Integrating Genomics and Metabolomics to Inform Breeding for Powdery Mildew Resistance in Grapevine

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF
THE UNIVERSITY OF MINNESOTA
BY

Soon Li Teh

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adrian D. Hegeman & James J. Luby

August 2018

Acknowledgements

Words are inadequate to express my utmost gratitude to my advisors, Professor Adrian D. Hegeman and Professor Jim J. Luby, for navigating me through my graduate education. I am indebted to their dedication and expertise in nurturing and 'fathering' me in the fields of metabolomics and genetics. Their gracious support allowed me to engage in collaborative research projects, participate in technical training, as well as travel to national and international conference meetings. It was a great privilege to be your student over the years!

I would like to acknowledge my advisory committee members, Professors Robert Blanchette, Jim Bradeen and Matthew Clark for their advice and guidance in various aspects of this project, particularly in areas of my research blind spots. Special thanks to Dr. Matthew Clark, an excellent scientist and a good friend, for his continued mentoring, editorial feedback, and ready availability.

Much of the work in this dissertation would not have been possible without the collaboration of Dr. Lance Cadle-Davidson (USDA-ARS), Professor Anne Fennell (SDSU), Dr. Jonathan Fresnedo-Ramírez (OSU), Dr. David Gadoury, Dr. Elizabeth Demmings, Bety Rostandy, and Mani Awale. Notably, Drs. Cadle-Davidson and Gadoury were gracious hosts who provided technical training on controlled fungal inoculation in Geneva, NY.

My heartfelt thanks to John and Jennifer Thull for their viticultural support and ready availability. Chapters 2 and 4 would not have existed without them! Many thanks to the members of Cohen-Gardner-Hegeman and Clark-Hokanson-Luby groups for their feedback and helpful discussion. In particular, the support of Dr. Nicholas Howard, John Tillman and Luke Haggerty was tremendous.

I owe my due respects to Dr. Alison Pawlus, Dr. Dana Freund and Dr. Arthur Eschenlauer for their selflessness in nurturing me in the areas of natural products, metabolomics, and statistical analysis, respectively. Their guidance helped me build a strong foundation in these areas.

I am also grateful for a conducive environment to pursue a graduate study that was made possible by Professor Emily Hoover, who leads the Department of Horticultural Science, as well as Lynne Medgaarden and the Director of Graduate Studies, who coordinate the Applied Plant Sciences.

Funding was largely provided by the University of Minnesota and the USDA-NIFA-SCRI. I am also appreciative of a number of awards that provided financial support and recognition, including the UMN Doctoral Dissertation Fellowship, William H. Alderman Graduate Award, and Daniels Fellowship in Horticulture.

I extend my cordial thanks to friends and family for always believing in me. My mom, my dad and my sisters have been a constant source of encouragement. To Huey Yee Lim, whose love is unconditional, thank you for your unwavering support throughout this journey and beyond.

Finally, I thank the Lord Almighty for the opportunity and privilege to pursue a path I dearly love. I am thankful for His constant love and faithfulness, especially when the path was arduous and gruesome.

Dedication

This dissertation is dedicated to farmers, particularly those in impoverished nations. Your unfailing drive to provide for the families has been a source of inspiration and reflection. May this work and my future endeavors be one step closer to bridging the chasm of poverty and malnourishment.

Abstract

Two powdery mildew resistance loci have been identified using pedigree-connected F_1 mapping families at the University of Minnesota grape breeding program. A consensus linkage map of the resistant parent (MN1264) was developed for genetic mapping. The resistance loci were mapped on chromosomes 2 and 15, with additive effects accounting for over 30% phenotypic variation. Marker haplotypes, hap+$_{chr2}$ and hap+$_{chr15}$, were constructed to trace the inheritance of resistance loci in grandparent-parent-progeny relationships. Both hap+$_{chr2}$ and hap+$_{chr15}$ in the resistant F_1 progeny were inherited from parent MN1264, that originated from grandparent 'Seyval blanc'. Additionally, two microsatellites markers (i.e., UDV-015b and VViv67) were identified to be associated with hap+$_{chr15}$, and can be applied for marker-assisted selection. In a follow-up study to characterize metabolic changes attributed to hap+$_{chr2}$ and hap+$_{chr15}$, a metabolomic experiment was conducted on whole-plant propagated grapes in a time-course response to *in vivo* inoculation. The use of several multivariate analyses systematically identified 52 biomarkers that were associated with hap+$_{chr2}$, and 12 biomarkers with hap+$_{chr15}$. In a temporal assessment of biomarkers, the discriminating metabolic changes distinguishing resistant and susceptible individuals appeared to be occurring from 24 to 48 hours after inoculation.

Table of Contents

Page

List of Tables ix

List of Figures x

Chapter 1 Literature Review 1

1.1 Introduction....1

1.2 Goal, Objectives and Significance....2

1.3 The Host: Grapevine 3

1.4 The Pathogen: *Erysiphe necator* 5

1.5 Genetics 7

1.6 Defense Response in Grapevine....9

1.7 Stilbene Phytoalexins: Metabolic Response against Powdery Mildew 9

1.8 The Tools: Genomic and Metabolomic Resources....12

Figures 15

Chapter 2 Genetic Dissection of Powdery Mildew Resistance in Interspecific Half-Sib Grapevine Families Using SNP-based Maps 18

2.1 Introduction....19

2.2 Materials and Methods 22

2.2.1 Plant Material 22

2.2.2 Quantification of Field Powdery Mildew Severity....23

2.2.3 DNA Extraction, Sequencing, and SNP Assignment....24

2.2.4 Linkage Maps 25
2.2.5 Construction of a Consensus Linkage Map 26
2.2.6 QTL Analysis 27
2.2.7 Haplotype Construction and Analysis with *Ren3* SSR 28
2.2.8 Haploblock Validation Using Controlled Inoculation 29
2.3 Results 29
2.3.1 Genotyping: Sequencing and SNP Assignment 29
2.3.2 Parental and Consensus Linkage Maps 30
2.3.3 Field Assessment of Powdery Mildew Severity 31
2.3.4 Analysis of Marker-Trait Association 31
2.3.5 Haplotypes Associated with Powdery Mildew Resistance 32
2.3.6 Haploblock Validation Using Controlled Inoculation 34
2.4 Discussion 35
2.4.1 Genotyping: Sequencing and SNP Assignment 35
2.4.2 Parental and Consensus Maps 36
2.4.3 Field Assessment of Powdery Mildew Severity 38
2.4.4 QTL Detection and Validation 38
2.4.5 Haplotypes Associated with Powdery Mildew Resistance 41
2.4.6 Co-localization with *Ren3*? 42
2.4.7 Chromosome 2: Ren10 and Potential Structural Dynamics in the Vicinity 43
2.5 Conclusions 44
Tables 46

Figures 52

Chapter 3 Genetic Analysis of Stilbenes in Grapevine Stems Indicates an mQTL Hotspot for Disease Resistance Motifs on Chromosome 18 62

3.1 Introduction 63

3.2 Materials and Methods 65

3.2.1 Sampling of Plant Material 65

3.2.2 Sample Preparation and Extraction 65

3.2.3 Metabolic Profiling with LC-HRMS 66

3.2.4 Data Pre-processing 66

3.2.5 Metabolic Quantitative Trait Locus (mQTL) Mapping 67

3.2.6 Integration with Functional Annotation Network 67

3.3 Results 68

3.3.1 Profiling of F_2 Mapping Family Revealed Metabolic Inheritance and Segregation 68

3.3.2 mQTL Mapping 69

3.3.3 An mQTL Hotspot on LG18 for Disease Resistance Proteins and Motifs 70

3.4 Discussion 71

3.4.1 Mapping Genomic Regions Using Variation in Metabolite Levels of the Progeny 71

3.4.2 No mQTL Was Detected for Most Metabolic Features 73

3.4.3 mQTL Hotspot on LG18 for Disease Resistance Proteins and Motifs 73
3.4.4 mQTL on LG18 Associated with Primary Metabolism 74
3.5 Conclusions 76
Tables 77
Figures 79
Chapter 4 Identification of Biomarkers Associated with Grapevine Powdery Mildew Resistance Loci in a Controlled Environment 83
4.1 Introduction 84
4.2 Materials and Methods 87
4.2.1 Plant Material 87
4.2.2 Propagation and Maintenance in a Controlled Environment (Main Chamber) 87
4.2.3 Maintenance of Powdery Mildew Inoculum (Inoculum Chamber) 88
4.2.4 *In vivo* Inoculation of Plants in a Controlled Environment (Inoculation Chamber) 88
4.2.5 Sample Preparation and Extraction 89
4.2.6 Untargeted Metabolomics 90
4.2.7 Data Processing 91
4.2.8 Data Analysis 91
4.2.9 Scaling and Transformation 92

4.2.10 Biomarker Discovery 92
4.3 Results 95
4.3.1 Double-Resistant-Haplotype Biomarkers across Generations ..95
4.3.2 Haplotype-Specific Biomarkers 96
4.3.3 Temporal Progression of Biomarkers 97
4.4 Discussion 99
4.4.1 Maintenance of Plants in a Controlled Environment 100
4.4.2 Metabolic Changes in Response to *in vivo* Inoculation 101
4.4.3 Discovery of Biomarkers Associated with Two Resistant Haplotypes 101
4.4.4 Discovery of Haplotype-Specific Biomarkers 102
4.4.5 Temporal Progression of Biomarkers 104
4.4.6 Recommendations and Reservations for Breeding Applications 106
4.5 Conclusions 107
Tables 108
Figures 113
Chapter 5 Conclusions 118
Figure 120
Bibliography 121

List of Tables

Page

Chapter 2

Table 2.1 Number of SNPs and genetic distance for each linkage group of the maternal genetic maps of GE0711/1009 and GE1025....46
Table 2.2 Summary of two-year QTL information 47
Table 2.3 Number of SNPs and genetic distance for each linkage group of the paternal genetic maps of GE0711/1009 and GE1025.....48
Table 2.4 Markers haplotypes for QTLs on LG 2 and LG 15.....49
Table 2.5 Statistical analysis of *Ren3* and *Ren10* prediction of powdery mildew response variables at 9-day post-inoculation following controlled inoculation with *Erysiphe necator* isolate NY19.....50

Chapter 3

Table 3.1 Summary of metabolic features, patterns of stilbene oligomers, mQTLs and functional annotations.....77

Chapter 4

Table 4.1 Summary of genotypes collected in the field for propagation in a controlled environment.....108
Table 4.2 Biomarkers associated with two resistant haplotypes 109
Table 4.3 Haplotype-specific biomarkers 111

List of Figures

Page

Chapter 1

Figure 1.1 Structure of *trans*-resveratrol, 3,5,4′-trihydroxy-*trans*-stilbene...15
Figure 1.2 A subset of stilbene oligomers 16

Chapter 2

Figure 2.1 Two-generation pedigrees and ancestry backgrounds of GE0711/1009 and GE1025.....52
Figure 2.2 Correlation between haplotypes and disease severity 53
Figure 2.3 Correlation between haplotypes and SSR allele sizes.....54
Figure 2.4 Two-dimensional histogram and scattered distribution of field powdery mildew severity in two years.....55
Figure 2.5 Distribution of field powdery mildew severity in two years.....56
Figure 2.6 QTL mapping on LG 2 and LG 15.....57
Figure 2.7 Maternal maps of GE0711/1009 and GE1025 58
Figure 2.8 Paternal map of GE0711/1009.....59
Figure 2.9 Paternal map of GE1025.....60
Figure 2.10 Distribution of sequence read fraction with respect to coverage read depth in the *Ren10* QTL region.....61

Chapter 3

Figure 3.1 Integrative workflow of genomics, metabolic profiling, mQTL mapping and functional annotations.....79
Figure 3.2 Total ion chromatograms for grape stem extracts of grandparent, parent, and a representative F_2 offspring.....80
Figure 3.3 Extracted ion chromatograms of two representative offspring for different patterns of stilbene oligomers 81
Figure 3.4 mQTL detection for metabolic features corresponding to different patterns of stilbene oligomerization.....82

Chapter 4

Figure 4.1 Biomarkers associated with dual resistant haplotypes.....113
Figure 4.2 Summary of four-way elliptic Venn diagram constructed in *post hoc* OPLS-DA analyses of each offspring.....114
Figure 4.3 A series of two sequential SUS plots for biomarker discovery ..115
Figure 4.4 Temporal progression of *k-means*-clustered features based on genotypes with common sampling time points.....116
Figure 4.5 Four-way elliptic Venn diagrams constructed for a range of $VIP4P_{OPLS}$ 117

Chapter 5

Figure 5.1 Summary of dissertation objectives.....120

Chapter 1

Literature Review

1.1 Introduction

The vast majority of the world's wine industry is based on the cultivation of a single grapevine species, *Vitis vinifera* L. However, this species is highly susceptible to various economically devastating pests and diseases, such as powdery mildew, which is caused by *Erysiphe necator* Schwein. To control powdery mildew, growers may rely on heavy use of chemical fungicides with up to ten fungicide applications per season. With modern grapevine cultivation moving toward sustainable practices, tremendous research efforts seek to understand the underlying grapevine defense mechanisms in order to decrease the need for chemical treatments. However, the pursuit of these objectives is impeded by numerous factors, such as genetic complexity of quantitative traits, grapevine's long generation time (three to five years), as well as costly, labor-intensive, and time-consuming maintenance of a grape breeding program. Recent advances in genomics and metabolomics provide a promising outlook for the future of grapevine breeding. The completion of reference genome sequences and physical maps, along with numerous genetic linkage maps further facilitate trait mapping and marker-assisted breeding. In addition to the advances in genetics and sequencing, the fields of functional genomics and systems biology have been rapidly growing. These approaches rely on the comprehensive profiling of large numbers of gene expression products. Since

metabolites most closely relate to phenotype, metabolomics provides a comprehensive and quantitative investigation of metabolites belonging to various chemical classes, including metabolites that play an important role in fighting pathogens.

1.2 Goal, Objectives and Significance

The overarching goal is to integrate the use of genomics and metabolomics to more precisely inform breeding for powdery mildew resistance in grapevine. The primary objectives within the framework of this dissertation are to characterize:

1. The genetic basis
    a. Identify genomic regions associated with powdery mildew resistance in a segregating population,
    b. Identify and trace the inheritance of resistant haplotype(s) in the pedigree background of the UMN grape breeding program, and
    c. Characterize molecular markers linked to resistance for MAB
2. The biochemical basis
    a. Characterize metabolic features that are linked to resistance (i.e., biomarkers), and
    b. Identify genomic regions associated with biomarker production of powdery mildew resistance.

This study will deepen our understanding of the underlying genetic and biochemical resistance to powdery mildew, thereby decreasing the need for chemical treatments, which are not compatible with modern sustainable viticulture.

The genomic studies will inform the current knowledge of the UMN grape breeding program through identification of resistance loci, and how resistant haplotype(s) are transmitted throughout complex pedigree backgrounds. Meanwhile, the metabolomic studies would elucidate the putative roles of secondary metabolites in plant defense. Finally, the complementary synergy of genomics and metabolomics will enable breeders to more reliably select genotypes for marker-assisted breeding through applications of genetic and biochemical markers.

1.3 The Host: Grapevine

Grapevine (*Vitis vinifera* L.; 2n = 38) is one of the earliest domesticated fruit crops and has been widely cultivated primarily for winemaking as well as fresh consumption, such as table grapes and raisins. This economically important perennial fruit crop is grown on over 7 million hectares worldwide, with production exceeding 77 million tons in 2016 (FAO, https://faostat3.fao.org/faostat/en/#data/QC/visualize). The genus *Vitis* consists of over 60 species that are spread mostly throughout the American (approximately 20 species) and Asian (approximately 40 species) centers of diversity (Alleweldt and Possingham, 1998). The genus can be divided into two major groups: the American group and

the Eurasian group. These groups differ significantly in their horticultural traits, thus making them attractive breeding partners. For example, *V. vinifera* belonging to the Eurasian group possesses excellent fruit quality, but lacks cold hardiness and disease resistance. On the other hand, the American species (e.g. *V. riparia* Michaux, *V. rupestris* Scheele, *V. labrusca* L., *V. aestivalis* Michaux, *V. berlandieri* Planchon) are cold hardy with generally good disease resistance, but do not possess the same level of fruit quality.

The single most cultivated grapevine species accounting for approximately 99% of world wine's production is *V. vinifera* ssp. *sativa*, which was domesticated from the wild ssp. *sylvestris*. Although the cultivated species produces hermaphroditic self-fertilizing flowers (due to selection by humans), wild grapevine species are dioecious containing either imperfect male or female flowers on different individual plants. Outbreeding by insect pollination and wind may have resulted in cultivars being highly heterozygous with many deleterious recessive mutations (Simmonds, 1976). Consequently, inbreeding depression is severe, with ensuing sterility after two to three generations of selfing. All *Vitis* species have 38 chromosomes (n = 19) and members of this genus can be readily interbred to form fertile interspecific crosses. Individuals of interspecific derivation are collectively called "hybrids" by viticulturists to indicate they are not pure *Vitis vinifera*. The high chromosome number suggests a paleopolyploid state of the genome (Lewis, 1980). In 2007, analysis of the grapevine genome sequence (487 Mb) revealed that the present-day grapevine haploid genome may have originated from the contribution of three ancestral genomes, although it

remains unclear whether the content was from a true hexaploidization event or through successive genome duplications (Jaillon et al., 2007).

In many production regions throughout the world, grapevines are predominantly grafted; thus, significant breeding efforts have been conducted to improve scions and rootstocks. While scion breeding focuses on the development of cultivars with excellent fruit quality that are adapted to biotic and abiotic stresses, rootstock breeding is aimed at improving protection against phylloxera and nematodes, as well as various soil conditions. The birth of modern grapevine breeding is closely intertwined with the arrival of a North American pest and pathogens in Europe. Successive waves of root louse, phylloxera (*Daktulosphaira vitifoliae* Fitch) as well as fungal pathogens such as powdery mildew, downy mildew and black rot were exported to European vineyards in the mid-19th-century, resulting in massive crop losses on the highly susceptible *V. vinifera* vines planted there (Owens, 2008; Reisch et al., 2012). Today, breeding programs continue to develop cultivars that are resistant to the major phylloxera and the pathogens mentioned above.

1.4 The Pathogen: *Erysiphe necator*

Caused by the obligate biotroph *Erysiphe necator* Schwein, powdery mildew is the most devastating and widespread grapevine disease. The fungus overwinters as cleistothecia, which are fungal fruiting bodies within cracks in the bark of the grapevine. When the temperature rises above 10 °C in the spring, the structures swell with water and discharge ascospores, which are then dispersed

to surrounding leaves and clusters. These ascospores germinate and start the growth of fungal colonies on the surface of the infected tissue. After colonies from primary infections have developed, the fungus produces conidia. These white powdery spores are blown by wind and do not require rain for dispersal or infection. New fungal colonies arising from these secondary infections produce even more conidia, as the repeating cycle of infection, spore production, spore dispersal and re-infection continues to spread the disease (New York State Integrated Pest Management, https://ecommons.cornell.edu/handle/1813/43121).

The fungus can infect all green tissues of the host by colonizing the epidermal cells using appresoria to penetrate the plant cuticle and cell wall (Gadoury et al., 2012). The symptom of the leaves is characterized as white powdery patches on the upper and lower leaf surfaces. When the fruits are infected, they may appear white and powdery, or dark and rusty, occasionally resulting in berry cracking and shriveling. Meanwhile, infected green shoots may have dark brown fungus lesions that remain as brown patches on the surface of dormant canes. Hence, the direct impacts of powdery mildew are often yield losses due to poor fruit set stemming from infected inflorescence, premature fruit drop, and diseased berries (Brown, 2008; https://www.ipm.ucdavis.edu/PMG/r302100311.html). In addition, the disease can result in indirect losses arising from premature defoliation, which depletes carbohydrates reserves and predisposes the vine to winter injury and reduced crop potential in the following season.

The levels of foliar resistance to powdery mildew vary both within and between *Vitis* species. Most *V. vinifera* cultivars are susceptible to powdery mildew, whereas other *Vitis* species exhibit varying levels of resistance (Cadle-Davidson et al., 2011). Previous findings have also shown the presence of race-specific resistance to *E. necator*, thereby suggesting that race-specificity could be widespread among resistant *Vitis* accessions. (Gadoury and Pearson, 1991; Cadle-Davidson et al., 2011).

1.5 Genetics

Genetic host resistance to powdery mildew is mainly derived from non-*vinifera* species, where the resistance varies from partial to complete resistance (Cadle-Davidson, 2011). Genetic resistance is typically confined to Asian and North American species, where the pathogen has co-evolved with the plant. The species include *V. aestivalis*, *V. berlandieri*, *V. cinerea*, *V. labrusca*, *V. lincecumii*, *V. riparia*, and *V. rupestris*, as well as *Muscadinia rotundifolia* Michx. (2n = 40).

At the cell surface, pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) involves recognition of an oomycete or fungal PAMP by plasma membrane receptor-like kinases to activate the PTI pathway, resulting in secretion of antimicrobial compounds and fortification of local cell walls. However, PTI can be suppressed by the pathogen through secretion of effectors to manipulate host responses in order to support compatibility. To counteract the activity of effectors, grapevine has evolved to recognize the PTI-suppressing effectors using resistance (*R*) genes, a strategy called effector-triggered

immunity (ETI). Consequently, this activates defense responses, including programmed cell death (PCD) to halt fungal development (Jones and Dangl, 2006).

Genetic mapping using bi-parental mapping populations is common in grapevine genetics, primarily for detection of quantitative trait loci (QTLs) linked to disease resistance. To date, thirteen *E. necator* major and minor resistance loci have been identified and mapped in grapevine (Akkurt et al., 2007; Barker et al., 2005; Bellin et al., 2009; Blanc et al., 2012; Blasi et al., 2011; Dalbó et al., 2001; Hoffmann et al., 2008; Mahanil et al., 2012; Marguerit et al., 2009; Merdinoglu et al., 2003; Moreira et al., 2011; Pap et al., 2016; Ramming et al., 2011; Riaz et al., 2011; Schwander et al., 2012; Teh et al., 2017; van Heerden et al., 2014; Venuti et al., 2013; Welter et al., 2007; Zendler et al., 2017; Zyprian et al., 2016). Although wild species contain resistance alleles for introgression, they often harbor an assortment of undesirable viticultural and fruit quality traits, potentially leading to off-flavored final products. Moreover, quantitative traits require introgression of several QTLs to reach a durable level of resistance.

Marker assisted selection helps overcome these limitations. The use of molecular markers allows for selection of resistant genotypes and stacking of several loci with minimal linkage drag. The importance of molecular markers to grapevine genetics is the driving force for the development of markers and genetic maps. The availability of physical maps and draft reference sequences for the grapevine genome facilitates our understanding of the distribution pattern

of candidate genes for disease resistance across the entire *V. vinifera* genome (Jaillon et al., 2007; Moroldo et al., 2008; Velasco et al., 2007).

1.6 Defense Response in Grapevine

In response to powdery mildew infection, various genes encoding key defense compounds such as PR proteins and stilbene synthase, are induced in both susceptible and resistant accessions of *V. aestivalis*, but the transcript levels of these defense-related genes start to decline after twelve hours post inoculation in disease-susceptible accessions of *V. vinifera* (Fung et al., 2008). The reduced expression level during the critical period of grapevine-powdery mildew interaction may indicate the introduction of effectors into host cells through haustoria to suppress the expression of key defense genes.

Ontogenic resistance to powdery mildew has been reported in berries of several *V. vinifera* and *V. labrusca* cultivars, where they are susceptible to *E. necator* until approximately two to three weeks post-bloom. Meanwhile, constitutive resistance to powdery mildew has been observed in genotypes highly enriched for North American *Vitis* background (Gadoury et al., 2003; Gee et al., 2008).

1.7 Stilbene Phytoalexins: Metabolic Response against Powdery Mildew

Derived from the general phenylpropanoid pathway, stilbenes represent a relatively restricted group of polyphenolic plant specialized-metabolites occurring in over 50 plant families, including Vitaceae (Rivière et al., 2012). The formation

of stilbenes is catalyzed by the stilbene synthase (STS) enzyme, via condensation of *p*-coumaroyl-CoA and three malonyl-CoA to produce the basic stilbene unit, *trans*-resveratrol (3,5,4′-trihydroxy-*trans*-stilbene) (Figure 1.1). This class of phenylpropanoids has aroused much attention due to their numerous health-promoting properties, including cardioprotective, neuroprotective, anticarcinogenic, antioxidant, and anti-inflammatory, among others (Rayne et al., 2008).

Although stilbenes have been well studied with regard to human health, relatively little is known about the potential roles of these compounds in plant defense. Found abundantly in grapevine, stilbenes also play a crucial role as phytoalexins, which are low-molecular-weight compounds that accumulate upon pathogen infection and display antibiotic activities (Adrian et al., 1997; Bavaresco et al., 1997; Pezet et al., 2004a; Pezet et al., 2004b; Schnee et al., 2008; Schnee et al., 2013). Within *Vitis*, stilbenes have been found in approximately 20 species, as well as interspecific hybrids. This class of compounds is mainly represented by resveratrol derivatives, including dimers, oligomers and glycosylated forms (Figure 1.2). A wide range of stilbenes are produced in the leaves, berries, stems, and flowers of *Vitis* species, and their expression has been shown to be higher in disease resistant cultivars than in susceptible cultivars (Pawlus et al., 2013). These observations have been supported by other studies showing that stilbene synthase and other enzymes related to stilbene oligomerization reactions are more strongly upregulated in disease resistant species.

Previous studies have shown that the levels of resveratrol and viniferins are associated with downy mildew, powdery mildew, and gray mold resistance (Bavaresco et al., 1997; Pezet et al., 2004b; Schnee et al., 2008). In addition, the antimicrobial activity of stilbenes has been demonstrated by the action of stilbenes toward major grapevine diseases. For example, a study by Adrian et al. (1997) showed *in vitro* inhibition of resveratrol on conidia germination and mycelial growth of *Botrytis cinerea*, the causal agent of gray mold disease. A separate study by Pezet et al. (2004a) demonstrated that after inoculation with *Plasmopara viticola*, leaves of a resistant cultivar (*V. vinifera* cv. Solaris) have higher concentrations of stilbenes at infection sites than those of a susceptible cultivar (*V. vinifera* cv. Chasselas).

Despite initial works in the 1980s and 1990s suggesting resveratrol as an important phytoalexin, these studies reported the ED_{50} (i.e., effective dose required for 50% mortality) of resveratrol at over 200 µg/mL, a level conflicting the definition of phytoalexin according to Müller and Börger (1940). More recently, Schnee et al. (2013) tested the activities of six isolated stilbene compounds (ampelopsin A, hopeaphenol, *trans*-resveratrol, ampelopsin H, *E*-**ε**-viniferin, and *E*-vitisin B) against *P. viticola*, *E. necator*, and *B. cineria*. Meanwhile, bioactivity was observed against *P. viticola*, but not against *E. necator* and *B. cineria*. More notably, results showed that *E*-vitisin B and hopeaphenol (tetrameric stilbenes) are most fungitoxic, while *trans*-resveratrol and ampelopsin A (monomer and dimer, respectively) are relatively benign to the pathogens *in vitro*.

The presence of a positive correlation between the production of stilbenes (resveratrol and its derivatives) and grapevine resistance to biotrophic pathogens has aroused the attention of grape breeders. In particular, grape breeder Dr. Pool (1981) pioneered an important quantitative disease screening technique to demonstrate the suitability of assessing differential stilbene production among grape clones and species (Pool et al., 1981). However, the correlation of stilbene levels and disease resistance could be more complex than just differential production of this class of compounds as a whole. More recently, studies have pointed to the types, rather than the amount of stilbenes that are implicated in grapevine resistance to the pathogens (Pezet et al., 2004a). Upon infection, downy mildew-susceptible cultivars appear to produce high amounts of resveratrol, but convert it rapidly into a less fungitoxic glucoside. In stark contrast, elevated levels of resveratrol produced by resistant cultivars are rapidly dimerized into more fungitoxic compounds, such as viniferins (Pezet et al., 2004a). Although similar results were observed in cultivars infected with powdery mildew, the mechanism of stilbene accumulation is rather different. Schnee and coworkers (2008) reported that fungitoxic viniferins, produced at the sites of infection, were present in higher amounts within resistant genotypes than susceptible genotypes.

1.8 The Tools: Genomic and Metabolomic Resources

In 2007, two draft genome sequences for *V. vinifera* were completed with annotations (Jaillon et al. 2007, Velasco et al. 2007). Building on these resources,

a single nucleotide polymorphism (SNP) genotyping project across thousands of accessions in the National Plant Germplasm System's (NGPS) *Vitis* species collection was completed by the United States Department of Agriculture-Agricultural Research Station (USDA-ARS) using a custom Illumina Infinium Genotyping Array for 6500 reliable SNPs (Myles et al., 2010). In September 2011, a 5-year multi-institution project, *Vitis*Gen (https://www.vitisgen.org/) was launched to accelerate grape cultivar improvement via next-generation sequencing (NGS) technology. Genotyping-by-sequencing (GBS), a next generation sequencing (NGS)-based method that enables simultaneous marker discovery and genotyping, has been implemented in grapevine to identify up to 50,000 SNP markers in each mapping population (Elshire et al., 2011).

While a genome sequence provides key information for prediction of gene functions, the use of genomics information alone is often insufficient to explain gene functions, which remain to be experimentally determined. Since a metabolome (i.e., a complete set of metabolites in an organism) closely correlates with the phenotypes of cells, analyses of the metabolome provide insights into understanding an organism's response to a perturbation. As a rapidly growing field of post-genomics research, metabolomics is a state-of-the-art analytical field that offers comprehensive investigations of metabolite variations in biological systems. It encompasses global metabolic profiling and multivariate statistical analysis, thereby providing quantitative and qualitative analyses of metabolites that depend on plant physiological conditions (Allwood et al., 2008; Dettmer et al., 2007).

The use of metabolomics tools for understanding plant response at the systems biology level has been rather limited in grapevine, especially when compared to other plant species such as *Arabidopsis thaliana*, potato, rice, and tomato (Carreno-Quintero et al., 2012; Keurentjes et al., 2006; Lisec et al., 2008; Matsuda et al., 2012; Schauer et al., 2006). Within grapevine, metabolomics has been mainly applied on grape and wine polyphenols, chiefly due to the importance of their sensory characteristics, as well as the health benefits of wine consumption (Flamini, 2013). Stilbene compounds, belonging to a subclass of polyphenols, have also received considerable interest for metabolic profiling, as these compounds have been implicated with disease resistance (Flamini et al., 2013; Mattivi et al., 2011; Pawlus et al., 2013; Vrhovsek et al., 2012). Currently, a database called *Grape Metabolomics* is under construction. This library contains information of grape metabolites that have been reported in the literature and found in electronic databases. To date, the library contains approximately 1000 putative compounds with molecular weights ranging from 100 to 1700 Da (Flamini, 2013; Flamini et al., 2013).

Figures

Figure 1.1 3,5,4′-trihydroxy-*trans*-stilbene, also known as *trans*-resveratrol, is the monomeric backbone unit of stilbenes that is catalyzed by the stilbene synthase (STS) enzyme via condensation of *p*-coumaroyl-CoA and three malonyl-CoA.

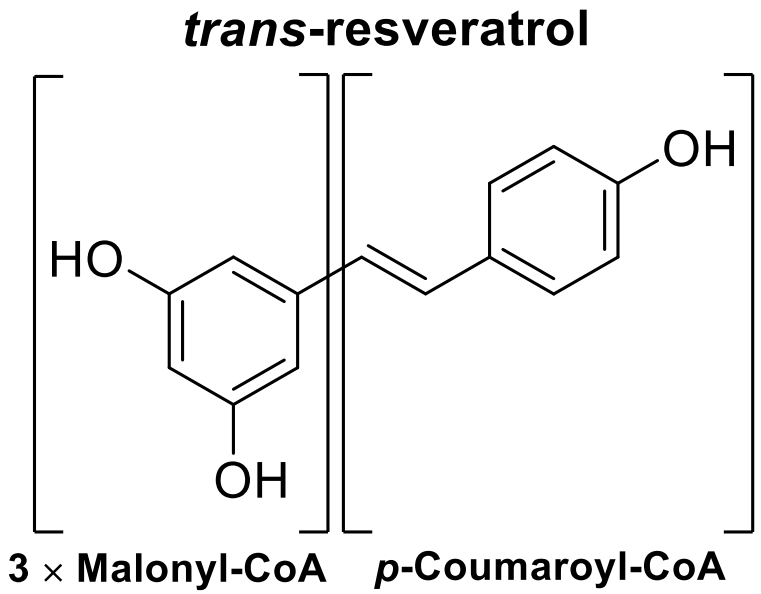
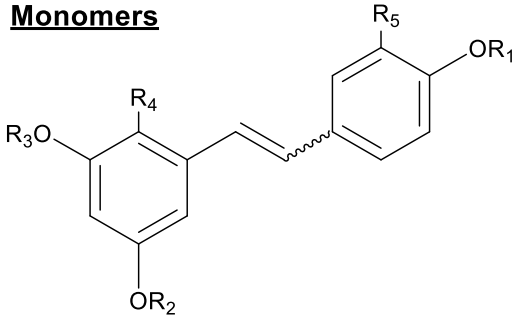


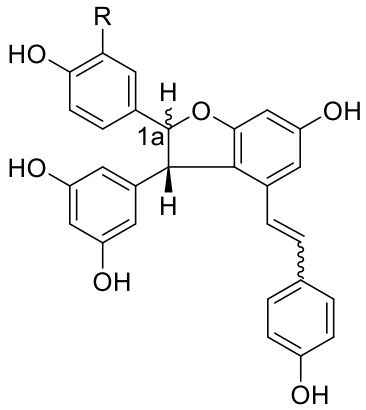
Figure 1.2 A subset of stilbene compounds ranging in structural complexity and oligomerization that has been identified within the genus *Vitis*.

Monomers

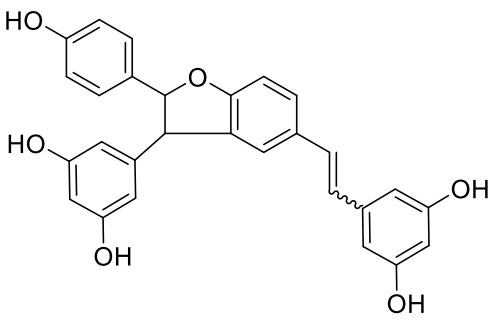


***E*-/*Z*-Resveratrol**: R_1, R_2, R_3, R_4, R_5 = H
***E*-/*Z*-Piceid**: R_2 = Glc; R_1, R_3, R_4, R_5 = H
***E*-Piceatannol**: R_1, R_2, R_3, R_4 = H; R_5 = OH
***E*-/*Z*-Resveratroloside**: R_1 = Glc; R_2, R_3, R_4, R_5 = H
***E*-/*Z*-Astringin**: R_2 = Glc; R_1, R_3, R_4 = H; R_5 = OH
***E*-Pterostilbene**: R_1, R_4, R_5 = H; R_2, R_3 = CH_3

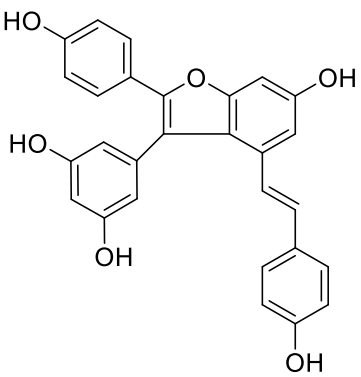
Dimers



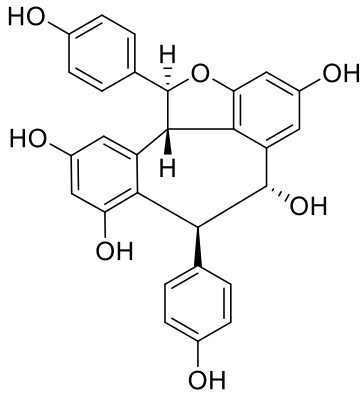
***E*-/*Z*-ε-Viniferin**: 1a = α; R = H
***E*-/*Z*-ω-Viniferin**: 1a = β; R = H
Scirpusin A: 1a = α; R = OH



***E*-/*Z*-δ-Viniferin**

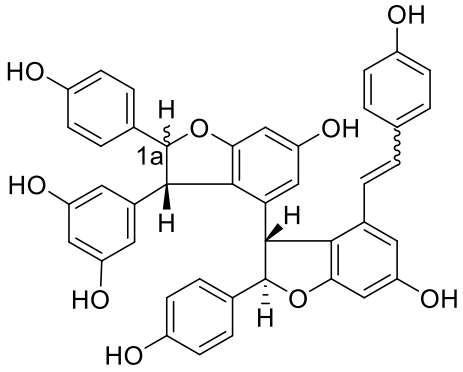


Viniferifuran

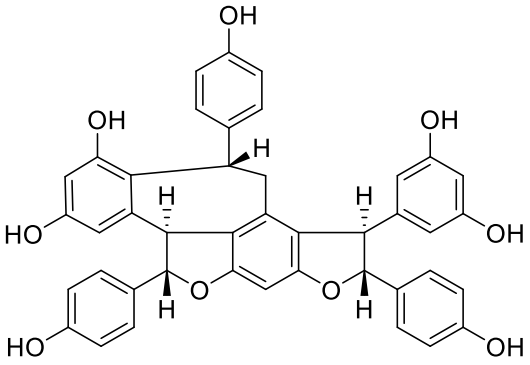


Ampelopsin A

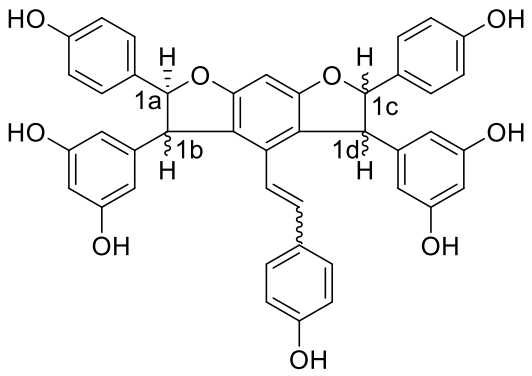
Trimers



***E-trans*-Miyabenol C**: 1a = α
***Z-trans*-Miyabenol C**: 1a = α
***E-cis*-Miyabenol C**: 1a = β

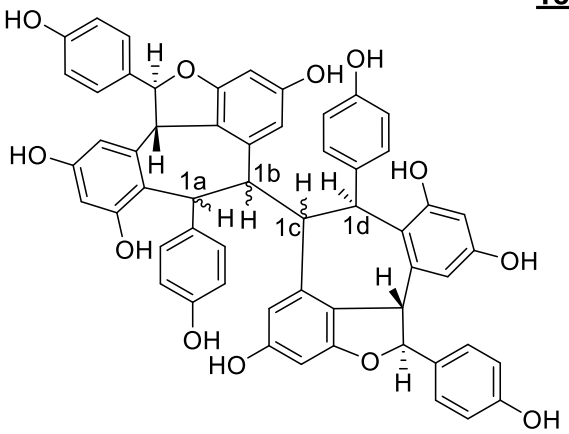


Vitisin E

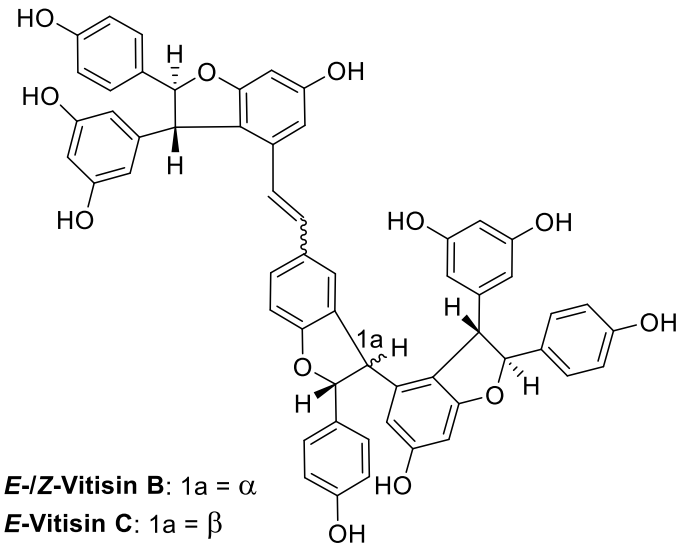


***E*-/*Z*-Ampelopsin E**: 1b, 1d = β; 1c = α
***E*-/*Z*-Amurensin B**: 1b, 1d = α; 1c = β
Amurensin C: 1a, 1b = dehydro; 1c, 1d = β
Amurensin D: 1a, 1b = dehydro; 1c = β; 1d = α

Tetramers



Hopeaphenol: 1a, 1b, 1c, 1d = α
Isohopeaphenol: 1a, 1b, 1c, 1d = β



***E*-/*Z*-Vitisin B**: 1a = α
***E*-Vitisin C**: 1a = β

Chapter 2

Genetic Dissection of Powdery Mildew Resistance in Interspecific Half-Sib Grapevine Families Using SNP-based Maps

Quantitative trait locus (QTL) identification in perennial fruit crops is impeded largely by their lengthy generation time, resulting in costly and labor-intensive maintenance of breeding programs. In a grapevine (genus *Vitis*) breeding program, although experimental families are typically unreplicated, the genetic backgrounds may contain similar progenitors previously selected due to their contribution of favorable alleles. In this study, we investigated the utility of joint QTL identification provided by analyzing half-sib families. The genetic control of powdery mildew was studied using two half-sib F_1 families, namely GE0711/1009 (MN1264 × MN1214; *N* = 147) and GE1025 (MN1264 × MN1246; *N* = 125) with multiple species in their ancestry. Maternal genetic maps consisting of 1,077 and 1,641 single nucleotide polymorphism (SNP) markers, respectively, were constructed using a pseudo-testcross strategy. Ratings of field resistance to powdery mildew were obtained based on whole-plant evaluation of disease severity. This two-year analysis uncovered two QTLs that were validated on a consensus map in these half-sib families with improved precision relative to the parental maps. Examination of haplotype combinations based on the two QTL regions identified strong association of haplotypes inherited from 'Seyval blanc', through MN1264, with powdery mildew resistance. This investigation also encompassed the use of SSR markers to establish a correlation between 206-bp (UDV-015b) and 357-bp (VViv67) fragment sizes with resistance-carrying

haplotypes. Our work is one of the first reports in grapevine demonstrating the use of SNP-based maps and haplotypes for QTL identification and tagging of powdery mildew resistance in half-sib families.

2.1 Introduction

Powdery mildew, caused by the biotrophic ascomycete *Erysiphe necator*, is a serious fungal disease of grapevine (*Vitis* spp.) worldwide (Gadoury et al. 2012). The fungus was inadvertently introduced to Europe through importation of grape cuttings from North America during the mid-nineteenth century. Subsequently, it was rapidly spread to other regions where grapevines were cultivated. Cultivated Eurasian grapevine (*Vitis vinifera* L.) is susceptible to the fungus, which can infect all green tissues as well as inflorescences and immature fruit, resulting in yield losses due to reduced photosynthetic leaf area, poor fruit set, premature fruit drop, and diseased berries (Pearson, 1988).

Control of powdery mildew typically relies on frequent application of fungicides, which is becoming increasingly prohibitive due to their cost and adverse effects on human health and the environment. With modern grapevine cultivation moving toward sustainable practices, grape breeding for disease resistance offers potentially effective and environmentally friendly control of powdery mildew. Marker-assisted breeding can expedite cultivar improvement for disease resistance by predicting the presence of resistance alleles when selecting parents or offspring.

Natural sources of disease resistance are typically found in regions where the pathogen and the host plant populations co-evolved. In particular, grapevine germplasm native to the North American and East Asian regions have been evaluated by breeders for sources of powdery mildew resistance. Thus far, the genetic research on powdery mildew resistance has elucidated QTLs from several genetic backgrounds, including *V. vinifera*, Asian *Vitis* species, and hybrids of North American *Vitis* species, as well as *Muscadinia rotundifolia* (Michx.) Small, also known as *Vitis rotundifolia* Michx. (Barker et al. 2005; Blanc et al. 2012; Dalbó et al. 2001; Hoffmann et al. 2008; Pap et al. 2016; Ramming et al. 2011; Riaz et al. 2011; Welter et al. 2007).

Genetic mapping in highly heterozygous, outbreeding crops like grapevine typically requires the use of a mapping strategy termed “pseudo-testcross” for construction of linkage maps (Grattapaglia and Sederoff 1994). In a conventional testcross where a heterozygous individual is crossed with an inbred that is homozygous at all loci, subsequent segregation can be attributed to the heterozygous parent. Meanwhile, a pseudo-testcross strategy implements crossing of two highly heterozygous individuals to identify markers that are fully informative with respect to the segregation attributed to the parent of interest. Despite being restricted by the number of usable markers with the approach, the inclusion of multiple mapping families provides an opportunity to harness markers heterozygous for both parents of a full-sib family for the eventual construction of a multiple-family consensus linkage map.

The advancement of whole genome sequencing and next-generation sequencing has propelled crop genomic and genetic research. In particular, genotyping-by-sequencing (GBS) offers an inexpensive and robust solution for simultaneous SNP discovery and genotyping through pooled, barcoded, reduced representation libraries, sequencing, and SNP assignment based on alignment of short reads (Elshire et al. 2011). The results are thousands of low-coverage markers that are usable for QTL mapping in bi-parental families (Davey et al. 2011). The work by Barba et al. (2014) demonstrated the first application of the GBS procedure in generating SNP markers to construct a high-resolution map for QTL mapping in grapevine.

Breeding programs in cold-climate regions typically utilize recurrent selection to improve traits, such as maturity, winter hardiness, fruit quality, and disease resistance. In perennial fruit crops such as grapevine, funding, labor, and space constraints due to plants' large size and long generation time may hamper such efforts. In this breeding method, clonal propagation and the long life of grapevines allows breeders to use desirable breeding parents to contribute their favorable alleles in multiple diverse crosses. Hence, the availability of multiple experimental families having pedigree-connected parents provides an opportunity to explore the genetic control of traits using a pseudo-testcross approach that combines information from parental meioses over multiple families. Genetic mapping studies using related existing families could be a useful tool in grapevine for (1) identifying reliable QTL, and (2) developing markers that are relevant to the breeding germplasm.

Hence, the main objectives of this investigation were to: (1) build SNP-based linkage maps for interspecific experimental *Vitis* families using pseudo-testcross strategy; (2) construct a consensus linkage map using markers commonly present in the maternal (MN1264) maps of these half-sib families; (3) identify and validate QTL for powdery mildew resistance in both parental and consensus maps; (4) develop SNP-based haplotypes that are associated with resistance; and (5) identify linked SSR markers that may be useful for future crossing or culling decisions in a breeding program.

2.2 Materials and Methods

2.2.1 Plant Material

Two bi-parental F_1 families were used in this study (Figure 2.1). A cross of MN1264 × MN1214 made in 2007 (GE0711; *N* = 53) and was repeated in 2010 (GE1009; *N* = 94). These crosses were treated as a single mapping family, GE0711/1009, for this experiment. The second mapping family, GE1025 was a cross of MN1264 × MN1246 made in 2010 (*N* = 125). Comprised of at least six *Vitis* species in their ancestry, including *V. vinifera*, *V. riparia*, *V. rupestris*, *V. labrusca*, *V. aestivalis*, and *V. berlandieri*, these F_1 families were grown at the University of Minnesota Horticultural Research Center (44°52'08.1"N 93°38'17.3"W). They were chosen as mapping populations for the USDA-funded *Vitis*Gen project because their pedigrees are connected and contain many key ancestors contributing winter hardiness and fruit quality to the University of Minnesota wine grape breeding program.

2.2.2 Quantification of Field Powdery Mildew Severity

Throughout the two-year investigation, no fungicide applications were used. Disease severity on each vine was evaluated toward the end of each growing season. In late August of 2014, each vine was assessed using a 7-point visual scale of whole-plant foliage (1: absent, 3: sporulation covers up to a third of the foliage, 5: sporulation covers up to two thirds of the canopy, 7: vast and dense sporulation covers over two thirds of the canopy; even numbers denote intermediate symptoms). In 2015, the leaves of each vine were evaluated using the 9-point International Plant Genetic Resources Institute (IPGRI) scale that was established by the Organisation Internationale de la Vigne et du Vin (IPGRI et al. 1997) on August 27 (1: absent, 3: low sporulation with limited patches < 2 cm in diameter, 5: moderate sporulation with patches of 2 – 5 cm in diameter, 7: vast sporulation and abundant mycelium, 9: extreme sporulation with unlimited patches of sporulation; even numbers denote intermediate symptoms).

In GE1025, field phenotype data were collected throughout the growing season after the first observation of powdery mildew. Plants were evaluated every seven to ten days until the disease severity of the mapping family reached a plateau. Observation of disease symptoms over the multiple time points allowed the construction of a disease progress curve (DPC) for QTL mapping to determine if area under disease progress curve (AUDPC) provided additional utility and resolution in QTL detection. In 2014, GE1025 plants were visually evaluated on July 24, July 31, August 14, August 26, September 5, September 12, September 19, and September 26. In 2015, field ratings were collected on

June 19, June 26, July 2, July 10, July 17, July 23, July 30, August 6, August 13, and August 20 (Figure 2.5).

2.2.3 DNA Extraction, Sequencing, and SNP Assignment

Methods from tissue sampling through SNP assignment were as described previously (Hyma et al. 2015). For each vine, a single small, newly expanding leaf (less than 1 cm diameter) was collected in one tube of a Costar 96-well cluster tube collection plate (Corning Life Sciences, Tewksbury, MA, USA). Leaf tissues were maintained at 4 °C from harvest until processing in the laboratory where two stainless steel genogrinder beads were placed in each tube and plates were frozen at -80 °C. Tissue grinding took place in a Geno/Grinder 2000 (OPS Diagnostics LLC, Lebanon NJ, USA) with 96-well plates agitated in pairs at 400× speed for 1 minute. Plates were then stored at –80 °C until processing with DNeasy 96-well DNA extraction kits (Qiagen, Valencia CA, USA). Modifications were made to the manufacturer's protocol to improve DNA quality and quantity as follows: 1) PVP-40 (2% w/v) was added to the AP1 lysis buffer prior to heating; and 2) visual inspection for complete re-suspension of the sample pellet of each 8-tube strip was added to the agitation step following AP1 addition.

GBS was performed as described by Elshire et al. (2011), integrating four 96-well plates across 384 barcodes for library preparation (Hyma et al. 2015) and sequencing of single-end 100-bp sequences using a HiSeq 2000 (Illumina Inc., San Diego, CA, USA). Raw data for the University of Minnesota germplasm (henceforth, UMN dataset) were processed based on an updated version of the

*Vitis*Gen databased previously employed by Hyma et al. (2015). Comprised of 409 maternal half-sib individuals and four progenitors, the UMN dataset was processed through TASSEL 3.0.139 GBS pipeline (Glaubitz et al. 2014) using the 12x.2 *V. vinifera* 'PN40024' reference genome (Adam-Blondon et al. 2011; Jaillon et al. 2007) from The French-Italian Public Consortium (https://urgi.versailles.inra.fr/Species/Vitis/Data-Sequences/Genome-sequences). For alignment, Burrows–Wheeler Aligner maximal exact match (BWA– MEM) was applied using with default parameters (Li and Durbin, 2009). The output consisted of variant call format (VCF) file version 4.1 (Danecek et al. 2011) including SNPs present in at least 40% of the progeny, and with a minor allele frequency (MAF) ≥ 0.1. Subsequently, the VCF file was filtered using vcftools ver. 1.12a (Danecek et al. 2011) and TASSEL (Bradbury et al. 2007) versions 3.0.139 and 4.3.13. Finally, the VCF file was sliced to include SNP data of the 277 F_1 individuals (in mapping families of GE0711/1009 and GE1025) and the three progenitors (MN1264, MN1214, and MN1246) in this study.

2.2.4 Linkage Maps

From the putative markers identified for GE0711/1009, GE1025 and the progenitors, a custom filtering process for alignment was applied, with minimum read depth of 6 and 85% completeness by site across progeny, and by progeny across sites. Subsequently, an additional filter for MAF between 0.05 and 0.35 was applied through vcftools to optimize the identification of markers for parental map construction. Results were output as a TASSEL hapmap file. Pseudo-testcross markers with segregation of 1:1 (heterozygous:homozygous) were

identified by a chi-squared (χ^2) goodness-of-fit test at $\alpha < 0.01$, coded and imported into JoinMap 4.1 (Van Ooijen 2006).

Using a two-way pseudo-testcross approach (Grattapaglia and Sederoff, 1994) and cross-pollination (CP) population in JoinMap, marker types of *lm×ll* and *nn×np* were retained for construction of maternal and paternal maps, respectively. Each parental linkage map was constructed separately. As an additional level of stringency, markers with over 10% missing data were discarded. Linkage groups (LGs) were constructed with a minimum threshold logarithm of odds (LOD) score of 6.0. LGs were grouped and numbered based on their corresponding physical chromosome numbers (Jaillon et al. 2007). Genetic maps and marker order were determined using the regression mapping algorithm with Haldane function and default parameters. Initial maps obtained in JoinMap were further analyzed in R/qtl software (Broman et al. 2003) using a four-way cross format. Markers were assessed for segregation distortion using geno.table command at a *p*-value of 1×10^{-7}, above which markers were discarded. Finally, map distance and marker order were recalculated in JoinMap.

2.2.5 Construction of a Consensus Genetic Map

Development of the maternal (MN1264) consensus map was conducted using LPmerge (Endelman and Plomion 2014). Contrary to the approach of minimizing an objective function based on the observed recombination frequencies between markers, this software employs a linear programming algorithm to minimize the mean absolute error between the consensus map and the genetic maps from each family. This minimization step results in preservation

of marker order in linkage maps. At instances where marker order is inconsistent between maps, LPmerge implements an additional algorithm to resolve ordinal conflicts.

The LPmerge software contains two key parameters that are set by the user: (1) maximum interval size between bins, and (2) weights applicable to each map. In our analysis, the maximum interval parameter was set at $K = 8$, while weights were adjusted based on family sizes due to more progeny providing better resolution. At each LG, the best consensus map was selected using two criteria: (1) minimization of the average root-mean-squared error (RMSE), and (2) map length comparable to the mean of the linkage maps. Final images of LGs were generated using MapChart 2.2 (Voorrips 2002) before they were exported to Adobe Illustrator CS2 for alignment and vector transformation.

2.2.6 QTL Analysis

QTL detection was conducted using multiple statistical analyses and software: Kruskal-Wallis (KW), interval mapping (IM), and composite interval mapping (CIM). KW and IM were performed on MapQTL 6 software (Van Ooijen 2009), which implemented a maximum likelihood mixture model. Meanwhile, CIM was analyzed in R/qtl software (Broman et al. 2003) using default parameters. The minimum LOD score required for QTL detection was approximately three, which was determined by the genome-wide LOD significance threshold ($\alpha = 0.05$) calculated using 1,000 permutations.

2.2.7 Haplotype Construction and Analysis with *Ren3* SSR

Based on the maternal consensus map, 12 and 22 SNP markers spanning the QTL (described in Results) on LG 15 and LG 2, respectively, were used to build the grapevine powdery mildew resistance haplotypes. With the use of SNPs from a pseudo-testcross approach, there are two unique haplotypes for each QTL spanned, thereby yielding a total of four haplotype combinations for the two QTLs of interest. Given the availability of GBS information for the maternal grandparents (i.e. MN1069 and 'Seyval blanc') of the progeny, the haplotypes were traced based on inheritance and were correspondingly assigned.

Analysis of variance (ANOVA) calculations were made to determine if a haplotype combination was associated with disease severity in each year. Where statistical difference was detected, Tukey's honest significant difference (HSD) test ($P < 0.05$) was conducted as a post-hoc analysis to test differences among the average effects of each haplotype combination.

The identification of powdery mildew resistance QTL on LG 15 raised the possibility of co-localization of this QTL with the previously reported *Ren3* (Welter et al. 2007). Two SSR markers (UDV-015b and VViv67) for powdery mildew resistance that were associated with the *Ren3* resistance locus were used to screen the mapping families and the progenitors. According to Welter et al. (2007), these markers were reported to be 3 cM apart and are located in the vicinity of *Ren3*. The amplified fragment sizes were used to trace the inheritance of allele sizes in the known grandparent-parent-offspring relationships (Figure 2.1).

2.2.8 Haploblock Validation Using Controlled Inoculation

In addition to obtaining field phenotype data based upon natural powdery mildew infection, controlled inoculation of GE1025 mapping family was carried out using *Erysiphe necator* isolate NY19 in two independent experiments (Experiment A: *N* = 119; Experiment B: *N* = 106). Protocols for the maintenance of pathogen isolate, leaf sampling, inoculation, and microscopic evaluation at 9-day post infection (9 dpi) were described in Cadle-Davidson et al. (2016). At 9 dpi, leaf samples were evaluated based on several responses: (i) number of hyphae intercepting a transect through a 200× field of view, henceforth referred to as hyphal transects; (ii) cube root transformation of hyphal transects, henceforth annotated as TransectT; and (iii) incidence of powdery mildew sporulation. Using the different haplotype combinations that were constructed based on the QTL regions, haploblock validation for each phenotypic response was carried out to determine the effect of each haplotype. Statistical analyses were conducted in JMP 12.0.1 (SAS Institute Inc. Cary, NC, USA).

2.3 Results

2.3.1 Genotyping: Sequencing and SNP Assignment

The SNP assignment conducted in TASSEL 3.0.139 produced a total of 526,775 SNP markers in the UMN dataset. The markers were the result of successful alignment of 595,270 GBS tags from a total of 886,142 (67.18%), with no alignment to multiple positions being reported in this dataset. With respect to the mapping families and progenitors in this study, the total number

of SNPs from the UMN dataset was retrieved for subsequent filtering. Preliminary filtering (≥ 40% marker presence in the progeny; MAF ≥ 0.1) yielded 103,072 markers for GE0711/1009 and 82,324 markers for GE1025. For parental maps, 5,464 SNP markers suitable for pseudo-testcross analysis were identified for GE0711/1009, and 5,738 SNPs for GE1025.

2.3.2 Parental and Consensus Linkage Maps

Maternal (MN1264) genetic maps of GE0711/1009 and GE1025 were constructed using 1,077 and 1,641 SNPs each, covering genetic distances of 2,142 and 2,005 cM, respectively, and physical distances of approximately 426.2 Mb (Figure 2.7). All significant markers (LOD score ≥ 6.0) were included in JoinMap 4.1 for linkage map construction, omitting unlinked markers due to weak linkages to other markers within the LG (Tables 2.1 and 2.3). Information of the paternal maps (MN1214 and MN1246) is summarized in Table 2.3, as well as in Figures 2.8 and 2.9. The use of the LPmerge software to build a consensus map provided LGs with comparable genetic lengths. The consensus map was constructed using 1,977 SNPs, covering a total genetic distance of 1,852.9 cM with an average interval of 0.94 cM between markers (Table 2.1).

Generally, the marker order between the consensus map and the parental maps was conserved. In most intervals, markers in the consensus map were co-linear to those in the two parental maps. Peculiarly, a region of 13 markers spanning from 7.2 cM to 39.9 cM on LG 15 of the maternal map of GE0711/1009 was mapped onto just a single locus at 24.4 cM on the consensus map. Since these markers were not in common with any SNP on GE1025, this could be a

rare instance where LPmerge was unable to resolve the recombination events that originated from GE0711/1009.

2.3.3 Field Assessment of Powdery Mildew Severity

Disease phenotypes of progeny in the two experimental families spanned nearly the complete rating scale in both years (Figure 2.4), with positive correlation between the two years (Pearson's: R^2 = 0.60; Spearman's: R^2 = 0.60). In 2014, the mean scores were 4.1 in GE0711/1009 (N = 147) and 4.7 in GE1025 (N = 125) based on a 7-point visual scale. Meanwhile in 2015, the mean scores for the families were 4.7 and 5.3, respectively, based on the 9-point IPGRI scale.

In the construction of a DPC, field evaluation of GE1025 started over a month earlier in 2015 than in 2014. In general, the vast majority of individuals were rated between one and three during the first 21 days of disease assessment. Field ratings of GE1025 appeared to level off approximately 55 days after the initial assessment.

2.3.4 Analysis of Marker-Trait Association

QTL mapping was conducted using IM and CIM on each maternal map, and resulted in identification of QTL on LGs 2 and 15 (Table 2.2; Figure 2.6). In GE0711/1009, a moderate QTL on LG 2, and a major QTL on LG 15 were detected using 2014 ratings and supported by significant LOD scores at chromosome- and genome-wide thresholds (α = 0.05). Though a QTL on LG 2 was again detected in GE0711/1009 using 2015 ratings, no significant QTL was detected on LG 15. In GE1025, a QTL was identified in the same genetic region of LG 15 in both years and explained a similar proportion of phenotypic variance,

regardless of the analysis methods. In contrast to GE0711/1009, a QTL was not detected on LG 2 in either year. In the QTL analysis using the consensus map, a minor QTL on LG 2 was detected in 2015, but not in 2014. Meanwhile, results of QTL analysis for LG 15 were consistent across years, where a major QTL was mapped to the region between 12.0 and 24.0 cM (CIM) with peak LOD score of approximately 10. As a whole, the use of the more sensitive CIM analysis method identified QTL that generally peaked at a similar genetic position in both years, and reported a finer two-LOD confidence interval overlapping the interval from IM.

The use AUDPC as a phenotype for QTL mapping corroborated the QTL findings based on a single time-point phenotype (Table 2.2). In both years, AUDPC in GE1025 allowed the detection of a major QTL on LG 15, explaining up to 27.0% of phenotypic variance. Similar to the QTL results from a single time-point phenotype, AUDPC in GE1025 also did not detect a QTL on LG 2.

2.3.5 Haplotypes Associated with Powdery Mildew Resistance

Haplotypes were constructed using 12 SNPs (covering ~6.6 cM) and 22 SNPs (covering ~24.3 cM) that spanned the LG 2 and LG 15 QTL regions, respectively. Linkage phases of the haplotypes for the maternal parent (MN1264) were determined; there were no observable recombination events from the grandparent progenitors, 'Seyval blanc' and MN1069. Considering only the maternal grandparent origin for LGs 2 and 15, four haplotype combinations are possible, namely $LG02_{MN1069}$: $LG15_{MN1069}$, $LG02_{Seyval\ blanc}$: $LG15_{MN1069}$, $LG02_{MN1069}$: $LG15_{Seyval\ blanc}$, and $LG02_{Seyval\ blanc}$: $LG15_{Seyval\ blanc}$.

Analyses of the four haplotype combinations revealed three statistically different classes for disease severity (Figure 2.2a; Table 2.4). Progeny with MN1069-inherited haplotypes consistently exhibited the highest mean disease severity score in both years. Meanwhile, genotypes with $LG02_{Seyval\ blanc}$: $LG15_{MN1069}$ had intermediate mean score between those with MN1069-inherited haplotypes and those with $LG02_{MN1069}$: $LG15_{Seyval\ blanc}$. Thus, phenotypic correlation with haplotype combinations for the two chromosomal segments was consistent in both years, where vines with 'Seyval blanc'-inherited haplotypes showed significantly lowest disease severity. Therefore, the evidence suggests the identification of a powdery mildew resistance in grape, henceforth called *Ren10*.

Screening each bi-parental mapping family with two SSR markers (UDV-015b and VViv67) with previously demonstrated linkage to *Ren3* (Welter et al. 2007) provided fragment sizes that were associated with powdery mildew resistance (Figure 2.3). In the assessment of GE1025 (*N* = 121) with marker UDV-015b, PCR products of MN1264 yielded fragment sizes of 184 / 206 while MN1246 (male parent) produced fragment sizes of 187 / 200, thereby resulting in four combinations of allele sizes in the offspring (Figure 2.1). ANOVA and Tukey's HSD test indicated that individuals containing the 206-bp fragment exhibited significantly reduced disease severity in both years (Figure 2.3). In GE0711/1009, meanwhile, MN1214 (male parent) produced fragment sizes of 190 / 200, which also resulted in four combinations of allele sizes in the offspring (*N* = 125). Though a similar trend was observed in both years, ANOVA and

Tukey's HSD test showed statistically significant association of the 206-bp fragment with resistance only when it was paired with the MN1214-derived 190-bp fragment.

Meanwhile, PCR products of VViv67 in GE1025 (*N* = 116) yielded fragment sizes of 357 / 379 in MN1264 and homozygous 384 / 384 in MN1246, providing two unique combinations of allele sizes in the offspring (Figure 2.1). In both years, ANOVA and Tukey's HSD test demonstrated an association between the 357-bp fragment and powdery mildew resistance. Meanwhile, PCR products of MN1214 (male parent) reported fragment sizes of 383 / 390, resulting in four combinations of allele sizes in the offspring of GE0711/1009 (*N* = 127). However, ANOVA calculations revealed statistically significant association of the 357-bp fragment with resistance only in year 2014, but not in 2015.

2.3.6 Haploblock Validation Using Controlled Inoculation

The phenotypic responses from two independent controlled inoculation experiments were used to compare the four haplotype combinations. Based on the TransectT results in Experiment A, offspring with the 'Seyval blanc'-inherited *Ren10* haplotype exhibited a statistically lower hyphal proliferation than offspring with either the 'MN1069'-inherited *Ren10* or the 'Seyval blanc'-inherited *Ren3* haplotype (Figure 2.2b; Table 2.5b).

The results from Hyphal Transects experiments (Table 2.5a) were comparable to TransectT, although heteroscedasticity analyses reported by Cadle-Davidson et al. (2016) indicated the need for a cube root transformation for a more uniform distribution of variance. In the case of sporulation, individuals

with the 'Seyval blanc'-inherited *Ren10* haplotype had a statistically lower incidence of sporulation than offspring with either the 'MN1069'-inherited *Ren10* or the 'Seyval blanc'-inherited *Ren3* haplotype (Table 2.5c). This corroborated the results of Hyphal Transects and TransectT.

Meanwhile, polymorphism for LG 15 haplotype showed no significant effect on either colonization (*p*-value = 0.1459) or sporulation (*p*-value = 0.4607). Results in Experiment B appeared to have a systematic shift (*p*-value < 1×10^{-4}), where values in all haplotype combinations were shifted upward by approximately 0.7 – 0.9 units. While significantly more disease was observed across the second experiment, the allele effect was consistent over both experiments, as shown for TransectT (Figure 2.2b), and all experiment × haplotype interaction terms were non-significant at $\alpha = 0.1$ (Table 2.5).

2.4 Discussion

2.4.1 Genotyping: Sequencing and SNP Assignment

The percentage of GBS tag alignment (~67%) and the number of putative SNP markers yielded (521,405 SNPs) for the 436 individuals considered in this study were similar to the reports of other families that were included in the *Vitis*Gen project and analyzed by Hyma et al. (2015). Although our investigation utilized a corrected version of the *V. vinifera* reference genome PN40024 12x.v2 (Adam-Blondon et al. 2011, Jaillon et al. 2007), several reads did not align likely due to divergent species-specific sequences, low complexity sequences and poor quality, as previously suggested by Barba et al. (2014) and Hyma et al.

(2015). The confidence in assigning heterozygote marker genotypes was dependent on the depth of sequence coverage (Hyma et al. 2015), a known issue happening with GBS while analyzing heterozygous individuals that affects our ability to obtain reliable pseudo-testcross markers for linkage map construction. In a bi-parental analysis of families with complex genetic backgrounds (as in our case, which includes at least six *Vitis* species), the optimal combination of GBS multiplexing, read depth, and the number of SNPs has to be assessed beforehand, as suggested for another heterozygous perennial species, *Populus trichocarpa* (Schilling et al. 2014).

2.4.2 Parental and Consensus Maps

In a cold-climate region, a grapevine breeding program may employ recurrent selection to select for winter hardiness and maturity time, as well as other important fruit quality traits. Due to the contribution of different trait alleles by various *Vitis* species, this selection method often results in experimental families with ancestry comprised of numerous *Vitis* species. To date, many grapevine genetic maps that have been published were based either on *V. vinifera* intra-specific crosses (Adam-Blondon et al. 2004; Doligez et al. 2002; Doligez et al. 2006; Riaz et al. 2004; Troggio et al. 2007), or hybrids of interspecific crosses involving up to three *Vitis* species (Barba et al. 2014; Bellin et al. 2009; Blasi et al. 2011; Dalbó et al. 2000; Di Gaspero et al. 2007; Doucleff et al. 2004; Fischer et al. 2004; Grando et al. 2003; Lodhi et al. 1995; Lowe and Walker 2006; Marguerit et al. 2009; Moreira et al. 2010; Salmaso et al. 2008;

Welter et al. 2007). This paper reports a grapevine genetic map based on hybrids of crosses between parents derived from at least six *Vitis* species.

The high heterozygosity in grapevine necessitates the use of a pseudo-testcross strategy for F_1 map construction, which builds each parental map independently. Comparisons of the two maternal genetic maps and the consensus map (Figure 2.7) revealed general similarity in map length, marker order and LG sizes, except for minor discrepancies, namely putative inversion segments at a few regions, particularly in LG 15 and LG16. Plausible explanations for these discrepancies include the use of relatively small family sizes and technical challenges of using GBS for genotypes with multiple species backgrounds. In addition, the region with several markers in inverted order may be attributed to the presence of a large cluster of NBS-LRR genes comprising *Ren3* (Welter et al. 2007).

In the maternal maps, a region on chromosome 15 exhibited significant segregation distortion, an observation previously reported by Lowe and Walker (2006), Welter et al. (2007), Riaz et al. (2011), Heerden et al. (2014) and Guo et al. (2014). In contrast, the paternal maps (Figures 2.8 and 2.9) had more prominent challenges, perhaps due to decreased sampling of parental meioses. In particular, the genetic map of MN1246 included two short unconnected fragments of LG 15, as well as the absence of LG 17. Meanwhile, the map of MN1214 contained several sizeable gaps, particularly in LG 1 and LG 8. These limitations could be due to biological reasons (e.g. lack of polymorphism due to high levels of homozygosity, or the presence of lethal alleles which results in a

lethal phenotype in homozygous condition) and/or technical reasons (e.g. sequencing quality, unsuitable reference genome for these families, missing data, or stringent calls).

2.4.3 Field Assessment of Powdery Mildew Severity

In a cold-climate region where breeding populations are potentially subjected to low temperature injury each year, winter injury to a vine or differences in overwintering success of the pathogen may affect the field evaluation of disease severity in the following growing season as well as weather conditions during the growing season. Here, resistance to powdery mildew displayed a relatively continuous variation in the two experimental F_1 families in both years, suggesting inheritance as a quantitative trait (Figure 2.4). However, in 2014 the distribution appeared distinctly bimodal, suggesting the presence of loci with major effects. Despite the year-to-year changes in the severity of symptoms and rating scales used, we reported consistent and reproducible QTL results.

2.4.4 QTL Detection and Validation

Pedigree-connected multiparental families can be comprised of several bi-parental families that are connected by a common parent. In comparison to single bi-parental family analyses, the use of joint analyses has proved to optimize QTL detection in carrot (Le Clerc et al. 2015), maize (Blanc et al. 2006; Steinhoff et al. 2011), oil palm (Billotte et al. 2010), perennial ryegrass (Pauly et al. 2012), and sugar beet (Schwegler et al. 2013). While a single bi-parental family takes into account only two alleles from each parent and allelic effects are

specific to the family, connected multiparental families can provide estimates for the allelic effects of different parental lines. As evidenced from our investigation, the joint analyses provided validation and higher precision than separate analyses for QTL locations, resulting in finer support intervals shown in both IM and CIM analyses.

In this study, we reported the identification of a major QTL on LG 15 and a moderate QTL on LG 2. A QTL named *Ren3* was previously identified on LG 15 (Welter et al. 2007), but to our knowledge, a QTL for powdery mildew resistance has not previously been reported on LG 2, and thus we name this resistance locus, Resistance to *Erysiphe necator 10* (*Ren10*). Justification for naming this locus is also supported by haploblock validation, which is described in the ensuing section. The QTLs on LGs 15 and 2 explained up to 29.4% and 13.3% of the total phenotypic variance, respectively. The relatively modest percentages of variance explained may be due to the experimental design of using a single unreplicated vine of each individual, as well as uneven distribution of inoculum, varying environmental conditions in the field, and other QTL not detected in our experimental design. Despite the detection of a major QTL on LG 15, this QTL was not detected in GE0711/1009 in year 2015. This could be attributed to the vine age and smaller sample size of GE0711.

Our QTL analyses also included the use of AUDPC as a phenotype. The collection of field ratings over multiple time points throughout the growing season was hypothesized to enhance QTL detection, since AUDPC would likely capture more phenotypic variance than that of a single time-point rating. However, the

QTL results in both years identified the same QTL as a single late-season rating, without providing additional utility or resolution. Therefore, AUDPC corroborated the single time-point ratings. However, for the conditions and goals of the breeding program addressed here, the use of AUDPC does not seem profitable due to the additional time and efforts incurred, and it did not provide further information about the genetic control of powdery mildew resistance.

Despite GE0711/1009 and GE1025 sharing a common maternal parent, LG 2 QTL was detected in the former, but not in the latter, which was likely attributed to the allelic contribution of the different male parents. The male parents of GE0711/1009 and GE1025 (i.e. MN1214 and MN1246, respectively) are half-sibs, sharing 'Frontenac' as a common parent. However, the non-common parents of MN1214 and MN1246 are *V. vinifera* cv. Cabernet Sauvignon and *Vitis* hybrid MN1200, respectively, where the former is highly susceptible to powdery mildew, while the latter is moderately resistant to the disease. Thus, our ability to detect the maternal QTL on LG 2 may be due to the enrichment for paternally-derived susceptible alleles in GE0711/1009 from the grandparent ‘Cabernet Sauvignon’.

Coincidentally, chromosomes 2, 15, and 16 belong to the same cluster of paralogs that demonstrated the contribution of three ancestral genomes to the present-day grapevine haploid genome (Jaillon et al. 2007). Thus, the two QTLs on LGs 2 and 15 may have arisen from either a hexaploidization event, or through two successive genome duplications. Subsequent to the polyploidization event, nonrandom loss of genomic content likely led to preferential retention of

genic regions that were associated with transcription factors, as well as expressions of novel morphologies and pathogen resistance (Doyle et al. 2008).

2.4.5 Haplotypes Associated with Powdery Mildew Resistance

Development of SNP-based haplotypes that are associated with QTL alleles is useful for quantifying their effects, and may provide DNA information for marker-assisted selection of parents or seedlings. Our analysis identified the desirable genetic contribution owing to a specific grandparent, 'Seyval blanc'. Based on the two QTLs, four unique maternal haplotype combinations were constructed spanning ~6.6 cM on LG 2 and ~24.3 cM on LG 15 using SNPs from the maternal maps. In two years of observations, the presence of 'Seyval blanc'-inherited haplotypes was significantly associated with decreased disease severity, while offspring with MN1069-inherited haplotypes exhibited significantly higher disease scores. This analysis reaffirmed the relatively larger effect of the QTL on LG 15 compared to the QTL on LG 2.

Meanwhile, laboratory controlled inoculation of the GE1025 mapping families provided strong support for the discovery of the QTL on LG 2. In two independent experiments measuring three phenotypic responses, individuals with the 'Seyval blanc'-inherited *Ren10* haplotype consistently exhibited reduced colonization and decreased sporulation. The use of a single isolate (*Erysiphe necator* NY19) that originated from New York *V. vinifera* vineyards (Cadle-Davidson, personal communication) allowed the detection of the LG 2 QTL, which was more difficult to detect in the field due to the larger effect of the LG15 QTL, and/or the genetics of the pathogen population in the natural field

environment. *Ren3* resistance is known to be ineffective in New York (Cadle-Davidson and Reisch, personal communication), and thus it is not surprising that *Ren3* was not detected using the NY19 isolate.

2.4.6 Co-localization with *Ren3*?

The detection of a major QTL on LG 15 raised a question of whether this QTL co-localizes with the previously-reported *Ren3* locus (Welter et al. 2007). Although we are provisionally designating it as *Ren3*, additional experiments such as complementation tests using both sources of resistance in this region will have to be conducted to address this question.

We screened our mapping families with two SSR markers, namely UDV-015b and VViv67 that were reportedly linked with *Ren3* (Welter et al. 2007) to establish a correlation between specific fragment sizes and disease severity in our families. The 206-bp fragment from UDV-015b and the 357-bp fragment from VViv67 were significantly associated with reduced mean powdery mildew severity. Though these correlations were consistently significant in GE1025 in both years, the resistant alleles in GE0711/1009 trended toward the expected effects in all cases, except the lack of association between fragment sizes and disease severity in 2015. Given the vine age of GE0711 mentioned above, and the use of field phenotypes, variance due to environment was likely larger than variance attributed to genetics. As a whole, our results have shown that the UDV-015b and VViv67 markers that were developed for the *Ren3* locus appear linked to resistance inherited from MN1264, and were ultimately derived from 'Seyval blanc'.

2.4.7 Chromosome 2: *Ren10* and Potential Structural Dynamics in the Vicinity

The identification of the *Ren10* powdery mildew resistance locus on chromosome 2 is intriguing, given the proximal location of several *Myb*-like genes controlling berry color, at a physical position of approximately 14.2 Mbp in the *Vitis vinifera* reference genome (Jaillon et al. 2007; Velasco et al. 2007; Walker et al. 2006). This region has been characterized by movement of the *Gret1* retrotransposon and by hemizygosity in some genotypes due to chromosomal replacement and/or deletion (Kobayashi et al. 2004; Kobayashi et al. 2005; Lijavetzky et al. 2006; Migliaro et al. 2014; Pelsy et al. 2015; Walker et al. 2007).

Here, the putative physical position of *Ren10*, which was likely introgressed from a North American species, resides in the QTL window between 8.5 and 18.7 Mbp (Table 2.4). In the GE1025 mapping family, a berry color locus was mapped to a physical position from 10.6 to 17.6 Mbp (Clark et al. 2016). To test whether the *Ren10* QTL spans structural variations related to a berry color locus, we re-sequenced MN1264 and MN1246, the parents of GE1025. The results indicate several regions of hemizygosity in this region (Figure 2.10). In particular, the physical position 10.5-12.5 Mbp is single copy in MN1246, suggesting hemizygosity, and is completely absent in MN1264, likely due to a homozygous deletion. This preliminary analysis is solely based on alignment to the *V. vinifera* reference genome, and thus additional genomic characterizations need to be conducted.

2.5 Conclusions

At the University of Minnesota grape breeding program, MN1264 is a quality hybrid wine grape that is unfortunately not suited for production due to the production of only female flowers. However, MN1264 has various favorable traits, such as winter hardiness, fungal resistance, and pest resistance, thus making it an ideal crossing parent for genetic studies within the context of a breeding program (Luby, personal communication). In an effort to study the genetic determinism of powdery mildew resistance, MN1264 was used to construct two segregating mapping families, which are described in this study.

Here, we present the first application of joint-family SNP-based analysis in grapevine to construct a consensus map for QTL identification and validation, the development of haplotypes, as well as the report of resistance-linked fragment sizes from SSR markers. First, the use of GBS provided over 10,000 uniformly distributed genome-wide SNPs, of which 2,732 were used to generate high-density genetic maps. Next, a robust consensus map was developed, and then QTL analyses were performed to corroborate marker-trait association test results using two years of field phenotypic data. We detected a major QTL on LG 15 that may co-localize with the previously-reported *Ren3* locus (Di Gaspero et al. 2005; Merdinoglu et al. 2005; Welter et al. 2007), which allowed the use of closely-linked SSR markers to screen the mapping families. In addition, a novel QTL (*Ren10*) with moderate effect was detected and this discovery was validated with a controlled inoculation study. Finally, the availability of high-quality SNP markers allowed development of haplotypes associated with the QTLs to identify

grandparental origin of powdery mildew resistance. In the case of *Ren3*, the use of linked SSR markers established a correlation between fragment sizes with resistance-carrying haplotypes that were not previously reported. This enables marker-assisted breeding for powdery mildew resistant individuals. Meanwhile for the *Ren10* QTL, fine mapping and additional marker development has to be made to convert the locus into reliable DNA markers before marker-assisted breeding can be applied effectively. The use of these molecular markers has the potential of helping breeders develop cultivars with more durable resistance to the worldwide economically damaging powdery mildew.

Within the larger context of our study, we demonstrated the advantage of dissecting complex genetic backgrounds that enabled the discovery of a novel resistance locus. Given the vast diversity in grape that has mostly been underexplored, future studies should explore other grape species to uncover potentially novel sources of resistance in an effort to develop resistant cultivars.

Tables

Table 2.1 Number of SNPs and total genetic distance (cM) of linkage groups (LGs) in the common parent MN1264 genetic maps of mapping families, GE0711/1009 (*N* = 147) and GE1025 (*N* = 125).

LG	GE0711/1009		GE1025		Consensus	
	Number of SNPs	Genetic length (cM)	Number of SNPs	Genetic length (cM)	Number of SNPs	Genetic length (cM)
1	26	32.5	53	29.2	54	34.0
2	59	97.4	67	74.5	93	90.2
3	31	95.0	47	91.6	51	56.3
4	87	145.1	99	87.5	148	130.1
5	75	109.7	141	120.4	153	109.7
6	51	78.6	76	98.6	90	97.0
7	87	132.4	133	137.7	165	137.7
8	78	104.7	89	80.3	126	84.7
9	48	112.1	60	82.1	73	82.1
10	40	71.9	57	78.1	65	77.7
11	30	97.9	47	92.6	52	93.8
12	63	108.9	78	91.9	112	104.4
13	80	125.5	118	144.4	149	70.9
14	54	140.9	135	115.3	151	139.4
15	64	129.8	111	111.9	123	122.7
16	36	100.6	49	144.7	61	97.6
17	37	113.2	91	94.4	94	106.6
18	82	238.9	113	189.6	127	106.6
19	49	107.3	77	140.6	90	111.4
Total	1077	2142.4	1641	2005.4	1977	1852.9

Table 2.2 Summary of two-year QTL information for powdery mildew resistance identified on LGs 2 and 15 for linkage maps of GE0711/1009 (*N* = 147) and GE1025 (*N* = 125), as well as the consensus map (*N* = 272).

LG	Map	Year	Analysis	LOD threshold		LOD_{max}	LOD_{max} position (cM)	Variance explained (%)	Confidence interval (cM) [LOD_{max} – 2]
				LG-specific[a]	Genome-wide[a]				
2	GE0711/1009	2014	IM	1.60	3.10	3.46	82.2	11.7	67.8 – 97.4
			CIM	–	3.68	3.49	83.0	10.1	77.0 – 86.1
		2015	IM	1.50	2.90	4.21	79.0	13.3	71.2 – 94.5
			CIM	–	3.73	5.38	79.0	13.3	74.8 – 84.0
	GE1025	2014	IM	4.10	4.10	–	–	–	–
			CIM	–	4.00	–	–	–	–
		2015	IM	6.50	6.50	–	–	–	–
			CIM	–	4.75	–	–	–	–
	Consensus	2014	IM	1.80	3.20	–	–	–	–
			CIM	–	3.54	–	–	–	–
		2015	IM	1.90	3.30	4.19	79.6	7.2	67.1 – 90.2
			CIM	–	3.58	3.91	74.8	7.0	69.0 – 78.0
15	GE0711/1009	2014	IM	1.60	3.10	5.58	7.2	17.3	0.0 – 58.5
			CIM	–	3.68	6.79	10.0	16.1	2.0 – 11.0
		2015	IM	1.60	2.90	–	–	–	–
			CIM	–	3.73	–	–	–	–
	GE1025	2014	IM	1.80	4.10	6.32	16.9	21.4	2.0 – 27.9
			CIM	–	4.00	8.08	16.9	20.1	12.0 – 21.2
		2015	IM	1.70	6.50	9.08	5.5	29.4	1.0 – 29.5
			CIM	–	4.75	10.30	14.0	28.5	9.0 – 18.0
	Consensus	2014	IM	1.60	3.20	10.21	18.5	17.1	5.0 – 24.4
			CIM	–	3.54	10.57	19.4	16.0	12.0 – 21.2
		2015	IM	1.70	3.30	9.11	16.5	15.7	2.0 – 37.2
			CIM	–	3.58	9.83	18.0	15.0	15.0 – 24.0
	GE1025 (AUDPC)	2014	IM	1.70	3.90	7.59	5.54	25.6	2.0 – 46.4
			CIM	–	4.14	6.96	13.2	23.6	5.5 – 15.0
		2015	IM	1.80	3.50	8.23	5.54	27.0	0.0 – 51.4
			CIM	–	4.04	8.21	5.54	27.0	1.0 – 10.5

[a] Estimated threshold values using a permutation test with 1,000 permutations at $\alpha = 0.05$.

Table 2.3 Number of SNPs and total genetic distance (cM) of linkage groups (LGs) in paternal genetic maps of GE0711/1009 (MN1214) and GE1025 (MN1246). The MN1246 map lacks LG 17, and contains two LG 15 segments that could not be merged.

LG	MN1214		MN1246	
	Number of SNPs	Total genetic distance (cM)	Number of SNPs	Total genetic distance (cM)
1	101	132.7	120	98.4
2	16	55.1	49	51.7
3	66	82.0	47	71.0
4	54	156.6	61	96.2
5	83	185.0	99	129.1
6	93	106.2	92	89.5
7	75	144.9	41	108.2
8	24	86.2	81	96.3
9	68	147.3	58	98.7
10	70	100.6	97	94.0
11	55	112.6	67	127.0
12	69	107.4	63	121.9
13	92	239.0	112	128.3
14	70	99.6	67	116.0
15	45	90.7	45[a]	50.5[a]
16	54	113.5	72	122.1
17	76	103.0	–[b]	–[b]
18	63	116.4	108	159.9
19	90	159.3	96	97.3
Total	1264	2338.1	1375	1856.1

[a] Numbers were expressed as a sum of two LG 15 segments.
[b] The paternal map of GE1025 was unable to identify a linkage group that corresponded to chromosome 17.

Table 2.4 Genotypes, QTL-spanning SNPs for LG 2 and LG 15, as well as field powdery mildew phenotypes for haplotype construction. SNPs from maternal genetic maps were used to trace the inheritance of resistant haplotypes from the maternal grandparents (i.e. 'Seyval blanc' and MN1069). At each locus, 'A' or 'B' was denoted to indicate inheritance from 'Seyval blanc' or MN1069, respectively. The order of SNPs is presented with respect to the maternal genetic maps.

Genotype	SNPs on LG 2												SNPs on LG 15								Field Phenotype	
	S2_8538108	S2_14108168	S2_14108170	S2_16681176	S2_17017148	S2_15954410	S2_16681183	S2_15052983	S2_17854965	S2_18455557	S2_18507119	S2_18702143	S15_2190957	S15_1931143	S15_102583	S15_202858	S15_98213	S15_98184	S15_584141	S15_1308526	2014[a]	2015[b]
GE0711_005	B	B	B	B	B	B	B	B	B	B	B	B	B	–[c]	B	B	B	B	B	B	7	8
GE0711_026	B	B	B	B	B	B	B	B	B	B	B	B	B	–	B	B	B	B	B	B	6	8
GE1009_048	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	7	9
GE1009_049	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	7	8
GE1025_044	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	7	9
GE1025_090	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	7	9
GE0711_004	A	A	A	A	A	A	A	A	A	A	A	A	B	–	B	B	B	B	B	B	5	5
GE0711_011	A	A	A	A	A	A	A	A	A	A	A	A	B	–	B	B	B	B	B	B	6	9
GE1009_052	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	4	3
GE1009_093	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	5	7
GE1025_018	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	5	8
GE1025_063	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	5	3
GE0711_020	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	2	5
GE0711_055	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	3	1
GE1009_020	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	3	5
GE1009_050	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	3	5
GE1025_081	B	B	B	B	B	B	B	–	B	B	B	B	A	A	A	A	A	A	A	A	3	3
GE1025_139	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	3	3
GE0711_033	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	2	1
GE0711_054	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	2	1
GE1009_030	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	2	1
GE1009_127	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	2	1
GE1025_053	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	3	1
GE1025_099	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	2	3

[a] Vines in 2014 were assessed using a 7-point scale of whole-plant foliage; [b] vines in 2015 with a 9-point IPGRI scale; [c] Missing data

Table 2.5 Statistical analysis of *Ren3* and *Ren10* prediction of powdery mildew response variables at 9-day post-inoculation following controlled inoculation with *Erysiphe necator* isolate NY19 in two independent experiments.

Table 2.5a Report of estimates and confidence intervals of the linear regression for the response variable Hyphal Transects (as a measure of hyphal proliferation), yielded from JMP 12.0.1 (Standard Least Squares personality). This indicates the statistical significance of the predictors Experiment and *Ren10*, and the lack of significance (at $\alpha \leq 0.1$) of *Ren3* and interaction terms in the two experiments with controlled inoculation. In the Term column, in brackets are shown the particular Experiment or origin of the resistance allele for which the estimates were calculated.

Term	Estimate	S.E.[a]	t Ratio	Prob > |t|	Lower 95%	Upper 95%
Intercept	56.63	4.08	13.88	<.0001*	48.63	64.64
Experiment[A]	-28.37	4.08	-6.95	<.0001*	-36.37	-20.37
Ren10[MN1069]	10.84	4.10	2.64	0.0083*	2.79	18.88
Ren3[MN1069]	-2.19	4.09	-0.54	0.5915	-10.21	5.82
Experiment[A]**Ren10*[MN1069]	-6.43	4.10	-1.57	0.1172	-14.47	1.62
Experiment[A]**Ren3*[MN1069]	-2.71	4.09	-0.66	0.5081	-10.72	5.31

[a] Standard error

Table 2.5b Report of estimates and confidence intervals of the linear regression for the response variable cube-root transformation of Hyphal Transects (as a measure of hyphal proliferation), yielded from JMP 12.0.1 (Standard Least Squares personality). This indicates the statistical significance of the predictors Experiment and *Ren10*, and the lack of significance (at $\alpha \leq 0.1$) of *Ren3* and interaction terms in the two experiments with controlled inoculation. In the Term column, in brackets are shown the particular Experiment or origin of the resistance allele for which the estimates were calculated.

Term	Estimate	S.E.[a]	t Ratio	Prob > |t|	Lower 95%	Upper 95%
Intercept	1.35	0.05	26.43	<.0001*	1.25	1.45
Experiment[A]	-0.40	0.05	-7.87	<.0001*	-0.50	-0.30
Ren10[MN1069]	0.21	0.05	4.08	<.0001*	0.11	0.31
Ren3[MN1069]	-0.07	0.05	-1.46	0.1459	-0.17	0.03
Experiment[A]**Ren10*[MN1069]	0.01	0.05	0.08	0.9347	-0.10	0.10
Experiment[A]**Ren3*[MN1069]	-0.04	0.05	-0.75	0.4557	-0.14	0.06

[a] Standard error

Table 2.5c Report of estimates and confidence intervals of the logistic regression for the response variable sporulation, yielded from JMP 12.0.1 (Standard Least Squares personality). This indicates the statistical significance of the predictors Experiment and *Ren10*, and the lack of significance (at $\alpha \leq 0.1$) of *Ren3* and interaction terms in the two experiments with controlled inoculation. In the Term column, in brackets are shown the particular Experiment or origin of the resistance allele for which the estimates were calculated.

Term	Estimate	S.E.[a]	χ^2	Prob > χ^2	Lower 95%	Upper 95%
Intercept	1.19	0.06	358.88	<.0001*	1.06	1.31
Experiment[A]	0.43	0.06	47.12	<.0001*	0.31	0.55
Ren10[MN1069]	-0.24	0.06	14.52	0.0001*	-0.36	-0.12
Ren3[MN1069]	0.05	0.06	0.54	0.4607	-0.07	0.17
Experiment[A]**Ren10*[MN1069]	-0.04	0.06	0.48	0.4884	-0.17	0.08
Experiment[A]**Ren3*[MN1069]	0.05	0.06	0.66	0.4160	-0.07	0.17

[a] Standard error

Figures

Figure 2.1 Two-generation pedigrees and ancestral species backgrounds of two mapping families, GE0711/1009 and GE1025. Pedigree information was obtained from the breeding records of *Vitis* International Variety Catalogue, VIVC (https://www.vivc.de/). Allele sizes from *Ren3* SSR markers were colorized to trace the inheritance of alleles.

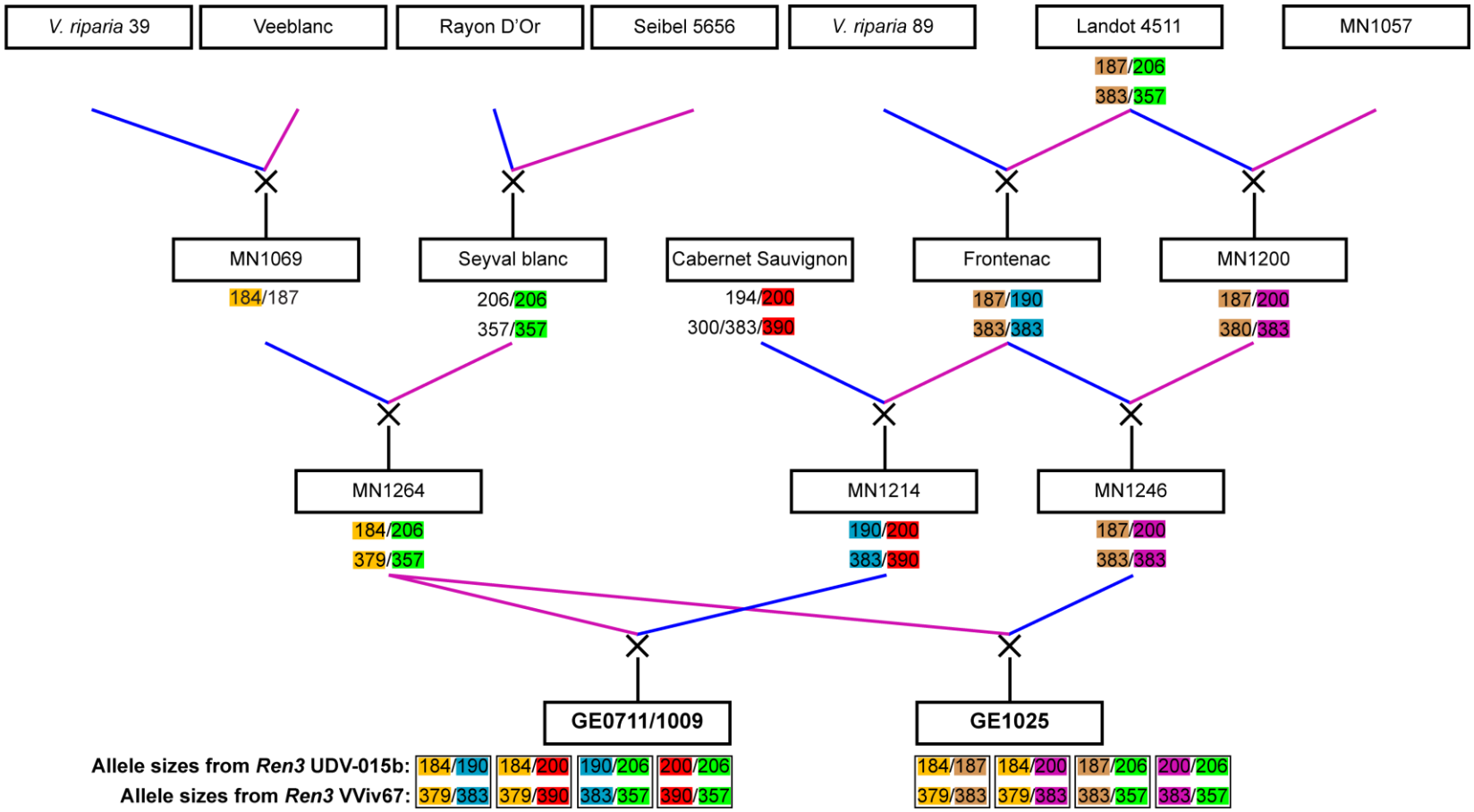


Figure 2.2 (A) Correlation between disease severity and four haplotype combinations in the combined maternal families of GE0711/1009 and GE1025 showed evidence of genotypes with 'Seyval blanc'-inherited haplotypes having the lowest disease severity mean score in both years. (B) Correlation between colonization (expressed in cube root transformed hyphal transect) and four haplotype combinations in GE1025 showed evidence of individuals with 'Seyval blanc'-inherited *Ren10* haplotype exhibiting the lowest mean of colonization in Experiment A. Similar trend was shown in Experiment B, but lacked statistical support. Different letters were assigned on score ranges to indicate statistical significance that was computed by Tukey's HSD test ($P < 0.05$).

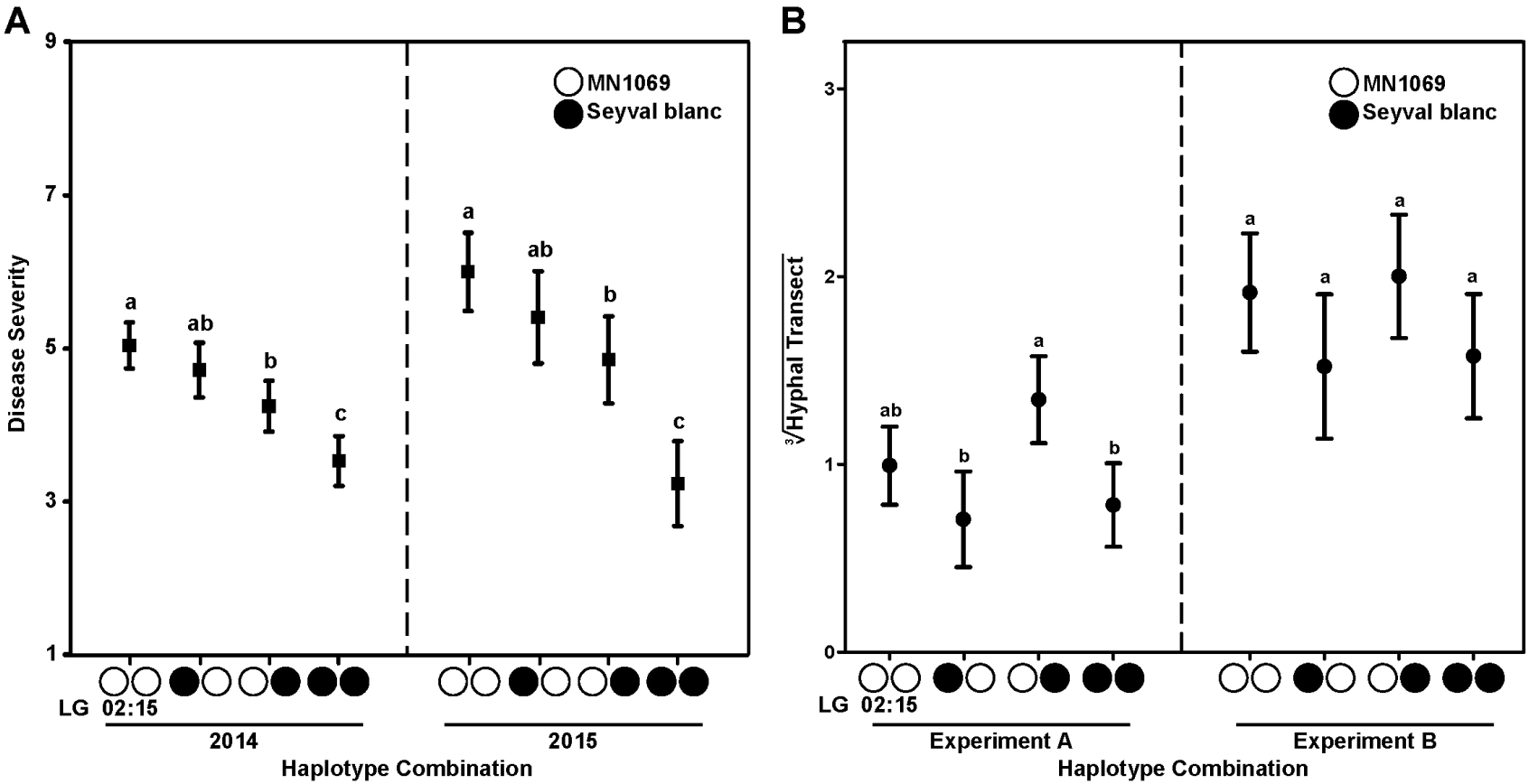


Figure 2.3 Correlation between disease severity and PCR fragment sizes of two SSR markers (UDV-015b and VViv67) that are tightly linked to the *Ren3* locus. Results were shown in two years of each bi-parental family: (A) GE0711/1009; (B) GE1025. Letters were assigned indicate statistical significance that was computed by Tukey's HSD test ($P < 0.05$).

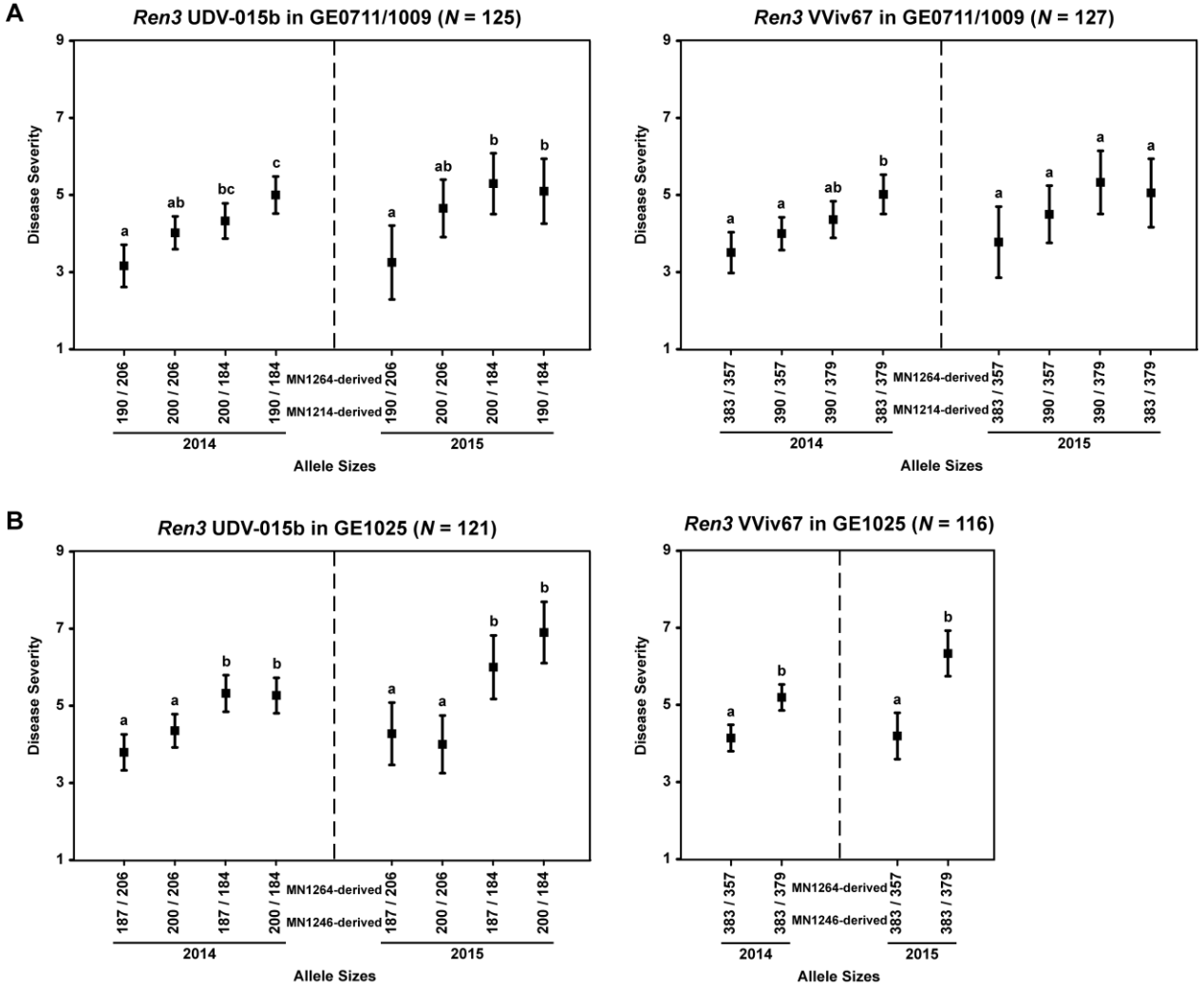


Figure 2.4 Two-dimensional histogram and scattered distribution of field powdery mildew severity in two years. Phenotypic information was jittered to show overlapping data points.

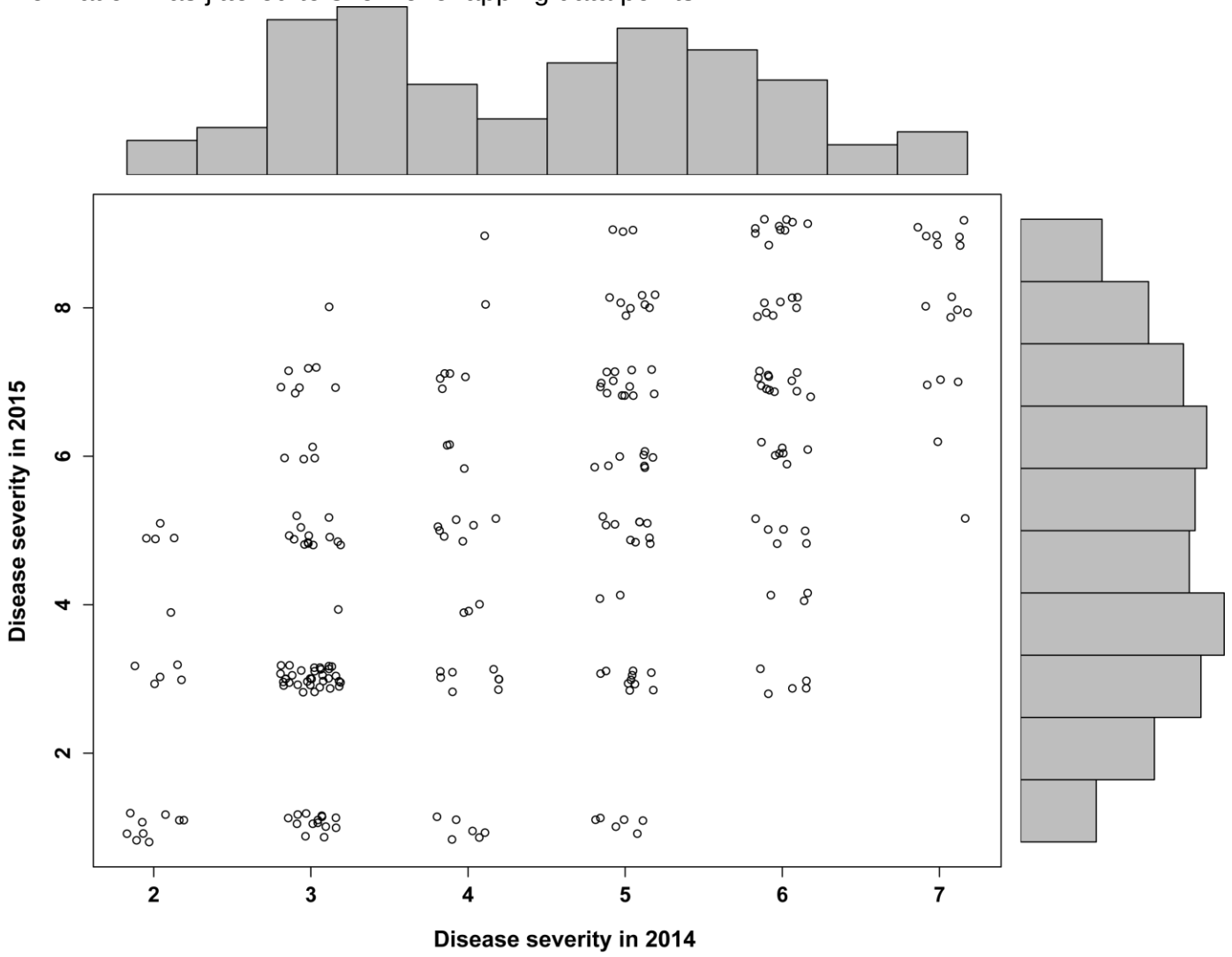
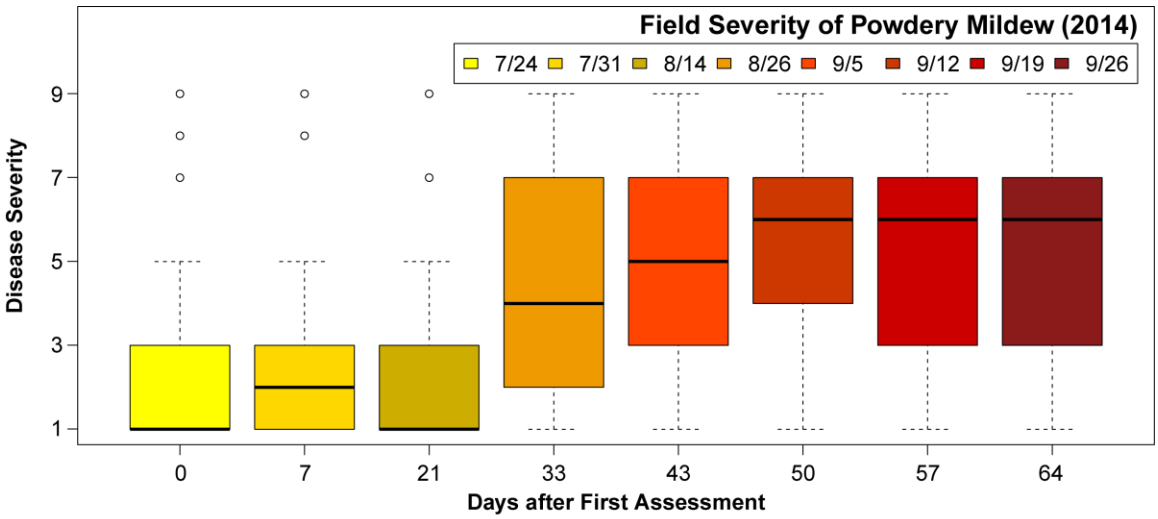


Figure 2.5 Distribution of field severity of powdery mildew in 2014 and 2015.



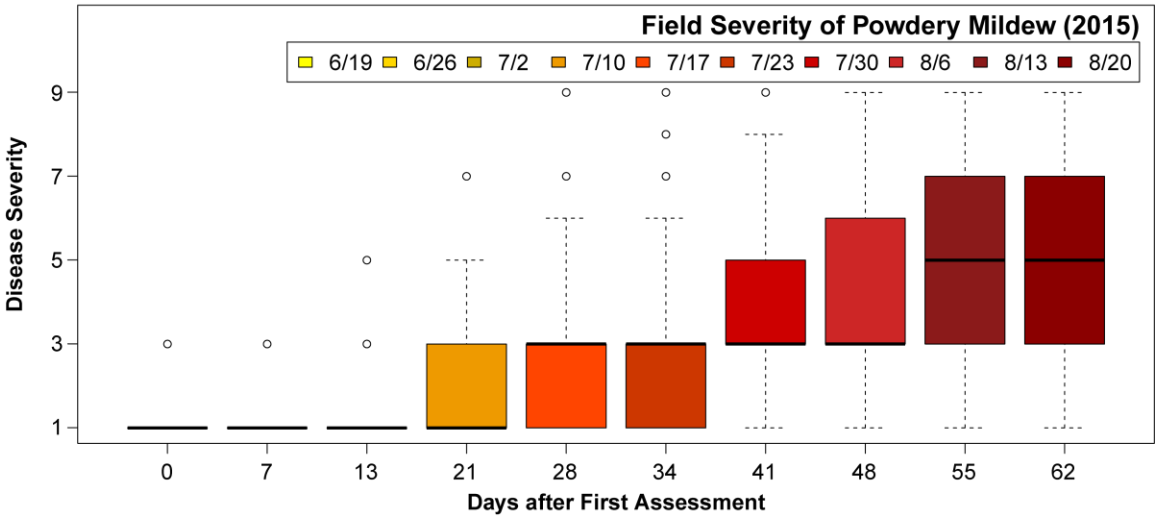


Figure 2.6 QTL results for powdery mildew resistance on LG 2 and LG 15 of the maternal maps of GE0711/1009 and GE1025, as well as the consensus map in both years using IM and CIM analyses. Genetic regions, where colorized, indicated intervals that exceeded a statistical threshold value of 3.0. LOD scores were color-coded as described in the legend.

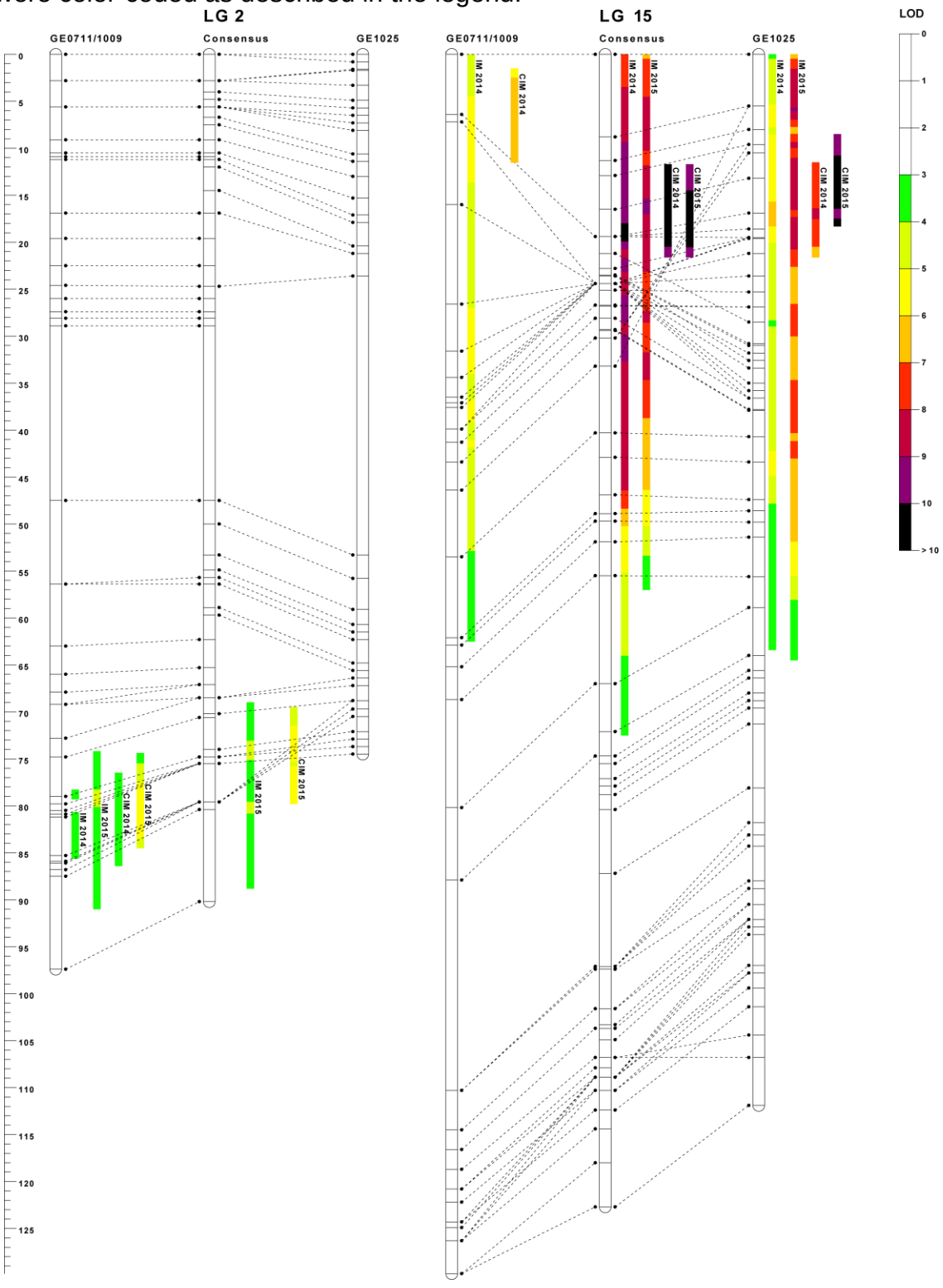


Figure 2.7 Genetic maps of MN1264. At each LG, the consensus map (middle) was constructed using maternal information from GE0711/1009 (left) and from GE1025 (right).

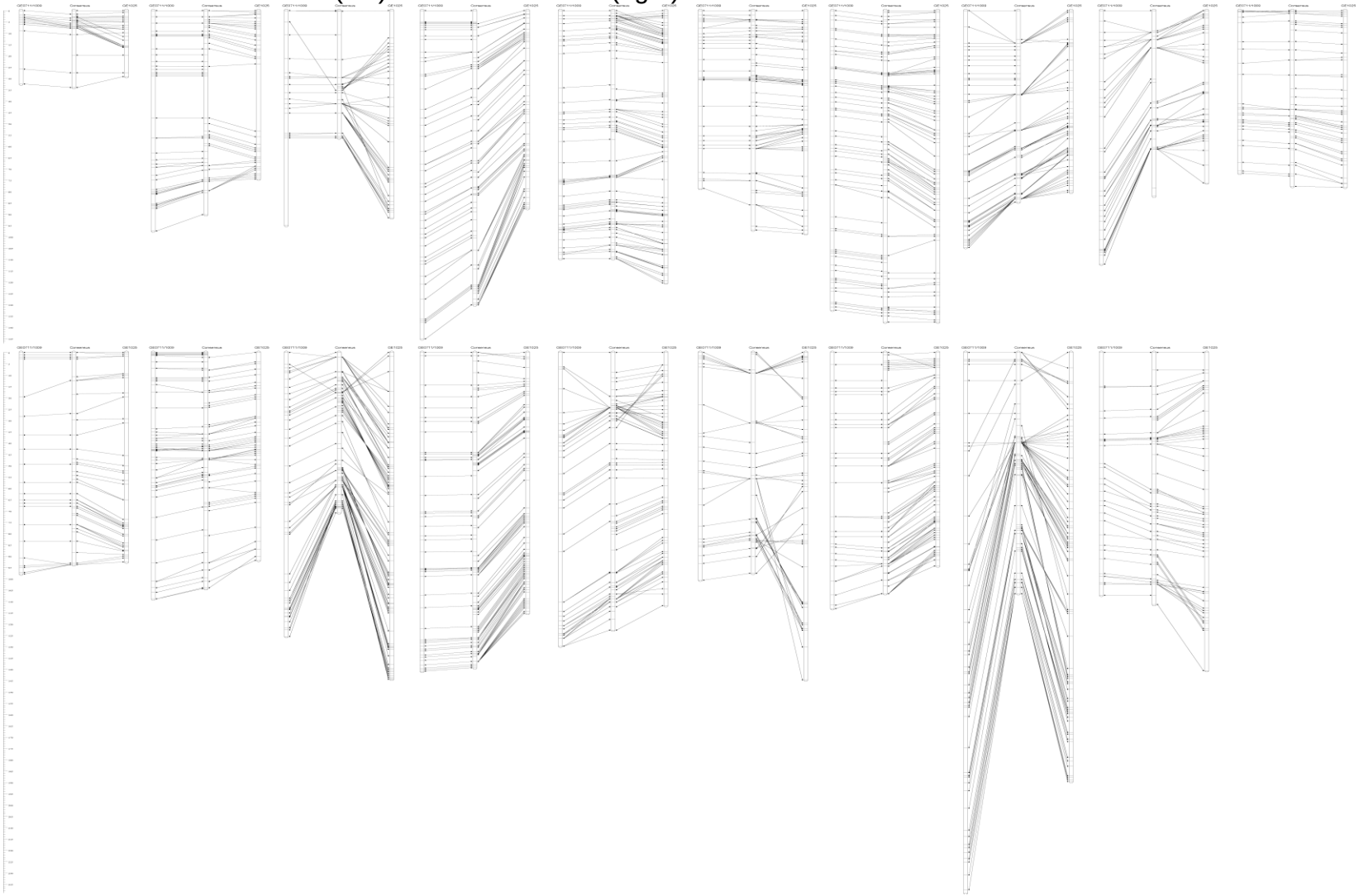


Figure 2.8 A genetic map of MN1214, or paternal map of GE0711/1009.

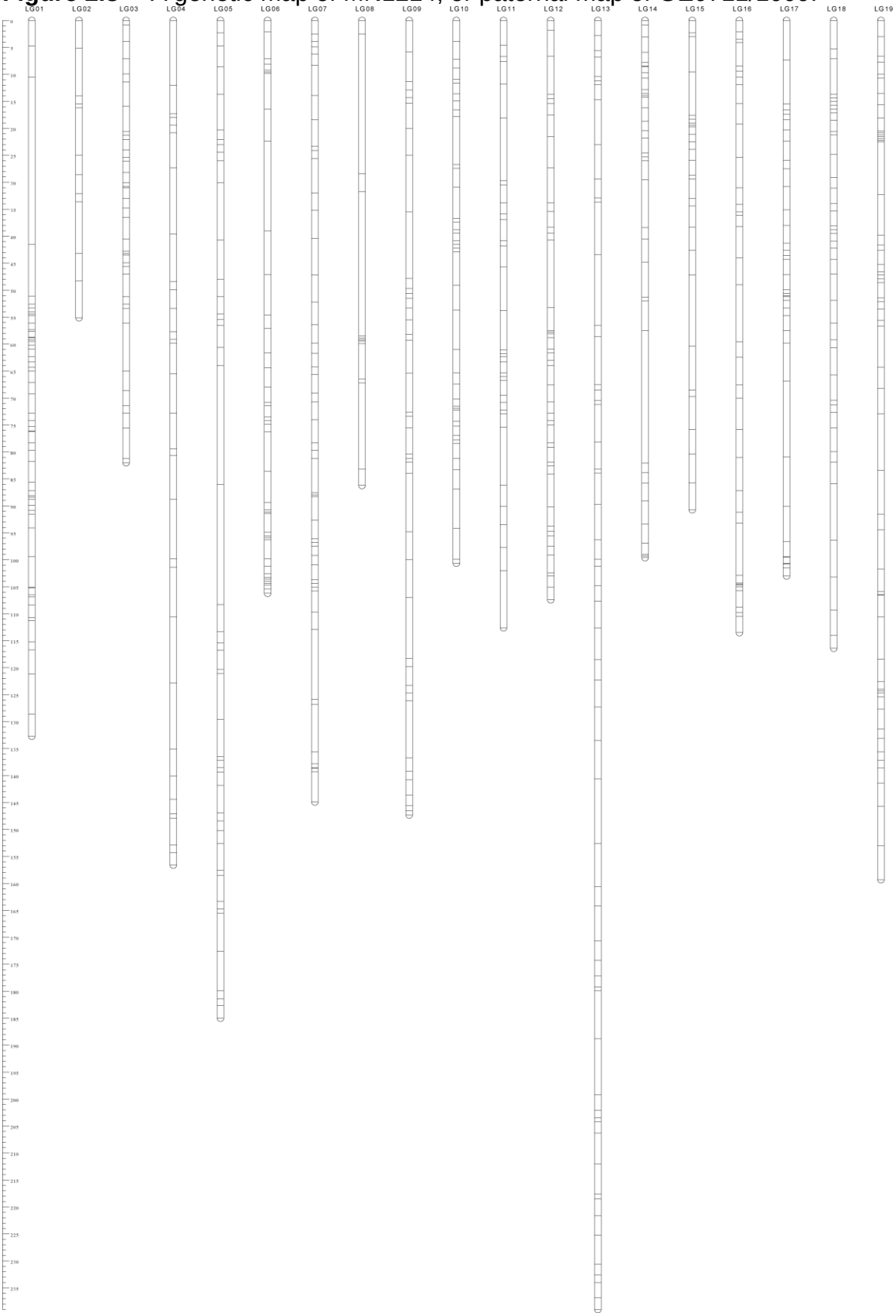


Figure 2.9 A genetic map of MN1246, or paternal map of GE1025. The map lacks LG 17, and contains two LG 15 segments that could not be joined.

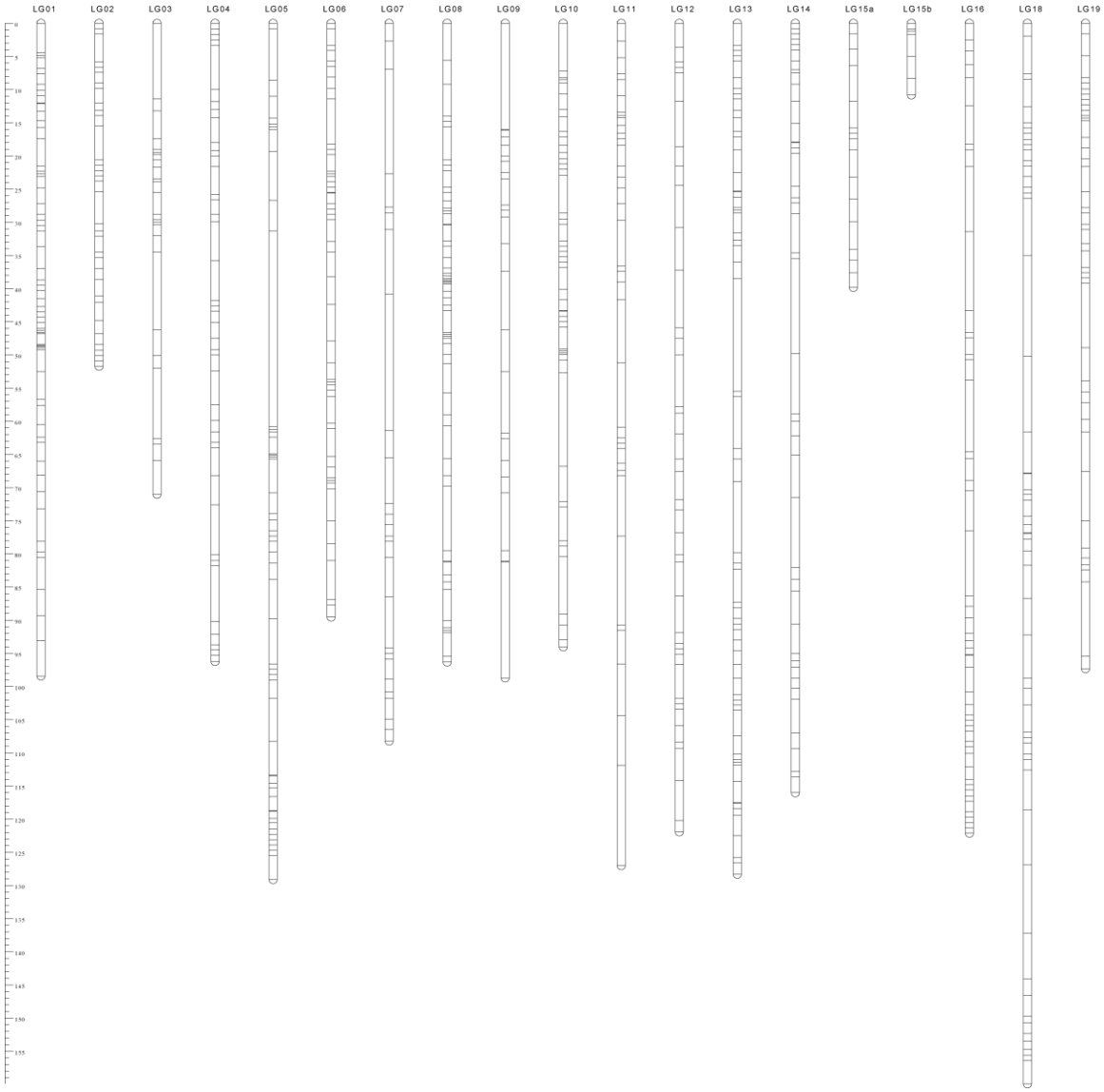
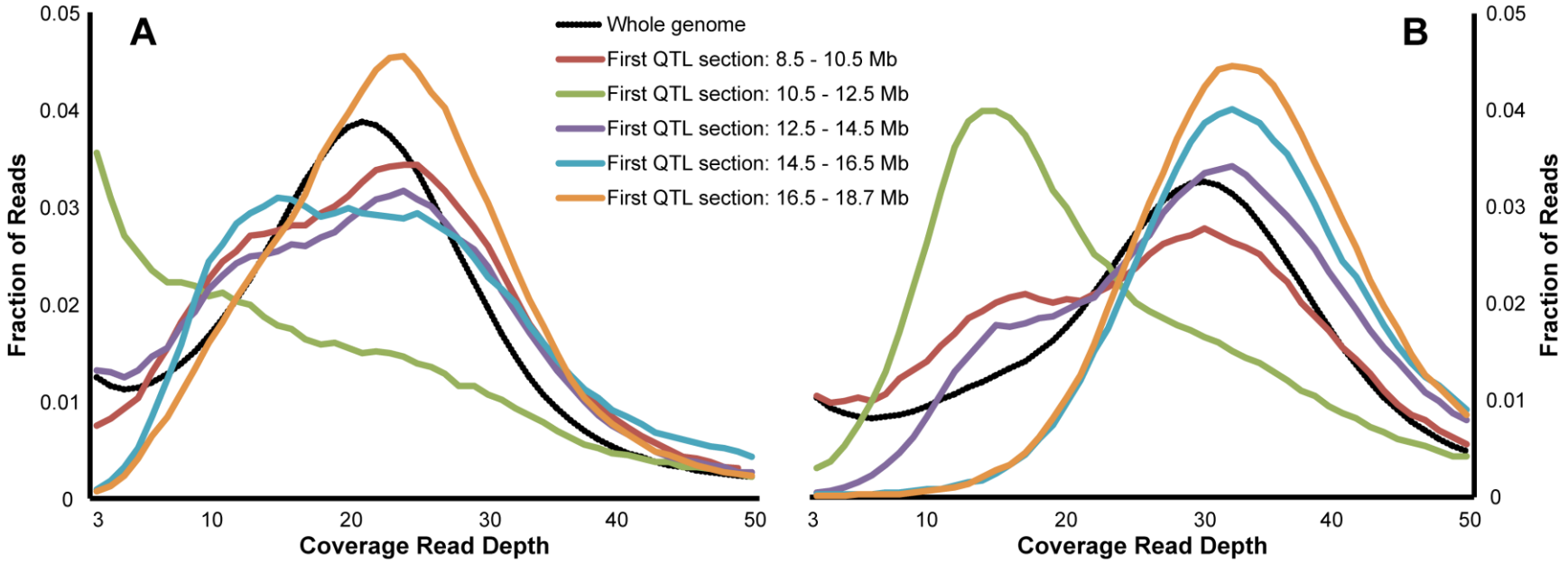


Figure 2.10 Distribution of the fraction of sequence reads with respect to the per base coverage read depth for the individuals: (A) MN1264; (B) MN1246 in the *Ren10* QTL region. Each solid line represents the per base coverage read depth distribution for a 2 Mbp section within the *Ren10* region. The black line represents the genome-wide per base coverage read depth distribution. The data come from a parallel project for the whole genome re-sequencing of parental germplasm within *Vitis*Gen.



Chapter 3

Genetic Analysis of Stilbenes in Grapevine Stems Indicates an mQTL Hotspot for Disease Resistance Motifs on Chromosome 18

Cultivated grapevine (*Vitis vinifera* L.) contains a wealth of phytochemicals, including stilbenes, which have received considerable attention due to health-promoting properties and phytoalexin activity. To date, the genetic basis of the quantitative variations for these potentially beneficial compounds has not been characterized. Here, metabolic quantitative trait loci (mQTLs) were mapped for compounds extracted from grapevine stems of a segregating F_2 population. Metabolic profiling of grapevine stems was performed using liquid chromatography-high resolution mass spectrometry (LC-HRMS), resulting in the detection of 1317 ions. Using previously assembled genetic maps, four mQTLs were identified using R/qtl software, one on linkage group (LG) 11 and three on LG18. Based on the VitisNet network database, functional annotation indicated a genomic hotspot on LG18 containing a suite of disease resistance motifs. In addition, another mQTL was localized on a genomic region approximately 15 Mbps upstream that is associated with genes involved in primary metabolism. In summary, this is one of the earliest reports of loci identified by metabolic profiling-based mQTL analysis using a segregating population in a perennial crop that has the potential to be applied for routine marker-assisted selection in a breeding program.

3.1 Introduction

Cultivated grapevine (*Vitis vinifera* L.) is a major perennial fruit crop worldwide due to its economic impact and historical significance. Grapes are used to produce wine, raisins, jam, juice and fresh fruit. In recent decades, grapes have also been recognized as a rich source of phytochemicals, such as phenolic compounds. A single glass (approximately 175 mL) of red wine that is obtained from the fermented extract of 100 to 140 berries contains up to 500 mg of polyphenolic compounds, depending on varieties and vinification methods (Lopez et al., 2001; Stervbo et al., 2007). Within the diversity of polyphenolics, stilbenes represent a relatively restricted group of phenols that is derived from the general phenylpropanoid pathway, and are found in several plant families, including Vitaceae (Pawlus et al., 2012). Stilbenes have received considerable attention due to: (1) epidemiological studies attributing moderate consumption of red wine to health benefits, and (2) biological activities as phytoalexins (Adrian et al., 2007; Anstey et al., 2009; Chao et al., 2008; Mattivi et al., 2011; Pezet et al., 2004a; Pezet et al., 2004b; Wang et al., 2006).

The production of stilbenes can be induced by abiotic (e.g., UV irradiation, mechanical injury), as well as biotic (e.g., fungal pathogens) stresses (Bavaresco and Fregoni, 2001; Bavaresco et al., 2009; Mattivi et al., 2011; Pezet et al., 2004a; Pezet et al., 2004b). Meanwhile, the levels of stilbene production vary across different plant parts, in terms of concentrations and types. For instance, resveratrol is induced in grape leaves and berries, but constitutive expression and accumulation of resveratrol and other stilbenes, which are hypothesized to

protect against fungal infection, take place primarily in stems and roots (Pawlus et al., 2012).

Due to the biological activities of stilbenes affecting both plant health and human health, the possibility to artificially manipulate stilbene production and levels is of great interest. Although breeding objectives in grapevine vary by regions and market targets, the overarching objective of many programs is to combine high-quality fruit traits with improved biotic (e.g., diseases and pests) and abiotic (e.g., climate and environmental adaptation) resistance.

The availability of myriad genetic resources has enabled routine application of genetic markers for parent and seedling selection, or marker-assisted breeding. In addition to genomic resources, metabolomics is a field that is receiving more attention in the area of crop breeding. Metabolomics is an analytical field that provides a comprehensive investigation of metabolite variations in a biological system. Coupled with resources from genomics and transcriptomics, metabolomics is developing as an integrative functional tool in crop breeding.

To date, although quantitative trait loci (QTLs) have been mapped for various traits in grapevine (e.g., disease resistance and plant morphology), genetic mapping in fruit crops that is based on metabolomics has been limited to flavor, aroma and nutritional traits (Tieman et al., 2006; McCallum et al., 2010; Eduardo et al., 2013; Chambers et al., 2014; Liu et al., 2016). The aim of the study is to elucidate the genetic basis of stilbene variability in grapevine stems by combining modern analytical tools with existing genetic information. In this study,

we highlighted a facile application of liquid chromatography-high resolution mass spectrometry (LC-HRMS)-based metabolic profiling to identify metabolic QTLs in a segregating experimental population without *a priori* knowledge of compounds.

3.2 Materials and Methods

3.2.1 Sampling of Plant Material

A segregating F_2 mapping family was derived from self-pollination of a single hermaphrodite F_1 individual (16-9-2), which was a hybrid from the initial F_0 cross of *V. riparia* 37 (USDA PI 588259) × 'Seyval' (Seyve-Villard 5-276), as previously described by Fennell et al. (2005). The experimental family was grown and maintained in Brookings, SD and was also used in other genetic studies (e.g., Yang et al. 2016). In early January, 2014 and mid-February, 2014, two to four cuttings of approximately three to five inches of dormant woody stems were collected for each individual from the field fora subset of this mapping family (N = 101). Samples were wrapped in aluminum foil to prevent moisture loss, shipped on dry ice to Saint Paul, Minnesota, and stored at -80 °C until subsequent analysis.

3.2.2 Sample Preparation and Extraction

Woody stems were lyophilized for 24 to 36 hours until all moisture was removed. The stems were subsequently ground with a conventional coffee grinder (Kuissential™ Ceramic Burr Coffee Grinder). The ensuing steps, including chemical extraction, sonication, centrifugation, evaporation, and

adjusted reconstitution, were performed based on a facile extraction method that was described by Pawlus et al. (2013) with some modifications (Figure 3.1).

3.2.3 Metabolic Profiling with LC-HRMS

Reconstituted extracts were analyzed on a hybrid quadruple Orbitrap Q Exactive mass spectrometer (Thermo Scientific, http://www.thermoscientific.com) equipped with a Dionex Ultimate 3000 HPLC (http://www.dionex.com). Samples were chromatographically resolved at a flow rate of 450 µL/min on a C_{18}-reverse-phase column (HSS T3, 2.1 mm i.d. × 100 mm, 1.8 µm particle size; Waters) by mixing mobile-phase solvent A (water with 0.1% formic acid) and solvent B (acetonitrile with 0.1% formic acid) to generate the following gradient: 0 to 1 min, 2% B; 1 to 10 min, 2 to 30% B; 10 to 12 min, 30 to 50% B; 12 to 20 min, 50 to 75% B; 20 to 22 min, 75 to 98% B; 22 to 23 min, 98 to 2% B; 23 to 27 min, 2% B. The mass spectrometer was operated in the positive/negative switching ionization mode over a full m/z scan range of 150 to 2000. The presence of monomers, dimers, trimers, and tetramers was monitored using the $[M+H]^+$ *m/z* values of 229, 445, 681, and 907, respectively. The HPLC-UV absorbance information was also collected throughout each separation, where traces at 280 nm and 306 nm (λ) were indicative of eluting stilbenes.

3.2.4 Data Pre-processing

Metabolomics data from LC-MS was processed in XCMS for peak alignment with parameter settings optimized for the analysis (method = centWave; peakwidth = 5,20; snthresh = 10; ppm = 3.0; mzdiff = 0.01). The aligned peaks were then subjected to grouping, retention time correction, and

regrouping with optimized parameters (bw = 2; mzwid = 0.015; minfrac = 0.1). Following feature grouping and correction, peak filling was performed using chromatography method in XCMS, allowing integration of the area under the curve of samples that might have been missed during the group step.

3.2.5 Metabolic Quantitative Trait Locus (mQTL) Mapping

An exhaustive search for literature-reported compounds belonging to the class of stilbenes was conducted. Using the list of known metabolites and their corresponding accurate masses, the XCMS-processed feature list was interrogated to match metabolic features based on accurate masses. A list of matched features was compiled for subsequent mQTL mapping.

Our present study utilized an integrated single nucleotide polymorphism (SNP)-based genetic map that was previously constructed and described by Yang et al. (2016). The objective of the study was to genetically map the putative compounds based on accurate masses. Targeted mQTL detection was conducted using composite interval mapping that was performed on R/qtl software (Broman et al. 2003). The minimum logarithm of odds (LOD) score for mQTL detection was 3.0, which was determined by the genome-wide LOD significance threshold (α = 0.10) calculated using 1,000 permutations.

3.2.6 Integration with Functional Annotation Network

Based on the association between putative metabolites and targeted loci, the physical positions of SNPs were identified and assigned using functional annotation derived from VitisNet. VitisNet is a grapevine molecular network database containing manual annotations of the grape genome including

individual genes, proteins and metabolites, as well as pathway names, thus providing biological meaning through the combination of genomic and metabolomic datasets (Grimplet et al. 2009).

3.3 Results

3.3.1 Profiling of F_2 Mapping Family Revealed Metabolic Inheritance and Segregation

The original F_2 mapping family consists of 424 offspring and has been used for QTL mapping of enological and environmental adaptation traits (Fennell et al., 2005; Yang et al., 2016). In this study, a subset (*N* = 101) from the pilot mapping family was used for metabolic profiling, as a precursor to mQTL mapping. Targeted metabolomic analysis of the grapevine stems using LC-HRMS yielded 1317 ions (unique retention time–*m/z* ion pairings) in the $[M+H]^+$ mode. Ions yielded in the $[M–H]^–$ mode were not used for mQTL mapping due to ionization issues.

Based on the total ion chromatograms (TIC), an inheritance pattern of chemical profiles is observed when comparing TICs of a grandparent (F_0: *V. riparia* (USDA PI 588259)), the parent (F_1: *V.* hybrid (16-9-2)), and a representative F_2 progeny (Figure 3.2). In addition, segregation of metabolic profiles among the F_2 progeny was exhibited when conducting extracted ion chromatograms for four distinct types of stilbene oligomers (Figure 3.3). The metabolic segregation of ions provided the foundation for mQTL mapping by treating each ion as a metabolic trait.

3.3.2 mQTL Mapping

In a classical QTL mapping experiment, two parents differing genetically for the trait of interest are crossed so that the measurable trait that segregates in the offspring family can be statistically associated with the underlying genetic markers to explain the genetic basis of variation for the trait. Despite the advancement and higher throughput of modern analytical tools, crop breeding and genetics have continued to rely on traditional traits. Reports of chemical profiles as metabolic traits for genetic mapping have been limited to aroma, flavor and nutritional traits (Tieman et al., 2006; McCallum et al., 2010; Eduardo et al., 2013; Chambers et al., 2014; Liu et al., 2016).

In this study, a list of accurate masses corresponding to stilbenes reported in the literature were compiled. Subsequently, extracted ion chromatograms were made for each metabolic feature/ion to acquire the intensity values of individuals in the mapping family (Figure 3.3). Each metabolic feature was treated as a metabolic trait for mQTL mapping.

From the list of 1317 ions/features, four metabolic features were detected for mQTLs (Figure 3.4; Table 3.1). These features belong to different types of stilbene oligomers. Feature M229.1423T503 (*m/z*: 229.1423; retention time: 503 seconds) is a monomeric stilbene that was mapped to chromosome 11 with an LOD score of 8.96. Meanwhile, an LOD of 6.26 was detected on chromosome 18 for a stilbene dimer of M453.1357T505. In addition, feature M681.2169T759 is a trimer that was strongly associated to a region on chromosome 18 with an LOD

of 8.93. Finally, a tetramer stilbene feature of M907.2767T765 was also mapped to chromosome 18 with an LOD of 3.18 (Table 3.1).

The discoveries of mQTL for these metabolic features were validated using two QTL detection methods – interval mapping and composite interval mapping (Figure 3.4). For both of these methods, the identified mQTLs exceeded the minimum LOD threshold scores that were required for QTL detection, further buttressing the significance of these detected mQTLs.

3.3.3 An mQTL Hotspot on LG18 for Disease Resistance Proteins and Motifs

Subsequent to the discovery of mQTLs, the VitisNet database was used to provide functional annotations at the corresponding genomic regions where mQTLs were detected. For the mQTL associated with feature M229.1423T503, there was no corresponding *Vitis* probe ID or reference region in the 12× draft genome sequence; hence, functional annotation at the neighboring 200 kbps was not identified.

The mQTL detected for feature M907.2767T765 was associated with probes that are localized to the physical position of 11.20 – 11.22 Mbps on chromosome 18 of the 12× draft genome sequence. Based on the functional annotation, proteins found in this genomic region include kinase, proteasome, glycosyl transferase, and dehydrogenase.

Another region on chromosome 18 that was detected as an mQTL is at a physical position of 26.39 Mbps. The mQTL for feature M453.1357T505 coincides with that of feature M681.2169T759. From the VitisNet database, this

region contains numerous genes with motifs for disease resistance proteins, particularly those of Toll/interleukin-1 receptor (TIR)-nucleotide-binding site (NBS)-leucine-rich repeat (LRR). As indicated in Table 3.1, over 50% of proteins encoded within this 200 kbps region are disease resistance proteins/motifs. Expanding beyond this region (300 kbps upstream and 500 kbps downstream), disease resistance motifs appear to predominate this 1 Mbps region.

3.4 Discussion

3.4.1 Mapping Genomic Regions Using Variation in Metabolite Levels of the Progeny

In general, genes that are involved in a biosynthesis pathway have been identified either through forward genetics (i.e., screening the phenotype to identify gene(s) responsible for a trait of interest), or reverse genetics (i.e., evaluating the phenotype subsequent to disruption of gene function). In traditional crop breeding, the former case has been routinely applied, resulting in QTL discovery, fine mapping, candidate gene cloning, and finally, functional identification. Despite the eventual identification of genes, it remains unclear which of the genes are responsible for the variation in metabolite levels across tissue types and genotypes. On the other hand, genetic mapping of metabolites provides an alternative solution to elucidate putative genes that underlie metabolic variations.

In this study, the combination of a genetic linkage map of 1449 SNPs and a metabolomic dataset of 1317 metabolic features using grapevine stems from

an F_2 family of 101 offspring afforded the discovery of four mQTLs. In addition, the use of a high-resolution mass spectrometer provided accurate-mass detection of ions from a complex chemical extract, thereby enabling high throughput analysis of the mapping family. Without *a priori* knowledge of compound identity, metabolic profiling of an experimental population can provide thousands of metabolic features to be evaluated in mQTL mapping. Similar to the conventional marker-trait association where marker information is derived from a genetic linkage map while trait information is acquired from a variation for phenotype of the trait, a metabolic feature can be treated as a 'trait' to assess marker-'trait' (feature) association. Given the variation in ion abundance/intensity, the distribution of feature levels may have to be transformed to prevent data skewedness in genetic mapping, which assumes a normal distribution. Although this was not required in our dataset, $\log_2$ and $\log_{10}$ transformations can be performed to normalize the distribution, if needed.

To gain more insights into the mQTL regions, we assessed the markers flanking peak QTLs to their physical positions on the 8×.v0 *V. vinifera* 'PN40024' reference genome (Jaillon et al., 2007) (Table 3.1). In addition, we utilized VitisNet to link the associated genomic-metabolomic regions with publicly available functional annotations (Grimplet et al., 2009; Grimplet et al., 2012). Through the integration of metabolomics information with VitisNet, we reported several putative biochemical compounds that are associated with the mQTL regions.

3.4.2 No mQTL Was Detected for Most Metabolic Features

From the list of 1317 metabolic features, a mQTL was not detected for the vast majority of features. This highlights the discriminating power of using metabolic features as traits for mQTL analysis. Despite the suitability of this method, there are several reasons why mQTLs may have been underreported (i.e., false negatives). First, complex genetic regulation of metabolites in a pathway could hamper the detection of mQTLs. The production of one metabolite is likely controlled by several small-effect loci, which may not be detected during the mapping analysis. In addition, the complex genetic backgrounds of an already highly heterozygous crop may help explain the presence of small-effect loci, rendering a lower-resolution mQTL mapping. The ancestry of the F_2 mapping family is comprised of at least six *Vitis* species, including *V. riparia*, *V. vinifera*, *V. rupestris*, *V. labrusca*, *V. aestivalis* and *V. berlandieri*. Therefore, the hybridization of genetic backgrounds likely dilutes the effects of an otherwise strong-effect mQTL. Finally, the ion signal of metabolic features could be above the noise threshold for some offspring, but within the noise threshold for other offspring, thereby resulting in no mQTL detection.

3.4.3 mQTL Hotspot on LG18 for Disease Resistance Proteins and Motifs

Integrating the mQTL mapping results and the VitisNet analysis, an mQTL hotspot on LG18 was detected for features M453.1357T505 and M681.2169T759. In both instances, the markers flanking the peak QTLs cover a physical distance of 13 kbps. Using a genetic window spanning 200 kbps of the mQTL hotspot, 16 putative genes/proteins were detected based on the published

genome of grape (Jaillon et al., 2007; https://urgi.versailles.inra.fr/Species/Vitis/Data-Sequences/Genome-sequences). Expanding beyond this mQTL region (300 kbps upstream and 500 kbps downstream) to cover a standard 1 Mbps window, 40 putative genes were observed, of which 20 are unique genes that are predominated by disease resistance proteins/motifs, such as TIR-NBS-LRR, TIR-NBS, LRR family protein and R protein L6 (Table 3.1).

From the network database, virtually all the proteins in this region were annotated as “R proteins from plant-pathogen interaction” or “diterpenoid biosynthesis”. Meanwhile, the corresponding functional categories were almost exclusively “response to stimuli”, “biotic stress response”, “plant-pathogen interaction”, and “transposable elements”. Taken together, the region on LG18 appears to be a hotspot for the biosynthesis of defense metabolites.

Based on the pedigree backgrounds of this mapping family, the F_2 progeny likely segregates for disease resistance, such as powdery mildew, downy mildew and black rot. However, given that the grapevine stems were collected from the vineyard, it remains to be determined whether the mQTL discovery was due to constitutive metabolite production, or inducible expression of defense compounds (i.e., phytoalexins).

3.4.4 mQTL on LG18 Associated with Primary Metabolism

The genomic region for mQTL associated with feature M907.2767T765 on LG18 appears to be populated with a suite of genes linked with primary metabolism. In combination with the observation of disease resistance motifs highlighted in the aforementioned mQTL hotspot (physical position: 26 Mbps), it

is likely that this region (11 Mbps) is also associated with plant-pathogen interaction. Secondary metabolism is of great interest in plant-pathogen interaction because phytopathogen infection induces a plant's defense program. However, less is known about the effects of pathogenic attack on primary metabolism. This is especially important because most attack (e.g, parasitic relationship) causes yield losses, without resulting in crop death. In particular, various aspects of photosynthesis, assimilate partitioning, and source-sink relationship are downstream physiological changes of the infected tissues that need to be investigated to understand the mechanisms and consequences of a plant-pathogen interaction (Berger et al., 2007; Bolton, 2009; Schwachtje and Baldwin, 2008).

Based on the database analysis of this region, some of the function annotations in this genomic window include signaling pathway, macromolecule transport, transcription regulation, ubiquitin-mediated proteolysis, nucleic acid metabolism, carbohydrate metabolism and glycosyl transference. In addition to genes related to primary metabolism, the annotation indicates a peroxidase gene that appears to be involved in the metabolism of phenylalanine, a precursor to various biosyntheses, such as phenylpropanoid and flavonoid pathways. Taken together, it is likely that the two mQTL regions (11 Mbps and 26 Mbps) may be linked with plant-pathogen interaction, both in primary and secondary metabolism.

3.5 Conclusions

In this study, we demonstrated the utility of combining analytical tools and large genomic datasets to study the genetic basis of metabolites. Despite the absence of compound identity and structures, the use of high-resolution mass spectrometer provided detection of various ions that can be treated as a metabolic trait and used in combination with genetic maps for mQTL mapping. Understanding the genetic controls of potentially bioactive compounds (e.g., stilbenes) can assist breeders and viticulturists to develop breeding strategies for selection of genotypes (e.g., seedlings, parents and existing cultivars) with increased levels of these biomarkers through marker-assisted breeding.

In subsequent experiments, structural characterization of the putative biomarkers should be carried out to decipher compound identity and structure. In addition, a control experiment of the mapping family should be set up to unambiguously delineate metabolites that are due to inducible response versus constitutive expression. Finally, the gold standard to prove the contribution of these putative biomarkers would require genetic complementation tests, where identified mQTL can be inserted into existing elite grapevine cultivars through selective backcrossing to assess if allelic insertion/enrichment would result in increased levels of these biomarkers.

Table

Table 3.1 Summary of metabolic features, patterns of stilbene oligomers, mQTL detection and statistics, as well as the corresponding functional annotations from the *Vitis*Net database.

Feature	Oligomer Type	Peak QTL Marker	LOD; Variance Explained	Markers Flanking Peak QTL	*Vitis* Probe ID	Chromosome Position on the 12× Draft Genome Sequence	Functional Annotation (within 200 kbps)
M229.1423T503	Monomer	–	8.96; 4.3%	S11_2824566 S11_3485916	–	–	–
M453.1357T505	Dimer	S18_26392346	6.26; 29.0%	S18_26379100 S18_26392343	VIT_18s0041g01550	chr18_26391778 _26396402	- TIR-NBS disease resistance - TIR-NBS-LRR disease resistance protein - R protein L6 - Taxadiene 5-alpha-hydroxylase - CYP706A12 - Tubulin, Beta 2 - Taxane 10-beta-hydroxylase - Leucine-rich repeat family protein - Retrovirus Pol polyprotein
M681.2169T759	Trimer	S18_26392346	8.93; 38.4%	S18_26379100 S18_26392343	VIT_18s0041g01550	chr18_26391778 _26396402	- TIR-NBS disease resistance - TIR-NBS-LRR disease resistance protein - R protein L6 - Taxadiene 5-alpha-hydroxylase - CYP706A12 - Tubulin, Beta 2 - Taxane 10-beta-hydroxylase - Leucine-rich repeat family protein - Retrovirus Pol polyprotein

M907.2767T765	Tetramer	S18_11206896	3.18; 7.8%	S18_10887948 S18_11381025	VIT_18s0001g13150 VIT_18s0001g13160	chr18_11201495 _11206144 chr18_11209086 _11221435	- Adaptor-related protein complex AP-4 - Mitogen-activated protein kinase 1 - Basic leucine-zipper 11 GBF6 - Protein translocase Tic20 - C3H2C3 ring-finger protein - Proteasome 26S AAA-ATPase subunit (RPT4) - Glycosyl transferase family 1 protein - GTP-binding protein hflX - BEL1-like homeodomain 10 - BEL1 homeotic protein 3 - Tetrapyrrole methylase - Ubiquitin-conjugating enzyme - MAPK (MPK9) - Cytokinin dehydrogenase 5 precursor - Lectin jacalin - Pentatricopeptide (PPR) repeat-containing protein - Beta-galactosidase BG1

Figures

Figure 3.1 Integrative workflow combining genomic dataset, metabolic profiling by LC-HRMS, mQTL mapping and functional annotation. Procedures with dotted outlines indicate works that were conducted and reported in the literature.

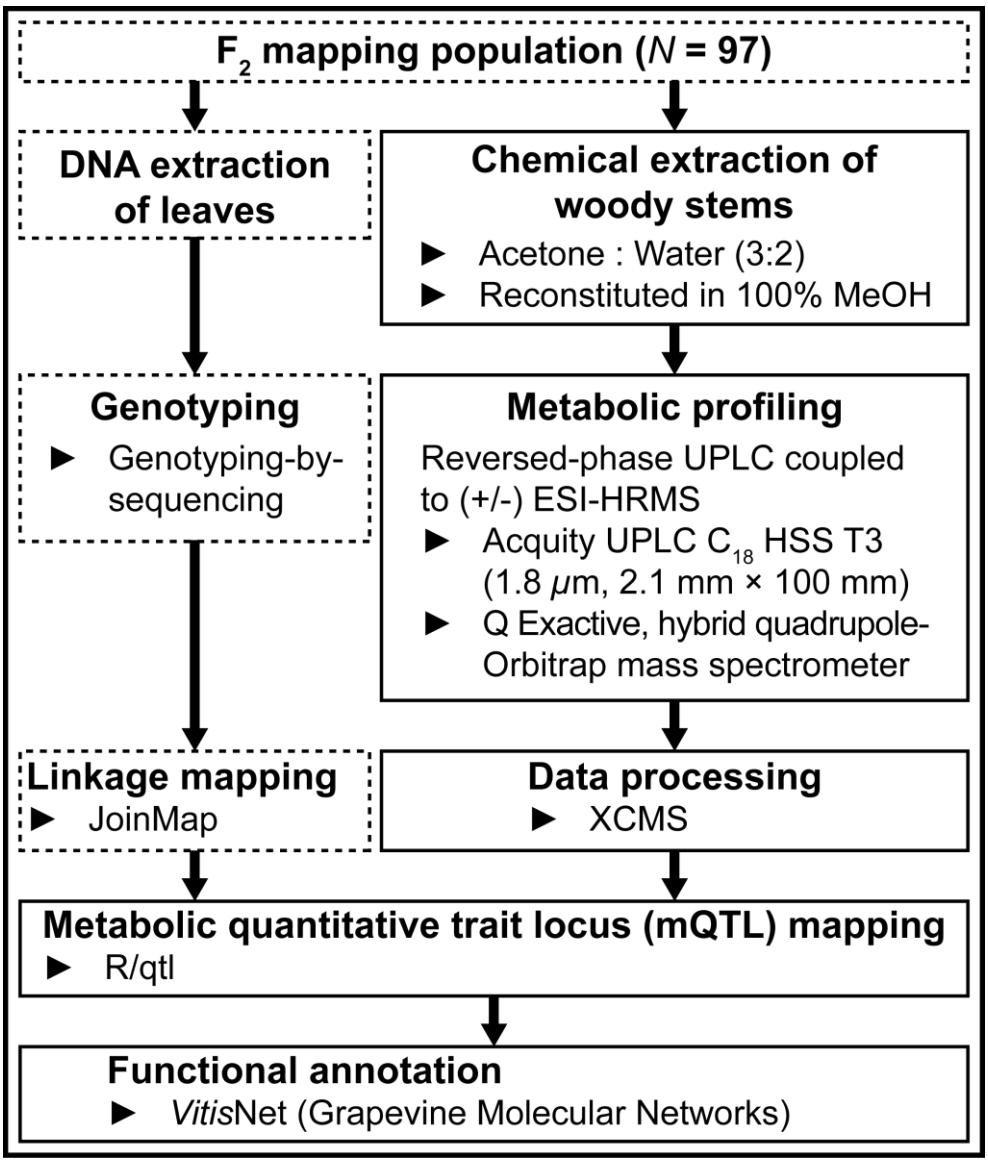


Figure 3.2 Total ion chromatograms for grape stem extracts of grandparent (F_0: *V. riparia* (USDA PI 588259)), parent (F_1: *V.* hybrid (16-9-2)), and a representative F_2 offspring that were analyzed by LC-HRMS.

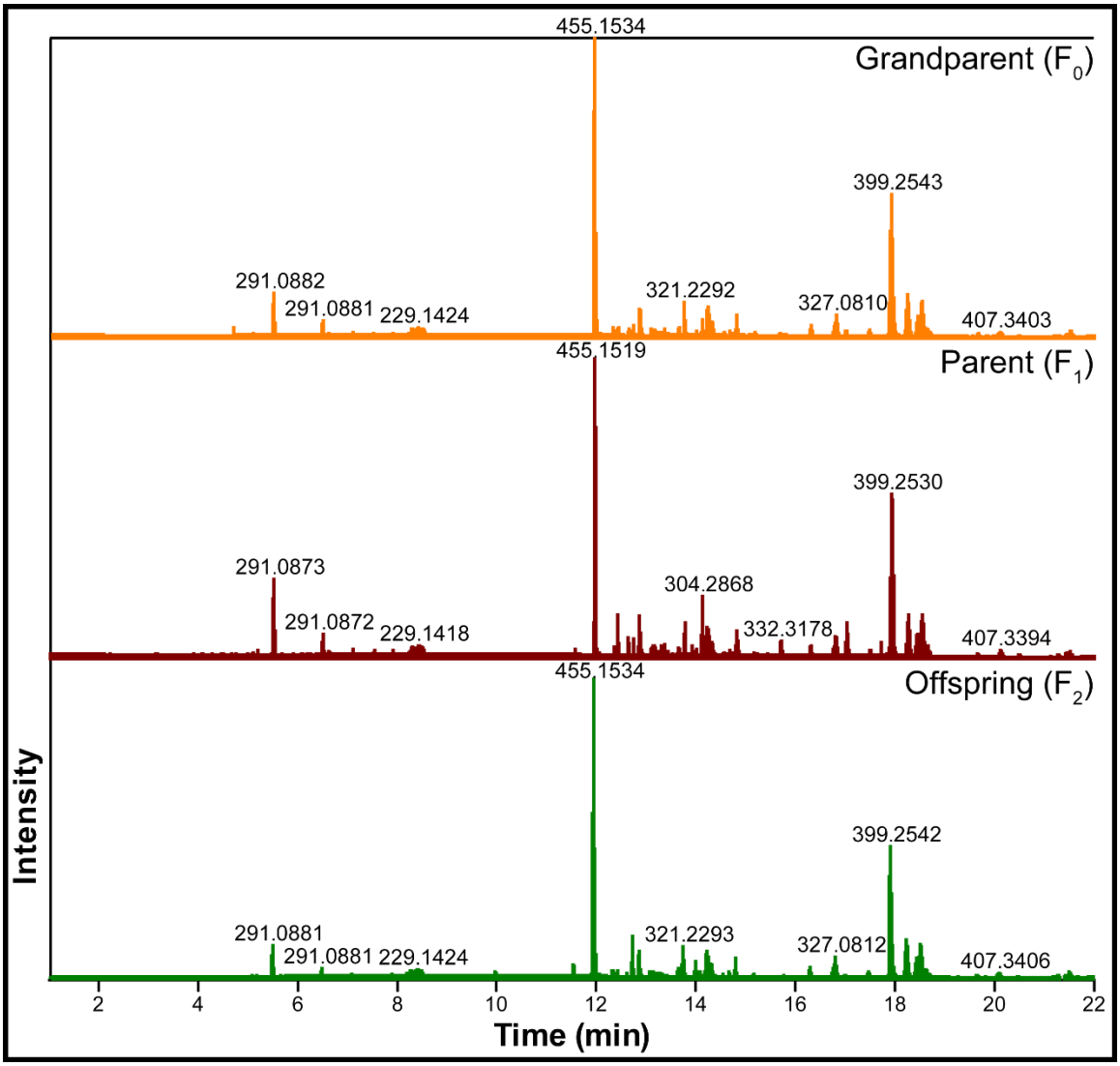


Figure 3.3 Extracted ion chromatograms of two representative offspring for different patterns of stilbene oligomers.

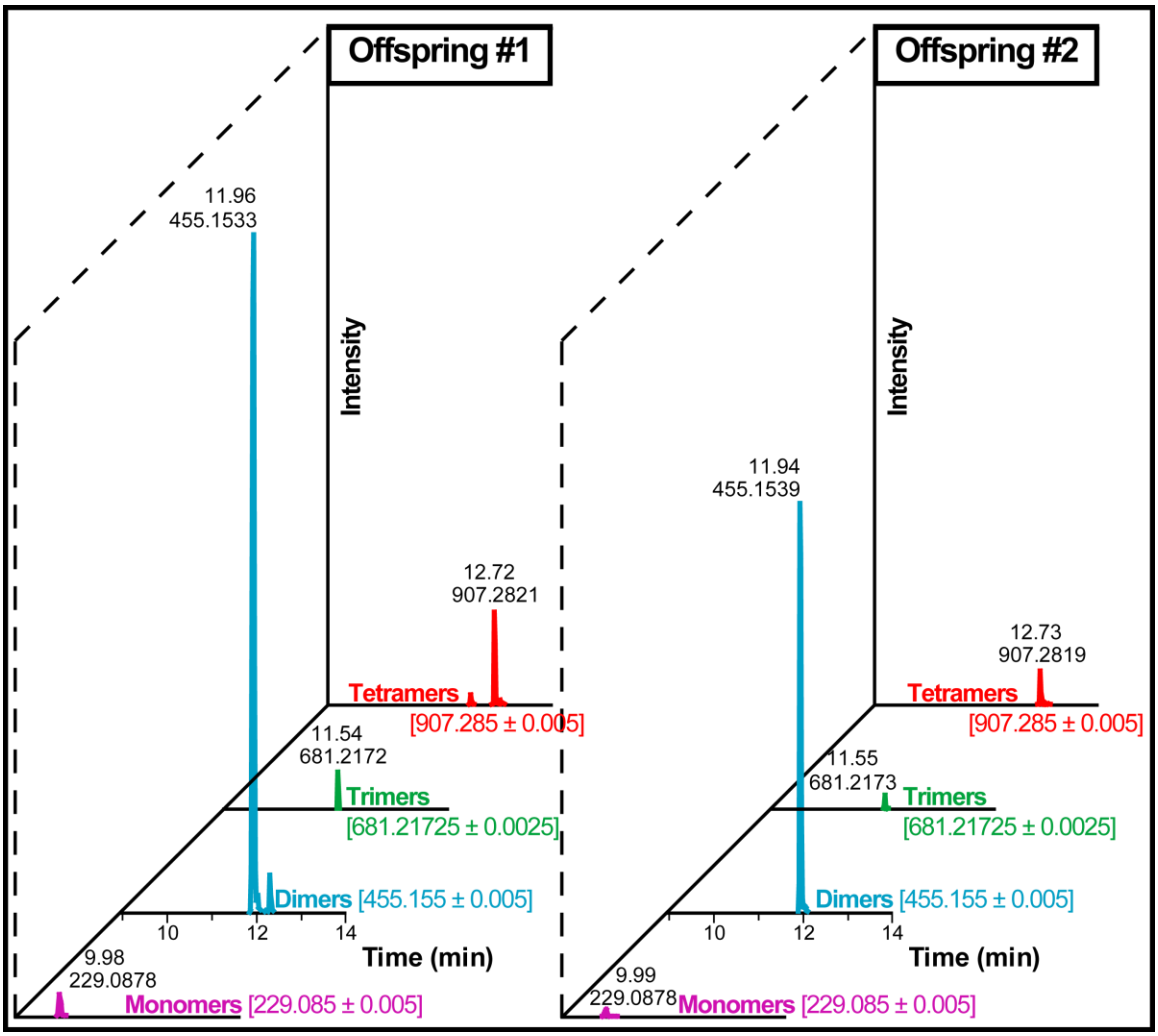
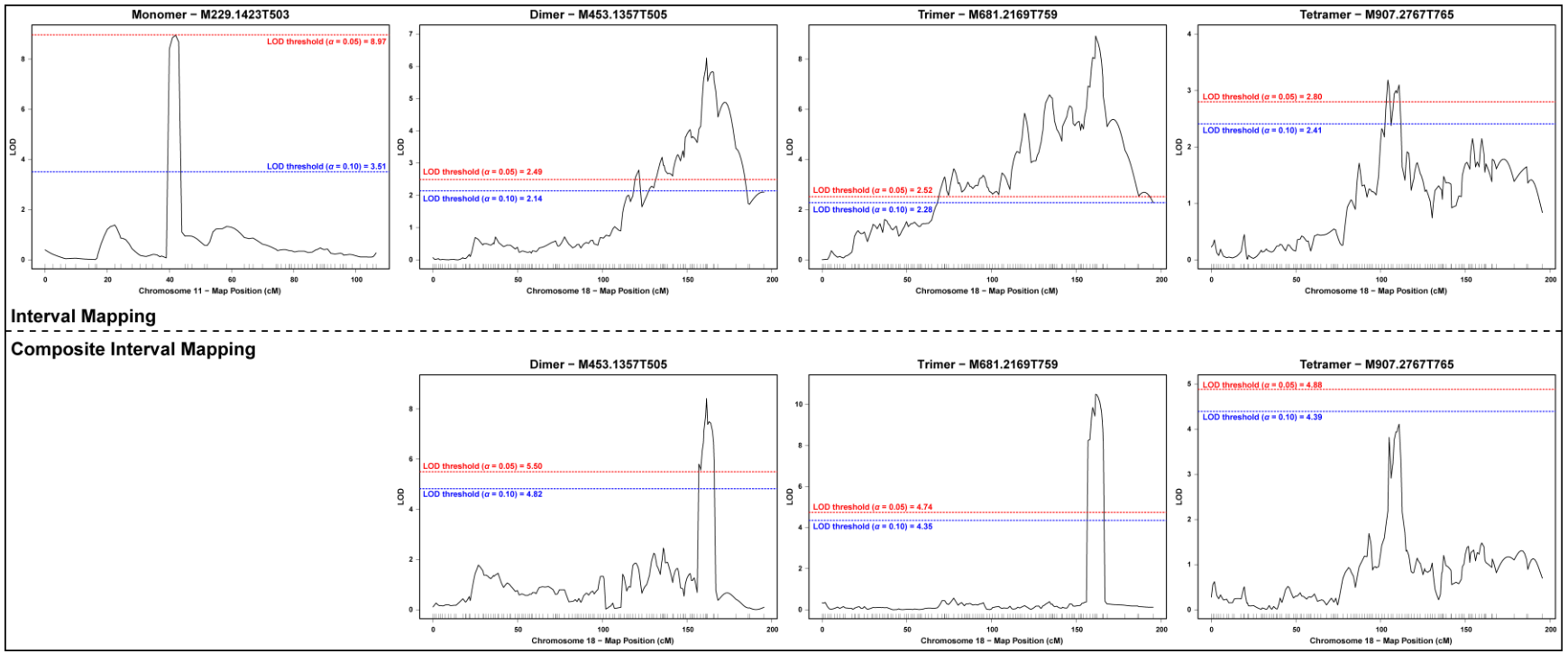


Figure 3.4 Detection of mQTLs for four features corresponding to different patterns of stilbene oligomerization. Each mQTL is shown for the chromosome where a marker-'trait' association was detected using two mapping analyses, interval mapping and composite interval mapping. The dotted lines indicate genome-wide LOD significance thresholds (α = 0.05 and α = 0.10) calculated using 1,000 permutations.



Chapter 4

Identification of Biomarkers Associated with Grapevine Powdery Mildew Resistance Loci in a Controlled Environment

Powdery mildew is a destructive fungal disease in cultivated grapevine, *Vitis vinifera*. Due to its economic importance to the grape industry worldwide, breeding for resistance through marker-assisted breeding is a promising alternative to frequent chemical applications. Previously, a study of families segregating for severity of powdery mildew infection at the University of Minnesota grape breeding program reported the identification of two resistance loci with additive effects accounting for over 30% phenotypic variation. Despite the importance of these resistance sources to the program, the associated metabolic changes to the plants after infection remain unknown. To shed light on this, a metabolomics experiment was conducted to monitor metabolic changes of propagated grapevines in a time-course response to *in vivo* inoculation. Additionally, multivariate analyses were employed to systematically identify biomarkers associated with resistance. From the study, 64 biomarkers were associated with the presence of two resistant haplotypes. In a separate analysis, 52 biomarkers were associated with the chromosome 2 resistant haplotype (henceforth, $hap+_{chr2}$), while 12 biomarkers were associated with the chromosome 15 resistant haplotype ($hap+_{chr15}$). In a time-course assessment of biomarkers, the discriminating metabolic changes distinguishing resistant and susceptible genotypes appeared to be occurring from 24 to 48 hours after

inoculation. Recommendations as well as reservations regarding the use of metabolic biomarkers in grapevine breeding are discussed.

4.1 Introduction

Powdery mildew is one of the most destructive fungal diseases of grapevine (*Vitis* spp.) worldwide (Gadoury et al., 2012). The pathogen *Erysiphe necator* was introduced from North America into Europe through importation of grape cuttings in the second half of the 19th century, resulting in the devastation of cultivated wine grapes and the wine industry. The vast majority of cultivated grapevine is *Vitis vinifera* L., a susceptible host to the fungus. Disease management strategies typically rely on frequent applications of fungicides that are not only harmful to humans and the environment, but are also increasingly cost-prohibitive due to chemical and labor expenditure (Christen et al., 2017; Rortais et al., 2017).

Alternatively, breeding for resistance is a promising strategy to cope with powdery mildew by introgressing elite cultivars with natural sources of resistance. DNA marker-assisted selection can expedite cultivar development for disease resistance by predicting the presence of resistant haplotype during parental and offspring selection.

However, DNA marker-assisted breeding (MAB) is a costly investment in crop breeding. For instance, an initial 5-year implementation of MAB is estimated to cost over $16 per seedling at the University of Minnesota apple breeding program (Luby & Shaw, 2001; Edge-Garza et al., 2015; Wannemuehler et al., 2018). In the case of perennial fruit breeding like grapevine, false positives,

arising from unreliable markers or a recombination between markers and the resistance gene, and the high heterozygosity of grapevine, would result in additional financial expenditures and space constraints due to the long juvenile period of grapevine. Hence, the use of genetic markers alone may be insufficient for a foolproof selection system.

Metabolomics can provide a second-layer solution to the hefty investment of DNA-based MAB. Since a metabolome closely correlates with the cellular phenotype, analyses of the metabolome may aid in understanding an organism's response to a perturbation that is otherwise undetected in a genomic experiment (Dettmer et al., 2007; Kuehnbaum & Britz-McKibbin, 2013; Aretz & Meierhofer, 2016). In addition, identification of biomarkers (i.e., chemical changes in response to a perturbation) can be harnessed as a metabolomics tool for MAB applications. Thus, the complementary use of genetic- and biochemical-based markers can provide an additional layer of verification in grapevine breeding.

While there have been ample reports of biomarker discovery, most studies have concentrated on staple food crops due to their economic importance and model plant species with relatively short generation cycle (Tarpley et al., 2005; Meyer et al., 2007; Deyanira et al., 2012; Degenkolbe et al., 2013; Marti et al., 2013). In rice (*Oryza sativa* L.), biomarkers associated with drought tolerance (Deyanira et al., 2012; Degenkolbe et al., 2013) and plant developmental stages (Tarpley et al., 2005) have been identified. In addition to abiotic-related biomarkers, there have also been works on biotic-related markers, albeit at a lower count. In particular, biomarkers associated with bacterial leaf blight and

Fusarium head blight have been reported in rice and barley, respectively (Sana et al., 2010; Bollina et al., 2011). Despite the reports of biomarkers relating to biotic and abiotic traits, there has not been any use of metabolomics to generate biomarkers for breeding purposes.

The research presented here follows from a previous genetic study that identified two powdery mildew resistance loci at the University of Minnesota grape breeding program (Teh et al., 2017). The resistance loci in two full-sibling populations were inherited from a common mother, MN1264, and originated from the maternal grandmother, 'Seyval blanc', which is an interspecific grape hybrid. The two loci exhibited additive effects that accounted for over 30% phenotypic variation in the populations studied and are important sources of resistance to the breeding program and the greater grape community. Despite their importance, the associated metabolic changes to the plants after infection remain unknown.

Our study attempted to dissect the metabolic changes that were associated to these resistance loci through the use of liquid chromatography-mass spectrometry. The high sensitivity of mass spectrometry in measuring dynamic changes of plant cells in response to perturbation necessitates that experimental plants be grown in a controlled environment, rather than being exposed to myriad uncontrolled field conditions. We monitored chemical ion (i.e., metabolic features) changes of asexually propagated grape plants in a controlled environment as a time-course response to *in vivo* inoculation with *E. necator*. The objectives of this work were to: 1) identify biomarkers associated with the

resistant haplotypes, 2) dissect time-dependent biomarkers during specific stages of the host defense, as well as 3) provide recommendations and reservations for the use of biochemical-based markers in breeding applications.

4.2 Materials and Methods

4.2.1 Plant Material

Using data from a study identifying powdery mildew resistance loci by Teh et al. (2017), nine individuals were selected based on the presence of resistant haplotypes on chromosomes 2 and 15 to identify biomarkers associated with these resistant haplotypes (Table 4.1). In December 2016, dormant grapevine stems were collected at the University of Minnesota Horticultural Research Center (44° 52′ 08.1″ N, 93° 38′ 17.3″ W).

Prior to propagation, dormant cane tissue was treated with a broad-spectrum biocide/fungicide. Briefly, two- to three-node cane segments of approximately six inches were submerged in a 4% solution of plant preservative mixture in an enclosed vacuum system for five minutes and vented; this process was repeated several times for thorough infiltration of the plant material with the biocide.

4.2.2 Propagation and Maintenance in a Controlled Environment (Main Chamber)

Sterilized cane segments were trimmed and treated with Hormex® rooting powder before placing them into plastic containers (diameter = 11 cm, height = 40 cm) with Burpee® Natural™ vermiculite and glass marbles at the bottom for

drainage. All reagents and materials were either autoclaved or surface sterilized with 1% (v/v) sodium hypochlorite, prior to experimental use.

Each container was loosely covered with a lid to prevent contamination and drying. The containers were transferred into a growth chamber and positioned using a randomized complete block design. The chamber settings were as follows: 18 °C, 50% humidity, 16-hour daylength, and illumination with cool white fluorescent of approximately 380 $\mu mol \cdot m^{-2} \cdot s^{-1}$ of photosynthetically active radiation (PAR), measured on the floor of the chamber. Each plant was watered every alternate day and supplemented with 40 mL of nutrient medium based on the vendor's recommended concentration every week (~4.49 g/L; Chee & Pool C2D *Vitis* medium).

4.2.3 Maintenance of Powdery Mildew Inoculum (Inoculum Chamber)

Grape powdery mildew is an obligate biotroph, which necessitated the use of susceptible plants to maintain the supply of inoculum for infection. A separate chamber was dedicated for this purpose. The chamber settings were adjusted to 25 °C, ambient humidity of approximately 50%, 16-hour daylength, and 230 $\mu mol \cdot m^{-2} \cdot s^{-1}$ PAR of cool white fluorescence.

4.2.4 *In vivo* Inoculation of Plants in a Controlled Environment (Inoculation Chamber)

Plants were routinely monitored for growth and nutrient requirements. Over the next four months, several plants were discarded due to root decay and contamination with gray mold (*Botrytis* spp.). The remaining plants were randomized and reorganized for the time-course inoculation experiment. Due to

the genotype-specificity growth, a plant was deemed ready for inoculation once it has reached a height of 34 – 38 cm with an estimated leaf count of 12 to 16.

First, these plants were transferred to the inoculation chamber, which had the same environmental settings as the “main chamber”. Next, several powdery mildew colonized leaves of seven- to ten-day-old were excised from the plants and shaken in a 20 mL of 0.001% Tween® 20 sterile water to prepare a conidial suspension, as previously described by Cadle-Davidson et al. (2016). Subsequently, the inoculum concentration was adjusted to 10^5 conidia/mL before spraying onto designated plants in the inoculation chamber using a handheld atomizer. Meanwhile, plants labeled uninoculated were sprayed with sterile water using a separate handheld atomizer containing no spores.

Leaf sampling was conducted at the designated time (i.e., hpi) for inoculated plants, or once the sprayed liquid had evaporated (approximately an hour later) for uninoculated plants. The fourth fully expanded leaves beneath the apex were detached, and two 1.1 cm diameter discs were excised with a cork borer and placed into a microcentrifuge tube on dry ice. Samples were stored at -80 °C until extraction.

4.2.5 Sample Preparation and Extraction

The samples were extracted according to the methods by Fiehn et al. (2008) with some modifications. Briefly, extraction was carried out in 400 µL -80 °C pre-chilled HPLC-grade methanol with two 1.6-mm tungsten carbide beads at approximately 80 g-force for 5 minutes using an automated tissue homogenizer, Geno/Grinder® 2010 (SPEX® SamplePrep, Metuchen, NJ). The

disrupted tissue samples and beads were subjected to centrifugation at 16,000 g-force for 2 minutes (Eppendorf® 5415C Centrifuge, Hamburg, Germany), and the supernatant was transferred into a pre-weighed glass vial and was evaporated to dryness. The dried extract was then reconstituted in methanol and adjusted to a concentration of 1 mg/mL (residue mass/volume methanol added).

4.2.6 Untargeted Metabolomics

Liquid chromatography-mass spectrometry (LC-MS) analyses were performed on a hybrid quadruple Orbitrap (Q Exactive mass spectrometer, Thermo Scientific, San Jose, CA) equipped with a Dionex UltiMate® 3000 HPLC (Thermo Fisher Scientific, Waltham, MA). Reconstituted extracts were diluted ten-fold, and 1 µL of each diluted extract was injected via an autosampler equipped with a cooled sample compartment (10 ˚C). LC was performed at a flow rate of 450 µL/min using a C_{18}-reverse-phase column (HSS T3, 2.1 mm inner diameter × 100 mm length, 1.8 µm particle size; Waters®, Milford, MA) using a gradient generated by mixing mobile-phase solvent A (water with 0.1% formic acid) and solvent B (acetonitrile with 0.1% formic acid) using the following time and % solvent B sequence: 0 to 1 min, 2% B; 1 to 10 min, 2 to 30% B; 10 to 12 min, 30 to 50% B; 12 to 20 min, 50 to 75% B; 20 to 22 min, 75 to 98% B; 22 to 23 min, 98 to 2% B; 23 to 27 min, 2% B. Water and formic acid were Optima™ HPLC-grade (Fisher Chemical, Waltham, MA); acetonitrile was HPLC-grade (Sigma-Aldrich, Saint Louis, MO). The mass spectrometer was operated in the positive/negative switching ionization mode over a scan range of m/z 150 -2000 at a resolution of 35,000. Predictive automatic gain control was employed with a

target value of 3.0×10^6 and a maximum injection time of 200 ms. The flow rate for sheath gas, auxiliary gas, and sweep gas was 60, 25, and 1, respectively. The capillary temperature was 350 °C, while the auxiliary gas heater temperature was 50 °C. The spray voltage was 3.6 kV, and the S-lens radio frequency level was set to 55. Samples were analyzed in a randomized order, including interspersed pooled samples for quality control and peak alignment purposes.

4.2.7 Data Processing

A general workflow for data processing was carried out using Progenesis QI software (Waters, Milford, MA) with several modifications. Here, the positive mode LC-MS profile data were imported for chromatographic alignment and feature extraction (i.e., peak picking). Data from the pooled samples were manually selected as references for chromatographic alignment. Feature extraction parameters were optimized manually for feature assignment to the largest number of compound ions. Chromatographic peak width was set to 0.02 min, based on the peak width of several features visualized on Xcalibur™ software version 2.1 (Thermo Scientific, Waltham, MA). Adduct forms were selected within the Progenesis peak picking parameter settings to include $[M+H]^+$, $[M+2H]^{2+}$, $[M+3H]^{3+}$, $[M+H-H_2O]^+$, $[M+H-2H_2O]^+$, $[2M+H]^+$, $[M+ACN+H]^+$, $[2M+ACN+H]^+$ and $[M+CH_3OH+H]^+$.

4.2.8 Data Analysis

All processed data were exported to an in-house implementation of Workflow4Metabolomics (Giacomoni et al., 2014) for data visualization and statistical analysis. Partial least squares (PLS) regression was conducted on

each genotype, particularly for those with more than two time points, to determine the relationship of the data with respect to time, which was the *Y*-vector. Additionally, orthogonal partial least squares-discriminant analysis (OPLS-DA) was carried out with inoculation as the *Y*-vector, to cluster variations attributed to inoculation. Because OPLS-DA modeling provided a geometrical rotation, the between-class variation was separated in the predictive component (i.e., inoculation) while the within-class variation was captured in the orthogonal component.

4.2.9 Scaling and Transformation

Multivariate analyses and projection models are sensitive to noise and artifacts, which are prevalent throughout metabolomics data. The data matrix was $\log_2$ transformed into a more Gaussian-type distribution to reduce the weight of outliers and skewed distribution. Throughout the analysis, features with a coefficient of variation (CV) within replicates exceeding 50% were discarded from further analysis.

4.2.10 Biomarker Discovery

Biomarker identification was accomplished using four methods. The genotypes used in this study were selected based on the presence of resistant haplotypes on chromosomes 2 and 15, henceforth denoted as hap+$_{chr2}$ and hap+$_{chr15}$. In addition, these genotypes represent three generations, thereby showing the inheritance of the resistant alleles from grandparent to parent, and from parent to offspring.

Method 1 examined features that were associated with the presence of two resistant haplotypes across genotypes spanning three generations. The correlation loading scores from three OPLS-DA analyses of resistant genotypes (i.e., 'Seyval blanc'$_{+/+}$ and MN1264$_{+/+}$, MN1264$_{+/+}$ and G53$_{+/+}$, 'Seyval blanc'$_{+/+}$ and G53$_{+/+}$) were tallied and compared with those of susceptible genotypes (i.e., MN1069$_{-/-}$ and MN1246$_{-/-}$, MN1246$_{-/-}$ and G33$_{-/-}$, MN1069$_{-/-}$ and G33$_{-/-}$). The summation of correlation scores alone could yield values that were skewed by one or two scores. Thus, these correlation loading scores were also ranked and tallied. The combination of unranked and ranked scores was used to select features that were: 1) positively correlated; and 2) negatively correlated between resistant and susceptible genotypes.

Method 2 examined features associated with resistant haplotypes within the four full-sibling individuals. An OPLS-DA analysis was conducted for the dataset of each offspring containing uninoculated (i.e., 0 hpi) and inoculated (i.e., 48 hpi) samples, where inoculation was the *Y*-vector. The importance or influence of variables in the OPLS-DA projection was estimated by the predictive and orthogonal variable influence on projection (VIP4P$_{OPLS}$ and VIP4O$_{OPLS}$, respectively). While the correlation loading scores span from a theoretical minimum of -1 to a maximum of 1, the predictive VIP4P$_{OPLS}$ is a non-negative metric that assesses the relevance of features in explaining variations correlated to the *Y*-vector, which in this experiment was inoculation. Though the cutoff value for predictive VIP4P$_{OPLS}$ can be arbitrary, a value between 0.83 and 1.21 has been recommended as a threshold for variable selection (Chong & Jun, 2005;

Mehmood et al., 2012). To complement the recommendation, a wider range of threshold values from 0.6 to 2.0 was tested (Figure 4.5). For each offspring, a list of features was retained at a specified $VIP4P_{OPLS}$. The differing lists of features retained for each offspring were used to map a four-way elliptic Venn diagram (Figure 4.2), where features were delineated into four major categories: 1) features unique to a genotype, 2) features shared by two genotypes, 3) features shared by three genotypes, and 4) features shared by all genotypes.

Method 3 relied on the use of data visualization to compare the outcome of two or more OPLS-DA models. Shared and unique structure (SUS) plots were used to portray the correlation profiles from two models in a single two-dimensional plot (Wiklund et al., 2008). In this study, an SUS plot was constructed to compare the correlation values of metabolic features between the doubly-resistant offspring (i.e., $G53_{+/+}$) and the doubly-susceptible offspring (i.e., $G33_{-/-}$). Features close to the diagonal (imaginary slope, m = 1 for $y = mx + b$) were positively correlated between the two offspring, while off-diagonal features were negatively correlated. Using the off-diagonal features, an additional round of an SUS plot was carried out on singly resistant offspring, namely $G18_{+/-}$ and $G81_{-/+}$ for resistant haplotype on chromosome 2 and 15, respectively, to delineate features associated with one resistant haplotype versus the other.

Finally, Method 4 examined temporal patterns in feature variation to identify biomarkers. The three combinations of time points were: 1) two genotypes with 0, 24, 48 and 96 hpi, 2) three genotypes with 0, 24 and 48 hpi, and 3) three genotypes with 0, 48 and 96 hpi. For each combination, *k*-means

clustering was conducted to partition features that exhibited the same patterns over time. Results were reported as a ratio that was normalized with *t* = 0 hpi, over time. This analysis omitted the use of ‘Cabernet Sauvignon’ due to its genetic dissimilarity from the rest of the genotypes.

4.3 Results

4.3.1 Double-resistant-haplotype Biomarkers across Generations

The *Vitis* hybrids used in this study and summarized in Table 4.1 include three genotypes with a single copy of the resistant haplotype at each locus: ‘Seyval blanc’$_{+/+}$ (grandparent), MN1264$_{+/+}$ (parent), and G53$_{+/+}$ (offspring), and three genotypes with susceptible haplotypes at both loci: MN1069$_{–/–}$ (grandparent) MN1246$_{–/–}$ (parent), and G33$_{–/–}$ (offspring). In the first biomarker discovery approach (using *Method 1*), OPLS-DA was used to identify features that changed during the course of infection for each resistant and susceptible genotype. Highly-correlated features (in the top 1% of total correlation scores) were then compared between these lists from susceptible and resistant genotypes to identify biomarkers. Based on the selection, 36 features were identified that co-varied in both resistant and susceptible individuals in response to inoculation. As highlighted in Figure 4.1, the levels of these features increased at 48 hours post inoculation (hpi). Meanwhile, 41 biomarkers were found to decrease significantly at 48 hpi in all resistant genotypes, but remained constant or increased marginally in all susceptible genotypes. On the other hand, 23

biomarkers were exhibited the opposite trend, where the levels decreased significantly in all susceptible genotypes, but remained constant in all resistant genotypes (Table 4.2).

4.3.2 Haplotype-Specific Biomarkers

To identify biomarkers associated with the presence of resistant haplotypes for genotypes with similar genetic backgrounds (*Method 2*), we focused on metabolic changes occurring during infection (between t = 0 hpi and 48 hpi) in four full-sibling offspring ($G33_{-/-}$, $G53_{+/+}$. $G18_{+/-}$, and $G81_{-/+}$). From the OPLS-DA analysis results of each genotype, a four-way elliptic Venn diagram was constructed to visualize features that were genotype-exclusive or were shared between/among genotypes, based on different $VIP4P_{OPLS}$ thresholds (Figure 4.5). The optimal cutoff value in this experiment was 1.2, which was within the recommended range of 0.83 to 1.21 (Chong & Jun, 2005; Mehmood et al., 2012). Among the 15 segments on the diagram shown in Figure 4.2, two segments of particular interest were: 1) region shared between $G18_{+/-}$ and $G53_{+/+}$, as well as 2) region shared between $G53_{+/+}$ and $G81_{-/+}$. The former region indicates features changing in response to inoculation that are associated with chromosome 2 resistant haplotype, while the latter region indicates those associated with chromosome 15 resistant haplotype. In the former segment, 401 features were identified, while the latter segment contained 608 features. Within each list, an additional step was carried out to determine features that were positively correlated between the pairs of genotypes. Subsequently, 52 biomarkers were identified to be associated with chromosome 2 resistant

haplotype, while 12 biomarkers were detected to be associated with chromosome 15 resistant haplotype (Table 4.3a).

Another biomarker discovery approach (*Method 3*) utilized a series of SUS plot visualizations, which began with the use of correlation scores of $G33_{-/-}$ and $G53_{+/+}$ for their common features (N = 844) at $VIP4P_{OPLS}$ = 1.2 (Figure 4.3). 746 features resided along the diagonal (imaginary slope, m = 1 for $y = mx + b$), while 98 features were situated off-diagonal. Then, correlation scores of $G18_{+/-}$ and $G81_{-/+}$ for the off-diagonal features (N = 98) at $VIP4P_{OPLS}$ = 1.2 were used for the next round of SUS plotting, yielding 46 off-diagonal features, 17 along-diagonal features and 35 features that resided in the proximity of the origin. After removing features with within-replicate CV exceeding 50%, 20 off-diagonal features, 7 along-diagonal features, and 22 close-to-origin/axes features were identified (Table 4.3b).

4.3.3 Temporal Progression of Biomarkers

The identification of biomarkers as a function of temporal response (*Method 4*) was conducted using *k*-means clustering and reported as a ratio normalized with t = 0 hpi. Here, the clustering of features is reported in separate sections based on the combination of hpi and the haplotype-associated features (Figure 4.4). As indicated above, features associated with the resistant haplotype on chromosome 2 and that of chromosome 15 are denoted as $hap+_{chr2}$ and $hap+_{chr15}$, respectively.

a) $MN1264_{+/+}$ and $G81_{-/+}$ for $hap+_{chr2}$ at 0, 24, 48 and 96 hpi

Two main clustering patterns were observed here. In the first cluster containing nine features, the levels of inoculated MN1264$_{+/+}$ samples were consistently higher than those of G81$_{–/+}$, over time. In another cluster containing 4 features, the levels of both MN1264$_{+/+}$ G81$_{–/+}$ declined sharply at 24 hpi, but increased gradually from 24 to 96 hpi.

b) MN1264$_{+/+}$ and G81$_{–/+}$ for hap+$_{chr15}$ at 0, 24, 48 and 96 hpi

With the exception of two features, all features fell into two distinct clusters. In one cluster, the levels of features in inoculated MN1264$_{+/+}$ samples were consistently higher than uninoculated samples over time, while the reverse trend was observed in G81$_{–/+}$. Conversely, in the other cluster, the levels of features in inoculated G81$_{–/+}$ were consistently higher than uninoculated G81$_{–/+}$ over time, while the reversal was observed in MN1264$_{+/+}$. In general, the divergence of features increased rapidly from 0 to 48 hpi, and then declined from 48 to 96 hpi.

c) MN1264$_{+/+}$, G53$_{+/+}$ and G81$_{–/+}$ for hap+$_{chr2}$ at 0, 24 and 48 hpi

A sizable cluster of 28 features increased significantly more in inoculated G53$_{+/+}$ compared to inoculated G81$_{–/+}$ and MN1264$_{+/+}$. Within this cluster, G53$_{+/+}$ either showed a continuous steep elevation throughout, or a gradual incline followed with a steep increase. In addition, G81$_{–/+}$ in some features exhibited an initial decline from 0 to 24 hpi, before a gradual increase from 24 to 48 hpi.

d) MN1264$_{+/+}$, G53$_{+/+}$ and G81$_{–/+}$ for hap+$_{chr15}$ at 0, 24 and 48 hpi

Two clustering patterns were observed here. In one cluster, the levels of features in inoculated MN1264$_{+/+}$ increased over time, while the opposite trend

was observed in the two offspring, G53$_{+/+}$ and G81$_{–/+}$. In another cluster, the levels of inoculated G53$_{+/+}$ and MN1264$_{+/+}$ increased gradually over time, while the levels of G81$_{–/+}$ exhibited an initial spike exceeding 40× times from 0 to 24 hpi, but sharply declined from 24 to 48 hpi.

e) MN1264$_{+/+}$, G33$_{–/–}$ and G81$_{–/+}$ for hap+$_{chr2}$ at 0, 48 and 96 hpi

Two main clusters were captured. Features in one cluster showed a rapid upward spike that was followed by a gradual decline in the doubly resistant MN1264$_{+/+}$, which was in stark contrast to the susceptible G33$_{–/–}$; meanwhile, the response of singly resistant G81$_{–/+}$ was intermediate. Conversely, features in another cluster showed an increase in the levels of G33$_{–/–}$ that was followed by a decline, while G81$_{–/+}$ and MN1264$_{+/+}$ exhibited an initial decline, or maintained relatively constant levels throughout.

f) MN1264$_{+/+}$, G33$_{–/–}$ and G81$_{–/+}$ for hap+$_{chr15}$ at 0, 48 and 96 hpi

A distinct clustering pattern was observed where the levels of inoculated G33$_{–/–}$ declined after inoculation, while the levels of inoculated G81$_{–/+}$ and MN1264$_{+/+}$ were elevated after infection.

4.4 Discussion

To date, the application of metabolite-based biomarkers in fruit crops has focused mainly in aroma, flavor and nutritional traits, but has been limited in disease resistance for breeding purposes (Chambers et al., 2014; Liu et al., 2016). Prior to the application of biomarkers for disease resistance breeding, it is important to first establish an understanding of the plant and the

pathological/physiological traits of interest. With respect to this study, the interaction between grapevine and *E. necator* is still poorly understood, especially with regards to metabolic effects. As a follow-up study to our previous genetic identification of powdery mildew resistance loci (Teh et al., 2017), we have initiated a metabolomics approach to reveal metabolic differences between resistant and susceptible genotypes in response to *E. necator* inoculation, in order to identify biomarkers for use in MAB.

4.4.1 Maintenance of Plants in a Controlled Environment

The use of highly sensitive instruments to measure dynamic changes of plant cells in response to infection requires experimental plants to be grown in a controlled environment. Maintenance of propagated grapevine plants in chambers does, however, present several challenges. In preliminary studies, where plants were grown on trays in a chamber, inadvertent cross-infestation with leaf tunneling insects and powdery mildew were significant issues. Therefore, the experiment was designed to minimize contamination and to prevent localized contaminants from spreading to neighboring plants by placing each individual plant in a tall (40 cm) plastic container. During the initial three months of maintaining propagated plants of the same developmental stage from different genotypes, approximately 60% of the plants were retained in the study with the remainder being culled due to health concerns (mostly root decay and some gray mold (*Botrytis* spp.) contamination).

4.4.2 Metabolic Changes in Response to *in vivo* Inoculation

Recently, a metabolomics study of grape defense response to downy mildew identified 53 biomarkers that were modulated in response to infection (Chitarrini et al., 2017). However, this, along with virtually all studies of grape fungal diseases have relied on artificial inoculation of excised leaf tissues, instead of *in vivo* whole-plant inoculation. Metabolic changes arising from physical injury of tissue excision may confound metabolic effects attributed solely to the host-pathogen interaction. Additionally, physical wounding may aid the pathogen in colonizing the host through the sites of injury. Hence, this experiment was designed to conduct *in vivo* inoculation prior to leaf excision, instead of the conventional reversed way.

4.4.3 Discovery of Biomarkers Associated with Two Resistant Haplotypes

The first set of biomarkers identified was specific to 0 and 48 hpi comparing the metabolic responses to inoculation of three fully resistant vs. three susceptible genotypes. Although grapevine response to fungal diseases is genotype-specific, 48 hpi was chosen as a common time point for all genotypes. As reported in previous studies, a metabolic response is most pronounced around the 48-hpi mark, which is preceded by a transcriptional response at 12 to 24 hpi (Pezet et al., 2004a; Pezet et al., 2004b; Fung et al., 2008; Gadoury et al., 2012).

Here, 36 features were detected at 48 hpi when their levels were highly elevated in both resistant and susceptible individuals (Figure 4.1; Table 4.2). These features are likely metabolites involved in a general recognition of the

fungal pathogen-associated molecular patterns (PAMP), but may not be discriminating with regards to a resistance response. Meanwhile, 64 features were observed for a differential response between hap+$_{chr2}$/hap+$_{chr15}$ versus hap–$_{chr2}$/ hap–$_{chr15}$ genotypes. From the list, 23 biomarkers were identified where their levels remained relatively constant in the resistant genotypes, but were significantly reduced in all susceptible genotypes at 48 hpi, indicating a plausible role of resistance during the initial stage of inoculation. In susceptible individuals, these compounds were likely converted into compounds with low phytotoxicity. Alternatively, the other 41 biomarkers exhibited somewhat constant levels in susceptible plants compared to a significant decline in resistant plants at 48 hpi, which may indicate that these metabolites serve as precursors to potentially phytotoxic molecules to facilitate a resistant host response.

4.4.4 Discovery of Haplotype-Specific Biomarkers

Despite the identification of these 64 biomarkers, the epistatic interaction between the two resistant haplotypes was unresolved due to the comparison of genotypes with hap+$_{chr2}$/hap+$_{chr15}$ versus hap–$_{chr2}$/ hap–$_{chr15}$. Here, the dissection of haplotype-specific biomarkers was focused on four full-sibling offspring, differing at the presence of resistant haplotypes, among other genetic dissimilarities. From the four-way elliptic Venn diagram (Figure 4.2), the segment shared exclusively by G18$_{+/–}$ and G53$_{+/+}$ was of interest due to their common hap+$_{chr2}$. The segment contains features with a significant change in levels after inoculation. However, these changes could be negatively correlated between the two genotypes. By selecting features that were positively correlated between

them, 52 features were chosen as candidate biomarkers that were associated with hap+$_{chr2}$. Meanwhile, 12 features were selected as candidate biomarkers associated with hap+$_{chr15}$, from a segment shared exclusively by G53$_{+/+}$ and G81$_{-/+}$. Biomarkers in segment hap+$_{chr2}$ likely correspond to metabolites that are different from those of segment hap+$_{chr15}$.

The *post hoc* analysis of the four-way Venn diagram provided several important validations. For example, the aforementioned 36 features (see the first set of biomarkers) fall in the segment shared by all genotypes, where the levels changed significantly from 0 to 48 hpi. Another segment of interest was the region shared between G18$_{+/-}$ and G81$_{-/+}$. Features in this region likely correspond to metabolites sharing the same resistance pathway in the metabolic response to powdery mildew infection.

Another way to dissect haplotype-specific biomarkers relied on a set of Shared and Unique Structure plots (Figure 4.3; Table 4.3b), which resulted in two groups of features: 1) features negatively correlated between G33$_{-/-}$ and G53$_{+/+}$ but positively correlated between G18$_{+/-}$ and G81$_{-/+}$; and 2) features negatively correlated between G33$_{-/-}$ and G53$_{+/+}$ and negatively correlated between G18$_{+/-}$ and G81$_{-/+}$. Features in the first group may correspond to metabolites sharing the same pathway. Meanwhile, features in the second group may operate in different pathways in response to infection.

Throughout the analysis of haplotype-specific biomarkers, a noteworthy caveat is that there is a reasonable likelihood of a biomarker being unrelated to a resistant haplotype but correlated, by chance, because an unrelated feature level

change was common to two genotypes but not observed in the other two genotypes. This is important in this case because of the genetic similarity of these four offspring. Unlike mutant lines with target insertion/deletion site(s), these four offspring have the same parents and will have inherited similar genetic information except where it has been altered through recombination.

4.4.5 Temporal Progression of Biomarkers

Thus far, 64 haplotype-specific biomarkers have been reported. However, a multiple time-course response of biomarkers has not been investigated, which is important for linking metabolic and physiological events to the course of the infection or a resistance response. In the initial stage of infection, following inoculation and germination of conidiospores, hyphae of *E. necator* pierce the host cuticle and epidermal cell wall, leading to an establishment of intimate contact between the pathogen haustoria and the host plasma membrane (Gadoury et al., 2012). During the early stages of host-pathogen interaction, metabolites that are identified as differential biomarkers that increase in abundance in resistant genotypes may play pivotal roles, such as interfering with the pathogen invasion of the cell membranes or the initial establishment of haustoria.

In this study, all individuals were sampled at t = 0 hpi and 48 hpi (Table 4.1). Some individuals were sampled at additional time points, including 24 hpi and 96 hpi. To conduct a comparison for the temporal response of biomarkers, groups of individuals sharing common sampling time points were compared. First, the analysis of biomarkers specific to all time points was shared by two

groups of individuals, namely MN1264$_{+/+}$ and G81$_{–/+}$. For features associated with hap+$_{chr15}$, a change in levels was most pronounced at 24 hpi for both genotypes, likely indicating that these metabolites were sensitive to pathogen invasion. It is also worth noting that G81$_{–/+}$ exhibited a steeper change, compared to MN1264$_{+/+}$, suggesting that G81$_{–/+}$ relied solely on defense compounds associated with hap+$_{chr15}$, while MN1264 had the luxury of using both defense mechanisms associated with hap+$_{chr2}$ and hap+$_{chr15}$.

The second analysis of biomarkers is specific to 0, 24 and 48 hpi for three genotypes, namely MN1264$_{+/+}$, G53$_{+/+}$, and G81$_{–/+}$. For features associated with hap+$_{chr2}$, G53$_{+/+}$ exhibited an increase much more significantly than MN1264$_{+/+}$ and G81$_{–/+}$ at 48 hpi, suggesting that these biomarkers were being accumulated due to their phytotoxicity. However, although MN1264$_{+/+}$ share the same resistant haplotypes with G53$_{+/+}$, it exhibited patterns resembling G81$_{–/+}$ rather than G53$_{+/+}$. On the other hand, for features associated with hap+$_{chr15}$, G81$_{–/+}$ showed an upward spike of 40-400× immediately after infection, while genotypes with two resistant haplotypes remained relatively constant. It is plausible that G81$_{–/+}$ contains an additional resistance gene that is not attributed to hap+$_{chr2}$ and hap+$_{chr15}$.

Finally, the third analysis of biomarkers is specific to 0, 48 and 96 hpi for three genotypes, namely MN1264$_{+/+}$, G33$_{–/–}$, and G81$_{–/+}$. In one set of features that is associated with hap+$_{chr2}$, the levels of G33$_{–/–}$ were elevated at 48 hpi, which was in stark contrast to MN1264$_{+/+}$, while the levels of G81$_{–/+}$ were intermediate. This suggests that most of the metabolic effort for mounting a

defense against pathogen infection occurred by 48 hpi, before returning to pre-inoculation levels.

4.4.6 Recommendations and Reservations for Breeding Applications

Taken together, we identified groups of biomarkers that were: 1) associated with hap+$_{chr2}$ and hap+$_{chr15}$ (*Method 1*; Table 4.2), 2) haplotype-specific (*Methods 2 and 3*; Table 4.3), and 3) time-course dependent (*Method 4*; Figure 4.4). After selecting features with a coefficient of variation below 25% a total of 148 of the most reliable and consistent biomarkers associated with *E. necator* resistance in grapevine have been identified. The stringency was imposed due to the importance of a reliable biomarker for breeding applications.

In classical plant breeding, three of the main merits of a desirable molecular marker are: 1) polymorphic, 2) highly heritable, and 3) logistically favorable. In the case of a biochemical-based marker (termed biomarker), it is highly polymorphic due to its quantitative nature in chemical analysis. However, a biomarker that is elicited in an inducible response may not be highly heritable due to the influence of the environment. Nonetheless, measuring a biomarker in response to a perturbation in a controlled environment can be used as a signature for biomarker-assisted selection. Although biomarkers are polymorphic and can be highly heritable, the economic feasibility for biomarker use is unfavorable at the present time, until the development of a targeted workflow for routine biomarker analysis. Based on the current fee structure of the Center for Mass Spectrometry and Proteomics at the University of Minnesota (https://cbs.umn.edu/cmsp/newfees), the hourly fee for quantitative metabolomics

is quoted at $94.60, excluding the time, labor and financial expenditure for sample extraction and data analysis. As the use of metabolomics becomes more ubiquitous, we may witness a significant price drop, as was seen in the price of genomic sequencing over the past 10 years.

4.5 Conclusions

In this study, we carried out a metabolomics experiment to reveal changes in the metabolome of resistant and susceptible genotypes in a time-course response to *in vivo* inoculation, in order to detect biomarkers attributed to resistant haplotypes on chromosomes 2 and 15. We demonstrated the use of various multivariate analyses to systematically identify biomarkers. In addition, the study was aimed at unraveling the effects of haplotype-specific biomarkers and understanding the temporal response of biomarkers.

The discovery of biomarkers is precursory to structural characterization of these candidate biomarkers so that compound structure, and ultimately, compound identity can be elucidated. Additionally, subsequent experiments can focus on studying the metabolic changes of primary and secondary metabolites in a time-dependent manner to characterize the host response to *E. necator*.

Tables

Table 4.1 Summary of genotypes collected in a controlled environment. Inoculation in the chamber was carried out in a time-course manner, where hpi denotes hours post inoculation.

Genotype	hap_{chr2}	hap_{chr15}	Generation/Remarks	Replicates for Time-course Inoculation			
				0 hpi	24 hpi	48 hpi	96 hpi
'Seyval blanc'	+	+	Resistant grandparent	4	0	4	0
MN1069	–	–	Susceptible grandparent	5	0	5	0
MN1264	+	+	Resistant parent	5	5	5	4
MN1246	–	–	Susceptible parent	4	0	5	0
G18	+	–	$hap+_{chr2}$ offspring	4	0	4	0
G33	–	–	Susceptible offspring	6	0	8	5
G53	+	+	Resistant offspring	6	6	5	0
G81	–	+	$hap+_{chr15}$ offspring	4	4	5	4
'Cabernet Sauvignon'	–	–	Susceptible check	6	6	5	0

Table 4.2 Biomarkers associated with two resistant haplotypes (*Method 1*).

Biomarker	Time (min)	*m/z*	Adduct Form	Associated with
T01.66M291.1045	1.66	291.1045	$[M+H]^+$	
T01.77M268.1037	1.77	268.1037	$[M+H]^+$	
T02.77M196.1550	2.77	196.1550	$[M+H]^+$	
T03.43M335.1306	3.43	335.1306	$[M+H]^+$	
T04.28M374.2012	4.28	374.2012	$[M+H]^+$	
T04.28M379.1566	4.28	379.1566	$[M+H]^+$	
T04.90M418.2280	4.90	418.2280	$[M+H]^+$	
T04.90M423.1835	4.90	423.1835	$[M+H]^+$	
T04.90M611.1579	4.90	611.1579	$[M+H]^+$	
T06.05M183.1014	6.05	183.1014	$[M+H]^+$	
T07.09M773.1877	7.09	773.1877	$[M+H]^+$	
T07.44M604.3552	7.44	604.3552	$[M+H]^+$	
T07.52M757.1906	7.52	757.1906	$[M+H]^+$	
T07.54M398.0747	7.54	398.0747	$[2M+H]^+$	
T07.73M313.6709	7.73	313.6709	$[2M+H]^+$	
T07.73M624.2662	7.73	313.1404	$[2M+H]^+$	
T08.17M309.1304	8.17	309.1304	$[M+H]^+$	
T08.27M241.1048	8.27	241.1048	$[M+H]^+$	
T08.42M379.1794	8.42	379.1794	$[2M+H]^+$	
T08.48M691.3970	8.48	692.4026	$[M+H]^+$	
T08.55M487.7919	8.55	487.7919	$[2M+H]^+$	Resistant genotypes
T08.59M151.0247	8.59	151.0247	$[M+H]^+$	
T12.79M219.2107	12.79	219.2107	$[M+H]^+$	
T13.43M185.1156	13.43	185.1156	$[M+H]^+$	
T13.72M297.1837	13.72	297.1837	$[2M+H]^+$	
T13.72M311.2203	13.72	311.2203	$[M+H]^+$	
T14.46M392.1740	14.46	392.1740	$[M+H]^+$	
T14.50M284.2936	14.50	284.2936	$[M+H]^+$	
T14.50M304.3106	14.50	304.3106	$[M+H]^+$	
T14.60M353.2673	14.60	353.2673	$[M+H]^+$	
T15.09M296.1923	15.09	296.1923	$[M+H]^+$	
T17.44M787.4588	17.44	787.4588	$[M+H]^+$	
T17.52M286.2730	17.52	286.2730	$[M+H]^+$	
T18.12M295.1873	18.12	295.1873	$[M+H]^+$	
T18.66M238.6374	18.66	238.6374	$[M+H]^+$	
T19.17M627.4655	19.17	627.4655	$[M+H]^+$	
T19.81M597.4566	19.81	597.4566	$[M+H]^+$	
T21.14M255.6449	21.14	255.6449	$[M+H]^+$	
T22.46M548.2278	22.46	548.2278	$[2M+H]^+$	
T23.24M182.1304	23.24	365.2645	$[M+H]^+$	
T23.93M734.5135	23.93	734.5135	$[M+H]^+$	

T08.61M401.7229	8.61	401.7229	$[2M+H]^+$	Susceptible genotypes
T08.61M762.4268	8.61	382.2207	$[2M+H]^+$	
T08.79M806.4492	8.79	404.2319	$[2M+H]^+$	
T08.95M443.2724	8.95	443.2724	$[2M+H]^+$	
T11.47M240.1957	11.47	240.1957	$[M+H]^+$	
T11.48M257.2221	11.48	257.2221	$[M+H]^+$	
T14.05M376.2581	14.05	376.2581	$[M+H]^+$	
T14.05M394.2888	14.05	394.2888	$[M+H]^+$	
T18.29M524.3723	18.29	524.3723	$[M+H]^+$	
T18.43M334.3094	18.43	334.3094	$[M+H]^+$	
T18.83M313.2606	18.83	278.2468	$[M+H]^+$	
T19.09M278.2468	19.09	278.2468	$[M+H]^+$	
T19.64M336.3249	19.64	336.3249	$[M+H]^+$	
T20.08M394.2804	20.08	394.2804	$[M+H]^+$	
T20.51M352.3197	20.51	352.3197	$[M+H]^+$	
T20.53M306.2783	20.53	306.2783	$[M+H]^+$	
T20.80M418.2861	20.80	418.2861	$[M+H]^+$	
T22.51M334.3090	22.51	334.3090	$[M+H]^+$	
T23.05M222.1463	23.05	222.1463	$[M+H]^+$	
T23.31M290.2083	23.31	290.2083	$[M+H]^+$	
T23.31M318.2906	23.31	319.2977	$[M+H]^+$	
T23.96M495.3298	23.96	495.3298	$[M+H]^+$	
T23.90M377.3026	23.90	377.3026	$[M+H]^+$	

Table 4.3 Haplotype-specific biomarkers from: a) OPLS-DA followed by a four-way elliptic Venn diagram (*Method 2*); and b) OPLS-DA followed by SUS analysis (*Method 3*).

(A) Method 2					(B) Method 3			
Biomarker	**Time (min)**	***m/z***	**Adduct Form**	**Associated with**	**Biomarker**	**Time (min)**	***m/z***	**Adduct Form**
T02.24M221.0352	2.24	221.0352	$[M+H]^+$		T04.22M595.1457	4.22	595.1457	$[M+H]^+$
T04.16M399.8410	4.16	399.8410	$[M+H]^+$		T05.38M359.8182	5.38	359.8182	$[M+H]^+$
T05.78M340.9137	5.78	340.9137	$[M+H]^+$		T05.48M221.0349	5.48	221.0349	$[M+H]^+$
T07.09M773.1877	7.09	773.1877	$[M+H]^+$		T06.38M396.9019	6.38	396.9019	$[3M+H]^+$
T12.73M465.8601	12.73	465.8601	$[M+H]^+$		T06.66M289.0701	6.66	289.0701	$[M+H]^+$
T13.09M465.8600	13.09	465.8600	$[M+H]^+$		T06.66M459.0928	6.66	459.0928	$[M+H]^+$
T13.32M307.2249	13.32	307.2249	$[M+H]^+$		T07.54M525.2901	7.54	525.2901	$[2M+H]^+$
T13.34M323.1824	13.34	323.1824	$[M+H]^+$		T07.54M526.2934	7.54	526.2934	$[M+H]^+$
T13.49M172.1698	13.49	172.1698	$[M+H]^+$		T07.62M272.9225	7.62	272.9225	$[2M+H]^+$
T13.53M264.1930	13.53	264.1930	$[M+H]^+$		T08.11M477.0325	8.11	477.0325	$[M+H]^+$
T14.88M303.1939	14.88	303.1939	$[M+H]^+$		T08.40M415.2563	8.40	415.2563	$[2M+H]^+$
T14.99M282.1690	14.99	282.1690	$[M+H]^+$		T08.55M469.1833	8.55	469.1833	$[M+H]^+$
T17.03M528.1328	17.03	528.1328	$[M+H]^+$		T09.01M404.8167	9.01	404.8167	$[M+H]^+$
T17.93M383.2030	17.93	383.2030	$[M+H]^+$		T09.24M413.8814	9.24	413.8814	$[M+H]^+$
T18.23M242.1018	18.23	242.1018	$[M+H]^+$	$hap+_{chr2}$	T09.62M297.9189	9.62	297.9189	$[M+H]^+$
T18.63M542.1489	18.63	542.1489	$[M+H]^+$		T12.75M186.0092	12.75	186.0092	$[M+H]^+$
T18.65M547.1268	18.65	547.1268	$[M+H]^+$		T14.32M437.1931	14.32	437.1931	$[M+H]^+$
T18.83M242.1027	18.83	242.1027	$[M+H]^+$		T14.32M446.2533	14.32	446.2533	$[M+H]^+$
T19.21M499.2279	19.21	499.2279	$[M+H]^+$		T14.32M460.2695	14.32	460.2695	$[M+H]^+$
T19.36M517.3720	19.36	517.3720	$[2M+H]^+$		T14.32M432.2374	14.32	432.2374	$[M+H]^+$
T19.90M581.1089	19.90	581.1089	$[M+H]^+$		T14.32M415.2112	14.32	415.2112	$[M+H]^+$
T20.47M242.1033	20.47	242.1033	$[M+H]^+$		T17.68M323.2196	17.68	323.2196	$[M+H]^+$
T21.93M528.7512	21.93	528.7512	$[M+H]^+$		T19.88M540.1394	19.88	540.1394	$[M+H]^+$
T22.09M337.2336	22.09	337.2336	$[M+H]^+$		T21.97M437.7206	21.97	437.7206	$[M+H]^+$
T22.21M803.3300	22.21	803.3300	$[M+H]^+$		T22.46M548.2278	22.46	548.2278	$[2M+H]^+$
T22.23M438.5021	22.23	438.5021	$[M+H]^+$		T23.01M310.1164	23.01	621.266	$[M+H]^+$
T22.30M253.1423	22.30	253.1423	$[M+H]^+$		T23.74M374.2582	23.74	374.2582	$[M+H]^+$
T22.58M184.9184	22.58	184.9184	$[M+H]^+$					
T22.70M279.1331	22.70	279.1331	$[M+H]^+$					

T22.75M935.3892	22.75	935.3892	$[M+H]^+$	
T22.89M430.1744	22.89	431.1817	$[M+H]^+$	
T23.07M233.0783	23.07	233.0783	$[M+H]^+$	
T23.10M267.1333	23.10	267.1333	$[M+H]^+$	
T23.17M407.1614	23.17	407.1614	$[M+H]^+$	
T23.17M539.2206	23.17	539.2206	$[M+H]^+$	
T23.17M935.3899	23.17	935.3899	$[M+H]^+$	
T23.17M671.2761	23.17	671.2761	$[M+H]^+$	
T23.51M233.0784	23.51	233.0784	$[M+H]^+$	
T23.53M267.1335	23.53	267.1335	$[M+H]^+$	
T23.53M279.1342	23.53	279.1342	$[M+H]^+$	
T23.62M249.1099	23.62	249.1099	$[M+H]^+$	$hap+_{chr2}$
T23.63M261.1088	23.63	261.1088	$[M+H]^+$	
T23.69M205.1194	23.69	205.1194	$[M+H]^+$	
T23.71M199.1321	23.71	164.1182	$[M+H]^+$	
T23.74M1331.5764	23.74	1331.5764	$[M+H]^+$	
T23.74M1128.3146	23.74	1128.3146	$[M+H]^+$	
T23.78M215.1254	23.78	215.1254	$[M+H]^+$	
T23.79M527.3877	23.79	527.3877	$[2M+H]^+$	
T23.84M265.1173	23.84	265.1173	$[M+H]^+$	
T23.84M265.1407	23.84	265.1407	$[M+H]^+$	
T23.86M251.1253	23.86	251.1253	$[M+H]^+$	
T23.86M263.1248	23.86	263.1248	$[M+H]^+$	
T08.27M504.7664	8.27	504.7664	$[2M+H]^+$	
T09.78M243.1011	9.78	243.1011	$[M+H]^+$	
T09.78M509.8037	9.78	509.8037	$[2M+H]^+$	
T10.55M243.1011	10.55	243.1011	$[M+H]^+$	
T12.97M315.1581	12.97	315.1581	$[M+H]^+$	
T13.17M193.1591	13.17	193.1591	$[3M+H]^+$	$hap+_{chr15}$
T13.28M257.1623	13.28	257.1623	$[M+H]^+$	
T14.54M535.8479	14.54	568.8547	$[M+H]^+$	
T15.22M274.1586	15.22	274.1586	$[M+H]^+$	
T18.49M590.4262	18.49	590.4262	$[M+H]^+$	
T20.93M441.8812	20.93	442.8885	$[M+H]^+$	
T21.75M477.2239	21.75	477.2239	$[M+H]^+$	

Figures

Figure 4.1 Biomarkers associated with dual resistant haplotypes, comparing the metabolic response of resistant and susceptible individuals. Represented here is a subset of 10 biomarkers from the full list of biomarkers in Table 4.2. 5 green lines correspond to biomarkers identified from pairwise OPLS-DA where resistant individuals exhibited a decline in metabolic levels at 48 hpi, while red lines correspond to those where susceptible individuals showed a decline at 48 hpi.

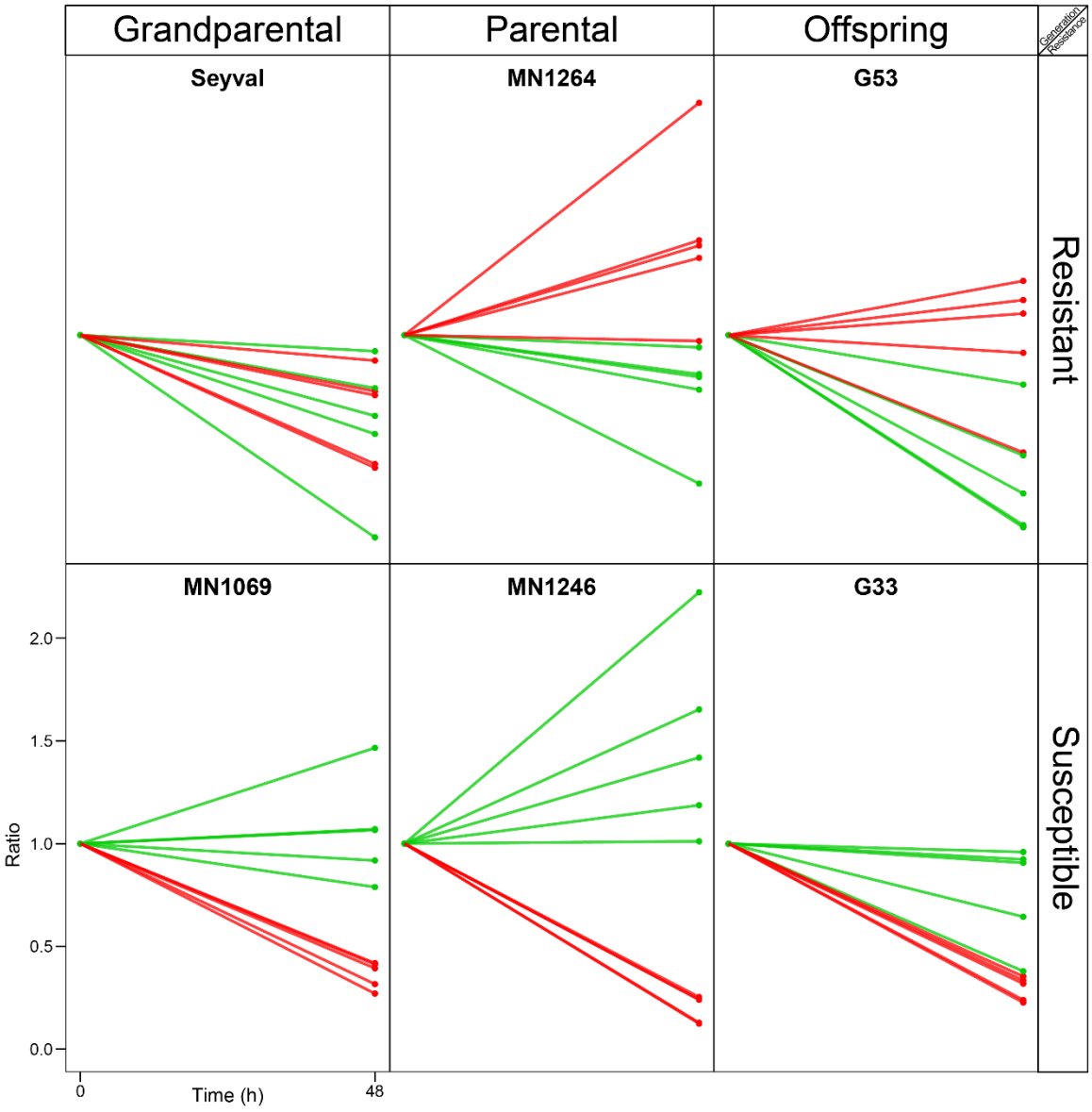


Figure 4.2 Summary of four-way elliptic Venn diagram constructed in *post hoc* OPLS-DA analyses for each of the four offspring differing for haplotype combinations.

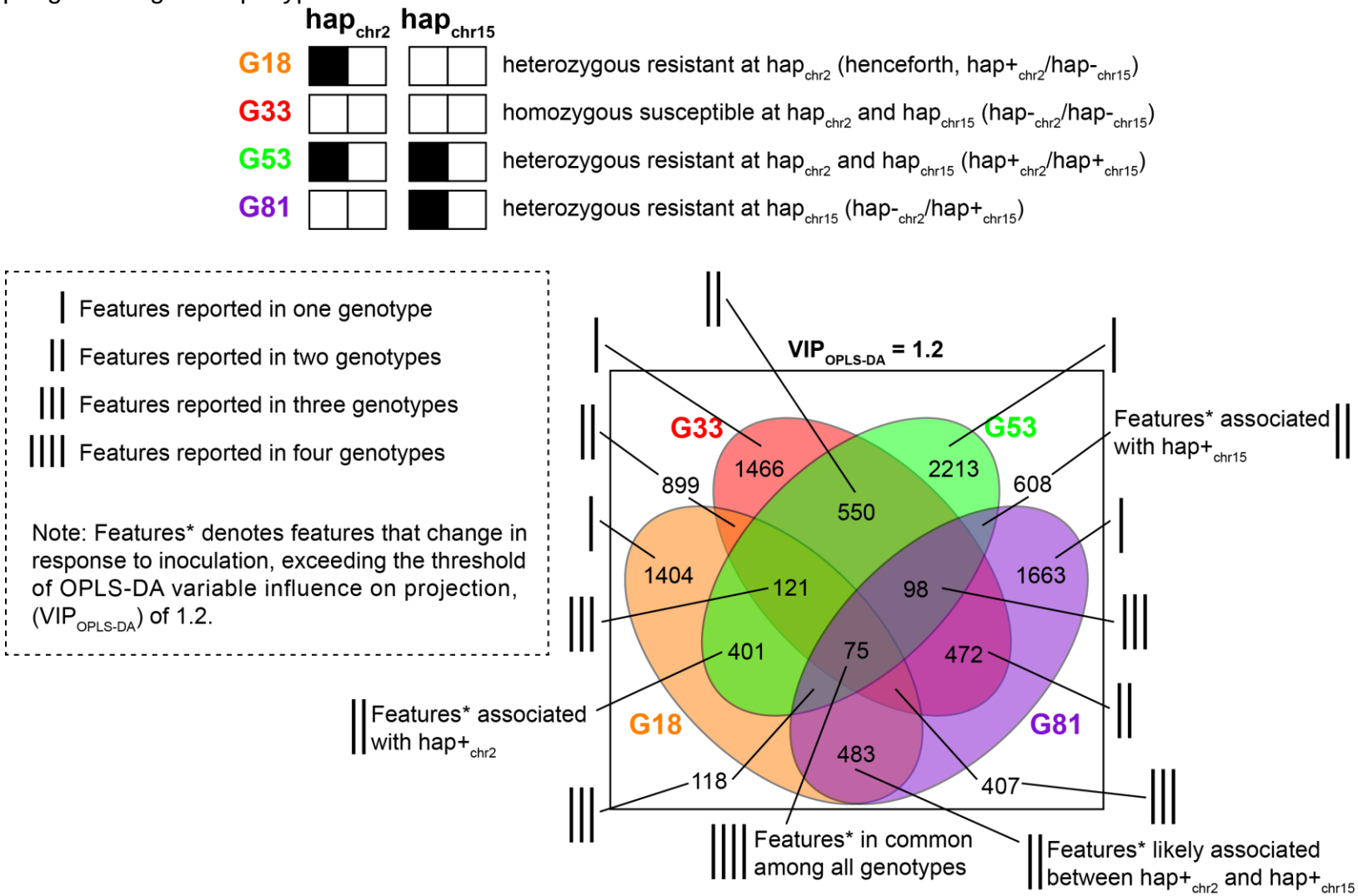


Figure 4.3 A series of two sequential shared and unique structure (SUS) plots to uncover biomarkers. Each point corresponds to a metabolic feature. Red points and points residing inside the red-border rectangles were discarded. The remaining black points in the bottom SUS plot marked as biomarkers.

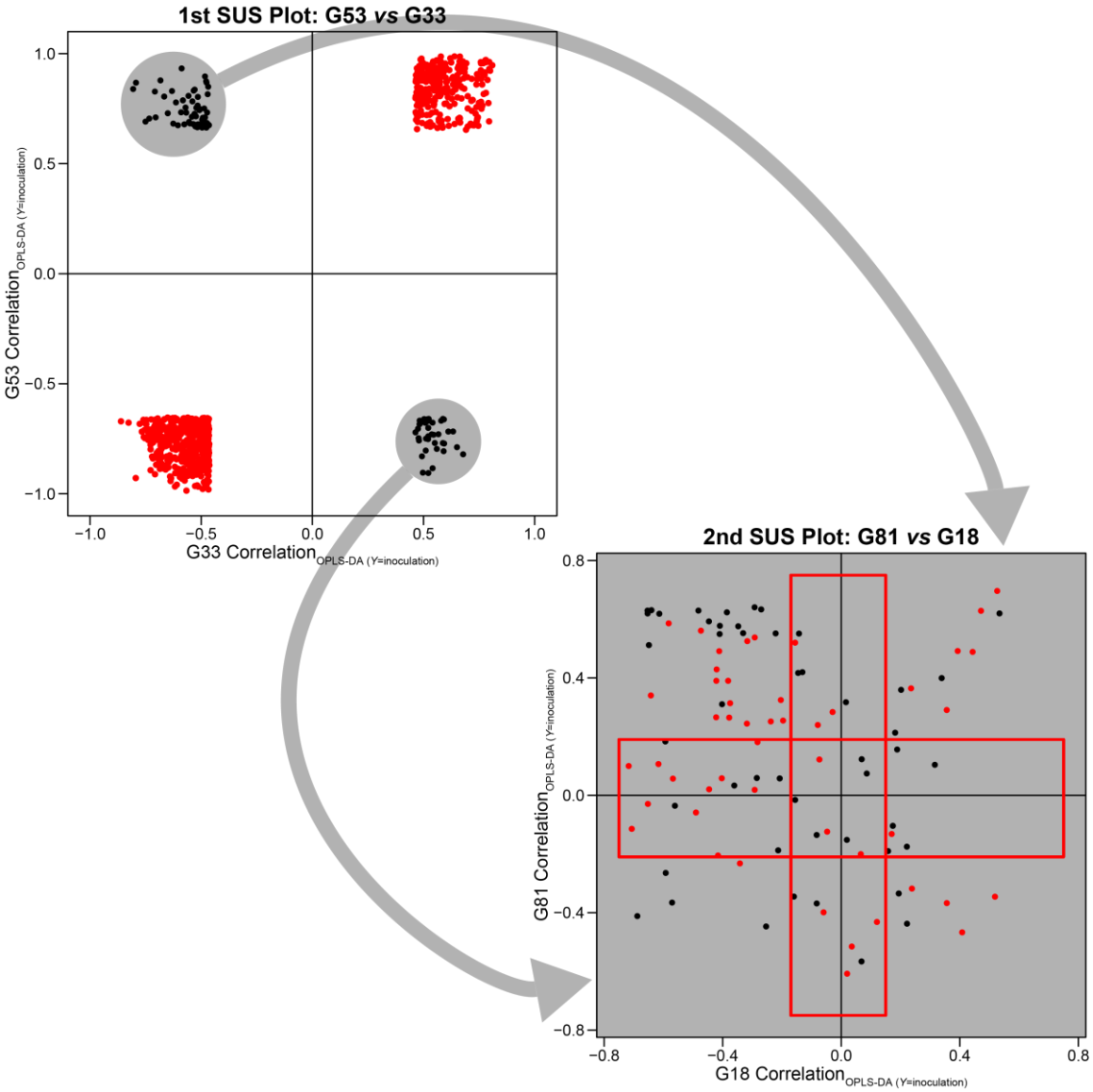


Figure 4.4 Temporal progression for a subset of *k-means*-clustered representative features based on genotypes with common sampling time points. Identical alphanumeric designations refer to the same feature of different genotypes. (A & B) Two clusters of features for MN1264$_{+/+}$, G33$_{-/-}$, and G81$_{-/+}$ at 0, 48 and 96 hpi; (C) One cluster of features for MN1264$_{+/+}$, G53$_{+/+}$, and G81$_{-/+}$ at 0, 24 and 48 hpi; (D & E) Two clusters of features for MN1264$_{+/+}$ and G81$_{-/+}$ at four time points.

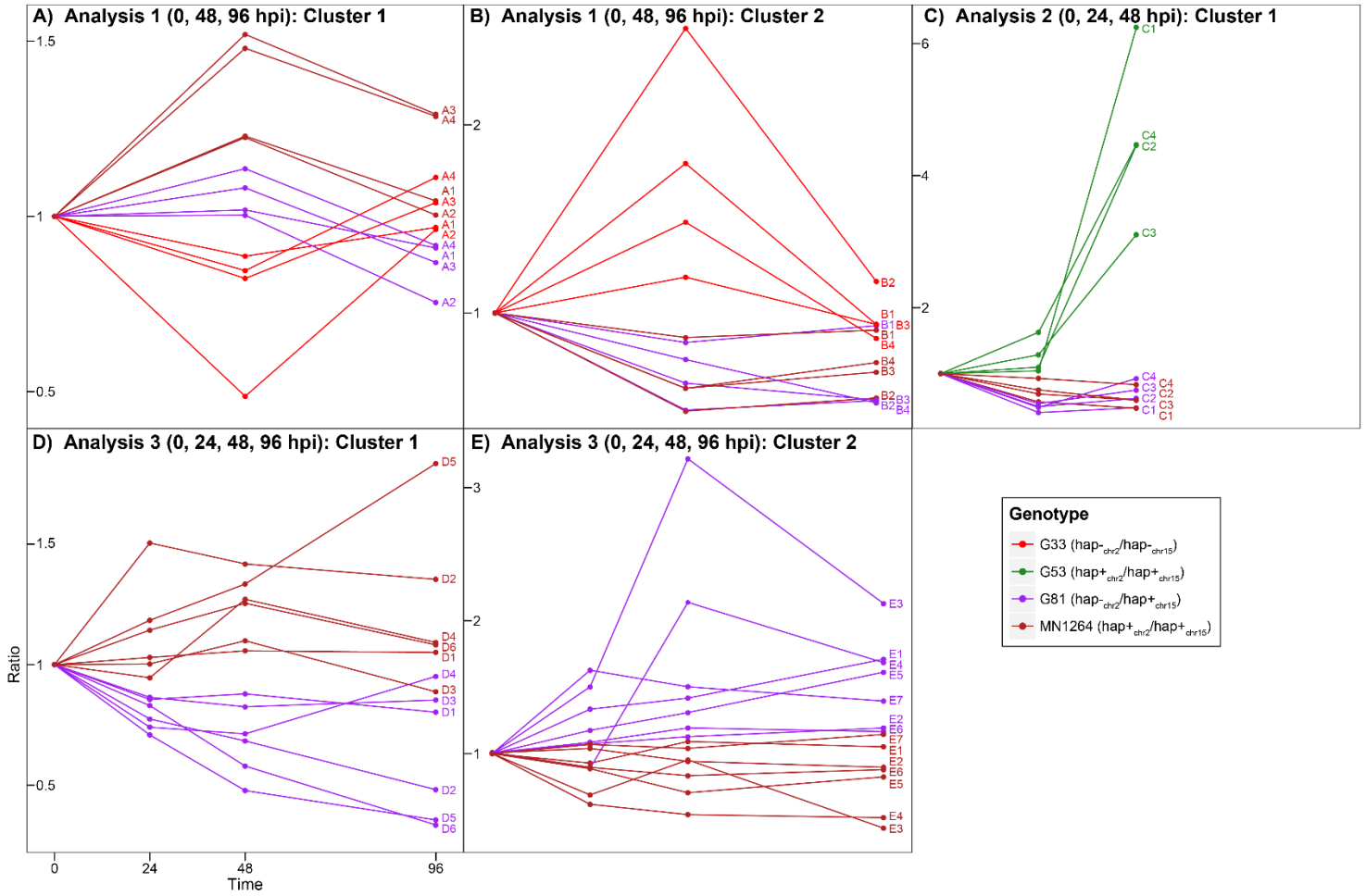
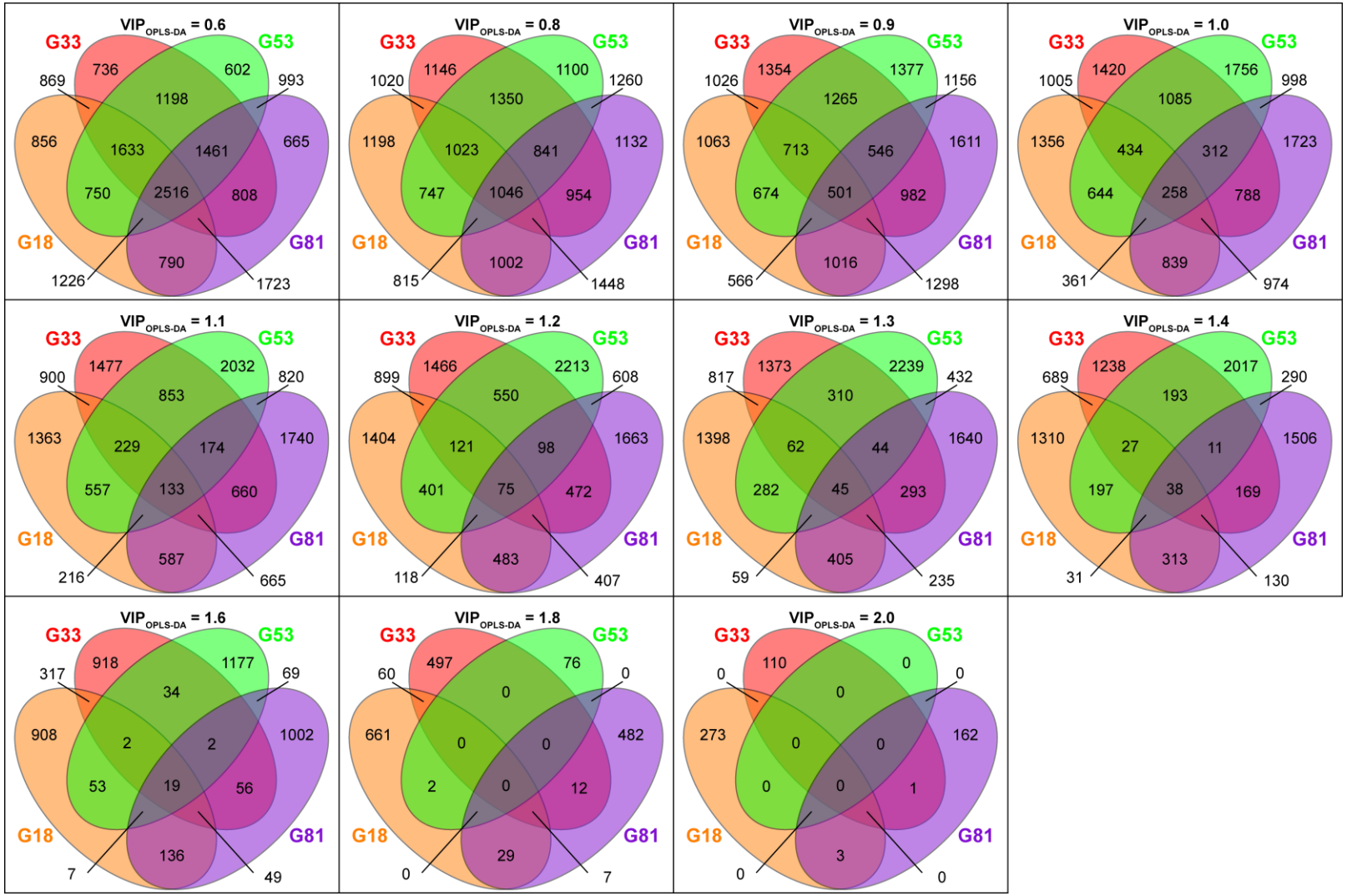


Figure 4.5 Four-way elliptic Venn diagrams constructed from a range of VIP4P$_{OPLS}$ to determine the suitability of a threshold value for feature selection.



Chapter 5

Conclusions

Powdery mildew is one of the most economically important diseases in cultivated grapevine worldwide. Instead of relying on frequent fungicide treatments for disease management, breeding for resistance is a promising alternative to manage powdery mildew by introgressing natural sources of resistance into elite cultivars. Marker-assisted selection (MAS) can expedite the development of resistant cultivars by predicting the presence of resistant haplotype during selection. However, MAS is a costly investment in perennial crop breeding, with an estimated cost of $16 per apple seedling during the initial five-year implementation (Luby & Shaw, 2001; Edge-Garza et al., 2015; Wannemuehler et al., 2018). Additionally, false positives during selection would result in additional financial expenditures and space constraints due to the long juvenility of grapevine. Hence, the use of markers discovered through genomic and genetic analyses alone may be insufficient for a foolproof MAS. Metabolomics can complement genomics as a second-layer solution for MAS. The combination of genomics and metabolomics can provide a systems biology understanding of a plant metabolic response to a perturbation, and can lead to biomarker discovery for future breeding applications.

To shed lights on the systems biology of grapevine resistance to powdery mildew, this dissertation is parsed into three sections: 1) genetic mapping, 2) metabolomic mapping, and 3) biomarker discovery (Figure 5).

In the genetic mapping study, two pedigree-connected F_1 mapping families were employed to identify resistance loci. The field phenotypic evaluation of powdery mildew severity over two years, coupled with the construction of high-quality SNP-based maps, enabled the identification of two resistance loci on chromosomes 2 and 15 with additive effects accounting for over 30% phenotypic variation. The construction of marker haplotypes, hap+$_{chr2}$ and hap+$_{chr15}$, traced the inheritance of the resistance loci from parent MN1264, and grandparent 'Seyval blanc'. In addition, two microsatellite markers, UDV-015b and VViv67, were identified to be associated with hap+$_{chr15}$, and can be recommended for MAS applications.

In the metabolomic mapping study, the genetic basis of a class of phytoalexins called stilbenes was characterized. The use of LC-HRMS for metabolic profiling of grapevine stems and a previously assembled F_2 map facilitated the discovery of four metabolic QTLs: one on chromosome 11 and three on chromosome 18. The integration of mQTL results with a functional annotation database, *Vitis*Net revealed a 200-kbp genomic hotspot in the vicinity of the chromosome 18 mQTL containing a suite of disease resistance motifs.

In the biomarker discovery study, a metabolomic experiment was carried out on whole-plant propagated grapes in a temporal response to *in vivo* inoculation to characterize metabolic changes attributed to resistant haplotypes during fungal infection. The use of statistical multivariate analyses systematically identified 52 biomarkers that were associated with hap+$_{chr2}$, and 12 biomarkers with hap+$_{chr15}$. In a temporal assessment of biomarkers, the discriminating

metabolic changes distinguishing resistant and susceptible individuals appeared to be occurring from 24 to 48 hours after inoculation.

Taken together, the studies presented here demonstrated the complementation of genomics and metabolomics, as well as their combined utility in characterizing the genetic and metabolic responses of grapevine to powdery mildew (Figure 5.1). Additionally, the discoveries of genetic markers and biochemical-based markers (termed biomarkers) can be applied as a two-layer marker-assisted selection, especially when the cost of routine metabolomic analysis becomes economically feasible.

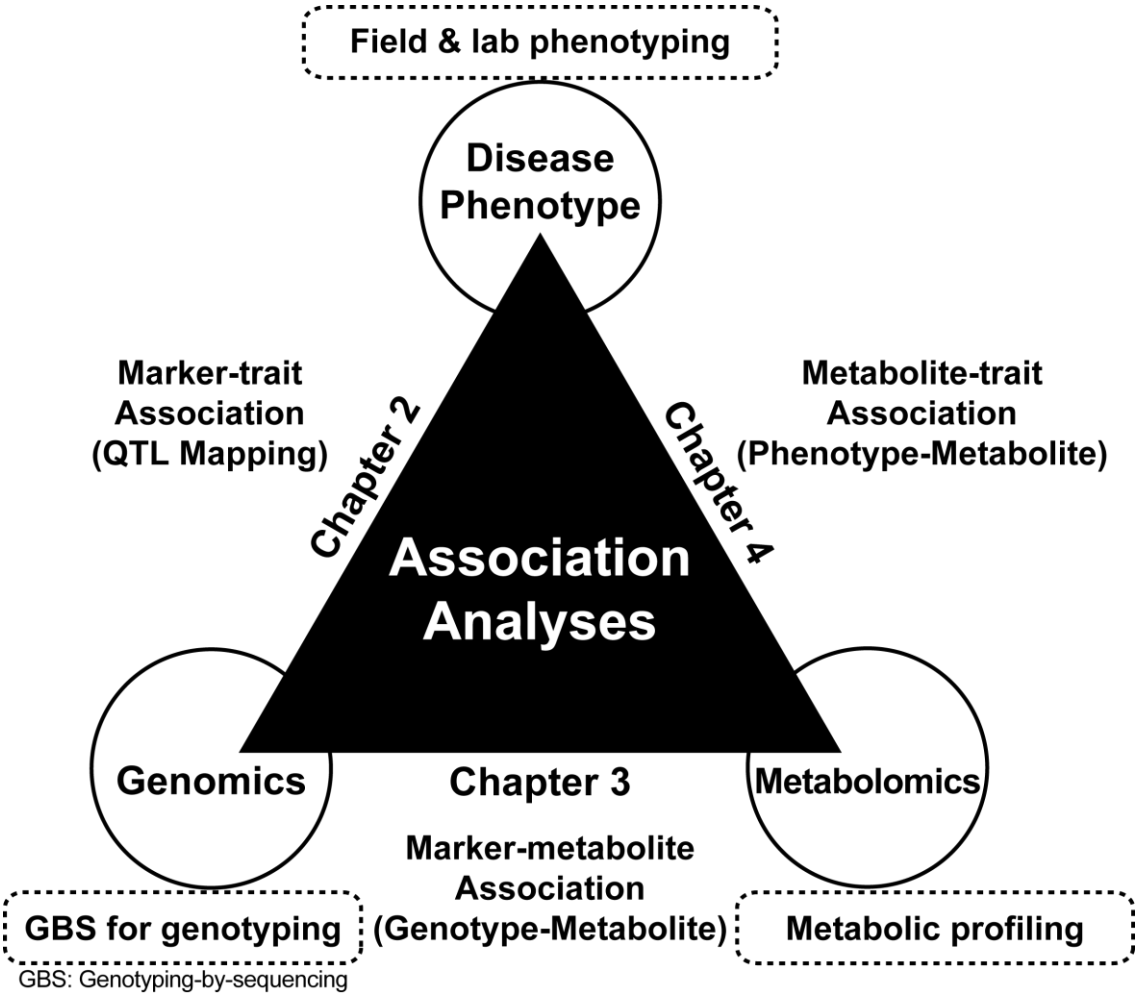


Figure 5.1 Summary of dissertation objectives.

Bibliography

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. *Theoretical and Applied Genetics*, *109*(5), 1017-1027.

Adam-Blondon, A. F., Jaillon, O., Vezzulli, S., Zharkikh, A., Troggio, M., Velasco, R., & Martinez-Zapater, J. (2011). Genome sequence initiatives. *Genetics, Genomics, and Breeding of Grapes*, *211*, 234.

Adrian, M., Jeandet, P., Veneau, J., Weston, L. A., & Bessis, R. (1997). Biological activity of resveratrol, a stilbenic compound from grapevines, against *Botrytis cinerea*, the causal agent for gray mold. *Journal of Chemical Ecology*,*23*(7), 1689-1702.

Alleweldt, G., & Possingham, J. V. (1988). Progress in grapevine breeding. *Theoretical and Applied Genetics*, *75*(5), 669-673.

Allwood, J. W., Ellis, D. I., & Goodacre, R. (2008). Metabolomic technologies and their application to the study of plants and plant–host interactions. *Physiologia Plantarum*, *132*(2), 117-135.

Anstey, K. J., Mack, H. A., & Cherbuin, N. (2009). Alcohol consumption as a risk factor for dementia and cognitive decline: meta-analysis of prospective studies. *The American journal of Geriatric psychiatry*, *17*(7), 542-555.

Arends, D., Prins, P., Jansen, R. C., & Broman, K. W. (2010). R/qtl: high throughput multiple QTL mapping. *Bioinformatics*, *26*(23), 2990-2992.

Aretz, I., & Meierhofer, D. (2016). Advantages and pitfalls of mass spectrometry based metabolome profiling in systems biology. *International journal of molecular sciences*, *17*(5), 632.

Barba, P., Cadle-Davidson, L., Harriman, J., Glaubitz, J. C., Brooks, S., Hyma, K., & Reisch, B. (2014). Grapevine powdery mildew resistance and susceptibility loci identified on a high-resolution SNP map. *Theoretical and applied genetics*, *127*(1), 73-84.

Barker, C. L., Donald, T., Pauquet, J., Ratnaparkhe, M. B., Bouquet, A., Adam Blondon, A. F., Thomas, M. R., & Dry, I. (2005). Genetic and physical mapping of the grapevine powdery mildew resistance gene, *Run1*, using a bacterial artificial chromosome library. *Theoretical and Applied Genetics*, *111*(2), 370-377.

Bavaresco, L., Petegolli, D., Cantü, E., Fregoni, M., Chiusa, G., & Trevisan, M.

(1997). Elicitation and accumulation of stilbene phytoalexins in grapevine berries infected by *Botrytis cinerea*. *Vitis*, *36*(2), 77-83.

Bavaresco, L., & Fregoni, C. (2001). Physiological role and molecular aspects of grapevine stilbenic compounds. In *Molecular biology & biotechnology of the grapevine* (pp. 153-182). Springer, Dordrecht.

Bavaresco, L., Fregoni, C., de Macedo Basto, M. V. Z., & Vezzulli, S. (2009). Physiology & molecular biology of grapevine stilbenes: an update. In *Grapevine Molecular Physiology & Biotechnology* (pp. 341-364). Springer, Dordrecht.

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 localised necrosis at the infection site. *Theoretical and applied genetics*, *120*(1), 163-176.

Berger, S., Sinha, A. K., & Roitsch, T. (2007). Plant physiology meets phytopathology: plant primary metabolism and plant–pathogen interactions. *Journal of experimental botany*, *58*(15-16), 4019-4026.

Bernardo, R. (2010). *Breeding for quantitative traits in plants* (p. 86). Woodbury: Stemma Press.

Billotte, N., Jourjon, M. F., Marseillac, N., Berger, A., Flori, A., Asmady, H., et al. (2010). QTL detection by multi-parent linkage mapping in oil palm (Elaeis guineensis Jacq.). *Theoretical and Applied Genetics*, *120*(8), 1673-1687.

Blanc, G., Charcosset, A., Mangin, B., Gallais, A., & Moreau, L. (2006). Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. *Theoretical and Applied Genetics*, *113*(2), 206-224.

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. *Theoretical and Applied Genetics*, *125*(8), 1663-1675.

Blasi, P., Blanc, S., Wiedemann-Merdinoglu, S., Prado, E., Rühl, E. H., 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. *Theoretical and applied genetics*, *123*(1), 43-53.

Bollina, V., Kushalappa, A. C., Choo, T. M., Dion, Y., & Rioux, S. (2011).

Identification of metabolites related to mechanisms of resistance in barley against *Fusarium graminearum*, based on mass spectrometry. *Plant molecular biology*, *77*(4-5), 355.

Bolton, M. D. (2009). Primary metabolism and plant defense—fuel for the fire. *Molecular plant-microbe Interactions*, *22*(5), 487-497.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633-2635.

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

Brown, M. (2008). Midwest Grape Production Guide. *Journal of Extension*, *46*(6).

Cadle-Davidson, L., Chicoine, D. R., & Consolie, N. H. (2011). Variation within and among *Vitis* spp. for foliar resistance to the powdery mildew pathogen *Erysiphe necator*. Plant Dis. 95:202-211.

Cadle-Davidson, L., Mahanil, S., Gadoury, D. M., Kozma, P., & Reisch, B. I. (2011). Natural infection of *Run1*-positive vines by naïve genotypes of *Erysiphe necator*. *Vitis*, *50*(4), 173-175.

Cadle-Davidson, L., Gadoury, D., Fresnedo-Ramírez, J., Yang, S., Barba, P., Sun, Q., Demmings, E. M., Seem, R., Schaub, M., Nowogrodzki, A., Kasinathan, H., Ledbetter, C., & Reisch, B. I. (2016). Lessons from a phenotyping center revealed by the genome-guided mapping of powdery mildew resistance loci. *Phytopathology*, *106*(10), 1159-1169.

Carreno-Quintero, N., Acharjee, A., Maliepaard, C., Bachem, C. W., Mumm, R., Bouwmeester, H., Bouwmeester, H., Visser, R. G., & Keurentjes, J. J. (2012). Untargeted metabolic quantitative trait loci analyses reveal a relationship between primary metabolism and potato tuber quality. *Plant physiology*, *158*(3), 1306-1318.

Chambers, A. H., Pillet, J., Plotto, A., Bai, J., Whitaker, V. M., & Folta, K. M. (2014). Identification of a strawberry flavor gene candidate using an integrated genetic-genomic-analytical chemistry approach. *BMC genomics*, *15*(1), 217.

Chao, C., Slezak, J. M., Caan, B. J., & Quinn, V. P. (2008). Alcoholic beverage intake and risk of lung cancer: the California Men's Health Study. *Cancer Epidemiology and Prevention Biomarkers*, *17*(10), 2692-2699.

Chitarrini, G., Soini, E., Riccadonna, S., Franceschi, P., Zulini, L., Masuero, D.,

Vecchione, A., Stefanini, M., Di Gaspero, G., Mattivi, F., & Vrhovsek, U. (2017). Identification of Biomarkers for Defense Response to Plasmopara viticola in a Resistant Grape Variety. *Frontiers in plant science*, *8*, 1524.

Chong, I. G., & Jun, C. H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and intelligent laboratory systems*, *78*(1-2), 103-112.

Christen, V., Rusconi, M., Crettaz, P., & Fent, K. (2017). Developmental neurotoxicity of different pesticides in PC-12 cells in vitro. *Toxicology and applied pharmacology*, *325*, 25-36.

Clark, M. D., Susko, A. Q., & Teh, S. L. (2016) Development of digital image analysis protocol for high-throughput phenotyping of fruiting traits in *Vitis*. Poster P1120. International Plant & Animal Genome XXIV, San Diego, CA https://pag.confex.com/pag/xxiv/webprogram/Paper20979.html Accessed 20 September 2016

Dalbó, M. A., Ye, G. N., Weeden, N. F., Steinkellner, H., Sefc, K. M., & Reisch, B. I. (2000). A gene controlling sex in grapevines placed on a molecular marker-based genetic map. *Genome*, *43*(2), 333-340.

Dalbó, M. A., Ye, G. N., Weeden, N. F., Wilcox, W. F., & Reisch, B. I. (2001). Marker-assisted selection for powdery mildew resistance in grapes. *Journal of the American Society for Horticultural Science*, *126*(1), 83-89.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*(7), 499.

Degenkolbe, T., Do, P. T., Kopka, J., Zuther, E., Hincha, D. K., & Köhl, K. I. (2013). Identification of drought tolerance markers in a diverse population of rice cultivars by expression and metabolite profiling. *PLoS One*, *8*(5), e63637.

Dettmer, K., Aronov, P. A., & Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass spectrometry reviews*, *26*(1), 51-78.

Deyanira, Q. M., Estrada-Luna, A. A., Altamirano-Hernández, J., Peña-Cabriales,

J. J., de Oca-Luna, R. M., & Cabrera-Ponce, J. L. (2012). Use of trehalose metabolism as a biochemical marker in rice breeding. *Molecular breeding*, *30*(1), 469-477.

Di Gaspero, G., Cipriani, G., Marrazzo, M. T., Andreetta, D., Castro, M. J. P., 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. *Molecular Breeding*, *15*(1), 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. *Theoretical and Applied Genetics*, *114*(7), 1249-1263.

Doligez, A., Bouquet, A., Danglot, Y., Lahogue, F., Riaz, S., Meredith, C., Edwards, K. J., & This, P. (2002). Genetic mapping of grapevine *(Vitis vinifera* L.) applied to the detection of QTLs for seedlessness and berry weight. *Theoretical and Applied Genetics*, *105*(5), 780-795.

Doligez, A., Adam-Blondon, A. F., Cipriani, G., Di Gaspero, G., Laucou, V., Merdinoglu, D., Meredith, C. P., Riaz, S., Roux C., & This, P. (2006). An integrated SSR map of grapevine based on five mapping populations. *Theoretical and applied genetics*, *113*(3), 369-382.

Doucleff, M., Jin, Y., Gao, F., Riaz, S., Krivanek, A. F., & Walker, M. A. (2004). A genetic linkage map of grape, utilizing *Vitis rupestris* and *Vitis arizonica*. *Theoretical and applied genetics*, *109*(6), 1178-1187.

Doyle, J. J., Flagel, L. E., Paterson, A. H., Rapp, R. A., Soltis, D. E., Soltis, P. S., & Wendel, J. F. (2008). Evolutionary genetics of genome merger and doubling in plants. *Annual review of genetics*, *42*, 443-461.

Edge-Garza, D. A., Luby, J. J., & Peace, C. (2015). Decision support for cost efficient and logistically feasible marker-assisted seedling selection in fruit breeding. *Molecular breeding*, *35*(12), 223.

Eduardo, I., Chietera, G., Pirona, R., Pacheco, I., Troggio, M., Banchi, E., Bassi, D., Rossini, L., Vecchietti, A., & Pozzi, C. (2013). Genetic dissection of aroma volatile compounds from the essential oil of peach fruit: QTL analysis and identification of candidate genes using dense SNP maps. *Tree Genetics & Genomes*, *9*(1), 189-204.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one*, *6*(5), e19379.

Endelman, J. B., & Plomion, C. (2014). LPmerge: an R package for merging genetic maps by linear programming. *Bioinformatics*, *30*(11), 1623-1624.

Fennell, A., Mathiason, K., & Luby, J. (2005). Genetic segregation for indicators of photoperiod control of dormancy induction in Vitis species. *Acta horticulturae*, 689, 533-540, doi:10.17660/ActaHortic.2005.689.66.

Fiehn, O., Wohlgemuth, G., Scholz, M., Kind, T., Lee, D. Y., Lu, Y., Moon, S., & Nikolau, B. (2008). Quality control for plant metabolomics: reporting MSI-compliant studies. *The Plant Journal*, *53*(4), 691-704.

Fischer, B. M., Salakhutdinov, I., Akkurt, M., Eibach, R., Edwards, K. J., Töpfer, R., & Zyprian, E. M. (2004). Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. *Theoretical and Applied Genetics*, *108*(3), 501-515.

Flamini, R. (2013). Recent applications of mass spectrometry in the study of grape and wine polyphenols. *ISRN Spectroscopy*, *2013*.

Flamini, R., De Rosso, M., De Marchi, F., Dalla Vedova, A., Panighel, A., Gardiman, M., Maoz, I., & Bavaresco, L. (2013). An innovative approach to grape metabolomics: stilbene profiling by suspect screening analysis. *Metabolomics*, *9*(6), 1243-1253.

Fu, J., Swertz, M. A., Keurentjes, J. J., & Jansen, R. C. (2007). MetaNetwork: a computational protocol for the genetic study of metabolic networks. *Nature protocols*, *2*(3), 685-694.

Fung, R. W., Gonzalo, M., Fekete, C., Kovacs, L. G., He, Y., Marsh, E., McIntyre, L. M., Schachtman, D. P., & Qiu, W. (2008). Powdery mildew induces defense-oriented reprogramming of the transcriptome in a susceptible but not in a resistant grapevine. *Plant physiology*, *146*(1), 236-249.

Gadoury, D. M., & Pearson, R. C. (1991). Heterothalism and pathogenic specialization in *Uncinula necator*. *Phytopathology*, *81*(10), 1287-1293.

Gadoury, D. M., Seem, R. C., Ficke, A., & Wilcox, W. F. (2003). Ontogenic resistance to powdery mildew in grape berries. *Phytopathology*, *93*(5), 547-555.

Gadoury, D. M., Cadle-Davidson, L., Wilcox, W. F., Dry, I. B., Seem, R. C., & Milgroom, M. G. (2012). Grapevine powdery mildew (*Erysiphe necator*): a fascinating system for the study of the biology, ecology and epidemiology of an obligate biotroph. *Molecular plant pathology*, *13*(1), 1-16.

Gee, C. T., Gadoury, D. M., & Cadle-Davidson, L. (2008). Ontogenic resistance

to *Uncinula necator* varies by genotype and tissue type in a diverse collection of *Vitis* spp. *Plant Disease*, *92*(7), 1067-1073.

Giacomoni, F., Le Corguillé, G., Monsoor, M., Landi, M., Pericard, P., Pétéra, M., Duperier, C., Tremblay-Franco, M., Martin, J., Jacob, D., Goulitquer, S., Thévenot, E. A., & Caron, C., (2014). Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics. *Bioinformatics*, *31*(9), 1493-1495.

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*(4), 1121-1137.

Grimplet, J., Cramer, G. R., Dickerson, J. A., Mathiason, K., Van Hemert, J., & Fennell, A. Y. (2009). VitisNet: "Omics" integration through grapevine molecular networks. *PloS one*, *4*(12), e8365.

Grimplet, J., Van Hemert, J., Carbonell-Bejerano, P., Díaz-Riquelme, J., Dickerson, J., Fennell, A., Pezzotti, M., & Martínez-Zapater, J. M. (2012). Comparative analysis of grapevine whole-genome gene predictions, functional annotation, categorization and integration of the predicted gene sequences. *BMC Research notes*, *5*(1), 213 doi:10.1186/1756-0500-5-213.

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PloS one*, *9*(2), e90346.

Grando, M. S., Bellin, D., Edwards, K. J., Pozzi, C., Stefanini, M., & Velasco, R. (2003). Molecular linkage maps of *Vitis vinifera* L. and *Vitis riparia* Mchx. *Theoretical and applied genetics*, *106*(7), 1213-1224.

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. *Theoretical and Applied Genetics*, *116*(3), 427-438.

Hyma, K. E., Barba, P., Wang, M., Londo, J. P., Acharya, C. B., Mitchell, S. E., 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*(8), e0134880.

IPGRI, U. (1997). OIV. *Descriptores para la Vid (Vitis spp.) Unión Internacional para la Protección de las Obtenciones Vegetales, Géneve, Switzerland*.

Jaillon, O., Aury, J. M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, *449*(7161), 463-467.

Jones, J. D., & Dangl, J. L. (2006). The plant immune system. *Nature*, *444*(7117), 323-329.

Jürges, G., Kassemeyer, H. H., 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 Biology*, *11*(6), 886-898.

Kamoun, S. (2006). A catalogue of the effector secretome of plant pathogenic oomycetes. *Annu. Rev. Phytopathol.*, *44*, 41-60.

Keurentjes, J. J., Fu, J., De Vos, C. R., Lommen, A., Hall, R. D., Bino, R. J., van der Plas, L. H. W., Janse, R. C., Vreugdenhil, D., & Koornneef, M. (2006). The genetics of plant metabolism. *Nature genetics*, *38*(7), 842-849.

Kobayashi, S., Goto-Yamamoto, N., & Hirochika, H. (2004). Retrotransposon induced mutations in grape skin color. *Science*, *304*(5673), 982-982.

Kobayashi, S., Goto-Yamamoto, N., &Hirochika, H. (2005) Association of *VvMybA1* gene expression with anthocyanin production in grape (*Vitis vinifera* L.) skin-color mutants. *J Jpn Soc Hortic Sci 74*:196–203

Kortekamp, A., & Zyprian, E. (2003). Characterization of Plasmopara-resistance in grapevine using *in vitro* plants. *Journal of plant physiology*, *160*(11), 1393-1400.

Kortekamp, A. (2006). Expression analysis of defence-related genes in grapevine leaves after inoculation with a host and a non-host pathogen. *Plant physiology and biochemistry*, *44*(1), 58-67.

Kuehnbaum, N. L., & Britz-McKibbin, P. (2013). New advances in separation science for metabolomics: resolving chemical diversity in a post-genomic era. *Chemical Reviews*, *113*(4), 2437-2468.

Le Clerc, V., Marques, S., Suel, A., Huet, S., Hamama, L., Voisine, L., Auperpin, E., Jourdan, M., Barrot, L., Prieur, R., & Briard, M. (2015). QTL mapping of carrot resistance to leaf blight with connected populations: stability across years and consequences for breeding. *Theoretical and applied genetics*, *128*(11), 2177-2187.

Lewis, W. H. (1980). Polyploidy in angiosperms: dicotyledons. In *Polyploidy* (pp. 241-268). Springer US.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics*, *25*(14), 1754-1760.

Lijavetzky, D., Ruiz-García, L., Cabezas, J. A., De Andrés, M. T., Bravo, G., Ibáñez, A., Carreño, J., Cabello, F., Ibáñez, J., & Martínez-Zapater, J. M. (2006). Molecular genetics of berry colour variation in table grape. *Molecular Genetics and Genomics*, *276*(5), 427-435.

Lisec, J., Meyer, R. C., Steinfath, M., Redestig, H., Becher, M., Witucka-Wall, H., Riehn, O., Törjék, O., Selbig, J., Altmann, T., & Willmitzer, L. (2008). Identification of metabolic and biomass QTL in *Arabidopsis thaliana* in a parallel analysis of RIL and IL populations. *The Plant Journal*, *53*(6), 960-972.

Liu, Z., Alseekh, S., Brotman, Y., Zheng, Y., Fei, Z., Tieman, D. M., Giovannoni, J., Fernie, A. R., & Klee, H. J. (2016). Identification of a Solanum pennellii chromosome 4 fruit flavor and nutritional quality-associated metabolite QTL. *Frontiers in plant science*, *7*, 1671.

Lodhi, M. A., Ye, G. N., Weeden, N. F., Reisch, B. I., & Daly, M. J. (1995). A molecular marker based linkage map of *Vitis*. *Genome*, *38*(4), 786-794.

Lopez, M., Martınez, F., Del Valle, C., Orte, C., & Miro, M. (2001). Analysis of phenolic constituents of biological interest in red wines by high-performance liquid chromatography. *Journal of Chromatography A*, *922*(1-2), 359-363.

Lowe, K. M., & Walker, M. A. (2006). Genetic linkage map of the interspecific grape rootstock cross Ramsey (*Vitis champinii*) × Riparia Gloire (*Vitis riparia*). *Theoretical and applied genetics*, *112*(8), 1582-1592.

Luby, J. J., & Shaw, D. V. (2001). Does marker-assisted selection make dollars and sense in a fruit breeding program? *HortScience*, *36*(5), 872-879.

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. *Theoretical and Applied Genetics*, *118*(7), 1261-1278.

Marti, G., Erb, M., Boccard, J., Glauser, G., Doyen, G. R., Villard, N., Robert, C. A., Turlings, T. C. J., Rudaz, S., & Wolfender, J. L. (2013). Metabolomics reveals herbivore-induced metabolites of resistance and susceptibility in maize leaves and roots. *Plant, Cell & Environment*, *36*(3), 621-639.

Matsuda, F., Okazaki, Y., Oikawa, A., Kusano, M., Nakabayashi, R., Kikuchi, J., Yonemaru, J., Ebana, K., Yano, M., & Saito, K. (2012). Dissection of genotype–phenotype associations in rice grains using metabolome quantitative trait loci analysis. *The Plant Journal*, *70*(4), 624-636.

Mattivi, F., Vrhovsek, U., Malacarne, G., Masuero, D., Zulini, L., Stefanini, M., Moser, C., Velasco, R., & Guella, G. (2011). Profiling of resveratrol oligomers, important stress metabolites, accumulating in the leaves of hybrid *Vitis vinifera* (Merzling × Teroldego) genotypes infected with *Plasmopara viticola*. *Journal of agricultural and food chemistry*, *59*(10), 5364-5375.

McCallum, S., Woodhead, M., Hackett, C. A., Kassim, A., Paterson, A., & Graham, J. (2010). Genetic and environmental effects influencing fruit colour and QTL analysis in raspberry. *Theoretical and applied genetics*, *121*(4), 611-627.

Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, *118*, 62-69.

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. *Molecular Breeding*, *15*(4), 349-366.

Meyer, R. C., Steinfath, M., Lisec, J., Becher, M., Witucka-Wall, H., Törjék, O., Fiehn, O., Eckardt, Ä., Willmitzer, L., Selbig, J., & Altmann, T. (2007). The metabolic signature related to high plant growth rate in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, *104*(11), 4759-4764.

Migliaro, D., Crespan, M., Muñoz-Organero, G., Velasco, R., Moser, C., & Vezzulli, S. (2014). Structural dynamics at the berry colour locus in *Vitis vinifera* L. somatic variants. *Australian journal of grape and wine research*, *20*(3), 485-495.

Moreira, F. M., Madini, A., Marino, R., Zulini, L., Stefanini, M., Velasco, R., Kozma, R., & Grando, M. S. (2011). Genetic linkage maps of two interspecific grape crosses (Vitis spp.) used to localize quantitative trait loci for downy mildew resistance. *Tree Genetics & Genomes*, *7*(1), 153-167.

Moroldo, M., Paillard, S., Marconi, R., Fabrice, L., Canaguier, A., Cruaud, C., De Berardinis, V., Guichard, C., Brunaud, V., Le Clainche, I., Scalabrin, S., Testolin, R., Di Gaspero, G., Morgante, M., & Adam Blondon, A. F. (2008).

A physical map of the heterozygous grapevine 'Cabernet Sauvignon' allows mapping candidate genes for disease resistance. *BMC Plant Biology*, *8*(1), 66.

Müller, K. O. & Börger H. (1940). Experimentelle Untersuchungen über die *Phytophthorainfestans*-Resistenz der Kartoffel. *Arb. Biol. Reichsanst. Land Forstwirtsch.* 23:189–31

Myles, S., Chia, J. M., Hurwitz, B., Simon, C., Zhong, G. Y., Buckler, E., & Ware, D. (2010). Rapid genomic characterization of the genus *Vitis*. *PloS one*, *5*(1), e8219.

Owens, C. L. (2008). Grapes. In *Temperate Fruit Crop Breeding* (pp. 197-233). Springer Netherlands.

Pap, D., Riaz, S., Dry, I. B., Jermakow, A., Tenscher, A. C., Cantu, D., Oláh, R., & Walker, M. A. (2016). Identification of two novel powdery mildew resistance loci, Ren6 and Ren7, from the wild Chinese grape species Vitis piasezkii. *BMC plant biology*, *16*(1), 170.

Pauly, L., Flajoulot, S., Garon, J., Julier, B., Béguier, V., & Barre, P. (2012). Detection of favorable alleles for plant height and crown rust tolerance in three connected populations of perennial ryegrass (*Lolium perenne* L.). *Theoretical and Applied Genetics*, *124*(6), 1139-1153.

Pawlus, A. D., Waffo-Téguo, P., Shaver, J., & Mérillon, J. M. (2012). Stilbenoid chemistry from wine and the genus Vitis, a review. *Journal International des Sciences de la Vigne et du Vin*, *46*(2), 57-111.

Pawlus, A. D., Sahli, R., Bisson, J., Rivière, C., Delaunay, J. C., Richard, T., Gomès, E., Bordenave, L., Waffo-Téguo, P., & Mérillon, J. M. (2013). Stilbenoid profiles of canes from *Vitis* and *Muscadinia* species. *Journal of agricultural and food chemistry*, *61*(3), 501-511.

Pearson, R. C., & Goheen, A. C. (1988). Compendium of Grape Diseases. The American Phytopathological Society.

Pelsy, F., Dumas, V., Bévilacqua, L., Hocquigny, S., & Merdinoglu, D. (2015). Chromosome replacement and deletion lead to clonal polymorphism of berry color in grapevine. *PLoS genetics*, *11*(4), e1005081.

Pezet, R., Gindro, K., Viret, O., & Spring, J. L. (2004a). Glycosylation and oxidative dimerization of resveratrol are respectively associated to sensitivity and resistance of grapevine cultivars to downy mildew. *Physiological and Molecular Plant Pathology*, *65*(6), 297-303.

Pezet, R., Gindro, K., Viret, O., & Richter, H. (2004b). Effects of resveratrol, viniferins and pterostilbene on *Plasmopara viticola* zoospore mobility and disease development. *Vitis*, *43*(3), 145-148.

Poland, J. A., & Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *The Plant Genome*, *5*(3), 92-102.

Pool, R. M., Creasy, L. L., & Frackelton, A. S. (1981). Resveratrol and the viniferins, their application to screening for disease resistance in grape breeding programs. *Vitis, 20, 136-145*.

Ramming, D. W., Gabler, F., Smilanick, J., Cadle-Davidson, M., Barba, P., Mahanil, S., & Cadle-Davidson, L. (2011). A single dominant locus, *Ren4*, confers rapid non-race-specific resistance to grapevine powdery mildew. *Phytopathology*, *101*(4), 502-508.

Rayne, S., Karacabey, E., & Mazza, G. (2008). Grape cane waste as a source of *trans*-resveratrol and *trans*-viniferin: High-value phytochemicals with medicinal and anti-phytopathogenic applications. *Industrial crops and products*, *27*(3), 335-340.

Reisch, B. I., Owens, C. L., & Cousins, P. S. (2012). Grape. In *Fruit breeding* (pp. 225-262). Springer US.

Riaz, S., Dangl, G. S., Edwards, K. J., & Meredith, C. P. (2004). A microsatellite marker based framework linkage map of *Vitis vinifera* L. *Theoretical and Applied Genetics*, *108*(5), 864-872.

Riaz, S., Tenscher, A. C., Ramming, D. W., & Walker, M. A. (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. *Theoretical and applied genetics*, *122*(6), 1059-1073.

Riedelsheimer, C., Lisec, J., Czedik-Eysenberg, A., Sulpice, R., Flis, A., Grieder, C., Altmann, T., Stitt, M., Willmitzer, L., & Melchinger, A. E. (2012). Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proceedings of the National Academy of Sciences*, *109*(23), 8872-8877.

Rivière, C., Pawlus, A. D., & Mérillon, J. M. (2012). Natural stilbenoids: distribution in the plant kingdom and chemotaxonomic interest in Vitaceae. *Natural product reports*, *29*(11), 1317-1333.

Rortais, A., Arnold, G., Dorne, J. L., More, S. J., Sperandio, G., Streissl, F., Szentes, C., & Verdonck, F. (2017). Risk assessment of pesticides and other stressors in bees: principles, data gaps and perspectives from the

European Food Safety Authority. *Science of the Total Environment*, *587*, 524-537.

Salmaso, M., Malacarne, G., Troggio, M., Faes, G., Stefanini, M., Grando, M. S., & Velasco, R. (2008). A grapevine (*Vitis vinifera* L.) genetic map integrating the position of 139 expressed genes. *Theoretical and applied genetics*, *116*(8), 1129-1143.

Sana, T. R., Fischer, S., Wohlgemuth, G., Katrekar, A., Jung, K. H., Ronald, P. C., & Fiehn, O. (2010). Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen *Xanthomonas oryzae pv. oryzae*. *Metabolomics*, *6*(3), 451-465.

Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Carrari, F., Pleban, T., Perez-Melis, A., Bruedigam, C., Kopka, J., Willmitzer, L., Zamir, D., & Fernie, A. R. (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature biotechnology*, *24*(4), 447-454.

Schilling, M. P., Wolf, P. G., Duffy, A. M., Rai, H. S., Rowe, C. A., Richardson, B. A., & Mock, K. E. (2014). Genotyping-by-sequencing for *Populus* population genomics: an assessment of genome sampling patterns and filtering approaches. *PLoS One*, *9*(4), e95292.

Schnee, S., Viret, O., & Gindro, K. (2008). Role of stilbenes in the resistance of grapevine to powdery mildew. *Physiological and Molecular Plant Pathology*,*72*(4), 128-133.

Schnee, S., Queiroz, E. F., Voinesco, F., Marcourt, L., Dubuis, P. H., Wolfender, J. L., & Gindro, K. (2013). *Vitis vinifera* canes, a new source of antifungal compounds against *Plasmopara viticola*, *Erysiphe necator*, and *Botrytis cinerea*. *Journal of agricultural and food chemistry*, *61*(23), 5459-5467.

Schwachtje, J., & Baldwin, I. T. (2008). Why does herbivore attack reconfigure primary metabolism? *Plant physiology*, *146*(3), 845-851.

Schwegler, D. D., Liu, W., Gowda, M., Würschum, T., Schulz, B., & Reif, J. C. (2013). Multiple-line cross quantitative trait locus mapping in sugar beet (*Beta vulgaris* L.). *Molecular breeding*, *31*(2), 279-287.

Simmonds, N. W. (1976). *Evolution of crop plants*. Longman Group Ltd.

Steinhoff, J., Liu, W., Maurer, H. P., Würschum, T., Friedrich, C., Longin, H., Friedrich, C., Ranc, N., & Reif, J. C. (2011). Multiple-line cross quantitative trait locus mapping in European elite maize. *Crop science*, *51*(6), 2505-2516.

Stervbo, U., Vang, O., & Bonnesen, C. (2007). A review of the content of the putative chemopreventive phytoalexin resveratrol in red wine. *Food Chemistry*, *101*(2), 449-457.

Tarpley, L., Duran, A. L., Kebrom, T. H., & Sumner, L. W. (2005). Biomarker metabolites capturing the metabolite variance present in a rice plant developmental period. *BMC plant biology*, *5*(1), 8.

Teh, S. L., Fresnedo-Ramírez, J., Clark, M. D., Gadoury, D. M., Sun, Q., Cadle-Davidson, L., & Luby, J. J. (2017). Genetic dissection of powdery mildew resistance in interspecific half-sib grapevine families using SNP-based maps. *Molecular Breeding*, *37*(1), 1.

Tieman, D. M., Zeigler, M., Schmelz, E. A., Taylor, M. G., Bliss, P., Kirst, M., & Klee, H. J. (2006). Identification of loci affecting flavour volatile emissions in tomato fruits. *Journal of experimental botany*, *57*(4), 887-896.

Troggio, M., Malacarne, G., Coppola, G., Segala, C., Cartwright, D. A., Pindo, M., Stefanini, M., Mank, R., Moroldo, M., Morgante, M., Grando, M. S., & Velasco, R. (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.

Unger, S., Büche, C., Boso, S., & Kassemeyer, H. H. (2007). The course of colonization of two different *Vitis* genotypes by *Plasmopara viticola* indicates compatible and incompatible host-pathogen interactions. *Phytopathology*, *97*(7), 780-786.

Van Ooijen, J. W. (2006). JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. *Kyazma BV, Wageningen*, *33*(10.1371).

Van Ooijen, J. W., & Kyazma, B. (2011). MapQTL 6: software for the mapping of quantitative trait loci in experimental populations of diploid species. Wageningen, The Netherlands.

Velasco, R., Zharkikh, A., Troggio, M., Cartwright, D. A., Cestaro, A., Pruss, D., et al. (2007). A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PloS one*, *2*(12), e1326.

Vergara, C., von Baer, D., Mardones, C., Wilkens, A., Wernekinck, K., Damm, A., Macke, S., Gorena, T., & Winterhalter, P. (2012). Stilbene levels in grape cane of different cultivars in southern Chile: determination by HPLC-DAD-MS/MS method. *Journal of agricultural and food chemistry*, *60*(4), 929-933.

Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of heredity*, *93*(1), 77-78.

Vrhovsek, U., Malacarne, G., Masuero, D., Zulini, L., Guella, G., Stefanini, M., Velasco, R., & Mattivi, F. (2012). Profiling and accurate quantification of *trans*-resveratrol, *trans*-piceid, *trans*-pterostilbene and 11 viniferins induced by *Plasmopara viticola* in partially resistant grapevine leaves. *Australian journal of grape and wine research*, *18*(1), 11-19.

Walker, A. R., Lee, E., & Robinson, S. P. (2006). Two new grape cultivars, bud sports of Cabernet Sauvignon bearing pale-coloured berries, are the result of deletion of two regulatory genes of the berry colour locus. *Plant molecular biology*, *62*(4-5), 623-635.

Walker, A. R., Lee, E., Bogs, J., McDavid, D. A., Thomas, M. R., & Robinson, S. P. (2007). White grapes arose through the mutation of two similar and adjacent regulatory genes. *The Plant Journal*, *49*(5), 772-785.

Wang, J., Ho, L., Zhao, Z., Seror, I., Humala, N., Dickstein, D. L., Thiyagarajan, M., Percival, S. S., Talcott S. T., & Pasinetti, G. M. (2006). Moderate consumption of Cabernet Sauvignon attenuates Aβ neuropathology in a mouse model of Alzheimer's disease. *The FASEB Journal*, *20*(13), 2313-2320.

Wannemuehler, S. (2018). *Cost-Benefit Analysis of Marker-Assisted Selection in Rosaceous Fruit Crop Breeding Programs* (Master's Thesis). University of Minnesota, Saint Paul, MN.

Welter, L. J., Göktürk-Baydar, N., Akkurt, M., Maul, E., Eibach, R., Töpfer, R., & Zyprian, E. M. (2007). Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). *Molecular Breeding*, *20*(4), 359-374.

Wiklund, S., Johansson, E., Sjöström, L., Mellerowicz, E. J., Edlund, U., Shockcor, J. P., Gottfries, J., Moritz, T., & Trygg, J. (2008). Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Analytical chemistry*, *80*(1), 115-122.

Yang, S., Fresnedo-Ramírez, J., Sun, Q., Manns, D. C., Sacks, G. L., Mansfield, A. K., Luby, J.J., Londo, J.P., Reisch, B.I., Cadle-Davidson, L.E., & Fennell, A. Y. (2016). Next generation mapping of enological traits in an F2 interspecific grapevine hybrid family. *PloS one*, *11*(3), e0149560.