of Missouri-Columbia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

By

SURYA DATTA SAPKOTA

Dr. Chin-Feng Hwang, Dissertation Supervisor

December 2017

© Copyright by Surya Datta Sapkota 2017

All Rights Reserved

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

HIGH DENSITY MAPPING AND QUANTITATIVE TRAIT LOCI ANALYSIS FOR FUNGAL DISEASES IN VITIS AESTIVALIS-DERIVED 'NORTON' AND BREEDING FOR LOW PALMITIC ACID IN GLYCINE MAX

Presented by Surya Datta Sapkota

A candidate for the degree of

DOCTOR OF PHILOSOPHY OF PLANT BREEDING, GENETICS AND GENOMICS

And hereby certify that, in their opinion, it is worthy of acceptance.

____

Dr. Chin-Feng Hwang

____

Dr. Kristin D. Bilyeu

____

Dr. James T. English

____

Dr. James E. Schoelz

____

Dr. Dong Xu

DEDICATION

To

Khanda Nanda Sapkota (Father),

Chandrakali Sapkota (Mother),

Chetu Kumari Ayer Sapkota (Wife), and

Family members

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Chin-Feng Hwang for not only being an excellent mentor for my PhD but also for his continuous support, advising, and encouragement during the period. I am sincerely thankful for the opportunity to join the research group, without which my PhD would not have completed.

I would like to acknowledge my committee members; Drs. Kristin Bilyeu, James English, Jim Schoelz, and Dong Xu for their suggestions and support during the period. My especial thanks to Dr. Bilyeu for providing me an opportunity to work in her lab, and for the guidance and suggestions during my PhD. I am also thankful to Li-Ling Chen (Research Specialist II, Missouri State University) for her continuous guidance, support and feedback. Many thanks to Christie Cole, Paul Little, Drs. Jo and Gillman, members from Dr. Bilyeu lab, for their support during my research.

I also like to thank current and past lab members of Dr. Hwang lab for helpful insights in various aspects of my research. My especial thanks to Steven J Schneider for his constructive comments and suggestions on manuscript, and to the State Fruit Experiment Station, Missouri State University, Mountain Grove MO for providing a pleasant ambience to do research.

Lastly, I would like to thank my parents, family members, and my beautiful wife without which I could not accomplish this. Thank you very much for your inspiration, patience and the support.

Chetu, I am blessed to be the part of your life, and you were always there to share my excitement and sorrow. Your support always motivates me, and brings happiness to my life.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ……………………………………………………...……….ii

LIST OF FIGURES ………………………………………………………...……….v

LIST OF TABLES vii

CHAPTER 1

LITERATURE REVIEW….. ……………………………………………………….1

1.1 References………………………………………………………….....29

CHAPTER 2

A PHENOTYPIC STUDY OF BOTRYTIS BUNCH ROT RESISTANCE IN *VITIS AESTIVALIS*-DERIVED 'NORTON' GRAPE……………………………..…...41

2.1 Figure Legends……………………………………………………….....48

2.2 References………………………………………………………….......53

CHAPTER 3

CONSTRUCTING A GENETIC LINKAGE MAP OF *VITIS AESTIVALIS*-DERIVED 'NORTON' USING SIMPLE SEQUENCE REPEAT MARKERS………..…...55

3.1 Introduction……………………………………………………………56

3.2 Materials and Methods…………………………………………...……58

3.3 Results……………………………………………………………..…...61

3.4 Discussion……………………………………………………………..63

3.5 Figure Legends……………………………………………...……...66

3.6 References…………………………...……………………………...74

CHAPTER 4

CONSTRUCTION OF HIGH DENSITY LINKAGE MAPS AND DETECTION OF DOWNY MILDEW RESISTANCE LOCUS IN A *VITIS AESTIVALIS*-DERIVD 'NORTON'…………………………………………………..……………..80

4.1 Introduction……………………………………………………...….81

4.2 Materials and Methods……………………………………………………...84
4.3 Results…………………………………………...…………………….91
4.4 Discussion……………………………………..……………….....94
4.5 Figure Legends…………………………………….……….............99
4.6 References………………………………………………………....120

CHAPTER 5
IDENTIFICATION OF NEW FATB1A MUTATIONS IN MUTANT SOYBEAN
LINES WITH LOW PALMITIC ACID IN THE SEED OIL ………………....126
5.1 Introduction………………………………………………………......127
5.2 Materials and Methods……………………………………...……..134
5.3 Results…………………………………………………...……...137
5.4 Discussion……………………………………………...…….139
5.5 Figure Legends…………………………………………..….........142
5.6 References……………………………………………..………..151

APPENDIX…………………………………………………………………...…..156

VITA……………………………………………………………………………170

LIST OF FIGURES

2.1 Comparisons of (A) disease incidence and (B) disease severity between 'Norton' and 'Cabernet Sauvignon' (CS) in 2012 and 2013…………………………………...50

2.2 Botrytis bunch rot exhibited by two different grape genotypes under laboratory conditions: 'Norton' (left) and 'Cabernet Sauvignon' (right)……………..……..51

2.3 Weather conditions (A) Relative humidity and (B) Temperature during phenotyping in 2012 and 2013……………………………………………..…………..52

3.1 SSR marker-based genetic map of *Vitis aestivalis*-derived 'Norton (N)', consensus (C) and *V. vinifera* 'Cabernet Sauvignon (CS)'………………………..……...70

4.1 Segregation of downy mildew disease in the mapping population obtained by crossing Norton with Cabernet Sauvignon……………………………..…....103

4.2 Marker segregation types (SSRs and SNPs) based on JoinMap4.1 format used for developing combined linkage map by combining both parental maps……..…105

4.3 SNPs and SSRs distribution across 19 linkage groups of grapevines derived by crossing *V. aestivalis*-derived 'Norton' with *V. vinifera* 'Cabernet Sauvignon'106

4.4 Localization of genetic determinants of downy mildew resistance in *V. aestivalis*-derived 'Norton' using SSR markers………………………..…………..…...107

4.5 Combined (SSRs and SNPs) *V. aestivalis*-derived 'Norton' and *V. vinfiera* 'Cabernet Sauvignon' linkage maps…………………………..………………………..109

4.6 Marker-Trait association using Generalized Linear Model (GLM) in TASSE.118

5.1 Fatty acid distribution among different mutant soybean lines……………………...145

5.2 Sequence alignment of nucleotides between 'Williams 82 (W82)' and FATB1a mutant in line HUDSON 17448.1 A) Detection of SNP B) Amino acid change, and C) Web logo output…………………………………..……...................146

5.3 Sequence alignment of nucleotides between 'Williams 82' (W82) and FATB1a mutant in line HUDSON 19668 A) Detection of SNP B) Amino acid change, and, C) Web logo output…………………………………………………………......147

5.4 Sequence alignment of nucleotides between 'Williams 82' (W82) and FATB1a mutant in line 1668 A) Detection of SNP B) Amino acid change, and C) Web logo output…………………………………………………………………..148

5.5 Genomic sequence of FATB1a gene………………………………………..……...149

5.6 Segregation of palmitic acid content in a mapping population generated by crossing HUDSON lines with LG046000……………………………………………..150

LIST OF TABLES

2.1 Average disease incidence and severity between two grape cultivars; Norton and Cabernet Sauvignon (CS) under different stages…………………………….....49

3.1 SSR markers tested for polymorphism and segregation pattern on the Norton (N) and Cabernet Sauvignon (CS) mapping population………..………………………...67

3.2 Distribution of SSR markers among Norton, Cabernet Sauvignon and Consensus maps and the map coverage…………………………………..……...........…68

3.3. Genome coverage and genetic distance comparisons between *Vitis aestivalis*-derived-'Norton' and *V. Vinifera* reference linkage maps………..…………….69

4.1 Genetic map from F1 population derived from the cross *V. aestivalis*-derived 'Norton' with *V. vinifera* 'Cabernet Sauvignon' (CS)……………………………………100

4.2 Summary of the QTLs in F1 population derived from the cross *V. aestivalis*-derived 'Norton' with *V. vinifera* 'Cabernet Sauvignon'………………..………………101

4.3 A comparison between previously constructed SSR based genetic map and SSR plus SNP integrated map……………………………………...………………..102

4.4 Pearson correlation coefficient of the phenotype between assays for years 2014 and 2015…………………………………………………………………………..119

5.1 Primer sets used to amplify the FATB1a gene in mutant lines…………………..…143

5.2 Average fatty acid content in mutant and control lines from a trail in Columbia MO for year …2016……………………………………………………...……..144

CHAPTER 1

LITERATURE REVIEW

Introduction

Grapevines are woody perennials in the genus *Vitis* L. that have been under cultivation for centuries. Archeological records suggest cultivation of domesticated grapevines- *Vitis vinifera* subsp. *vinifera* dated back 6,000-8,000 years ago in the Near East from its wild progenitor, *V. vinifera* subsp. sylvestries. The genus *Vitis* is diverse and consists of 60 inter-fertile wild *Vitis* species distributed in Asia, North America and Europe under different climatic conditions. Cultivated grapevine species- *V. vinifera* alone have more than 6,000 cultivars of which only less than 400 are of commercial economic importance (This et al. 2004).

Grapevines are one of the most important horticulture crops with great economic importance. Its fruit mostly transforms into valuable beverages such as wine and spirits, in addition are a source of fresh fruits, juice and raisins. According to the Food and Agriculture Organization of United Nations, (http://www.fao.org/faostat/en/#data/QC) 2014, production of grapes reached ~ 70 million tons, and the area harvested over ~ 8 million hectares. The world's viticulture is mostly dominated by the single most cultivated species- *V. vinifera* (> 98% of total wine production) followed by hybrids that have *V. vinifera* parentage. This dominant species is mostly cultivated for wine and fresh grapes while the rest is used for breeding rootstocks and disease resistance. For example, North American *V. rupestries*, *V. riparia,* or *V. berlandieri*, are used as breeding materials for rootstocks due to their resistance against grapevine pathogens, such as phylloxera and mildews (Terral et al. 2009).

Traditional grape breeding is long-term and a labor intensive process. Furthermore, most of the traits in grapevines that are of economic importance are

complex and quantitatively inherited. These traits are difficult to incorporate into commercially cultivated grapevines just by using classical breeding. These hindrances signify the importance of molecular breeding, and the use of modern biotechnological tools to expedite grape breeding programs. The use of molecular marker techniques expands the possibilities for understanding the genetics of trait inheritance, mapping, gene tagging, and ultimately the ability to use beneficial genes in grape breeding (Eibach et al. 2007).

Vitis aestivalis-derived 'Norton' Grape

In general, cultivated grapevines are divided into 3 groups: a) North American cultivars- derived from native North American grape species that include cultivars derived from *V. labrusca* (Concord, Niagara, Catawba, Delaware), *V. aestivalis* (Norton/Cynthiana), and other grape species (Wolf 2008), b) mixed hybrid cultivars- grapes that have complex parentage consisting of North American hybrids, French hybrids, and *V. vinifera* cultivars; for example, Cayuga White, Reliance, Chardonel, Seyval Blanc, Vidal Blanc, Vignoles, Chambourcin, and many others (Wolf 2008), and c) *V. vinifera*, the most dominant commercial species, is grown in areas adjacent to large bodies of water, area with mild winters and moderate weather patterns (Mehmel and Heerden 2010). *V. aestivalis*-derived 'Norton', a grape of American origin developed by Dr. Daniel Norborne Norton as a hybrid seedling in a Virginia garden, is a popular wine grape in the Midwest and Eastern United states. Norton is believed to be the result of an unintentional cross between the lost grapevine 'Bland' (a cultivar hypothesized to have *V. labrusca* and *V. vinifera* parentage) and the native grapevine, *V. aestivalis* (Ambers

2013). Microsatellite or simple sequence repeat (SSR) marker analysis showed that Norton contains alleles from the grape species *V. vinifera* and *V. aestivalis.* This supports the possibility that Norton is a *V. aestivalis* and *V. vinifera* hybrid (Stover et al. 2009). The names Norton and Cynthiana have been used interchangeably for the same cultivars. In Missouri, it is referrd to as Norton whereas in Arkansas, it is referred to as Cynthiana. Despite reports of phenotypic differnces between these two cultivars, a genome-wide SSR marker analysis showed both cultivars are acutally the same (Harris 2012; Hammers et al. 2017).

V. aestivalis-derived 'Norton' is an attractive cultivar for viticulturist and wine growers for its quality of grapes, disease resistance and hardy nature. Norton is resistant to most of the fungal diseases such as powdery mildew, downy mildew and Botrytis bunch rot, which cause extensive damages in the vineyards (Sapkota et al. 2015). It also displays tolerance to Pierce's disease (Kamas et al. 2000) and is resistant to Phylloxera (Hedrick 1908). In spite of its great disease resistance characteristics, it is difficult to propagate from hardwood cuttings (Keeley et al. 2003) and is highly sensitive to sulfur spray (Bordelon et al. 2007).

Norton produce a dry, red wine with high titratable acidity (8.5 to 13 g/L), malate (~ 6g/L), and potassium (~6g/L) as well as has a high pH (>3.5). The malic acid content is 2-3 times higher than in other grapes, which is a problem for winemakers (Main and Morris 2004). This grape has gained considerable importance in the Midwest and Eastern United states due to its various useful characteristics. In Missouri, Norton is the major cultivated grape comprising 16.1% of the total acreage (274.9 acres) (Missouri Grape and Wine Facts, 2014).

Flowering in Grapevines

Though the majority of the cultivated grapevines are hermaphrodite, wild grapevines (*V. vinifera* ssp. sylvestris and *Vitis* sp.) are generally dioecious plants, requiring cross-pollination for fertilization and fruit set. Grapevines take several years to transit from the vegetative to the reproductive stage; however, viticulture practices and environmental conditions can alter the timeline. Practices like changing viticulture managements and environmental conditions to induce early flowering are also common (Srinivasan and Mullins 1979). Formation of inflorescence and flowers in grapevines is characterized in a well-defined three-step process: a) formation of anlagen: these are the meristematic protuberances that arise from apices of latent buds. Anlagen are uncommitted primordia which later change into inflorescence primordial, tendril primordial or shoot primordial, b) formation of inflorescence primordial: these are anlagen that undergo repeated branching to form a conical structure of branched primordial and, c) formation of flowers: this step consists of differentiation of inflorescence primordial to form an individual flower. Inflorescence and tendrils are homologous organs and either one may develop from the primordial organ depending upon environment conditions and hormonal level. In temperate regions, flowering takes place in two different growing seasons separated by dormancy. In the first growing season, lateral meristems give rise to inflorescence meristems and branches before the bud enters into dormancy. In the subsequent season, additional inflorescence and branch meristems form that later give rise to typical raceme or a bunch structure which changes into a terminal flower (Boss et al. 2003; Poupin et al., 2011; Vasconcelos et al. 2009).

Flowering is the result of combined genetic, hormonal and environmental factors. Numerous regulatory proteins at the transcriptional level tightly regulate flowering initiation and development. Even though several studies have been performed in model plant species such as Arabidopsis, establishing the function of genes in grapevine is challenging due to difficulties in obtaining mutants and transgenic plants (Searle et al. 2006). The existence of florigen, a flowering-inducing molecule, is thought to be responsible for inducing floral meristems. *Flowering locus T* (*FT*) and photoperiodic induction protein - *Constans* (*CO*) is identified as the florigen candidate in Arabidopsis (Kotake et al. 2003). These factors are regulated by hormones such as GA or other pathways related to photoperiod and vernalization. A genome-wide study in grapevine showed at least 17 genes belonging to the *CO*-like family. Two genes from this study, *V. vinifera Constans* (*VvCO*) and *V. vinifera CO-LIKE* homologous (*VvCOL1)* are expressed in latent buds and coordinate with the expression of grape orthologues (*LEAFY) LFY* and *SUPPRESSOR OF OVEREXPRESSION OF CO 1* (*SOC1)*, the genes for flowering (Almada et al. 2009). Similarly, Carmona et al. (2007) described *FT/ (TERMINAL FLOWER 1) TFL1* gene family, and products from this gene family acting as promoters or repressors of the flowering process.

Generation of Mapping Population

Generation of a segregating mapping population is the primary requirement for linkage mapping. Mapping populations should be derived from sexual reproduction i.e. fertilizing male and female gametes from parents that have contrasting traits. Although, the size of the mapping population depends on crop species, a population size of about

200 individuals is considered an ideal population for linkage mapping, and a higher population size is required for fine mapping (Mohan et al.1997).

There are different ways to create mapping populations, depending on the pollination nature of plant species. In self-pollinating species, the mapping population is generated from homozygous parents (inbreds) which are cross-fertilized followed by a series of selfings to many generations giving rise to recombinant inbred lines (RILS). However, in cross-pollinated species, the way a mapping population is generated is different due to complicated situations. Cross-pollinated species cannot tolerate self-pollination due to inbreeding depression, have more complicated ploidy, and parents are heterozygous, haploid or homozygous to begin with (Wu et al. 1992). In these species, the mapping population is generated from a cross between heterozygous/haploid/homozygous parents. In many cases, F_1 hybrids (pseudo-cross), F2, or backcross (BC) population is taken into consideration for gentic mapping. Another way to create a mapping population is by generating a double haploid (DH). DH is produced by the induction of chromosome doubling from the pollen grains. Although, this method is efficient, it can only be applied in crop species that are amenable to tissue culture (rice, barley, wheat) (Collard et al. 2005).

Grape Breeding for Disease Resistance

Unlike European and a few commercial hybrids that are suitable for wine or fresh grape production, most of the American and wild Asian grapes are used for breeding disease resistance. The overall goal of grapevine breeding is to generate new cultivars by combining both quality and disease resistance characteristics using various genetic

resources. Traditional grapevine breeding, which persisted for hundreds of years, relies on phenotypic differences among the parents based on the visual parameters. This breeding method was able to generate various cultivars with good quality profiles and considerable degree of disease resistance, needs support from molecular markers to achieve a high degree of success. Fruit crops including grapevines have a long crop cycle, growing grapes is a labor and time-consuming task and factors such as lack of knowledge on inheritance of important traits and the genomic information, are responsible for slow progress in grapevine breeding. Nonetheless, recent developments in genetic mapping techniques, identification of quantitative trait loci (QTLs), development of genetic markers linked with various traits, and availability of genomic sequence are significant achievements in breeding grapevines (Adam-Blondon et al. 2004; Fischer et al. 2004; Riaz et al. 2008; Jason et al. 2007).

Molecular Markers

Molecular/DNA markers have been indispensable for molecular breeding. These are the segments of DNA with a known location on the chromosome. In general, molecular markers are not themselves genes but they are closely associated. Genetic markers represent genetic differences among the individuals or species, called polymorphism. Unlike morphological and biological markers, DNA markers are stable, and are not influenced by environment. They are selectively neutral and are thought to have arisen from DNA mutations such as substitution, rearrangement or errors in replication of tandemly repeated DNA (Collard et al. 2005). Molecular marker are very useful in plant genetics and breeding. They are a prerequisite for gene mapping and

tagging, segregation analysis, genetic diagnosis, forensic examination, phylogenetic analysis and other biological operations (Semagna et al. 2006; Lem and Lallemand 2003). Molecular markers are popular in crop breeding due to their abundance, and ability to locate the gene of interest. Based on the method of detection, they can be divided into the following categories; 1) hybridization based: restriction fragment length polymorphism (RFLP), 2) polymerase chain reaction (PCR) based: random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR), and 3) DNA sequence based: single nucleotide polymorphism (SNP). RFLP, RAPD and AFLP are the markers that were common in the past. RFLP markers are regarded as the first shot in the genome revolution (Dodgson et al. 1997) in which polymorphism between the species is detected after digestion using restriction endonucleases. The digested fragments vary in number and size among the individuals, population and species. These fragments can be viewed by various methods, such as; southern blot analysis or detection through PCR provided flanking sequence are known (Liu et al. 2004). RAPD markers were developed in 1990 to randomly amplify segments of nuclear DNA using PCR. They are produced with single primers of arbitrary nucleotide sequence, usually about 10 nucleotides long. Genetic variation and polymorphism between the species are assessed by the presence and absence of each product i.e. they are dominant in nature. In grape, fagments size from 1-13 (200-2500bp) are amplified and routinely visualized through agarose gel electrophoresis (Dalbo et al. 2000). Most of the genetic linkage maps from mid to late 1990’s were created from RAPD markers (Grattapaglia and Sederoff 1994; Hemmat et al. 1994) that were later replaced by other markers. AFLP is a PCR based multi-locus fingerprinting, which

overcomes the weakness of RFLP and RAPD. During AFLP analysis, a small amount of DNA is digested with rare cutting restriction enzymes, like EcoRI, MseI. Digested fragments are ligated with known small oligonucleotides, called adapters. Complementary primers with the adapter sequence are used to amplify fragments via PCR followed by selective observation through gel electrophoresis. In highly heterozygous species like grape, the addition of 3-4, extra nucleotides to each primer maximizes the number of readable fragments (Marques et al. 1998). The power of AFLP analysis is extremely high for revealing genomic polymorphism. Multiple numbers of loci can be detected with high reproducibility and relative economy per marker basis. Though these markers were popular in the past, these are labor intensive to genotype, time consuming, dominate in nature and are limitedly available (Marques et al. 1998). Currently, SSR and SNP markers are popular markers for grape breeding.

Simple Sequence Repeat (SSR) Markers

SSR markers are short tandemly repeated DNA (2-6 bp) that are highly polymorphic between individuals and species (Dangl et al. 2001). They are easy to generate, and highly reproducible across laboratories. Due to high a degree of variability, SSRs are excellent markers for genotyping, analysis of pedigree and genetic variation or Kinship studies, and for marker-assisted selection (MAS) (Bowers et al. 1999).

Eukaryotic genomes contain large number of SSRs that can be used for constructing high-density genetic maps. The advantages of SSR markers over other are: 1) multiple SSR alleles can be detected at a single locus (multi-allelic), 2) SSR are evenly distributed across the genome, co-dominant in nature, 3) requires only small amount of

DNA, 4) reproducible across the laboratories, transferable across the species, and 5) can be detected using semi-automated techniques (Robinson et al. 2004). SSR markers have been extensively used in grape breeding including table grapes (Sanchez-Escribano et al. 1999; Dangl et al. 2001), wine grapes (Bowers et al. 1996) and rootstocks (Lin and Walker 1998). They are also used to identify parents-progeny relationship. For example, use of SSR markers to identify the parents of 'Chardonnay' cultivar as 'Pinot Noir' and 'Gouais blanc' (Browers et al. (1999).

SSR markers can be generated from short stretches of DNA coding regions of DNA sequence called expressed sequence tag (EST). These markers are derived from the transcripts, which are useful for assaying the functional diversity in natural populations or germplasm collections as well as to anchor markers for comparative mapping and evolutionary studies (Varshney et al. 2005). Availability of ESTs greatly accelerate the systematic identification of SSRs and corresponding marker development. EST-derived SSRs have been well documented in some plant species including Arabidopsis (Depeiges et al. 1995), sugarcane (Cordeiro et al. 2001), cereal species (Kantety et al. 2002), and rubber tree (Feng et al. 2009). Though, EST-SSR markers are reported to be less polymorphic, they are highly transferable due to their greater DNA sequence conservation in transcribed regions (Huang et al. 2011). Like genomic SSR, in the past a number of study have used EST-SSR markers for genotyping and mapping in various crops including grape. Huang et al. (2011) used EST-SSR markers, through mining and validating, for genotyping and mapping *V. vinifera* and *V. rotundifolia* cultivars. Similarly, an analysis of transferability of EST-SSR markers in grapes (Akkak et al.

2006), and analysis of SSRs derived from grape ESTs (Scott et al. 2000) are other examples.

With the advancement of grape breeding, hundreds of microsatellite markers have been developed, and most of them are publicly available from the *Vitis* Microsatellite Consortium (VMC) coordinated by Agrogene in France. In addition, a number of microsatellite markers were developed by VVI series in France. Among these microsatellite markers, a set of six (VVS2, VVMD5, VVMD7, VVMD27, VrZag62, VrZAG79) markers are regarded as universal markers, and are frequently used in grapevine study, especially in determining genetic variability and polymorphism between species (Jakse et al. 2013).

Single Nucleotide Polymorphism (SNPs)

Unlike SSR markers, that are available in limited number in the genome, a constraint to saturate regions in the linkage map with larger gap, thousands of SNPs markers can be generated to fulfill the gap. Though SSR markers are multi-allelic, revealing higher information per locus, difficulties arise during merging multiple data from different platforms and curating the allele sizes in the database. In addtion, gel-based SSRs are labor intensive, and automated fragment analyzer is expensive with limited applications. These low throughput and higher cost nature of SSR marker increases the need for SNPs (Thompson et al. 2014).

SNPs are single nucleotide base difference between individuals or DNA sequences. Unlike traditional markers, this high-throughput marker system can be generated in huge number for high-density mapping. SNPs are bi-allelic, and are thought

to be arising from substitutions/point mutations or due to insertions/deletions of single nucleotides in the genome. The frequency in which SNPs are detected is one SNP per 100-300bp in plants. These regions could be coding or non-coding sequence of genes or intergenic regions and the frequency in which they are detected depends on the region of chromosome (Batley and Edwards 2007).

SNPs are flexible, efficient, and cost-effective in terms of data management. They are straightforward to merge data across groups due to their bi-allelic nature, though bioinformatics data management and curation might be needed during the process. There are many platforms that have been designed to generate SNPs. Early generation SNPs relied on gel-based methods such as cleaved amplified polymorphic sequence (CAPS) markers or allele-specific amplification methods. With expansion of technology and advancement in sequencing, large-scale SNPs are generated through sophisticated sequencing and genotyping platforms. These genotyping platforms such as Genotyping-by-Sequencing (GBS), Restricted site Associated DNA sequencing (RAD-seq) are optimized for speed, efficiency, robustness and cost-effectiveness (Thompson et al. 2014).

GBS is a simple highly multiplexed genotyping platform with reduced representation libraries for Illumina next generation sequencing. This technique is able to generate large number of SNPs for genetic analyses and genotyping (Elshire et al. 2011). The major advantages of GBS are low cost, reduced sampling handling, fewer PCRs, and purification steps, no size fractionation, no reference sequence limits, efficient barcoding, and easiness to scale up (Davey et al. 2011). A methyl sensitive restriction enzyme digest (REs) protocol provides a greater degree of complex reduction with uniform library for

sequencing. By choosing appropriate REs (crop specific), repetitive region of the genomes can be avoided, and lower copy regions can be targeted with two to three-fold higher efficiency. This simplifies computational challenges especially in species with high level of heterozygosity. This approach was first demonstrated in maize and barley recombinant inbred populations followed by other crop and animal species (Elshre et al. 2011).

GBS has increasing become popular in many crop species including grapevines as a genotyping means by generating thousands of markers to a whole genome profiling. This method provides rapid and low-cost tool to genotype breeding populations, allowing plant breeders for efficient genetic linkage analysis, GWAS and genomic diversity studies, molecular marker discovery, and genomic selection (GS) under large scale of plant breeding populations (Poland and Rife 2012). With advancement, different crops have been included, and the system is optimized to generate millions of SNPs. For example, a comprehensive genotyping of 2,815 maize inbred accessions showed more than 681,257 SNP markers that are distributed across the entire genome and re-sequencing of 31 soybean genotypes yielded more than 205,614 SNPs for soybean breeding (Romay et al. 2013; Lam et al. 2010). GBS is also becoming a tool for genomic diversity studies (Lu et al. 2013; Fu et al., 2014) and genetic linkage analysis (Heffner et al. 2009; Poland et al. 2012). By integrating molecular markers from previous studies, GBS provides an excellent platform for molecular breeding in crop species that lack reference genome sequences (Bus et al. 2012; Truong et al. 2012; Yang et al. 2012).

Though, there are numerous benefits of GBS, this system possess few drawbacks: 1) difficulties to align markers data with reference sequence from each locus in highly

heterozygous and polyploidy species, 2) mutation in the restriction site causes failure to amplify reads and detect SNPs, 3) difficulties to decide REs due to their methyl sensitive nature, 4) bioinformatics challenges, and 5) transferability of these SNPs across other species (He et al. 2014).

Genotyping and Linkage Map Construction

To construct genetic linkage maps, one of the primary requirements is to genotype entire mapping population using molecular markers. To accomplish this, markers that reveal differences between the parents have to be identified, also called polymorphism (Young 1994). The level of polymorphism differs among the species. In general, polymorphisms are higher in cross-pollinated species than in self-pollinated species. Polymorphic markers are used to genotype the entire mapping population (genotyping). For this, DNA is extracted from each individual in the mapping population, and is amplified with a list of markers to generate genotype data. Based on the type of markers, a segregation ratio can be calculated, which is usually expressed in a Mendelian fashion, a basis for linkage map construction (Xu et al.1997).

For the construction of the linkage map, allelic size from the genotype data is decoded into software format according to population type (DH, CP, BC and RILs). Genotypic coding depends on the nature of the software used. There are many kinds of software available for linkage mapping like JoinMap4.1 (Van Ooijen 2006), Mapmaker/EXP (Lander et al. 1987), MapManager QTX (Manly et al. 2001), R/QTL (Broman et al. 2003). Each of them has their merits and demerits. Linkage between the markers is calculated using odds ratios-the ratio of linkages versus no linkage, and is

usually expressed as the logarithm of the odds (LOD) score (Risch 1992). In general, LOD value of >3 is used to construct linkage maps. This value can be adjusted based on the number of markers desired in the final linkage groups. The ower the LOD value, the more chances linkage can be detected with greater numbers of markers in the map.

Two major factors determine good linkage groups: the frequency of recombination responsible for random distribution of markers, and the number of polymorphic markers that are evenly distributed. The number of genotypes taken into consideration during genotyping determines the accuracy of genetic distance and marker order in the linkage map. The mapping distance (number of recombination fractions) is measured in centiMorgans (cM), where 1 cM equals a 1% recombination frequency. Recombination rate in meiosis is determined by the distribution of the markers, the higher the distance between the markers, the greater the chance of recombination. However, this distance may not directly be related to the physical distance of DNA between two markers. There are various mapping functions for linkage mapping. Two commonly used are the Kosambi mapping function and the Haldane mapping function (Collard et al. 2005).

Phenotyping

After genotyping, another integral part of QTL analysis is phenotyping, i.e. to phenotype a genotypic population for the trait of interest. A phenotype could be disease resistance, berry quality, flower sex, cold hardiness, dormant rooting or other depending on need. A mapping population can segregate for a single trait or multiple traits. In perennial crops, like grapevines it is not always possible to construct a mapping

population for each of the segregating traits. If the population segregates for more than one trait, this is beneficial. However, a separate approach is needed for destructive or semi-destructive bioassays during phenotyping. In this study, we phenotyped two major diseases of grapevines- downy mildew and Botrytis bunch rot.

Downy Mildew of Grapevines

Grapevine downy mildew is the major destructive diseases of grapevines and occurs throughout worldwide in viticultural areas with high rainfall. Downy mildew of grapevine is cause by an obligate oomycete *Plasmopara viticola* that attacks all green parts of the plants including leaves, tendrils, shoots, inflorescences and bunches. Infection of vines start when hyphal tips from the germinating zoospores lay on an open structure such as stomata that further penetrates to form a network of mycelia. Further development of hyphae give rise to a feeding structure called haustoria that draw nutrients from plants. Fruiting structures; sporangia on sporangiophores arise after four to six days of infection through stomatal openings, and the sporangia are released for further infection (Gindro et al. 2003). Under favorable conditions (wet weather with optimal temperature), a white, downy fungal mass of mycelium appears on the lesions causing further destruction. Infection of plant parts with downy mildew reduces functional green parts of the plant, resulting in stunted growth or death (Moriondo et al. 2005).

Genotypes that display natural resistance to downy mildew are mainly confined to North American and few Asian *Vitis* species like *V. riparia*, *V. rupestries*, *V. lincecumii*, *V. aestivalis*, *V. berlandieri*, *V. labrusca* and *Muscadania rotundifolia* (Alleweldt and Possingham 1988). However, both resistant and susceptible *Vitis* species are colonized by

P. viticola. Development of the pathogen is rapidly restricted in resistant species. Various mechanisms of resistance have been elucidated so far against downy mildew including presence of physical barriers such as hairs and stomatal closures, accumulation of phenolic antimicrobial compounds, increase perioxidase activity, accumulation of pathogenesis-related proteins and a hypersensitive response (Wan et al. 2007; Moreira et al. 2011). The resistance mechanisms so far characterized have shown that the resistance is quantitatively inherited and is governed by many genes (Moreira et al. 2011).

So far, more than fifteen minor and major *P. viticola* QTLs have been identified and mapped to various chromosomes including 4, 5, 7, 8, 9, 12, 14, 17 and 18. A resistance locus *Rpv1* (Merdinoglu et al. 2003) was identified in a cross between a susceptible *V. vinfera* 'Syrah' and the resistant '28-8-78' (*M. rotundifolia* derived) and was mapped on chromosome 12, explaining 73% of the total phenotypic variation. Resistance loci *Rpv2*, originated from *M. rotundifolia,* was mapped on chromosome 18 explaining about 76% of the total observed phenotypic variation (Wiedeman-Merdinoglue et al. 2006). A major resistance locus *Rpv3* was identified in a population of a resistant parent 'Reagent' cross with susceptible parent 'Lemberger', further downstream of *Rvp2* on chromosome 18 (Fischer et al. 2004; Welter et al. 2007). Welter et al. (2007) localized *Rpv4*, a minor resistance locus on chromosome 4. Similarly, resistance loci *Rpv5* and *Rpv6* were identified in a cross between *V. vinifera* 'Cabernet Sauvignon' and *V. riparia* 'Gloire de Montpellier' on chromosome 9 and 12, respectively (Marguerit et al. 2009). A minor resistance locus, *Rpv7* was found on chromosome 7 of 'Bianca' explaining 12.7% of the total phenotypic explanation (Bellin et al. 2009). Blasi et al. (2011) and Moreira et al. (2011) identified resistance loci *Rpv8* and *Rpv9* explaining

86% and 21.1% of the phenotypic variance observed respectively. Using *V. amurensis*-derived 'Solaris' as a resistance source, a major locus *Rpv10* and a minor locus *Rpv11* was localized on chromosome 9 and 5 explaining 50% and 12.1 % of the total phenotypic variance observed (Schwander et al. 2012). A major resistance locus associated with markers UDV014 and UDV370 was identified on chromosome 14 from *V. amurensis* and was designated as *Rpv12*; explain 78.7% of total phenotypic variance observed. Moreira et al. (2011) found another resistance locus on chromosome 12, *Rpv13* in a *V. vinifera* 'Moscato Bianco' x *V. riparia* population. Similarly, Ochssner et al. (2016) identified a resistance locus on linkage group 5 from a mapping population 'Gf.V3125' cross 'Borner' flanked by marker GF05-13. A recent study on an Asian *Vitis* species, *V. piasezkii* showed a resistance locus on linkage group 18 (Pap et al.-personal communication) Several other minor resistance loci and QTLs have been also reported on chromosomes 8, 12, 15 and 17 (Blasi et al. 2011; Moreira et al. 2011; van Heerden et al. 2014).

Botrytis Bunch Rot of Grapevines

Botrytis bunch rot, also called grey mold, is one of the most important fungal diseases that is responsible for damaging more than 230 host plants of agricultural and horticultural importance (Viret et al. 2004). In grapes, *Botrytis cinerea* Pers., a necrotrophic pathogen, is responsible for causing bunch rot. Common Botrytis rot, also called grey mold or slipskin, occurs in all major grape-producing regions of the world. Botrytis can grow on any plant materials that are succulent such as young shoots, flower parts or bunches. Grape cultivars with tight clusters and thin-skinned berries under a

heavy canopy are severely affected by *Botrytis*, especially during wet growing seasons. In addition, this fungus is also responsible for post-harvest loss or storage losses of grapes picked for fresh market.

B. cinerea pathogen completes its lifecycle by producing small, dark, hard, resting structures called sclerotia during winter in debris on the vineyard floor. These sclerotia are resistant to low temperature injury and germinate in spring under suitable weather conditions (moisture in the form of fog or dew and a temperature of 59-77 ^{0}F) producing conidia for further infection. Sporulation may occur on debris on the vineyard floor, cluster remains from the previous year or mummified berries hanging on the vines. Ripe berries on tight clusters as well as damaged berries are easily colonized by *Botrytis*. Late season infections are most severe when there is high relative humidity, free moisture in the fruit surface and temperature in the rage of 59 to 82 ^{0}F. Infected berries crack and spores germinate to quickly cover the berry (Bettiga et al. 1989).

This disease is controlled by a combination of various practices including vineyard management techniques such as; canopy management and leaf removal (Gubler et al. 1991; English et al. 1989), application of growth regulators (Hed et al. 2011), and application of chemical fungicides. Cultural practices including training and pruning, cluster thinning and cluster zone leaf removal help to increase airflow with sunlight exposure in vines and clusters, minimizing the effect of *Botrytis*. Hed (2009) found that application of gibberellic acid during bloom has a significant impact on reducing compactness (berries per centimeter) as a prudent strategy to control bunch rot. Chemical fungicides have been the most widely used strategy to reduce the impact of disease (Leroux et al. 2002). A standard fungicide spray program consists of three preventive

applications of fungicide in different phonological stages: at the end of flowering, bunch closure and beginning of berry ripening. Fungicides consist of fenhexamid, fludioxonil and pyrimethanil are applied in three different stages, respectively (Pillonel and Meyer 1997). However, current widely used mechanical methods of bunch rot control are costly and require considerable time. Similarly, chemical control of *Botrytis* is impeded by its effect on human and animal health and by the development of resistant strains. Several classes of resistance were recorded in European vineyards after the introduction of benzimidazoles and dicarboximides (Leroux et al. 1999, 2002). Thus, it is recommended to regularly alternate treatments, either with different mechanisms of action and/or with different appropriate pesticide mixtures. However, it is possible to find some level of pesticides residues in fresh grapes and wines due to late spray (Ortelli et al. 2004; Cabras and Angioni 2000). Therefore, it is utmost necessary to characterize and identify resistant germplasms against Botrytis bunch rot for breeding purposes.

There is limited information available on the genetics of Botrytis bunch rot resistance. Breeding for resistance against *B. cinerea* has been difficult and unrewarding in most crops. Fewer initiatives have been taken in few crops like tomato and chickpea for resistance mapping. A QTL mapping on chickpea showed resistance to Botrytis is controlled by few genes. A single dominant gene '*Bor1*' for Botrytis resistance was identified on chickpea (Tiwari et al. 1985). Three QTLs were identified in a segregating F_2 tomato population, and were introgressed to breed for resistance to *B. cinerea* (Finkers et al. 2007). One of the genotypes identified in this population displayed reduction in Botrytis as high as 85%, contributed by several QTLs. Herzog et al. (2015) studied grape berry cuticle as a source of novel phenotypic trait that shows resistance against *B.*

cinerea. This study was able to identify a preliminary QTL on chromosome 17, explaining 20.3 % of the total phenotypic variation. Grape berry skin features including biochemical composition, the ripening stage and morphology were found playing important role to susceptibility of *B. cinerea* infection. Presence of cuticle and epicuticular waxes on the berry surface showed differences in berry cracking response in presence of warm temperatures, high air humidity and water, affecting susceptibility to *B. cinerea* (Herzog et al. 2015). Deytieux-Belleau et al. (2009) noted grape berry skin features to ontogenic resistance to *B. cinerea*. The temporal, development changes in various morphological and biochemical features of grape berry are related to berry susceptibility. This could be due to the presence of constitutive antifungal compounds in active concentrations in immature berries (Pezet et al. 2004), or higher phenolic (Sarig et al. 1997) or tannin content (Goetz et al. 1999) or a combination thereof.

Quantitative Trait Loci (QTLs) Analysis

QTL analysis is based on the principle of detecting an association between a trait of interest/phenotype and the genotypes of markers. Markers are used to divide the segregating population into different groups based on the segregation pattern and marker types, and the significant differences between the groups are identified with respect to the phenotypes (Tanksley 1993; Young 1996). There are different methods of QTL detection: 1) Singe marker analysis (SMA), 2) Simple interval mapping (SIM) and 3) Composite interval mapping (CIM). SMA is the simplest way to detect association between the marker and the phenotype. This can be accomplished using simple statistics such as a *t*-test, analysis of variance (ANOVA) and linear regression. This method does not require a

linkage map; however, the disadvantage of this method is that the QTL is less likely to be detected if it is located farther apart. This is due to a higher probability of recombination, and it can be avoided with a greater number of DNA markers (Tanksley 1993).

SIM requires a linkage map. It analyses the interval between adjacent linked markers instead of single marker each time (Lander and Botstein 1989). This method is statistically more powerful and avoids the chances of recombination. Mapping program such as MapMaker, QGene (Nelson 1997), and JoinMap use this method to identify QTLs.

A recently developed and more popular method for QTL detection is CIM. This method combines SIM with linear regression and includes additional markers for statistical analysis. This method is more powerful and precise compared to the previous two methods, especially when linked QTLs are involved (Jasen and Stam 1994).

QTL detection produces a LOD figure with respect to a linkage map. The higher the LOD value, the better the association. A QTL is said to be detected if its value exceeds a specified threshold level. Threshold value is determined mostly by permutation tests (Churchill and Doerge 1994). Usually, a pair of the most tightly linked markers on either side of a QTL peak (flanking markers) are reported for future MAS. In perennial species like grapes, mapping population are usually constructed in such a way that, the population segregates for more than one trait. This is beneficial, since multiple QTLs using the same linkage group can be identified (Beattie et al. 2003). Detection of a QTL is determined by numerous factors including genetic properties of QTLs that control traits, environmental effects, population size and experimental error. Phenotypes with higher effects are usually detected, and small effects may fall below the threshold line.

The environment has huge effects on expression of quantitative traits. Therefore, it is necessary to replicate experiments across sites and over time (George et al. 2003). Population size plays an important role, as accuracy increases with larger population sizes, and this increases the likelihood that the QTL will be detected. Larger population sizes increase the power of statistics, the estimate of genetic effects and the confidence intervals (Beavis 1998). Experimental errors such as missing marker genotypes and errors in phenotypic values play key roles in QTL detection. Genotypic error and missing data can affect order and distance between the makers in linkage maps (Hackett 2002).

Detected QTLs can be categorized as ‘Major or ‘Minor’ based on the proportion of phenotypic variation explained (R^2). Major QTLs act as single dominate genes and explained the majority of the phenotypic variation whereas minor QTLs usually account for low phenotypic variation and are environmentally sensitive. A confidence interval is determined between two flanking markers, usually in cM to identify the flanking distance. It is usually expressed in ‘one-LOD support interval’ or ‘two-LOD support interval’. Confidence intervals are species specific, and can be altered by two different ways: 1) by increasing the number of markers for genotyping, and 2) by increasing population size to find more number of recombinants (Mohan et al. 1997).

It is necessary to cross-compare newly constructed linkage map within and among the species. This is conducted by identifying common markers between the linkage maps. Common markers are also used to construct a consensus map by combining two parental maps. Consensus maps are extremely useful to locate tightly linked markers to QTLs (Gardiner et al. 1993). Cross comparison between the maps is also necessary to identify conservation between the maps- called synteny. This is done by identifying the common

markers between the maps and their consistency in order. As higher the conservation in order of marker increases, it indicates a higher degree of synteny between maps (Paterson et al. 1991).

QTL Confirmation and Marker-assisted Selection (MAS)

In many cases, identified QTLs are rarely confirmed due to various constraints such as a lack of research funding, time consumption and crop complexity. Ideally, QTL mapping should be independently confirmed and validated. QTLs can be confirmed by: 1) using a separate mapping population from the same parental genotypes or closely related genotypes. In some cases, larger populations can be divided into groups and mapped separately to confirm QTL position and effects (Melchinger et al. 1998). 2) to use near isogenic lines (NILs). NILs are generated by backcrossing the F_1 to the recurrent parent in a number of generations. F_1 individuals are generated by crossing a donor parent (with the specific trait of interest) to a recurrent parent (elite parent commercially popular). The backcross genotypes can be selfed to generate F_2 individuals for genotyping. By comparing mean trait values of particular NIL lines with the recurrent parent, the effects of QTLs can be confirmed (Bernacchi et al. 1998; Van Berloo et al. 2001).

MAS is the process of selecting individuals based on the genotype of the marker. For this, a marker identified in the preliminary genetic mapping requires further testing and development. This can be done by developing high-resolution maps, validating markers and possibly by marker conversion.

Sparsely distributed markers may not be suitable for MAS. This could be because even closely associated markers may not be tightly linked and there is a higher chance of recombination. Recombination reduces the reliability and usefulness of a marker for MAS. A skeleton linkage map can be fine-tuned by adding more makers (same or different types) or by increasing the population size (Michelmore 1995; Mohan et al. 1997). There is no universal number for the number of markers; however, a population size of more than 1,000 individuals used for fine mapping could reduce the confidence interval to <1cM (Blair et al. 2003).

Markers should be validated by testing their effectiveness in determining target phenotype in an independent population with a different genetic background before implementation. This is called marker validation. Only validated markers are able to predict a phenotype accurately for routine screening in MAS (Li et al. 2001; Sharp et al. 2001).

In some instances, markers need to be converted to another type of marker when there are problems of reproducibility; the marker technique is complicated, time-consuming or expensive (RAPDs, RFLPs or AFLPs). This can be overcome by developing sequence characterized amplified regions (SCARs) or sequence-tagged sites (STSs) derived by cloning and sequencing (Jung et al. 1999; Paran and Michelmore 1993). This conversion to PCR based markers is simple, less time consuming and cheaper (Brondani et al. 2003).

The advantage of MAS is that a large number of plants can be screened prior to genotyping and field evaluations. MAS saves time that would be taken up in complex field trials, eliminates unwanted genotypes, allows selection of genotypes at seedling

stage, eliminates unreliable phenotypic evaluation, prevents linkage drag that comes from the donor parent, and is valuable for selection of genotypes that are complex to evaluate phenotypically and have low heritability.

Genetic Linkage Mapping and QTL Analysis in Grapevines

Grapevines, a perennial highly heterozygous species with some extent of outcrossing, typically require a different method for genetic mapping called 'pseudo-testcross'. Conventional genetic mapping is conducted by generating F_2/RILs/backcross/double haploid population using parents of different genetic backgrounds. Meanwhile, a pseudo-testcross strategy implements crossing of two highly heterozygous individuals to identify markers that segregate for either of the parents or are double heterozygous. These markers are fully informative with respect to the segregation attributed to the parent of interest. Despite difficulties in linkage map construction, markers heterozygous for both parents can be utilized to harness additional genetic gain (Grattapaglia and Sederoff 1994).

With the availability of molecular markers, several genetic maps have been constructed in grapevines (Dalbo et al. 2000; Doligez et al. 2002; Grando et al. 2003; Adam-Blondon et al. 2004; Fischer et al. 2004; Riaz et al. 2004; Doligez et al. 2006; Lowe and Walker 2006; Di Gaspero et al. 2007; Troggio et al. 2007; Welter et al. 2007; Vezzuli et al. 2008; Bellin et al. 2009; Marguerit et al. 2009; Moreira et al. 2011; Blasi et al. 2011; Blanc et al. 2012; Hammers et al. 2017). Initial maps were developed using AFLP and RAPD markers, and these were later improved with addition of SSR markers. Most maps recently have been constructed using SSR markers including an integrated

map from the International Grape Genome Program (http://www.vitaceae.org/index.php/Maps_and_Markers). With the availability of genomic sequences, physical maps have also been developed in various cultivars such as the *V. vinifiera* grapevine reference genome for a nearly homozygous selection, 'PN40024' (Jaillon et al. 2007), 'Cabernet Sauvignon' (Moroldo et al. 2008) and 'Pinot noir' (Velasco et al. 2007). Next generation sequencing, particularly GBS is employed to generate a high-density linkage map using SNPs (Wang et al. 2012; Barba et al. 2014; Yang et al. 2016). GBS offers an inexpensive and robust solution for SNP discovery through reduced representation libraries. GBS generates thousands of markers that are suitable for QTL mapping in bi-parental families (Elshire et al. 2011).

Several of the maps that were constructed previously have been used to detect QTLs associated with fungal diseases such as powdery mildew and downy mildew as well as pest resistance (Di Gaspero et al. 2007; Fischer et al. 2004; Riaz et al. 2004, 2006, 2008; Xu et al. 2008;). QTLs associated with a range of agronomic traits, e.g. berry size, seed number, mean and total seed fresh and dry weight, berry weight, inflorescences per shoot, flowering date, time and duration, and veraison have also been identified (Constantini et al. 2008; Dalbo et al. 2000; Doligez et al. 2002; Fanizza et al. 2005). Recently, there are a few reports on QTLs on berry chemistry, anthocyanin content, sugar and acid production and aromas (Chen et al. 2015; Sevini et al. 2003; Yang et al. 2016).

References

Adam-Blondon A-F, Roux C, Claux D, Butterlin G, Merdinoglu D, This P (2004) Mapping 245 SSR markers on the *Vitis vinifera* genome: a tool for grape genetics. Theor Appl Genet 109:1017-1027

Akkak A, Lembo A, Torello Marinoni D, Gribaudo I, Botta R (2006) Analysis and transferability of EST-SSR markers in grapes. IX Intl Con on Grape Genet and Breed 827:59-62

Alleweldt G, Possingham J (1988) Progress in grapevine breeding. Theor Appl Genet 75:669-673

Almada R, Cabrera N, Casaretto JA, Ruiz-Lara S, Villanueva EG (2009) *VvCO* and *VvCOL1*, two CONSTANS homologous genes, are regulated during flower induction and dormancy in grapevine buds. Plant Cell Rep 28:1193-1203

Ambers CP (2013) A historical hypothesis on the origin of the Norton grape. J Wine Res 24:85-95

Barba P, Cadle-Davidson L, Harriman J, Glaubitz JC, Brooks S, Hyma K, Reisch B (2014) Grapevine powdery mildew resistance and susceptibility loci identified on a high-resolution SNP map. Theor Appl Genet 127:73-84

Batley J, Edwards D (2007) SNP applications in plants. Association mapping in plants. Springer 95-102

Beattie AD, Larsen J, Michaels TE, Pauls KP (2003) Mapping quantitative trait loci for a common bean (*Phaseolus vulgaris* L.) ideotype. Genome 46:411-422

Beavis WD (1998) QTL analyses: power, precision, and accuracy. Molecular dissection of complex traits 145-162

Bellin D, Peressotti E, Merdinoglu D, Wiedemann-Merdinoglu S, Adam-Blondon A-F, Cipriani G, Morgante M, Testolin R, Di Gaspero G (2009) Resistance to *Plasmopara viticola* in grapevine 'Bianca'is controlled by a major dominant gene causing localized necrosis at the infection site. Theor Appl Genet 120:163-176

Bernacchi D, Beck-Bunn T, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley S (1998) Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. Theor Appl Genet 97:381-397

Bettiga L, Gubler W, Marois J, Bledsoe A (1989) Integrated control of Botrytis bunch rot of grape. California agriculture 43:9-11

Blair MW, Garris AJ, Iyer AS, Chapman B, Kresovich S, McCouch SR (2003) High resolution genetic mapping and candidate gene identification at the *xa5* locus for bacterial blight resistance in rice (*Oryza sativa* L.). Theor Appl Genet 107:62-73

Blanc S, Wiedemann-Merdinoglu S, Dumas V, Mestre P, Merdinoglu D (2012) A reference genetic map of *Muscadinia rotundifolia* and identification of *Ren5*, a new major locus for resistance to grapevine powdery mildew. Theor Appl Genet 125:1663-1675

Blasi P, Blanc S, Wiedemann-Merdinoglu S, Prado E, Rühl EH, Mestre P, Merdinoglu D (2011) Construction of a reference linkage map of *Vitis amurensis* and genetic mapping of *Rpv8*, a locus conferring resistance to grapevine downy mildew. Theor Appl Genet 123:43-53

Bordelon BP, Ellis MA, Foster R (2007) Midwest commercial small fruit and grape spray guide. Purdue University Extension Publication ID-169: West Lafayette, IN

Boss PK, Buckeridge EJ, Poole A, Thomas MR (2003) New insights into grapevine flowering. Funct Plnt Bio 30:593-606

Bowers J, Dangl GS, Vignani R, Meredith C (1996) Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). Genome 39:628-633

Bowers JE, Dangl GS, Meredith CP (1999) Development and characterization of additional microsatellite DNA markers for grape. Am J of Enol and Vitic 50:243-246

Broman, K. W., Wu, H., Sen, Ś., & Churchill, G. A. (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19(7), 889-890

Brondani C, Rangel PHN, Borba TCO, Brondani RPV (2003) Transferability of microsatellite and sequence tagged site markers in *Oryza* species. Hereditas 138:187-192

Bus A, Hecht J, Huettel B, Reinhardt R, Stich B (2012) High-throughput polymorphism detection and genotyping in *Brassica napus* using next-generation RAD sequencing. BMC Genomics 13:281

Cabras P, Angioni A (2000) Pesticide residues in grapes, wine, and their processing products. J of Agric Food Chem 48:967-973

Carmona MJ, Calonje M, Martínez-Zapater JM (2007) The *FT/TFL1* gene family in grapevine. Plnt Mol Biol 63:637-650

Chen J, Wang N, Fang L-C, Liang Z-C, Li S-H, Wu B-H (2015) Construction of a high-density genetic map and QTLs mapping for sugars and acids in grape berries. BMC Plant Biol 15:28

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963-971

Collard B, Jahufer M, Brouwer J, Pang E (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica 142:169-196

Cordeiro GM, Henry RJ (2001) Evaluation of microsatellites (Simple Sequence Repeats) as genetic markers in sugarcane. Proc Intl Soc Sugar Cane Technol 627-629

Costantini L, Battilana J, Lamaj F, Fanizza G, Grando MS (2008) Berry and phenology-related traits in grapevine (*Vitis vinifera* L.): from quantitative trait loci to underlying genes. BMC Plant Biol 8:38

Dalbó M, Ye G, Weeden N, Steinkellner H, Sefc K, Reisch B (2000) A gene controlling sex in grapevines placed on a molecular marker-based genetic map. Genome 43:333-340

Dangl GS, Mendum ML, Prins BH, Walker MA, Meredith CP, Simon CJ (2001) Simple sequence repeat analysis of a clonally propagated species: a tool for managing a grape germplasm collection. Genome 44:432-438

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Rev Genet 12:499

Depeiges A, Goubely C, Lenoir A, Cocherel S, Picard G, Raynal M, Grellet F, Delseny M (1995) Identification of the most represented repeated motifs in *Arabidopsis thaliana* microsatellite loci. Theor Appl Genet 91:160-168

Deytieux-Belleau C, Geny L, Roudet J, Mayet V, Donèche B, Fermaud M (2009) Grape berry skin features related to ontogenic resistance to *Botrytis cinerea*. Eur J of Plant Pathol 125:551

Di Gaspero G, Cipriani G, Adam-Blondon A-F, Testolin R (2007) Linkage maps of grapevine displaying the chromosomal locations of 420 microsatellite markers and 82 markers for R-gene candidates. Theor Appl Genet 114:1249-1263

Dodgson JB, Cheng HH, Okimoto R (1997) DNA marker technology: a revolution in animal genetics. Poult Sci 76:1108-1114

Doligez A, Adam-Blondon A-F, Cipriani G, Di Gaspero G, Laucou V, Merdinoglu D, Meredith C, Riaz S, Roux C, This P (2006) An integrated SSR map of grapevine based on five mapping populations. Theor Appl Genet 113:369-382

Doligez A, Bouquet A, Danglot Y, Lahogue F, Riaz S, Meredith C, Edwards K, This P (2002) Genetic mapping of grapevine (*Vitis vinifera* L.) applied to the detection of QTLs for seedlessness and berry weight. Theor Appl Genet 105:780-795

Eibach R, Zyprian E, Welter L, Topfer R (2007) The use of molecular markers for pyramiding resistance genes in grapevine breeding. Vitis 46:120

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS One 6:e19379

English J, Thomas C, Marois J, Gubler W (1989) Microclimates of grapevine canopies associated with leaf removal and control of Botrytis bunch rot. Phytopathology 79:395-401

Fanizza G, Lamaj F, Costantini L, Chaabane R, Grando M (2005) QTL analysis for fruit yield components in table grapes (*Vitis vinifera*). Theor Appl Genet 111:658-664

Feng S, Li W, Huang H, Wang J, Wu Y (2009) Development, characterization and cross-species/genera transferability of EST-SSR markers for rubber tree (*Hevea brasiliensis*). Mol Breed 23:85-97

Finkers R, van Heusden AW, Meijer-Dekens F, van Kan JA, Maris P, Lindhout P (2007) The construction of a *Solanum habrochaites* LYC4 introgression line population and the identification of QTLs for resistance to *Botrytis cinerea*. Theor Appl Genet 114:1071-1080

Fischer BM, Salakhutdinov I, Akkurt M, Eibach R, Edwards K, Töpfer R, Zyprian E (2004) Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. Theor Appl Genet 108:501-515

Fu Y-B (2014) Genetic diversity analysis of highly incomplete SNP genotype data with imputations: an empirical assessment. G3: Genes Genom Genet 4:891-900

Gardiner J, Coe E, Melia-Hancock S, Hoisington D, Chao S (1993) Development of a core RFLP map in maize using an immortalized F2 population. Genetics 134:917-930

George M, Prasanna B, Rathore R, Setty T, Kasim F, Azrai M, Vasal S, Balla O, Hautea D, Canama A (2003) Identification of QTLs conferring resistance to downy mildews of maize in Asia. Theor Appl Genet 107:544-551

Gindro K, Pezet R, Viret O (2003) Histological study of the responses of two *Vitis vinifera* cultivars (resistant and susceptible) to *Plasmopara viticola* infections. Plant Physiol Biochem 41:846-853

Goetz G, Fkyerat A, Métais N, Kunz M, Tabacchi R, Pezet R, Pont V (1999) Resistance factors to grey mold in grape berries: identification of some phenolics inhibitors of *Botrytis cinerea* stilbene oxidase. Phytochemistry 52:759-767

Grando MS, Bellin D, Edwards K, Pozzi C, Stefanini M, Velasco R (2003) Molecular linkage maps of *Vitis vinifera* L. and *Vitis riparia* Mchx. Theor Appl Genet 106:1213-1224

Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137:1121-1137

Gubler W, Bettiga L, Heil D (1991) Comparisons of hand and machine leaf removal for the control of Botrytis bunch rot. American journal of enology and viticulture 42:233-236

Hackett CA (2002) Statistical methods for QTL mapping in cereals. Plant Mol Biol 48:585-599

Hammers M, Sapkota S, Chen L-L, Hwang C-F (2017) Constructing a genetic linkage map of *Vitis aestivalis*-derived “Norton” and its use in comparing Norton and Cynthiana. Mol Breed 37:64

Harris J (2012) The great grape debate. The Midwest Wine Grower 4:19-21

Hed B, Ngugi HK, Travis JW (2009) Relationship between cluster compactness and bunch rot in Vignoles grapes. Plant Dis 93:1195-1201

He, J., Zhao, X., Laroche, A., Lu, Z. X., Liu, H., & Li, Z. (2014) Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. Front Plant Sci 5

Hed B, Ngugi HK, Travis JW (2011) Use of gibberellic acid for management of bunch rot on Chardonnay and Vignoles grape. Plant Dis 95:269-278

Hedrick UP (1908) The grapes of New York. State of New York, Dept. of Agriculture 15th Annual Report, Volume III, part II. J.R. Lyon Co. State Printers, Albany

Heffner EL, Sorrells ME, Jannink J-L (2009) Genomic selection for crop improvement. Crop Sci 49:1-12

Hemmat M, Weedon N, Manganaris A, Lawson D (1994) Molecular marker linkage map for apple. J Hered 85:4-11

Herzog K, Wind R, Töpfer R (2015) Impedance of the grape berry cuticle as a novel phenotypic trait to estimate resistance to *Botrytis cinerea*. Sensors 15:12498-12512

Huang H, Lu J, Ren Z, Hunter W, Dowd SE, Dang P (2011) Mining and validating grape (*Vitis* L.) ESTs to develop EST-SSR markers for genotyping and mapping. Mol Breed 28:241-254

Jaillon O, Aury J-M, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463

Jakše J, Štajner Na, Tomić L, Javornik B (2013) Application of Microsatellite Markers in Grapevine and Olives. The Mediterranean Genetic Code-Grapevine and Olive. InTech

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447-1455

Jung G, Skroch P, Nienhuis J, Coyne D, Arnaud-Santana E, Ariyarathne H, Marita J (1999) Confirmation of QTL associated with common bacterial blight resistance in four different genetic backgrounds in common bean. Crop Sci 39:1448-1455

Kamas J, Black M, Appel D, Wilson L (2000) Management of Pierce's disease in Texas. Texas Agr Ext Serv L 5383

Kantety RV, La Rota M, Matthews DE, Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. Plant Mol Biol 48:501-510

Keeley K, Preece JE, Taylor BH (2003) Increased rooting of 'Norton'grape cuttings using auxins and gibberellin biosynthesis inhibitors. HortScience 38:281-283

Kotake T, Takada S, Nakahigashi K, Ohto M, Goto K (2003) Arabidopsis *TERMINAL FLOWER 2* gene encodes a heterochromatin protein 1 homolog and represses both *FLOWERING LOCUS T* to regulate flowering time and several floral homeotic genes. Plant and Cell Physiol 44:555-564

Kruglyak ELL (1996) Genetic dissection of complex traits. Nature genetics 12:357-358

Lam H-M, Xu X, Liu X, Chen W, Yang G, Wong F-L, Li M-W, He W, Qin N, Wang B (2010) Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nat Genet 42:1053-1059

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185-199

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174-181

Lem P, Lallemand J (2003) Grass consensus STS markers: an efficient approach for detecting polymorphism in Lolium. Theor Appl Genet 107:1113-1122

Leroux P, Chapeland F, Desbrosses D, Gredt M (1999) Patterns of cross-resistance to fungicides in *Botryotinia fuckeliana* (*Botrytis cinerea*) isolates from French vineyards. Crop Prot 18:687-697

Leroux P, Fritz R, Debieu D, Albertini C, Lanen C, Bach J, Gredt M, Chapeland F (2002) Mechanisms of resistance to fungicides in field strains of *Botrytis cinerea*. Pest Manag Sci 58:876-888

Li Z, Jakkula L, Hussey R, Tamulonis J, Boerma H (2001) SSR mapping and confirmation of the QTL from PI96354 conditioning soybean resistance to southern root-knot nematode. Theor Appl Genet 103:1167-1173

Lin H, Walker MA (1998) Identifying grape rootstocks with simple sequence repeat (SSR) DNA markers. Am J Enol Vitic 49:403-407

Liu Y-G, Huang N (1998) Efficient amplification of insert end sequences from bacterial artificial chromosome clones by thermal asymmetric interlaced PCR. Plant Mol Biol Rep 16:175-175

Lowe K, Walker M (2006) Genetic linkage map of the interspecific grape rootstock cross Ramsey (*Vitis champinii*) × Riparia Gloire (*Vitis riparia*). Theor Appl Genet 112:1582-1592

Lu F, Lipka AE, Glaubitz J, Elshire R, Cherney JH, Casler MD, Buckler ES, Costich DE (2013) Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. PLoS Genet 9:e1003215

Main GL, Morris JR (2004) Leaf-removal effects on Cynthiana yield, juice composition, and wine composition. Am J Enol Vitic 55:147-152

Manly, K. F., Cudmore Jr, R. H., & Meer, J. M. (2001) Map Manager QTX, cross-platform software for genetic mapping. Mamm Genome 12:930-932

Marguerit E, Boury C, Manicki A, Donnart M, Butterlin G, Némorin A, Wiedemann-Merdinoglu S, Merdinoglu D, Ollat N, Decroocq S (2009) Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. Theor Appl Genet 118:1261-1278

Marques C, Araujo J, Ferreira J, Whetten R, O'malley D, Liu B-H, Sederoff R (1998) AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. Theor Appl Genet 96:727-737

Mehmel, T. O., & Van Heerden, T. O. (2010) Effect of climate and soil water status on Cabernet Sauvignon (*Vitis vinifera* L.) grapevines in the Swartland region with special reference to sugar loading and anthocyanin biosynthesis

Melchinger AE, Utz HF, Schön CC (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383-403

Merdinoglu D, Wiedeman-Merdinoglu S, Coste P, Dumas V, Haetty S, Butterlin G, Greif C (2003) Genetic analysis of downy mildew resistance derived from *Muscadinia rotundifolia*. Acta Hort

Michelmore R (1995) Molecular approaches to manipulation of disease resistance genes. Annu Rev Phytopathol 33:393-427

Mohan M, Nair S, Bhagwat A, Krishna T, Yano M, Bhatia C, Sasaki T (1997) Genome mapping, molecular markers and marker-assisted selection in crop plants. Mol Breed 3:87-103

Moreira FM, Madini A, Marino R, Zulini L, Stefanini M, Velasco R, Kozma P, Grando MS (2011) Genetic linkage maps of two interspecific grape crosses (*Vitis* spp.) used to localize quantitative trait loci for downy mildew resistance. Tree Genet Genomes 7:153-167

Moriondo M, Orlandini S, Giuntoli A, Bindi M (2005) The effect of downy and powdery mildew on grapevine (*Vitis vinifera* L.) leaf gas exchange. J Phytopath 153:350-357

Moroldo M, Paillard S, Marconi R, Fabrice L, Canaguier A, Cruaud C, De Berardinis V, Guichard C, Brunaud V, Le Clainche I (2008) A physical map of the heterozygous grapevine 'Cabernet Sauvignon' allows mapping candidate genes for disease resistance. BMC Plant Biol 8:66

Nelson JC (1997) QGENE: software for marker-based genomic analysis and breeding. Mol Breed 3:239-245

Ortelli D, Edder P, Corvi C (2004) Multiresidue analysis of 74 pesticides in fruits and vegetables by liquid chromatography–electrospray–tandem mass spectrometry. Analytica Chimica Acta 520:33-45

Paran I, Michelmore R (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet 85:985-993

Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES, Tanksley SD (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. Genetics 127:181-197

Pezet R, Gindro K, Viret O, Richter H (2004) Effects of resveratrol, viniferins and pterostilbene on *Plasmopara viticola* zoospore mobility and disease development. Vitis 43:145-148

Pillonel C, Meyer T (1997) Effect of phenylpyrroles on glycerol accumulation and protein kinase activity of *Neurospora crassa*. Pest Manag Sci 49:229-236

Poland J, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, Dreisigacker S, Crossa J, Sánchez-Villeda H, Sorrells M (2012) Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome 5:103-113

Poland JA, Rife TW (2012) Genotyping-by-sequencing for plant breeding and genetics. Plant Genome 5:92-102

Riaz S, Dangl G, Edwards K, Meredith C (2004) A microsatellite marker based framework linkage map of *Vitis vinifera* L. Theor Appl Genet 108:864-872

Riaz S, Krivanek A, Xu K, Walker M (2006) Refined mapping of the Pierce's disease resistance locus, *PdR1*, and Sex on an extended genetic map of *Vitis rupestris*× *V. arizonica*. Theor Appl Genet 113:1317

Riaz S, Tenscher A, Rubin J, Graziani R, Pao S, Walker M (2008) Fine-scale genetic mapping of two Pierce's disease resistance loci and a major segregation distortion region on chromosome 14 of grape. Theor Appl Genet 117:671

Risch N, Giuffra L (1992) Model misspecification and multipoint linkage analysis. Hum Hered 42:77-92

Robinson AJ, Love CG, Batley J, Barker G, Edwards D (2004) Simple sequence repeat marker loci discovery using SSR primer. Bioinformatics 20:1475-1476

Romay MC, Millard MJ, Glaubitz JC, Peiffer JA, Swarts KL, Casstevens TM, Elshire RJ, Acharya CB, Mitchell SE, Flint-Garcia SA (2013) Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol 14:R55

Sanchez-Escribano E, Martin J, Carreno J, Cenis J (1999) Use of sequence-tagged microsatellite site markers for characterizing table grape cultivars. Genome 42:87-93

Sapkota S, Chen L-L, Schreiner K, Ge H, Hwang C-F (2015) A phenotypic study of Botrytis bunch rot resistance in *Vitis aestivalis*-derived 'Norton' grape. Trop Plant Pathol 40:279-282

Sarig P, Zutkhi Y, Monjauze A, Lisker N, Ben-Arie R (1997) Phytoalexin elicitation in grape berries and their susceptibility to *Rhizopus stolonifer*. Physiol Mol Plant Pathol 50:337-347

Schwander F, Eibach R, Fechter I, Hausmann L, Zyprian E, Töpfer R (2012) *Rpv10*: a new locus from the Asian *Vitis* gene pool for pyramiding downy mildew resistance loci in grapevine. Theor Appl Genet 124:163-176

Scott KD, Eggler P, Seaton G, Rossetto M, Ablett EM, Lee LS, Henry RJ (2000) Analysis of SSRs derived from grape ESTs. Theor Appl Genet 100:723-726

Searle I, He Y, Turck F, Vincent C, Fornara F, Kröber S, Amasino RA, Coupland G (2006) The transcription factor *FLC* confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. Genes Dev 20:898-912

Semagn, K., Bjørnstad, Å., & Ndjiondjop, M. N. (2006) An overview of molecular marker methods for plants. Afr J Biotechnol 5: 25

Sevini F, Marino R, Grando M, Moser S, Versini G (2003) Mapping candidate genes and QTLs for aroma content in grape. I International Symposium on Grapevine Growing, Commerce and Research 652:439-446

Sharp P, Johnston S, Brown G, McIntosh R, Pallotta M, Carter M, Bariana H, Khatkar S, Lagudah E, Singh R (2001) Validation of molecular markers for wheat breeding. Aust J Agric Res 52:1357-1366

Srinivasan C, Mullins MG (1979) Flowering in *Vitis*: Conversion of tendrils into inflorescences and bunches of grapes. Planta 145:187-192

Stover E, Aradhya M, Yang J, Bautista J, Dangl GS (2009) Investigations into the origin of 'Norton' grape using SSR markers. Proceedings of the annual meeting of the Florida State Horticultural Society

Tanksley SD (1993) Mapping polygenes. Annu Rev Genet 27:205-233

Terral J-F, Tabard E, Bouby L, Ivorra S, Pastor T, Figueiral I, Picq S, Chevance J-B, Jung C, Fabre L (2009) Evolution and history of grapevine (*Vitis vinifera*) under domestication: new morphometric perspectives to understand seed domestication syndrome and reveal origins of ancient European cultivars. Annals Botany 105:443-455

Tewari S, Pandey M, Pandya B, Chaube H, Tripathi H (1985) Inheritance of resistance to Botrytis gray mold in chickpea. International chickpea newsletter

This P, Jung A, Boccacci P, Borrego J, Botta R, Costantini L, Crespan M, Dangl G, Eisenheld C, Ferreira-Monteiro F (2004) Development of a standard set of microsatellite reference alleles for identification of grape cultivars. Theor Appl Genet 109:1448-1458

Thomson MJ (2014) High-throughput SNP genotyping to accelerate crop improvement. Plant Breed Biotechnol 2:195-212

Troggio M, Malacarne G, Coppola G, Segala C, Cartwright DA, Pindo M, Stefanini M, Mank R, Moroldo M, Morgante M (2007) A dense single-nucleotide polymorphism-based genetic linkage map of grapevine (*Vitis vinifera* L.) anchoring Pinot Noir bacterial artificial chromosome contigs. Genetics 176:2637-2650

Truong HT, Ramos AM, Yalcin F, de Ruiter M, van der Poel HJ, Huvenaars KH, Hogers RC, van Enckevort LJ, Janssen A, van Orsouw NJ (2012) Sequence-based genotyping for marker discovery and co-dominant scoring in germplasm and populations. PLoS One 7:e37565

van Berloo R, Aalbers H, Werkman A, Niks RE (2001) Resistance QTL confirmed through development of QTL-NILs for barley leaf rust resistance. Mol Breed 8:187-195

van Heerden CJ, Burger P, Vermeulen A, Prins R (2014) Detection of downy and powdery mildew resistance QTL in a ‘Regent’×‘RedGlobe’ population. Euphytica 200:281-295

Van Ooijen J (2006) JoinMap 4. Software for the calculation of genetic linkage maps in experimental populations Kyazma BV, Wageningen, Netherlands 33

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23:48-55

Vasconcelos MC, Greven M, Winefield CS, Trought MC, Raw V (2009) The flowering process of *Vitis vinifera*: a review. Am J Enol Vitic 60:411-434

Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, FitzGerald LM, Vezzulli S, Reid J (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PloS One 2:e1326

Vezzulli S, Troggio M, Coppola G, Jermakow A, Cartwright D, Zharkikh A, Stefanini M, Grando MS, Viola R, Adam-Blondon A-F (2008) A reference integrated map for cultivated grapevine (*Vitis vinifera* L.) from three crosses, based on 283 SSR and 501 SNP-based markers. Theor Appl Genet 117:499-511

Viret O, Keller M, Jaudzems VG, Cole FM (2004) *Botrytis cinerea* infection of grape flowers: light and electron microscopical studies of infection sites. Phytopathology 94:850-857

Wan Y, Schwaninger H, He P, Wang Y (2007) Comparison of resistance to powdery mildew and downy mildew in Chinese wild grapes. Vitis 46:132

Wang N, Fang L, Xin H, Wang L, Li S (2012) Construction of a high-density genetic map for grape using next generation restriction-site associated DNA sequencing. BMC Plant Biol 12:148

Welter LJ, Göktürk-Baydar N, Akkurt M, Maul E, Eibach R, Töpfer R, Zyprian EM (2007) Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). Mol Breed 20:359-374

Wiedemann-Merdinoglu S, Prado E, Coste P, Dumas V, Butterlin G, Bouquet A, Merdinoglu D (2006) Genetic analysis of resistance to downy mildew from *Muscadinia rotundifolia*. 9th International Conference on Grape Genetics and Breeding 2006

Wolf TK, Natural Resource A, Extension ESC (2008) Wine Grape Production Guide for Eastern North America. Natural Resource, Agriculture, and Engineering Service (NRAES) Cooperative Extension

Wu K, Burnquist W, Sorrells M, Tew T, Moore P, Tanksley S (1992) The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor Appl Genet 83:294-300

Xu K, Riaz S, Roncoroni N, Jin Y, Hu R, Zhou R, Walker M (2008) Genetic and QTL analysis of resistance to *Xiphinema index* in a grapevine cross. Theor Appl Genet 116:305-311

Xu Y, Zhu L, Xiao J, Huang N, McCouch S (1997) Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). Mol Gen Genet 253:535-545

Yang H, Tao Y, Zheng Z, Li C, Sweetingham MW, Howieson JG (2012) Application of next-generation sequencing for rapid marker development in molecular plant breeding: a case study on anthracnose disease resistance in *Lupinus angustifolius* L. BMC Genomics 13:318

Yang S, Fresnedo-Ramírez J, Sun Q, Manns DC, Sacks GL, Mansfield AK, Luby JJ, Londo JP, Reisch BI, Cadle-Davidson LE (2016) Next generation mapping of enological traits in an F2 interspecific grapevine hybrid family. PloS One 11:e0149560

Young N (1996) QTL mapping and quantitative disease resistance in plants. Annu Rev Phytopathol 34:479-501

Young ND (1994) Constructing a plant genetic linkage map with DNA markers. DNA-based markers in plants. Springer 39-57

CHAPTER 2

A PHENOTYPIC STUDY OF BOTRYTIS BUNCH ROT RESISTANCE IN *VITIS AESTIVALIS*-DERIVED 'NORTON' GRAPE

Sapkota, S., Chen, L. L., Schreiner, K., Ge, H., & Hwang, C. F. (2015) A phenotypic study of Botrytis bunch rot resistance in *Vitis aestivalis*-derived 'Norton' grape. Trop Plant Pathol 40: 279-282

Botrytis bunch rot, also called grey mold, is a major threat to grape growing areas worldwide where periods of high relative humidity (> 90 %) persist for a long time or where wetness and cool temperature (58 to 82 °F) coincides with bloom and ripening (Broome et al., 1995; Cadle-Davidson, 2008). The disease is caused by a necrotrophic fungus, *Botrytis cinerea* (telemorph; *Botryotinia fuckeliana*), and affects fruit quality and yield. This disease is responsible for major postharvest loss that occurs primarily during storage and transportation; however, *B. cinerea* infection during bloom and after veraison is also not uncommon (Keller et al., 2003; Holz et al., 2003). It has been reported that the susceptibility of ripening berries increases as berry defenses decline in the post-veraison period (Pezet et al., 2003). The current strategy for control of Botrytis bunch rot majorly relies on a combination of canopy management and fungicide usage. Vineyard postharvest fungicides are the most common method for controlling disease (Smilanick et al., 2010). However, multidrug resistant *B. cinerea* populations have been recently identified and reported in the vineyards of France and Germany (Kretschmer et al., 2009). Physical methods like the use of fruit zone leaf removal, shoot positioning, and timely training and pruning have reduced disease severity, but these techniques are expensive, laborious, and less effective than fungicides (Percival et al., 1994; Kulakiotu et al., 2004). Thus, the development of resistant varieties would reduce the dependence of viticulture on pesticide inputs and have significant environmental and economic benefits.

V. aestivalis-derived 'Norton', the official grape of the State of Missouri is grown in many US regions where *V. vinifera* production requires extensive pesticide use for fungal disease management (Ambers and Ambers, 2004). Total Norton grape bearing acreage in Missouri is approximately 16% of the total grape growing area or 274.9 acres

(Missouri Grape Facts, Missouri Grape and Wine Board, 2014; http://missouriwine.org/wp-content/uploads/2012/12/2014-grape-facts.pdf). Norton is a grape of American origin that was created as a hybrid seedling between *V. vinifera* and *V. aestivalis* (Stover et al., 2009; Ambers, 2013). It produces a dark, full-bodied premium red wine. Growers throughout the Midwest have observed Norton to be cold hardy and resistant to several fungal pathogens including powdery mildew, downy mildew, Botrytis bunch rot, and black rot. In addition to fungal disease resistance, Norton also displays tolerance to Pierce's disease (Kamas et al. 2004)(Kamas et al., 2004) and is resistant to Phylloxera (Hedrick, 1908). An additional prominent characteristic of Norton grape is high levels of anthocyanins that are associated with health benefits (Hogan et al., 2009). The disease resistance and cold hardiness of Norton have enhanced its attractiveness to wine growers due to increasing concerns regarding environmental protection and pesticide avoidance. However, there are no documented reports of Botrytis bunch rot on Norton grapes.

Pathogen stress imposed by *B. cinerea* can be unpredictable and sporadic. Another complication in disease evaluation is the fact that several berry rot pathogens may simultaneously infect vineyard clusters. To improve the efficiency of disease assessment, we used disease incidence and severity as an indicator of berry-pathogen interaction under *in vitro* conditions. The objective of this study was to develop a protocol to improve the accuracy, repeatability, and speed of inoculated Norton and Cabernet Sauvignon berries under laboratory conditions. To optimize the technique, post-veraison berries at Eichhorn-Lorenz (E-L) 36 to 40 stages (Coombe, 1995) were selected; the assay was then validated and refined for two years, 2012 and 2013.

Norton and Cabernet Sauvignon plants grown and maintained in a vineyard at the Missouri State Fruit Experiment Station (MSFES), Mountain Grove, MO were used in this study. Randomly selected berries from each of the two cultivars were detached by cutting the pedicel from the rachis for each E-L system stage 36 to 40 (Coombe, 1995). The harvested berries were first washed with tap water, surface-sterilized by immersion in 0.5% sodium hypochlorite solution for 10 minutes and soaked and rinsed thoroughly using distilled water.

A *B. cinerea* isolate was originally obtained and purified from diseased Cabernet Sauvignon berries at MSFES. To produce sufficient conidia for inoculation, the culture was grown on freshly prepared potato dextrose media (39g/L; Cole-Parmer, Vernon Hills, IL) in the dark at room temperature. Conidia were collected from 2-week-old cultures by placing agar slices containing mycelium in 1% Sabouraud Maltose Broth buffer (SMB) (Difco, Sparks, MD), filtered through sterile cheesecloth to remove mycelia fragments and vortexed to release the spores. Density of the conidial suspension was determined with a hemocytometer under a microscope. The spore suspension was adjusted to $1*10^5$ spores/mL in SMB and transferred to a bottle sprayer. Berries were placed in metallic racks in plastic containers after sterilization and were uniformly sprayed with the spore suspension. To insure a high humidity for disease development, a moist sterile tissue paper was placed underneath the metallic rack. Moisture loss was minimized by covering the container with a lid. The infected berries were kept at room temperature in the dark until disease symptoms developed.

Berry infection was evaluated based on the appearance of mycelium and conidia on berry surfaces. Berries were monitored for 10 days. Disease incidence (percentage of

infected berries i.e. number of symptomatic berries out of total berries that were evaluated) and severity (proportion of each berry infected by *B. cinerea*) were assessed as described in Broome et al. (1995). Disease severity was evaluated as a visual semi-quantitative scale of notation ranging from 1 to 5. 1- no disease with no fungus establishment (highly resistant), 2- fungus establishment but growth restricted, 3- fungus growth with less than 1/3rd of the berries, 4- sufficient fungus growth with more than 50% of a berry covered and 5- full growth of fungus and berries completely covered by fungal hyphae (highly susceptible). The t-test was used to determine whether there is a significant difference between the two cultivars using five replications of 24 berries at each of the 5 stages.

Five post-veraison stages of Norton and Cabernet Sauvignon were analyzed for grape-*B. cinerea* interactions. Five 24-berry replications for each stage were used in the 2012 and 2013 experiments; the temperature and humidity during the growing seasons are shown in Supplementary Figures 1A and 1B. The average disease incidence for Norton was 1.7% (ranging from 0.0 to 4.0%) in 2012 and 12.8% (ranging from 8.0 to 17.0%) in 2013. The average disease incidence for Cabernet Sauvignon was 95.0% (ranging from 83.0 to 100.0%) in 2012; and 99.2% (ranging from 96.0 to 100.0%) in 2013 (Figure 1A). The average disease severity for Norton was 0.7% (ranging from 0.0 to 2.0%) in 2012 and 6.7% (ranging from 3.0% to 10.8%) in 2013. The average disease severity for Cabernet Sauvignon was 93.2% (ranging from 81.0 to 100.0%) in 2012 and 95.7% (ranging from 87.0 to 100.0%) in 2013 (Figure 1B). Table 1 summarizes the disease incidence and severity of the two test cultivars at five different developmental stages. There were no significant differences ($P>0.05$) in Botrytis bunch rot incidence or

severity within either Norton or Cabernet Sauvignon in the separate test stages. Figure 2 showed that the greatest difference between the two cultivars occurs 10 days post-inoculation infection of berries at E-L stage 40.

Studies of inheritance can be challenging because traits can be affected by multiple genes, the interactions between genes, and by the interactions between genes and environmental factors. Trait phenotyping is limited by environmental interactions and epistasis; both can mask the value of alleles and of individuals of interest. Thus, quantitative trait loci (QTLs) characterized in one environment may behave differently in a different environment (Wang et al., 1999). In addition, assessment of disease under field conditions becomes difficult due to disease complexes. To address these problems, an *in vitro* assay for Botrytis bunch rot was established using two grape cultivars, Norton and Cabernet Sauvignon, in 5 different post-veraison developmental stages. Use of metallic racks in a closed plastic container provided a simple and inexpensive tool to study *Vitis*-*B. cinerea* interactions. We found that inoculum concentration of $1*10^5$ conidia/ml is sufficient to show the phenotypic differences in disease development between these two cultivars. This assay greatly increases the probability of identifying variation due to genetics rather than environmental factors.

To date, there have been no documented reports of Botrytis bunch rot resistance in Norton. The low expression of the disease at post-veraison in Norton berries revealed it to be highly resistant to fungal growth and disease development. To our knowledge, this is the first demonstration of grape-*B. cinerea* interactions in both resistant and susceptible responses under laboratory conditions. A significant difference in disease incidence and severity between these two cultivars suggests that these two cultivars differ

in genetic makeup to respond to a local *B. cinerea* strain. In addition, the phenotypic differences between these two cultivars were measurable 10 days after fungal inoculation. Although there were no significant differences between stages for Botrytis bunch rot infection within a cultivar, a maximum infection was observed on berries that are at harvest (E-L stage 39 and 40) providing evidence that these berries are at the optimum stage to study disease reaction. The resistance reaction in Norton could be due to various factors, and this assay will be a valuable tool in phenotyping the hybrid population of Norton and Cabernet Sauvignon to lead to the identification of the possible causes.

This study was designed to test the differences in the resistance reactions exhibited by two different grape varieties, Norton and Cabernet Sauvignon, by establishing a unique phenotyping technique under laboratory conditions. This technique will be used for further investigation of the segregation pattern in a mapping population between Norton and Cabernet Sauvignon in order to localize the resistance-related allele(s) against Botrytis bunch rot. In addition, this experiment demonstrates that Norton is resistant to Botrytis bunch rot, providing scientific evidence to validate growers' experiences. This novel source of resistance may be a valuable asset for future grape breeding.

Figure Legends

Fig. 1 Comparisons of (A) disease incidence and (B) disease severity between Norton and Cabernet Sauvignon (CS) in 2012 and 2013. Bars represent standard errors of means. The t-test indicated a significant difference between the two cultivars with a $P < 0.05$.

Fig. 2 Botrytis bunch rot exhibited by two different grape genotypes under laboratory conditions: Norton (left) and Cabernet Sauvignon (right). Photo was taken 10 days post-inoculation.

Supplementary Fig. 1 Weather conditions (A) Relative humidity and (B) Temperature during phenotyping in 2012 and 2013.

Table 1 Average disease incidence and severity between two grape cultivars; Norton and Cabernet Sauvignon (CS) under different stages.

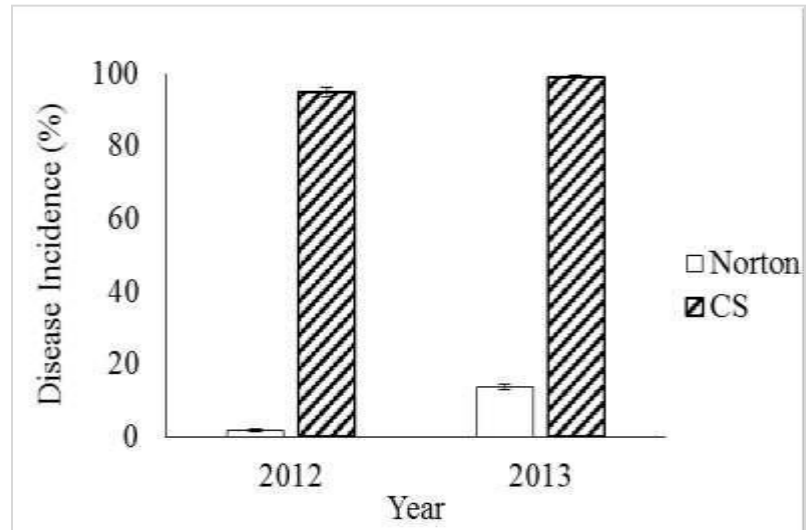
	Disease Incidence				Disease Severity			
Harvesting Stage	**2012**		**2013**		**2012**		**2013**	
	Norton	**CS**	**Norton**	**CS**	**Norton**	**CS**	**Norton**	**CS**
E-L Stage 36	4.2	100.0	8.3	100.0	1.7	96.7	3.3	96.7
E-L Stage 37	0.0	83.3	12.5	95.8	0.0	80.8	5.8	86.7
E-L Stage 38	0.0	91.7	16.7	100.0	0.0	91.7	6.7	96.7
E-L Stage 39	4.1	100.0	16.7	100.0	1.6	96.7	10.8	100.0
E-L Stage 40	0.0	100.0	10.0	100.0	0.0	100.0	6.7	98.3

Values are mean for five replicates for each stage between two cultivars for both years.

Significant differences (two-sampled t test) are designed at P>0.05.

Figure 1

A.



B.

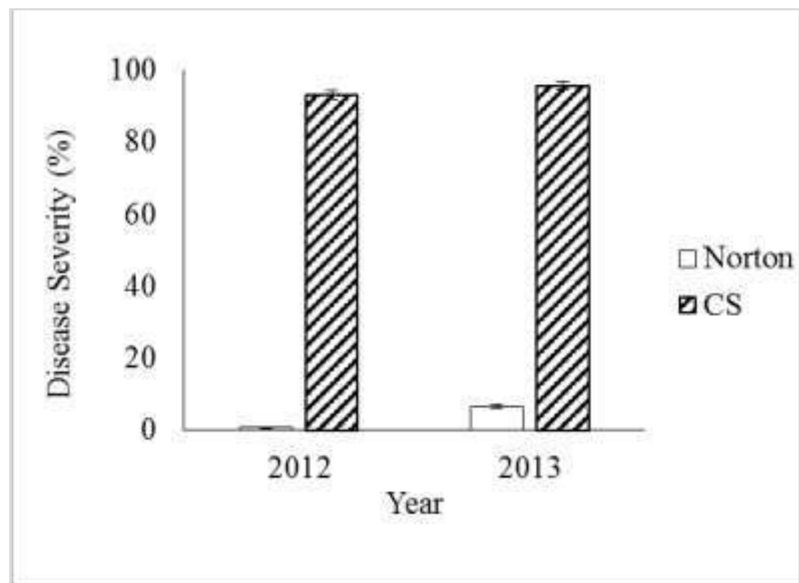
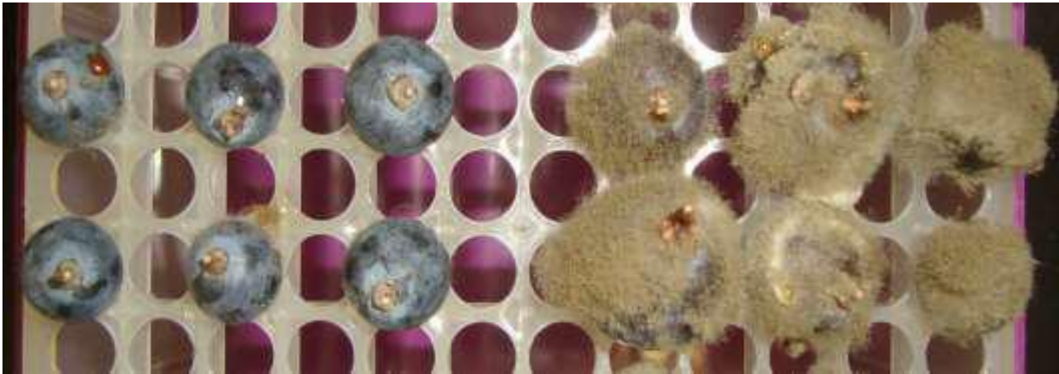
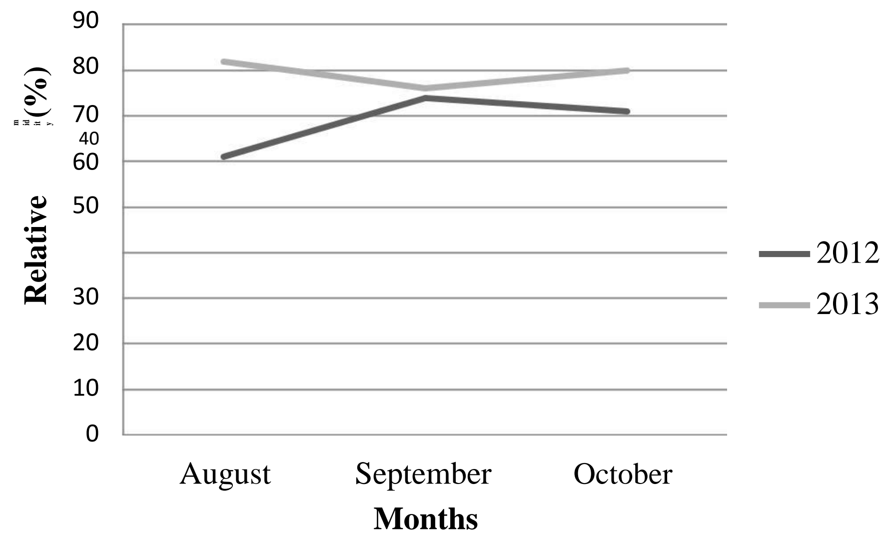


Figure 2

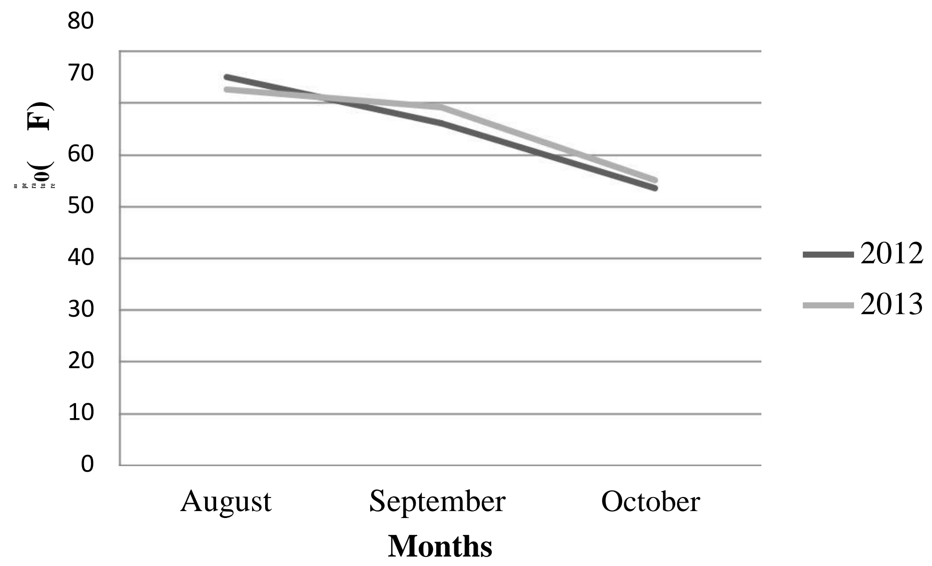


Supplementary Figure 1

A



B.



References

Ambers RK, Ambers CP (2004) Dr. Daniel Norborne Norton and the origin of the Norton grape. Amer Wine Soc J 36:77-87

Ambers CP (2013) A historical hypothesis on the origin of the Norton grape. J Wine Res 24:85-95

Broome J, English J, Marois J, Latorre B, Aviles J (1995) Development of an infection model for Botrytis bunch rot of grapes based on wetness duration and temperature. Phytopathology 85:97-102

Cadle-Davidson L (2008) Monitoring pathogenesis of natural *Botrytis cinerea* infections in developing grape berries. Am J Enol Vitic 59:387-395

Coombe BG (1995) Growth Stages of the Grapevine: Adoption of a system for identifying grapevine growth stages. Aust J Grape Wine R 1:100-110

Hedrick UP (1908) The grapes of New York. State of New York, Dept. of Agriculture 15th Annual Report, Volume III, part II. J.R. Lyon Co., Albany, NY

Hogan S, Zhang L, Li J, Zoecklein B, Zhou K (2009) Antioxidant properties and bioactive components of Norton (*Vitis aestivalis*) and Cabernet Franc (*Vitis vinifera*) wine grapes. Food Sci Technol 42:1269-1274

Holz G, Gutschow M, Coertze S, Calitz FJ (2003) Occurrence of *Botrytis cinerea* and subsequent disease suppression at different positions on leaves and bunches of grape. Plant Dis 87:351–358

Kamas J, Appel D, Black M, Hellman E, Lauziere I, Mitchell F, Morano L, Wendel L (2004) Unraveling Pierce’s Disease in its native environment. Wine Business Monthly 11:35-38

Keller M, Viret O, Cole M (2003) *Botrytis cinerea* infection in grape flowers: Defense reaction, latency and disease expression. Phytopathology 93:316-322

Kretschmer M, Leroch M, Mosbach A, Walker AS, Fillinger S, Mernke D, Schoonbeek H, Pradier JM, Leroux, P Waard, M Hahn, M (2009) Fungicide driven evolution and molecular basis of multidrug resistance in field population of grey mould fungus *Botrytis cinerea*. PLoS Patho 5:12

Kulakiotu EK, Thanassoulopoulos CC, Sfakiotakis EM (2004) Postharvest biological control of *Botrytis cinerea* on kiwifruit by volatiles of [Isabella] Grapes. Phytopathology 94:1280-1285

Percival D, Fisher K, Sullivan J (1994) Use of fruit zone leaf removal with *Vitis vinifera* L. cv. Riesling grapevines. II. Effect on fruit composition, yield, and occurrence of bunch rot (*Botrytis cinerea* Pers.: Fr.). Am J Enol Vitic 45:133-140

Pezet R, Viret O, Perret C, Tabacchi R (2003) Latency of *Botrytis cinerea* Prs.: Fr. and biochemical studies during growth and repening of two grape berry cultivars, respectively susceptible and resistant to grey mould. J Phytopathol 151:208-214

Smilanick JL, Mansour MF, Gabler FM, Margosan DA, Hashim-Buckey J (2010) Control of postharvest gray mold of table grapes in the San Joaquin Valley of California by fungicides applied during the growing season. Plant Dis 94:250-257

Stover E, Aradhya M, Yang J, Bautista J, Dangl GS (2009) Investigations into the origin of Norton grape using SSR markers. Proceedings of the Florida State Horticultural Society 122:19-24

Wang DL, Zhu J, Li ZK, Paterson AH (1999) Mapping QTLs with epistatic effects and QTL× environment interactions by mixed linear model approaches. Theor Appl Genet 99:1255-1264

CHAPTER 3

CONSTRUCTING A GENETIC LINKAGE MAP OF *VITIS AESTIVALIS*-DERIVED 'NORTON' USING SIMPLE SEQUENCE REPEAT MARKERS

Hammers, M., Sapkota, S., Chen, L. L., & Hwang, C. F. (2017) Constructing a genetic linkage map of *Vitis aestivalis*-derived "Norton" and its use in comparing Norton and Cynthiana. Mol Breed 37: 64

Introduction

Vitis aestivalis-derived 'Norton', the official grape of the State of Missouri, is a historic cultivar and remains economically important to the Central Appalachian Mountain region and South-Central Great Plains area of the United States (Morton 1985). It is grown in many US regions where *V. vinifera* (the European grape used for most winemaking worldwide; e.g., Cabernet Sauvignon) production would require extensive pesticide use for fungal disease management (Stover et al. 2009; Sapkota et al. 2015). In addition to fungal disease resistance, Norton displays tolerance to Pierce's disease (Kamas 2014) and resistance to Phylloxera (Hedrick 1908). It is reported to be a grape of American origin that was produced as a hybrid seedling in the Virginia garden of Dr. Daniel Norborne Norton. The cultivar is believed to be the result of an unintentional cross between the lost grapevine 'Bland' (a cultivar hypothesized to have *V. labrusca* and *V. vinifera* parentage) and the native grapevine, *V. aestivalis* (Ambers 2013). Microsatellite, or simple sequence repeat (SSR), marker analysis showed that Norton contains alleles from the grape species *V. vinifera* and *V. aestivalis.* This supports the possibility that Norton is a *V. aestivalis* and *V. vinifera* hybrid (Stover et al. 2009).

Norton and Cynthiana grape cultivars are both described to be largely derived from *V. aestivalis* (Parker et al. 2005; Stover et al. 2009). However, the precise origin of the two cultivars can only be hypothesized. Norton and Cynthiana vines have become increasingly popular in Missouri and Arkansas, respectively. Early records report that Norton was introduced into Missouri vineyards in the late 1840s while Cynthiana was introduced in the late 1850s (Husmann 1883). Since that time, it has been speculated that Norton and Cynthiana are actually the same cultivar because many phenotypic

similarities have been noted between them. For example, both display similar cluster, berry, and peduncle sizes (Main and Morris 2004). They also display resistance to many fungal diseases, such as powdery mildew, downy mildew, and a variety of berry rots, which have caused the most extensive amount of damage to vineyards across the world (Harris 2012). Other resemblances include the difficulty of establishing roots from dormant hardwood cuttings (Keeley et al. 2003) and a high sensitivity to sulfur spray (Bordelon et al. 2007). Both vines are cold hardy, withstanding temperatures as low as -32 °C, and require a long growing season (~125 days) to fully ripen (Dami et al. 2005). The two cultivars produce a dry, red wine with high titratable acidity (8.5 to 13 g/L) that may be attributed to the amount of malic acid present within the fruit (Main and Morris 2004). Some phenotypic differences also exist between Norton and Cynthiana. For instance, differences in the ideal soil type have been noted. Although they both grow well in sandy soils, Cynthiana thrives in a loamy soil better than Norton, and Norton thrives in a clay soil better than Cynthiana (Harris 2012). Some researchers have accepted Norton and Cynthiana as the same cultivar, but many growers and winemakers still assert that distinctions exist in their respective viticultural performance and enological quality (Morris and Main 2010). Reisch et al. (1993) provided preliminary evidence that Norton and Cynthiana are genetically indistinguishable using isozyme analysis, but only five banding patterns were used for identification. Similarly, in a study by Parker et al. (2009), four microsatellite loci were used to identify Norton and Cynthiana as genetically synonymous cultivars. A genome-wide assessment using microsatellites may help to confirm or refute conclusions drawn from these previous studies.

SSR markers are widely used to distinguish cultivars and to assess genetic relationships among them (Bautista et al. 2008). They are very valuable in molecular breeding because they are PCR-derived, polymorphic, and co-dominant (Merdinoglu et al. 2005). SSR markers are often used in *V. vinifera* genetic analyses (Adam-Blondon et al. 2004) but have become increasingly used in other grapevine species due to their high interspecies transferability (Doligez et al. 2006; Li et al. 2013). They have been implemented for rootstock identification (Lin and Walker 1998), surveys of germplasm (Giannetto et al. 2010), comparisons of cultivars (Lefort and Roubelakis-Angelakis 2001) and breeding for disease resistance (Riaz et al. 2009). In addition, several SSR-based linkage maps have been developed to identify quantitative trait loci (QTLs) linked with various traits for marker-assisted selection (Doucleff et al. 2004; Fischer et al. 2004; Hoffmann et al. 2008; Bellin et al. 2009; Blasi et al. 2011; Zhang et al. 2009; Riaz et al. 2006, 2009 and 2011; van Heerden et al. 2014). However, these resources are lacking for *V. aestivalis* and the current genetic analyses are extremely limited.

This study details the construction of a genetic linkage map for *V. aestivalis*-derived 'Norton' grape based on microsatellite markers.

Materials and Methods

Plant material

Crosses of *V. aestivalis*-derived 'Norton' x *V. vinifera* 'Cabernet Sauvignon' were made in 2005 and resulted in 93 hybrid progeny, of which 19 were from a reciprocal cross. This population was expanded to 310 genotypes in 2011 as described by Adhikari et al. (2014). Briefly, young seedlings were transferred to half-gallon containers

following germination until they reached the 8 to 10-leaf stage. These seedlings were then transferred to a shade house for two weeks and then into an open area for a week before being transplanted to the vineyard. Vines were planted at 10 foot between-row and 8 foot within-row spacing and were trained to a high bilateral cordon system. All of the vines were maintained at the Missouri State Fruit Experiment Station (MSFES) in Mountain Grove, Missouri. For this study, the mapping population consisted of 183 progeny by combining 90 genotypes from 2011 with the 93 generated in 2005.

Marker selection and identification of polymorphic markers

The primer pairs flanking SSR loci were obtained from previously published literature (Table 1), NCBI database uniSTS (http://www.Ncbi.nlm.nih.gov), *Vitis*-EST database (http://cgf.ucdavis.edu) and Grape Genome Browser-Genoscope (http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/). Primers were synthesized by Sigma–Aldrich (St. Louis, MO). To identify polymorphic markers for genetic map construction, SSR markers were first tested on two of the parents and six of their progeny.

DNA extraction, PCR amplification and SSR genotyping

DNA extraction and PCR amplification were carried out as previously described by Adhikari et al. (2014). Three primers were used in a PCR reaction: an M13-tailed 5’-TGTAAAACGACGGCCAGT-3’ forward primer, a reverse primer and a WellRED (D2-black, D3-green or D4-blue) labeled M13 sequence. This not only allowed for the fluorescent labeling of PCR products but also pooling for capillary electrophoresis

separation using the Beckman Coulter CEQ 8000 (Beckman Coulter, Inc., Brea CA). PCR amplification was confirmed using a 1.5 % agarose gel followed by fragment analysis using a GenomeLab GeXP capillary sequencer (Beckman Coulter, Brea, CA). Fragment sizes were double checked, and any ambiguous genotypes were re-amplified and re-analyzed or labeled as unknown.

Linkage map construction

To construct genetic maps of *V. aestivalis*-derived 'Norton', *V. vinifera* 'Cabernet Sauvignon' and the consensus map, polymorphic markers were scored using JoinMap 4.1 (van Ooijen 2006). Initially, parental maps were constructed using an integrated two-way pseudo-testcross strategy described by Grattapaglia and Sederoff 1994. The consensus map was generated by merging the two parental maps with segregating markers for cross-pollinated populations (aa x ab, ab x aa, ab x ab, ab x ac and ab x cd). Segregation of markers were evaluated for goodness-of-fit for both observed and expected Mendelian ratios 1:1 ($P \leq 0.05$) using a Chi-square method to identify distorted markers. Distorted markers were included if they did not hinder linkage map calculations and order of the markers. Markers were grouped using recombination frequency parameters (start 0.25, end 0.05) and were further confirmed with a test for independence using a logarithm of odds (LOD) value of 6.0. Marker order within the group was calculated using a regression mapping algorithm until the third round and verified with the maximum likelihood method. Map units in centimorgan (cM) were generated using Kosambi's mapping function (Kosambi 1944). The linkage groups were numbered 1 to 19 according to internationally acknowledged grapevine reference genetic maps (Doligez et al. 2006).

The final parental and consensus maps were aligned using the software Map Chart (Voorrips 2002).

Results

Marker analysis including segregation distortion

The parents and six randomly chosen hybrid progeny were first tested with 1,157 microsatellite flanking primer pairs. Of these, 859 of the primer pairs produced PCR products and a further subset of 413 markers were identified as polymorphic. These markers were used to screen the entire mapping population of 183 genotypes (Table 1). The remaining 446 SSR markers were unusable because they were either monomorphic or displayed an unclear banding pattern. Among the polymorphic markers screened, 159 (38.5%) primer pairs were fully informative (ab x cd), 123 (29.8%) primer pairs segregated for Norton (ab x aa), 34 (8.2%) primer pairs segregated for Cabernet Sauvignon (aa x ab), 3 (<1 %) primer pairs were double heterozygous segregating with two alleles (ab x ab), and 94 (22.8%) primer pairs segregated with three alleles (ab x ac) for both parents.

Parent and consensus maps

Of the 413 polymorphic SSR markers, 376 anchored on the Norton linkage map and 279 on the Cabernet Sauvignon linkage map (Table 2; Fig. 1). Three markers (VMC2E11, UDV-108 and UDV-093) that were polymorphic for Norton did not link with the map. Similarly, in Cabernet Sauvignon, only 279 out of 290 polymorphic SSR markers were mapped. The markers VMC2H10, VMC3C9, VMC3F8, VVIB66, FAM40,

UDV-090, VVCS1H018018F1-1, FAM54, VVIU04, UDV-108 and VVCh14-37 were not mapped on Cabernet Sauvignon linkage groups. The two maps were combined to develop an integrated genetic linkage map. In the consensus map, 411 out of 413 SSR markers were aligned into 19 linkage groups. Two SSR markers (FAM54 and UDV-108) did not map to any linkage group in the consensus map.

Map coverage and estimation of genomic size

Table 2 summarizes the distribution of SSR marker coverage among Norton, Cabernet Sauvignon and the consensus maps. In Norton, 19 linkage groups spanned a total genetic distance of 1496.6 cM with an average marker distance of 4.0 cM. The largest number of markers were anchored on linkage group 14 (45) covering a genetic distance of 96.1 cM and the fewest number were anchored on linkage group 15 (7) with a genetic distance of 56.6 cM. In Cabernet Sauvignon, a total genetic distance of 1579.7 cM was covered by 19 linkage groups with an average marker distance of 5.7 cM. The largest number of markers were linked on chromosome 4 (27) and the fewest were linked on chromosome 15 (5). In the consensus map, 19 linkage groups spanned a total genetic distance of 1678.6 cM with an average marker distance of 4.1 cM. Linkage group 18 was the largest (134.6 cM) with a total of 40 SSR markers (9.7 %) and linkage group 15 was the smallest (55.9 cM) with only 7 SSR markers (1.7 %). In general, marker order was consistent between the Norton and Cabernet Sauvignon maps and the consensus map with only small inversions. There are 26 gaps that are > 10 cM in the consensus map. The largest gap of 23.1 cM is on linkage group 11. Linkage groups 1, 4, 9, 14 and 18 possess no gap greater than 10 cM (Table 2; Fig. 1).

Segregation ratios were tested using Chi-square for both the parental and consensus maps. The Chi-square test showed distortion from the normal segregation ratio ($P \leq 0.05$) in 78 (20.6 %) SSR markers on the Norton map, 75 (25.9 %) markers on the Cabernet Sauvignon map, and 131 (31.7 %) markers on the consensus map (Table 2). The distorted markers were included in the final genetic linkage map. Table 3 shows a comparison of the Norton linkage map with the internationally acknowledged *V. vinifera* reference map (Doligez et al. 2006). The Norton map covered 93.0 % of the reference map. Furthermore, the average genetic ratio was 1.0 when the common markers between the maps were compared.

Discussion

The Norton grape has experienced a resurgence in popularity as a wine grape, especially in Southern and Midwestern states including Missouri, Arkansas and Virginia (Ambers and Ambers 2004; Kliman 2010). Despite having multiple useful traits, very little information is available about the genetics of its disease resistance, cold hardiness and berry quality. In view of this, a mapping population of 183 individuals was constructed from a cross between Norton and Cabernet Sauvignon. A linkage mapping strategy was able to generate 19 different linkage groups corresponding to the 19 chromosomes of the reference genome. We present here the first genetic linkage map of *V. aestivalis*-derived 'Norton' using SSR markers.

Four hundred and thirteen (35.8 %) SSR markers were used for the construction of the genetic linkage map. A segregation analysis showed 78 (20.6%) markers distorted in Norton and 75 (25.9%) markers distorted in Cabernet Sauvignon. In the consensus

map, 131 (31.7%) markers were distorted from the normal segregation ratio ($P \leq 0.05$) which is slightly lower than the map from grapevine V3125 x 'Börner' published by Zhang et al. 2009 (33.6%). However, this is higher than percentages previously observed in *V. vinifera* and *Vitis* hybrid crosses reported by Doligez et al. 2006 (7-11%), Di Gaspero et al. 2007 (3-13%), Troggio et al. 2007 (20.3%) and Salmaso et al. 2008 (19-20%) as well as in wide intra- and interspecific crosses reported by Grando et al. 2003 (22.4%), Lowe and Walker 2006 (16.0%) and Riaz et al. 2006 (17%). The predominant distorted markers were fully informative and heterozygous on both parents (ab x cd). The unbalanced segregation for these loci could be due to various reasons including discordance of parents and major chromosomal rearrangements (Guo et al. 2014). Out of the 19 linkage groups, the largest number of markers were distorted on linkage group 5 (18) followed by linkage group 1 (10). Linkage groups 2, 13, 15 and 19 did not possess any distorted markers. Though the cause(s) of skewed distribution is still unclear, Riaz et al. (2006) pointed out the possibility that segregation distortion regions may influence specific linkage groups. The availability of the Norton genome sequence as well as the phenotypic effect of the genes in those regions may help to understand any underlying mechanism.

In addition, a comparison of linkage groups between the parental and consensus maps shows that linkage groups 3 and 9 in Norton are nearly half the size than in the other two genetic maps. In Cabernet Sauvignon, linkage groups 6, 8 and 15 are much shorter than they are in the other two maps (Fig. 1). The difference in chromosome size and the presence of distorted markers could be attributed to difficulties in recombination events between the homologous chromosomes (Welter et al. 2007; Zhang et al. 2009).

Paterson et al. (1990) and Causse et al. (1994) had also reported a negative correlation between genetic distance and recombination events in different studies. Integration of parental maps reduced the number of gaps greater than 10 cM (Table 2; Fig. 1). The presence of larger gaps may either be due to fewer recombination events or unavailability of sufficient polymorphic markers in those regions. The marker order of this genetic map was consistent with the reference genetic map from five different *V. vinifera* crosses (Doligez et al. 2006) and to maps involving resistance donor parents (Zhang et al. 2009; Moreira et al. 2011). Slight differences were observed on linkage groups 8, 9 and 13 suggesting no major chromosomal rearrangements. These results further indicate a high level of synteny among *Vitis* species.

Figure Legends

Fig. 1 SSR marker-based genetic map of *Vitis aestivalis*-derived 'Norton (N)', consensus (C) and *V. vinifera* 'Cabernet Sauvignon (CS).' The parental maps are on either side with consensus in the middle. Chromosomes are numbered from 1 to 19 according to the internationally acknowledged grapevine reference genetic map (Doligez et al. 2006), with prefix N, C and CS. Dotted lines link common markers among the genetic maps representing macrosynteny.

Table 2 SSR markers tested for polymorphism and segregation pattern on the Norton (N) and Cabernet Sauvignon (CS) mapping population

Marker series	No. tested	No. Amplified	Useful for this map	N (abxaa)	CS (aaxab)	N and CS (abxab)	N and CS (abxac)	N and CS (abxcd)	References
A, B, C, CB, CD, EE, GB,NS, Rgamu, STS,Vamu	84	69	19	10	2	0	4	3	NCBI database UniSTS
Chr	4	4	2	1	1	0	0	0	Blasi et al. 2011
AF, CF, ctg,	107	87	45	15	1	0	14	15	EST database; Riaz et al. 2006
EST	8	8	4	2	0	0	0	2	Sevini et al. 2003
FAM	136	116	55	21	8	1	14	11	Huang et al. 2011
C_, PSCT, SC8, SSR, VVCN, VVCS	125	98	56	17	6	0	12	21	Genoscope
SCU	16	9	3	3	0	0	0	0	Scott et al. 2000
VMC, VMCNG	356	203	105	17	8	1	28	51	*Vitis* Microsatellite Consortium
VVC	8	7	6	1	1	0	1	3	Decroocq et al. 2003
VVch	48	34	10	8	0	0	2	0	Riaz et al. 2009
VVI	138	118	53	17	4	0	9	23	Merdinoglu et al. 2005
VVS	9	6	3	0	0	0	1	2	Thomas and Scott, 1993
VVMD	22	18	10	2	0	1	2	5	Bowers et al. 1996 and 1999
VrZAG	9	8	7	3	0	0	1	3	Sefc et al. 1999
UDV, VRIP	87	74	35	6	3	0	6	20	Di Gaspero et al. 2005 and 2007
Total	1157	859 (74.2%)	413 (35.7%)	123	34	3	94	159	

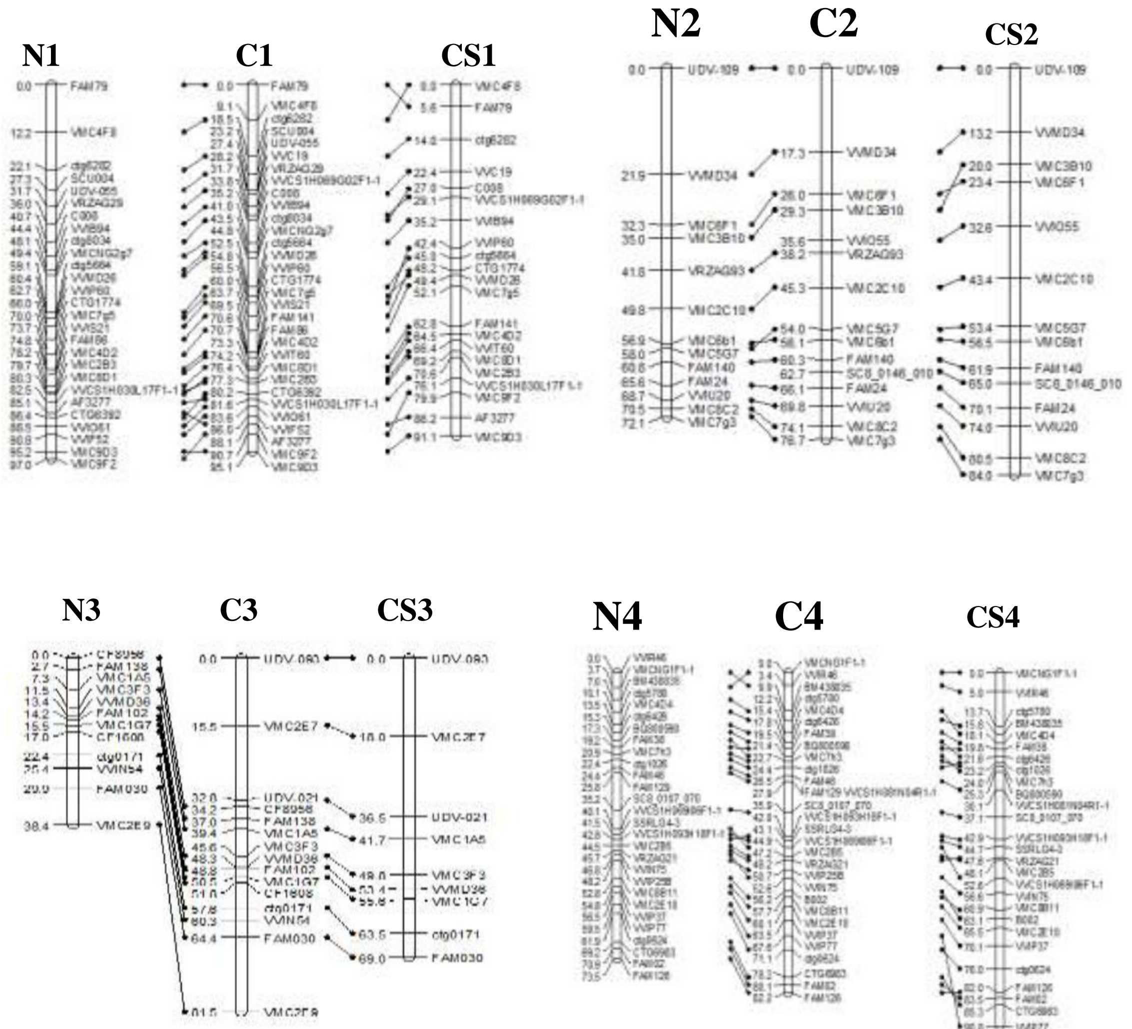
Table 3 Distribution of SSR markers among Norton, Cabernet Sauvignon and Consensus maps and the map coverage

	Norton			Cabernet Sauvignon			Concensus			No. of gaps >10 cM
LGs	No. of SSR markers	Length (cM)	% of total markers	No. of SSR markers	Length (cM)	% of total markers)	No. of SSR markers	Length (cM)	% of total markers	
1	27	97	7.2	21	91.1	7.5	31	95.1	7.5	0
2	13	72.1	3.5	14	84	5	15	76.7	3.6	1
3	12	38.4	3.2	9	69	3.2	15	81.5	3.6	3
4	28	73.5	7.4	27	90.8	9.7	30	82.2	7.3	0
5	24	80.2	6.4	22	84.9	7.9	26	84.9	6.3	1
6	14	76.7	3.7	9	55.3	3.2	18	91.6	4.4	2
7	24	100.6	6.4	22	124.7	7.9	27	113	6.6	1
8	23	82.6	6.1	11	56.6	3.9	24	88.3	5.8	1
9	11	50.8	2.9	11	106	3.9	16	97.9	3.9	0
10	16	78.8	4.3	12	65.5	4.3	16	82.4	3.9	3
11	23	91.5	6.1	15	92.2	5.4	23	101.8	5.6	2
12	15	72.5	4	13	87.2	4.7	15	83.4	3.6	2
13	13	70.2	3.5	9	61.9	3.2	13	66.1	3.2	2
14	45	96.1	12	24	127.1	8.6	48	104.7	11.7	0
15	7	56.6	1.9	5	29.6	1.8	7	55.9	1.7	2
16	19	99.4	5.1	15	77.4	5.4	19	92.6	4.6	2
17	12	63.8	3.2	10	71.2	3.6	14	73.4	3.4	3
18	36	122.9	9.6	20	132.2	7.2	40	134.6	9.7	0
19	14	72.9	3.7	10	73.3	3.6	14	72.6	3.4	1
Total mapped markers	376	1496.6		279	1579.7		411	1678.6		
Total SSR markers	379			290			413			
Ave. marker distance (cM)		4			5.7				4.1	
No. of distorted loci (P ≤0.05)	78(20.6%)								131(31.7%)	

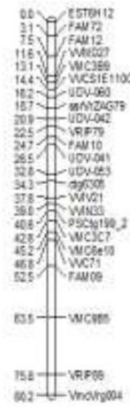
Table 3 Genome coverage and genetic distance comparisons between *Vitis aestivalis*-derived Norton and *V. Vinifera* reference linkage maps

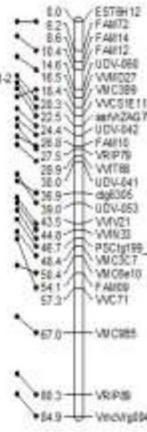
Markers common between maps			Genetic distance between common marker			Total genetic distance			
LGs	Start Marker	End Marker	*V. vinifera* reference map	*V. aestivalis*-derived Norton map	Genetic ratio	*V. vinifera* reference map	*V. aestivalis*-derived Norton map	Genetic ratio	Physical distance ratio
1	VMC4F8	VMC9D3	87.5	83.0	0.95	87.5	97.0	1.11	1.17
2	UDV-109	VMC7g3	79.7	76.7	0.96	79.7	72.1	0.90	0.94
3	VMC1A5	VMC2E9	51.2	31.1	0.61	70.3	38.4	0.55	0.90
4	VVIR46	VVIP77	66.8	59.2	0.89	90.9	73.5	0.81	0.91
5	VVMD27	VMC9B5	52.8	51.9	0.98	83.4	80.2	0.96	0.98
6	UDV085	VVIM43	57.9	44.3	0.77	82.5	76.7	0.93	1.22
7	VVMD7	VVIV04	87.9	100.6	1.14	102.7	100.6	0.98	0.86
8	VMC2F12	VMC2H10	95.1	75.0	0.79	112.7	82.6	0.73	0.93
9	VMC6D12	VVIV37	8.3	11.9	1.43	104.1	50.8	0.49	0.34
10	VVIH01	Vrzag25	28.3	26.7	0.94	83.7	78.8	0.94	1.00
11	VMC2A12	VVIB19	28.8	26.8	0.93	75.1	91.5	1.22	1.31
12	VMC8G6	VMC8G9	60.9	51.3	0.84	81.9	72.5	0.89	1.05
13	VMCNG4E10.1	VMC3B12	23.5	25.1	1.07	101.1	70.2	0.69	0.65
14	VMCNG1E1	VVIN94	87.8	83.8	0.95	94.8	96.1	1.01	1.06
15	VVIV67	VMC5G8	22.2	31.2	1.41	37.9	56.6	1.49	1.06
16	UDV013	VMC5A1	63.6	80.1	1.26	92.4	99.4	1.08	0.85
17	VMC3C11.1	VMC9G4	44.5	42.5	0.96	58.0	63.8	1.10	1.15
18	VMC3E5	VVMD17	84.7	88.9	1.05	131.5	122.9	0.93	0.89
19	VMC9A2.1	VVIV33	49	59.6	1.22	76.6	72.9	0.95	0.78
Total			1080.5	1049.7		-1647.1	1496.6	-	-
Average					1.0			0.93	0.95

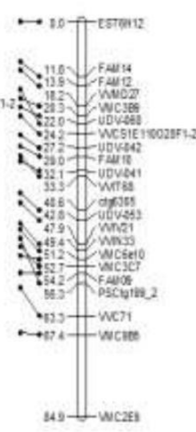
Figure 1

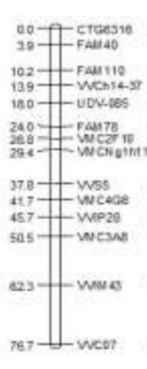


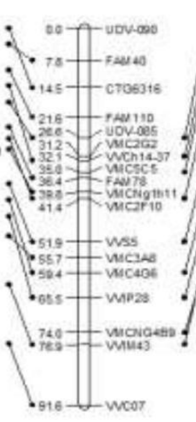
N5 **C5** **CS5** **N6** **C6** **CS6**

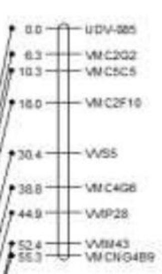






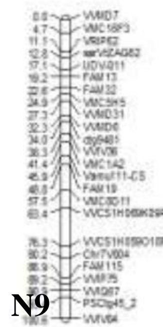


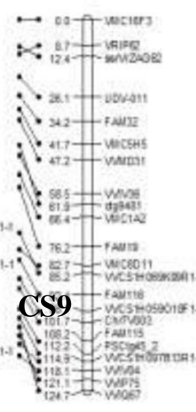


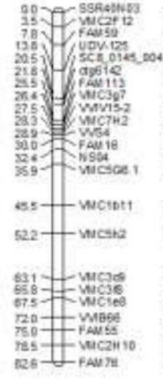


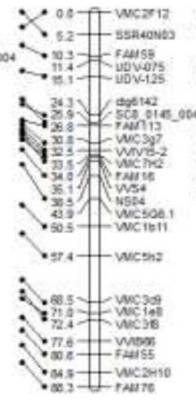
…N7 **C7…………..CS7** **N8** **C8** **CS8**

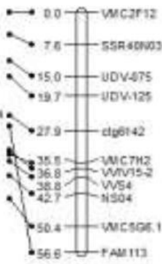
N9 **C9** **CS9**

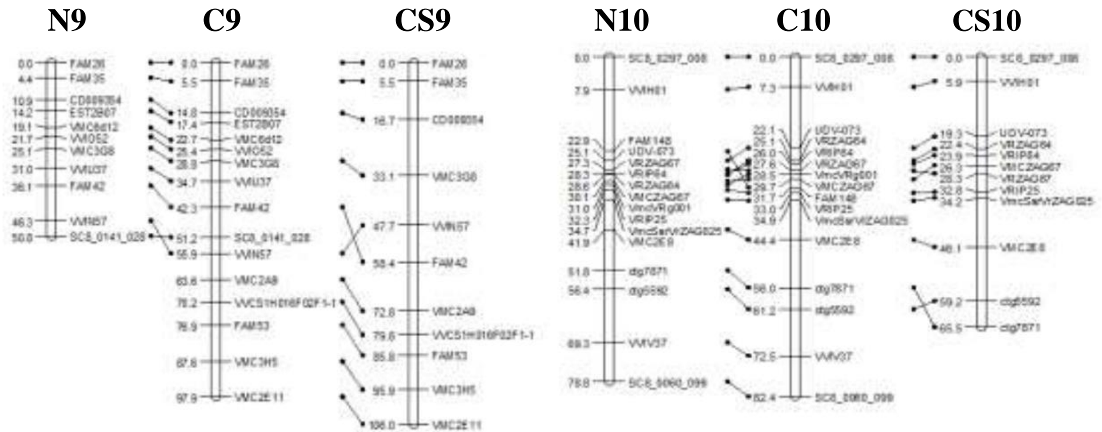


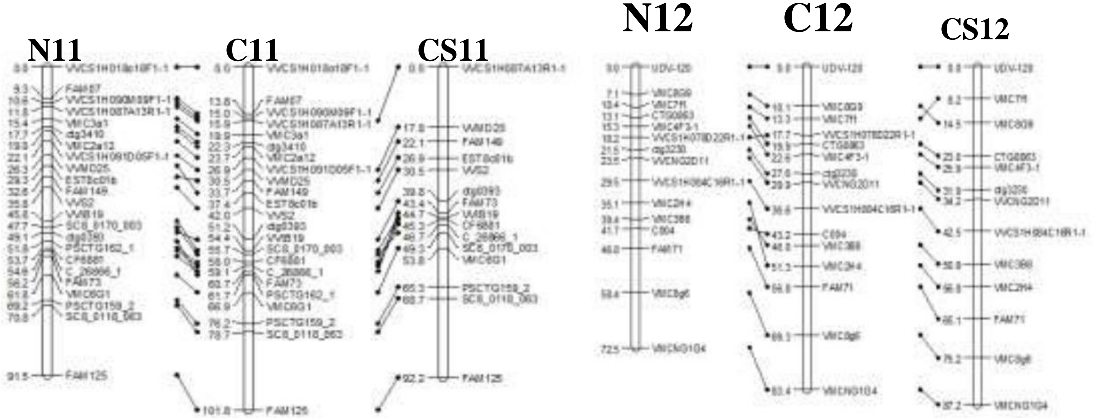


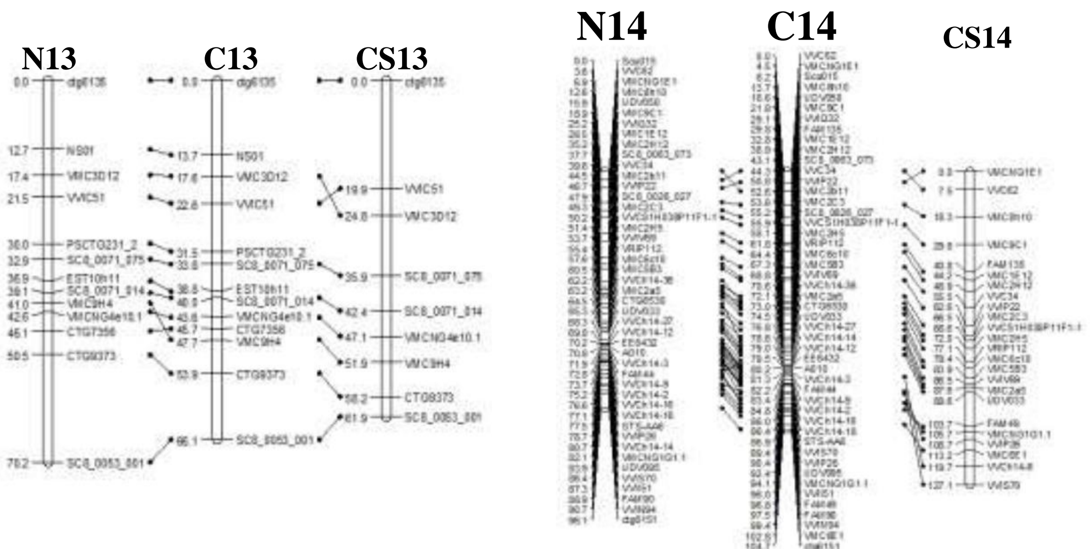


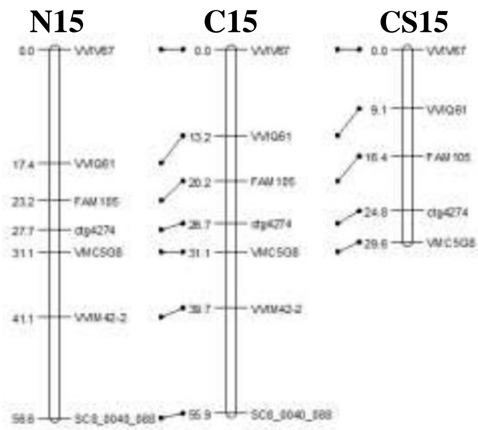


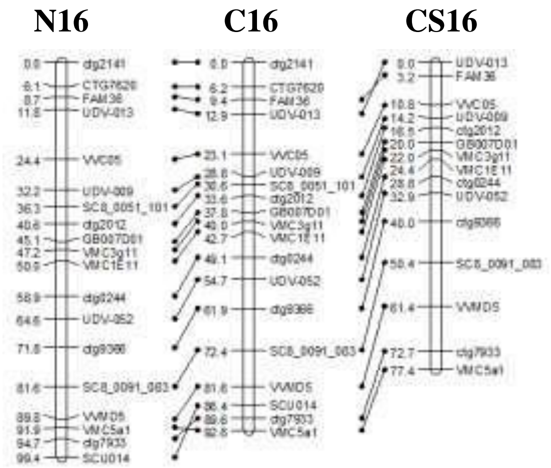


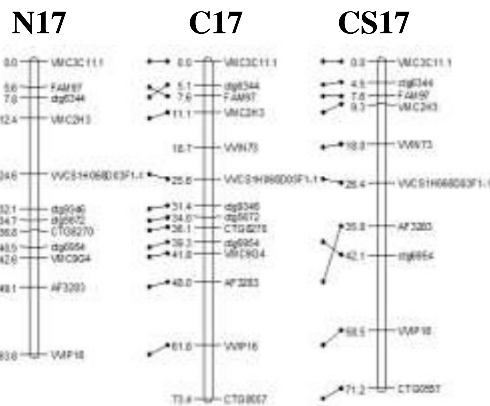

N9
C9
CS9
N10
C10
CS10

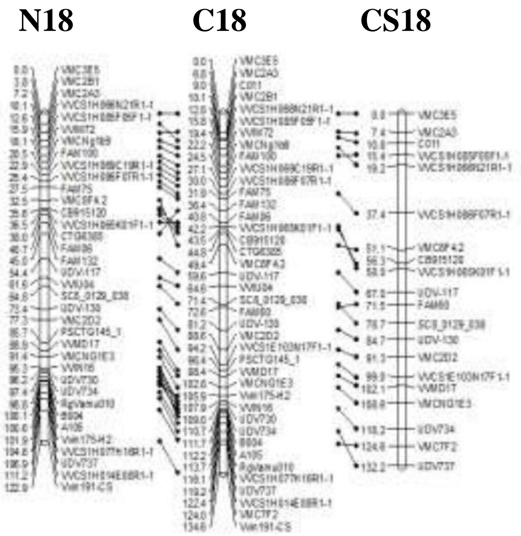

N11
C11
CS11
N12
C12
CS12

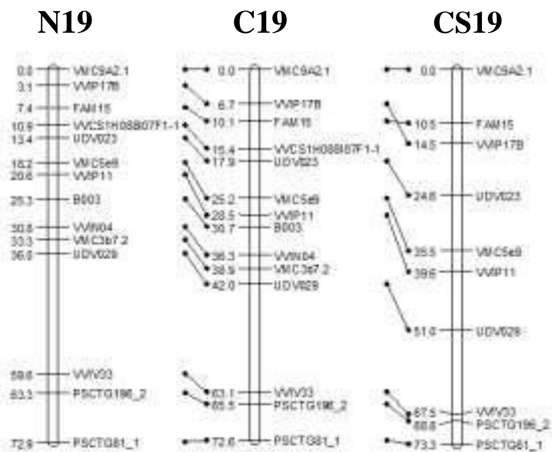

N13
C13
CS13
N14
C14
CS14


N15
C15
CS15


N16
C16
CS16


N17
C17
CS17


N18
C18
CS18


N19
C19
CS19

References

Adam-Blondon A-F, Roux C, Claux D, Butterlin G, Merdinoglu D, This P (2004) Mapping 245 SSR markers on the *Vitis vinifera* genome: a tool for grape genetics. Theor Appl Genet 109:1017-1027

Adhikari P, Chen L-L, Chen X, Sapkota SD, Hwang C-F (2014) Interspecific hybrid identification of *Vitis aestivalis*-derived 'Norton'-based populations using microsatellite markers. Sci Hortic 179:363-366

Ambers RK, Ambers CP (2004) Dr. Daniel Norborne Norton and the origin of the Norton grape. Amer Wine Soc J 36:77-87

Ambers CP (2013) A historical hypothesis on the origin of the Norton grape. J Wine Res 24:85-95

Bautista J, Dangl GS, Yang J, Reisch B, Stover E (2008) Use of genetic markers to assess pedigrees of grape cultivars and breeding program selections. Am J of Enol Vitic 59:248-254

Bellin D, Peressotti E, Merdinoglu D, Wiedemann-Merdinoglu S, Adam-Blondon AF, Cipriani G, Morgante M, Testolin R, Di Gaspero G (2009) Resistance to *Plasmopara viticola* in grapevine 'Bianca' is controlled by a major dominant gene causing localised necrosis at the infection site. Theor Appl Genet 120:163-176

Blasi P, Blanc S, Wiedemann-Merdinoglu S, Prado E, Ruhl EH, Mestre P, Merdinoglu D (2011) Construction of a reference linkage map of *Vitis amurensis* and genetic mapping of *Rpv8*, a locus conferring resistance to grapevine downy mildew. Theor Appl Genet 123:43-53

Bordelon BP, Ellis MA, Foster R (2007) Midwest commercial small fruit and grape spray guide. Purdue University Extension Publication ID-169: West Lafayette, IN

Bowers JE, Dangl GS, Vignani R, Meredith CP (1996) Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). Genome 39:628-633

Bowers JE, Dangl GS, Meredith CP (1999) Development and characterization of additional microsatellite DNA markers for grape. Am J Enol Vitic 50:243-246

Causse MA, Fulton TM, Cho YG, Ahn SN, Chunwongse J, Wu K, Xiao J, Yu Z, Ronald PC, Harrington SE (1994) Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138:1251-1274

Dami I, Bordelon B, Ferree DC, Brown M, Ellis MA, Williams RN, Doohan D (2005) Midwest grape production guide. Ohio State University Extension 919:19

De Andrés M, Benito A, Perez-Rivera G, Ocete R, Lopez M, Gaforio L, Muñoz G, Cabello F, Martínez Zapater Jm, Arroyo-García R (2012) Genetic diversity of wild grapevine populations in Spain and their genetic relationships with cultivated grapevines. Mol Ecol 21:800-816

Decroocq V, Fave M, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor Appl Genet 106:912-922

Di Gaspero G, Cipriani G, Marrazzo MT, Andreetta D, Castro MJP, Peterlunger E, Testolin R (2005) Isolation of (AC) n-microsatellites in *Vitis vinifera* L. and analysis of genetic background in grapevines under marker assisted selection. Mol Breed 15:11-20

Di Gaspero G, Cipriani G, Adam-Blondon A-F, Testolin R (2007) Linkage maps of grapevine displaying the chromosomal locations of 420 microsatellite markers and 82 markers for R-gene candidates. Theor Appl Genet 114:1249-1263

Doligez A, Adam-Blondon AF, Cipriani G, Di Gaspero G, Laucou V, Merdinoglu D, Meredith CP, Riaz S, Roux C, This P (2006) An integrated SSR map of grapevine based on five mapping populations. Theor Appl Genet 113:369-382

Doucleff M, Jin Y, Gao F, Riaz S, Krivanek AF, Walker MA (2004) A genetic linkage map of grape, utilizing *Vitis rupestris* and *Vitis arizonica*. Theor Appl Genet 109:1178-1187

Fatahi R, Ebadi A, Bassil N, Mehlenbacher S, Zamani Z (2015) Characterization of Iranian grapevine cultivars using microsatellite markers. Vitis 42:185-192

Fischer B, Salakhutdinov I, Akkurt M, Eibach R, Edwards K, Töpfer R, Zyprian E (2004) Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. Theor Appl Genet 108:501-515

Giannetto S, Caruana R, La Notte P, Costacurta A, Crespan M (2010) A survey of Maltese grapevine germplasm using SSR markers. Am J Enol Vitic 61:419-424

Grando MS, Bellin D, Edwards K, Pozzi C, Stefanini M, Velasco R (2003) Molecular linkage maps of *Vitis vinifera* L. and *Vitis riparia* Mchx. Theor Appl Genet 106:1213-1224

Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137:1121-1137

Guo Y, Lin H, Liu Z, Zhao Y, Guo X, Li K (2014) SSR and SRAP marker-based linkage map of *Vitis vinifera* L. Biotechnol Biotechnol Equip 28:221-229

Harris J (2012) The great grape debate. The Midwest Wine Grower 4:19-21

Hedrick UP (1908) The grapes of New York. State of New York, Dept. of Agriculture 15th Annual Report, Volume III, part II. J.R. Lyon Co. State Printers, Albany

Hoffmann S, Di Gaspero G, Kovacs L, Howard S, Kiss E, Galbacs Z, Testolin R, Kozma P (2008) Resistance to *Erysiphe necator* in the grapevine 'Kishmish vatkana' is controlled by a single locus through restriction of hyphal growth. Theor Appl Genet 116:427-438

Huang XC, Quesada MA, Mathies RA (1992) DNA sequencing using capillary array electrophoresis. Anal Chem 64:2149-2154

Huang H, Lu J, Ren Z, Hunter W, Dowd SE, Dang P (2011) Mining and validating grape (*Vitis* L.) ESTs to develop EST-SSR markers for genotyping and mapping. Mol Breed 28:241-254

Husmann G (1883) American grape growing and wine making. Orange Judd Co., New York

Kamas J (2014) Growing grapes in Texas: from the commercial vineyard to the backyard vine. Texas A&M University Press, College Station, TX

Keeley K, Preece JE, Taylor BH (2003) Increased rooting of 'Norton'grape cuttings using auxins and gibberellin biosynthesis inhibitors. HortScience 38:281-283

Kliman T (2010) The wild vine: a forgotten grape and the untold story of American wine. Clarkson Potter, NY, USA

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172-175

Lefort F, Roubelakis-Angelakis KK (2001) Genetic comparison of Greek cultivars of *Vitis vinifera* L. by nuclear microsatellite profiling. Am J Enol Vitic 52:101-108

Li C, Erwin A, Pap D, Coleman C, Higgins AD, Kiss E, Kozma P, Hoffmann S, Ramming DW, Kovács LG (2013) Selection for *Run1-Ren1* dihybrid grapevines using microsatellite markers. Am J Enol Vitic 64:152-155

Lin H, Walker MA (1998) Identifying grape rootstocks with simple sequence repeat (SSR) DNA markers. Am J Enol Vitic 49:403-407

Lopes MS, Sefc KM, Dias EE, Steinkellner H, Machado MLC, Machado AC (1999) The use of microsatellites for germplasm management in a Portuguese grapevine collection. Theor Appl Genet 99:733-739

Lowe KM, Walker MA (2006) Genetic linkage map of the interspecific grape rootstock cross Ramsey (*Vitis champinii*)× Riparia Gloire (*Vitis riparia*). Theor Appl Genet 112:1582-1592

Main GL, Morris JR (2004) Leaf-removal effects on Cynthiana yield, juice composition, and wine composition. Am J Enol Vitic 55:147-152

Martín JP, Borrego J, Cabello F, Ortiz JM (2003) Characterization of Spanish grapevine cultivar diversity using sequence-tagged microsatellite site markers. Genome 46:10-18

Martinez L, Cavagnaro P, Masuelli R, Zuniga M (2006) SSR-based assessment of genetic diversity in South American *Vitis vinifera* varieties. Plant Sci 170:1036-1044

Merdinoglu D, Butterlin G, Bevilacqua L, Chiquet V, Adam-Blondon A-F, Decroocq S (2005) Development and characterization of a large set of microsatellite markers in grapevine (*Vitis vinifera* L.) suitable for multiplex PCR. Mol Breed 15:349-366

Moreira FM, Madini A, Marino R, Zulini L, Stefanini M, Velasco R, Kozma P, Grando MS (2011) Genetic linkage maps of two interspecific grape crosses (*Vitis* spp.) used to localize quantitative trait loci for downy mildew resistance. Tree Genet & Genomes 7:153-167

Morris JR, Main GL (2010) An investigation of training system, pruning severity, spur length, and shoot positioning on Cynthiana/Norton grapes. Am J Enol Vitic 61:445-450

Morton LT (1985) Winegrowing in eastern America: an illustrated guide to viniculture east of the Rockies. Cornell University Press, Ithaca and London

Parker LD, Bordallo P, Colova VM (2005) Tracing the pedigree of Cynthiana grape by DNA microsatellite markers. Proc Fla State Hort Soc 118:200-204

Parker LD, Bordallo PN, Colova VM (2009) Phylogenetics analysis of North American native Cynthiana/Norton grape variety using DNA microsatellite markers. Acta Hort 827:225-228

Paterson AH, DeVerna JW, Lanini B, Tanksley SD (1990) Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato. Genetics 124:735-742

Regner F, Stadlbauer A, Eisenheld C (2000) Molecular markers for genotyping grapevine and for identifying clones of traditional varieties. Acta Hort 546:331-341

Reisch BI, Goodman RN, Martens M-H, Weeden NF (1993) The relationship between Norton and Cynthiana, red wine cultivars derived from *Vitis aestivalis*. Am J Enol Vitic 44:441-444

Riaz S, Krivanek A, Xu K, Walker MA (2006) Refined mapping of the Pierce's disease resistance locus, *PdR1*, and Sex on an extended genetic map of *Vitis rupestris*× *V. arizonica*. Theor Appl Genet 113:1317-1329

Riaz S, Tenscher AC, Graziani R, Krivanek AF, Ramming DW, Walker MA (2009) Using marker-assisted selection to breed Pierce's disease-resistant grapes. Am J Enol Vitic 60:199-207

Riaz S, Tenscher A, Ramming D, Walker MA (2011) Using a limited mapping strategy to identify major QTLs for resistance to grapevine powdery mildew (*Erysiphe necator*) and their use in marker-assisted breeding. Theor Appl Genet 122:1059-1073

Salmaso M, Malacarne G, Troggio M, Faes G, Stefanini M, Grando MS, Velasco R (2008) A grapevine (*Vitis vinifera* L.) genetic map integrating the position of 139 expressed genes. Theor Appl Genet 116:1129-1143

Sapkota S, Chen L-L, Schreiner K, Ge H, Hwang C-F (2015) A phenotypic study of *Botrytis* bunch rot resistance in *Vitis aestivalis*-derived 'Norton' grape. Trop Plant Pathol 40:279-282

Sawler J, Reisch B, Aradhya MK, Prins B, Zhong GY, Schwaninger H, Simon C, Buckler E, Myles S (2013) Genomics assisted ancestry deconvolution in grape. PloS One 8:e80791

Scott KD, Eggler P, Seaton G, Rossetto M, Ablett EM, Lee LS, Henry RJ (2000) Analysis of SSRs derived from grape ESTs. Theor Appl Genet 100:723-726

Sefc K, Steinkellner H, Glössl J, Kampfer S, Regner F (1998) Reconstruction of a grapevine pedigree by microsatellite analysis. Theor Appl Genet 97:227-231

Sefc KM, Regner F, Turetschek E, Glossl J, Steinkellner H (1999) Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. Genome 42:367-373

Sevini F, Marino R, Segala C, Grando MS (2003) Development and transferability of grape EST-SSR markers suitable for mapping in *Vitis* spp. In: XLVII Italian society of agricultural genetics-SIGA annual congress: proceedings, Verona, Italy, 24-27 September 2003:5.57. ISBN 88-900622-4-X

Stover E, Aradhya M, Yang J, Bautista J, Dangl GS (2009) Investigations into the origin of 'Norton' grape using SSR markers. Proc Fla State Hort Soc 122:19-24

This P, Jung A, Boccacci P, Borrego J, Botta R, Costantini L, Crespan M, Dangl G, Eisenheld C, Ferreira-Monteiro F (2004) Development of a standard set of microsatellite reference alleles for identification of grape cultivars. Theor Appl Genet 109:1448-1458

Thomas M, Scott N (1993) Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). Theor Appl Genet 86:985-990

Troggio M, Malacarne G, Coppola G, Segala C, Cartwright DA, Pindo M, Stefanini M, Mank R, Moroldo M, Morgante M (2007) A dense single-nucleotide polymorphism-based genetic linkage map of grapevine (*Vitis vinifera* L.) anchoring Pinot Noir bacterial artificial chromosome contigs. Genetics 176:2637-2650

van Heerden CJ, Burger P, Vermeulen A, Prins R (2014) Detection of downy and powdery mildew resistance QTL in a 'Regent'×'RedGlobe'population. Euphytica 200:281-295

van Ooijen JW (2006) JoinMap 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen, Netherlands

Voorrips R (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77-78

Welter LJ, Göktürk-Baydar N, Akkurt M, Maul E, Eibach R, Töpfer R, Zyprian EM (2007) Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). Mol Breed 20:359-374

Zhang J, Hausmann L, Eibach R, Welter LJ, Töpfer R, Zyprian EM (2009) A framework map from grapevine V3125 (*Vitis vinifera* 'Schiava grossa'×'Riesling')× rootstock cultivar 'Börner'(*Vitis riparia*× *Vitis cinerea*) to localize genetic determinants of phylloxera root resistance. Theor Appl Genet 119:1039-1051

CHAPTER 4

CONSTRUCTION OF HIGH DENSITY LINKAGE MAPS AND DETECTION OF DOWNY MILDEW RESISTANCE LOCUS IN A *VITIS AESTIVALIS*-DERIVD 'NORTON'

Introduction

Globally, fungal diseases are one of the most common problems in vineyards. These diseases are responsible for major damage of grapevines leading to reduced yield and wine quality. Among them, powdery mildew, downy mildew, *Botrytis* bunch rot and black rot are the most threatening. Downy mildew, caused by the oomycete *Plasmopara viticola*, is one of the most widespread and destructive, particularly in viticultural areas where warm and humid conditions persist during growing seasons. This disease is endemic to North American vineyards and was introduced to Europe during the 19th century by way of vine exchange between the continents (Moreira et al. 2011). Downy mildew attacks young seedlings at a very early stage as well as leaves, rachis and clusters.

The oomycete *P. viticola* is able to colonize plants and complete its life cycle in both resistant and susceptible *Vitis* species. However, in most of the resistant species (*V. riparia*, *V. cinera*, *V. labrusca*, *V. rupestries*, *V. berlandieri* and *V. lincecumii*), a defense response is exhibited, after haustoria establish around the mesophyll cells, that retards sporangial growth. Non *Vitis* Genus, *Muscadania rotundifolia* possess a stronger resistance reaction by efficiently obstructing hyphal growth and suppressing visible symptoms (Bellin et al. 2009). However, in many resistant Asian cultivars, hyphal growth is restricted only on outer side of the leaf lamina, preventing stomatal penetration and release of viable sporangiophores (Jurges et al. 2009). In susceptible species, *P. viticola* colonizes the mesophyll cells when the germinating zoospores penetrate the stomatal openings with protruding hyphal tips. These give rise to a network of intercellular mycelia with haustoria for further infection via sporangiophores. A study conducted by

Cadle-Davidson (2008) observed a range of variation in downy mildew severity across *Vitis* species. Accessions of *V. cinerea* and *Vitis* x *champinii* showed the lowest severity, whereas the interspecific hybrids *V. acerifolia* and *Vitis* x *novaeangliae* showed the highest disease severity.

Grapevines, being a woody perennial species with a long life cycle, require a longer time to introgress desired traits, and thus it takes more time to study complex quantitative inheritance. Genetic maps are key tools for studying quantitative traits, especially in grapevines, which have higher degrees of heterozygosity and inbreeding depression. During the last decade, a significant effort has been applied in constructing simple sequence repeat (SSR)-based linkage maps (Doucleff et al. 2004; Fischer et al. 2004; Hoffmann et al. 2008; Bellin et al. 2009; Blasi et al. 2011; Riaz et al. 2004, 2006 and 2008). These maps have been utilized to map major and minor resistance loci against various diseases including downy mildew of grapevines. To date, more than fifteen loci (*Rpv1-Rpv15*) have been identified on chromosomes 4, 5, 7, 8, 9, 12, 14, 17 and 18 that confer resistance to downy mildew of grapevines under different genetic backgrounds (Merdinoglu et al. 2003; Fischer et al. 2004; Welter et al. 2007; Bellin et al. 2009; Marguerit et al. 2009; Moreira et al. 2011; Schwander et al. 2012; Venuti et al. 2013; Ochssner et al., 2016; Zyprian et al. 2016). However, the majority of linkage maps consisted of a limited number of SSR markers (less than 500) with high inter-marker distance. These low-resolution maps are less useful for marker-trait association for marker-assisted selection (MAS).

Recently, with advances in next generation sequencing (NGS) techniques, there has been a shift in genomic research towards creating enormous amounts of molecular

markers. Researchers have been moving from amplified fragment length polymorphisms (AFLPs) and SSR markers to direct sequence variation in the genome, including single nucleotide polymorphism (SNPs) (Zhang et al. 2013; Zhou et al. 2014; Pootakham et al. 2015). Unlike AFLPs and SSRs, SNP markers are evenly distributed with wider genome coverage and are abundant in plant genomes (one in every 100-300 bp, Edwards et al. 2007). Despite being bi-allelic in nature, the ease of discovery and lower genotyping cost has increased their application in linkage and association mapping (Ball et al. 2010). Though high throughput sequencing techniques are more feasible and cost-effective in plants with smaller genomes, adoption of reduced representation libraries (RRLs) has benefited many plant species with larger and complex genomes including grapevines (Barba et al. 2014; Hyma et al. 2015;).

Genotyping-by-sequencing (GBS) is a direct sequencing approach through the adoption of RRLs. GBS is a simple, highly multiplexed genotyping technique that utilizes methyl sensitive restriction enzymes (REs) for reduction of genome complexity and target enrichment. Rapid adoption of the technique was motivated by an efficient barcoding system that allowed pooling of hundreds of samples in a single sequencing lane, target enrichment with long range PCR-amplification of specific genomic regions, use of molecular inversion probes, and various DNA hybridization/sequencing capture methods (Elshire et al. 2011). GBS library development is simplified in comparison with other systems (restriction-site associated DNA, RAD) since it requires less DNA, avoids random shearing and size selection, and is completed in a couple of steps through PCR amplification of the pooled library. Similarly, it allows simultaneous discovery of SNPs after alignment with the reference genome and genotyping through pooled barcoded

RRLs generating sufficient markers for linkage mapping and QTL studies (Davey et al. 2011; Barba et al. 2014). GBS has successfully generated high-density linkage maps in various model and non-model species (Pootakham et al. 2015; Ward et al. 2013; He et al. 2014; Huang et al. 2014; Wang et al. 2015; Bielenberg et al. 2015).

Here, we present the use of GBS techniques to generate the first high-density linkage map of *Vitis aestivalis*-derived 'Norton.' Using 2,072 markers (SSRs and SNPs), high-density linkage maps of both parents as well as a consensus were constructed. The map in combination with phenotypic data was able to identify a downy mildew resistance QTL. Construction of high-density linkage maps in species that lack a reference genome will facilitate genome assembly in addition to providing tools for genome-assisted breeding.

Materials and Methods

Plant Materials and DNA Extraction

A *Vitis aestivalis*-derived 'Norton' based population was used in this study. *V. aestivalis*-derived 'Norton' is reported to be a grape of American origin that shares alleles from the grape species *V. vinifera* and *V. aestivalis* (Stover et al. 2009). In 2005, crosses of *V. aestivalis*-derived 'Norton' x *V. vinifera* 'Cabernet Sauvignon' were made that resulted in 93 hybrid progeny, of which 19 were from a reciprocal cross. This population was expanded to 310 genotypes in 2011 as described by Adhikari et al. (2014). In this study, 159 progeny (73 were from 2005 and 86 were from 2011) maintained at the Missouri State Fruit Experiment Station (MSFES) in Mountain Grove, Missouri were used as a mapping population.

For SSR genotyping, genomic DNA was extracted from young leaves as previously described by Adhikari et al. (2014). Extracted DNA was normalized to 10ng/ul and stored in -80^{0} C until needed. For SNP genotyping, leaf samples were collected according to the sampling protocol recommended by VitisGen (www.vitisgen.org). Briefly, from each vine, a single newly expanded leaf, smaller than dime size, was collected in each well of a Costar 96-well cluster tube collection plate (Corning Life Sciences, Tewksbury, MA, USA). These tubes were sealed and shipped to the genotyping center according to the prescribed protocol. Genomic DNA was extracted using DNeasy 96-well DNA extraction kits (Qiagen, Valencia CA, USA) following grinding of the leaf samples using Geno/Grinder 2000 (OPS Diagnostics LLC, Lebanon NJ, USA) for a minute. Slight modifications to the manufacture's protocol and quality control check were made as described by Hyma et al. (2015) to improve DNA quality and quantity. Eluted DNA was quantified using QuantiFlor dsDNA System (Promega) and were stored until further processing.

Genotyping

SSR amplification, genotyping and linkage map construction

SSR amplification, genotyping and linkage map construction was conducted as described by Hammers et al. (2017). Previously, a total of 413 SSR markers were used to construct an integrated linkage map of 'Norton' and 'Cabernet Sauvignon'. This linkage map was utilized to develop a high-density linkage map by combining with SNPs obtained from GBS.

Genotyping-by-sequencing (GBS)

GBS was carried out at Cornell University Biotechnology Resource Center (BRC; Ithaca, NY, USA) as a part of mapping population under VitisGen, following the protocol described by Elshire et al. (2011). Briefly, genomic DNA from both parents and progeny were digested with methyl sensitive *ApeKI* restriction enzyme followed by ligation with unique barcode and a common adaptor. Each adaptor contained a three base overhang for ligation with *ApeKI* digested DNA. These 100 base pair (bp) reads were bridge amplified using PCR in a flow cell and were sequenced using Illumina HiSeq® 2000 platform. Parental DNA were sequenced three times (with separate barcode for each) whereas F_1 hybrids were run once. All of the 100 bp reads were converted into 64 bp reads after trimming off the barcode along with excess 3'-nucleotides. Reads containing large numbers of missing data were filtered and the remaining reads were aligned to the 12x *V. vinifera* 'PN40024' reference genome using the Burrows-Wheelers alignment tool (BWA) with default parameters (Li and Durbin 2009; Barba et al. 2014). Further, a variant call format (VCF) file was generated with a list of SNPs for each genotype, and was uploaded into TASSEL 5.0 for analysis. SNPs, with minor allele frequency (MAF) $\geq$ 0.1 and a minimum count of 10, were kept and all other missing sites and taxa as well as insertions, deletions and polymorphism (INDELs) were removed. A total of 40,724 SNPs were retained after additional stringent filtering. An association between downy mildew and SNPs genotypes was tested using a general linear model (GLM) function in TASSEL 5.0 (Bradbury et al. 2007).

Linkage map construction

To construct the genetic linkage map, 40,724 SNPs were filtered down to 3,700 SNPs. During the process, parental genotypes were merged, markers were filtered to 50% completeness of data with taxa minimum proportion of 0.1. This hap map file was run with Connor's code for identifying markers with different segregation patterns (aa x ab, ab x aa, ab x ab, ab x ac and ab x cd). This left a total of 3,825 markers for linkage map construction. A genotype file from SNPs was concatenate with the SSR file obtained previously to create a combined loci file for JoinMap4.1 yielding 4,236 markers. In addition, during the process, taxa number were reduced to 159 from 182 by excluding reciprocal crosses, and 'Norton' cross 'Cabernet Sauvignon' progeny 115, 117, 149 and 185 due to their high proportion of missing genotypes. In JoinMap 4.1, both .loc file and .map were uploaded using load data function of JoinMap4.1 under file tab. Segregation of markers were evaluated for goodness-of-fit for both observed and expected Mendelian ratios 1:1 ($P \leq 0.05$) using a χ^2 method to identify distorted markers. Highly distorted markers ($P \leq 0.001$) were excluded from the genotypes. Markers were grouped using independence LOD value of 25.0 and were further confirmed using recombination frequency parameters (start 0.25, end 0.05). Marker order within the group was calculated using a regression mapping algorithm until the third round and verified with the maximum likelihood method. Map units in centimorgan (cM) were generated using Kosambi's mapping function. Linkage groups were numbered 1 to 19 according to internationally acknowledged grapevine reference genetic maps (Doligez et al. 2006). For the construction of consensus map, markers mapped in both parental maps and heterozygous markers (abxab) were used. The final parental and consensus map charts

were constructed using software Map Chart (Voorrips 2002). The linkage map constructed from JoinMap4.1 was further analyzed with R/QTL software (Broman et al. 2003) using a four-way cross format. Marker order in SSR map and combined (SSR+SNP) maps were compared and contrasted.

Phenotyping of Downy mildew resistance

DM isolates maintenance

Leaves in oil leaf stage that showed DM infection and few sporangia were collected from 'Cabernet Sauvignon' in the Genomics vineyard at MSFES. These leaves were incubated overnight at room temperature with 100% relative humidity inside zip block bags to produce sporangia. To generate and maintain enough sporangia for phenotyping, 4th to 5th expanded leaves from susceptible cultivars of *V. vinifera,* kept in the green house or shade house were collected. These leaves were surface sterilized with 0.6% sodium hypochlorite for 2 minutes and then washed with deionized water three times. Sterilized leaves were kept adaxial side up in a petri dish containing 1% (w/v) agar. Sporangial suspension was drop-inoculated onto the leaf surface and incubated inside a humidified plastic bag overnight at room temperature. Petri-dishes were taken from plastic bags and sealed with parafilm after excess water was removed. The dishes were kept under 14 hrs of light at room temperature for 7 to 10 days. Mixtures of downy mildew from these collections were used to phenotype the mapping population in 2014 and 2015.

Sampling, leaf preparation and artificial inoculation

In each replicate, four leaf samples were collected from either side of each vine (trained as high wire bilateral codon) grown in genomics vineyards. Leaves were surface sterilized as described above and from each leaf; two leaf discs were created with 10mm cork. A total of 8 leaf discs from each vine were kept in 1% agar plate with adaxial side up. Leaf discs were inoculated with 50 µl *P. viticola* suspension at 150,000 sporangia per ml and were kept in phenotyping conditions as described above.

Evaluation of phenotype

Downy mildew progression was monitored daily until symptoms appeared to quantify genotype reaction to the pathogen. The level of infection was scored based on OIV descriptor 452-1 (Organisation Internationale de la Vigne et du Vin 1984) and as described by Kono and Sato (2015) as 1-5 with 1 = no downy mildew symptoms and sporangia, 2 = single or few sporangia and no visible symptoms, 3 = medium, somewhat strong sporulation, 4 = high sporulation with abundant mycelia and sporulation much bigger and 5 = very high sporulation, leaf disc completely covered with sporangia and mycelia as well as presence or absence of visible necrosis on the leaf disc. Disease progression and sporulation were evaluated 6, 8, 10 and 12 days after inoculation, and were used to detect QTL by combining with the genotype from the same mapping population. All leaf discs were evaluated visually as well as with microscope by two different people to make sure sporulation and disease progression was consistent across the genotypes. Leaf discs that were rotten or damaged were assigned as missing during computation.

Downy mildew evaluation under field condition

Using the same mapping population, a phenotypic analysis of *P. viticola* infection was carried out in the vineyard established at MSFES. The experimental vineyard was maintained with standard practices, i.e. vines are planted at 10 foot between-row and 8 foot within-row spacing and were trained to a high bilateral cordon system. Vineyard management practices including training and pruning are conducted in a timely manner except that application of fungicides were withheld. Disease severity was assessed three times during 2014 and 2015 growing seasons for an entire vine and individual leaf. Phenotypes that showed clear differences between the parents and the progeny were considered for further analysis. Averaging individual scores evaluated during different times generated final rating for QTL analysis.

QTL analysis

QTL analysis was performed using both interval mapping (IM) and multiple QTL mapping (MQM) mapping of MapQTL6.0. Locus file and map file were obtained from JoinMap4.1 after linkage map construction. These data were combined with trait file to detect best possible association. LOD threshold significant value was obtained using 1,000 permutation of the phenotyping data. Best cofactor combination was identified after several rounds of interval mapping to generate MQM peak. An additional run was carried out using composite interval mapping in R/QTL to find possible maker-trait association using a combined map.

Results

Phenotypic evaluation of downy mildew resistance

In years 2014 and 2015, the mapping population's response to DM was evaluated under both lab and field conditions. Phenotypic observations were conducted four times under lab assay and three times in the field. Under lab conditions, DM resistance was scored both by visual observation and examining under microscope 8 days after inoculation. Segregation of downy mildew was categorized from 1 to 5, with one being most resistant (no downy mildew) and five as most susceptible. During evaluation, a majority of the susceptible parents had a ranking of five and leaves were displaying severe DM. The susceptible leaf discs were fully covered with DM sporangia. The resistance parent did not have any DM under both field and lab conditions. The DM resistance phenotype segregated quantitatively with a continuous variation irrespective of the assays used (Fig. 1). Though there were differences in distribution patterns among two assays, the results were significantly correlated to each other (Supp. Table 1). The highest correlation was observed between the individual's vines under field conditions (0.89).

High-density linkage map

During linkage map construction polymorphic markers segregated for various types as; 159 (7.67%) markers were fully informative (ab x cd), 96 markers (4.63%) segregated with three alleles (ef x eg), 975 markers were double heterozygous and segregated with two alleles (hk x hk) for both parents, 335 (16.16%) markers segregated for 'Norton' (lm x ll), and remaining 507 (24.46%) markers segregated for 'Cabernet

Sauvignon' (nn x np) (Fig. 2). Distorted markers ($P< 0.05$) were seen on almost every linkage group. These markers were included in the final genetic linkage map unless they hindered linkage map construction.

Genetic linkage analysis clustered 2,072 (407 SSRs and 1,665 SNPs) markers on consensus linkage groups 1-19 that correspond to the haploid chromosome number in grapevines. The map encompassed 2,203.5cM, with linkage groups ranging from 80.9cM (LG10) to 180.6cM (LG7). The number of combined markers per linkage group ranged from 176 (LG14) to 58 (LG15) with an average of 108.68 (Table 1; Fig. 3). The integrated maternal and paternal map anchored 1,330 and 1,770 markers with genetic distance of 1,956.0cM and 1,983.2cM respectively (Supp. Fig. 1). The average number of markers in 'Norton' is 70 and in 'Cabernet Sauvignon' is 93.15 per linkage groups. In 'Norton', linkage group 14 harbored the highest number of markers (114, 45 SSRs and 69 SNPs) whereas linkage group 15 harbored the lowest number of markers (42, 7 SSRs and 35 SNPs). Similarly, in 'Cabernet Sauvignon', linkage group 18 harbored the highest number of markers (137, 19 SSRs and 118 SNPs) and linkage group 15 harbored the lowest number of markers (45, 5 SSRs and 40 SNPs). Average inter-marker distance in 'Norton', 'Cabernet Sauvignon' and consensus linkage maps were 1.5, 1.1 and 1.1 cM respectively (Table 1; Fig. 3; Supp. Fig. 1).

Comparison with previously published SSR linkage map

In comparison to a previously developed SSR linkage map (Hammers et al. 2017), this combined high-density linkage map showed a 5-fold higher number of markers that covered a greater genetic distance (Table 3). Average inter-marker distance in the

combined map is 0.9 with only 4 gaps greater than 10cM (LGs 9, 11, 15 and 18). Whereas in SSR maps, there is an average of 4.1cM inter-marker distance and 26 gaps that are larger than 10cM. The present high-density map possesses higher genome coverage than previously constructed SSR maps, as well as other maps available in grape and wine research community.

QTL mapping for downy mildew resistance

Significant QTLs were detected for downy mildew resistance in different environmental conditions, using different parameters. Firstly, phenotypic data was associated with genotypes using SSR linkage map. Analysis showed, a significant QTL on linkage group 18 flanked by markers VVCS1H077H16R1-1 AND UDV737 (Table 2, Figure 4 A). A permutation test with 1000 permutations was calculated and identified a significant threshold value of 3.2 for various phenotypic scores. A MQM analysis confirmed a QTL on linkage group 18 explaining 33.8% of total phenotypic variation observed (R^2) for downy mildew with a maximum LOD score of 16.42. However, even though flanking markers were same, slight variations were observed on LOD and R^2 scores with the phenotypes of different years and conditions. The shortest possible genetic distance between two flanking markers was 3.5 cM.

QTL detected with SSR linkage map was further confirmed with combined high-density linkage map using R/QTL. Before combining both type of markers, an independent marker-trait association was carried-out by combining SNPs and DM phenotype using GLM on TASSEL. A significant association was observed on linkage group 18 (Supp. Fig. 2). A genome wide scan for QTL using R showed a significant QTL

on linkage group 18 flanked by previously detected SSR markers with an additional SNP between them (Fig. 4 B). A LOD threshold of 4.93 was detected using 1000 permutation tests, and QTL explained 42.5% of total phenotypic variation with a LOD value of 24.9.

Discussion

Despite Norton being a popular wine grape in Southern and Midwestern states, little is known about its genetics of disease resistance. This study uses a quantitative genetic approach to obtain a segregating mapping population using *V. aestivalis*-derived 'Norton' as a resistance source. To date, several genetic maps have been published in the grape and wine research society. However, most of these maps are based on limited SSR markers with greater inter-marker distance. Here, we report the first high-density genetic map by combining both SSRs and SNPs, as well as QTL associated with downy mildew resistance.

Using GBS, 40,724 SNPs were generated for Norton cross Cabernet Sauvignon mapping population. However, GBS generated data are accompanied by a large number of erroneous SNP calling especially for heterozygous species like grapes due to complexity of genome, errors in alignment, and lack of representation and sequencing techniques (Barba et al. 2014; Spindel et al. 2013). Therefore, markers were set to filter down from original 40,724 to 3,700. A further subset of 2,072 were clustered into 19 different linkage groups including 407 SSR markers after additional stringent filtering due to the capacity limitation of JoinMap 4.1, a maximum of approximately 3000 markers. Among 2,072 markers, 975 (47.05%) were heterozygous, and 1,097 (52.9%) did

not follow normal Mendelian ratios (*P* ≤0.05). Distorted markers were included in the final maps unless the genetic maps construction hindered.

Results showed that incorporation of SNPs into the previously constructed SSR maps changes the maker order and synteny. The physical position of the SNPs did not coincide with genetic position, with additional changes in the order of SSR markers in the final linkage map. Wang et al. (2012) and Ganal et al. (2012) pointed out several possible reasons behind this, including errors in the *V. vinifera* 'PN40024' reference genome sequence, genetic diversity between *Vitis* cultivars or species-speciation leading to different micro-structures on chromosomes, or errors in alignment step. Further, limited synteny in cross comparison can be also due to presence of larger number of markers that are heterozygous and distorted, which could result in errors in mapping possible due to inversions in marker order.

The result of merging two different maker types generated saturated map with greater genome coverage and shorter inter-marker distance. The average distance between adjacent markers across in all 19 linkage groups was significantly lower (0.9cM) than each of the individual component maps. Even though, in consensus map there were 4 gaps that were larger than 10 cM, this is much lower than the previously constructed SSR map using the same population (26 gaps, Hammer et al. 2017) as well as other maps available in grape and wine research community (Fischer et al. 2004; Blasi et al. 2011; Barba et al. 2014), making this the most saturated genetic map in *Vitis* to date.

Downy mildew oomycetes are obligate biotrophs that require living hosts for growth and expansion. This makes them difficult study under controlled condition without the presence of living tissue (Heath and Skalamera 1997). The leaf disc assay

seems to be a viable method for studying downy mildew resistance, especially when field conditions are not optimum for disease development. Even though this assay is labor intensive, its use helps to incorporate a large number of genotypes within a short time interval, particularly helpful for bigger populations. Brown et al. (1999), Diez-Navajas et al. (2008) and Bellin et al. (2009) pointed out several advantages of leaf disc assay over others; as it helps to detect localized necrosis across infection sites in the resistant cultivars. This phenotype is rarely seen under field conditions, and could be confused with necrotic spots caused by other biotic and abiotic stresses. In addition, it reduces environmental variance by narrowing down the confidence interval of QTLs, and disease progression can be monitored at precise time intervals to detect sequence of host-pathogen interactions. Dissection of phenotypes like this helps to understand the basic underlying mechanism under resistance or susceptibility for quantitative resistance breeding. In our experiment, downy mildew segregated quantitatively from resistance to susceptible phenotypes. In resistant genotypes, there was a range of resistance response, from no growth of DM to limited growth. In contrast, susceptible genotypes were colonized by sporangia. The phenomenon was consistent on both years, exhibiting a high degree of co-relation.

To date, two different loci that conferred resistance to downy mildew have been identified in LG18. A resistance locus *Rpv2* was identified on a cross between the susceptible *V. vinifera* parent ‘Cabernet Sauvignon’ and the resistant parent ‘8624’ (derived from *M. rotundifolia*). This locus explains 76% of the total phenotypic variation observed and was also detected in S1 population of *M. rotundifolia* ‘Regale’ (Blanc et al. 2012; Weidemann-Merdinoglu et al. 2006). Similarly, a major resistance locus, *Rpv3* was

identified on chromosome 18 of the grapevine cultivars 'Regent' and 'Bianca' (Fischer et al. 2004; Welter et al. 2007; Bellin et al. 2009; Di Gaspero et al. 2012; van Heerden et al. 2014 and Zyprian et al. 2016). In 'Regent', this locus was flanked by SSR marker UDV 112 and was later confirmed in 'Regent' x 'Red Globe' population adjacent to SSR marker VMC7F2/GF18-08 (van Heerden et al. 2014). Marker VMC7F2 is *Rpv3* linked in 'Bianca', corresponds to position 26,896,989 of chromosome 18 in the 12X reference genome sequence. And the other *Rpv3* flanking marker, UDV 305, corresponds to position 24,868,359 on chromosome 18 in the same reference genome sequence (Jaillon et al. 2007). The resistance locus conferred by *V. aestivalis*-derived 'Norton' is located on the lower arm of LG 18, flanked by SSR markers VVCS1H077H16R1-1 and UDV737, correspond to positions 24,636,885 and 26,050,244, respectively. Therefore, the identified QTL shared most of the genomic region of *Rpv3*. Di Gaspero et al. (2012) studied selective sweep in *Rpv3* and had generated seven different haplotypes, which have been retained in downy mildew resistance breeding in grapevines. However, none of the haplotypes were present in 'Norton/Cynthiana'. Downy mildew resistance in Norton could be due to presence of unique locus in the same region or *Rpv3* paralogues that have not been identified yet. The availability of Norton genome sequence as well as phenotypic effect of the genes on those regions may help to generate relevant information for resistance breeding.

Flanking SSR markers VVCS1H077H16R1-1 and UDV737 are 3.5 cM apart. Based on the latest version of the grapevine genome sequence (http://www.genoscope.cns.fr/cgibin/ggb/vitis/12X/gbrowse/vitis/), this region is 1.41 Mb and contains around 388 predicted genes. Di Gaspero et al. (2007) reported that *Rpv3*,

which resides on the lower arm of chr18, is rich in TIR-NBS-LRR genes inherited from North American grape species. A further study is needed to dissect the QTL, flanked by SSR markers VVCS1H077H16R1-1 and UDV737, to identify genes that contribute resistance in 'Norton'.

Further, this high-density map was able to reduce the confidence interval to a shorter genomic distance (3.5cM), compared to previously published genetic map on 'Regent' 17-35cM (Welter et al. 2007) or 16.6cM (van Heerden et al. 2014). Using general linear model (GLM) in TASSEL, marker trait association detected several significantly associated SNPs on LG 18 for downy mildew. However, fewer markers were retained after additional filtering and there were 10 SNPs that were present in combined linkage map generated by R/QTL. SNP S_{18} – 33,162,624 was the nearest downy mildew-associated marker within two flanking SSR markers. This SNP was able to narrow down the flanking region to less than 1cM

To our knowledge, this is the first high-density linkage map combining both SSR and SNP markers. We were able to present the first high-density combined map for *V. aestivalis*-derived 'Norton' and localized downy mildew resistance locus on chromosome 18. Linked markers in this study can be further tested for their predictability of resistance to downy mildew in different genetic backgrounds. Such validated markers are useful for marker-assisted selection to accelerate breeding as well as characterize germplasm. In addition, these markers can be used for positional cloning of the region to study possible candidate genes for resistance, and understand cellular pathways involved.

Figure Legends

Fig. 1 Segregation of downy mildew disease in the mapping population obtained by crossing Norton with Cabernet Sauvignon, a) segregation from 1-5 (as described in Kono et al. 2016) b) phenotyping year 2014 lab c) phenotyping year 2014 field d) phenotyping year 2015 lab and, e) phenotyping year 2015 field. The assay was obtained using leaf disc 8 DAI under control environment condition. Disease severity was assessed using both visual and quantitative measures.

Fig. 2 Marker segregation types (SSRs and SNPs) based on JoinMap4.1 format used for developing combined linkage map by combining both parental maps.

Fig. 3 SNPs and SSRs distribution across 19 linkage groups of grapevines derived by crossing *V. aestivalis*-derived 'Norton' with *V. vinifera* 'Cabernet Sauvignon'. Distances are in cM. Linkage groups were obtained by R/QTL software using a 4-way cross format for composite interval mapping (CIM).

Fig. 4 A) Localization of genetic determinants of downy mildew resistance in *V. aestivalis*-derived 'Norton' using SSR markers. The LOD threshold of 3.2 ($P < 0.05$) was determined after 1,000 permutations. A significant QTL was obtained on linkage group 18 flanked by markers VVCS1H077H16R1-1 and UDV737 using MapQTL6.0 and, B) Whole genome scan of QTL for downy mildew resistance in *V. aestivalis*-derived 'Norton' based population. The LOD threshold of 4.93 ($P < 0.05$) was determined after 1,000 permutations. A significant QTL was obtained on linkage group 18 associated with the SNP s18_33146891.

Supplementary Fig. 1 Combined (SSRs and SNPs) *V. aestivalis*-derived 'Norton' and *V. vinfiera* 'Cabernet Sauvignon' linkage maps.

Supplementary Fig 2 Marker-Trait association using generalized linear model (GLM) in TASSEL.

Table 1 Genetic map from F1 population derived from the cross *V. aestivalis*-derived 'Norton' with *V. vinifera* 'Cabernet Sauvignon' (CS)

	Number of Markers									Genetic Size (cM)			Avg. Inter-marker Distance (cM)		
	Norton			CS			Integrated			Norton	CS	Integrated	Norton	CS	Integrated
LGs	SSR	SNP	Total	SSR	SNP	Total	SSR	SNP	Total						
LG01	27	53	80	21	79	100	31	99	130	117.1	102.7	109.5	1.5	1.0	0.8
LG02	13	34	47	14	61	75	15	64	79	100.5	97.3	105.1	2.1	1.3	1.3
LG03	13	37	50	9	58	67	15	71	86	88.7	93.8	127.6	1.8	1.4	1.5
LG04	27	25	52	27	93	120	29	67	96	78.3	125	104.3	1.5	1.0	1.1
LG05	24	69	93	22	84	106	26	104	130	94.9	102.9	108.8	1.0	1.0	0.8
LG06	12	31	43	12	87	99	18	95	113	101.8	112.7	104.1	2.4	1.1	0.9
LG07	24	77	101	22	104	126	27	125	152	139.6	115.9	180.6	1.4	0.9	1.2
LG08	23	44	67	4	39	43	24	82	106	147.9	67.6	132.1	2.2	1.6	1.2
LG09	11	33	44	11	39	50	16	50	66	61.3	108.3	123.3	1.4	2.2	1.9
LG10	16	28	44	12	40	52	16	47	63	141.8	139.2	80.9	3.2	2.7	1.3
LG11	22	31	53	14	67	81	23	73	96	91.7	89.7	102.9	1.7	1.1	1.1
LG12	15	95	110	13	139	152	15	96	111	91.2	78.9	107.8	0.8	0.5	1.0
LG13	13	56	69	9	77	86	12	86	98	74.8	99.9	93.4	1.1	1.2	1.0
LG14	45	69	114	24	118	142	46	130	176	113	125.4	135.6	1.0	0.9	0.8
LG15	7	35	42	5	40	45	7	51	58	81.6	75	103.0	1.9	1.7	1.8
LG16	19	68	87	15	87	102	19	103	122	101.4	104.7	114.1	1.2	1.0	0.9
LG17	12	39	51	10	64	74	14	76	90	96.8	101.5	123.2	1.9	1.4	1.4
LG18	35	60	95	19	118	137	40	123	163	114.5	128.5	123.5	1.2	0.9	0.8
LG19	14	74	88	10	103	113	14	123	137	119.1	114.2	123.7	1.4	1.0	0.9
Total	372	958	1330	273	1497	1770	407	1665	2072	1956.0	1983.2	2203.5	1.5	1.1	1.1

Table 2 Summary of the QTLs in F_1 population derived from the cross *V. aestivalis*-derived 'Norton' with *V. vinifera* 'Cabernet Sauvignon'

Trait	Year	Linkage Group (LG)	Significance threshold (α = 0.05	SSR Flanking Markers	Physical Location	Peak position	LOD score	R^2 (%)	Confidence interval (LODmax-2)(cM)
Downy Mildew	2014 Lab	18	4.5	VVCS1H077H16R1-1 and UDV737	24,636,888-26,050,244	118.05	16.42	33.8	116.05-120.67
	2014 Field	18	4.6	VVCS1H077H16R1-1 and UDV737	24,636,888-26,050,244	118.05	9.99	22.2	113.66-121.67
	2015 Lab	18	4.6	VVCS1H077H16R1-1 and UDV737	24,636,888-26,050,244	118.05	9.84	21.9	105.86-121.17
	2015 Field	18	4.5	VVCS1H077H16R1-1 and UDV737	24,636,888-26,050,244	118.05	9.53	21.3	113.23-121.17

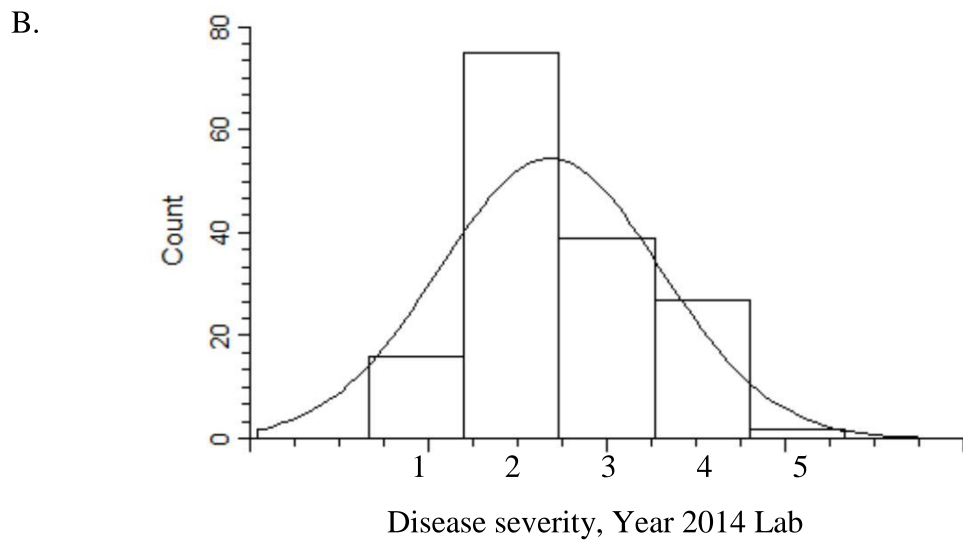
Table 3 A comparison between previously constructed SSR based genetic map and SSR plus SNP integrated map

Features	SSR map	SSR plus SNP map
Size of mapping population	182	159
Number of markers	411	2065
Number of linkage groups	19	19
Map size (cM)	1678.6	2203.5
Average Inter-marker distance (cM)	4.1	0.9
Number of gaps > 10 cM	26	4

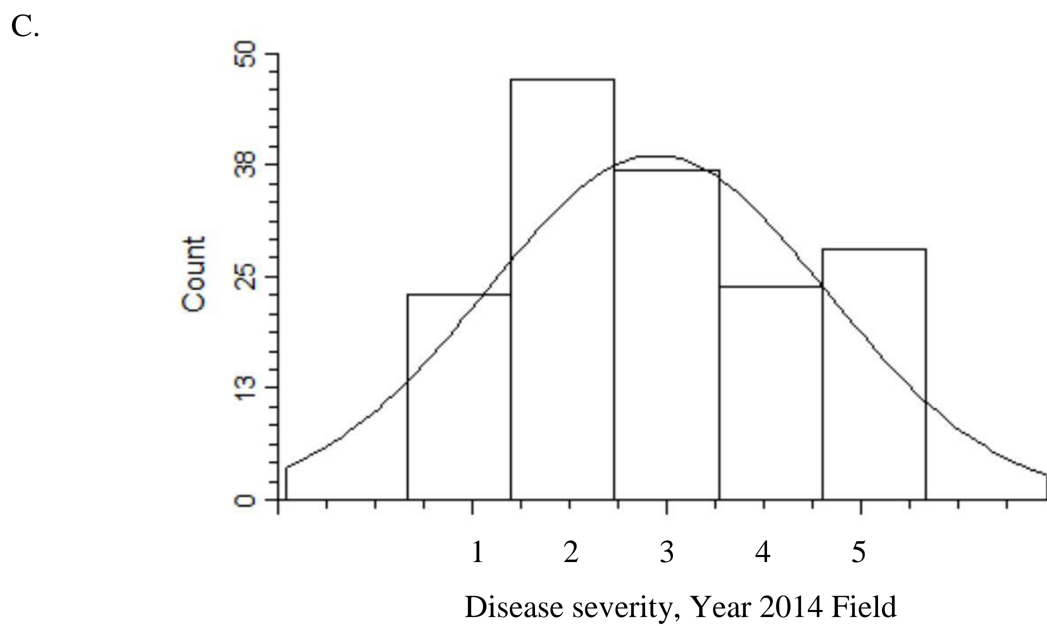
Figure 1

A.

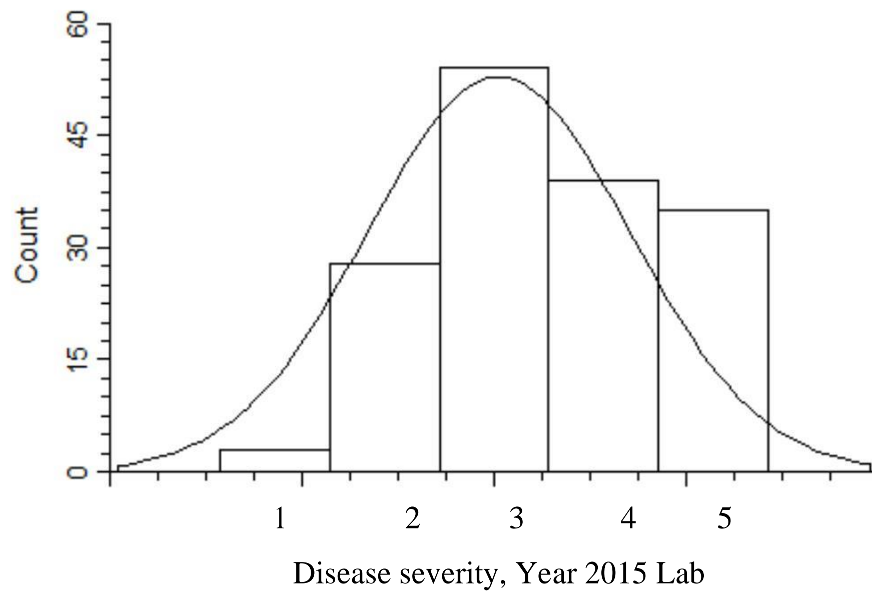
B.



C.



D.



E.

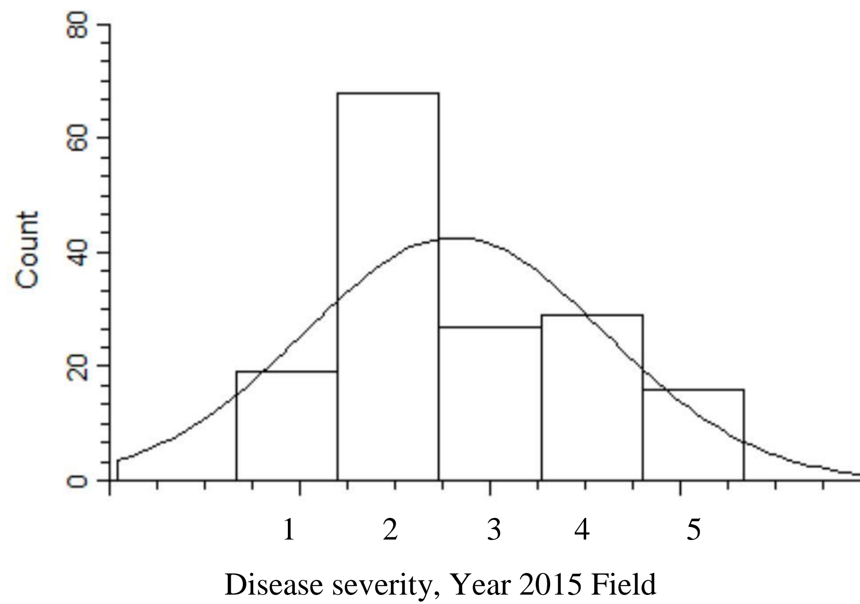


Figure 2

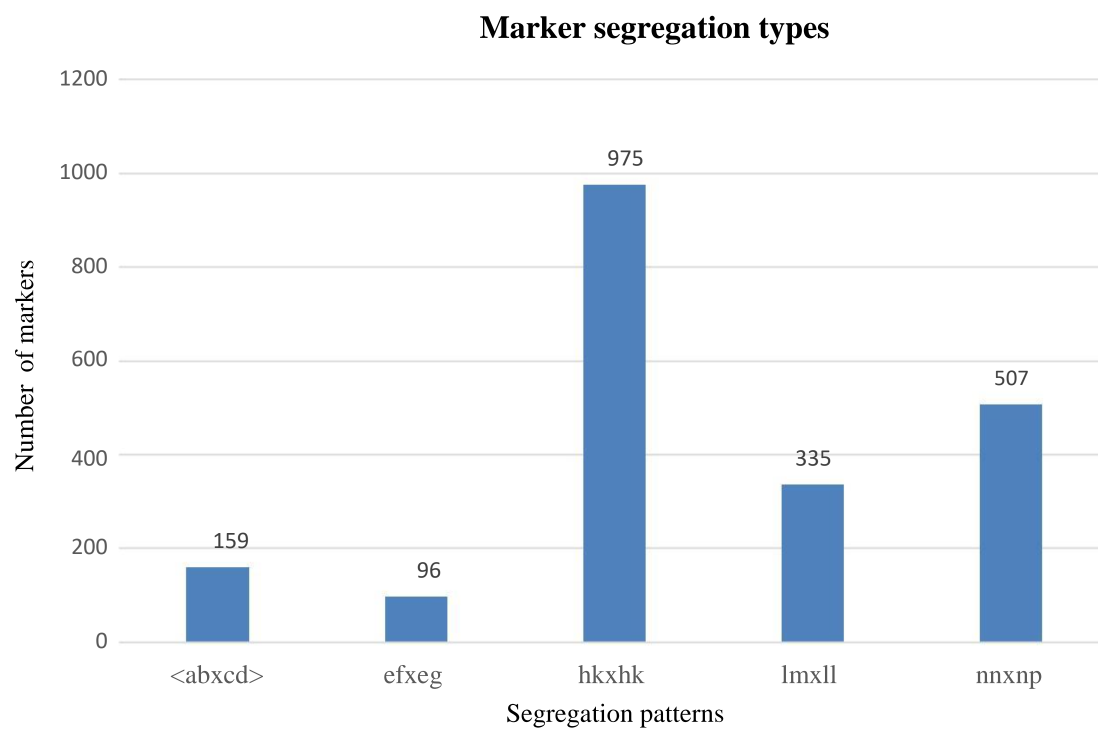


Figure 3

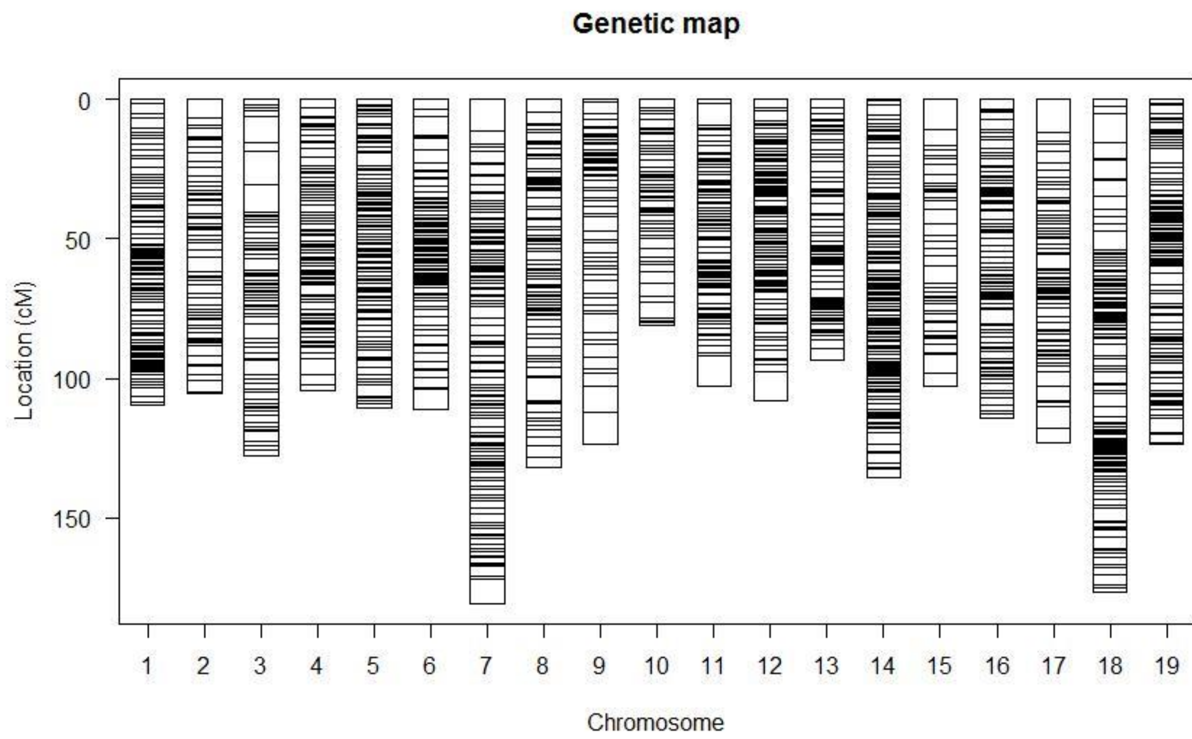
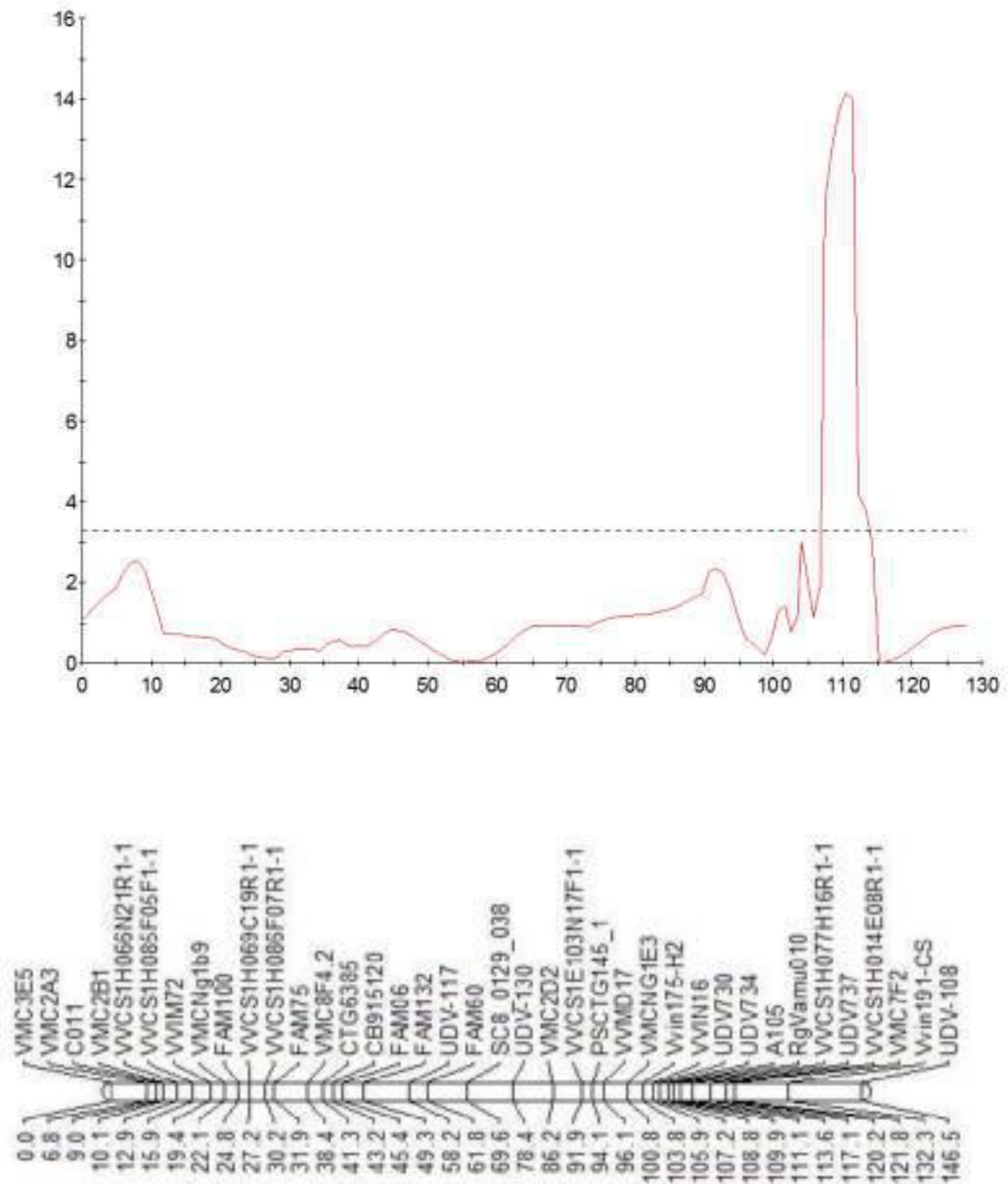
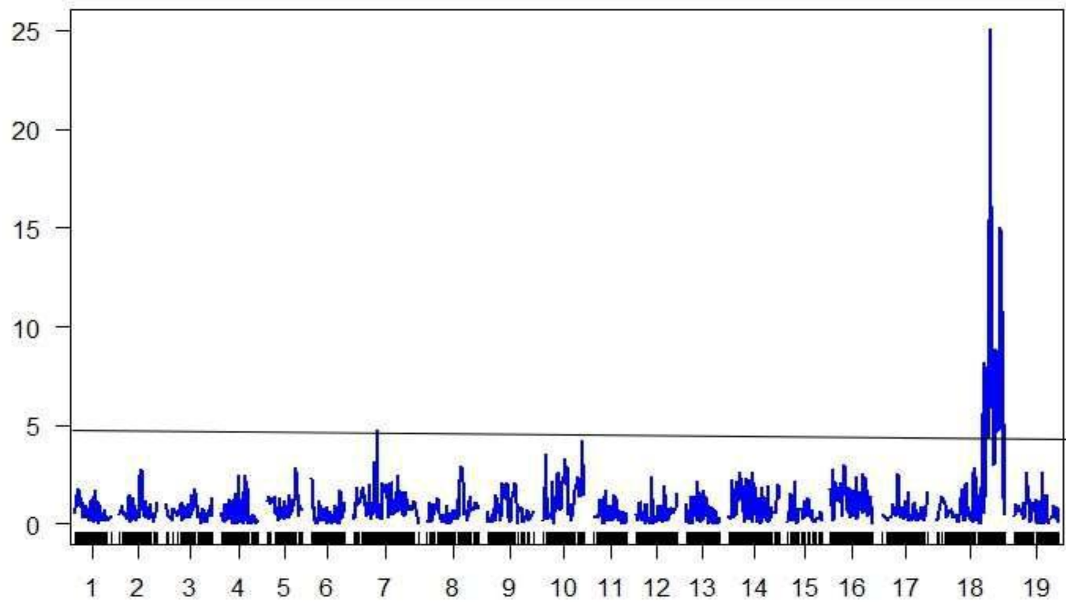


Figure 4

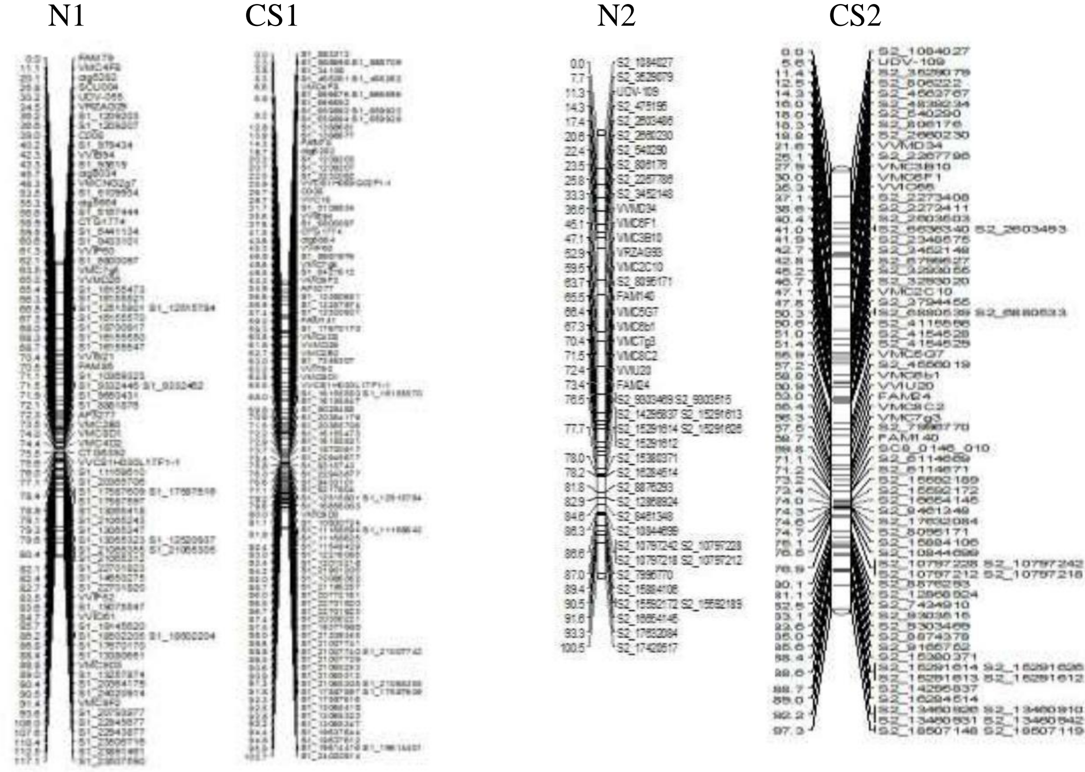
A.

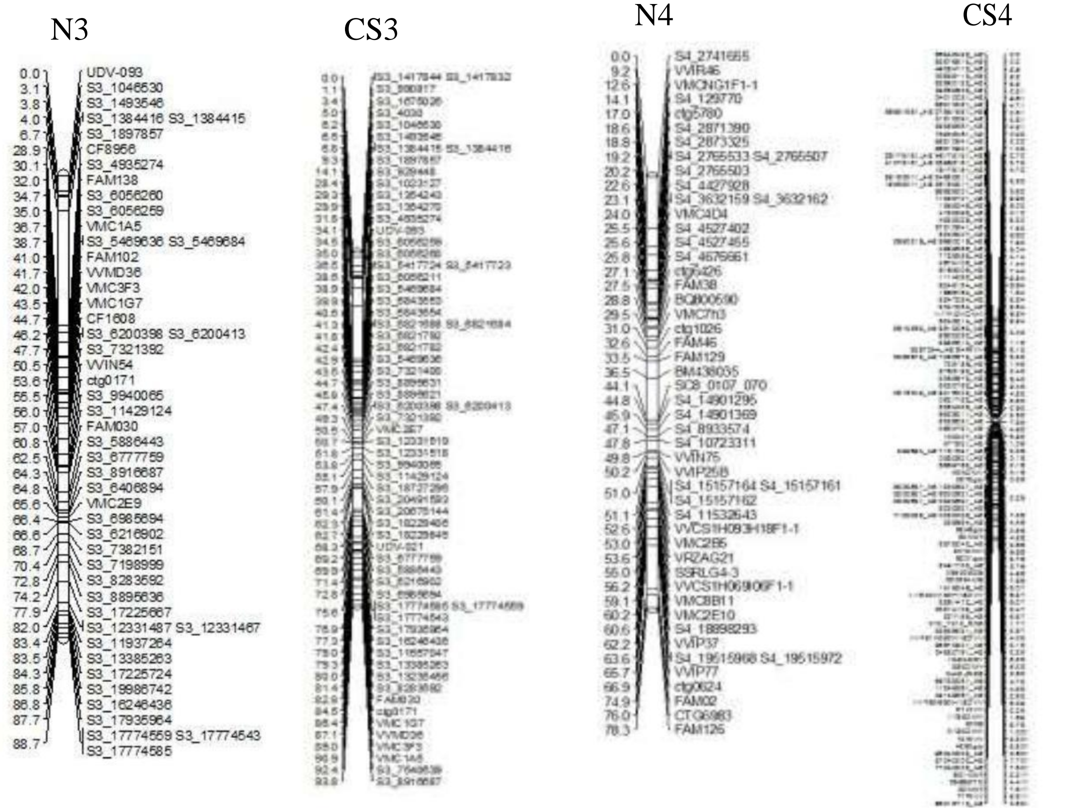


B.



Supplementary Figure 1





N5

CS5

N6

cM	Marker
0.0	S6_18491003
5.9	S6_17977656
9.5	S6_14925609
11.8	S6_15514212
13.2	S6_15476004
17.8	S6_14627054 S6_14627011
18.6	S6_14358972
22.8	S6_12868269
32.8	S6_6848839
34.1	S6_6834811
34.2	S6_6834834
35.7	S6_6813190
36.7	S6_6497212
44.0	FAM40
46.2	CTG6316
48.3	S6_7010559
49.5	S6_1655934
52.3	S6_13817176
54.7	FAM110
56.8	VVCh14-37
58.6	S6_593508
60.5	S6_4664411
61.7	S6_4664327
63.3	UDV-085
65.9	S6_6491894 S6_6491892
69.2	FAM78
71.0	VMC2F10
72.6	VMCNg1h11
74.8	S6_6752074
75.8	S6_6491879
76.3	S6_7775546
77.6	S6_6752147 S6_6752134
77.6	S6_6752098 S6_6752130
83.9	VVS5
87.4	VMC4G6
91.0	VVIP28
93.1	S6_15545533
96.5	VMC3A8
101.8	S6_18992168

CS6

N7 CS7 N8 CS8

0.0 VMC3f8
1.4 VMC3c9
2.5 S8_17359996
4.3 VVIB66
6.0 S8_17360001 S8_17360002
6.9 S8_17860765 S8_17860782
11.1 S8_22044674
12.1 S8_21964700 S8_21954678
14.0 S8_20805816
15.7 VMC2H10
22.3 S8_20308519 S8_20308500
27.0 S8_19502017
28.7 S8_21708841
41.9 S8_17140345
43.5 S8_16801799
45.4 S8_17357121
46.9 S8_19984443
49.2 S8_19589319 S8_19589311
49.7 S8_19483426
50.1 S8_19483414
50.5 S8_19961468
51.0 S8_19589344
51.5 S8_18585599
51.7 S8_19446588
52.0 S8_18588809
52.2 S8_18216602
52.6 S8_18216601
52.9 S8_18326828
53.9 S8_18667393
55.6 S8_21588409
55.9 S8_21059740
57.0 S8_21318808
57.4 S8_21775756
57.9 S8_21727868
59.6 S8_20168956
61.5 S8_20877411
62.7 S8_19525054
67.6 S8_20238739

N9

0.0 S9_10531895
1.9 S9_10400490
3.2 S9_10495050
11.2 VVIU37
13.8 S9_6230019 SC8_0141_028
16.5 FAM26
18.4 S9_5740727 S9_5740730
19.5 FAM35
22.8 VMC3G8
23.4 S9_5147162
23.7 S9_4635894
23.9 S9_6855824
24.4 S9_4871890
25.0 VVIO52
25.4 S9_2648924
26.6 S9_1097404
26.8 S9_910176
26.9 S9_910143
27.2 S9_1097398
28.6 CD009354
30.1 S9_2531690
30.6 EST2B07
31.4 S9_1153762
33.6 VMC6d12
35.1 S9_1911211
36.5 S9_9634405
36.9 S9_5651118
37.5 S9_5651106
37.9 S9_5651112
39.1 S9_2895444
39.8 S9_5651147
41.3 FAM42
41.9 S9_9251328
46.9 S9_2678523 S9_2678532
52.0 VVIN57
52.3 S9_915820
55.3 S9_777818
55.7 S9_463190
57.6 S9_655440 S9_655479
61.3 S9_915809

CS9

0.0 S9_1954807
1.6 S9_2794986
2.5 S9_2648924
3.0 S9_1153762
4.1 S9_1097404 S9_1097398
4.6 S9_1163621
5.3 S9_910143 S9_910176
7.5 S9_1591540 S9_1591584
8.9 S9_2531690
10.5 S9_1911211
11.8 S9_3422506
12.8 S9_2895444
15.7 S9_4714947
17.0 S9_5651118 S9_5651106
S9_5651112
18.0 S9_5651147
19.2 S9_5734771
20.3 S9_5640663
22.1 S9_6190286
22.6 S9_6817594
24.2 S9_6855824
25.3 S9_7616256
29.3 S9_2678532 S9_2678523
31.2 CD009354
31.4 S9_9634405
31.9 S9_9634402
32.8 S9_10019124
33.4 S9_9251328
37.7 FAM26
40.3 FAM35
41.3 S9_463190
44.9 S9_6230019
49.9 VMC3G8
52.9 S9_5147162
59.6 VVIN57
65.7 FAM42
73.5 S9_10531895
80.1 VMC2A9
85.5 S9_22437349
86.9 S9_22356390
89.1 FAM53
92.9 S9_22437516
99.6 VMC3H5
105.1 VMC2E11
108.3 VVCS1H016F02F1-1

N10

0.0 SC8_0297_008
4.6 S10_247457
4.7 S10_429263
10.0 VVIH01
21.4 S10_5220730
23.7 S10_6502831
25.2 S10_4760630 S10_4760610
26.1 UDV-073
27.6 S10_5075775
27.9 S10_5075810
29.2 VRZAG64
29.4 VRIP64
30.4 VRZAG67
31.1 VMCZAG67
31.7 VmcVRg001
33.2 VRIP25
35.2 FAM148
36.9 VmcSsrVrZAG025
43.2 VMC2E8
50.1 ctg7871
52.1 S10_10985361 S10_10985375
53.1 S10_11781406
55.8 ctg5592
60.4 S10_10985370
65.2 VVIV37
66.0 S10_15002491
67.8 S10_13885879
76.5 SC8_0060_099
82.8 S10_12312759 S10_12312744
S10_12312709
87.5 S10_11658068
105.5 S10_7368549
106.3 S10_6815794
117.9 S10_4125263 S10_4125257
124.2 S10_4047118
125.5 S10_3578397
134.2 S10_1377161
136.1 S10_620801
138.6 S10_241961
141.8 S10_356904

CS10

0.0 VMC2E8
13.7 VMCZAG67
16.5 VmcSsrVrZAG025
17.4 VRIP25
20.7 VRZAG67
23.2 VRIP64
24.6 VRZAG64
27.2 UDV-073
30.9 S10_3947810
32.1 S10_4047118
38.4 SC8_0297_008
42.1 VVIH01
46.1 S10_4125263 S10_4125257
49.6 S10_722448
50.7 S10_356904
50.8 S10_880949
51.6 S10_528940
52.7 S10_429263
53.5 S10_247457
55.8 S10_1876982
57.0 S10_1876979
59.4 S10_2188642
60.9 S10_2326510
68.8 S10_5220730
69.6 S10_5075775
70.2 S10_5075810
72.5 S10_5646638
73.6 S10_4760630 S10_4760610
74.9 S10_5599639
76.1 S10_5584825
77.5 S10_5694195
79.5 S10_6502831
83.8 S10_8090993
84.1 S10_8100724
85.1 S10_8400818 S10_8400782
86.2 S10_8101891
90.3 S10_8230177
96.2 S10_11781406
98.1 S10_10985361 S10_10985375
103.0 S10_10985370
106.1 S10_15002491
108.5 S10_13885879
113.3 S10_20840475
124.3 S10_12312759 S10_12312744
S10_12312709
134.2 ctg7871
139.2 ctg5592

N11

CS11

N12

CS12

N13

cM	Marker
0.0	ctg6135
3.9	S13_6330207
5.8	S13_4006500
12.1	S13_6875679
13.3	S13_6616341 S13_6616335
15.1	NS01
16.2	S13_6788015
17.1	S13_6788016
17.4	S13_6875596
17.7	S13_6786667 S13_6786626
19.0	S13_6788018
22.0	VVIC51
23.0	S13_7795575
24.2	VMC3D12
25.4	S13_4247365
27.5	S13_3966084
27.8	S13_4932886
29.2	S13_10588708 S13_10588724
31.6	S13_4932889
33.2	S13_15175476
33.9	PSCTG231_2
34.4	S13_6087648
35.9	S13_18783444
36.4	S13_5862721 S13_5862718
39.0	SC8_0071_075
39.5	SC8_0071_014
41.2	S13_9453287
41.3	S13_15153556
42.3	EST10h11
43.6	S13_16895005
45.3	VMCNG4e10.1
46.0	VMC9H4
46.6	S13_6982612
48.1	S13_18443346
48.8	S13_18443341
49.5	S13_9390173
49.6	CTG9373
50.6	CTG7358
50.8	S13_14828357
52.8	S13_21407962
53.0	S13_11400257
54.0	S13_15188451 S13_15188452
54.8	S13_12860217
56.0	S13_25833119
56.7	S13_24745499
57.3	S13_21408001 S13_21408035
	S13_15746220
57.6	S13_25012003
59.4	S13_15746219 S13_20229112
59.8	S13_20229106
60.4	S13_20229078
66.7	S13_22315429
67.0	S13_20558610
67.3	S13_21992867 S13_21992854
	S13_21992895
67.4	S13_22472744
67.6	S13_22229167
71.9	S13_24059910
72.3	S13_23514471 S13_23514461
74.8	SC8_0053_001

CS13

cM	Marker
0.0	S13_2476572
2.0	S13_3367257
7.2	S13_4898438
9.6	S13_5255339 S13_5255336
11.5	S13_6330207
12.8	S13_4006500
19.5	S13_6788016
20.9	S13_6786667 S13_6786626
21.9	S13_6875596 S13_6788015
24.4	S13_6616341 S13_6616335
24.7	S13_6875679
25.3	S13_6788018
25.4	S13_6875496
25.6	S13_6788017
26.8	S13_7405478
28.6	S13_7795575
32.4	S13_4900317
32.6	S13_8706959
33.4	S13_10588724 S13_10588708
35.1	S13_9922886 S13_9922884
35.7	S13_9453288
37.3	S13_9665605
38.2	S13_14783676
38.8	ctg6135
41.4	S13_9453287
42.0	S13_4932886
43.7	S13_15153556 S13_4932889
45.5	S13_15175476
45.9	S13_18225176
48.2	S13_21407964
49.1	S13_18431186
49.7	S13_18443341 S13_18443346
	S13_18431201 S13_18431162
51.1	S13_18431144
51.9	S13_18822148
52.5	S13_16895005
53.2	S13_21357147
54.2	S13_20346548
56.0	S13_21408001 S13_21408035
57.0	VVIC51
57.5	S13_21407962
57.6	VMC3D12
58.2	S13_25833119
58.5	S13_24954440
59.0	S13_24745499
60.4	S13_25012003
61.6	S13_20229078 S13_20229106
	S13_20229112
62.2	S13_9390173
62.9	S13_6987467
64.7	S13_18783444
66.3	S13_15746220
67.2	S13_15746219
68.7	SC8_0071_075
71.1	S13_15189465
71.7	S13_15189082
72.3	S13_15189452
73.2	S13_14828357
73.6	S13_15189451
75.8	S13_18485927
78.6	SC8_0071_014
81.7	VMC9H4
83.6	VMCNG4e10.1
88.3	SC8_0053_001
91.8	S13_20558610
92.6	S13_22472744 S13_21992854
	S13_21992867 S13_21992895
92.8	S13_22229167
93.7	S13_22315429
95.0	CTG9373
99.3	S13_24059910
99.9	S13_23514471 S13_23514461

N14

cM	Marker
0.0	S14_1962783
2.1	S14_5189013
4.5	S14_524948
4.9	S14_524921 S14_524954
7.3	S14_1093935
8.7	S14_2138144
9.1	S14_1775818
9.4	S14_1895142
10.0	S14_2769607
10.6	S14_2695916 S14_2695934
12.3	S14_3271410
13.0	S14_768842
13.5	S14_1824209
14.4	S14_1668571
15.4	S14_524281
16.7	S14_3063941
19.1	S14_1746167
19.4	S14_1050765
19.5	S14_1746153
19.6	S14_1746098
21.1	S14_2810292
22.5	VMCNG1E1
23.8	VVC62
25.7	Scu015
29.1	VMC8h10
30.1	S14_3537862
33.4	UDV050
35.1	VMC9C1
38.3	S14_6896386
39.1	VVC32
40.4	S14_6071813
42.5	VMC1E12
43.9	S14_7277648
44.1	S14_6071832
44.9	S14_9443321
47.9	VMC2H12
49.4	SC8_0063_073
51.8	VVC34
56.8	VMC2G11
57.6	SC8_0026_027
58.1	VVIP22
59.7	VMC2C3
60.6	VMC2H5
61.1	VVCS1H038P11F1-1
62.3	S14_7925041
63.5	S14_22173109
64.7	VRIP112
66.7	VVIV69
67.1	VMC6c10
68.4	VMC6E3
69.6	S14_8660842 S14_8660841
70.4	S14_7744899
70.9	VMC2a5
71.8	VVCh14-36
72.3	CTG6530
72.8	UDV033
73.6	S14_15508303
73.8	S14_15688484
75.6	VVCh14-27
76.0	S14_10846623
76.7	VVCh14-12 EE6432
77.4	A010
78.1	VVCh14-3
78.3	S14_12454821
78.9	FAM44
79.8	VVCh14-2
80.3	S14_12454830
81.4	VVCh14-9
82.1	VVCh14-16
82.3	S14_13756328
82.7	VVIP26
83.0	VVCh14-18
83.5	STS-AA6
85.1	S14_22639682
85.4	S14_22893114
85.6	S14_17968361
86.3	VMCNG1G1.1
86.5	VVCh14-14
87.6	S14_27336815
87.8	S14_27779977
88.2	S14_17023290 S14_26365183
	S14_26365150
88.9	S14_25419941
89.9	S14_25356079
90.3	S14_22577662
91.2	UDV095
91.8	S14_18245647
92.4	S14_22774456
92.6	S14_24623239
93.3	S14_23931094 S14_22746109
94.3	VVIS70
94.9	S14_25421530
95.5	VVII51
96.3	FAM90
97.0	S14_26527597
97.8	VVIN94
99.8	S14_20529109
101.7	ctg6151
102.4	S14_27369286 S14_27369289
104.3	S14_27285739
104.8	S14_27405330
105.7	S14_27405381
107.1	S14_27925231
108.1	S14_27369275
110.8	S14_22996881
111.8	S14_24479995
113.0	S14_23909056

CS14

N15

Position	Marker
0.0	S15_16486381
3.4	S15_16295792
6.5	S15_15690659
7.4	S15_15827646
7.8	S15_15572693 S15_15572694
7.9	S15_15609845
9.2	S15_16040752
17.4	S15_13472399
18.2	S15_13472392
19.1	S15_13450011
19.8	S15_13472386
21.2	S15_13144458
22.5	S15_11852906
24.6	S15_12643700
25.5	S15_16351352
28.3	VVIV67
29.9	S15_10989263
43.0	S15_10293255
44.9	VVIQ61
46.4	S15_12022184
46.9	S15_12022134
47.8	S15_10966290
48.0	S15_10966285
48.4	S15_10966228 S15_10966229
48.7	S15_10966331
49.3	S15_10305982 S15_10305990
	S15_10305995 S15_10305969
50.2	S15_10966206
50.7	S15_12022191
54.6	FAM105
59.0	S15_15251404
60.7	S15_15161477
61.5	ctg4274
63.6	VMC5G8
72.7	VVIM42-2
81.6	S15_19173600
84.3	S15_19348020
86.0	SC8_0040_088

CS15

Position	Marker
0.0	S15_13801050
0.7	S15_15178892
2.2	S15_15520566
2.6	S15_15251404
4.1	S15_15178912
4.6	S15_15161477
5.0	S15_14994574
8.1	S15_14634276
10.9	S15_12655908
12.2	S15_12566396
14.5	S15_12022191
15.2	S15_12022184
15.3	S15_12022134
15.6	S15_10966331
15.8	S15_11563394 S15_11563398
	S15_11563409
16.9	S15_10966285 S15_10966290
17.3	S15_10966228 S15_10966206
	S15_10966229
18.9	S15_10305969 S15_10305990
	S15_10305995 S15_10305982
20.9	S15_10442819
23.2	S15_10293255
37.9	S15_10989263
39.2	VVIV67
45.3	S15_16351352
46.0	S15_12643700
47.7	VVIQ61
51.1	FAM105
53.9	S15_12454820 S15_12454825
57.9	S15_10989283
60.5	ctg4274
62.2	VMC5G8
66.2	S15_15572693 S15_15572694
66.4	S15_15609845
69.7	S15_13805537
71.8	S15_16295792
75.0	S15_16486381

N16

Position	Marker
0.0	ctg2141
1.4	S16_970677
3.7	S16_786198
5.2	CTG7620
7.5	FAM36
9.0	S16_1299197
13.7	UDV-013
17.1	S16_8079740 S16_8079722
18.3	S16_6795751
21.0	S16_1805480
21.9	S16_6790465
25.4	VVC05
28.6	S16_1746326 S16_1746300
29.6	S16_1746334
32.4	S16_4699826
32.8	S16_3097998
33.9	S16_3097847
35.1	S16_3101339 S16_3097843
35.5	SC8_0051_101
36.0	UDV-009
38.3	S16_13270028
38.9	S16_13661631
39.7	S16_9287718
40.1	S16_13005963
40.8	S16_12455696
41.1	S16_11014580
42.1	S16_16793298
42.3	ctg2012
42.8	S16_16104091
43.7	S16_12257986 S16_12258009
43.9	S16_12257997 S16_12258010
	S16_12257988
44.4	S16_12258001
44.9	S16_15464284
45.7	GB007D01
46.9	S16_16104060 S16_16104094
48.4	S16_13197404
48.6	VMC1E11
49.3	VMC3g11
49.9	S16_16104100
51.4	S16_16567151 S16_16567133
	S16_16567159
52.2	S16_17343656 S16_17343683
54.7	UDV-052
56.6	S16_18940745
57.9	S16_17577722 S16_17577720
58.8	S16_15724607 S16_15724640
59.4	S16_18789949 S16_18789945
61.4	S16_16925729
61.8	S16_17312550
63.3	S16_16925713
65.1	ctg9366
65.4	S16_17325382
65.7	S16_17349324
67.8	ctg0244
68.9	S16_22125322 S16_22125320
	S16_22125319
71.1	S16_20589018
72.5	S16_20603496 S16_20603505
73.2	S16_20963009
77.1	S16_22579391
77.8	S16_22719837
79.3	S16_22360876
80.9	SC8_0091_083
82.7	S16_22369472
84.8	S16_19664230
88.0	S16_22076504
90.2	VVMD5
92.4	VMC5a1
93.9	ctg7933
94.9	S16_23083254
96.7	S16_22778350
99.1	SCU014
101.4	S16_22671653

CS16

Position	Marker
0.0	S16_786166
3.1	S16_970677
5.2	S16_1311782
5.8	S16_1340418
6.8	S16_786198
8.2	S16_1299197
11.8	S16_2895138
12.5	S16_2895089 S16_2895137
16.1	S16_4641509
23.3	S16_8079722
23.9	S16_8079740
24.8	S16_6795751
25.6	S16_8000284
26.5	S16_1805480
26.7	S16_6790465
28.3	S16_1805428
28.4	S16_10319333
29.5	S16_17223274
31.2	S16_13382903
32.7	S16_9192348
33.7	S16_12759311
35.3	UDV-013
37.6	S16_14019783 S16_14019787
37.8	FAM36
38.7	S16_14618354
39.6	S16_14066449
41.8	S16_15782010
42.9	S16_4699826
43.8	S16_3097998
43.9	S16_15781967
44.0	S16_3097847 S16_3097843
	S16_3101339
45.2	S16_15724607
45.4	S16_22778350
45.6	S16_15724640
47.3	S16_22941361
47.6	VVC05
48.0	S16_20124329
48.6	S16_19341600
49.4	UDV-009
50.6	S16_19296559
51.4	S16_20754229 ctg2012
52.5	S16_14020368
52.9	S16_20754264
53.5	GB007D01
53.7	S16_20820946
54.2	S16_20759866
54.9	VMC1E11 S16_20124334
55.6	ctg0244
55.8	S16_17624312
56.4	VMC3g11
57.3	S16_13661631
57.5	S16_16104091
57.6	S16_23083254
58.3	S16_22671653
58.6	S16_17343656
59.0	S16_16567151 S16_16567159
59.3	S16_16567133
59.6	S16_17343683
59.7	S16_16104100
60.1	S16_22076504
61.1	S16_16925729
61.2	S16_17325382
61.4	S16_16925713
61.6	S16_15464284
62.0	S16_17312550
62.7	S16_17349324
63.0	S16_19664230
63.8	S16_16567156
63.9	S16_16567158
64.3	S16_16104060 S16_16104094
64.4	S16_22011497
66.0	UDV-052
66.9	S16_17577720 S16_17577722
68.5	S16_20603505
69.4	S16_20603496
71.7	S16_18789945
72.2	S16_18789949
72.6	S16_18940745
73.9	ctg9366
75.5	S16_22360876
76.9	S16_19223438
78.8	S16_22369472
80.2	S16_20589018
81.6	SC8_0091_083
84.9	S16_20963009
86.9	S16_22125319 S16_22125322
	S16_22125320
93.7	S16_22719837
94.0	S16_22579391
97.0	ctg7933
100.3	VMC5a1
104.7	VVMD5

N17

0.0 S17_2222040
1.6 S17_2222037
4.5 S17_2724874
8.8 S17_3691435 S17_3691417
S17_3691421
14.4 S17_4983586 S17_4983624
S17_4983610
15.5 S17_5362279
19.7 S17_4839283
23.4 S17_6438209
25.1 S17_7125497
26.0 S17_6516667
28.3 S17_7700361
29.7 S17_7915342
31.6 S17_8691103
32.0 S17_7949016
32.5 S17_7646241
32.9 S17_7646228
33.4 S17_7646271
34.6 S17_8587123
38.1 VMC3C11.1
40.8 S17_4413107
41.8 S17_4278880
43.1 S17_3546199 S17_3546213
46.7 ctg6344
49.1 FAM97
50.5 S17_14263328
51.3 VVIP16
51.7 VMC2H3
52.6 S17_14858272
53.7 S17_1271705
54.9 S17_1308386
56.0 S17_1271694
60.6 S17_9053830
63.1 S17_16543788 S17_16543790
64.1 VVCS1H068D03F1-1
66.7 S17_11216146
71.6 ctg5672
72.9 CTG8270
74.6 ctg9346
77.0 S17_9036779
79.0 VMC9G4
79.7 ctg6954
82.8 S17_9053908
84.3 S17_7713800
87.0 AF3283
96.8 S17_9864260

CS17

N18

0.0 S18_348563
6.8 S18_1338055
7.8 S18_1344763
12.9 S18_1887519
13.7 S18_1887373
S18_1887407 S18_1887413
15.6 S18_1887441
17.7 VMC3E5
21.5 VMC2B1
24.2 S18_324355
24.9 S18_2127190
25.6 VVCS1H066N21R1-1
26.7 VMC2A3
27.6 S18_2435617
28.4 VVCS1H085F05F1-1
31.1 VVIM72
32.0 FAM100
33.5 VMCNg1c9
35.3 VVCS1H069C19R1-1
36.8 VVCS1H085F07R1-1
37.9 S18_5172757
38.4 S18_5172695
39.4 S18_7144620
41.0 FAM75
41.7 S18_7890792
42.3 S18_7712093
44.5 VMC8F4.2
46.6 VVCS1H065K01F1-1
47.7 CTG6385
48.6 CB915120
49.4 FAM06
50.2 S18_10296599 S18_10296559
50.6 S18_10826679 S18_10826681
51.5 S18_10908083
52.3 S18_11698183
53.2 S18_10909076
54.8 S18_11698180
55.7 FAM132
57.8 S18_10767219
64.9 UDV-117
71.7 SC8_0129_036
75.6 S18_7099621
76.8 UDV-130
77.8 S18_7148382
84.9 S18_22629194
86.9 S18_10540653
87.5 PSCTG145_1
87.7 S18_10262657
89.9 S18_8965216
90.4 VVMD17
92.4 S18_9546763
92.8 UDV737
93.6 Win175H2
94.3 VMCNG1E3
95.1 S18_10262625
96.0 S18_31952418
97.3 UDV730
98.3 UDV734
98.9 VVIN16
99.3 B004
99.4 S18_11175559
100.1 At05
101.0 S18_12801626
101.2 RgVamu010
101.8 S18_23489438
101.9 S18_11741279
103.3 VVCS1H077H16R1-1
105.3 S18_23212346
105.8 S18_12801601
106.6 S18_30104163
106.9 S18_30104169
107.5 S18_30104133
109.0 S18_26224285
109.6 S18_26224288 S18_26224276
S18_18221948 S18_18221937
109.9 S18_18221942
112.4 VVCS1H014E08R1-1
113.5 VMC2D2
115.8 S18_23675023 S18_23675024
122.0 S18_22644157
123.8 Win191-C5
127.3 S18_24683073
129.2 S18_32027583
130.3 S18_26695034
131.4 S18_28413517
133.6 S18_31115345
137.6 S18_27670052
146.9 S18_29644804 S18_29644811

CS18

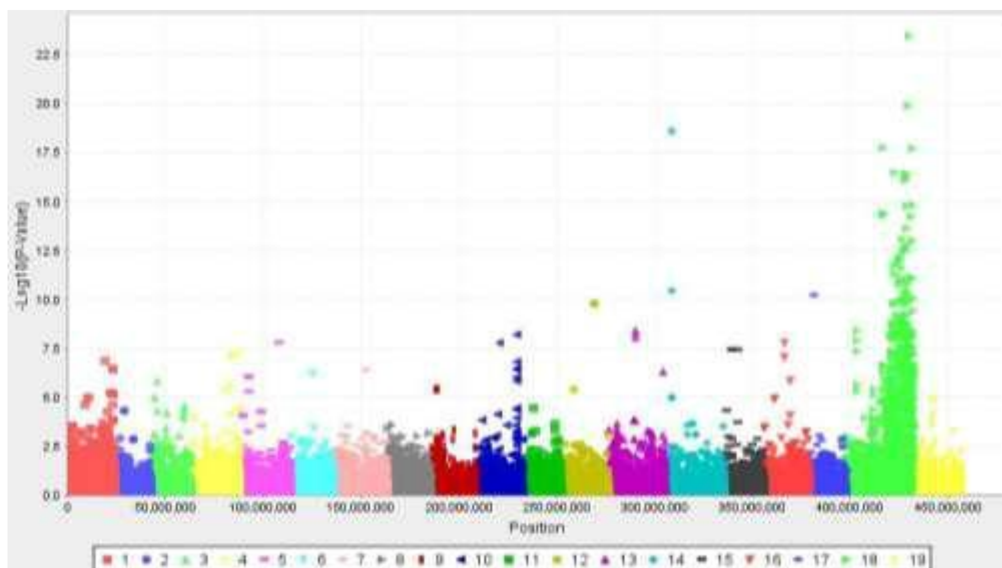
N19

cM	Markers
0.0	S19_20806519 S19_20623436 S19_20623434
1.1	S19_15795023
1.5	S19_15845890
2.1	S19_15845854
3.8	S19_12247362
4.9	S19_11545372 S19_11545366
6.2	S19_17865482
6.8	S19_17865463 S19_12939583
7.8	S19_9979889
9.1	S19_9255211
9.5	S19_9588221
10.3	S19_8876944
11.6	S19_8668226
17.9	S19_8246828
21.7	S19_7389690
25.4	S19_7135250
27.2	S19_4109460
31.0	S19_4499005 S19_4499006
34.5	S19_25611 S19_25607 S19_25615
37.7	S19_3144337
41.6	S19_2220202
41.8	S19_988547
43.1	S19_525605
43.4	S19_473588
44.1	S19_891486
47.5	S19_125216
48.4	S19_3327491
49.7	S19_1981810
50.2	S19_1981817
50.6	S19_1908113
50.8	VVIP17B
51.0	S19_1908099
53.0	S19_675975
53.7	S19_2534436 S19_2534486
56.7	S19_1307916
57.6	S19_3502952
58.4	S19_2154072 S19_2154075
58.8	S19_2254983 S19_2254967 S19_2254974 S19_2254999
59.7	S19_2788266
60.0	S19_2538619
60.4	FAM15
61.4	UDV023
62.2	VVCS1H088I07F1-1
63.2	S19_4232720
63.8	S19_4471321 S19_4471269 S19_4471272
64.2	S19_4232755 S19_4484823
67.3	VVIP11
68.1	VMC5e9
68.9	S19_4471263
69.5	S19_5454634
71.0	VMC9A2.1
73.6	B003
74.6	S19_6509681
77.3	VMC3b7.2
78.9	VVIN04
79.9	UDV029
84.2	S19_7756264
86.2	S19_7836327
94.7	S19_9208658 S19_9208609
100.1	VVIV33
102.2	PSCTG196_2
103.8	S19_14660027
104.9	S19_14660037 S19_14660001
105.8	S19_16915901
107.5	S19_16042021
108.6	S19_19245503 S19_19245508
109.7	S19_19257848
110.4	S19_19245580
112.3	S19_16042612
119.1	PSCTG81_1

CS19

cM	Markers
0.0	S19_19428126 S19_19428125
0.7	S19_16042612
2.6	S19_16182263
3.1	S19_19245503 S19_19245508
4.0	S19_19257848
5.8	S19_16754839
6.1	S19_16754829
6.8	S19_18512469
7.4	S19_14660037 S19_14660001
7.7	S19_14740619
7.9	S19_14660027
8.8	S19_11545353 S19_16915901
9.1	S19_10666451
9.5	S19_12204437
9.9	S19_16042021
10.7	S19_18732061
11.8	S19_19245580
15.3	S19_19257842
18.7	S19_9208658 S19_9208609
28.2	S19_7836327
29.5	S19_7756264
33.2	S19_6813931
34.4	S19_6813808
34.7	S19_6813824
35.3	S19_6813872
38.1	S19_5736842
39.8	S19_5260131
40.3	S19_6509681
41.3	S19_3347088
42.0	S19_5853013
42.9	S19_5736816
43.4	S19_5737700
44.1	S19_5427711 S19_5427683
44.5	S19_5454634
45.5	S19_5258336
47.4	S19_4471269
47.5	S19_4471321 S19_4471272
47.7	S19_4484823
48.0	S19_4232755 S19_4232720
48.8	S19_3756539
49.1	S19_3502952
49.7	S19_3521182
50.2	S19_4471263
50.6	S19_2858483
51.4	S19_3049790
51.9	S19_2788266
52.5	S19_2538619
52.6	S19_1896369
53.0	S19_2134423 S19_2134441
53.1	S19_2134437
53.2	S19_2156353 S19_2156382
53.8	S19_1922682
54.2	S19_1720644
54.8	S19_2154072
55.2	S19_2254974
55.5	S19_2154075 S19_2254983
55.7	S19_2254967 S19_2254999
56.1	S19_5260127
56.5	S19_5260115
57.1	S19_427644
57.7	S19_802776
58.0	S19_473594
58.6	S19_1307916
59.2	S19_480867
60.6	S19_2534486 S19_2534436
61.3	S19_3183639
62.0	S19_675975
62.3	S19_1729592
64.6	S19_2220202
66.0	S19_1908099
66.5	S19_1908113
66.7	S19_1981817 S19_1981810
67.4	S19_125216
67.7	S19_3327491
70.7	VMC9A2.1
74.6	FAM15
75.5	S19_2220196
77.5	VVIP17B
78.0	UDV023
79.7	S19_7135250
81.8	VMC5e9
82.7	VVIP11
84.5	S19_3144337 S19_3144362
85.7	S19_4087440
87.0	S19_8246828
88.6	S19_4499006 S19_4499005
91.7	UDV029
94.4	S19_7389690
100.5	PSCTG81_1
102.3	VVIV33
103.1	PSCTG196_2
105.3	S19_9979889
105.9	S19_12939583
108.5	S19_15845854
109.0	S19_15845890
110.1	S19_15795023
114.2	S19_21678578

Supplementary Figure 2



Supplementary Table 1

Pearson correlation coefficient among phenotypic years and assays for downy mildew

	Year 2014 Lab	**Year 2014 Field**	**Year2015 Lab**	**Year2015 Field**
Year 2014 Lab	1.000000	0.692772	0.596511	0.715882
Year 2014 Field	0.692772	1.000000	0.527501	0.892929
Year 2015 Lab	0.596511	0.527501	1.000000	0.539328
Year 2015 Field	0.715882	0.892929	0.539328	1.000000

References

Adam-Blondon A-F, Roux C, Claux D, Butterlin G, Merdinoglu D, This P (2004) Mapping 245 SSR markers on the *Vitis vinifera* genome: a tool for grape genetics. Theor Appl Genet 109:1017-1027

Adhikari P, Chen L-L, Chen X, Sapkota SD, Hwang C-F (2014) Interspecific hybrid identification of *Vitis aestivalis*-derived 'Norton'-based populations using microsatellite markers. Sci Hort 179:363-366

Ball AD, Stapley J, Dawson DA, Birkhead TR, Burke T, Slate J (2010) A comparison of SNPs and microsatellites as linkage mapping markers: lessons from the zebra finch (*Taeniopygia guttata*). BMC Genomics 11:218

Barba P, Cadle-Davidson L, Harriman J, Glaubitz JC, Brooks S, Hyma K, Reisch B (2014) Grapevine powdery mildew resistance and susceptibility loci identified on a high-resolution SNP map. Theor Appl Genet 127:73-84

Bellin D, Peressotti E, Merdinoglu D, Wiedemann-Merdinoglu S, Adam-Blondon A-F, Cipriani G, Morgante M, Testolin R, Di Gaspero G (2009) Resistance to *Plasmopara viticola* in grapevine 'Bianca'is controlled by a major dominant gene causing localized necrosis at the infection site. Theor Appl Genet 120:163-176

Bielenberg DG, Rauh B, Fan S, Gasic K, Abbott AG, Reighard GL, Okie WR, Wells CE (2015) Genotyping by sequencing for SNP-based linkage map construction and QTL analysis of chilling requirement and bloom date in peach [*Prunus persica* (L.) Batsch] PLoS One 10:e0139406

Blanc S, Wiedemann-Merdinoglu S, Dumas V, Mestre P, Merdinoglu D (2012) A reference genetic map of *Muscadinia rotundifolia* and identification of *Ren5*, a new major locus for resistance to grapevine powdery mildew. Theor Appl Genet 125:1663-1675

Blasi P, Blanc S, Wiedemann-Merdinoglu S, Prado E, Rühl EH, Mestre P, Merdinoglu D (2011) Construction of a reference linkage map of *Vitis amurensis* and genetic mapping of *Rpv8*, a locus conferring resistance to grapevine downy mildew. Theor Appl Genet 123:43-53

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633-2635

Broman KW, Wu H, Sen Ś, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889-890

Brown, M. V., Moore, J. N., Fenn, P., & McNew, R. W. (1999) Comparison of leaf disk, greenhouse, and field screening procedures for evaluation of grape seedlings for downy mildew resistance. HortScience 34:331-333

Cadle-Davidson L (2008) Variation within and between *Vitis* spp. for foliar resistance to the downy mildew pathogen *Plasmopara viticola*. Plant Dis 92:1577-1584

Dalbó M, Ye G, Weeden N, Steinkellner H, Sefc K, Reisch B (2000) A gene controlling sex in grapevines placed on a molecular marker-based genetic map. Genome 43:333-340

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Rev Genet 12:499

Di Gaspero G, Cipriani G, Adam-Blondon A-F, Testolin R (2007) Linkage maps of grapevine displaying the chromosomal locations of 420 microsatellite markers and 82 markers for R-gene candidates. Theor Appl Genet 114:1249-1263

Di Gaspero, G., Copetti, D., Coleman, C., Castellarin, S. D., Eibach, R., Kozma, P., & Kovács, L. (2012) Selective sweep at the *Rpv3* locus during grapevine breeding for downy mildew resistance. Theor Appl Genet 124:277-286

Diez-Navajas, A. M., Wiedemann-Merdinoglu, S., Greif, C., & Merdinoglu, D. (2008) Nonhost versus host resistance to the grapevine downy mildew, *Plasmopara viticola*, studied at the tissue level. Phytopathology 98:776-780

Doligez A, Adam-Blondon A-F, Cipriani G, Di Gaspero G, Laucou V, Merdinoglu D, Meredith C, Riaz S, Roux C, This P (2006) An integrated SSR map of grapevine based on five mapping populations. Theor Appl Genet 113:369-382

Doligez A, Bouquet A, Danglot Y, Lahogue F, Riaz S, Meredith C, Edwards K, This P (2002) Genetic mapping of grapevine (*Vitis vinifera* L.) applied to the detection of QTLs for seedlessness and berry weight. Theor Appl Genet 105:780-795

Doucleff M, Jin Y, Gao F, Riaz S, Krivanek A, Walker M (2004) A genetic linkage map of grape, utilizing *Vitis rupestris* and *Vitis arizonica*. Theor Appl Genet 109:1178-1187

Edwards, D., Forster, J.W., Cogan, N.O.I., Batley, J. and Chagné, D. (2007) Chapter 4: single nucleotide polymorphism discovery in plants. In Association Mapping in Plants (Oraguzie, N.C., Rikkerink, E., Gardiner, S.E. and De Silva, N.H., eds) New York, USA: Springer 53–76

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS One 6:e19379

Fischer BM, Salakhutdinov I, Akkurt M, Eibach R, Edwards K, Töpfer R, Zyprian E (2004) Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. Theor Appl Genet 108:501-515

Galeano, C. H., Fernández, A. C., Gómez, M., & Blair, M. W. (2009) Single strand conformation polymorphism based SNP and Indel markers for genetic mapping and synteny analysis of common bean (*Phaseolus vulgaris* L.). BMC Genomics 10:629

Ganal, M. W., Polley, A., Graner, E. M., Plieske, J., Wieseke, R., Luerssen, H., & Durstewitz, G. (2012) Large SNP arrays for genotyping in crop plants. J Biol Sci 37:821-828

Grando MS, Bellin D, Edwards K, Pozzi C, Stefanini M, Velasco R (2003) Molecular linkage maps of *Vitis vinifera* L. and *Vitis riparia* Mchx. Theor Appl Genet 106:1213-1224

Hammers M, Sapkota S, Chen L-L, Hwang C-F (2017) Constructing a genetic linkage map of *Vitis aestivalis*-derived “Norton” and its use in comparing Norton and Cynthiana. Mol Breed 37:64

He J, Zhao X, Laroche A, Lu Z-X, Liu H, Li Z (2014) Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. Front Plant Sci 5

Heath, M. C., & Skalamera, D. (1997) Cellular interactions between plants and biotrophic fungal parasites. Adv Bot Res 24:195-225

Hoffmann S, Di Gaspero G, Kovács L, Howard S, Kiss E, Galbács Z, Testolin R, Kozma P (2008) Resistance to *Erysiphe necator* in the grapevine ‘Kishmish vatkana’is controlled by a single locus through restriction of hyphal growth. Theor Appl Genet 116:427-438

Huang Y-F, Poland JA, Wight CP, Jackson EW, Tinker NA (2014) Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. PloS One 9:e102448

Hyma KE, Barba P, Wang M, Londo JP, Acharya CB, Mitchell SE, Sun Q, Reisch B, Cadle-Davidson L (2015) Heterozygous mapping strategy (HetMappS) for high resolution genotyping-by-sequencing markers: a case study in grapevine. PLoS One 10:e0134880

Jaillon, O., Aury, J. M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., & Vezzi, A. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463

Jürges G, Kassemeyer HH, Dürrenberger M, Düggelin M, Nick P (2009) The mode of interaction between *Vitis* and *Plasmopara viticola* Berk. & Curt. Ex de Bary depends on the host species. Plant Biol 11:886-898

Kono A, Sato A, Reisch B, Cadle-Davidson L (2015) Effect of detergent on the quantification of grapevine downy mildew sporangia from leaf discs. HortScience 50:656-660

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754-1760

Lowe K, Walker M (2006) Genetic linkage map of the interspecific grape rootstock cross Ramsey (*Vitis champinii*)× 'Riparia Gloire' (*Vitis riparia*). Theor Appl Genet 112:1582-1592

Marguerit E, Boury C, Manicki A, Donnart M, Butterlin G, Némorin A, Wiedemann-Merdinoglu S, Merdinoglu D, Ollat N, Decroocq S (2009) Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. Theor Appl Genet 118:1261-1278

Merdinoglu D, Wiedeman-Merdinoglu S, Coste P, Dumas V, Haetty S, Butterlin G, Greif C (2003) Genetic analysis of downy mildew resistance derived from *Muscadinia rotundifolia*. Acta Hort

Moreira FM, Madini A, Marino R, Zulini L, Stefanini M, Velasco R, Kozma P, Grando MS (2011) Genetic linkage maps of two interspecific grape crosses (*Vitis* spp.) used to localize quantitative trait loci for downy mildew resistance. Tree Genet Genomes 7:153-167

Ochßner I, Hausmann L, Töpfer R (2016) *Rpv14*, a new genetic source for *Plasmopara viticola* resistance conferred by *Vitis cinerea*. Vitis 55:79-81

Pootakham W, Ruang-Areerate P, Jomchai N, Sonthirod C, Sangsrakru D, Yoocha T, Theerawattanasuk K, Nirapathpongporn K, Romruensukharom P, Tragoonrung S (2015) Construction of a high-density integrated genetic linkage map of rubber tree (*Hevea brasiliensis*) using genotyping-by-sequencing (GBS). Front Plant Sci 6

Riaz S, Dangl G, Edwards K, Meredith C (2004) A microsatellite marker based framework linkage map of *Vitis vinifera* L. Theor Appl Genet 108:864-872

Riaz S, Krivanek A, Xu K, Walker M (2006) Refined mapping of the Pierce's disease resistance locus, *PdR1*, and Sex on an extended genetic map of *Vitis rupestris*× *V. arizonica*. Theor Appl Genet 113:1317

Riaz S, Tenscher A, Rubin J, Graziani R, Pao S, Walker M (2008) Fine-scale genetic mapping of two Pierce's disease resistance loci and a major segregation distortion region on chromosome 14 of grape. Theor Appl Genet 117:671

Schwander F, Eibach R, Fechter I, Hausmann L, Zyprian E, Töpfer R (2012) *Rpv10*: a new locus from the Asian *Vitis* gene pool for pyramiding downy mildew resistance loci in grapevine. Theor Appl Genet 124:163-176

Stover E, Aradhya M, Yang J, Bautista J, Dangl GS (2009) Investigations into the origin of 'Norton' grape using SSR markers. Proceedings of the annual meeting of the Florida State Horticultural Society

Troggio M, Malacarne G, Coppola G, Segala C, Cartwright DA, Pindo M, Stefanini M, Mank R, Moroldo M, Morgante M (2007) A dense single-nucleotide polymorphism-based genetic linkage map of grapevine (*Vitis vinifera* L.) anchoring Pinot Noir bacterial artificial chromosome contigs. Genetics 176:2637-2650

van Heerden CJ, Burger P, Vermeulen A, Prins R (2014) Detection of downy and powdery mildew resistance QTL in a 'Regent'×'RedGlobe' population. Euphytica 200:281-295

Venuti S, Copetti D, Foria S, Falginella L, Hoffmann S, Bellin D, Cindrić P, Kozma P, Scalabrin S, Morgante M (2013) Historical introgression of the downy mildew resistance gene *Rpv12* from the Asian species *Vitis amurensis* into grapevine varieties. PLoS One 8:e61228

Vezzulli S, Troggio M, Coppola G, Jermakow A, Cartwright D, Zharkikh A, Stefanini M, Grando MS, Viola R, Adam-Blondon A-F (2008) A reference integrated map for cultivated grapevine (*Vitis vinifera* L.) from three crosses, based on 283 SSR and 501 SNP-based markers. Theor Appl Genet 117:499-511

Voorrips R (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77-78

Wang L, Wan ZY, Bai B, Huang SQ, Chua E, Lee M, Pang HY, Wen YF, Liu P, Liu F (2015) Construction of a high-density linkage map and fine mapping of QTL for growth in Asian seabass. Sci Rep 5

Wang, H., Misztal, I., Aguilar, I., Legarra, A., & Muir, W. M. (2012) Genome-wide association mapping including phenotypes from relatives without genotypes. Genet Res 94:73-83

Ward JA, Bhangoo J, Fernández-Fernández F, Moore P, Swanson J, Viola R, Velasco R, Bassil N, Weber CA, Sargent DJ (2013) Saturated linkage map construction in *Rubus idaeus* using genotyping by sequencing and genome-independent imputation. BMC Genomics 14:2

Welter LJ, Göktürk-Baydar N, Akkurt M, Maul E, Eibach R, Töpfer R, Zyprian EM (2007) Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). Mol Breed 20:359-374

Welter LJ, Göktürk-Baydar N, Akkurt M, Maul E, Eibach R, Töpfer R, Zyprian EM (2007) Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). Mol Breed 20:359-374

Wiedemann-Merdinoglu S, Prado E, Coste P, Dumas V, Butterlin G, Bouquet A, Merdinoglu D (2006) Genetic analysis of resistance to downy mildew from *Muscadinia rotundifolia*. 9th International Conference on Grape Genetics and Breeding 2006

Zhang Y, Wang L, Xin H, Li D, Ma C, Ding X, Hong W, Zhang X (2013) Construction of a high-density genetic map for sesame based on large scale marker development by specific length amplified fragment (SLAF) sequencing. BMC Plant Biol 13:141

Zhou X, Xia Y, Ren X, Chen Y, Huang L, Huang S, Liao B, Lei Y, Yan L, Jiang H (2014) Construction of a SNP-based genetic linkage map in cultivated peanut based on large scale marker development using next-generation double-digest restriction-site-associated DNA sequencing (ddRADseq). BMC Genomics 15:351

Zyprian E, Ochßner I, Schwander F, Šimon S, Hausmann L, Bonow-Rex M, Moreno-Sanz P, Grando MS, Wiedemann-Merdinoglu S, Merdinoglu D (2016) Quantitative trait loci affecting pathogen resistance and ripening of grapevines. Mol Genet Genomics 291:1573-1594

CHAPTER 5

IDENTIFICATION OF NEW FATB1A MUTATIONS IN MUTANT SOYBEAN LINES WITH LOW PALMITIC ACID IN THE SEED OIL

Introduction

Origin and Domestication of Soybean

Soybean (*Glycine max* (L.) Merr.) is one of the major edible legumes which is cosmopolitan in distribution. This crop is first thought to be domesticated in the eastern half of North China in the eleventh century B.C. or later from the wild soybean *Glycine soja* Sieb. and Zucc. (Xu et al. 2002). The ancient crop was used in many Asian countries as a major food crop; however, the crop was initially expanded to other continents as a forage crop. Soybean was introduced in the United States in 1804 and quickly became an important crop particularly in the South and Midwest in the mid-20th century. Currently, the US is the leading producer of soybean in the world followed by Brazil and Argentina (Hymowitz et al. 1983).

Uses of Soybean

Soybean is an important source of material for various edible and industrial products that are of daily use, and great economic importance. Soybean seed consists of 40% protein, 20% oil, 35% carbohydrate and 5% ash, in addition to various health beneficial isoflavones and minerals (Lui 1997). Soybean is mostly consumed as seed oil and is the most widely consumed edible oil in the United States, 54% of total oil consumption. The majority of this oil is used for salad/cooking oil and frying/baking, representing ~53% and ~ 21% of soybean oil utilization, respectively (http://soystats.com/soybean-oil-u-s-vegetable-oils-consumption/). In addition, it is the leading source of biodiesel (80% of the domestic biodiesel production in US), meal for

human and animals, as well as constituents of various industrial products such as cosmetics, plastics, inks, pesticides, lubricants and many others.

Processed products like soymilk and tofu are other major uses of soybean for human consumption. Soymilk, a beverage made by soaking dry soybean seeds followed by grinding, contains significant amounts of protein with less fat (Soyfood.com, USDA). Tofu, also called soybean curd, is rich in high quality proteins and B-vitamins with low sodium. This is one of the important sources of dietary meals in the East and Southeast Asia (Hymowitz and Newell 1981). In addition, other well-known products from soybean are soy sauce and miso. Soy meal, a high protein diet generated after extracting oil from soybean seed, is a popular feed for farm animals like poultry, pork, and cattle. This diet is not only rich in protein (44-49%), but also easy to digest, and releases amino acids that are ideal for nonruminants like swine and poultry. Soy meal is rich in amino acids like lysine, tryptophan, threonine, isoleucine, and valine- the amino acids that are seriously deficient in corn, grain sorghum, and other cereal gains that are commonly fed to animals (Fickler et al. 1995). These amino acids alone meet the nutritional requirements of farm animals

Soybean seed composition

In general, soybean requires 108 to 144 days from seed germination to the recovery of matured seed. However, this is highly dependent on numerous factors such as maturity group, genotypes, and growth conditions. Seed formation is a series of events from seed fill through a process of cell division and cell elongation for formation of cellular organelles to the accumulation of oils and proteins (Carson and Lersten 1987). At

physiological maturity, a typical soybean seed contains 20% oil, 40% proteins, and 35% carbohydrates, although this varies greatly with genotypes and environmental conditions. For example, changes in environmental conditions can result in either an extension or premature termination of seed fill compared to the normal process. An extended period of seed fill results in larger seed with increased protein and decreased oil and carbohydrate, and the opposite is true for harsh environmental conditions such as drought (Egli et al. 1978; Sato and Ikeda 1979). Similarly, low temperature during seed development results in a greater accumulation of unsaturated fatty acid and starch, lower total oil content and a decreased accumulation of oleic acid (Graham and Patterson 1982).

Seed carbohydrate

Soybean seed contains 35% carbohydrates, which is non-starch polysaccharides such as cellulose, hemicelluloses, and pectin. The majority of soluble carbohydrate at seed maturity is either sucrose (41-68%), stachyose (12-35%) or raffinose (5-16%). Stachyose and raffinose occur late in seed development, prior to or during seed dehydration (Amuti and Pollard 1997). Though these are major carbohydrates present in seed, other carbohydrates such as fructose, glucose, and galactose are also present in trace amount (Schweizer et al. 1978). Larger amounts of raffinose and stachyose in soybean seed are problematic for feeds due to their inability to be digested.

Seed protein

The average cultivated soybean contains about 40% of protein, with many variations among the accessions. Proteins are in the form of globulins, 11S glycinin and

7S beta-conglycinin that contain many of the amino acids such as Isoleucine, Leucine, Lysine, Methionine, Cysteine, Phenylanine, Tyrosine, Threonine, Tryptophan and Valine (Carrera et. al 2011). Efforts have been made to incorporate high protein contents into commercially cultivated varieties through various measures including backcross breeding. However, typical strong negative relationships have been noted between protein content and yield (Hartwing and Hinson 1972).

Soybean seed oil

Soybean seed contains approximately 21% of oil, of which triglycerides are the major component (Hartwing and Kilen 1991). This widely consumed vegetable oil consists of five major fatty acids: 12% palmitic (16:0), 4% stearic (18:0), 23% oleic (18:1), 53% linoleic (18:2), and 8% linolenic (18:3) acids (Fehr 2007). Among them, 16:0 and 18:0 fractions are saturated fatty acids and constitute 16% of the soybean oil. The relative composition of these saturated and unsaturated fatty acids determines the quality of oil in soybean. Decreased amounts of saturated palmitic acid and increased amounts of unsaturated oleic acid is considered optimum for human health. Higher oleic acid is desirable, since this monounsaturated fatty acid not only helps to improve shelf life but also reduces the need for hydrogenation (Wilson 2004; Pham et al. 2010). Minimal amounts of linoleic and linolenic acid are also considered beneficial to improve oxidative stability (Mozaffarian et al. 2006).

In addition, soybean is a source of various vitamins (B-vitamins), tocopherols and minerals (Ghani et al. 2016). It also has various isoflavone compounds, phytosterols, phospholipids, saponins, ferritins (Kanchana et al. 2015).

Fatty acid modification on soybean

Soybean seed oil is known to be heritable quantitatively and is affected by genotype, environment and their interaction (Burton 1987). The key goal for improving soybean oil quality has been to increase its oxidative stability by reducing its linolenate content, improving functionality and nutritional quality. Linoleic acid oxidizes readily, resulting in off-flavours, rancidity, and reduced performance (Frankel 1980). In addition, there is great interest in reducing saturated content of soybean oil due to health concerns (Ascherio et al. 1999).

Oil biosynthesis in soybean takes place by two different stages: fatty acid biosynthesis and triacylglycerol assembly. The biosynthesis of 16:0, 18:0, and 18:1 occurs in acyl-carrier protein (ACP) in the plastid and biosynthesis of 18:2 and 18:3 occurs on glycerolipids following acylation of 18:1 on the glycerol-3-phophate backbone in the endoplasmic reticulum or plastid (Stahl et al. 2004).

In the past, various researche studies were conducted to alter fatty acid profiles in soybean seed through measures like induced and natural mutations, generating transgenic lines, and combining one or more genes thorugh conventional crosses. Several genes and quantitative loci have been discovered that affect fatty acid composition in soybean (Shi et al. 2015; Guo et al. 2006; Gillman et al. 2009 and 2014; Pham et al. 2012 and 2013; Lee et al. 2012; Flores et al. 2008; Chappell et al. 2007; Bilyeu et al. 2003, 2005 and 2006; Lee et al. 2007; Clemente et al. 2009; Wilcox et al. 1984; Aghoram et al. 2006). Most of these studies utilized modern molecular genetic tools to alter oil content as per end-user preferences to improve industrial uses, and in other products. Creation of

mutants is an important step to create breeding material with altered oil contents. Oil composition of a mutant and the parent is compared under different environmental conditions, and are used for successive breeding programs. There are various mutants available for reduced levels of 18:2 and 18:3 (Wilcox et al. 1984; Chappell et al. 2006; Bilyeu et al. 2011), increased levels of 18:0 and 18:1 (Graef et al. 1985; Pham et al. 2010), and reduced levels of 16:0 (Erickson et al. 1988, Wilcox and Cavins 1990). Though creation of mutants to alter oil content is an effective approach, this method has a few shortcomings. It can be difficult to detect mutants in genes whose loss is lethal or does not result in a phenotype. Furthermore, it can be difficult to detect mutants when, homologous genes code for the same function, and mutants may also affect both membrane and oil lipid composition. Sometimes, the induction of a mutant may have a penalty for agronomic performance that is difficult to accept in the market-place (Miquel 1994).

Among the five common fatty acids, palmitic acid (16:0) accounts for 12% of the total fatty acid composition, and is a predominate saturated fatty acid in soybean oil. Consumption of a diet rich in saturated fats increases heart disease, a leading cause of death in United States (www.heart.org). The American heart association recommends limiting the consumption of saturated fats and replacing them with mono- and poly-unsaturated fatty acids. As palmitic acid is a major contributor for saturated fatty acids, reducing palmitic acid content in newly identified soybean lines would be a major achievement to improve oil quality for human consumption.

Previously, numerous attempts have been practiced to generate soybean lines with low palmitic acid by breeding, mutation or transgenic approach. Since natural genetic

variability for low palmitic acid is limited, the most effective method for modifying fatty acid composition involved development of mutant lines via ethyl methane sulfonate (EMS) (Gillman and Bilyeu 2012). In the past, lines C1726 (8.5% 16:0) and ELLP2 (7.1% 16:0) were developed using EMS mutagenesis of the cultivars Century and Elgin87, respectively. Line A22 (7.8% 16:0) was developed by *N-nitroso-N-methyl* urea mutagenesis of A1937, whereas line J3 (5.7% 16:0) was developed from cultivar 'Bay' using X-radiation (Cardinal et al. 2007). Additional lines were developed through a recurrent selection method- N87-2122-4, N79-2077, N94-2575, and C1943. Soybean line N87-2122-4 has palmitic acid level of 5.3% and is an F6-derived line from a cross between N78-2245 and N79-2077 (Burton et al. 1994 and 1998). Genetic studies have identified at least two independent mutant loci, *fap1* and *fap3* that result in reduction of palmitic acid content to ~8-9% and ~7-8%, respectively (Erickson et al. 1988, Schnebly et al. 1994). By combining these two mutants, along with other minor modifier genes, lines with < 4.0 % were developed (Fehr 2007).

Similarly, previous studies have identified fap_{nc}, a major genetic locus encoding 16:0- acyl carrier protein (ACP) thioesterase activity that is known to reduce seed palmate with a deletion on *FATB* gene. Cardinal et al. (2007) identified a specific deletion on *FATB* isoform called *GmFATB1a* by isolating full-length cDNAs of three of the four unique FATB genes that are homozygous for fap_{nc} expressed in soybean. This deletion is also found to influence 18:0 content in seed oil; most likely, the encoded enzyme also acts on 18:0-ACP substrates. De Vries et al. (2011) identified a mutant of *fap3 (A22)* allele, allelic to fap_{nc} that is responsible for reduced palmitate in seed oil. One study showed a single nucleotide polymorphism (SNP) in the region that differentiates it

from *fap3,* resulting in a nonconservative amino acid substitution that is likely to be detrimental to the function of 16:0-ACP thioesterase. In another study in mutant line RG3, a new null allele of *fap3-ug*/GmFATB1A (derived from a line ELLP2) was identified that is responsible for reducing palmitic acid from 12.2% to 6.6%. This line also had a splice site mutation in the *fap1/KASIII* gene candidate (Gillman et al. 2014; Cardinal et al. 2014). Recently, Thapa et al. (2016) located a new alleles of FATB1a (a distinct mutation on FATB1a-Glyma.05G012300) on two different lines that co-segregated with a reduced palmitic acid phenotype and was responsible for a 30% reduction in palmitic acid. Two lines; FATB1a $_{G180D}$ and FATB1a$_{splice}$ carry missense and splice mutations respectively with similar palmitic acid level.

In this study, we were able to identify and confirm three mutant lines that carried independent *FATB1a* mutations that showed low levels of palmitic acid in the seed oil. After further testing, these new genetic resources can be utilized in future breeding programs to create new cultivars with low palmitic acid content.

Materials and Methods

Plant materials and phenotyping

Initially, four different mutant HUDSON lines were obtained from Dr. Hudson, USDA/ARS and Department of Agronomy, Purdue University., West Lafayette, IN. Briefly, these lines were obtained by screening more than 5000 N-nitro-N-methyl urea (NMU)-mutagenized soybean lines using fatty acid profiling by gas chromatography as described by Thapa et al. (2016). These HUDSON lines were inbred mutant individuals (M4 generation) harvested and phenotyped separately. A mapping population was

developed from these mutant lines through crosses with another parent ‘LG04-6000’ in 2013. The F_1 population obtained was advanced to F_3 generation in Costa Rica and was used for fatty acid analysis.

Mutant line 1668 was a remutagenesis of elevated stearic acid mutant 194D originally derived from an EMS induced ‘Williams 82’ mutant population (Gillman et al. 2014).

These lines along with control ‘William 82’ were grown under field conditions with individual plants as replicates in a completely randomized design at University of Missouri South Research Farm. To analyze fatty acids, a composite of 5 seeds from each replicate was chosen. Fatty acid analysis was performed on individual seed as described by Bilyeu et al. (2005). The concentration of palmitic acid in the sample was determined as a percentage of the total fatty acids of the seed by lipid gas chromatography of fatty acid methyl esters of extracted oil. Standard fatty acid mixtures (Animal or vegetable Oil reference Mixture 6, AOACS) were used as calibration reference standards.

DNA extraction and PCR amplification

Genomic DNA was extracted using the manufacturer’s protocols for the DNeasy Plant Mini Kit (Qiagen, Valencia, CA USA). For each sample, 10 soybean seeds were grounded into a fine powder using a mortar and pestle in the presence of liquid nitrogen, and 0.02 g of this ground powder was transferred to a micro centrifuge tube. Samples were then processed according to the DNeasy Plant Mini Kit protocol. All the extracted samples were kept at -20^{0}C before further processing. *FATB1a*/Glyma05g08060 gene specific primers were designed using Primer3Plus software

(http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). Because of the large gene size (>4kb), 5 different amplification primers were designed, covering the entire gene (Table 1). Primers were confirmed to be gene specific by blasting against the unmasked *Glycine max* 'Williams 82' genomic sequence (www.phytozome.net/soybean) to avoid unwanted amplification. PCR amplification was performed using Ex Taq according to the manufacturer's protocol (Takara, Otsu, Shiga, Japan) in a thermocycler with following conditions: denaturation at 95^{0}C for an initial 5-min followed by 35 cycles of 95^{0}C for 20s, annealing at 60^{0}C for 20s and an extension step at 72^{0}C for 30s followed by additional extension for 5 min with same temperature. PCR products were run on 1.5% agarose gels to check amplification of appropriate size.

Sequencing and confirmation of mutation

PCR products were purified using a QIAquick PCR purification kit (Qiagen). During the process, the amount of PCR products was adjusted according to the intensity of band. Following purification, the products were Sanger sequenced at the DNA Core Facility at the University of Missouri-Columbia using the primers that were designed for amplification. Sequencing was done from both sides (forward and reverse); to make sure it covered the entire region. The resultant sequence was imported into Multiple Sequence Alignment-CLUSTALW (http://www.genome.jp/tools-bin/clustalw), visualized with Box Shade (http://embnet.vital-it.ch/software/BOX_form.html) and was evaluated for variant nucleotides between mutated lines and the reference 'Williams 82' sequence (http://www.phytozome.net/soybean). To identify whether there are any changes in

amino acids, protein translation of the obtained sequence was conducted using ExPaSy protein translation tools (http://web.expasy.org/translate/).

Results

Identification of FATB1a mutations

Fatty acid profiles of all HUDSON and the control 'W82' lines were obtained through composite seed from the replicated trials (Table 2; Fig. 1). Using two different mutagenesis experiments, the study was able to identify six different mutant lines with low saturated fatty acid levels including palmitic acid. The maximum effect was about a 40% reduction in palmitic acid in line HUDSON 17448.1 compared to the palmitic acid content in 'Williams 82', a control. Reduced palmitic acid levels in the mutants were responsible for the reduction in total saturated fats in the seed oil. However, none of the new mutant lines achieved the level of saturated fatty acids present in breeding line 'SS04-2564' resulting from a FATB1a missense mutation plus the *fap1* mutation (Cardinal et al. 2014; De Vries et al. 2011).

Out of the four different HUDSON mutant lines, HUDSON lines; 17448.1, 17591.12 and 19668 have significantly low levels of palmitic acid; 4.6, 6.2 and 6.4% respectively. These lines were genotyped to identify possible mutations in the FATB1a gene. During the process, a full length genomic GmFATB1a sequence was amplified and sequenced. This study was able to identify a FATB1a mutation in two different HUDSON lines, possibly responsible for reducing palmitic acid level in seed oil.

Analysis of the sequence of the HUDSON 17448.1 mutant line revealed a nucleotide change from G to A in the coding region of FATB1a (compared to the

reference 'Williams 82' genome sequence) (Fig. 2A and Fig. 5). This single base pair change resulted in a change in amino acid from aspartic acid to asparagine (Fig. 2B). Re-amplification of the FATB1a in the mutation region confirmed the presence of a single base pair change. However, we did not detect any SNPs in other amplified regions. The induced mutation in HUDSON 17448.1 lies in a highly conserved region of the acyl-ACP thioesterase protein family (Fig. 2C).

Similarly, sequence analysis of the HUDSON 19668 mutant line detected a nucleotide change from G to A, as that of HUDSON 17448.1 line, in the coding region of FATB1a after comparing with reference 'Williams 82' genome sequence (Fig. 3A). The single base pair change resulted in a change in amino acid from glycine to aspartic acid (Fig. 3B). Re-amplification confirmed the changes in the amino acid, ultimately changing the gene function. Conservation of the amino acid in the region was studied using web logo (http://weblogo.berkeley.edu/logo.cgi); Fig. 3C showed the induced mutation in the region is semi-conserved.

However, the study was unable to identify any sequence change on FATB1a or KAS III genes on mutant lines HUDSON 17591.12 and 19513 low. This was tested through PCR amplification of the gene and Sanger sequencing of the products. In addition, whole genome sequencing of both lines and bulked segregant analysis of HUDSON 17591.12 lines failed to identify any mutation that correlates with fatty acid changes.

Mutant line 1668, obtained through remutageneis of an elevated stearic acid mutant in a 'Williams 82' background showed a reduced level of palmitic acid to 6.6%; however, this was achieved without decreasing the average saturated fatty acid content in

the seed oil (17.0%). As with the previous mutants, DNA from this line was amplified to check any possible FATB1a or KAS III mutation. This analysis showed a single nucleotide change from G to A in a coding region of FATB1a gene (Fig. 4A). The SNP change in the gene is responsible for changing the amino acid code of the region from arginine to glutamate (Fig. 4B). Changes in the sequence was confirmed with reamplifcation. A multiple sequence alignment of the amino acid in the region showed the amino acid is vital for gene function i.e. highly conserved among the related species (Fig. 4C).

Mapping populationa were generated by crossing mutants HUDSON 17448.1 and 19668 with the 'LG04-6000'. A phenotypic analysis of F_3 seeds showed palmitic acid segregated across lines with values ranging from 3-6 to 11.9% in HUDSON 17448.1 and 5.3 to 12.5% in HUDSON 19668 mutant line (Fig. 6). A similar mapping population was developed for mutant line HUDSON 17591.12 and is segregated for low palmitic acid (Fig. 6), even though whole genome sequencing and bulked segregant analysis failed to highlight possible causative mutations. Complementation tests with populations of the mutant lines as one parent and the low palmitic acid parent SS04-2564, containing *fap3* W231L FATB1a alleles and *fap1* KAS III splice-site alleles, confirmed allelism for HUDSON 17448.1 and HUDSON 19668 (data not shown).

Discussion

Fatty acid manipulation in soybean oil by altering oil composition is an important breeding objective in the US. The combination of saturated and unsaturated fatty acid in soybean oil ultimately determines the oil quality. Reduction in saturated fatty acid and

particularly in the levels of palmitic acid in the human diet is recommended to improve cardiovascular health (Thapa et al. 2016). Therefore, it is necessary to identify genetic methods to reduce palmitic acid levels in oils that are intended for human consumption. In this study, we were able to identify and confirm novel mutation in FATB1a in three different lines that carry low palmitic acid levels. Line HUDSON 17448.1 had very low palmitic acid levels, in addition to total saturated fatty acids, so the allele of FATB1a that was identified allele may hold particular value.

In the past, several studies have been conducted to generate soybean lines with reduced palmitic acid content through chemical mutagenesis, recurrent selection, and hybridization (Erickson et al. 1988; Wilcox and Cavins 1990; Li et al. 2002; Thapa et al. 2016; Gillman et al. 2014). In soybean, where natural germplasm is limited, creation of mutants provides an important source of novel genetic variation for various traits. Utilizing induced mutations and a candidate gene approach, we were able to identify three novel mutation in FATB1a gene that altered soybean oil composition. Palmitic acid is significantly low in the mutant lines, providing a new genetic resource for breeding for oil composition.

Changes in the amino acid composition in the FATB1a gene are responsible for changes in gene function to code for other kinds of proteins. Amino acid position 352 in HUDSON 17448.1 and 304 in 1668 are highly conserved across the FATB1a gene family, which is likely to alter the structure or function of FATB1a, resulting in reduced or abolished enzyme activity. Similarly, amino acid position 460 in HUDSON 19668 is moderately conserved, underlying its importance for gene function as well.

Using three of the mutants, mapping populations were generated (F_3) to study seed oil composition. A phenotypic study showed segregation on oil phenotypes including palmitic acid. These lines segregated for low to medium to high palmitic acid, which can be utilized for further breeding program to generate lines that have low palmitic acid. A single gene appears to account for the three categories of palmitic acid level inherited from both mutant lines. However, the study failed to identify any causal mutation that had altered changes in fatty acid compostion in mutant line HUDSON 17591.12. A further study is needed to identify a novel candidate for reduced palmitic acid.

In conclusion, in this study we were able to identify independent mutations on FATB1a in three different mutant lines. This study highlights the importance of a simple molecular candidate gene investigation composed of PCR and sequencing; to detect novel genetic mutations at the molecular level. The identified SNPs allow for the development of molecular markers and rapid assays that can be used for maker-assisted selection in future breeding programs for low palmitic acid in soybean seed oil.

Figure Legends

Fig.1 Fatty acid distribution among different mutant soybean lines and the control 'William 82". mutant line hudson 19513 low lacks replication, and so is considered as zero replicated during analysis

Fig. 2 A Sequence alignment of nucleotides between 'Williams 82' (W82) and FATB1a mutant in line HUDSON 17448.1
B Amino acid changes (in red) between 'Williams 82' (W82) and FATB1a mutant in line HUDSON 17448.1
C. Web logo output of the amino acid conservation FATB1a mutant. The overall height of the stack indicates the sequence conservation whereas height of each symbol indicates the relative frequency of each amino acid at the position. The arrow indicates the position that was identified.
Grey highlighting indicates changes in nucleotide sequence. The under lined amino acid in red is changed from the SNP, D (aspartic acid) to N (asparagine)

Fig. 3 A Sequence alignment of nucleotides between 'Williams 82' (W82) and FATB1a mutant in line HUDSON 19668
B Amino acid changes (in red) between 'Williams 82' (W82) and FATB1a mutant in line HUDSON 19668
C. Web logo output of the amino acid conservation FATB1a mutant. The overall height of the stack indicates the sequence conservation whereas height of each symbol indicates the relative frequency of each amino acid at the position. The arrow indicates the position that was identified.
Grey highlighting indicates changes in nucleotide sequence. The under lined amino acid in red is changed from the SNP, G (glycine) to D (aspartic acid)

Fig. 4 A Sequence alignment of nucleotides between 'Williams 82' (W82) and FATB1a mutant in line 1668
B Amino acid changes (in red) between 'Williams 82' (W82) and FATB1a mutant in line 1668
C Web logo output of the amino acid conservation FATB1a mutant. The overall height of the stack indicates the sequence conservation whereas height of each symbol indicates the relative frequency of each amino acid at the position. The arrow indicates the postion that was identified.
Grey highlighting indicates changes in nucleotide sequence. The under lined amino acid in red is changed from the SNP, R (arginine) to Q (gultamine)

Fig. 5 Genomic sequence of FATB1a gene. Color code representating identified SNPs in all three mutant lines. Exons are indicated with gray highlighting.

Fig. 6 Segregation of palmitic acid content in a mapping population generated by crossing HUDSON lines with LG046000.

Table 1 Primer sets used to amplify the FATB1a gene in HUDSON mutant lines

S.N.	Primer Sets	Type
	TTTGCCTATTTGCATTTCTC	Forward
1	**AACAGCCCTATCTTAGTCAAAG**	Reverse
	CACTGTAATTCAATTTCATA	Forward
2	**TAACATGGATACATTGCAAATC**	Reverse
	AACCAAAATGCTTTGGAAATTT	Forward
3	**ATGGAAGATTCAAAATTGTGC**	Reverse
	GTCATGACTGAAGTTTCAGGT	Forward
4	**TAAATTAATGACAGCAATTAGA**	Reverse
	TCTAATTGCTGTCATTAATTTA	Forward
5	**GAAGGTGTATATATAAAGAC**	Reverse

Table 2 Average fatty acid contents in mutant and control lines from a trial in Columbia MO for year 2016

Mutant Genotypes	16:0	18:0	18:1	18:2	18:3	Total sats	AVG sats	Palmitic acid (16:0)
HUDSON 17448.1	4.7	3.4	26.8	57.9	7.2	8.2	8.0	4.6
HUDSON 17448.1	4.5	3.5	28.4	57.1	6.5	7.9		
HUDSON 19497	10.6	4.1	23.7	54.6	6.9	14.7	14.6	10.3
HUDSON 19497	10.0	4.4	24.9	53.7	6.9	14.4		
HUDSON 17591.12	6.2	4.1	30.2	53.0	6.5	10.3	10.1	6.2
HUDSON 17591.12	6.2	4.0	23.5	58.4	8.0	10.1		
HUDSON 17591.12	6.3	3.5	29.1	53.7	7.4	9.8		
HUDSON 19668	5.7	4.2	58.5	28.5	3.1	9.9	10.1	6.4
HUDSON 19668	6.5	3.6	73.8	12.1	4.1	10.1		
HUDSON 19668	6.9	3.5	24.1	57.9	7.5	10.4		
19513 low	6.6	3.6	27.4	56.0	6.5	10.2	10.2	6.6
1668	6.4	10.8	22.7	52.8	7.4	17.2	17.0	6.6
1668	6.9	10.3	20.7	53.7	8.4	17.2		
1668	6.6	9.9	21.8	54.3	7.5	16.5		
W82	11.6	3.8	24.9	52.2	7.4	15.5	15.5	11.6
W82	11.5	3.8	25.2	52.8	6.8	15.3		
W82	11.7	4.1	23.4	53.4	7.5	15.7		
SS03-2564	4.5	3.1	26.5	57.7	8.2	7.6	7.4	4.2
SS03-2564	4.4	3.6	26.8	57.4	7.8	8.0		
SS03-2564	3.9	2.8	27.3	58.0	7.9	6.7		
194D (636)	10.4	**9.1**	18.4	54.2	7.8	19.5	20.4	10.5
194D (636)	10.6	**10.4**	17.7	53.9	7.4	21.0		
194D (636)	10.4	**10.3**	20.2	51.6	7.5	20.8		

Figure 1

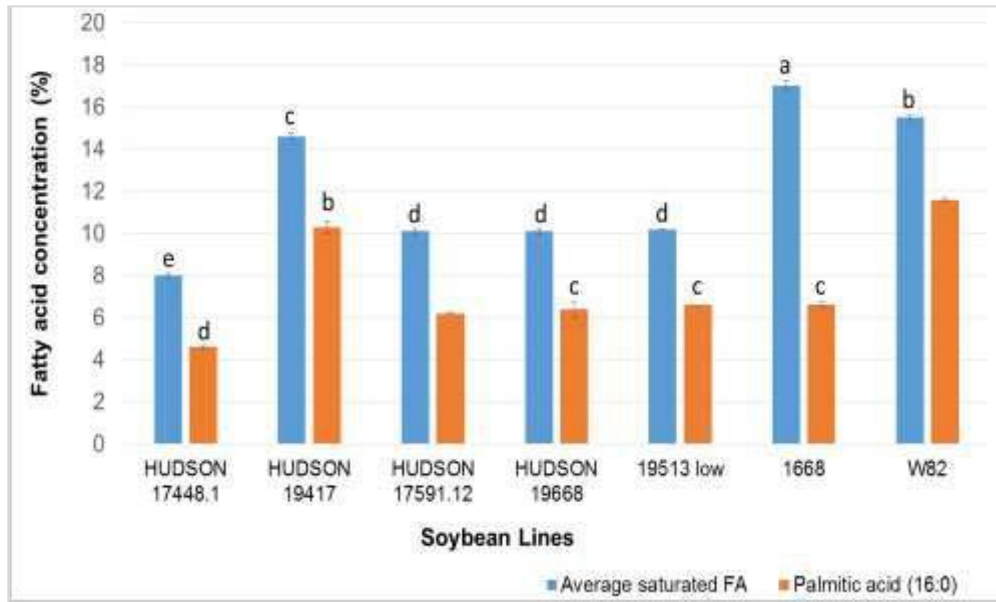


Figure 2

A. **HUDSON 17448.1**

W82	GATCTAGATATCAATCAGCATGTCAACAATGTGAAGTACATTGGCTGGATT
Mutant/FATB1a	GATCTAAATATCAATCAGCATGTCAACAATGTGAAGTACATTGGCTGGATT

B.

W82	**DL<u>D</u>INQHVNN**
Mutant/FATB1a	**DL<u>N</u>INQHVNN**

C.

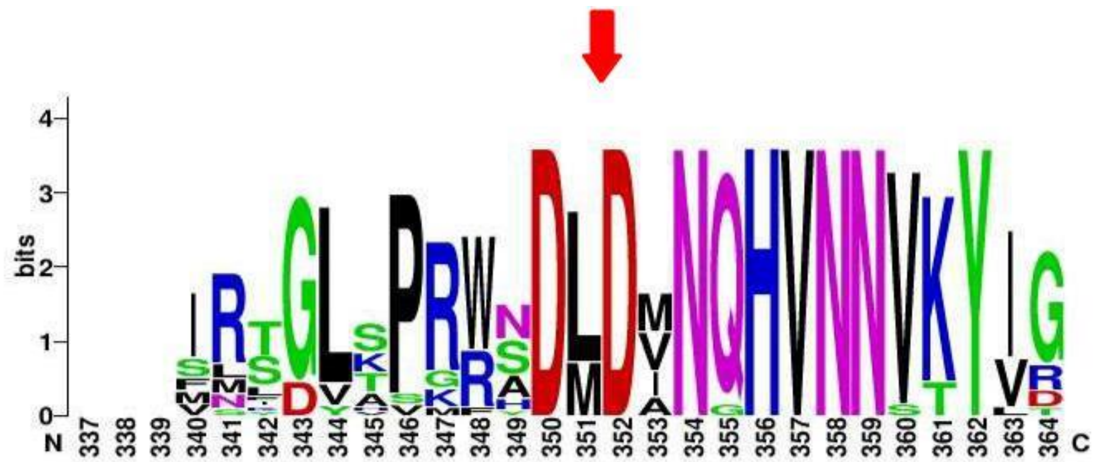


Figure 3

A. **HUDSON 19668**

W82 . AGCATTTGCTTCGACTGGAAAATGGTGCTGAGATTGTGAGGGGCAGGAC
Mutant/FATB1a AGCATTTGCTTCGACTGGAAAATGGTGCTGAGATTGTGAGGGACAGGAC

B.

W82 **GAEIVRGRTEWRPK**
Mutant/FATB1a **GAEIVRDRTEWRPK**

C.

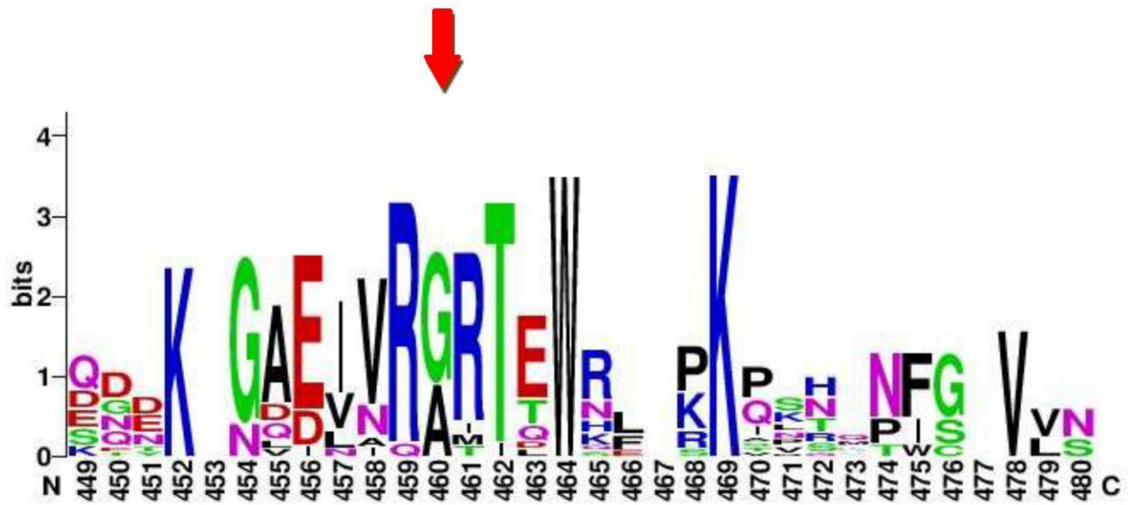


Figure 4

A. **1668**

W82 TCATGATGAATAAGCTAACACGGAGGCTGTCTAAAATTCCAGAAGAAGTCA
Mutant/FATB1a TCATGATGAATAAGCTAACACAGAGGCTGTCTAAAATTCCAGAAGAAGTCA

B.

W82 **MMNKLTRRLS KIPEEVRQE**
Mutant/FATB1a **MMNKLTQRLS KIPEEVRQE**

C.

Figure 5

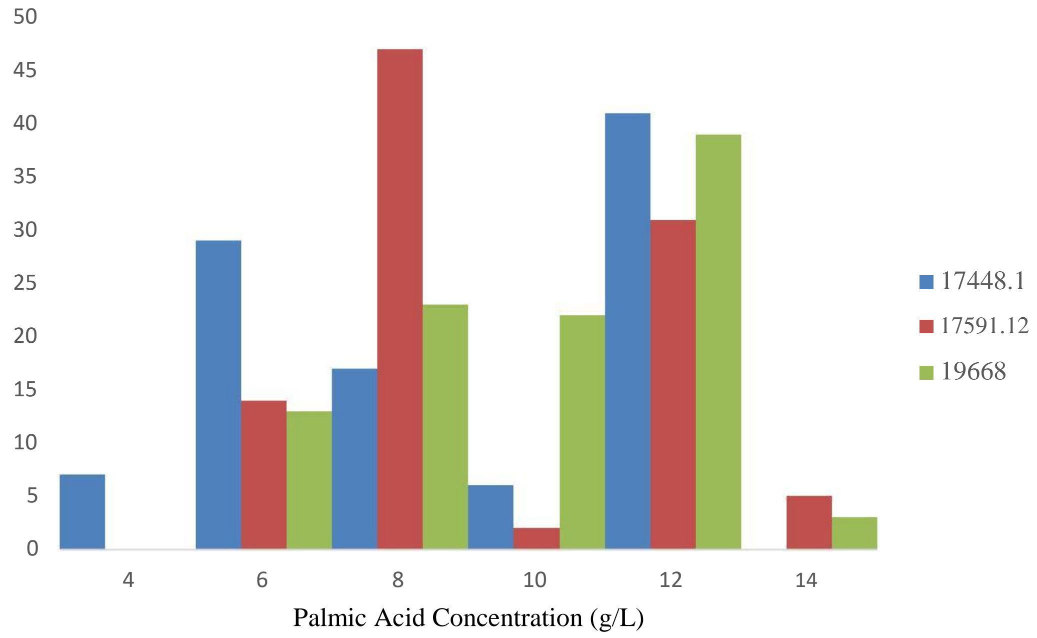
GmFATB **Acyl-ACP thioesterase**

This family consists of various acyl-acyl carrier protein (ACP) thioesterases (TE) these terminate fatty acyl group extension via hydrolysing an acyl group on a fatty acid.

>position=Gm05:7993029..7997071 (+ strand)

GCTATTCATTCATTCATTCCTCTTTCTCTCTGATCGCAAACTGCACCTCTACGCTCCACTCTTCTCATTTTCTCTTCCTTTC
TCGCTTCTCAGATCCAACTCCTCAGATAACACAAGACCAAACCCGCTTTTTCTGCATTTCTAGACTAGACGTTCTACCGG
AGAAGGTTCTCGATTCTTTTCTCTTTTAACTTTATTTTTAAAATAATAATAATGAGAGCTGGATGCGTCTGTTCGTTGTG
AATTTCGAGGCAATGGGGTTCTCATTTTCGTTACAGTTACAGATTGCATTGTCTGCTTTCCTCTTCTCCCTTGTTTCTTTG
CCTTGTCTGATTTTTCGTTTTTATTTCTTACTTTTAATTTTTGGGGATGGATGTTTTTTCTGCATTTTTTCGGTTTGCGATG
TTTTCAGGATTCCGATTCCGAGTCAGATCTGCGCCGGCTTATACGACGAATTTGTTCTTATTCGCAACTTTTCGCTTGATT
GGCTTGTTTTACCTCTGGAATCTCACACGTGATCAAATAAGCCTGCTATTTTAGTTGAAGTAGAATTTGTTCTTTATCGG
AAAGAATTCTATGGATCTGTTCTGAAATTGGAGCTACTGTTTCGAGTTGCTATTTTTTTTAGTAGTATTAAGAACAAGTT
TGCCTTTTATTTTACATTTTTTTCCTTTGCTTTTGCCAAAAGTTTTTATGATCACTCTCTTCTGTTTGTGATATAACTGATG
TGCTGTGCTGTTATTATTTGTTATTTGGGGTGAAGTATAATTTTTTGGGTGAACTTGGAGCATTTTTAGTCCGATTGATTT
CTCGATATCATTTAAGGCTAAGGTTGACCTCTACCACGCGTTTGCGTTTGATGTTTTTTCCATTTTTTTTTTATCTCATAT
CTTTTACAGTGTTTGCCTATTTGCATTTCTCTTCTTTATCCCCTTTCTGTGGAAGGTGGGAGGGAAAATGTATTTTTTTTT
TCTCTTCTAACTTGCGTATATTTTGCATGCAGCGACCTTAGAAATTCATTATGGTGGCAACAGCTGCTACTTCATCATTT
TTCCCTGTTACTTCACCCTCGCCGGACTCTGGTGGAGCAGGCAGCAAACTTGGTGGTGGGCCTGCAAACCTTGGAGGAC
TAAAATCCAAATCTGCGTCTTCTGGTGGCTTGAAGGCAAAGGCGCAAGCCCCTTCGAAAATTAATGGAACCACAGTTGT
TACATCTAAAGAAAGCTTCAAGCATGATGATGATCTACCTTCGCCTCCCCCCAGAACTTTTATCAACCAGTTGCCTGATT
GGAGCATGCTTCTTGCTGCTATCACAACAATTTTCTTGGCCGCTGAAAAGCAGTGGATGATGCTTGATTGGAAGCCACG
GCGACCTGACATGCTTATTGACCCCTTTGGGATAGGAAAAATTGTTCAGGATGGTCTTGTGTTCCGTGAAAACTTTTCTA
TTAGATCATATGAGATTGGTGCTGATCGTACCGCATCTATAGAAACAGTAATGAACCATTTGCAAGTAAGTCCGTCCTC
ATACAAGTGAATCTTTATGATCTTCAGAGATGAGTATGCTTTGACTAAGATAGGGCTGTTTATTTAGTCACTGTAATTCA
ATTTCATATATAGATAATATCATTCTGTTGTTACTTTTCATACTATATTTATATCAACTATTTGCTTAACAACAGGAAACT
GCACTTAATCATGTTAAAAGTGCTGGGCTTCTTGGTGATGGCTTTGGTTCCACGCCAGAAATGTGCAAAAAGAACTTGA
TATGGGTGGTTACTCGGATGCAGGTTGTGGTGGAACGCTATCCTACATGGTTAGTCATCTAGATTCAACCATTACATGT
GATTTGCAATGTATCCATGTTAAGCTGCTATTTCTCTGTCTATTTTAGTAATCTTTATGAGGAATGATCACTCCTAAATAT
ATTCATGGTAATTATTGAGACTTAATTATGAGAACCAAAATGCTTTGGAAATTTGTCTGGGATGAAAATTGATTAGATA
CACAAGCTTTATACATGATGAACTATGGGAAACCTTGTGCAACAGAGCTATTGATCTGTACAAGAGATGTAGTATAGCA
TTAATTACATGTTATTAGATAAGGTGACTTATCCTTGTTTAATTATTGTAAAAATAGAAGCTGATACTATGTATTCTTTG
CATTTGTTTTCTTACCAGTTATATATACCCTCTGTTCTGTTTGAGTACTACTAGATGTATAAAGAATGCAATTATTCTGAC
TTCTTGGTGTTGGGTTGAAGTTAGATAAGCTATTAGTATTATTATGGTTATTCTAAATCTAATTATCTGAAATTGTGTGT
CTATATTTGCTTCAGGGGTGACATAGTTCAAGTGGACACTTGGGTTTCTGGATCAGGGAAGAATGGTATGCGCCGTGAT
TGGCTTTTACGTGACTGCAAAACTGGTGAAATCTTGACAAGAGCTTCCAGGTAGAAATCATTCTCTGGAATTTTCCTTCC
CCTTTCCTTCTGCTTCAAGCAAATTTTAAGATGTGTATCTTAATGTACTTGATGGTGATTGGGCACAATTTTGAATCTTCC
ATACATTTTAAAAGTTATGGAACCCTTTCTTTTCCTTCTTAAGATGCAAATTTGTCATGACTGAAGTTTCAGGTAATCAT
TTGCATTTTGCAGTGTTAAAAAAGATAATGAACTACACATTTATTATATTTTGCAGGCAAAAACCTCTAATTAAACAAA
CTGAACATTGTATCTTAGTTTATTTATCAGACTTTATCATGTGTACTGATGCATCACCTTGGAGCTTGTAATGAATTACA
TATTAGTATTTTCTGAACTGTTTGTTATGGTTTTGGTGATCTACAGTGTTTGGGTCATGATGAATAAGCTAACACGGAGG
CTGTCTAAAATTCCAGAAGAAGTCAGACAGGAGATAGGATCTTATTTTGTGGATTCTGATCCAATTCTGGAAGAGGATA
ACAGAAAACTGACTAAACTTGACGACAACACAGCGGATTATATTCGTACCGGTTTAAGTGTATGTCAACTAGTTTTTTT
TCTAATTGCTGTCATTAATTTATTTTCTCAAATTATTTCAGATGTTGTTTTCTAATTAGTTTACATAATGCATCTTC
ATTTTGCAGCCTAGGTGGAGTGATCTAGATATCAATCAGCATGTCAACAATGTGAAGTACATTGGCTGGATTCTGGAGG
TATTTTTCTGTTCTTGTATTCTAATCAACTGCAATCCATGTTAGTTCTTTAACCAAAGGACTGTCTTTTGATTGTTGCAGA
GTGCTCCACAGCCAATCTTGGAGAGTCATGAGCTTTCTTCCATGACTTTAGAGTATAGGAGAGAGTGTGGTAGGGACAG
TGTGCTGGATTCCCTGACTGCTGTATCTGGGGCCGACATGGGCAATCTAGCTCACAGCGGGCATGTTGAGTGCAAGCAT
TTGCTTCGACTGGAAAATGGTGCTGAGATTGTGAGGGGCAGGACTGAGTGGAGGCCCAAACCTGTGAACAACTTTGGT
GTTGTGAACCAGGTTCCAGCAGAAAGCACCTAAGATTTTGAAATGGTTAACGATTGGAGTTGCATCAGTCTCCTTGCTA
TGTTTAGACTTATTCTGGTCCCTGGGGAGAGTTTTGCTTGTGTCTATCCAATCAATCTACATGTCTTTATATATATACACC
TTCTAATTTGTGATACTTTGGTGGGTAAGGGGGAAAAGCAGCAGTAAATCTCATTCTCATTGTAATTAGCTACTGCTGT
ATTCTCTCTTTCTGCTGCTCCATATTTCATTTCATCTCTGATTGCGCTACTGCTAGGCTGTCTTCAATATTTAATTGCTTGT
TCAAAATAGCTAGGCATGTATATTATTATTCTTTTTCTCTTGGCCCAATTAAAGATGCAATTTTCCTTGTGAACACAGCA
TAATTCTTATTCTTATTTATTTTTGTATATCCTGTATGCAAGAATGACTTGTCCTCCAATACAACTGTGATTTTGTATGCT
CCAGCTTGTATTTTATGCCAGCTTCCCACATGGGAATTATTGTGC

Figure 6



References

Aghoram, K., Wilson, R. F., Burton, J. W., & Dewey, R. E. (2006) A mutation in a 3-keto-acyl-ACP synthase II gene is associated with elevated palmitic acid levels in soybean seeds. Crop Sci 46:2453-2459

Amuti, K. S., & Pollard, C. J. (1977) Soluble carbohydrates of dry and developing seeds. Phytochemistry 16:529-532

Ascherio, A., Katan, M. B., Zock, P. L., Stampfer, M. J., & Willett, W. C. (1999) Trans fatty acids and coronary heart disease. N Engl J Med 340:1994-1998

Bilyeu, K. D., Palavalli, L., Sleper, D. A., & Beuselinck, P. R. (2003) Three microsomal omega-3 fatty-acid desaturase genes contribute to soybean linolenic acid levels. Crop Sci 43:1833-1838

Bilyeu, K., Palavalli, L., Sleper, D., & Beuselinck, P. (2005) Mutations in soybean microsomal omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds. Crop Sci 45:1830-1836

Bilyeu, K., Palavalli, L., Sleper, D. A., & Beuselinck, P. (2006) Molecular genetic resources for development of 1% linolenic acid soybeans. Crop Sci 46:1913-1918

Bilyeu, K., Gillman, J. D., & LeRoy, A. R. (2011) Novel mutant allele combinations produce soybeans containing 1% linolenic acid in the seed oil. Crop Sci 51:259-264

Burton, J. W. (1987) Quantitative genetics: results relevant to soybean breeding. Agronomy (USA)

Burton, J. W., Harlow, C., & Theil, E. C. (1998) Evidence for reutilization of nodule iron in soybean seed development. J Plant Nutr 21:913-927

Burton, J. W., Wilson, R. F., & Brim, C. A. (1994) Registration of N79-2077-12 and N87-2122-4, two soybean germplasm lines with reduced palmitic acid in seed oil. Crop Sci 34: 313-313

Cardinal, A. J., Burton, J. W., Camacho-Roger, A. M., Yang, J. H., Wilson, R. F., & Dewey, R. E. (2007) Molecular analysis of soybean lines with low palmitic acid content in the seed oil. Crop Sci 47:304-310

Cardinal, A. J., Whetten, R., Wang, S., Auclair, J., Hyten, D., Cregan, P., & Upchurch, G. (2014) Mapping the low palmitate fap1 mutation and validation of its effects in soybean oil and agronomic traits in three soybean populations. Theor Appl Genet 127: 97-111

Carlson JB, Lersten NR (1987) Reproductive morphology. In JR Wilcox, ed, Soybeans: Improvement, Production, and Uses, ‘Ed 2. Agronomy Society of America, Madison, WI 95-134

Carrera, C. S., Reynoso, C. M., Funes, G. J., Martínez, M. J., Dardanelli, J., & Resnik, S. L. (2011) Amino acid composition of soybean seeds as affected by climatic variables. Pesq Agropec Bras 46:1579-1587

Chappell, A. S., & Bilyeu, K. D. (2006) A GmFAD3A mutation in the low linolenic acid soybean mutant C1640. Plant Breeding 125:535-536

Chappell, A. S., & Bilyeu, K. D. (2007) The low linolenic acid soybean line PI 361088B contains a novel mutation. Crop Sci 47:1705-1710

Clemente, T. E., & Cahoon, E. B. (2009) Soybean oil: genetic approaches for modification of functionality and total content. Plant Physiol 151:1030-1040

De Vries, B. D., Fehr, W. R., Welke, G. A., & Dewey, R. E. (2011) Molecular characterization of the mutant (A22) allele for reduced palmitate concentration in soybean. Crop Sci 51:1611-1616

Egli, D. B., Leggett, J. E., & Wood, J. M. (1978) Influence of soybean seed size and position on the rate and duration of filling. Agron J 70:127-130

Erickson, E. A., Wilcox, J. R., & Cavins, J. F. (1988) Inheritance of altered palmitic acid percentage in two soybean mutants. J Hered 79:465-468

Erickson, E. A., Wilcox, J. R., & Cavins, J. F. (1988) Inheritance of altered palmitic acid percentage in two soybean mutants. J Hered 79:465-468

Fehr, W. R. (2007) Breeding for modified fatty acid composition in soybean. Crop Sci 47

Fickler, J., Kirchgessner, M., & Roth, F. X. (1995) The effect of dietary arginine supply on the N balance of piglets, 4: Importance of non-essential amino acids for protein retention. J Anim Physiol Anim Nutr

Flores, T., Karpova, O., Su, X., Zeng, P., Bilyeu, K., Sleper, D. A., & Zhang, Z. J. (2008) Silencing of GmFAD3 gene by siRNA leads to low α-linolenic acids (18: 3) of fad3-mutant phenotype in soybean [*Glycine max* (Merr.)]. Transgenic Res 17:839-850

Frankel, E. N. (1980) Soybean oil flavor stability. Handbook of soy oil processing and utilization 229-244

Ghani, M., Kulkarni, K. P., Song, J. T., Shannon, J. G., & Lee, J. D. (2016) Soybean Sprouts: A review of nutrient composition, health benefits and genetic variation. Plant Breed Biotechnol 4:398-412

Gillman, J. D., Pantalone, V. R., & Bilyeu, K. (2009) The low phytic acid phenotype in soybean line CX1834 is due to mutations in two homologs of the maize low phytic acid gene. Plant Genome 2:179-190

Gillman, J. D., & Bilyeu, K. (2012) Genes and alleles for quality traits on the soybean genetic/physical map. Designing Soybeans for 21st Century Markets 67-96

Gillman, J. D., Tetlow, A., Hagely, K., Boersma, J. G., Cardinal, A., Rajcan, I., & Bilyeu, K. (2014) Identification of the molecular genetic basis of the low palmitic acid seed oil trait in soybean mutant line RG3 and association analysis of molecular markers with elevated seed stearic acid and reduced seed palmitic acid. Mol Breed 34:447-455

Gillman, J. D., Stacey, M. G., Cui, Y., Berg, H. R., & Stacey, G. (2014) Deletions of the SACPD-C locus elevate seed stearic acid levels but also result in fatty acid and morphological alterations in nitrogen fixing nodules. BMC Plant Biol 14:143

Graef, G. L., Fehr, W. R., & Hammond, E. G. (1985) Inheritance of three stearic acid mutants of soybean. Crop Sci *25*:1076-1079

Graham, D., & Patterson, B. D. (1982) Responses of plants to low, nonfreezing temperatures: proteins, metabolism, and acclimation. Annu Rev Plant Physiol 33:347-372

Guo, B., Sleper, D. A., Lu, P., Shannon, J. G., Nguyen, H. T., & Arelli, P. R. (2006) QTLs associated with resistance to soybean cyst nematode in soybean: meta-analysis of QTL locations. Crop Sci 46:595-602

Hartwig, E. E., & Hinson, K. (1972) Association between chemical composition of seed and seed yield of soybeans. Crop Sci 12:829-830

Hartwig, E. E., & Kilen, T. C. (1991) Yield and composition of soybean seed from parents with different protein, similar yield. Crop Sci 31:290-292

Hymowitz, T., & Newell, C. A. (1981) Taxonomy of the genus *Glycine*, domestication and uses of soybeans. Econ Bot 35:272-288

Hymowitz, T., & Harlan, J. R. (1983) Introduction of soybean to North America by Samuel Bowen in 1765. Econ Bot 37:371-379

Kanchana, P., Santha, M. L., & Raja, K. D. (2015) A review on *Glycine max* (L.) Merr.(soybean). World J Pharm Res 5:356-371

Lee, J. D., Bilyeu, K. D., & Shannon, J. G. (2007) Genetics and breeding for modified fatty acid profile in soybean seed oil. J Crop Sci Biotech 10:201-210

Lee, J. H., & Cho, K. M. (2012) Changes occurring in compositional components of black soybeans maintained at room temperature for different storage periods. Food Chem 131:161-169

Li, Z., & Nelson, R. L. (2002) RAPD marker diversity among cultivated and wild soybean accessions from four Chinese provinces. Crop Sci 42:1737-1744

Liu, J., Bollinger, D. W., Ledoux, D. R., Ellersieck, M. R., & Veum, T. L. (1997) Soaking increases the efficacy of supplemental microbial phytase in a low-phosphorus corn-soybean meal diet for growing pigs. J Anim Sci 75:1292-1298

Miquel, M. F. (1994) High-oleate oilseeds fail to develop at low temperature. Plant Physiol 106:421-427

Mozaffarian, D., Katan, M. B., Ascherio, A., Stampfer, M. J., & Willett, W. C. (2006) Trans fatty acids and cardiovascular disease. N Engl J Med 354:1601-1613

Pham, A. T., Lee, J. D., Shannon, J. G., & Bilyeu, K. D. (2010) Mutant alleles of FAD2-1A and FAD2-1B combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol 10:195

Pham, A. T., Shannon, J. G., & Bilyeu, K. D. (2012) Combinations of mutant FAD2 and FAD3 genes to produce high oleic acid and low linolenic acid soybean oil. Theor Appl Genet 125:503-515

Pham, A. T., McNally, K., Abdel-Haleem, H., Boerma, H. R., & Li, Z. (2013) Fine mapping and identification of candidate genes controlling the resistance to southern root-knot nematode in PI 96354. Theor Appl Genet 126:1825-1838

SATO, K., & IKEDA, T. (1979) The Growth responses of soybean plant to photoperiod and temperature: IV. The effect of temperature during the ripening period on the yield and characters of seeds. Jpn J Crop Sci 48:283-290

Schnebly, S. R., Fehr, W. R., Welke, G. A., Hammond, E. G., & Duvick, D. N. (1994) Inheritance of reduced and elevated palmitate in mutant lines of soybean. Crop Sci 34:829-833

Schweizer, T. F., Horman, I., & Würsch, P. (1978) Low molecular weight carbohydrates from leguminous seeds; a new disaccharide: galactopinitol. J Sci Food Agric 29:148-154

Shi, Z., Liu, S., Noe, J., Arelli, P., Meksem, K., & Li, Z. (2015) SNP identification and marker assay development for high-throughput selection of soybean cyst nematode resistance. BMC Genomics 16:314

Stahl, A. (2004) A current review of fatty acid transport proteins (SLC27). Pflügers Archiv 447:722-727

Thapa, R., Carrero-Colon, M., Crowe, M., Gaskin, E., & Hudson, K. (2016) Novel alleles confer an elevated oleic acid phenotype in soybean seeds. Crop Sci 56:226-231

Wilcox, J. R., & Cavins, J. F. (1990) Registration of C1726 and C1727 soybean germplasm with altered levels of palmitic acid. Crop Sci 30

Wilcox, J. R., Cavins, J. F., & Nielsen, N. C. (1984) Genetic alteration of soybean oil composition by a chemical mutagen. J Am Oil Chem Soc 61:97-100

Wilson, R. F. (2004) Seed composition. Soybeans: improvement, production, and uses, (soybeansimprove) 621-677

Xu, D., Abe, J., Gai, J., & Shimamoto, Y. (2002) Diversity of chloroplast DNA SSRs in wild and cultivated soybeans: evidence for multiple origins of cultivated soybean. Theor Appl Genet 105:645-653

APPENDIX

Bulked Segregant Analysis (BSA)

Bulked segregant analysis is a genotyping technique that involves screening for differences between two pooled DNA samples derived from a segregating population originated from a single cross. This is a rapid method for identifying markers in the specific regions of the genome (Michelmore et al. 1991). During the process, segregating bulk of each individual is created by pooling DNA from many samples (usually 10-15) for genotyping. In each pool, the individuals are identical for all other genes except for the trait or gene of interest. The pools are analyzed, and markers that are polymorphic between the pools will be genetically linked to the trait/loci used to construct the pools. BSA is highly effective with high marker density and accurate allele frequency estimation within bulks (Magwene et al. 2011). This can be used for any populations with significant phenotypic differences for the target trait among individuals. Samples can be collected from segregating individuals obtained from bi- or multiparents or from variants from any populations of a species including those with diverse genetic background (Zhou et al. 2016).

With the availability of next generation sequencing techniques and rapid development of molecular marker assays, BSA has also witnessed many improvements (Brauer et al. 2006; Schneerberger et al. 2009; Doitsidou et al. 2010). BSA has been successfully applied in the budding yeast *Saccharomyces cerevisiae* to uncover the genetic basis of Mendelian traits (Birkeland et al. 2010, Wenger et al. 2010) and multi-gene traits (Ehrenreich et al. 2010). Similarly, with the availability of these advanced sequencing techniques, the pooled DNA analysis can be used for any two contrasting groups of individuals from any populations, not just from bi-parental mapping population

(Xu et al. 2008). The individuals with extreme phenotypes form natural populations can be bulked for whole genome sequencing, genotyping-by-sequencing or genome wide association study (GWAS) (Bastide et al. 2013; Turner et al. 2010).

However, number of factors including the sampling of segregants and measurement techniques affect BSA. Sampling of segregants or bulk error can be minimized by increasing sample size for bulk and the population. The other source of variation includes DNA sequencing techniques used; including library preparation, sequencing chemistry, sequence coverage, post alignment of reads, and allele/base calling algorithms. This can be minimized by adopting a standard set of protocols (Magwene et al. 2011).

There are few studies that have utilized BSA as a means of genotyping for marker-trait association in soybean. An *Rpp3* locus was identified using BSA to soybean rust through GoldenGate assay (Hyten et al. 2009). Meksem et al. (2001) developed high-density genetic maps and identified QTLs that confers resistance to soybean cyst nematode. BSA was utilized to improve seed quality in soybean by increasing protein content (Dordevic et al. 2008). Moreover, a recent study utilizes next generation sequencing from BSA pools to identify two qualitative genes that are responsible for governing cotyledon color (Song et al. 2017).

Whole Genome Sequencing (WGS)

With recent advances in sequencing techniques, it has driven a revolution in genomic analysis of genomes of animals and plants, and their use for future application. Whole genome sequencing provides the most comprehensive collection of an individual's

genetic variation. Several whole genome sequencing projects are in effect aiming to unveil novel insights in genomic evolution, and diversity including effects of domestication and human selection (Kim et al. 2010, Lam et al. 2010, and Ling et al. 2013). There are a number of studies that utilize sequencing techniques to identity novel candidates for various traits, including some studies on soybean (Maroof et al. 2008; Lam et al. 2010; Hyten et al. 2010; Li et al. 2013; Qi et al. 2014). Use of next generation techniques in combination with other high throughput SNP detection assays would be able to discover larger number of SNPs that are associated with possible candidates, including genes that are responsible for reduce palmitic acid.

Materials and Methods

Bulked Segregant Analysis

Mutant soybean line HUDSON 17591.12 was used to genotype through BSA. During the process, DNA from high and low bulks were pooled together and sent for GBS (Table 1 and 2, Fig. 1). DNA was extracted, and quantified to meet certain standard (100ng) according to the protocol provided by the GBS center based at Cornell University. A DNA sample of 30 uL for each genotype was sent for genotyping. Results obtained after genotyping were analyzed.

Whole Genome Sequencing

Three of the HUDSON mutant lines (17591.12, 19513 low and 19513 high) that segregate for low palmitic acid with control were sent for whole genome sequencing to

find any possible candidates responsible for low palmitic acid. The sequenced data was aligned and annotated using CLC genomics work bench.

Results

Initially, markers linked with low palmitic acid were identified and confirmed through whole genome sequence data. Gene annotation was able to identify four possible candidates on chromosome 8 that were more likely responsible for low saturated fatty acid (Appendix Table 6). However, PCR amplification of the *FATa* gene on the region failed to detect any SNPs. However, a possible annotation error was identified for the gene that prevented a complete analysis. A further data analysis is needed to identify other possible candidates in the region. Appendix table 3-5 summarizes the mapping analysis for three different HUDSON lines that were sequenced.

References

Bastide, H., Betancourt, A., Nolte, V., Tobler, R., Stöbe, P., Futschik, A., & Schlötterer, C. (2013) A genome-wide, fine-scale map of natural pigmentation variation in Drosophila melanogaster. PLoS Genet 9:e1003534

Birkeland, S. R., Jin, N., Ozdemir, A. C., Lyons, R. H., Weisman, L. S., & Wilson, T. E. (2010) Discovery of mutations in *Saccharomyces cerevisiae* by pooled linkage analysis and whole-genome sequencing. Genetics 186:1127-1137

Brauer, M. J., Christianson, C. M., Pai, D. A., & Dunham, M. J. (2006) Mapping novel traits by array-assisted bulk segregant analysis in *Saccharomyces cerevisiae*. Genetics 173:1813-1816

Doitsidou, M., Poole, R. J., Sarin, S., Bigelow, H., & Hobert, O. (2010) *C. elegans* mutant identification with a one-step whole-genome-sequencing and SNP mapping strategy. PloS One 5:e15435

Đorđević, V., Miladinović, J., Balešević-Tubić, S., Đukić, V., & Mikić, A. (2008) Bulk segregation analysis and improvement of seed quality in soybean. In Grain Legumes Technology Transfer Platform (GL-TTP) Workshop, 2, Novi Sad (Serbia). Institute of Field and Vegetable Crops

Ehrenreich, I. M., Torabi, N., Jia, Y., Kent, J., Martis, S., Shapiro, J. A., & Kruglyak, L. (2010) Dissection of genetically complex traits with extremely large pools of yeast segregants. Nature 464:1039

Hyten, D. L., Cannon, S. B., Song, Q., Weeks, N., Fickus, E. W., Shoemaker, R. C., & Cregan, P. B. (2010) High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. BMC Genomics 11:38

Hyten, D. L., Smith, J. R., Frederick, R. D., Tucker, M. L., Song, Q., & Cregan, P. B. (2009) Bulked segregant analysis using the GoldenGate assay to locate the locus that confers resistance to soybean rust in soybean. Crop Sci 49:265-271

Kim, M. Y., Lee, S., Van, K., Kim, T. H., Jeong, S. C., Choi, I. Y., & Kim, W. Y. (2010) Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. Proc Natl Acad Sci 107:22032-22037

Lam, H. M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F. L., & Li, J. (2010) Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nature Genet 42:1053-1059

Li, Y. H., Zhao, S. C., Ma, J. X., Li, D., Yan, L., Li, J., & Chang, R. Z. (2013) Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. BMC Genomics 14:579

Magwene, P. M., Willis, J. H., & Kelly, J. K. (2011) The statistics of bulk segregant analysis using next generation sequencing. PLoS Comput Biol 7:e1002255

Meksem, K., Pantazopoulos, P., Njiti, V. N., Hyten, L. D., Arelli, P. R., & Lightfoot, D. A. (2001) 'Forrest'resistance to the soybean cyst nematode is bigenic: saturation mapping of the *Rhg1*and *Rhg4* loci. Theor Appl Genet 103:710-717

Michelmore, R. W., Paran, I., & Kesseli, R. V. (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci 88:9828-9832

Qi, X., Li, M. W., Xie, M., Liu, X., Ni, M., Shao, G., & Isobe, S. (2014) Identification of a novel salt tolerance gene in wild soybean by whole-genome sequencing. Nat commun 5

Schneeberger, K., Ossowski, S., Lanz, C., Juul, T., Petersen, A. H., Nielsen, K. L., & Andersen, S. U. (2009) SHOREmap: simultaneous mapping and mutation identification by deep sequencing. Nature Methods 6:550-551

Song, J., Liu, Z., Hong, H., Ma, Y., Tian, L., Li, X., & Qiu, L. J. (2016) Identification and validation of loci governing seed coat color by combining association mapping and bulk segregation analysis in soybean. PloS One 11: e0159064

Turner, T. L., Bourne, E. C., Von Wettberg, E. J., Hu, T. T., & Nuzhdin, S. V. (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. Nature Genet 42:260-263

Wenger, J. W., Schwartz, K., & Sherlock, G. (2010) Bulk segregant analysis by high-throughput sequencing reveals a novel xylose utilization gene from *Saccharomyces cerevisiae*. PLoS Genet 6: e1000942

Xu, Y., Wang, J., & Crouch, J. H. (2008) Selective genotyping and pooled DNA analysis: An innovative use of an old concept. In Recognizing Past Achievement, Meeting Future Needs, Proceedings of the 5th International Crop Science Congress

Zou, C., Wang, P., & Xu, Y. (2016) Bulked sample analysis in genetics, genomics and crop improvement. Plant Biotechnol J 14:1941-1955

Appendix Table 1 Low and high pools for BSA for line HUDSON 17448.1

Genotypes	16:00	18:00	18:01	18:02	18:03	
144	5.6	3.7	26.7	56.6	7.4	
109	5.7	3.1	20.7	63.1	7.5	
146	5.7	4.0	31.5	51.6	7.2	
155	5.7	3.1	26.1	57.4	7.7	
103	5.7	3.2	21.8	61.7	7.6	
134	5.7	3.2	32.0	52.3	6.9	
106	5.7	3.2	19.8	63.2	8.0	
131	5.8	3.2	28.6	55.0	7.4	
159	5.9	3.6	22.3	60.5	7.9	
124	5.9	3.5	24.0	59.1	7.5	low pool
133	10.4	3.6	17.5	59.7	8.7	
154	10.4	4.0	24.8	53.2	7.6	
160	10.4	4.3	22.9	54.5	7.9	
147	10.6	3.7	23.2	55.5	6.9	
145	10.7	3.8	23.0	55.6	6.9	
138	10.9	3.9	24.6	53.3	7.3	
120	11.0	3.9	20.0	57.1	8.0	
121	11.1	4.0	19.8	57.2	8.0	
102	11.1	4.0	18.9	57.9	8.0	
98	12.0	3.8	21.2	55.5	7.5	high pool

Appendix Table 2 DNA concentrations and adjustment according to GBS protocol for BSA

Parent 1	17591.12	Mutant parent derived from W82		
Parent 2	LG04-6000			
Genotypes	**DNA concentration**	**Concentration Needed**	**Volume needed**	
144	15.279		17.64349761	
109	38.351		7.029151782	
146	48.238		5.588436502	
155	15.279		17.64349761	
103	15.279		17.64349761	Low pool
134	18.575		14.512786	
106	5.3915	50	50	
131	15.279		17.64349761	
159	15.279		17.64349761	
124	18.575		14.512786	
			179.8606483	
133	25.167		23.80696944	
154	21.871		27.39472361	
160	35.055		17.09171302	
147	11.983	50	50	High pool
145	25.167		23.80696944	
138	48.238		12.42070567	
120	15.279		39.21395379	
121	11.983		50	
102	35.055		17.09171302	
98	15.279		39.21395379	
			300.0407018	

Appendix Table 3 Summary of mapping reads from whole genome sequencing on line HUDSON 17591.12

Name	Consensus length	Total read counts	Single reads	Reads in pairs	Average coverage	Reference sequence	Reference length
Chr01 mapping	54574144	7395639	1783533	5612106	15.021527	Chr01	56831624
Chr02 mapping	46725922	6041870	1478246	4563624	14.294104	Chr02	48577505
Chr03 mapping	43383733	5470543	1295645	4174898	13.765347	Chr03	45779781
Chr04 mapping	50404461	6707148	1623218	5083930	14.803059	Chr04	52389146
Chr05 mapping	40297773	5342784	1322066	4020718	14.587199	Chr05	42234498
Chr06 mapping	48586499	6196623	1506529	4690094	13.900239	Chr06	51416486
Chr07 mapping	42671155	5433929	1305615	4128314	14.012894	Chr07	44630646
Chr08 mapping	46090106	5771476	1393876	4377600	13.85054	Chr08	47837940
Chr09 mapping	48073068	6280042	1521386	4758656	14.447848	Chr09	50189764
Chr10 mapping	49527741	6391507	1496905	4894602	14.294937	Chr10	51566898
Chr11 mapping	33373855	4094207	958817	3135390	13.540462	Chr11	34766867
Chr12 mapping	38160500	5027129	1194761	3832368	14.434115	Chr12	40091314
Chr13 mapping	43324673	5326082	1291330	4034752	13.295604	Chr13	45874162
Chr14 mapping	47179968	6313773	1527257	4786516	14.850434	Chr14	49042192
Chr15 mapping	49315055	6610337	1602001	5008336	14.713635	Chr15	51756343
Chr16 mapping	36207222	4631603	1093903	3537700	14.086767	Chr16	37887014
Chr17 mapping	39964485	5133223	1221299	3911924	14.190733	Chr17	41641366
Chr18 mapping	55550491	7404948	1795758	5609190	14.730526	Chr18	58018742
Chr19 mapping	48469396	6335438	1509184	4826254	14.418577	Chr19	50746916
Chr20 mapping	46117596	6238003	1568581	4669422	15.061725	Chr20	47904181

Appendix Table 4 Summary of mapping reads from whole genome sequencing on line HUDSON 19513 high

Name	Consensus length	Total read count	Single reads	Reads in pairs	Average coverage	Reference sequence	Reference length
Chr01 mapping	54671894	11402213	2459631	8942582	19.466848	Chr01	56831624
Chr02 mapping	46850485	9412431	2065043	7347388	18.780052	Chr02	48577505
Chr03 mapping	43548070	8311223	1795437	6515786	17.606102	Chr03	45779781
Chr04 mapping	50381870	9944469	2186163	7758306	18.402205	Chr04	52389146
Chr05 mapping	40379258	8256940	1824542	6432398	18.967411	Chr05	42234498
Chr06 mapping	48607989	9495130	2106396	7388734	17.912137	Chr06	51416486
Chr07 mapping	42685085	8370647	1818983	6551664	18.174449	Chr07	44630646
Chr08 mapping	46246743	9014306	1941546	7072760	18.251683	Chr08	47837940
Chr09 mapping	48080592	9509902	2100280	7409622	18.37697	Chr09	50189764
Chr10 mapping	49682252	9926159	2049409	7876750	18.690354	Chr10	51566898
Chr11 mapping	33489994	6402979	1336715	5066264	17.856235	Chr11	34766867
Chr12 mapping	38257192	7811162	1660320	6150842	18.886421	Chr12	40091314
Chr13 mapping	43531761	8359294	1796818	6562476	17.643756	Chr13	45874162
Chr14 mapping	47254734	9762660	2118238	7644422	19.301852	Chr14	49042192
Chr15 mapping	49274999	9846431	2246299	7600132	18.414032	Chr15	51756343
Chr16 mapping	36301405	7158361	1506823	5651538	18.334822	Chr16	37887014
Chr17 mapping	40094308	7929573	1657651	6271922	18.464503	Chr17	41641366
Chr18 mapping	55523812	10909810	2472892	8436918	18.21823	Chr18	58018742
Chr19 mapping	48394227	9391847	2070363	7321484	17.950569	Chr19	50746916
Chr20 mapping	46203435	9513784	2055160	7458624	19.255829	Chr20	47904181

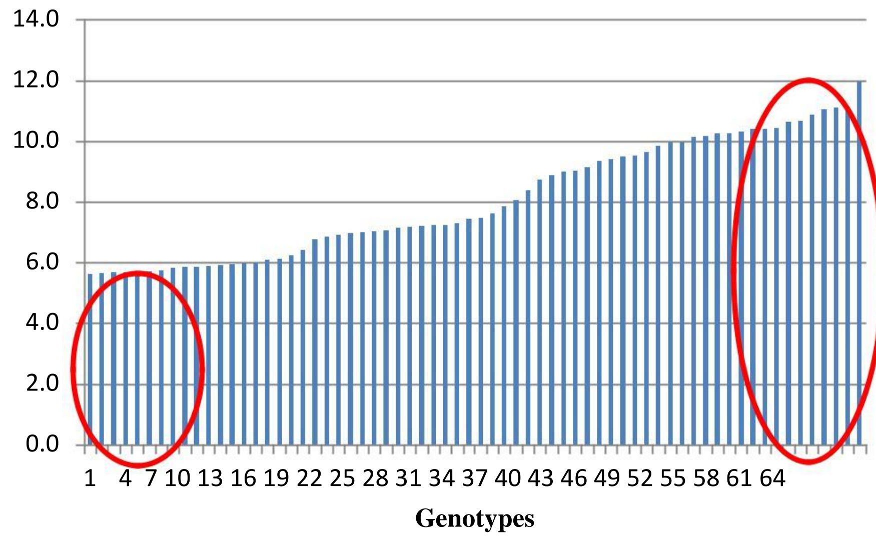
Appendix Table 5 Summary of mapping reads from whole genome sequencing on line HUDSON 19513 low

Name	Consensus length	Total read count	Single reads	Reads in pairs	Average coverage	Reference sequence	Reference length
Chr01 mapping	53950249	6187595	1498461	4689134	13.018858	Chr01	56831624
Chr02 mapping	46061705	5013802	1239856	3773946	12.281379	Chr02	48577505
Chr03 mapping	42653376	4535245	1082327	3452918	11.815967	Chr03	45779781
Chr04 mapping	49808136	5616226	1359806	4256420	12.846246	Chr04	52389146
Chr05 mapping	39753607	4445903	1108033	3337870	12.573328	Chr05	42234498
Chr06 mapping	47855911	5158295	1268517	3889778	11.980491	Chr06	51416486
Chr07 mapping	41966578	4502900	1092682	3410218	12.014331	Chr07	44630646
Chr08 mapping	45389990	4762062	1161002	3601060	11.829104	Chr08	47837940
Chr09 mapping	47429955	5251657	1281735	3969922	12.517298	Chr09	50189764
Chr10 mapping	48853681	5310042	1252976	4057066	12.29804	Chr10	51566898
Chr11 mapping	32855672	3387571	804025	2583546	11.594969	Chr11	34766867
Chr12 mapping	37632394	4202792	1008474	3194318	12.489847	Chr12	40091314
Chr13 mapping	42614436	4388430	1079876	3308554	11.341267	Chr13	45874162
Chr14 mapping	46582042	5292850	1291814	4001036	12.892898	Chr14	49042192
Chr15 mapping	48685085	5539335	1352347	4186988	12.766873	Chr15	51756343
Chr16 mapping	35717979	3876976	923122	2953854	12.204896	Chr16	37887014
Chr17 mapping	39369273	4256243	1021407	3234836	12.185512	Chr17	41641366
Chr18 mapping	54852583	6205908	1512090	4693818	12.778019	Chr18	58018742
Chr19 mapping	47860393	5300380	1272864	4027516	12.488026	Chr19	50746916
Chr20 mapping	45537114	5167836	1277626	3890210	12.904579	Chr20	47904181

Appendix Table 6 List of possible candidates responsible for low palmitic acid

Gm08	phytozome v9_0	gene	43775538	43779231	Glyma.08 g318400	ID=351 503	agi_genecode=GDSL-like Lipase%2FAcylhydrolase superfamily protein
Gm08	phytozome v9_0	gene	43765532	43771537	Glyma.08 g318300	ID=351 469	agi_genecode=tetraacyl disaccharide 4%5C%27-kinase family protein
Gm08	phytozome v9_0	gene	44154085	44155155	Glyma.08 g323100	ID=351 942	agi_genecode=MBOAT (membrane bound O-acyl transferase) family protein
Gm08	phytozome v9_0	gene	46347921	46352990	Glyma.08 g349200	ID=355 553	pfam=Acyl-ACP thioesterase

Appendix Figure 1



VITA

Surya Datta Sapkota was born in one of the remote places of Nepal, Mallaj Parbat. He came from a farming background, as his parents spent most of the time in the farm, a source of living for his family. At the age of 10, he with his family members migrated to Chitwan, Nepal to access good education. He completed his secondary level education from Sagarmatha Secondary Boarding School, Jamunapur Chitwan and high school from Apex Acadamy, Kshetrapur Chitwan.

Inspired by his parents, he joined the only agriculture university in the country, Institute of Agriculture and Animal Science, Rampur Chitwan and obtained the Bachelor of Science in Agriculture (B.Sc. Ag.) degree in 2008. After that, he speb a few years working with farmers to address their food security needs through FAO/UN. In 2011, with an ambition of obtaining a higher degree, he came to the U.S. and joined Master's in Plant Science at Missouri State University, Springfield Missouri. He then completed his Masters degree in 2013 and joined University of Missouri and Missouri State University's collaborative PhD program in the same year. In December 2017, he obtained his doctorate degree in Plant, Insect and Microbial Sciences with an emphasis on Plant Breeding, Genetics and Genomics.

Surya is very passionate about agriculture, loves interacting with farmers, identifying problems, asking research questions, and attempting to solve them. In his free time, he enjoys being with his family and friends.