

UC Davis

UC Davis Electronic Theses and Dissertations

Title

Genomic Dissection and Genome Editing of Pierce's Disease Resistance and Powdery Mildew Susceptibility in Grapevine

Permalink

<https://escholarship.org/uc/item/08s5j9jr>

ISBN

9798193262794

Author

Zaccheo, Mirella

Publication Date

2026-09-08

Peer reviewed|Thesis/dissertation

Genomic Dissection and Genome Editing of Pierce's Disease Resistance and Powdery Mildew
Susceptibility in Grapevine

By

MIRELLA ZACCHEO
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Horticulture and Agronomy

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Dario Cantu, Chair

Caroline Roper

Luis Diaz Garcia

Committee in Charge

2026

Genomic Dissection and Genome Editing of Pierce's Disease Resistance and Powdery Mildew Susceptibility in Grapevine

Abstract

The European grape (*Vitis vinifera*) is used worldwide for wine making. However, its production is constantly challenged by a range of pests and pathogens that cause substantial losses in yield and fruit quality. Among these, Pierce's disease (PD) and powdery mildew (PM) pose a serious threat to the wine industry in California. The management of both diseases currently relies on intensive pesticide applications, which raises concerns regarding environmental sustainability, application costs and the emergence of resistant pathogen populations. Wild North-American *Vitis* species are not suitable for winemaking but represent a valuable reservoir of genetic diversity and natural disease resistance that can be leveraged to improve susceptible *V. vinifera* cultivars. Breeding efforts have focused on introducing some of these favorable traits into susceptible varieties, but the process is lengthy and can result in linkage drag. Advancements in grapevine genomics and genome editing offer the unprecedented opportunity to identify, characterize and utilize these traits to develop disease-resistant cultivars. This dissertation aims to characterize the resistance locus associated with PD in wild *Vitis* species and target susceptibility genes associated with PM using genome editing. Together, these approaches provide complementary strategies to understand the molecular basis of disease resistance and contribute to the development of improved vines.

Chapter 1 provides a comprehensive overview of recent advancements in grapevine and pathogen genomics aimed at achieving durable disease resistance. This research emphasizes the critical role of diploid genome assemblies in dissecting disease resistance loci, enabling the precise identification of candidate genes suitable for gene stacking and precision breeding. Additionally, it highlights the importance of understanding pathogen evolution, including mechanisms of adaptation, effector discovery, and resistance

breakdown. This knowledge is essential not only for informed resistance breeding strategies, but also for identifying novel gene targets for genome editing.

Chapter 2 focuses on the identification of candidate PD-resistance genes in a wild grape species, *Vitis arizonica* b40-14, that segregates for resistance using genome-wide association studies (GWAS). Significant associations between genotype and phenotype were identified across multiple genomic regions. However, one of the strongest association peaks was located on chromosome 14, between the SSR markers defining the Pierce's disease resistance locus 1 (*PdR1*) locus, further supporting this region as a candidate resistance locus. The integration of gene expression data revealed four expressed candidate genes encoding for extracellular receptors potentially involved in PD resistance. These included two genes coding for leucine-rich repeat receptor-like proteins (RLPs), one coding for a leucine motif receptor-like kinase (LysM-RLK) and one coding for a leucine-rich repeat receptor-like kinase (LRR-RLK).

Chapter 3 explores variation within the *PdR1* locus across different *V. arizonica* hybrids accessions. This study integrated multiple layers of information to refine the *PdR1* region in *V. arizonica* b40-14 and identified candidate genes associated with PD. Genome assemblies generated from HiFi reads for four *V. arizonica* hybrids enabled the generation of highly contiguous and complete haplotypes. Among these, four haplotypes were associated with PD resistance and three were not associated with the resistant phenotype. One resistant haplotype was sequenced from a PD-resistant backcross individual that inherited resistance from *V. arizonica* b40-14, the wild grape accession used in Chapter 2. To investigate sequence diversity and identify variants specific to *PdR1*, sequence graph approaches were employed. Results showed a substantial sequence and structural divergence between PD-resistant haplotypes and haplotypes not associated with PD resistance, as well as variation among the resistant haplotypes. By integrating these findings with gene expression data, it was possible to narrow down the candidate genes in the *PdR1* refined region to two genes coding for LRR-RLPs, which were advanced to functional validation.

Chapter 4 builds upon the findings from Chapter 2 and 3 and aims to validate the function of PD candidate resistance genes. To achieve this, a dual complementary platform was developed. The candidate genes were overexpressed in the susceptible *V. vinifera* cv. Thompson Seedless (TS), and the same genes were

knocked-out in the resistant *V. arizonica* b40-14. For overexpression in TS, *Agrobacterium*-mediated transformation was used to generate transgenic plants. Gene knock-out in *V. arizonica* was achieved using through CRISPR/Cas9 editing of protoplasts, which is an extremely challenging yet valuable approach.

Transformation was successful for two out of the four PD candidate genes. Ten putatively transgenic TS lines were obtained, five for each of the two candidate genes that could be transformed. For the other two genes, transformation was unsuccessful, suggesting that constitutive overexpression under the CaMV35S promoter may have detrimental effects. For the knockout experiment, preliminary results showed that the sgRNAs designed to target the candidate genes display a wide range of editing efficiencies and that the type, frequency and pattern of mutation is strongly guide-dependent. Putatively edited *V. arizonica* embryos were regenerated from protoplasts for all the genes of interest and currently at plantlets stage.

Chapter 5 explores an alternative strategy for inducing resistance by targeting a family of susceptibility genes associated with PM, known as *Mildew Resistance Locus O* (*MLO*) genes. Using DNA-free genome editing, we generated *MLO*-knockout lines carrying different mutation patterns across three *MLO* genes: *VvMLO3*, *VvMLO7* and *VvMLO11*. Edited plants regenerated from protoplasts were genotyped and the absence of CRISPR components was verified confirming that the plants are not transgenic. Eight edited lines were evaluated for susceptibility to PM and carried diverse combination of mutations across the three *MLO* loci. Phenotypic assessment was conducted both visually, by scoring pathogen growth, and quantitatively, by measuring pathogen accumulation in the edited plants at multiple time points. Among the edited lines, one showed overall partial reduced pathogen accumulation compared to the non-edited susceptible control.

ACKNOWLEDGMENTS

I would like to thank Professor Dario Cantu for giving me this opportunity and for his guidance throughout my PhD. I am grateful to all the members of the Cantu Lab, especially to Rosa and Andrea. I cannot thank you enough for everything you taught me, and for showing me that perseverance can take you far. I am also grateful to all my coworkers and collaborators, who have helped me to become both the scientist and the person I am today.

I would like to express my gratitude to my dissertation and QE committee, the Horticulture and Agronomy Graduate Group, and my cohort for their support, encouragement, and for making this journey both rewarding and enjoyable.

To my Mom and Dad, thank you for raising me as a resilient and independent woman and for encouraging me to follow my dreams, even when it meant being thousands of miles away from each other. This degree belongs to you as much as it does to me, and it reflects all the sacrifices you made to help me get here.

To Marvin, Amigo, you believed in me long before grad school was even on my mind. You encouraged me to do things I would never imagine I could do. I will always be grateful for your support.

To Peter, thank you for the countless hours spent on the phone, for standing by me through my lowest moments and celebrating my highest ones. After all these years, you are the proof that true friendship is not measured by distance.

To Paul, I met you halfway through this PhD, and somehow you made everything feel lighter and better. Thank you for your patience, for being there through every high and low, and for believing in me and loving me unconditionally.

To Casa Menlo, thank you for making California feel like home. Living with all of you has been a blessing, and I will always cherish the great memories we created together.

To Graziana and Francisco, thank you for being incredible mentors and for showing and reminding me that this PhD is only the beginning of an exciting career. To all the Bayer mentors, colleagues, fellow co-ops and the friends I met in St. Louis, thank you for making those six months an unforgettable experience filled with learning, laughter, and lasting friendships.

Finally, thank you to all my friends, near and far, and to everyone I met during my years in Davis. Each of you made these six years richer, happier and more meaningful. I will carry all the memories, the lessons learned, and the friendships with me wherever life takes me.

DEDICATION

To the ones who believed in me,
more than I ever believed in myself

TABLE OF CONTENTS

Title page	i
Abstract	ii
Acknowledgments	v
Dedication	vii
Table of contents	viii
CHAPTER 1: Advances in grape and pathogen genomics toward durable grapevine disease resistance.....	1
Abstract	2
Introduction	2
The genetic basis of disease resistance.....	3
Pathogen adaptation to resistant grapevines.....	5
Conclusions and perspectives.....	8
Supplementary data	10
References	11
CHAPTER 2: Multigenic resistance to <i>Xylella fastidiosa</i> in wild grapes (<i>Vitis</i> spp.) and its implications within a changing climate.....	14
Abstract	15
Introduction	16
Results	16
Discussion	21
Methods	24
Data availability	26
References	26
CHAPTER 3: Haplotype graph analysis of <i>PdR1</i> uncovers resistance diversity to Pierce’s disease in <i>Vitis arizonica</i> and its hybrids	30
Abstract	31
Introduction	31
Materials and methods.....	32
Results	33
Discussion	38
Data availability	39
Literature cited	39

CHAPTER 4: Dual genome editing strategy development to validate Pierce's disease candidate resistance genes	41
Abstract	41
Introduction	41
Materials and methods.....	43
Results	48
Discussion	62
Conclusions	65
References	66
Supplementary data	69
 CHAPTER 5: Knock-out of <i>VvMLO3</i>, <i>7</i>, <i>11</i> by RNP-mediated genome editing leads to reduced susceptibility to grape powdery mildew	72
Abstract	72
Introduction	72
Materials and methods.....	74
Results	77
Discussion	86
Conclusions	88
References	88

CHAPTER 1

Advances in grape and pathogen genomics toward durable grapevine disease resistance

EXPERT VIEW

Advances in grape and pathogen genomics toward durable grapevine disease resistance

Manon Paineau¹, Mirella Zaccheo¹, Mélanie Massonnet¹, and Dario Cantu^{1,2,*}

¹ Department of Viticulture and Enology, University of California Davis, Davis, CA 95616, USA

² Genome Center, University of California Davis, Davis, CA 95616, USA

* Correspondence: dacantu@ucdavis.edu

Received 13 August 2024; Editorial decision 28 October 2024; Accepted 1 November 2024

Editor: Nathalie Ollat, Institut National de la Recherche Agronomique, France

Abstract

The future sustainability of viticulture depends on the development of grapevine cultivars with genetic resistance to diseases such as powdery mildew, downy mildew, and Pierce's disease. Recent advances in grape and pathogen genomics have dramatically improved our approach to durable disease resistance. The availability of diploid genome references for wild species, combined with the ability to phase resistance haplotypes and conduct genome-wide association and expression analyses, has greatly enhanced our ability to dissect genetic resistance loci. This progress is yielding candidate genes that will form the foundation for precise breeding, gene stacking, and genome editing in grape improvement programs. As resistance genes are deployed in vineyards, pathogen populations evolve to adapt and evade these defenses, posing ongoing challenges. Understanding the adaptive mechanisms of grapevine pathogens in response to resistant cultivars is crucial. Grape pathogenomics is advancing rapidly, marked by the sequencing of many pathogen genomes, the discovery of effectors, including the first ones responsible for disease resistance breakdown, and the development of graph-based pangenomes. These advancements offer valuable insights into pathogen evolution and inform strategies for sustainable disease management. Together, these genomic tools and insights are paving the way for developing resilient grapevine varieties, ensuring the long-term sustainability of viticulture.

Keywords: Disease resistance, effectors, grape downy mildew, grape powdery mildew, pathogenomics, phased genome assembly, Pierce's disease, resistant gene stacking.

Introduction

Global viticulture predominantly relies on *Vitis vinifera* ssp. *vinifera* (*Vv vinifera*), a species domesticated in the Near East ~8000 years ago (Zhou *et al.*, 2017; Dong *et al.*, 2023). Although *Vv vinifera* has been subjected to extensive selection for local adaptation, it remains highly susceptible to various diseases caused by fungi, oomycetes, and bacteria. This is primarily due to its

lack of co-evolution with many of these pathogens, particularly those introduced from North America. In contrast, wild grape populations (*Vitis* spp.), which have evolved alongside these pathogens over millennia, possess multiple forms of genetic resistance, making them more resilient to diseases compared with *Vv vinifera*.

© The Author(s) 2024. Published by Oxford University Press on behalf of the Society for Experimental Biology.
 This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Wild *Vitis* spp. populations span a wide range of climatic, soil, and ecological conditions in the temperate regions of the Northern Hemisphere, particularly in North America and Asia (Cantu *et al.*, 2024). North America is also considered the origin for the most economically significant grapevine pathogens. For example, while all domesticated grapevines are susceptible to Pierce's disease (PD), some wild grape accessions from North America exhibit PD resistance, likely to be due to persistent exposure to *Xylella fastidiosa* in their native habitats (Morales-Cruz *et al.*, 2023). Similarly, *Plasmopara viticola* [causal agent of downy mildew (DM)] and *Erysiphe necator* [causal agent of powdery mildew (PM)] are also believed to have originated in North America (Brewer and Milgroom, 2010; Fontaine *et al.*, 2021). Both pathogens were introduced to European vineyards, where only *Vv vinifera* was grown, in the 1870s. The absence of inherent resistance in European vineyards led to devastating epidemics. The discovery of PM resistance in wild *V. vinifera* ssp. *sylvestris* (*Vv sylvestris*; Lukšić *et al.*, 2022) and other *Vv vinifera* accessions from the Caucasus and Central Asia (Riaz *et al.*, 2013) suggests that genetic resistance might have been present in Central Asia for a long period, potentially before *Vv vinifera* domestication (Riaz *et al.*, 2020).

To exploit these natural sources of genetic resistance, extensive efforts have been made to collect plants from natural *Vitis* spp. populations, establish germplasm collections, and characterize these resources (Khoury *et al.*, 2019; Margaryan *et al.*, 2023). Breeding programs have been introducing genetic disease resistance into elite material to develop *Vv vinifera* cultivars that are inherently resistant or tolerant to disease. Additionally, the role of susceptibility (*S*) genes, which facilitate pathogen infection, has been explored in *V. vinifera* due to their importance in other pathosystems (Garcia-Ruiz *et al.*, 2021; Pirrello *et al.*, 2021). Disrupting *S* genes has shown promise in increasing grapevine resistance to DM and PM (Pessina *et al.*, 2016; Wan *et al.*, 2020; Marcianò *et al.*, 2021; Giacomelli *et al.*, 2023). Disease-resistant grapevine varieties are becoming more popular as regulatory and consumer pressure favor more sustainable and environmentally friendly viticulture practices (Vecchio *et al.*, 2022; Moura *et al.*, 2024). While genetic mapping has played a key role in enabling marker-assisted selection (Vezzulli *et al.*, 2019), recent and ongoing efforts are focused on identifying the individual genes responsible for the identified quantitative trait loci (QTL; Massonnet *et al.*, 2022; Morales-Cruz *et al.*, 2023; Zou *et al.*, 2023). Understanding the molecular basis of genetic resistance to disease is key to prioritizing breeding efforts. Identifying genes responsible for disease resistance allows for the design of perfect markers, which can accelerate breeding and reduce genetic drag (Foria *et al.*, 2022). Candidate resistance (*R*) and *S* genes also serve as targets for genome editing when this technology becomes viable for the industry.

The durability of *R* genes is a significant concern for viticulture, where developing new cultivars can take decades. It often takes only a few years for pathogens to overcome the deployed

resistance in the field due to their fast reproduction cycles, which include both sexual and asexual phases, and their highly dynamic genome architectures (Jones *et al.*, 2014; Dussert *et al.*, 2019; Castillo and Almeida, 2023; Zaccaron *et al.*, 2023; Garcia *et al.*, 2024). As a result, *R* genes are likely to remain effective for only a limited time. To enhance the durability of genetic resistance, breeders build new cultivars by stacking (or pyramiding) multiple *R* genes together (Merdinoglu *et al.*, 2018; Dormatey *et al.*, 2020). Optimizing gene stacking strategies requires a thorough understanding of the molecular mechanisms underlying resistance, which helps reduce redundancies and enhance the functional complementarity of the stacked *R* genes (Michelmore *et al.*, 2013). Genomic advances are just beginning to shed light on how grape pathogens overcome genetic resistance, offering insights that could improve the durability and effectiveness of disease resistance in grapevines (Paineau *et al.*, 2024a).

The genetic basis of disease resistance

Limitations in early grape genomics: the PN40024 genome reference does not carry resistance loci

About 90 disease resistance loci have been identified in grapes, covering major diseases such as PM, DM, PD, black rot, bunch rot, and Phomopsis cane and leaf spot (Supplementary Table S1). So far, only three *R* genes have been functionally characterized, all encoding Toll/interleukin-1 receptor-like nucleotide-binding site leucine-rich repeat (TIR-NBS-LRR) proteins (Feechan *et al.*, 2013; Foria *et al.*, 2020). The identification of *R* genes in grapes was delayed for over a decade due to the lack of an appropriate reference genome for studying genetic resistance. Released in 2007, the haploid consensus genome of *Vv vinifera* cv. PN40024 was among the first plant genomes assembled at chromosome scale, and has served as a genome reference for thousands of studies (Jaillon *et al.*, 2007). This genome was developed using a nearly homozygous line (~93–96%; Velt *et al.*, 2023). Despite the high level of homozygosity, later studies found that the PN40024 genome assembly contained a mix of haplotypes for some heterozygous loci (Zhou *et al.*, 2019; Shi *et al.*, 2023). Additionally, due to its *Vv vinifera* parentage, it was shown early on that the PN40024 genome does not contain the genes involved in disease resistance (Feechan *et al.*, 2013). Although the structure of the PN40024 genome was recently improved to a telomere-to-telomere and gap-free assembly with representation of phased heterozygous haplotypes (Shi *et al.*, 2023), it still represents a susceptible grape species and cannot serve as a reference for *R* gene identification.

Advances in grape genomics: diploid genome references and new solutions to study *R* loci

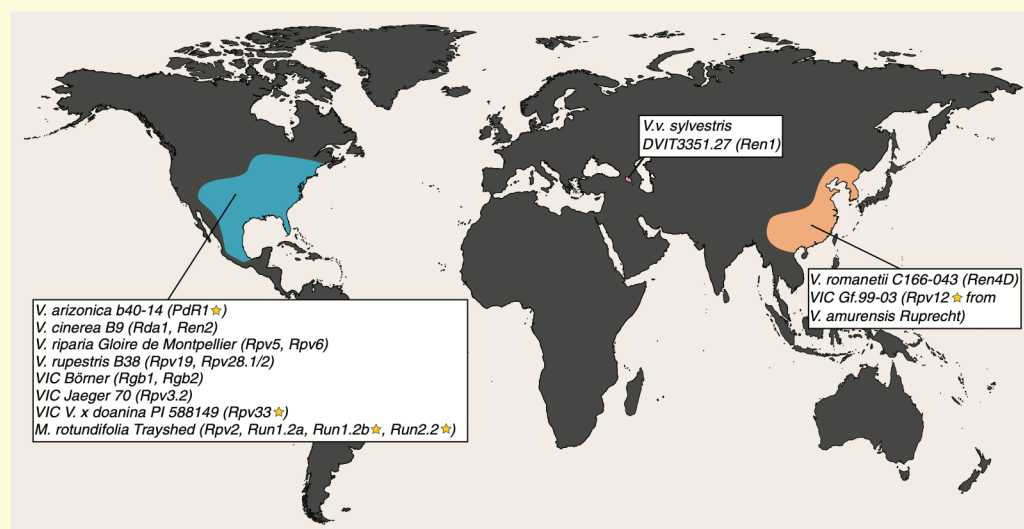
Thanks to the recent advances of long-read sequencing and haplotype-aware assembly tools (Chin *et al.*, 2016; Minio *et al.*, 2022a, b), the first diploid genomes of disease-resistant grapes

have been generated within the past 2 years (Box 1). This marks a significant turning point in the identification of *R* genes in grapes, although challenges remain, such as (i) the large size of the genomic regions corresponding to *R* loci, defined by genetic association and limited by recombination frequencies; (ii) the phasing of the two haplotypes at the *R* loci; and (iii) the accurate annotation of the *R* genes.

The first *R* locus examined in a diploid genome was the PM *R* locus *Run1.2* of *Mirabilis rotundifolia* Trayshed (Cochetel et al., 2021). Analysis of the gene content in nucleotide-binding site leucine-rich repeat (NLR) genes showed a significant expansion of the TIR-NBS-LRR genes compared

with the PM-susceptible *Vv vinifera* cv. Cabernet Sauvignon. However, three challenges remained in studying the *Run1.2* locus. First, the genetic markers flanking the locus were considerably distant from each other (~4 Mb), causing the locus to be fragmented into multiple contigs. Second, since the two haplotypes of *Run1.2* are associated with PM resistance, phasing the two haplotypes was necessary to identify *R* genes specific to each haplotype. Massonnet et al. (2022) used short sequencing reads from a cross possessing only one haplotype of *Run1.2* (*Run1.2b*) for reconstructing and haplotyping the genomic region. Comparison of the two haplotypes of *Run1.2* showed several heterozygous structural variations affecting

Box 1. Key developments in wild grape diploid genome assemblies for the identification of disease resistance genes



Genome assemblies of 12 grape accessions possessing one or multiple *R* loci have been generated so far. These comprise seven *Vitis* species, four *Vitis* interspecific crosses (VIC), and an *M. rotundifolia* accession, from North America, East-Asia, and Armenia (reviewed in Cantu et al., 2024). Combining the diploid genome assembly with haplotype resolution, gene annotation, and gene expression profiling has led to the identification of candidate *R* genes in five *R* loci (indicated by yellow stars in the figure):

- Morales-Cruz et al. (2023) found four extracellular receptor genes expressed in the *PdR1c* locus: two receptor-like protein genes and two receptor-like kinase genes.
- Frommer et al. (2023) identified two CC-NBS-LRR genes in *Rpv12* with a constitutive expression.
- Zou et al. (2023) described three NLR genes specific to the *Rpv33* locus.
- Massonnet et al. (2022) identified several NLRs exhibiting constitutive expression in *Run1.2b* and *Run2.2*. The most highly expressed NLRs included one TIR-NBS, two TIR-NBS-LRR, and one CC-NBS-LRR gene in *Run1.2b*, and one TIR-NBS-LRR gene in *Run2.2*.

the NLR gene content and the domain composition of the encoded NLRs, emphasizing the need to separate the two haplotypes for identifying *R* genes in wild grapes. Third, the accurate annotation of the NLR genes which make up the *R* loci was crucial for assessing their expression and potential role in PM resistance. The annotation of the NLR genes has been a challenge in plants due to their multidomain nature, high intrafamily sequence homology, and low expression. Several tools have been created to extract NLR genes based on the co-occurrences of NLR domains, such as RGAugury (Li et al., 2016), DRAGO 2 (Osuna-Cruz et al., 2017), NLR-annotator (Steuernagel et al., 2020), and NLRtracker (Kourelis et al., 2021). Manual refinement of the NLR gene models by combining the identification of NLR domains with the tool NLR-annotator (Steuernagel et al., 2020) and visualization of the alignments of RNA-seq data showed that 60% of the NLR gene models which make up the two loci were not correctly predicted. This highlights that *R* loci are complex genomic regions that require meticulous dissection and manual curation. Notably, the gene *MrRUN1* from *M. rotundifolia* G52 was not found among the two haplotypes of the *Run1.2* locus from *M. rotundifolia* Trayshed. This supports the necessity of producing the genome assembly of the disease-resistant grape accession when identifying *R* genes. By profiling the expression of the NLR genes which make up the *Run1.2b* and *Run2.2* loci, multiple genes with constitutive expression were identified; one TIR-NBS, two TIR-NBS-LRR, and one CC-NBS-LRR gene in *Run1.2b*, and one TIR-NBS-LRR gene in *Run2.2* were proposed as candidate genes.

Different approaches have been used for phasing and dissecting the two DM resistance loci, *Rpv12* and *Rpv33* (Frommer et al., 2023; Zou et al., 2023). For *Rpv12*, long sequencing reads from an accession that is heterozygous at *Rpv12* were binned into two haplotypes using short sequencing reads of its parents. This process yielded two fully phased haplotypes at the *Rpv12* locus, despite the large size of the region (~1.9 Mb; Frommer et al., 2023). Zou et al. (2023) assembled the genome of *Vitis doaniana* PI 588149, which carries the *Rpv33* locus. Using rhAmpSeq and skim-seq markers, they reduced the size of the region from 9 Mbp to 0.46 Mbp and identified the haplotype carrying the *Rpv33* locus. For both loci, candidate genes were identified by comparing the resistance haplotypes with the alternative haplotypes associated with disease susceptibility, along with gene expression profiling. This led to the identification of two CC-NBS-LRR genes in *Rpv12* and three NLR genes in *Rpv33*. Further analysis of additional DM *R* loci in the appropriate genomes will uncover additional candidate genes, such as, for example, those responsible for *Rpv29*, *Rpv30*, and *Rpv31* in *V. vinifera* Mgaloblishvili, which so far only have been localized by genome-wide association study (GWAS) on PN40024 12X.v2 (Sargolzaei et al., 2020).

Similarly, gene candidates for *Pierce's disease resistance 1* (*PdR1*) in *Vitis arizonica* were identified using a chromosome-scale diploid genome reference of *V. arizonica* b40-14, which

carries *PdR1c* (Box 2). Analysis of manually curated gene models in the phased haplotypes and allele-specific gene expression pointed to four extracellular receptors as potential *PdR1c* *R* genes (Morales-Cruz et al., 2023). Similar approaches should be applied to additional *V. arizonica* accessions resistant to PD. Resistance in multiple accessions has been mapped to the same region on chromosome 14, suggesting that *PdR1* might be the sole locus for PD resistance in *V. arizonica* and its hybrids (Riaz et al., 2008, 2018; Huerta-Acosta et al., 2022). However, the polymorphic allelic profile of the simple sequence repeat (SSR) markers flanking *PdR1* suggests that resistance could be conferred by different allelic forms of the same gene or different *R* genes. Although combining multiple allelic forms of *PdR1* through traditional crosses is challenging due to their location on the same chromosome region, targeted knock in approaches could stack them together (Ainley et al., 2013; Rajput et al., 2023). The precise insertion of multiple *R* genes into a single locus could then be easily introgressed into various genetic backgrounds through crosses. These may be stacked with additional regions associated with PD resistance, yet to be dissected, which have been recently identified by GWAS (Morales-Cruz et al., 2023).

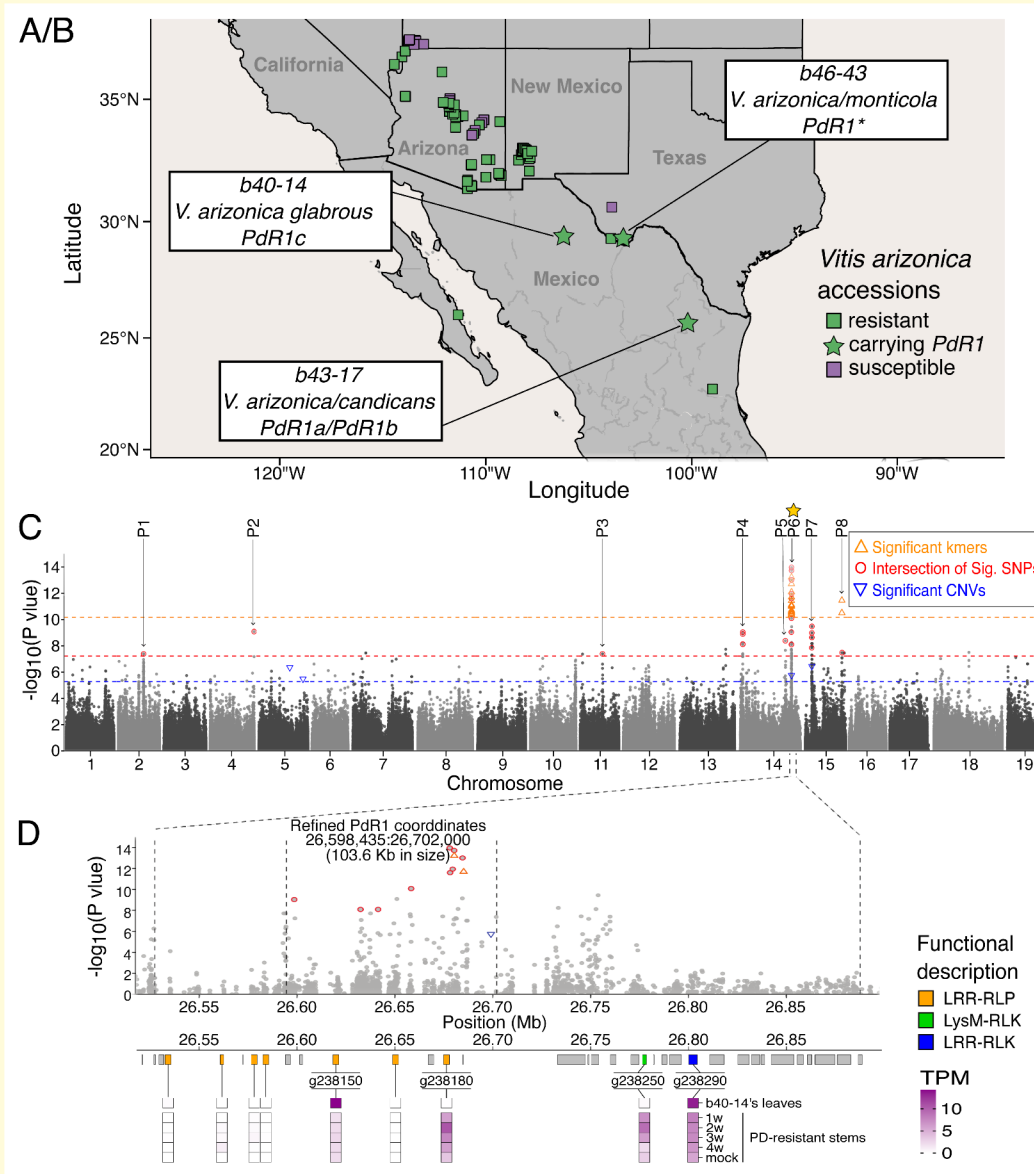
Pathogen adaptation to resistant grapevines

Understanding how pathogens overcome grapevine disease resistance is essential for achieving durable resistance. In the past decade, several grapevine pathogen genomes have been released (Table 1; Supplementary Table S2). Analyzing these genomes has begun to elucidate the mechanisms by which pathogens evade grapevine defenses and cause disease. However, several challenges still impede progress in this field. While improved isolation and sampling techniques now enable the analysis of large populations, essential for GWAS and QTL analyses (Sánchez-Vallet et al., 2018), culturing obligate biotrophic pathogens, such as *E. necator*, remains resource-intensive and often limits the application of genetic association approaches. Additionally, despite advancements, genetic transformation techniques are still not available for most pathogens (Gosavi et al., 2020).

Pathogen evasion of resistance: the case of *Plasmopara viticola* and *Rpv3.1*

Plasmopara viticola populations have already overcome several DM *R* loci deployed in vineyards, including *Rpv3.1*, *Rpv3.2*, *Rpv10*, and *Rpv12* (Heyman et al., 2021; Wingerter et al., 2021; Paineau et al., 2022). *Plasmopara viticola* is known for its rapid evolutionary potential due to a combination of vegetative and sexual reproduction, as well as a highly dynamic and unstable genomic architecture (Chen et al., 2007; Blum et al., 2010; Delmas et al., 2016). Multiple *P. viticola* genomes have

Box 2. *Vitis arizonica* genetic diversity contributes to Pierce's disease resistance in the Southwestern USA and Mexico



Downloaded from <https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erae450/7863762> by guest on 22 June 2025

Box 2. Continued

- (A) *Vitis arizonica glabrous* b40-14 was previously identified to carry *PdR1c*, one of the allelic forms of *PdR1* (Riaz et al., 2018). The improved genome assembly of b40-14 was instrumental to find candidate genes associated with resistance to PD.
- (B) Huerta-Acosta et al. (2023) identified PD resistance in *V. arizonica/monticola* hybrid b46-43, a wild grape from Texas. This accession differs from *V. arizonica/candicans* b43-17, highlighting the necessity of comparing the genomic regions carrying *PdR1*.
- (C) *PdR1* is the major resistance locus (represented by a yellow star; Riaz et al., 2018), but minor QTLs are associated with resistance (Morales-Cruz et al., 2023).
- (D) Morales-Cruz et al. (2023) conducted a GWAS to identify genomic regions associated with PD resistance. The primary region identified was *PdR1*, though other loci were also found. Within the *PdR1* locus, four expressed putative surface receptor genes were identified and are currently considered candidate genes for *PdR1*.

(Figures adapted from Morales-Cruz et al., 2023)

been assembled (Table 1; Supplementary Table S2), facilitating the *in silico* discovery of numerous putative effectors (Mestre et al., 2012, 2016; Combiér et al., 2019; Lan et al., 2019; Lei et al., 2019; Liu et al., 2021; Ma et al., 2021; Fu et al., 2024). Recently, the integration of the diploid genome assembly of an avirulent *P. viticola* strain, GWAS, and genotyping by whole-genome sequencing led to the discovery of the genetic basis of *P. viticola* breakdown of *Rpv3.1* (Box 3; Paineau et al., 2024a). The *Rpv3.1* resistance breakdown locus (*avrRpv3.1*) was identified by GWAS using 123 *P. viticola* strains. Using the diploid genome assembly of a strain heterozygous at the *avrRpv3.1* locus, two effectors were found to be responsible for avirulence on *Rpv3.1*. Resistance breakdown occurred when strains lost both effectors on both homologous chromosomes. Multiple *avrRpv3.1* alleles were identified among isolates with varying degrees of virulence on *Rpv3.1*, suggesting that resistance breakdown evolved independently multiple times at the same locus. This study emphasizes the need to use the diploid genome of an avirulent strain as reference to identify avirulence loci and understand the genomic mechanisms behind resistance breakdown.

Genomic signatures of adaptation in *Erysiphe necator*

Populations of the ascomycete *E. necator*, the causal agent of PM, have been shown to adapt quickly to selective pressures, particularly from fungicides (Miller and Gubler, 2004). Genetic modifications leading to fungicide resistance have been described (Délye et al., 1997), also using whole-genome analysis approaches. For example, genome structure analyses determined that increased gene copy numbers of the dimethyl inhibitor fungicide target *EnCYP51* provides additively increasing tolerance to this class of fungicides (Jones et al., 2014). The *E. necator* genes associated with genetic resistance breakdown have not been described yet, although PM resistance breakdown has been reported (Feechan et al., 2015). Multiple genome assemblies of *E. necator* (Table 1; Supplementary Table

S2) have enabled the identification of candidate secreted effector proteins (CSEPs) potentially involved in suppressing grapevine defense responses and therefore targets of grape *R* genes (Jones et al., 2014; Zaccaron et al., 2023). CSEP080 and CSEP087 were shown to translocate to the host chloroplast and plasma membrane where they suppress hydrogen peroxide accumulation and promote pectin degradation (Mu et al., 2022, 2023). Broader characterization of the molecular mechanisms underlying grapevine–*E. necator* interactions could help determine how some strains overcome disease resistance. However, no sequenced strains have been linked to *R* locus breakdown yet. Screening multiple *E. necator* strains on PM-resistant grape accessions and developing more genomic resources, such as genome assemblies of virulent and avirulent isolates, are necessary to identify avirulence genes linked to resistance breakdown.

Genomic insights into the evolution of other grapevine pathogens

Adaptive genomic structural variability has also been observed in other grapevine pathogens (Massonnet et al., 2018a, b; Garcia et al., 2024). A total of 93 genomes of causal agents of grapevine diseases, representing 19 species, have been sequenced and assembled. This extensive genomic characterization provides a comprehensive view of the vast arsenal of putative virulence factors and the first insights into how these pathogens evolve (Table 1; Supplementary Table S2). Comparative genome and transcriptome analyses of *Phaeoacremonium minimum*, a causative agent of Esca, revealed extensive structural variation among isolates (Massonnet et al., 2018b). These variations include presence/absence polymorphisms involving blocks of multiple adjacent virulence factors, often linked to secondary metabolism. This suggests that the gain or loss of secondary metabolism functions may have an adaptive effect on fitness. Graph-based pangenomes of this pathogen, as well as other grape pathogens such as *Eutypa lata* and *Neofusicoccum parvum*, highlighted dispensable genomic regions enriched with gypsy

Table 1. Summary of the genomes from causal agents of grape diseases

Grape disease	Species	Number of isolates	DNA sequencing data	
			Long reads	Short reads
Black rot	<i>Phyllosticta ampellicida</i> (syn. <i>Guignardia bidwellii</i>)	1		1
Bunch rot; gray mold	<i>Botrytis cinerea</i>	1	1	
Downy mildew	<i>Plasmopara viticola</i>	3	2	2
GTD— <i>Botryosphaeria</i> dieback	<i>Botryosphaeria dothidea</i>	1	1	
	<i>Diplodia seriata</i>	1		1
	<i>Dothiorella viticola</i>	1		1
	<i>Lasiodiplodia exigua</i>	1		1
	<i>Lasiodiplodia missouriana</i>	1		1
	<i>Lasiodiplodia theobromae</i>	1		1
	<i>Neofusicoccum australe</i>	1		1
	<i>Neofusicoccum mediterraneum</i>	1		1
	<i>Neofusicoccum parvum</i>	16	1	15
	<i>Neoscytalidium dimidiatum</i>	1		1
GTD—Esca	<i>Phaeoacremonium minimum</i>	6	1	5
GTD—Esca complex	<i>Fomitiporia mediterranea</i>	2	1	1
	<i>Fomitiporia polymorpha</i>	1	1	
	<i>Inonotus vitis</i>	1	1	
	<i>Phaeoniella chlamydospora</i>	1		1
	<i>Stereum hirsutum</i>	1		1
	<i>Tropicoporus texanus</i>	1	1	
GTD— <i>Eutypa</i> dieback	<i>Eutypa lata</i>	54	3	51
GTD—Phomopsis dieback	<i>Diaporthe ampelina</i>	1		1
Phylloxera	<i>Daktulosphaira vitifoliae</i>	1	1	1
Pierce's disease	<i>Xylella fastidiosa</i> ssp. <i>fastidiosa</i>	180	5	181
Powdery mildew	<i>Erysiphe necator</i>	6	1	6
Root-knot nematode	<i>Meloidogyne javanica</i>	1		1
White rot	<i>Coniothyrium diplodiella</i>	1		1

Details can be found in [Supplementary Table S2](#). Abbreviations: GTD, grapevine trunk disease.

transposable elements and virulence factors under positive selection, such as polyketide synthase genes in *E. lata* and *P. minimum*, and glycosyltransferases in *N. parvum* (Garcia et al., 2024). Whether the observed genomic structural variability plays a role in host adaptation remains to be established. However, the broad variability in virulence observed among isolates is likely to be due to differences in virulence factor content, properties, and expression during infection. Although *R* loci for all but one of the grapevine trunk pathogens are yet to be discovered and utilized in breeding programs (Barba et al., 2018), research on trunk diseases should expand to understanding the molecular basis of resistance. This is particularly important as these pathogens exhibit more necrotrophic behavior compared with the better-studied biotrophic pathogens.

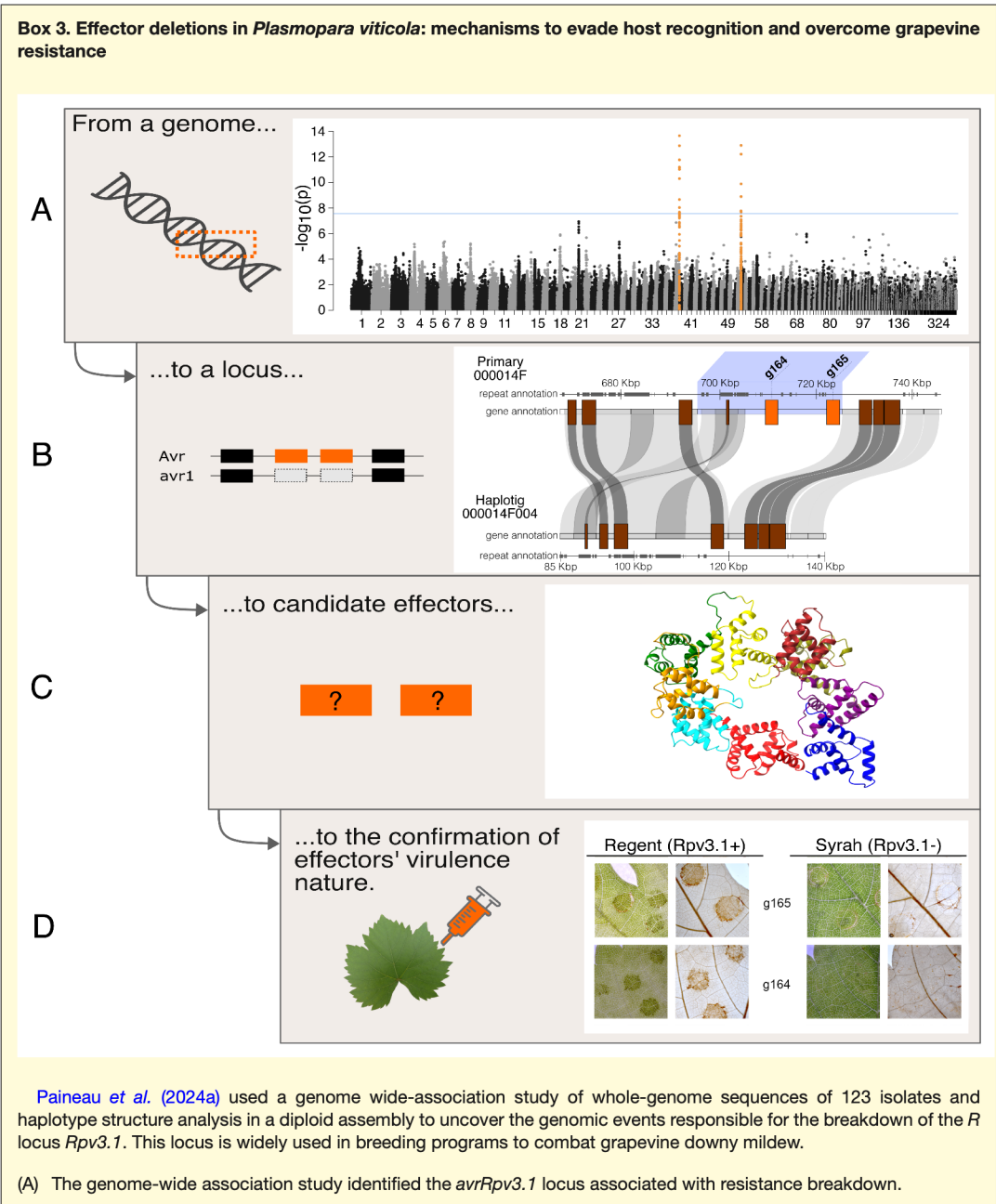
Although there have been no reports of *PdR1* breakdown by *X. fastidiosa*, there is reasonable concern that widespread adoption of grape cultivars carrying only *PdR1* could drive the evolution of virulence in *X. fastidiosa* populations, especially since PD can only be controlled through genetic resistance. *Xylella fastidiosa* populations have been shown to undergo frequent homologous recombination, which allows rapid incorporation

of circulating alleles (Castillo and Almeida, 2023). This process might facilitate quick adaptation to new hosts, such as hypothetically *PdR1*-resistant cultivars, particularly since genes associated with host defense evasion and colonization have higher homologous recombination rates (Castillo and Almeida, 2023).

Conclusions and perspectives

Advancements in grape genomics, from the first consensus haploid genome to phased diploid chromosome-scale genomes, have enabled the identification of 14 candidate *R* genes across five *R* loci. With the decreasing cost of long-read DNA sequencing and the development of solutions for haplotype phasing, we can expect many additional grape *R* loci to be fully described using similar approaches. However, only the availability of large genetic mapping populations and fine-mapping will accelerate the identification of candidate genes.

Beyond grape genomics, discovering new *R* loci to pyramidize a maximum of functionally diverse *R* loci in new cultivars is essential for achieving durable and robust disease



Box 3. Continued

- (B) A diploid-aware reassembly of the strain genome used as reference for the GWAS, *P. viticola* INRA-Pv221 revealed structural variations at the *avrRpv3.1* locus, involving the deletion of two putative effectors.
- (C) These two putative effectors code for proteins with LWY-fold structural modules, a domain specific to oomycete effectors.
- (D) In the presence of these effectors, resistant plants recognize them and trigger cell death. Their deletion allows the pathogen to evade recognition, enabling infection of the plant.

(Figures adapted from Paineau *et al.*, 2024a)

resistance. Maintaining and expanding germplasm collections that capture broad genetic variability among *Vitis* spp. is invaluable for researching new *R* loci (Cantu *et al.*, 2024). With the development of high-throughput phenotyping approaches, reliable phenotypic data from large germplasm collections, ideally grown in multiple locations, combined with whole-genome sequencing data can be used for pangenome-wide association analysis to rapidly identify new *R* loci. Using this approach, a new locus associated with salt tolerance and its underlying gene candidates were recently identified in the pangenome of North American wild grape species by studying individuals from complex wild populations comprising several *Vitis* spp. and hybrids (Cochetel *et al.*, 2023).

An important step in grape *R* gene characterization is the testing of candidate genes through stable knock in in a susceptible *Vv vinifera* background and knockout in the source of the *R* locus. Knockout using CRISPR/Cas9 and knockdown through RNAi have proven effective in assessing the role of PM and DM *S* genes (Pessina *et al.*, 2016; Wan *et al.*, 2020; Marcianò *et al.*, 2021; Giacomelli *et al.*, 2023). However, transgenesis in grapes is challenging due to low efficiency of *in vitro* regeneration of transformed plants. Recent groundbreaking research has shown promising advances in the acceleration of regeneration of genome-edited plants from embryogenic calli-derived protoplasts from multiple genotypes, including *Vitis* spp. (Tricoli and Debernardi, 2024) and recalcitrant cultivars (Gambino *et al.*, 2024). Testing candidate *R* genes across different genetic backgrounds will require protocol development and optimization for multiple *Vitis* spp.

Due to the reported breakdown of several grape *R* loci, it is crucial to understand how pathogens adapt and overcome these loci to achieve durable resistance. Like grape genomics, pathogenomics has advanced significantly in recent years. Although the number of genome assemblies of grape pathogens is increasing, it is essential to develop genome assemblies of isolates that have overcome host resistance, as well as closely related avirulent strains, to uncover the molecular mechanisms behind host resistance breakdown. The growing availability of genomic resources for grapevine and its pathogens enables advanced methods such as co-GWAS, which integrates host and pathogen genomic and phenotypic data to identify co-evolving loci (reviewed in Märkle *et al.*, 2021). Additionally,

constructing pathogen pangenomes has proven valuable in studying genomic variability among strains (Garcia *et al.*, 2024).

Overall, the integration of advanced genomics and pathogenomics approaches will be critical for developing durable disease-resistant grape varieties. Understanding the molecular basis of host resistance and pathogen adaptation will guide breeding strategies and ensure long-term sustainability in viticulture.

Supplementary data

The following supplementary data are available at [JXB online](#).

Table S1. Disease resistance loci identified in grapes.

Table S2. Grape pathogen genome assemblies.

Acknowledgements

We thank Drs Sarah Cookson, Nathalie Ollat, and Eric Gomès, editors of the Special Issue on 'Grapevine', for inviting us to contribute this article. We also appreciate the valuable feedback by the two anonymous reviewers.

Author contributions

MP, MZ, MM, and DC: co-wrote the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Funding

This work was funded by the US Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA) Award #2022-51181-38240 and partially supported by the Ray Rossi Endowment in Viticulture and Enology and the E.J. Gallo Winery.

Data availability

Supplementary Tables S1 and S2 are available at Zenodo at <http://doi.org/10.5281/zenodo.14009076> (Paineau *et al.*, 2025b).

References

- Ainley WM, Sastry-Dent L, Welter ME, *et al.* 2013. Trait stacking via targeted genome editing. *Plant Biotechnology Journal* **11**, 1126–1134.
- Barba P, Lillis J, Luce RS, Travadon R, Osier M, Baumgartner K, Wilcox WF, Reisch BI, Cadle-Davidson L. 2018. Two dominant loci determine resistance to Phomopsis cane lesions in F1 families of hybrid grapevines. *Theoretical and Applied Genetics* **131**, 1173–1189.
- Blum M, Waldner M, Gisi U. 2010. A single point mutation in the novel *PvCesA3* gene confers resistance to the carboxylic acid amide fungicide mandipropamid in *Plasmopara viticola*. *Fungal Genetics and Biology* **47**, 499–510.
- Brewer MT, Milgroom MG. 2010. Phylogeography and population structure of the grape powdery mildew fungus, *Erysiphe necator*, from diverse *Vitis* species. *BMC Evolutionary Biology* **10**, 268.
- Cantu D, Massonnet M, Cochetel N. 2024. The wild side of grape genomics. *Trends in Genetics* **40**, 601–612.
- Castillo AI, Almeida RPP. 2023. The multifaceted role of homologous recombination in a fastidious bacterial plant pathogen. *Applied and Environmental Microbiology* **89**, e0043923.
- Chen W-J, Delmotte F, Cervera SR, Douence L, Greif C, Corio-Costet M-F. 2007. At least two origins of fungicide resistance in grapevine downy mildew populations. *Applied and Environmental Microbiology* **73**, 5162–5172.
- Chin C-S, Peluso P, Sedlazeck FJ, *et al.* 2016. Phased diploid genome assembly with single-molecule real-time sequencing. *Nature Methods* **13**, 1050–1054.
- Cochetel N, Minio A, Guarracino A, *et al.* 2023. A super-pangenome of the North American wild grape species. *Genome Biology* **24**, 290.
- Cochetel N, Minio A, Massonnet M, Vondras AM, Figueroa-Balderas R, Cantu D. 2021. Diploid chromosome-scale assembly of the *Muscadinia rotundifolia* genome supports chromosome fusion and disease resistance gene expansion during *Vitis* and *Muscadinia* divergence. *G3 Genes|Genomes|Genetics* **11**, jkab033.
- Combiér M, Evangelisti E, Piron M-C, Rengel D, Legrand L, Shenhav L, Bouchez O, Schornack S, Mestre P. 2019. A secreted WY-domain-containing protein present in European isolates of the oomycete *Plasmopara viticola* induces cell death in grapevine and tobacco species. *PLoS One* **14**, e0220184.
- Delmas CEL, Fabre F, Jolivet J, Mazet ID, Cervera SR, Delière L, Delmotte F. 2016. Adaptation of a plant pathogen to partial host resistance: selection for greater aggressiveness in grapevine downy mildew. *Evolutionary Applications* **9**, 709–725.
- Délye C, Laigret F, Corio-Costet MF. 1997. A mutation in the 14 alpha-demethylase gene of *Uncinula necator* that correlates with resistance to a sterol biosynthesis inhibitor. *Applied and Environmental Microbiology* **63**, 2966–2970.
- Dong Y, Duan S, Xia Q, *et al.* 2023. Dual domestications and origin of traits in grapevine evolution. *Science* **379**, 892–901.
- Dormatey R, Sun C, Ali K, Coulter JA, Bi Z, Bai J. 2020. Gene pyramiding for sustainable crop improvement against biotic and abiotic stresses. *Agronomy* **10**, 1255.
- Dussert Y, Mazet ID, Couture C, Gouzy J, Piron M-C, Kuchly C, Bouchez O, Rispe C, Mestre P, Delmotte F. 2019. A high-quality grapevine downy mildew genome assembly reveals rapidly evolving and lineage-specific putative host adaptation genes. *Genome Biology and Evolution* **11**, 954–969.
- Feechan A, Anderson C, Torregrosa L, *et al.* 2013. Genetic dissection of a TIR-NB-LRR locus from the wild North American grapevine species *Muscadinia rotundifolia* identifies paralogous genes conferring resistance to major fungal and oomycete pathogens in cultivated grapevine. *The Plant Journal* **76**, 661–674.
- Feechan A, Kocsis M, Riaz S, Zhang W, Gadoury DM, Walker MA, Dry IB, Reisch B, Cadle-Davidson L. 2015. Strategies for *RUN1* deployment using *RUN2* and *REN2* to manage grapevine powdery mildew informed by studies of race specificity. *Phytopathology* **105**, 1104–1113.
- Fontaine MC, Labbé F, Dussert Y, Delière L, Richart-Cervera S, Giraud T, Delmotte F. 2021. Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*. *Current Biology* **31**, 2155–2166.
- Foria S, Copetti D, Eisenmann B, *et al.* 2020. Gene duplication and transposition of mobile elements drive evolution of the *Rpv3* resistance locus in grapevine. *The Plant Journal* **101**, 529–542.
- Foria S, Magris G, Jurman I, Schwöpe R, De Candido M, De Luca E, Ivanišević D, Morgante M, Di Gaspero G. 2022. Extent of wild-to-crop interspecific introgression in grapevine (*Vitis vinifera*) as a consequence of resistance breeding and implications for the crop species definition. *Horticulture Research* **9**, uhab010.
- Frommer B, Müllner S, Holtgräwe D, Viehöver P, Huettel B, Töpfer R, Weisshaar B, Zyprian E. 2023. Phased grapevine genome sequence of an *Rpv12* carrier for biotechnological exploration of resistance to *Plasmopara viticola*. *Frontiers in Plant Science* **14**, 1180982.
- Fu Q, Chen T, Wang Y, *et al.* 2024. *Plasmopara viticola* effector PvCRN20 represses the import of VvDEG5 into chloroplasts to suppress immunity in grapevine. *New Phytologist* **243**, 2311–2331.
- Gambino G, Nuzzo F, Moine A, *et al.* 2024. Genome editing of a recalcitrant wine grape genotype by lipofectamine-mediated delivery of CRISPR/Cas9 ribonucleoproteins to protoplasts. *The Plant Journal* **119**, 404–412.
- García JF, Morales-Cruz A, Cochetel N, Minio A, Figueroa-Balderas R, Rolshausen PE, Baumgartner K, Cantu D. 2024. Comparative pangenomic insights into the distinct evolution of virulence factors among grapevine trunk pathogens. *Molecular Plant-Microbe Interactions* **37**, 127–142.
- García-Ruiz H, Szurek B, den Ackerveken GV. 2021. Stop helping pathogens: engineering plant susceptibility genes for durable resistance. *Current Opinion in Biotechnology* **70**, 187.
- Giacomelli L, Zeilmaker T, Giovannini O, Salvagnin U, Masuero D, Franceschi P, Vrhovsek U, Scintilla S, Rouppe van der Voort J, Moser C. 2023. Simultaneous editing of two *DMR6* genes in grapevine results in reduced susceptibility to downy mildew. *Frontiers in Plant Science* **14**, 1242240.
- Gosavi G, Yan F, Ren B, Kuang Y, Yan D, Zhou X, Zhou H. 2020. Applications of CRISPR technology in studying plant–pathogen interactions: overview and perspective. *Phytopathology Research* **2**, 21.
- Heyman L, Höfle R, Kicherer A, Trapp O, Ait Barka E, Töpfer R, Höfte M. 2021. The durability of quantitative host resistance and variability in pathogen virulence in the interaction between European grapevine cultivars and *Plasmopara viticola*. *Frontiers in Agronomy* **3**, 39.
- Huerta-Acosta KG, Riaz S, Tenschler A, Walker MA. 2022. Genetic characterization of Pierce's disease resistance in a *Vitis arizonica/monticola* wild grapevine. *American Journal of Enology and Viticulture* **74**, 0740003.
- Jaillon O, Aury JM, Noel B, *et al.*; French-Italian Public Consortium for Grapevine Genome Characterization 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**, 463–467.
- Jones L, Riaz S, Morales-Cruz A, Amrine KC, McGuire B, Gubler WD, Walker MA, Cantu D. 2014. Adaptive genomic structural variation in the grape powdery mildew pathogen, *Erysiphe necator*. *BMC Genomics* **15**, 1081.
- Khoury CK, Greene SL, Krishnan S, Miller AJ, Moreau T. 2019. A road map for conservation, use, and public engagement around North America's crop wild relatives and wild utilized plants. *Crop Science* **59**, 2302–2307.
- Kourelis J, Sakai T, Adachi H, Kamoun S. 2021. RefPlantNLR is a comprehensive collection of experimentally validated plant disease resistance proteins from the NLR family. *PLoS Biology* **19**, e3001124.
- Lan X, Liu Y, Song S, Yin L, Xiang J, Qu J, Lu J. 2019. *Plasmopara viticola* effector PvRXLR131 suppresses plant immunity by targeting plant receptor-like kinase inhibitor BKI1. *Molecular Plant Pathology* **20**, 765–783.
- Lei X, Lan X, Ye W, Liu Y, Song S, Lu J. 2019. *Plasmopara viticola* effector PvRXLR159 suppresses immune responses in *Nicotiana benthamiana*. *Plant Signaling & Behavior* **14**, 1682220.

- Li P, Quan X, Jia G, Xiao J, Cloutier S, You FM. 2016. RGAugury: a pipeline for genome-wide prediction of resistance gene analogs (RGAs) in plants. *BMC Genomics* **17**, 852.
- Liu J, Chen S, Ma T, Gao Y, Song S, Ye W, Lu J. 2021. *Plasmopara viticola* effector PvRXLR53 suppresses innate immunity in *Nicotiana benthamiana*. *Plant Signaling & Behavior* **16**, 1846927.
- Lukić K, Zdunić G, Hančević K, Mihaljević MZ, Mucalo A, Maul E, Riaz S, Pejić I. 2022. Identification of powdery mildew resistance in wild grapevine (*Vitis vinifera* subsp. *sylvestris* Gmel Hegl) from Croatia and Bosnia and Herzegovina. *Scientific Reports* **12**, 2128.
- Ma T, Chen S, Liu J, Fu P, Wu W, Song S, Gao Y, Ye W, Lu J. 2021. *Plasmopara viticola* effector PvRXLR111 stabilizes VWRKY40 to promote virulence. *Molecular Plant Pathology* **22**, 231–242.
- Marcianò D, Ricciardi V, Marone Fassolo E, Passera A, Bianco PA, Failla O, Casati P, Maddalena G, De Lorenzis G, Toffolatti SL. 2021. RNAi of a putative grapevine susceptibility gene as a possible downy mildew control strategy. *Frontiers in Plant Science* **12**, 667319.
- Margaryan K, Töpfer R, Gasparyan B, Arakelyan A, Trapp O, Röckel F, Maul E. 2023. Wild grapes of Armenia: unexplored source of genetic diversity and disease resistance. *Frontiers in Plant Science* **14**, 1276764.
- Märkle H, John S, Cornille A, Fields PD, Tellier A. 2021. Novel genomic approaches to study antagonistic coevolution between hosts and parasites. *Molecular Ecology* **30**, 3660–3676.
- Massonnet M, Morales-Cruz A, Figueroa-Balderas R, Lawrence DP, Baumgartner K, Cantu D. 2018a. Condition-dependent co-regulation of genomic clusters of virulence factors in the grapevine trunk pathogen *Neofusicoccum parvum*. *Molecular Plant Pathology* **19**, 21–34.
- Massonnet M, Morales-Cruz A, Minio A, Figueroa-Balderas R, Lawrence DP, Travadon R, Rolshausen PE, Baumgartner K, Cantu D. 2018b. Whole-genome resequencing and pan-transcriptome reconstruction highlight the impact of genomic structural variation on secondary metabolite gene clusters in the grapevine Esca pathogen *Phaeoacremonium minimum*. *Frontiers in Microbiology* **9**, 1784.
- Massonnet M, Vondras AM, Cochetel N, Riaz S, Pap D, Minio A, Figueroa-Balderas R, Walker MA, Cantu D. 2022. Haplotype-resolved powdery mildew resistance loci reveal the impact of heterozygous structural variation on NLR genes in *Muscadinia rotundifolia*. *G3 Genes|Genomes|Genetics* **12**, jkac148.
- Merdinoglu D, Schneider C, Prado E, Wiedemann-Merdinoglu S, Mestre P. 2018. Breeding for durable resistance to downy and powdery mildew in grapevine. *OENO One* **52**, 203–209.
- Mestre P, Carrere S, Gouzy J, Piron M-C, Labrousse DT de, Vincourt P, Delmotte F, Godiard L. 2016. Comparative analysis of expressed CRN and RXLR effectors from two *Plasmopara* species causing grapevine and sunflower downy mildew. *Plant Pathology* **65**, 767–781.
- Mestre P, Piron M-C, Merdinoglu D. 2012. Identification of effector genes from the phytopathogenic oomycete *Plasmopara viticola* through the analysis of gene expression in germinated zoospores. *Fungal Biology* **116**, 825–835.
- Michelmore RW, Christopoulou M, Caldwell KS. 2013. Impacts of resistance gene genetics, function, and evolution on a durable future. *Annual Review of Phytopathology* **51**, 291–319.
- Miller TC, Gubler WD. 2004. Sensitivity of California isolates of *Uncinula necator* to trifloxystrobin and spiroxamine, and update on triadimefon sensitivity. *Plant Disease* **88**, 1205–1212.
- Minio A, Cochetel N, Massonnet M, Figueroa-Balderas R, Cantu D. 2022a. HiFi chromosome-scale diploid assemblies of the grape rootstocks 110R, Kober 5BB, and 101–14 Mgt. *Scientific Data* **9**, 660.
- Minio A, Cochetel N, Vondras AM, Massonnet M, Cantu D. 2022b. Assembly of complete diploid-phased chromosomes from draft genome sequences. *G3 Genes|Genomes|Genetics* **12**, jkac143.
- Morales-Cruz A, Aguirre-Liguori J, Massonnet M, et al. 2023. Multigenic resistance to *Xylella fastidiosa* in wild grapes (*Vitis* spp.) and its implications within a changing climate. *Communications Biology* **6**, 1–15.
- Moura MF, de Oliveira GL, Rodrigues CS, Paioli-Pires EJ. 2024. The role of plant breeding in grapevine production. In: Martins-Lopes P. ed. *Grapevine: from origin to the vineyard*. Advances in Botanical Research Vol. 110. Academic Press, 255–294.
- Mu B, Chen J, Wang H, Kong W, Fan X, Wen Y-Q. 2022. An effector CSEP087 from *Erysiphe necator* targets arginine decarboxylase VviADC to regulate host immunity in grapevine. *Scientia Horticulturae* **303**, 111205.
- Mu B, Teng Z, Tang R, Lu M, Chen J, Xu X, Wen Y-Q. 2023. An effector of *Erysiphe necator* translocates to chloroplasts and plasma membrane to suppress host immunity in grapevine. *Horticulture Research* **10**, uhad163.
- Osuna-Cruz CM, Paytuvi-Gallart A, Di Donato A, Sundesha V, Andolfo G, Aiese Cigliano R, Sanseverino W, Ercolano MR. 2017. PRGdb 3.0: a comprehensive platform for prediction and analysis of plant disease resistance genes. *Nucleic Acids Research* **46**, D1197.
- Paineau M, Mazet ID, Wiedemann-Merdinoglu S, Fabre F, Delmotte F. 2022. The Characterization of pathotypes in grapevine downy mildew provides insights into the breakdown of *Rpv3*, *Rpv10*, and *Rpv12* factors in grapevines. *Phytopathology* **112**, 2329–2340.
- Paineau M, Minio A, Mestre P, Fabre F, Mazet ID, Couture C, Legeai F, Dumartinet T, Cantu D, Delmotte F. 2024a. Multiple deletions of candidate effector genes lead to the breakdown of partial grapevine resistance to downy mildew. *New Phytologist* **243**, 1490–1505.
- Paineau M, Zaccaro M, Massonnet M, Cantu D. 2025b. Data from: Advances in grape and pathogen genomics toward durable grapevine disease resistance. Zenodo <https://doi.org/10.5281/zenodo.14009076>
- Pessina S, Lenzi L, Perazzolli M, Campa M, Dalla Costa L, Urso S, Valè G, Salamini F, Velasco R, Malnoy M. 2016. Knockdown of MLO genes reduces susceptibility to powdery mildew in grapevine. *Horticulture Research* **3**, 16016.
- Pirrello C, Zeilmaker T, Bianco L, Giacomelli L, Moser C, Vezzulli S. 2021. Mining grapevine downy mildew susceptibility genes: a resource for genomics-based breeding and tailored gene editing. *Biomolecules* **11**, 181.
- Rajput R, Naik J, Misra P, Trivedi PK, Pandey A. 2023. Gene pyramiding in transgenic plant development: approaches and challenges. *Journal of Plant Growth Regulation* **42**, 6038–6056.
- Riaz S, Boursiquot J-M, Dangl GS, Lacombe T, Laucou V, Tenschler AC, Walker MA. 2013. Identification of mildew resistance in wild and cultivated Central Asian grape germplasm. *BMC Plant Biology* **13**, 149.
- Riaz S, Huerta-Acosta K, Tenschler AC, Walker MA. 2018. Genetic characterization of *Vitis* germplasm collected from the southwestern US and Mexico to expedite Pierce's disease-resistance breeding. *Theoretical and Applied Genetics* **131**, 1589–1602.
- Riaz S, Menéndez CM, Tenschler A, Pap D, Walker MA. 2020. Genetic mapping and survey of powdery mildew resistance in the wild Central Asian ancestor of cultivated grapevines in Central Asia. *Horticulture Research* **7**, 1–15.
- Riaz S, Tenschler AC, Rubin J, Graziani R, Pao SS, Walker MA. 2008. Fine-scale genetic mapping of two Pierce's disease resistance loci and a major segregation distortion region on chromosome 14 of grape. *Theoretical and Applied Genetics* **117**, 671–681.
- Sánchez-Vallet A, Hartmann FE, Marcel TC, Croll D. 2018. Nature's genetic screens: using genome-wide association studies for effector discovery: nature's genetic screens. *Molecular Plant Pathology* **19**, 3–6.
- Sargolzaei M, Maddalena G, Bitsadze N, Maghradze D, Bianco PA, Failla O, Toffolatti SL, De Lorenzis G. 2020. *Rpv29*, *Rpv30* and *Rpv31*: three novel genomic loci associated with resistance to *Plasmopara viticola* in *Vitis vinifera*. *Frontiers in Plant Science* **11**, 562432.
- Shi X, Cao S, Wang X, et al. 2023. The complete reference genome for grapevine (*Vitis vinifera* L.) genetics and breeding. *Horticulture Research* **10**, uhad061.
- Steuernagel B, Witek K, Krattinger SG, et al. 2020. The NLR-Annotator tool enables annotation of the intracellular immune receptor repertoire. *Plant Physiology* **183**, 468–482.
- Tricoli DM, Debernardi JM. 2024. An efficient protoplast-based genome editing protocol for *Vitis* species. *Horticulture Research* **11**, uhad266.

- Vecchio R, Pomarici E, Giampietri E, Borrello M.** 2022. Consumer acceptance of fungus-resistant grape wines: evidence from Italy, the UK, and the USA. *PLoS One* **17**, e0267198.
- Velt A, Frommer B, Blanc S, et al.** 2023. An improved reference of the grapevine genome reasserts the origin of the PN40024 highly homozygous genotype. *G3 Genes|Genomes|Genetics* **13**, jkad067.
- Vezzulli S, Doligez A, Bellin D.** 2019. Molecular mapping of grapevine genes. In: Cantu D, Walker MA. eds. *Compendium of plant genomes. The grape genome*. Cham: Springer International Publishing, 103–136.
- Wan D-Y, Guo Y, Cheng Y, Hu Y, Xiao S, Wang Y, Wen Y-Q.** 2020. CRISPR/Cas9-mediated mutagenesis of *VvMLO3* results in enhanced resistance to powdery mildew in grapevine (*Vitis vinifera*). *Horticulture Research* **7**, 116.
- Wingarter C, Eisenmann B, Weber P, Dry I, Bogs J.** 2021. Grapevine *Rpv3*-, *Rpv10*- and *Rpv12*-mediated defense responses against *Plasmopara viticola* and the impact of their deployment on fungicide use in viticulture. *BMC Plant Biology* **21**, 470.
- Zaccaron AZ, Neill T, Corcoran J, Mahaffee WF, Stergiopoulos I.** 2023. A chromosome-scale genome assembly of the grape powdery mildew pathogen *Erysiphe necator* reveals its genomic architecture and previously unknown features of its biology. *mBio* **14**, e0064523.
- Zhou Y, Massonnet M, Sanjak JS, Cantu D, Gaut BS.** 2017. Evolutionary genomics of grape (*Vitis vinifera* ssp. *vinifera*) domestication. *Proceedings of the National Academy of Sciences, USA* **114**, 11715–11720.
- Zhou Y, Minio A, Massonnet M, Solares E, Lv Y, Beridze T, Cantu D, Gaut BS.** 2019. The population genetics of structural variants in grapevine domestication. *Nature Plants* **5**, 965–979.
- Zou C, Sapkota S, Figueroa-Balderas R, Glaubitz J, Cantu D, Kingham BF, Sun Q, Cadle-Davidson L.** 2023. A multitiered haplotype strategy to enhance phased assembly and fine mapping of a disease resistance locus. *Plant Physiology* **193**, 2321–2336.

CHAPTER 2

Multigenic resistance to *Xylella fastidiosa* in wild grapes (*Vitis* spp.) and its implications within a changing climate

Multigenic resistance to *Xylella fastidiosa* in wild grapes (*Vitis* spp.) and its implications within a changing climate

Abraham Morales-Cruz^{1,9} , Jonas Aguirre-Liguori^{2,9} , Mélanie Massonnet³, Andrea Minio³, Mirella Zaccheo³, Noe Cochetel³ , Andrew Walker³, Summaira Riaz⁴ , Yongfeng Zhou^{5,6} , Dario Cantu^{3,7}  & Brandon S. Gaut^{2,8} 

Xylella fastidiosa is a bacterium that infects crops like grapevines, coffee, almonds, citrus and olives. There is little understanding of the genes that contribute to plant resistance, the genomic architecture of resistance, and the potential role of climate in shaping resistance, in part because major crops like grapevines (*Vitis vinifera*) are not resistant to the bacterium. Here we study a wild grapevine species, *V. arizonica*, that segregates for resistance. Using genome-wide association, we identify candidate resistance genes. Resistance-associated kmers are shared with a sister species of *V. arizonica* but not with more distant species, suggesting that resistance evolved more than once. Finally, resistance is climate dependent, because individuals from low (<10 °C) temperature locations in the wettest quarter were typically susceptible to infection, likely reflecting a lack of pathogen pressure in colder climates. In fact, climate is as effective a predictor of resistance phenotypes as some genetic markers. We extend our climate observations to additional crops, predicting that increased pathogen pressure is more likely for grapevines and almonds than some other susceptible crops.

¹U.S. Department of Energy, Joint Genome Institute, Lawrence Berkeley National Lab, Berkeley, CA 94720, USA. ²Dept. of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA. ³Dept. of Viticulture and Enology, University of California, Davis, CA, USA. ⁴San Joaquin Valley Agricultural Center, United States Dept of Agriculture, Parlier, CA, USA. ⁵Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China. ⁶Present address: Agricultural Genomics Institute at Shenzhen, The Chinese Academy of Agricultural Sciences, No. 7 Pengfei Road, Shenzhen 518120, China. ⁷Present address: Dept. of Viticulture and Enology, One Shields Avenue, University of California Davis, Davis, CA 95616-5270, USA. ⁸Present address: Dept. of Ecology and Evolutionary Biology, 321 Steinhaus Hall UC Irvine, Irvine, CA 92617-2525, USA. ⁹These authors contributed equally: Abraham Morales-Cruz, Jonas Aguirre-Liguori. ✉email: zhouyongfeng@caas.cn; dacantu@ucdavis.edu; bgaut@uci.edu

Climate change is impacting crop yields by shifting temperatures, weather extremes, and water availability¹, thereby affecting the distribution of arable lands². There is, however, another important effect of climate change, which is the altered distribution of plant pathogens^{3,4}. One especially prominent pathogen is the bacterium *Xylella fastidiosa* (Wells). *X. fastidiosa* is a generalist that colonizes > 300 plant species^{5,6}, but it is pathogenic on major crops like citrus, coffee, almonds and grapevines (*Vitis vinifera* L. ssp. *vinifera*). Until recently, *X. fastidiosa* had been limited to the Americas, but human-mediated migration has led to its colonization of Europe, where it causes > ~\$100 M of damage per year to the olive (*Olea europaea* L.) industry⁷. This olive example illustrates that the bacterium is more than a persistent threat in the Americas; it is also an emerging and expanding global threat to Europe, the Middle East⁸ and beyond⁹. Accordingly, there are urgent needs to better understand the genetic mechanisms of plant resistance¹⁰, particularly in the wild where both pathogens and hosts evolve¹¹.

Thus far, studies of *X. fastidiosa*-mediated diseases have focused primarily on citrus^{12,13} and on Pierce's Disease (PD) in domesticated grapevines, but also with an increasing emphasis on olives^{14,15}. In grapevines, PD manifests by colonizing the xylem, leading to vascular blockages and eventual plant death after several years. In the course of infection, PD causes other detrimental symptoms, including marginal leaf necrosis, berry desiccation, irregular maturation of canes and abnormal petiole abscission¹⁶. The bacterium is spread from plant to plant by xylem-feeding insect vectors, which affect the severity and spread disease. The distribution of these insect vectors is being affected by changing climate¹⁷ and by anthropomorphic activity. One pertinent example is the glassy-winged-sharpsooter (GWSS; *Homalodisca vitripennis* Germar), which was introduced to Southern California in the late 1990s. The GWSS has a higher transmission efficiency compared to native vectors and fueled a large PD outbreak that has permanently altered viticulture in the region.

Although all domesticated grapevines are susceptible to PD, some wild relatives of grapevines segregate for PD resistance, likely reflecting the evolution of resistance in regions of persistent *X. fastidiosa* pressure¹⁸. Among wild grapevines, *Vitis arizonica* (Engelm.) merits particular interest because it exhibits strong resistance to PD and because it contains the only characterized plant locus to segregate for *X. fastidiosa* resistance, the *Pierce's disease resistance 1* (*PdR1*) locus^{19,20}. *PdR1* was identified by genetic mapping of a segregating family, defined by simple-sequence-repeat (SSR) markers, and backcrossed into susceptible grapevine cultivars to introduce resistance²¹. A recent study utilized BAC sequences of the region to identify candidate genes for resistance²². Two canonical leucine-rich receptor (LRR) loci were transformed into *V. vinifera*, but neither conferred resistance²². Additional candidate resistance genes have been identified based on comparative transcriptomics and proteomics in *V. vinifera*^{23,24}, olives¹⁵ and citrus (*Citrus reticulata* Blanco)²⁵.

Despite the enormous economic impact of *X. fastidiosa* infection, the genomic architecture of resistance has not yet been investigated in any species, and the genomic basis of resistance remains unclear. Here we address this shortcoming by performing genome-wide association (GWA) analyses for *X. fastidiosa* resistance in *V. arizonica*. In addition to identifying several novel candidate genes for resistance in *PdR1* and in other genomic regions, our work begins to fill another surprising gap. Although GWA and similar approaches are commonly used to study disease resistance in crops, surprisingly few studies have focused on the wild relatives of crops¹¹. [One notable exception is the wild relative of soybean, *Glycine soja* (Siebold & Zucc.)^{26,27}.] This dearth of studies is surprising both because crop wild relatives are

a proven and valuable source of resistance genes for crop improvement²⁸ and because studying resistance in wild samples may provide insights into the evolution of resistance and the ecological and climatic factors that shape resistance¹¹.

In this study, we generate landscape genomic data from a sample of *V. arizonica* from throughout its native range and perform GWA based on a resistance phenotype - i.e., bacterial load after experimental inoculation. In doing so, we identify several genomic regions, including the *PdR1* region, that are associated with resistance, and we identify candidate genes in these regions based on an improved *V. arizonica* reference genome. We combine GWA with several types of evidence - including population genetic analyses, gene expression assays, comparisons among wild *Vitis* species, data from *V. vinifera* cultivars bred for PD resistance and bioclimatic modeling - to address three sets of questions. First, which and how many genetic regions contribute to resistance, and what are some of the likely candidate resistance genes within these regions? Second, are these regions implicated in resistance across *Vitis* species and also in cultivars that were specifically bred for PD resistance? What do these inter-species analyses imply about the origin of resistance? Finally, does plant resistance correlate with climate? If so, what might this correlation imply about the potential effects of climate change? Overall, our work provides information about the genetics, evolution and ecology of PD resistance, all of which will help inform strategies to manage an economically damaging and expanding pathogen²⁹.

Results

Genome-wide associations for resistance to Pierce's Disease.

We studied the genetics of PD resistance in *V. arizonica* by combining three sources of information: an updated reference genome (accession b40-14, which is homozygous for PD resistance)³⁰, whole-genome resequencing data from 167 accessions sampled across the species' native range (Supplementary Fig. S1), and previously published PD resistance data measured in a common greenhouse environment on the same set of 167 accessions^{30,31}. We used PD resistance as a quantitative variable - i.e., the log-transformed number of colony forming units (CFUs/ml) 12-14 weeks after experimental *X. fastidiosa* ssp. *fastidiosa* (Wells) inoculations (Supplementary Data S1). However, following precedence³¹, we also characterized individual accessions as resistant if they had *X. fastidiosa* concentrations below 13.0 CFUs/ml. Based on this threshold, our sample contained 135 resistant and 32 susceptible individuals, with the susceptible individuals more common in the northern region of the geographic distribution (Fig. 1).

We first performed genome-wide association (GWA) analyses based on SNP variants. To do so, we mapped resequencing data to the reference haplotype of the phased diploid genome and then tested for associations between high-quality SNPs and PD resistance using LFMM2³² and EMMAX³³, both of which correct for genetic structure (see Methods). On the reference haplotype (hap 1), we identified 74 and 40 associated SNPs (Bonferroni $p < 0.05$) with the two methods, of which 25 were significant with both methods. We used these 25 SNPs to conservatively define eight peaks across five chromosomes (Fig. 2, Supplementary Figs. S2-S4, Supplementary Data S2). The most evident peaks were on chromosomes 14 and 15, with one of the former located between the SSR markers that define the *PdR1* locus. We also called SNPs independently to the second haplotype (hap2) and identified 11 significant SNPs in five peaks (Supplementary Figs. S5-S7, Supplementary Dataset 2). One of these peaks was also on chromosome 14 between the *PdR1* flanking markers.

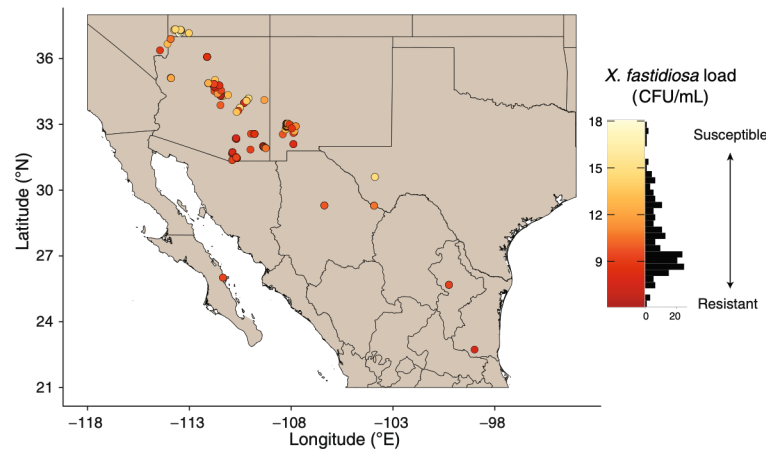


Fig. 1 *Vitis arizonica* sampling and phenotypes. A map of the Southwestern United States and Northern Mexico indicates sampling locations of the $n = 167$ *V. arizonica* accessions used in this study. The color of sample locations (circles) are colored according to their resistance phenotype, as measured by bacterial load (CFU/mL). The histogram of phenotypes (in CFU/mL) is to the right of the map. Map generation relied on information from GADM, a publicly available database (<http://gadm.org>).

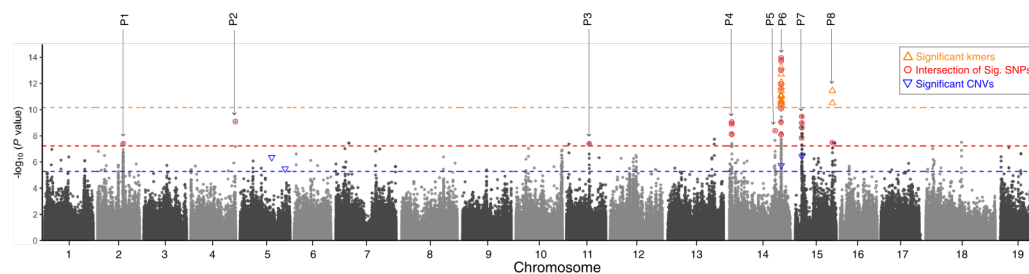


Fig. 2 A Manhattan plot of the *V. arizonica* genome showing markers associated with bacterial load. The plot denotes each of the 19 chromosomes for haplotype 1. Each circle represents a SNP with a corresponding p value, based on EMMAX genome-wide association analysis. The 25 SNPs that were detected in two separate GWA analyses are circled in red and define the 8 peaks of association, which are numbered as P1, P2, etc., and referred to in the text. In addition to SNPs, the locations of significantly associated kmers and CNVs are provided when they overlap with a SNP-defined peak. The colored horizontal lines represent the cut-off p -values ($P < 0.05$, Bonferroni corrected) for the different marker types. Significant ($P < 0.05$, Bonferroni corrected) kmers and CNVs are represented by red and blue triangles, respectively.

Previous studies have suggested that *PdR1* alleles differ in size among *V. arizonica* accessions³⁴, potentially implicating structural variants in resistance. We therefore investigated associations using copy number variants (CNVs), identifying 14,294 CNVs throughout the genome, of which 60 were in the 8 PD-significant peaks (Supplementary Data S3). The 60 CNVs included 19 deletions and 41 duplications, with means of 1.3 and 3.3 copies. We also performed GWA on the complete CNV set, finding four that were significantly (Bonferroni $p < 0.05$) associated with bacterial load (Fig. 2, Supplementary Figs. S8–S9). Two of these four CNVs mapped to two SNP-defined peaks: CNV10605 within *PdR1* (mean copy Number = 0.88, size = 6 kb, $R = -0.36$, $p = 1.77e-06$) and CNV10806 within peak 7 of chromosome 15 (mean copy number = 2.83, size = 17 kb, $R = -0.38$, $p = 3.61e-07$) (Fig. 2). The negative correlations for both CNVs indicated that a higher number of copies had lower bacterial loads and higher PD resistance. However, both CNVs had homology to long-terminal repeat transposable elements and so provided few insights into the functional basis of resistance. To further account for potential structural variation among accessions, we also

applied GWA to 31 bp kmers, using a reference-free approach³⁵ (Fig. 2, Supplementary Fig. S10). Of 115 significant kmers (Bonferroni $p < 0.05$) (Supplementary Data S4), 79 mapped to the reference genome (Supplementary Data S5) and 62 mapped uniquely to either hap1 or hap2. Among the uniquely mapped kmers, 57 of 62 were located on hap1 near *PdR1* and five were in peak 8 on chromosome 15 (Fig. 2). Altogether, CNV and kmer analyses corroborated four of the eight SNP-based peaks, added further evidence of the role of SVs in PD resistance, and confirmed *PdR1* as a candidate locus^{19,20}.

We examined annotated genes under the eight hap1 peaks, using peak boundaries defined by 100 kb windows, since genome-wide LD decayed to background levels ($r^2 < 0.05$) within this distance (Supplementary Fig. S11). The eight peaks included 124 genes, and several had annotations that implied a role in plant immunity (Supplementary Data S6). For example, peak 4 included a gene inferred to have calmodulin-binding function (g226310), which is involved in the regulation of plant disease response through changes in phytohormone biosynthesis^{36,37}. This peak also had a gene annotated as a “syntaxin of plants 41”

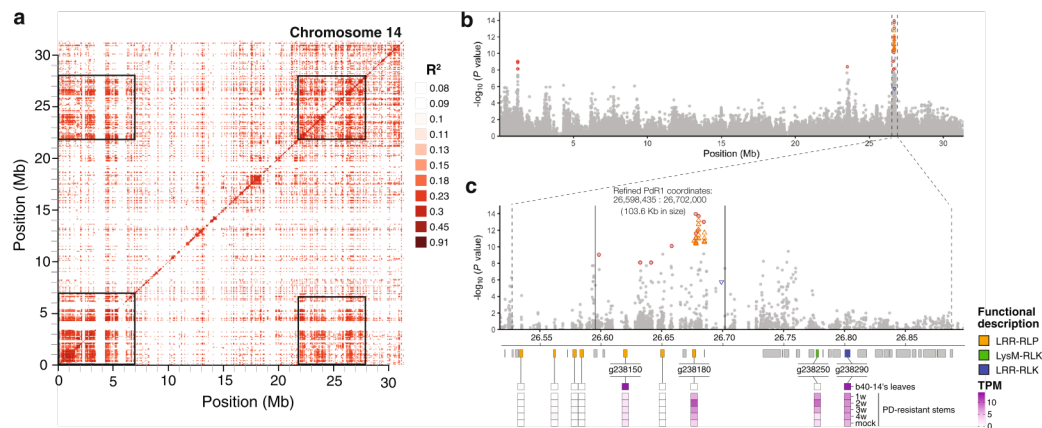


Fig. 3 Genetic analyses of the *PdR1* region. **a** A plot of linkage disequilibrium (LD) across chromosome 14, where darker colors represent higher levels of LD. The two dark squares on the diagonal include GWA peak 4 (on the left, from 0 to 7 Mb on the chromosome) and the *PdR1* peak that corresponds to *PdR1* (on the right, from 22 to 28 Mb on the chromosome). The two off-diagonal squares reflect LD between these two distinct regions. **b** A Manhattan plot of chromosome 14 indicating the locations of peak 4 and the *PdR1* region. **c** An expanded representation of the *PdR1* region showing the location of significant SNPs (red circles denote SNP significant with two GWA methods), significant kmers (green triangles) and CNVs (blue triangles). The dashed vertical lines represent the 361 kb region defined by SSR markers and the 106 kb region defined by the location of significant markers. The schematics below represent genes in *PdR1* and a summary of gene expression results. Genes are denoted by rectangles and colored if they are related to R genes, with the category of R gene indicated by its color. Expression information shows expression, in transcripts per million (TPM) for leaves, for stems during four stages after infection and for mock controls.

gene (g226360) that is homologous to genes that act in plant resistance against bacterial pathogens³⁸. At *PdR1* (peak 6), we identified 7 leucine-rich repeat receptor-like protein (LRR-RLP) genes, one LRR receptor-like protein kinase (RLK) gene, and one lysin motif (LysM) RLK gene, all gene types that are commonly involved in pathogen detection and initiate the plant response³⁹. Peak 7 contained two nucleotide-binding site leucine-rich repeat proteins (NBS-LRR; g243780, and g243820), genes that typically detect pathogens and initiate a host response, as well as a gene annotated as a Phloem protein 2-like (PP2) protein, which have antimicrobial properties in cucumber⁴⁰. We also identified eight genes of interest on chromosome 15 (peak 8). Two of those genes have functional annotations related to phytohormone interactions (g252710, Ethylene-responsive transcription factor CRF4, and g252790, Abscissic Acid Insensitive-like 1 or ABIL 1), and thus could in theory play a role in plant immunity. Another four genes in peak 8 were annotated as acidic endochitinases, which can provide defense against fungal pathogens⁴¹. Finally, we studied the potential gene function associated with the 36 significant kmers that did not map to the reference genome by assembling reads containing the kmers and then aligning the assemblies to the NCBI Transcript Reference Sequences ("refseq_rna"). Of the assembled contigs, 80% had high similarity to three specific Receptor-like proteins (RLPs) ("XM_010648495.2", "XM_034852027.1" and "XM_019224733.1"). Overall, the set of candidate genes suggest that multiple diverse functions may contribute to PD resistance, but with likely involvement of classic disease resistance (R) genes.

We mapped previously used SSR markers to define the *PdR1* locus as a 361 kb region on hap1 chromosome 14 (with a corresponding 360 kb region on hap2), but we further characterized the locus in three ways. First, we evaluated linkage disequilibrium (LD) across chromosome 14. We observed two large blocks (~7 Mb in size) in high LD that contained the three PD-significant peaks of chromosome 14 (peaks 4, 5, and 6), even though peaks 5 and 6 were located on opposite ends of the chromosome from peak 4 (Fig. 3a). This striking pattern may

simply reflect properties of our sample, but it also suggests that PD-related alleles co-segregate across peaks, implying that additive or epistatic interactions contribute to resistance in nature. Second, we focused on the location of significantly associated *PdR1* SNPs, which fell into a narrower 103.6 kb region containing six genes, three of which were RLPs (Fig. 3b, c).

Finally, we assayed gene expression in the region and manually confirmed gene annotations in this region with RNAseq data. One RLP gene (g238150) was expressed in b40-14 leaves, as was a receptor-like-kinase (RLK; g238290) that fell outside the 103.6 kb region (Fig. 3c). We also assayed gene expression in three resistant full-sibs that were inoculated with *X. fastidiosa* and a control (water). The stems above the inoculation site were sampled weekly for up to four weeks. Both g238150 and g238290 were expressed at higher levels than the control in at least one weekly stage, although not significantly so ($p = 0.98$ and $p = 0.54$ for g238150 and g238290). Two additional genes - an RLP (g238180) and an RLK (g239250) - also exhibited this pattern, and g238180 co-located with several associated kmers (Fig. 2). Altogether four candidate R genes (g238150, g238180, g238250, and g238290) under the *PdR1* peak were expressed in PD-resistant stems, but only g238150 and g238290 were found expressed in b40-14 leaves. All four of these genes were also present on the hap2 version of *PdR1*. Importantly, none of these four candidates were the closest homologs of the candidate genes that failed to confer resistance when transformed into *V. vinifera*²² (see Discussion).

The genetic basis of resistance in breeding. The complex LD pattern on chromosome 14 suggests that resistance may require genic action from more than one locus - i.e., multigenic (horizontal) resistance. To investigate this possibility, we examined the distribution of kmers across accessions (Supplementary Data S7). Among the 117 kmers associated with bacterial load, 99 were common among resistant accessions; they were found in

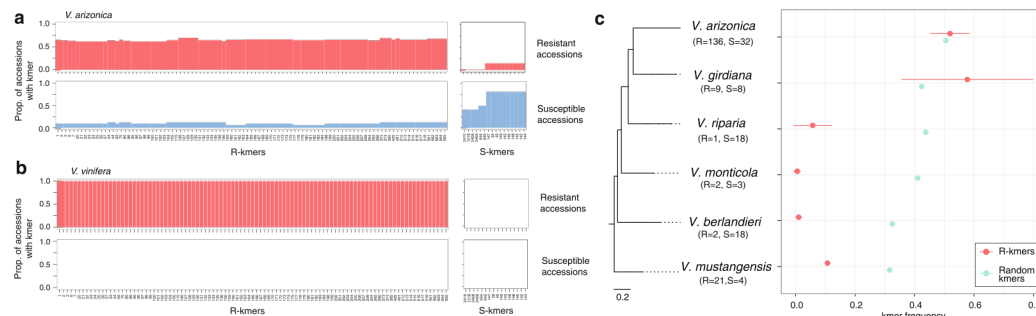


Fig. 4 The presence of resistance and susceptibility kmers in different data sets. **a** Analyses within the *V. arizonica* sample set. The top-left graph indicates the 99 different resistance (R-kmers) kmers across the x-axis, with their detection frequency across the resistant (CFU/mL < 13) accessions. The top-right graph plots the average detection frequency of susceptibility kmers (S-kmers). The bottom-left and bottom-right graph are similar, they but show R-kmer and S-kmer detection frequencies among susceptible accessions. **b** The same graphs as in **a**, but the top graphs plot R-kmer and S-kmer detection frequencies for the five *V. vinifera* cultivars bred for PD resistance by backcrossing to *V. arizonica*, while the bottom graphs represent susceptible *V. vinifera* cultivars. **c**. Plots of kmer frequencies in six *Vitis* species. The species phylogeny is shown on the left, with the average detection frequency of R-kmers shown in red dot. The gray dots represent average detection frequencies of randomly chosen kmers that had similar population frequencies in *V. arizonica* as the set of R kmers. Whiskers denote 95% confidence intervals.

65.0% of resistant plants, on average, but in only 9.5% of susceptible accessions. We labeled these kmers as resistant (R-kmers). In contrast, 16 kmers were detected in 67.2% and 10.1% of susceptible and resistant accessions, on average, suggesting associations with disease susceptibility (S-kmers). Interestingly, 10 of the 16 S-kmers mapped to a region on chromosome 15 that was ~12 kb upstream of a Jasmonic Acid-Amido Synthetase gene (*JARI*, g252600). In *Arabidopsis*, changes in the expression of *JARI* are associated with a reduction of host defenses⁴². We hypothesize that S-kmers are linked to variants that affect the expression of *JARI* and promote susceptibility to *X. fastidiosa*. An important goal for breeding may be to avoid these S-kmers⁴³.

We then investigated the genomic content of five resistant cultivars (Ambulo Blanc, Caminante Blanc, Camminare noir, Errante noir and Paseante noir) derived from backcrosses to *V. arizonica* (accession b43-17)⁴⁴ to test whether the basis of their resistance lay solely in *PdR1* or included additional genomic regions. After resequencing the five cultivars, we detected all 99 R-kmers in each cultivar but no S-kmers (Fig. 4b, Supplementary Data S8). In contrast, a control dataset from four susceptible *V. vinifera* cultivars (Cabernet Sauvignon cl. 08, Chardonnay cl. 04, Zinfandel cl. 03 and Petite Sirah) contained neither R-kmers nor S-kmers (Supplementary Data S8). Although our analyses used a reference (b40-14) that was not the source of PD resistance in backcrossed cultivars (b43-17), we found 56 kmers mapped to b40-14 hap1, 44 to hap2, and 53 to unplaced contigs. Importantly, the hap1 kmers mapped to both *PdR1* (51 kmers) and to peak 8 on chromosome 15 (5 kmers), suggesting these two regions contribute to resistance. As a complementary method, we scanned SNP heterozygosity in resistant cultivars, reasoning that backcrossed regions should be heterozygous for *V. arizonica*-specific alleles. As expected, this analysis revealed that portions of chromosome 14 were heterozygous across a region that encompassed Peak 5, Peak 6 (*PdR1*) and beyond (Supplementary Fig. S12). However, the proximal peak (peak 4) on chromosome 14 was also heterozygous (spanning from ~6.43Mbp to 6.59 Mbp on chromosome 14; Supplementary Fig. S12). Another prominent peak of heterozygosity on chromosome 9 did not correspond to peaks detected in our GWA. In short, both kmer and SNP analyses suggest that resistance backcrossed from *V. arizonica* encompassed multiple genomic regions.

Resistance markers across *Vitis* species. These observations raise additional questions about the evolution of PD resistance: Did PD resistance arise only once in wild *Vitis* and, if so, is there evidence for the involvement of multiple genic regions? These questions are especially pertinent because all North American wild *Vitis* species can hybridize, with their genomes containing relics of introgression events that are enriched for RLK and RLP genes³⁰. To address these questions, we extended kmer analyses to a population genomic sample of 105 individuals from six wild North American *Vitis*, all of which were assayed for resistance^{30,31} (Fig. 4c, Supplementary Data S8 & Supplementary Fig. S13). The six species were estimated to have a common ancestor ~25 million years ago (mya)³⁰.

We hypothesized that PD resistance was introgressed across species and therefore predicted that the same R-kmers were present across species. We found (as expected) that R-kmers were at significantly higher frequencies in a subset of resistant vs. susceptible individuals for *V. arizonica* (Welch Two Sample WTS t-test, $p = 4.50 \times 10^{-16}$), and also for its sister species, *V. girdiana* Munson ($p = 0.007$) (Supplementary Fig. S14). Interestingly, five of the R-kmers within *V. girdiana* mapped to the chromosome 15 peak, again suggesting a non-*PdR1* component to resistance. These five kmers were detected in ~67% (12/18) of the *V. girdiana* individuals. These data suggest that *V. arizonica* and *V. girdiana* share the basis for resistance, either due to introgression or (more parsimoniously) common ancestry. For the remaining four species, no resistant individuals had > 50% of R-kmers (Fig. 4c), with no difference in R-kmer frequency between resistant and susceptible accessions (Supplementary Fig. S14). In fact, we detected R-kmers less often in these species than for a set of random *V. arizonica* kmers chosen to have similar population genetic frequencies as the R-kmers. Contrary to our hypothesis, the R-kmer distribution in these more distant species provide no evidence that the genetic mechanism of PD resistance (or at least the kmers linked to resistance) was introgressed from *V. arizonica/V. girdiana* to the remaining four species.

Predicting PD resistance. Because our plant accessions were sampled across a geographic range (Fig. 1), we can use the resequencing data to investigate relationships to climate. We

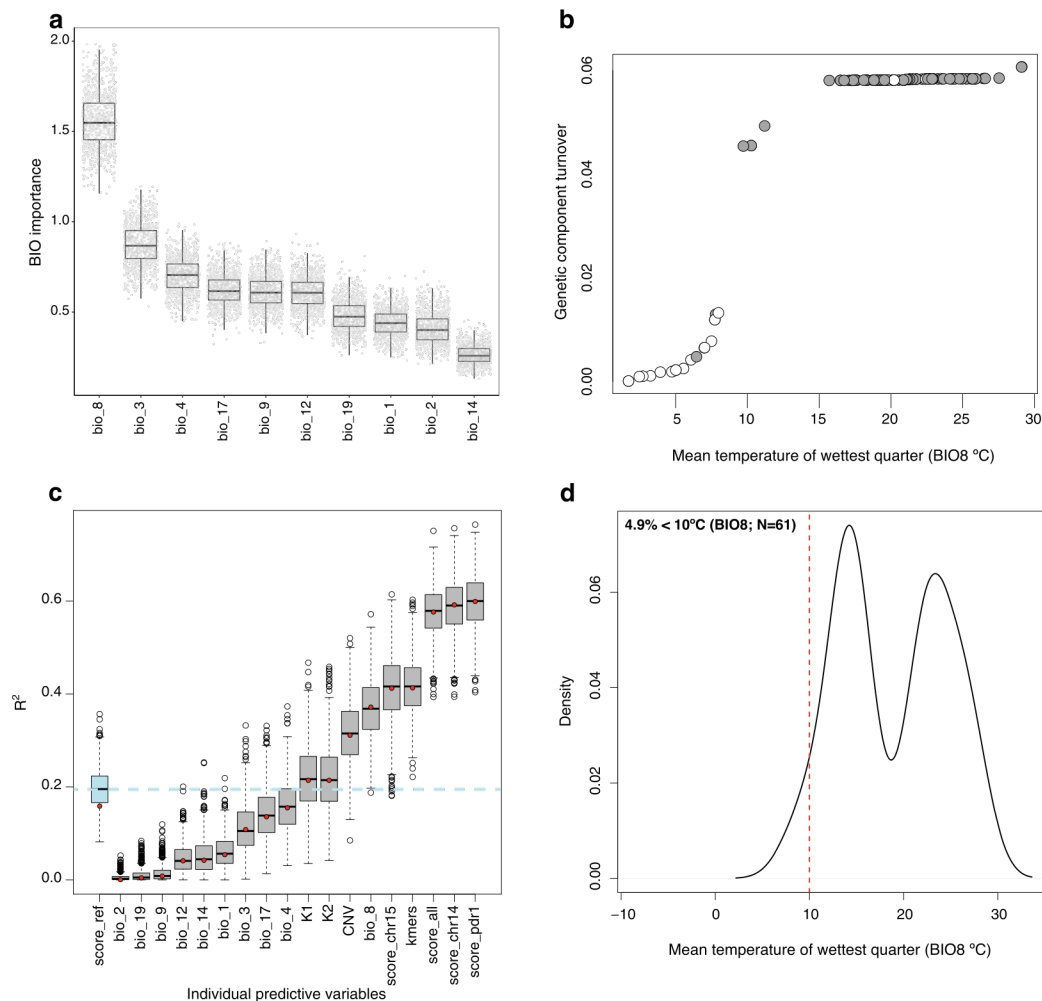


Fig. 5 Relationships among resistance, genetic markers and bioclimatic data. **a** The estimated relative importance, from GF modeling, of each of the bioclimatic variables tested. The y-axis is a measure of the importance of various variables to explain the model - i.e., the relative importance of each bioclimatic variable for predicting changes in allele frequency across the landscape. Each boxplot denotes the average inferred importance of the bioclimatic variable, with the whiskers plotting the standard deviation of 1000 separate analyses (gray dots). BIO8 was estimated to have the biggest impact on the model in all 1000 analyses. **b** The turnover function showing the temperature range of BIO8 on the x-axis and the change in the genetic composition on the y-axis. The circles represent individuals that are colored by resistance (gray) or susceptible (white). **c** Individual predictors in a linear model to predict resistance levels (CFU/ml). The label score_ref represents sets of 1000 randomly chosen sets of 25 SNPs; K1 and K2 are the proportion of the assignments to each admixture group for each individual. The other predictors include bioclimatic variables and genomic data, as listed in the text, each evaluated 1000 times with bootstrapped datasets. Each boxplot reports the second and third quartiles, with median values in the square and circles showing outliers. The barplot whiskers report standard deviation, and the dashed horizontal line reflects the median value of 1000 replicates of the R_{pd} score. **d** The density distribution of BIO8 for a global database of locations of *Xylella fastidiosa* detection.

utilized gradient forest (GF) to detect bioclimatic factors related to resistance. GF is a machine learning method that models the turnover in genetic composition and frequency across the climate landscape⁴⁵ while identifying bioclimatic variables that are important to the construction of the model. As is common for GF applications⁴⁶, we applied it to candidate SNPs, specifically the 25 SNPs associated with resistance. To test for robustness, we also

repeated GF analysis 1000 times. In all 1000 runs, GF identified BIO8 (Mean Temperature of Wettest Quarter) as the most important model contributor among 10 bioclimatic variables, followed by BIO3 (Isothermality), BIO4 (Temperature Seasonality) and BIO17 (Precipitation of Driest Quarter) (Fig. 5a). Moreover, the turnover function revealed a bias in which susceptible individuals were from locations where BIO8 was <10 °C

(Fig. 5b), which was confirmed by a significant pairwise comparison between resistant and susceptible individuals (Supplementary Fig. S15).

We performed two additional analyses to assess the generality between resistance and BIO8. First, we examined our complete dataset of all assayed *Vitis* individuals ($n = 275$) across six species. The dataset was highly skewed, because 30% of susceptible individuals – but only 1.6% of resistant individuals – had $\text{BIO8} < 10^\circ\text{C}$ (FET, $p = 8.2\text{e-}12$). This result held separately for *V. arizonica* (FET, $p = 8.2\text{e-}12$) and when *V. arizonica* was not included in the analysis (FET, $p = 0.03$). Second, we constructed a global dataset of known *X. fastidiosa* geographic locations that integrates across plant families and all *X. fastidiosa* subspecies⁴⁷. Unfortunately, the dataset had few exact locations suitable for analysis, leaving only 61 reputable observations. Of these, fewer than 5% had BIO8 values $< 10^\circ\text{C}$ (Supplementary Fig. S15), reflecting the previously reported relationship between temperature and *X. fastidiosa* presence^{6,48–50}.

Given an association between plant resistance and temperature, we explored whether genetic or climatic factors better predicted bacterial load in *V. arizonica*. We assessed individual predictors with linear models, focusing on 10 bioclimatic predictors and nine genomic predictors (Fig. 5c). The genomic predictors included kmers, CNVs, assignments into genetic groups (K1 and K2), randomly chosen SNPs, and SNPs associated with PD. We summarized SNPs associated with PD with S_{pd} , a measure that ranged from 0.0 to 1.0 and reflected the average proportion of alleles associated with resistance (where 0.0 is no resistance-associated alleles). Focusing on the performance of each of the 19 predictors, S_{pd} calculated across *PdR1* SNPs (10 total SNPs) had the strongest predictive power ($R^2 = 0.599$), followed by related S_{pd} scores based on all candidate SNPs on chromosome 14 (16 SNPs, $R^2 = 0.592$), all candidate SNPs across the genome (25 SNPs, $R^2 = 0.576$) and finally all candidate SNPs on chromosome 15 (6 SNPs, $R^2 = 0.412$) (Fig. 5c). Among the bioclimatic variables, BIO8 had an R^2 of 0.370 in the linear model, which was much higher than the median value for 1000 randomly chosen sets of SNPs ($R^2 = 0.196$) and similar to the predictive power of Kmers ($R^2 = 0.410$) and CNVs ($R^2 = 0.307$). Thus, BIO8 was a reasonable predictor of resistance, even in the absence of genetic data. Notably, the other bioclimatic variables that were identified by GF were not strongly predictive by themselves, e.g., BIO4, BIO17 and BIO3 had lower predictive power than random sets of SNPs (Fig. 5c).

Discussion

X. fastidiosa causes Pierce's Disease in domesticated grapevines (*V. vinifera*) and economically devastating diseases in other crops like citrus, olives, coffee and almonds¹⁶. A diverse body of work has investigated the basis of resistance across diverse crop species but has produced few plausible candidate resistance genes^{15,22,23,25}. To date, however, no studies of *X. fastidiosa* resistance have taken advantage of full-scale genomic approaches like GWA. Indeed, GWA studies in the wild relatives of crops are surprisingly rare, despite the importance of understanding the basis of resistance in ecological settings¹¹ and the transformative potential of such knowledge for crop breeding²⁸. Here we have applied GWA to resistance in *V. arizonica*, based on an improved reference genome, on resequencing data from 167 wild-sampled accessions and on phenotypic data measured from *X. fastidiosa* infection assays performed in the greenhouse^{30,31}. Together, these analyses have yielded information about genomic regions associated with PD resistance and also identified candidate genes within those regions. We have also studied the population frequencies of SNPs associated with resistance, using a machine-

learning framework to help identify the climatic features that correlate with resistance and representing a novel extension of this approach to study interactions between resistance and climate.

Not surprisingly, GWA identified several significant SNP and kmer markers on chromosome 14 between the genetic boundaries that define the *PdR1* locus (Fig. 2). This locus was originally identified by genetic mapping and QTL analyses^{19,51} and it was subsequently backcrossed into susceptible *V. vinifera* to produce resistant cultivars²¹. There had been no insights into causative genes that lie within this region until Aguero et al.²² transformed two R genes separately into *V. vinifera*. They failed to find evidence for enhanced resistance using over-expression assays, but our study of a different accession (b40-14) provides further insights. First, based on careful genomic annotation combined with association and gene expression analyses, we have identified a narrower *PdR1* region with two strong candidates (g238150 and g238180), along with two additional candidates (g238250 and g238290) within the traditional *PdR1* locus (Fig. 3). Second, we have mapped the two candidate genes tested in Aguero et al. (RGA14 and RGA18) from b43-17²² to our genome. The closest homologues in the *PdR1* region were genes g238170, which encodes an putative LRR protein with 99.15% amino acid identity to RGA14, and g238120, which encodes a putative LRR-RLP with 92.46% identity to RGA18. Both g238120 and g238170 are located within the *PdR1* locus, but they are either not expressed or lowly expressed in leaves and stems across our sample of three PD-resistant full-sibs (Supplementary Data S9). Our results thus suggest that these genes, like their homologues in b43-17, are unlikely to be involved in PD resistance.

Overall, our candidate genes, which are present on both b40-14 haplotypes, differ substantially from those identified in b43-17. It is possible, of course, that different genes confer resistance in different accessions, given that structural variants are common in *Vitis* genomes⁵² and that allelic heterogeneity for resistance is also common⁵³. A reasonable way to test our candidates would be to generate knock-outs in *V. arizonica*, but unfortunately transformation in *Vitis* is currently efficient only for a narrow set of *V. vinifera* cultivars⁵⁴. Hence, an important challenge for viticulture is to improve transformation techniques for application to more cultivars and to agronomically valuable wild species. A future goal is to transform *V. vinifera* with sets of our candidate genes both as single loci and in combination, because disease resistance genes often act in tandem and also because stacking resistant genes is often desirable.

It is possible that candidate genes from *PdR1* will not be sufficient to confer complete resistance because contributions from additional genomic regions are required. There is evidence consistent with and against this multigenic hypothesis. For example, in an early study of PD resistance among *Vitis* species, Mortensen (1968) performed controlled hybrid crosses and concluded that complementary gene action among three independent genes best explained his results⁵⁵. Similarly, although multiple QTL studies have identified *PdR1* as a major effect locus^{19,20,56,57}, they also hint that other regions may be important. For example, a recent study used interval mapping and found that *PdR1* was the only major QTL, explaining 55.5% of phenotypic variation⁵⁷, but phenotypic associations were also detected to regions of chromosomes 3, 7, 9, 10 and 15. Another study investigated the basis of resistance across 17 *V. arizonica* accessions and identified three accessions for which the *PdR1* region explained $< 10\%$ of phenotypic variation for resistance, suggesting that the possibility that resistance arises from non-*PdR1* locations in some accessions⁵⁶.

The value of GWA is that, unlike QTL studies, it integrates across an extensive population sample. Here we have done so in

the context of a platinum-level reference genome, which likely provides a more complete view of the basis for resistance. Our results are tentatively consistent with the idea that multiple genomic regions contribute to resistance. First, SNPs associated with resistance are found across the genome. We studied these significant SNPs further by measuring effect sizes - i.e., by estimating the slope of the phenotypic response to genotypic categories (Supplementary Fig. S16). We caution that this approach is subject to numerous caveats, including the Beavis effect and a lack of independence among SNPs, and thus must be viewed cautiously; ultimately effect estimation requires a second independent dataset⁵⁸. Nonetheless, we find that the SNPs within peak 4 produce ~1.2-fold higher slopes, on average, than those within the *PdR1* peak (peak 5), which have similar (1.02x) slopes, to the SNPs in peak 6 on chromosome 15. Moreover, the SNPs with the largest inferred effect are not in *PdR1*, although *PdR1* has the most associated SNPs. Thus, this simple measure of effect sizes provides little evidence that SNPs in *PdR1* explain resistance phenotypes more completely than significant SNPs from other peaks and regions. To further explore this multigenic theme, we have also applied a multi-locus mixed model, which can more accurately control for false positives than some other methods (see Methods). Although the associations to *PdR1* remained, the highest proportion of significant SNPs were in other regions of the genomes, including significant hits on chromosomes 2, 4, and 11 (Supplementary Data S10 and Supplementary Fig. S17). We note, however, that some of our peaks are not recapitulated with this method.

Second, the striking LD pattern on chromosome 14 suggests the possibility of additive or epistatic interactions between distinct regions on the chromosome (Fig. 3a). Another explanation for this pattern could be reduced recombination on chromosome 14. However, there is no evidence for reduced recombination from genetic maps and QTL mapping in F1 populations^{19,20}, which discriminate the *PdR1* region from the proximal end of the chromosome where peak 4 is located. The argument that this pattern arises from recombination is further discounted by the fact that the rest of the chromosome, including sequences intervening peaks 4 and 5, are not in strong LD (Fig. 3a). Ultimately, we cannot discriminate whether this unique pattern is caused by sampling phenomenon or even whether this pattern of LD contributes to our detection of peak 4. We speculate, however, that the pattern of LD is likely best explained by selection for resistance that acts on alleles in both regions. Clearly this speculation requires further investigation.

A third observation is based on the distribution of R-kmers across species. We found, for example, that *V. girdiana* and *V. arizonica* share R-kmers that map to at least two different chromosomes (14 and 15), suggesting that not only that multiple genomic regions contribute to resistance but that these regions may have contributed to resistance since the divergence of these two species ~ 23 million years ago³⁰ (Fig. 4b). Finally, we have also investigated the genomics of five resistant *V. vinifera* cultivars, representing the only grapevine cultivars bred for PD resistance to date. Because these cultivars were produced by backcrossing, we hypothesized that genomic regions contributing to resistance are highly heterozygous. The *PdR1* region was highly heterozygous, as expected, but so were additional peaks on chromosomes 8, 9, 14 and 15 (Fig. 4a). These patterns suggest that additional, non-*PdR1* regions may have been instrumental for backcrossing resistance into the *V. vinifera* background and continue to hint that non-*PdR1* regions contribute to the resistance phenotype.

Inferring the spatial distribution of disease resistance is critical for understanding its evolution and ecology⁵³. We have investigated resistance across the landscape of *V. arizonica* (Fig. 1) and

across six wild species that segregate for PD resistance. Given that all North American *Vitis* species are interfertile and that there is ample genomic evidence for historical introgression of resistance genes among species³⁰, we predicted that the genetic solution to PD would yield clear signals of introgression. As mentioned above, we find that R-kmers are commonly shared between *V. arizonica* and its closest species in our sample, *V. girdiana*, but not across the other species (Fig. 4c). It is possible that we cannot detect introgression events because causative loci and associated R-kmers have become uncoupled over evolutionary time. At a minimum, however, our results provide no evidence that PD resistance has introgressed from *V. arizonica* to the non-*girdiana* species in our sample, as speculated previously²². If true, this implies that other wild *Vitis* species have independently evolved mechanisms of PD resistance, such that further study of these species may provide additional insights into alternative genetic mechanisms of PD resistance. In this vein, one particularly interesting species is *V. mustangensis* (syn. *V. candidans*), which has not been widely utilized agronomically as a rootstock or for hybrid scion breeding, but it does segregate for PD resistance and also contains alleles that may be useful for viticulture in the context of climate change⁵⁹.

We have taken advantage of our geographic sampling to investigate correlates between resistance and climate, finding that resistance-associated SNP markers correlate with climatic variables, especially temperature in the wettest quarter (BIO8). More specifically, susceptible plants tend to be found below specific BIO8 thresholds. Based on the turnover function, which is inferred from both climatic and population genetic data, the BIO8 threshold lies between 8 °C and 10 °C (Fig. 5b). Focusing specifically on the higher of these two thresholds, we have found that the 10 °C threshold not only applies within *V. arizonica*, but also to our expanded sample of *Vitis* species and to a wider (although still quite limited) geographic sample of *X. fastidiosa* that summarizes across plant hosts and bacterial subspecies⁴⁷ (Fig. 5d). Somewhat remarkably, this simple climatic measure predicts bacterial load nearly as well as some genetic markers (e.g., kmers) and better than others (e.g., CNVs) (Fig. 5c). This is, to our knowledge, the first time that genomic data have been used to associate plant resistance with climate, yielding a useful bioclimatic predictor. Our findings are not without precedent, however, because temperature has previously been identified as a strong predictor of *X. fastidiosa* distribution and presence^{6,48,49,60}. Combining these observations, we suspect that individuals with low (<10 °C) BIO8 temperatures lack resistance because *X. fastidiosa* growth is hampered at low temperatures^{50,61} and/or because temperature affects its insect vectors^{17,62}. Plant resistance will not be favored by natural selection in regions where the pathogen does not persist, particularly if there is a fitness cost to resistance, as has been demonstrated for R-gene mediated resistance in *A. thaliana*⁶³.

Previous work has modeled the distribution of *X. fastidiosa* under climate change^{6,48,49,60,64}, but these models have not been informed by data on the distribution of plant resistance. To illustrate how such information may be useful, we have employed climate models to predict where BIO8 will shift in the future. More specifically, we have identified regions where BIO8 will transition across the thresholds of 10 °C (Fig. 6a), as informed by our analysis of genetic variants associated with plant resistance. By categorizing regions where BIO8 is predicted to move from below (or above) 10 °C in the present to above (or below) 10 °C in the future, we can identify regions where *X. fastidiosa* pressure is likely to expand or contract. We performed these categorizations across 54 climate models to consider uncertainty in global circulation models, shifts in greenhouse gas emissions and time (see Methods). Summarizing across models, we have found that most

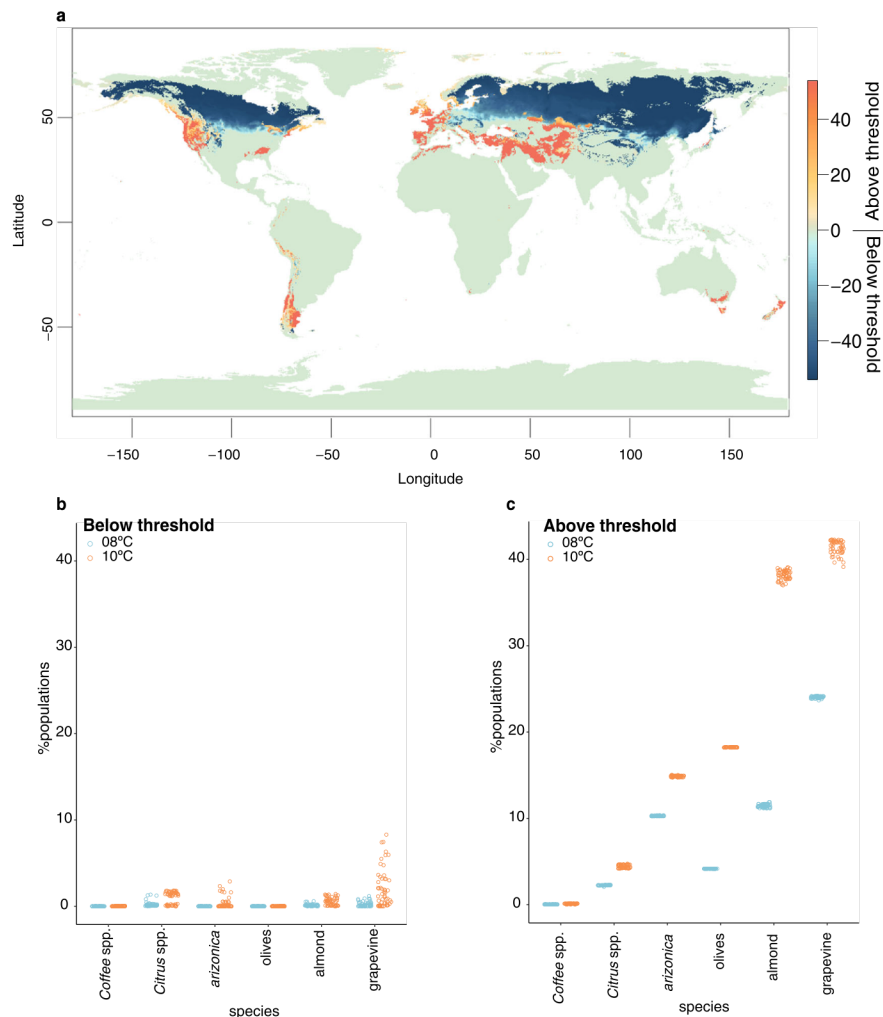


Fig. 6 Climate predictions and projections of the prevalence of *X. fastidiosa* for focal crops. **a** The map portrays the number of climate models (out of 54 total) that support movement across the BIO8 = 10 °C threshold. The warmer colors reflect regions that are moving from below (in the present) to above the threshold, while the cooler colors portray areas that are moving from above (in the present) to below the threshold. The intensity of color in the scale bar reflects the number of 54 climate models that agree with the threshold transition. The generation of this map, as well as the maps in Figures S1 and S18, relied on publicly available information from WorldClim2 (<https://www.worldclim.org/>) and CMIP6 (<https://pcmdi.llnl.gov/CMIP6/>). **b** A summary of the percentage of locations associated with movement from above the 8 °C or 10 °C threshold (in the present) to below that threshold for five crops and for *V. arizonica*. **c** A summary of the percentage of locations associated with movement from above the threshold (in the present) to below the threshold for five crops and *V. arizonica*. Both b and c are based on 6204 locations for coffee; 3386 locations for almonds, 1111 locations for *V. arizonica*; 5256 locations for *Citrus*; 174,713 locations for olives and 33,225 locations for grapevines. In both b and c, each dot represents an estimate based on one of the 54 climate models.

of the globe will not transition - i.e., it will remain either above or below the 10 °C (Fig. 6a) during the wettest quarters. Perhaps unsurprisingly⁶⁵, large portions of Canada, Eastern Europe, Russia and Northern Asia climate are predicted to move beneath the 10 °C BIO8 threshold, suggesting that these regions may be less likely to have pathogen pressure in the future.

There are, however, discrete areas of the Western Americas, Western Europe, Central Asia, Southern Australia and elsewhere

that are predicted to transition above the 10 °C (Fig. 6a) threshold in most models, suggesting increasing *X. fastidiosa* pressure in these regions. These models of course make numerous assumptions. Some are common to all climate models (e.g., a reliance on accuracy of the climate predictions) and to previous *X. fastidiosa* species distribution models (which assume *X. fastidiosa* is not dispersal limited and often ignore the potential of *X. fastidiosa* to evolve to new temperature regimes^{46,64}). Some are more specific

to this work, specifically that the BIO8 threshold has application to plant species beyond *V. arizonica*, but this assumption is supported by our analysis of data that includes different plant hosts (Fig. 5c). Finally, we note that there is some uncertainty in the appropriate BIO8 threshold, because the turnover function has inflection points along the range of 8 °C to 10 °C (Fig. 5b). Hence, we repeated these analyses with an 8 °C threshold, finding similar results (Supplementary Fig. S18). Importantly, our climate modeling illustrates how data about plant resistance can help inform the potential distribution of disease under a shifting climate.

In fact, we can proceed one step further by assessing the potential effects of climate on specific plant taxa: wild *V. arizonica* and five affected crops (grapes, coffee, almonds, citrus and olives). To do so, we first downloaded data on global locations where each species is grown, and we then used climate projections to estimate the proportion of locations that will transition over either the 8 °C threshold or the 10 °C BIO8 threshold under climate models (see Methods). Using this approach, we predict that few locations for these crops will transition below the 8 °C and the 10 °C thresholds. In citrus and grapes, for example, the average estimate across the 54 climate models is that only between 0.21% and 0.97% (of 7853) locations and between 0.24% and 2.10% (of 40,075) locations will transition below 8 °C or 10 °C, respectively, under climate projections (Supplementary Fig. S18 & Fig. 6b). Perhaps not unexpectedly, a much higher proportion of locations are estimated to exceed the 8 °C and 10 °C threshold over time. For example, we estimate that between 11% and 38% of regions of almond cultivation will transition above the 8 °C and 10 °C threshold, respectively (Fig. 6c). The corresponding numbers for grapevine cultivation are 24% and 41%, suggesting that roughly a quarter of viticulture will be subjected to additional *X. fastidiosa* pressure even under the less expansive value based on an 8 °C threshold. Similarly, 4% and 18% of olive growing locations are expected to transition above the 8 °C and 10 °C threshold, respectively, corroborating previous modeling work that predicts further spread of the bacterium amongst olive growing regions even under current climates⁷. We note that, like some other climate studies, our calculations treat *X. fastidiosa* subspecies similarly, even though there are differences in their host specificity^{6,66} and distributions^{7,67}. However, *X. fastidiosa* can move between crops, as has been shown, for example, between almonds and grapevines⁶⁸; hence we believe that a multi-crop perspective is critical.

Overall, our study has considered the genomic, climate and evolutionary context of resistance to a pathogen that is an emerging global pest and that already causes devastating economic damage to several crops. By studying the complex genetic architecture of PD resistance in a wild grapevine, we have implicated several genomic regions in resistance and identified genes that are fitting candidates for genetic introduction into susceptible crops. Furthermore, by comparing features across breeding lines and wild *Vitis* species we have uncovered hints about the origins of this critical trait. Finally, our work has highlighted the potential of combining landscape-scale resequencing data and climate models to predict the shifting pressures of a damaging plant pathogen. These results underscore the urgent need to identify additional *X. fastidiosa* management and containment methods, potentially via enhanced information about resistance mechanisms and genes.

Methods

Plant material and PD disease evaluations. We studied 167 accessions of *V. arizonica* collected across the southwestern states of the US and northern Mexico, covering the distribution of the species (Fig. 1 and Supplementary Fig. S1). The sampling location was available for all 167 *V. arizonica* accessions, which are part

of a living collection of Southwest *Vitis* accessions maintained at the University of California, Davis; all *Vitis* accessions used in this paper were from that collection. PD resistance in these 167 accessions were previously assayed in controlled greenhouse trials^{30,31}, following previously published protocols^{31,69}. Briefly, accessions were inoculated with *Xylella fastidiosa* subsp. *fastidiosa* (Wells) using a randomized complete block design that included at least two water-inoculated plants and susceptible *V. vinifera* cultivar Chardonnay plants as controls. Nineteen screens were carried out from 2011 to 2020, with a minimum of four biological replicates per accession. Disease severity was assessed 10 to 14 weeks after inoculation using categorical phenotypes for disease severity. Whole immunoglobulin (IgG) antiserum raised against *X. fastidiosa* in a rabbit were used to quantify the bacteria levels from stem tissue above the inoculation site from 12 to 14 weeks, using a double-antibody sandwich ELISA method as described previously⁶⁹. In this assay, absorbance values were converted to colonizing forming units per milliliter (CFU/ml) concentrations using a standard calibration curve. The ELISA data were log-transformed and statistical analysis was performed using JMP Pro14 software (Copyright 2020, SAS Institute Inc.) to determine the variability of ELISA for the reference control plants across experiments. Following precedent^{30,31}, we used the least squared means of CFU/ml across biological replicates as an indicator of disease resistance. Lower bacterial load reflects higher PD resistance.

Genomic reference, resequencing data and SNP calling. Version 1 of the *V. arizonica* sequence (b40-14 v1) was published previously^{30,70}, but this study relied on an updated version (b40-14 v2). Version 2 was a complete re-assembly based on the application of HaploSync⁷¹ combined with ~2000 rhAmpSeq *Vitis* markers⁷², independent of any single reference genome assembly. This new version of *V. arizonica* genome chromosome haplotypes were better phased, with fewer unplaced sequences (~65 Mb; 3035 gene loci). Correct phasing was assessed during the assembly quality control steps based on genome and locus-specific markers and gene content. The gene annotations were reported between versions 1 and 2, so that genes may have moved locations in the new assembly but no new annotations were performed. The genome contains 57,003 gene loci. This version 2 assembly represents the most contiguous genome assembly of any wild *Vitis* genome released to date and was used as the reference for all analyses. The genome is available for browsing at grapegenomics.com, and the long-read sequencing data are available from NCBI BioProject ID PRJNA593045.

Our *V. arizonica* resequencing dataset consisted of short-read, whole genome data from 167 *V. arizonica* individuals, for which a subset of $n = 20$ had been sequenced previously³⁰. The data for this subset was available from NCBI BioProject PRJNA731597. For the remaining 147 individuals, genomic DNA was extracted from leaf samples with the Qiagen DNeasy plant kit and assessed for quality and degradation on agarose gels and a BioAnalyzer (Agilent). Sequencing libraries were constructed with an insert size of ~300 bp using Illumina library preparation kits; the libraries were assessed for quality and quantified using a Kapa library quantification kit; libraries were then sequenced using the Illumina HiSeq 2500 platform with 2×150 bp paired reads to a target coverage of 10x. The raw sequencing data for this study were deposited in the Short Read Archive at NCBI under BioProject ID: PRJNA842753.

We filtered and evaluated raw reads from 167 individuals using Trimmomatic^{36,73} and FastQC⁷⁴. Reads were scanned in windows of four base pairs, cutting when the average quality per base dropped below 20. We removed leading and trailing bases of reads that had a quality below 3 and finally kept reads with sizes of 60 bp or higher after trimming. Filtered reads were then mapped to the reference genome (independently to Hap1 and Hap2) with the BWA-MEM algorithm⁷⁵ implemented in bwa-0.78. Joint SNP calling was conducted using the GATK v4.2.2.0 pipeline⁷⁶ for Hap1 and Hap2 independently. We first used the integrated version of Picard tools⁷⁷ to remove duplicated reads with the “MarkDuplicates” function, followed by the “AddOrReplaceReadGroups” function to label the reads for each individual. For the SNP prediction we used the HaplotypeCaller algorithm with a sample ploidy of 2 and a mapping base quality score threshold of 20 ($Q > 20$). We combined the VCF files of all individuals to make the final SNP calls using the “GenotypeGVCFs” function with default parameters. We then filtered raw SNPs with bcftools v1.9⁷⁸ (<https://samtools.github.io/bcftools/>) and vcftools v0.1.17 (<https://vcftools.github.io/>)⁷⁹. We kept SNP sites for downstream analysis if they were biallelic, had quality higher than 30, had a depth of coverage higher than five reads, had no more than three times the median coverage depth across accessions, and had no missing data among individuals. Additionally, the following expression was applied under the exclusion argument of the filter function in bcftools: “QD < 2.0 | FS > 60.0 | MQ < 40.0 | MQRankSum < -12.5 | ReadPosRankSum < -8.0 | SOR > 3.0”.

Population Structure and genome-wide associations. We transformed the VCF file into BEAGLE format using vcftools 0.1.17. We used BEAGLE files as input to evaluate the genetic structure of the *V. arizonica* using the NGSAdmix software included in the ANGSD package version 0.931-21-g13af014⁸⁰. We ran NGSAdmix for 1 to 10 ancestral populations (K), repeating analyses 10 times for each K value and including variants with a minimum minor allele frequency > 0.05 (Supplementary Fig. S19). We then employed the Cluster Markov Package Across K (ANGSD package version 0.931-21-g13af014 software⁸¹) to detect the K value with the highest likelihood of $K = 2$.

The NGSadmix results were used to help guide controlling for genetic structure in genome-wide association (GWA). We performed GWA to identify significant associations between SNP allelic frequencies and bacterial load (*Xylella fastidiosa* CFU's), using two different methods that control for population structure using different approximations. First, we used LFMM2, which uses latent factors to control for genetic structure⁸². The number of latent factors were chosen based on a visual observation of the screenplot of the percentage of variance explained by the loadings of the genetic PCA of all individuals. The PCA was obtained using all loci with no missing data and with the `prcomp` function in R⁸³. In total, we defined $K = 5$ latent factors (Supplementary Fig. S20). Next, we ran LFMM2 using a ridge penalty (function `lfmm_ridge`), and we controlled for genomic inflation factor (function `lfmm_test`, with `calibrate = "gifi"`). We corroborated that the p values had a flat distribution, and we corroborated that the genetic structure was well controlled based on a qqplot (Supplementary Figs. S4 & S7).

In addition, we used the variance component algorithm Efficient Mixed-Model Association eXpedited (EMMAX) version beta-07Mar2010⁸⁴. We converted the set of filtered SNPs with no missing data from VCF format to transposed ped format using PLINK version 1.90b6.16⁸⁵. Using the transposed ped files as input, we calculated the Balding-Nichols (BN) kinship matrix using the "emmax-kin" script and default parameters. Finally, we ran the associations using the SNPs data (as transposed ped), the BN matrix, and the phenotype as Least Squares Means of CFU/ml of *X. fastidiosa* for each accession. Both methods (EMMAX and LFMM2) were adjusted for multiple comparisons using the Bonferroni correction of the function "p.adjust" from the stats package version 4.1.2 in R. We focused only on SNPs and candidate regions that were detected by both methods.

Finally, we also used the R package mrMLM v 5.0.1⁸⁶ to perform a multi-locus genome-wide association. We used the FASTmrEMMA algorithm⁸⁷ within the mrMLM package with the same set of SNPs, bacterial load measurements and the kinship matrix described above to run multi-locus associations.

Kmer-based GWA. To perform GWA based on kmers, we followed a previously published pipeline³⁵ (<https://github.com/voicheck/kmersGWAS>). Briefly, we extracted all kmers and canonized (i.e. reverse complement is assumed to be the same kmer) kmers of 31 bp in size using KMC version 3⁸⁸. We extracted the kmers directly from the paired and unpaired filtered reads for each of the 167 *V. arizonica* samples independently. We compared the kmers across samples and created a table of kmers that were found in at least 5 individuals ("mac 5") and in each canonized/non-canonized form in at least 20% of individuals from which it appeared in ("p 0.2"). We used the script "emma_kinship_kmers" included in the pipeline with a MAF < 0.05 filter to create a kinship matrix based on the kmer table. Finally, we ran the kmer-GWAS with the script "kmers_gwas.py" and GEMMA version 0.98.5⁸⁹ with the kinship matrix, the kmer table, and the phenotype data as Least Squares Means of CFU/ml of *X. fastidiosa* for each accession. The script provided a list of 9991 kmers that passed a parametric test as an initial filter. To filter more stringently, we used the number provided in the file "pheno.tested_kmers", which was 967066440, to adjust the p -values with a Bonferroni correction using the program "p.adjust" from the stats package version 4.1.2 in R. To create a textual version of the presence/absence kmers of the significant Kmers we used the "filter_kmers" from the pipeline. Given 115 significant kmers, we mapped them to the *V. arizonica* genome using BLASTN⁹⁰ and a word size of 8 bp. We filtered all the BLASTN results and kept alignments, allowing a maximum of 1 mismatch.

We further explored the sequences of the 36 kmers that did not map to our reference genome. We first searched and extracted the reads matching these 36 kmers across the 167 *V. arizonica* accessions. We then used SPADES v3.15.4⁹¹ to assemble the matching reads into contigs independently for each kmer. Finally, we aligned the resulting contigs to the Reference RNA Sequences database (Refseq_rna) using BLASTN (visited on 06/10/2022) and recorded the top hit gene.

Copy Number variation analysis and associations. To identify Copy Number Variants (CNVs), we used the program CNVcaller version 2.0 (<https://github.com/JiangYuLab/CNVcaller>)⁹². CNVcaller uses normalized read-depth values across windows in the genome to identify CNVs and it is especially suited for large population data like our *V. arizonica* sample. First, we generated a duplicated window record file specifically from our genome reference *V. arizonica* b40-14 v2 using a window size of 2 kb and for each chromosome independently. We then analyzed the individual read depth in 2 kb non-overlapping windows using the "Individual.Process.sh" script and the alignment files of all 167 accessions in bam format with the PCR duplicated reads removed during the SNP calling pipeline (see section above). The script produces normalized values of read depth across the genome for each genotype. We then used the normalized read depth values of all genotypes as input to the script "CNV.Discovery.sh", excluding windows with a lower frequency of gain/loss individuals of 0.1 ("f 0.1") and with Pearson's correlation coefficient lower than 0.3 ("r 0.3"). Finally, we used "Genotype.py" to classify genotypes across the population according to their CNV profiles.

To explore the associations between CNVs and PD-resistance, we used the CNVcaller estimation of diploid copy number for each CNV and tested for correlations with *X. fastidiosa* bacterial levels, while taking into account the genetic structure of the *V. arizonica* population. We used the function "pcor.test" from the R library "pcor" v1.1⁹³ to run a partial correlation for each of the CNVs, using genotype assignment ($\hat{Q}$) values from the genetic structure analysis (see section

above) as the confounding variable. To account for multiple testing we imposed a Bonferroni correction and identified significant CNVs with adjusted $p < 0.05$.

Defining PD-associated peaks. We performed a total of four association analyses: LFMM2 based on SNPs, EMMAX based on SNPs, kmer-based GWA, and CNV-based GWA. From these analyses we defined eight GWA peaks of interest in the genome (Fig. 2). To define these peaks, we required that a peak contains at least one SNP that was significant with both LFMM2 and EMMAX. However, most peaks had multiple pieces of evidence - i.e., either more than one SNPs, significant kmers and/or CNV variants. When applying this logic, we focused only on kmers that mapped uniquely to the genome and so excluded 17 kmers that mapped to multiple places in the genome with the same identity and alignment length.

Analyses of PD-associated kmers in other *Vitis* species. We identified kmers associated with resistance in *V. arizonica* sample and then characterized their presence in three different resequencing datasets: i) a multiple species *Vitis* dataset, ii) a dataset generated from scion cultivars bred for PD resistance by backcrossing to the b43-17 accession of *V. arizonica*, and iii) a set of PD susceptible cultivars. The first dataset included 105 accessions from five species: *V. arizonica* ($n = 22$), *V. candicans* ($n = 24$), *V. berlandieri* ($n = 22$), *V. girdiana* ($n = 18$) and *V. riparia* ($n = 19$)³⁰. All of these accessions had been assayed for PD resistance, and categorized as resistant if CFU/mls were <13.0 at the time of assay. In this dataset, 20 *V. arizonica* were resistant ($n = 20$), 21 *V. candicans* ($n = 25$), 3 *V. berlandieri* ($n = 21$), 9 *V. girdiana* ($n = 17$), 2 *V. monticola* ($n = 5$) and 2 *V. riparia* ($n = 20$). The resistance assay data and sampling locations of the accessions are available³⁰. For the second dataset, we generated resequencing data for five PD resistant cultivars (Ambulo Blanc, Caminante Blanc, Camminare noir, Errante noir and Peseante noir). DNA extraction, library preparation and Illumina sequencing followed the protocols mentioned above, and the data were deposited into Bioproject: PRJNA842753. Finally, the "control" dataset of PD susceptible accessions was downloaded from public databases (Cabernet Sauvignon cl. 08: SRR3346862; Chardonnay cl. 04: SRR5627799; Zinfandel cl. 03: SRR8727823; and Petite Sirah: SRR12328988). For each of the datasets, we generated kmers of 31 bp for each sample as described above. We then searched for the presence of 115 associated kmers from *V. arizonica* using the "filter_kmers" script from the kmer-GWAS pipeline³⁵ (<https://github.com/voicheck/kmersGWAS>).

To compare the R-kmers with random sequences in other *Vitis* species, we first extracted 100,000 unique and random kmers from the *V. arizonica* population. We then calculated the frequency of these sequences across all individuals and selected kmers with similar frequencies as the R-kmers mean (0.52), resulting in 38,523 kmers. We then created 100 subsets of 99 random kmers from the set with similar frequencies as R-kmers. Finally, we searched for the presence of each set of other *Vitis* species and calculated the mean frequencies of the 99 kmers in each random set. We report the average and standard error of means across the 100 sets in Fig. 4c.

Linkage disequilibrium. We calculated the genome-wide LD decay across the *V. arizonica* population with the software PopLDdecay v3.40⁹⁴. We used the filtered SNPs of the 167 individuals from hap1, allowing a maximum distance of 1 Mb (Supplementary Fig. S11). We used the perl script "Plot_OnePop.pl" included in the package to create the decay graph.

To explore the LD landscape of the regions around PdR1 and chromosome 14 as a whole we used Tomahawk v0.7.0 (<https://github.com/mklarqvist/tomahawk>). We used as input the filtered SNPs of chromosome 14 for hap1 as input, containing the 167 *V. arizonica* accessions. We converted the VCF file into a custom format file ("twk") for the package and calculated the LD with the "calc" function. We then filtered LD values using the "view" function, keeping regions with $R^2 > 0.5$ and p values < 0.001. Given that we were interested in the LD at the chromosome-scale (~30 Mb) we used the "aggregate" function. Using this function we aggregated R^2 values in 1000 bins for both the x and y-axis, and used 5 as the minimum cut-off value in the reduction function. Finally, we used the "tomahawk" R package (<https://github.com/mklarqvist/tomahawk>) to create the chromosome-scale LD landscape plot using the aggregated R^2 values.

Functional annotation and refinement of PdR1 gene models. Gene models located within the two haplotypes of PdR1 were manually refined by visualizing alignments of RNA-seq reads from *V. arizonica* b40-14 leaves³⁰ using the Integrative Genomics Viewer (IGV) v2.4.14⁹⁵. RNA-seq reads were aligned onto the diploid genome of *V. arizonica* b40-14 using HISAT2 v2.1.0⁹⁶ and the following settings: --end-to-end --sensitive -k 50.

Predicted proteins of PdR1 genes were scanned with hmmssearch from HMMER v3.3.1 (<http://hmmer.org>) and the Pfam-A Hidden Markov Models (HMM) database⁹⁷ (downloaded on January 29th, 2021). Protein domains with an independent E-value less than 1.0 and an alignment covering at least 50% of the HMM were selected. Transmembrane helices were predicted with TMHMM2 v2.0c⁹⁸. Proteins with a predicted Leucine Rich Repeat (LRR) domain and a transmembrane helix were classified as LRR receptor-like proteins. Proteins having a predicted LRR or lysin motif (LysM), a kinase domain, and transmembrane helices were categorized as receptor-like kinases.

Predicted proteins of *PdR1* genes were aligned onto the predicted proteome of *A. thaliana* and the grape PN40024 (V1 annotation) using BLASTP v.2.2.28 +⁹⁹. Alignments with an identity greater than 30% and a reciprocal target:query coverage between 75% and 125% were kept. For each *V. arizonica* protein, best hit in the *A. thaliana* and PN40024 proteomes was determined using the highest product of identity, query coverage, and reference coverage. The sequences of the two ORFs (V.ari-RGA14 and V.ari-RGA18)²² were aligned onto the b40-14 genome using blastn v. 2.2.28 +⁹⁹.

Gene expression analyses. To evaluating the transcript abundance of candidate genes, plants from three PD-resistant genotypes and three PD-susceptible genotypes of the 07744 population ([*V. rupestris* Wichita refuge x *V. arizonica* b40-14] x *V. vinifera* Airen) were propagated in a controlled environment and inoculated with either *Xylella fastidiosa* or water. Pieces of green stem at 10, 20, 30 and 40 cm above the inoculation were collected from each plant at 1, 2, 3, and 4 weeks post-inoculation. Pieces of the green stem from each genotype were pooled together. Each genotype constitutes a biological replicate. All plant material was immediately frozen in liquid nitrogen after collection and ground into powder. Total RNA were extracted as described¹⁰⁰, using the Spectrum Plant Total RNA kit (Sigma-Aldrich). RNA quality and quantity were determined using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific) and a Bioanalyzer Chip RNA 7500 series II (Agilent Technologies). cDNA libraries were prepared using the Illumina TruSeq RNA sample preparation kit v.2 (Illumina, CA, USA) and sequenced in single-end 100-bp reads on an Illumina HiSeq4000. RNA-seq reads were parsed using Trimmomatic v.0.37⁷³ with the following settings: LEADING:7 TRAILING:7 SLIDINGWINDOW:10:20 MINLEN:36. Transcript abundance was evaluated with Salmon v.1.5.1¹⁰¹ with the parameters: --gcBias --seqBias --validateMappings. A transcriptome index file was built using a k-mer size of 31 and the combined transcriptsomes of *V. arizonica* b40-14 (v2.1), *V. vinifera* cv. Cabernet Sauvignon⁷⁰, with their genomes as decoy. Quantification files were imported using the R package tximport v.1.20.0¹⁰². The RNAseq data were deposited in NCBI Bioproject PRJNA956994.

Bioclimatic variables associated with resistance. To identify the association between SNPs associated with resistance and the environmental landscape, we applied gradient forest (GF)¹⁰³, which models the turnover in genetic composition across the landscape⁴⁵ and identifies both the bioclimatic variables that contribute to the construction of the model and the 'turnover function' - i.e., the change of genetic composition across the landscape^{45,104,105}. To estimate the GF model, we used the *gradient forest* package in R, using the 25 SNPs identified by LFMM2 and EMMAX as response variables and using bioclimatic variables as predictive variables. The 19 bioclimatic variables were filtered to retain any correlations < 0.80, based on a variance inflation factor calculated by corSelect from the R package *fuzzySim*¹⁰⁶. After filtering, we retained 10 of 19 bioclimatic variables (BIO1, BIO2, BIO3, BIO4, BIO8, BIO9, BIO12, BIO14, BIO17, BIO19). We performed the GF analysis using SNP frequencies from each individual (i.e., 0, 1 and 0.5 for heterozygotes)⁵⁹ and repeated the analysis 1000 times. We also plotted the turnover functions for each bioclimatic variable to show how populations with different resistance to PD are distributed across allele frequency change (Supplementary Fig. S15).

Predicting PD resistance. To determine whether genetic and environmental variables predicted PD resistance, we first ran a linear model individually for each bioclimatic and genetic variable to determine their individual predictive power, using the *lm* function in R. We performed 1000 bootstrap replicates for each model to estimate the variance in predictive ability. For the bioclimatic variables, we estimated the individual linear models between the 10 bioclimatic values from where the *arizonica* individuals were sampled and their bacterial load. For some genetic variables, we first estimated a polygenic score across SNPs that associated with bacterial load. This polygenic score, which we called the PD score (S_{pd}), was calculated as the proportion of alleles that contribute to resistance to PD in a given individual. The state of these contributing alleles was inferred from the GWA results. S_{pd} was designed after the population adaptive index, which measures the proportion of alleles that show patterns of local adaptation¹⁰⁷. S_{pd} assumes that alleles contribute equally to PD resistance and that a higher proportion of resistance alleles correlates with lower CFU in the individuals. S_{pd} ranges from 0 when no resistance alleles are present to 1 when all alleles across loci are homozygous for the resistant state. We estimated S_{pd} for different sets of SNPs: i) all candidate SNPs across the genome (25 SNPs), ii) all candidate SNPs on chromosome 14 (16 of the 25 SNPs); iii) all candidate SNPs in the region defined by *PdR1* (10 of 25 SNPs); iv) all candidate SNPs on chromosome 15 (6 of 25 SNPs). To control for potential ancestry effects, we also tested an analogous S_{pd} score, which we called the reference PD scores (R_{pd}). For R_{pd} we sampled 25 random reference SNP (SNPs that were not significant for any of the two GWA methods) and obtained a R_{pd} distribution based on 1000 R_{pd} values. Following the concept of S_{pd} , we also estimated a Kmer score (K_{pd}), which consisted of the proportion of kmers associated with resistance across populations. A value of 1 indicates that the individual has all the resistant kmers and none of the susceptible kmers (see Above), while a value of 0 indicates that the individual has all the susceptible kmers and none of the resistant

ones. Finally, we also estimated a CNV score (C_{pd}) that corresponds to the number of adaptive copy variants in an individual, where a higher number of CN variants indicates that the individual is more resistant to PD. For the S_{pd} , K_{pd} and C_{pd} scores, we estimated the fit between the scores and the bacterial load. Finally, for the genetic independent variables, we also analyzed the linear model between the assignments into genetic groups (K1 and K2) based on the admixture analyses and the concentration of bacterial load.

Climate modeling. We used BIO8 to model the future distribution of *Xylella fastidiosa*, assuming that 8 °C and 10 °C were predictive thresholds of potential *Xylella fastidiosa* presence and absence. We predicted two discrete types of locations: i) regions across the globe that currently have BIO8 > 8 °C or >10 °C in the present but predicted to have BIO8 < 8 °C and <10 °C in the future and ii) regions across the globe that currently have BIO8 < 8 °C and <10 °C in the present but predicted to have BIO8 > 8 °C and >10 °C in the future. To make these predictions, we downloaded the BIO8 data at a 2.5 minutes resolution from Worldclim 2¹⁰⁸ for the present and for 54 climatic models in the future to consider the uncertainty in future climate projections. These future climate models included five global circulation models (GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM1-2-HR, MRI-ESM2-0, UKESM1-0-LL), three-time periods (2041-2060; 2061-2080; 2081-2100) and 4 shared socioeconomic pathways (SSPs: 126, 245, 370, 585). To plot the areas where the climate is expected to change across the 10 °C threshold, for the present and the 54 future layers we set the raster layers to 1 and 0 if they were above or below the BIO8 temperature threshold. Next, we subtracted each future layer from the present layer. Areas with values of -1 indicate that a region in the future layer is expected to be under the threshold but is currently above the threshold. Regions with values of 0 indicated that the threshold remained unchanged in present and future layers. Finally, regions with values of +1 indicated that the future layer was expected to become above the threshold in the future. We performed this calculation for the 54 layers (or models) and created a sum_raster object by summing the 54 raster layers.

We tested how BIO8 will change in the future for regions where *V. arizonica*, *V. vinifera* (grapevines), *Citrus* sp., *Olea europaea* (olives), *Prunus amygdalus* (almond) and *Coffea* sp currently grow. For this, we first used the *gbif* function in the R *dismo* package¹⁰⁹ to download all the known locations of *V. arizonica*, grapevines, olives, almonds, coffee sp. and *Citrus* sp (*gbif.org*; download data: 2022-06-06). Next, for each species we used the *CoordinateCleaner* package in R¹¹⁰ to remove locations that 1) were duplicated; 2) had equal longitude and latitude; 3) were next to country centroids, capitals of countries, biological stations or *gbif* headquarters; 4) were in the sea; 5) were outliers based on the "quantile" option. Finally, we also removed locations if they were the only report in a given country, suggesting they may be outliers. After cleaning the data, we retained 6204 locations (from 11,834) for *Coffea* sp.; 3386 locations (from 9992) for almonds; 1111 locations (from 1155) for *V. arizonica*; 5256 locations (from 7853) for *Citrus* sp.; 174,713 locations (from 204,775) for olives; and 33,225 locations (from 47,075) for grapevines. For each species, we used the *extract* function in the *raster* R package¹¹¹ to obtain for each location whether climate is expected to cross the 8 °C and 10 °C BIO8 threshold. For each species we estimated the percentage of locations that are expected to move below and above the 8 °C and 10 °C threshold across the 54 layers.

Data availability

Raw genome resequencing data of the 147 new *V. arizonica* accessions and the PD resistance *V. vinifera* breeding lines are available at NCBI under BioProject ID PRJNA842753. The remaining 20 *V. arizonica* accessions, which were previously published, are available at NCBI under BioProject PRJNA731597. The RNAseq data were deposited in NCBI Bioproject PRJNA956994. The published long-read *V. arizonica* genome sequencing data were taken from NCBI BioProject ID PRJNA593045. The genome files and a genome browser can also be found at www.grapegenomics.com/pages/Vari/.

Code availability

Analyses in this paper relied on published methods, the sources for which has been cited throughout. Climate change analyses relied on previously published pipelines⁴⁶, which can be accessed here: https://static-content.springer.com/esm/art%3A10.1038%2F41559-021-01526-9/MediaObjects/41559_2021_1526_MOESM1_ESM.pdf.

Received: 23 November 2022; Accepted: 12 May 2023;

Published online: 30 May 2023

References

1. Zhao, C. et al. Temperature increase reduces global yields of major crops in four independent estimates. *Proc. Natl Acad. Sci. USA*. **114**, 9326–9331 (2017).

2. Ramankutty, N., Foley, J. A., Norman, J. & McSweeney, K. The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. *Glob. Ecol. Biogeogr.* **11**, 377–392 (2002).
3. Velásquez, A. C., Castroverde, C. D. M. & He, S. Y. Plant-Pathogen Warfare under Changing Climate Conditions. *Curr. Biol.* **28**, R619–R634 (2018).
4. Burdon, J. J. & Zhan, J. Climate change and disease in plant communities. *PLoS Biol.* **18**, e3000949 (2020).
5. European Food Safety Authority (EFSA). Update of the *Xylella* spp. host plant database - systematic literature search up to 30 June 2019. *EFSA J.* **18**, e06114 (2020).
6. Sicard, A. et al. *Xylella fastidiosa*: Insights into an Emerging Plant Pathogen. *Annu. Rev. Phytopathol.* **56**, 181–202 (2018).
7. Schneider, K. et al. Impact of *Xylella fastidiosa* subspecies *pauca* in European olives. *Proc. Natl Acad. Sci. USA.* **117**, 9250–9259 (2020).
8. Frem, M. et al. The potential direct economic impact and private management costs of an invasive alien species: *Xylella fastidiosa* on Lebanese wine grapes. *NeoBiota* **70**, 43–67 (2021).
9. Su, C.-C. et al. Pierce's disease of grapevines in Taiwan: Isolation, cultivation and pathogenicity of *Xylella fastidiosa*. *J. Phytopathol.* **161**, 389–396 (2013).
10. National Research Council, Division on Earth and Life Studies, Board on Agriculture and Natural Resources & Committee on California Agricultural Research Priorities: Pierce's Disease. *California Agricultural Research Priorities: Pierce's Disease*. (National Academies Press, 2004).
11. Bartoli, C. & Roux, F. Genome-Wide Association Studies In Plant Pathosystems: Toward an Ecological Genomics Approach. *Front. Plant Sci.* **8**, 763 (2017).
12. García, A. L., Torres, S. C. Z., Heredia, M. & Lopes, S. A. Citrus Responses to *Xylella fastidiosa* Infection. *Plant Dis.* **96**, 1245–1249 (2012).
13. De Souza, A. A., Takita, M. A., Amaral, A. M., do, Coletta-Filho, H. D. & Machado, M. A. Citrus responses to *Xylella fastidiosa* infection, the causal agent of citrus variegated chlorosis. *Tree Forestry Sci. Biotechnol.* **3**, 73–80 (2009).
14. Novelli, S. et al. Plant defense factors involved in *Olea europaea* resistance against *Xylella fastidiosa* infection. *J. Plant Res.* **132**, 439–455 (2019).
15. Giampetruzzi, A. et al. Transcriptome profiling of two olive cultivars in response to infection by the CoDiRO strain of *Xylella fastidiosa* subsp. *pauca*. *BMC Genom.* **17**, 475 (2016).
16. Rapicavoli, J., Ingel, B., Blanco-Ulate, B., Cantu, D. & Roper, C. *Xylella fastidiosa*: an examination of a re-emerging plant pathogen. *Mol. Plant Pathol.* **19**, 786–800 (2018).
17. Hoddle, M. S. The potential adventive geographic range of glassy-winged sharpshooter, *Homalodisca coagulata* and the grape pathogen *Xylella fastidiosa*: implications for California and other grape growing regions of the world. *Crop Protection* **23**, 691–699 (2004).
18. Ruel, J. J. & Andrew Walker, M. Resistance to Pierce's Disease in *Muscadinia rotundifolia* and Other Native Grape Species. *Am. J. Enol. Vitic.* **57**, 158–165 (2006).
19. Krivanek, A. F., Riaz, S. & Walker, M. A. Identification and molecular mapping of *PdR1*, a primary resistance gene to Pierce's disease in *Vitis*. *Theor. Appl. Genet.* **112**, 1125–1131 (2006).
20. Riaz, S., Krivanek, A. F., Xu, K. & Walker, M. A. Refined mapping of the Pierce's disease resistance locus, *PdR1*, and Sex on an extended genetic map of *Vitis rupestris* x *V. arizonica*. *Theor. Appl. Genet.* **113**, 1317–1329 (2006).
21. Riaz, S. et al. Using marker-assisted selection to breed Pierce's disease-resistant grapes. *Am. J. Enol. Vitic.* **60**, 199–207 (2009).
22. Agüero, C. B., Riaz, S., Tenschler, A. C., Bistué, C. & Walker, M. A. Molecular and functional characterization of two RGA type genes in the *PdR1b* locus for Pierce's disease resistance in *Vitis arizonica/candicans*. *Plant Cell Tissue Organ Cult.* <https://doi.org/10.1007/s11240-022-02366-6> (2022).
23. Zaini, P. A. et al. Molecular Profiling of Pierce's Disease Outlines the Response Circuitry of *Vitis vinifera* to *Xylella fastidiosa* Infection. *Front. Plant Sci.* **9**, 771 (2018).
24. Yang, L. et al. Proteomic analysis of grapevine stem in response to *Xylella fastidiosa* inoculation. *Physiol. Mol. Plant Pathol.* **75**, 90–99 (2011).
25. Rodrigues, C. M., de Souza, A. A., Takita, M. A., Kishi, L. T. & Machado, M. A. RNA-Seq analysis of Citrus reticulata in the early stages of *Xylella fastidiosa* infection reveals auxin-related genes as a defense response. *BMC Genom.* **14**, 676 (2013).
26. Leamy, L. J., Zhang, H., Li, C., Chen, C. Y. & Song, B.-H. A genome-wide association study of seed composition traits in wild soybean (*Glycine soja*). *BMC Genom.* **18**, 18 (2017).
27. Zhang, H., Song, Q., Griffin, J. D. & Song, B.-H. Genetic architecture of wild soybean (*Glycine soja*) response to soybean cyst nematode (*Heterodera glycines*). *Mol. Genet. Genom.* **292**, 1257–1265 (2017).
28. Migicovsky, Z. & Myles, S. Exploiting Wild Relatives for Genomics-assisted Breeding of Perennial Crops. *Front. Plant Sci.* **8**, 460 (2017).
29. Frem, M. et al. *Xylella fastidiosa* invasion of new countries in Europe, the Middle East and North Africa: Ranking the potential exposure scenarios. *NeoBiota* **59**, 77–97 (2020).
30. Morales-Cruz, A. et al. Introgression among North American wild grapes (*Vitis*) fuels biotic and abiotic adaptation. *Genome Biol.* **22**, 254 (2021).
31. Riaz, S., Tenschler, A. C., Heinitz, C. C., Huerta-Acosta, K. G. & Walker, M. A. Genetic analysis reveals an east-west divide within North American *Vitis* species that mirrors their resistance to Pierce's disease. *PLoS One* **15**, e0243445 (2020).
32. Caye, K., Jumentier, B., Lepeule, J. & François, O. LFMM 2: fast and accurate inference of gene-environment associations in genome-wide studies. *Mol. Biol. Evol.* **36**, 852–860 (2019).
33. Hao, Z., Gao, J., Song, Y., Yang, R. & Liu, D. Genome-wide hierarchical mixed model association analysis. *Brief. Bioinform.* **22**, bbab306 (2021).
34. Riaz, S. et al. Fine-scale genetic mapping of two Pierce's disease resistance loci and a major segregation distortion region on chromosome 14 of grape. *Theor. Appl. Genet.* **117**, 671–681 (2008).
35. Voickek, Y. & Weigel, D. Identifying genetic variants underlying phenotypic variation in plants without complete genomes. *Nat. Genet.* **52**, 534–540 (2020).
36. Levy, M., Wang, Q., Kaspi, R., Parrella, M. P. & Abel, S. Arabidopsis IQD1, a novel calmodulin-binding nuclear protein, stimulates glucanase accumulation and plant defense. *Plant J.* **43**, 79–96 (2005).
37. Lv, T. et al. The Calmodulin-Binding Protein IQM1 Interacts with CATALASE2 to Affect Pathogen Defense. *Plant Physiol.* **181**, 1314–1327 (2019).
38. Kalde, M., Nühse, T. S., Findlay, K. & Peck, S. C. The syntaxin SYP132 contributes to plant resistance against bacteria and secretion of pathogenesis-related protein 1. *Proc. Natl Acad. Sci. USA.* **104**, 11850–11855 (2007).
39. Liu, P.-L., Du, L., Huang, Y., Gao, S.-M. & Yu, M. Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. *BMC Evol. Biol.* **17**, 47 (2017).
40. Du, C. et al. Prokaryotic expression, purification, physicochemical properties and antifungal activity analysis of phloem protein PP2-A1 from cucumber. *Int. J. Biol. Macromol.* **194**, 395–401 (2022).
41. Samac, D. A., Hironaka, C. M., Yallaly, P. E. & Shah, D. M. Isolation and Characterization of the Genes Encoding Basic and Acidic Chitinase in *Arabidopsis thaliana*. *Plant Physiol.* **93**, 907–914 (1990).
42. Jiang, Y. et al. Overexpression of Poplar PtrWRKY89 in Transgenic Arabidopsis Leads to a Reduction of Disease Resistance by Regulating Defense-Related Genes in Salicylate- and Jasmonate-Dependent Signaling. *PLoS One* **11**, e0149137 (2016).
43. Zaidi, S. S.-E.-A., Mukhtar, M. S. & Mansoor, S. Genome Editing: Targeting Susceptibility Genes for Plant Disease Resistance. *Trends Biotechnol.* **36**, 898–906 (2018).
44. Grapevine plant named 'Camminare Noir'. <https://patents.justia.com/patent/PP32929>.
45. Fitzpatrick, M. C. & Keller, S. R. Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecol. Lett.* **18**, 1–16 (2015).
46. Aguirre-Liguori, J. A., Ramírez-Barahona, S. & Gaut, B. S. The evolutionary genomics of species' responses to climate change. *Nat. Ecol. Evolution* **5**, 1350–1360 (2021).
47. European Food Safety Authority (EFSA). Update of the *Xylella* spp. host plant database. *EFSA J.* **16**, e05408 (2018).
48. Purcell, A. H. Environmental Therapy for Pierce's Disease of Grapevines. *Plant Disease.* **64**, 388 (1980).
49. Lieth, J. H., Meyer, M. M., Yeo, K.-H. & Kirkpatrick, B. C. Modeling cold curing of Pierce's disease in *Vitis vinifera* 'Pinot Noir' and 'Cabernet Sauvignon' grapevines in California. *Phytopathology* **101**, 1492–1500 (2011).
50. Bosso, L., Di Febbraro, M., Cristinzio, G., Zoia, A. & Russo, D. Shedding light on the effects of climate change on the potential distribution of *Xylella fastidiosa* in the Mediterranean basin. *Biol. Invasions* **18**, 1759–1768 (2016).
51. Doucleff, M. et al. A genetic linkage map of grape, utilizing *Vitis rupestris* and *Vitis arizonica*. *Theor. Appl. Genet.* **109**, 1178–1187 (2004).
52. Zhou, Y. et al. The population genetics of structural variants in grapevine domestication. *Nat. Plants* **5**, 965–979 (2019).
53. Karasov, T. L., Shirsekar, G., Schwab, R. & Weigel, D. What natural variation can teach us about resistance durability. *Curr. Opin. Plant Biol.* **56**, 89–98 (2020).
54. Zhang, X.-M., Wu, Y.-F., Li, Z., Song, C.-B. & Wang, X.-P. Advancements in plant regeneration and genetic transformation of grapevine (*Vitis* spp.). *J. Integr. Agric.* **20**, 1407–1434 (2021).
55. Mortensen, J. A. The inheritance of resistance to Pierce's disease in *Vitis*. *J. Am. Soc. Hortic. Sci.* **92**, 331–337 (1968).
56. Riaz, S., Huerta-Acosta, K., Tenschler, A. C. & Walker, M. A. Genetic characterization of *Vitis* germplasm collected from the southwestern US and Mexico to expedite Pierce's disease-resistance breeding. *Theor. Appl. Genet.* **131**, 1589–1602 (2018).
57. Huerta-Acosta, K. G., Riaz, S., Tenschler, A. & Walker, M. A. Genetic Characterization of Pierce's Disease Resistance in a *Vitis arizonica/monticola* Wild Grapevine. *Am. J. Enol. Vitic.* **74**, 0740003 (2022).

58. Burghardt, L. T., Young, N. D. & Tiffin, P. A guide to genome-wide association mapping in plants. *Curr. Protoc. Plant Biol.* **2**, 22–38 (2017).
59. Aguirre-Liguori, J. A., Morales-Cruz, A. & Gaut, B. S. Climate Change, Crop Wild Relatives and Genomic Data: The persistence and agronomic utility of five *Vitis* species. *bioRxiv* 2021.12.10.472174 <https://doi.org/10.1101/2021.12.10.472174> (2021).
60. Raffini, F. et al. From Nucleotides to Satellite Imagery: Approaches to Identify and Manage the Invasive Pathogen *Xylella fastidiosa* and Its Insect Vectors in Europe. *Sustainability* **12**, 4508 (2020).
61. Hopkins, D. L. & Purcell, A. H. *Xylella fastidiosa*: Cause of Pierce's Disease of Grapevine and Other Emergent Diseases. *Plant Dis.* **86**, 1056–1066 (2002).
62. Godefroid, M. et al. Climate tolerances of *Philaenus spumarius* should be considered in risk assessment of disease outbreaks related to *Xylella fastidiosa*. *J. Pest Sci.* **95**, 855–868 (2022).
63. Tian, D., Traw, M. B., Chen, J. Q., Kreitman, M. & Bergelson, J. Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature* **423**, 74–77 (2003).
64. Godefroid, M., Cruaud, A., Streito, J.-C., Rasplus, J.-Y. & Rossi, J.-P. Forecasting future range shifts of *Xylella fastidiosa* under climate change. *Plant Pathol.* <https://doi.org/10.1111/ppa.13637> (2022).
65. Cohen, J., Agel, L., Barlow, M., Garfinkel, C. I. & White, I. Linking Arctic variability and change with extreme winter weather in the United States. *Science* **373**, 1116–1121 (2021).
66. Batareseh, T. N., Morales-Cruz, A., Ingel, B., Caroline Roper, M. & Gaut, B. S. Using genomes and evolutionary analyses to screen for host-specificity and arms-race dynamics in the plant pathogen *Xylella fastidiosa*. *Appl. Env. Microbiol.* **88**, e01220–e01222 (2022).
67. Godefroid, M., Cruaud, A., Streito, J.-C., Rasplus, J.-Y. & Rossi, J.-P. *Xylella fastidiosa*: climate suitability of European continent. *Sci. Rep.* **9**, 8844 (2019).
68. Moralejo, E. et al. Phylogenetic inference enables reconstruction of a long-overlooked outbreak of almond leaf scorch disease (*Xylella fastidiosa*) in Europe. *Commun. Biol.* **3**, 560 (2020).
69. Krivanek, A. F. & Walker, M. A. Vitis Resistance to Pierce's Disease Is Characterized by Differential *Xylella fastidiosa* Populations in Stems and Leaves. *Phytopathology* **95**, 44–52 (2005).
70. Massonnet, M. et al. The genetic basis of sex determination in grapes. *Nature Communications* **11**, 2902 (2020).
71. Minio, A., Cocheteu, N., Vondras, A., Massonnet, M. & Cantu, D. Assembly of complete diploid phased chromosomes from draft genome sequences. *G3* <https://doi.org/10.1093/g3journal/jkac143> (2022).
72. Zou, C. et al. Haplotyping the *Vitis* collinear core genome with rhAmpSeq improves marker transferability in a diverse genus. *Nat. Commun.* **11**, 413 (2020).
73. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
74. Andrews, S. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
75. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* 1303.3997 (2013).
76. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297–1303 (2010).
77. Institute, B. Picard tools. (2016).
78. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
79. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
80. Kornelissen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356 (2014).
81. Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. & Mayrose, I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **15**, 1179–1191 (2015).
82. Fritch, E., Schoville, S. D., Bouchard, G. & François, O. Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. *Molecular Biology and Evolution* **30**, 1687–1699 (2013).
83. Kassambara. Principal Component Analysis in R: prcomp vs princomp. *Statistical Tools for High-Throughput Data Analysis*.
84. Kang, H. M. et al. Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42**, 348–354 (2010).
85. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
86. Zhang, Y.-W. et al. mrMLM v4.0.2: An R Platform for Multi-locus Genome-wide Association Studies. *Genom. Proteom. Bioinforma.* **18**, 481–487 (2020).
87. Wen, Y.-J. et al. Methodological implementation of mixed linear models in multi-locus genome-wide association studies. *Brief. Bioinform* **19**, 700–712 (2018).
88. Kokot, M., Dlugosz, M. & Deorowicz, S. KMC 3: counting and manipulating k-mer statistics. *Bioinformatics* **33**, 2759–2761 (2017).
89. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012).
90. Zhang, J. & Madden, T. L. PowerBLAST: a new network BLAST application for interactive or automated sequence analysis and annotation. *Genome Res* **7**, 649–656 (1997).
91. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
92. Wang, X. et al. CNVcaller: highly efficient and widely applicable software for detecting copy number variations in large populations. *Gigascience* **6**, 1–12 (2017).
93. Kim, S. ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. *Commun. Stat. Appl. Methods* **22**, 665–674 (2015).
94. Zhang, C., Dong, S.-S., Xu, J.-Y., He, W.-M. & Yang, T.-L. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **35**, 1786–1788 (2019).
95. Robinson, P. & Others. Integrative genomics viewer (IGV): Visualizing alignments and variants. in *Computational Exome and Genome Analysis* 233–245 (Chapman and Hall/CRC, 2017).
96. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* **12**, 357–360 (2015).
97. El-Gebali, S. et al. The Pfam protein families database in 2019. *Nucleic Acids Res* **47**, D427–D432 (2019).
98. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
99. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
100. Massonnet, M. et al. Ripening Transcriptomic Program in Red and White Grapevine Varieties Correlates with Berry Skin Anthocyanin Accumulation. *Plant Physiol.* **174**, 2376–2396 (2017).
101. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
102. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* **4**, 1521 (2015).
103. Ellis, N., Smith, S. J. & Pitcher, C. R. Gradient forests: calculating importance gradients on physical predictors. *Ecology* **93**, 156–168 (2012).
104. Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M. & Keller, S. R. Genomic Prediction of (Mal)Adaptation Across Current and Future Climatic Landscapes. <https://doi.org/10.1146/annurev-ecolsys-020720-042553> (2020).
105. Waldvogel, A.-M. et al. Evolutionary genomics can improve prediction of species' responses to climate change. *Evolution Letters* **4**, 4–18 (2020).
106. Barbosa, A. M. fuzzySim: applying fuzzy logic to binary similarity indices in ecology. *Methods Ecol. Evol.* **6**, 853–858 (2015).
107. Bonin, A., Nicole, F., Pompanon, F., Maud, C. & Taberlet, P. Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conserv. Biol.* **21**, 697–708 (2007).
108. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
109. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. dismo: Species distribution modeling. <http://cran.r-project.org/web/packages/dismo/index.html> (2022).
110. Zizka, A. et al. CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods Ecol. Evol.* **10**, 744–751 (2019).
111. Hijmans, R. J. The raster package. <https://rspatial.org/raster/pkg/RasterPackage.pdf> (2021).

Acknowledgements

The authors are grateful for data generation by R. Gaut and the Genomics High Throughput Facility at UC Irvine. D. Seymour, C. Roper and three reviewers provided valuable comments. This work was funded by the NSF grant #1741627 to B.S.G., M.A.W. and D.C., and partially supported by funds to D.C. from the E&J. Gallo Winery and the Louis P. Martini Endowment in Viticulture. J.A.-L. was supported by a UC-Mexico postdoctoral fellowship.

Author contributions

A.M.-C.: performed GWA and other analyses. Wrote the paper. J.A.-L.: performed GWA and climate analyses. Wrote the paper. M.M.: performed gene expression analyses and candidate gene annotation within *PdRI*. A.M.: performed analyses to generate the updated genome reference. M.Z.: contributed to gene expression analyses. N.C.: contributed to genomic analyses and genome annotations. A.W.: sampled the materials and supervised phenotypic assays. S.R.: performed phenotypic assays and contributed materials and data from previous work. Y.Z.: collated samples, collated data and

performed initial SNP analyses. D.C.: provided ideas, data, supervision and funding. Edited the paper. B.S.G.: provided ideas, data, supervision and funding. Wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-04938-4>.

Correspondence and requests for materials should be addressed to Yongfeng Zhou, Dario Cantu or Brandon S. Gaut.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Diane Saunders and Christina Karlsson Rosenthal.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

CHAPTER 3

Haplotype graph analysis of *PdR1* uncovers resistance diversity to Pierce's Disease in *Vitis arizonica* and its hybrids

Haplotype graph analysis of *PdR1* uncovers resistance diversity to Pierce's disease in *Vitis arizonica* and its hybrids

Mélanie Massonnet ¹, Mirella Zaccheo ¹, Noé Cochetel ¹, Rosa Figueroa-Balderas ¹, Summaira Riaz,²
Dario Cantu ^{1,3,*}

¹Department of Viticulture and Enology, University of California Davis, Davis, CA 95616, United States

²U.S. Department of Agriculture-Agricultural Research Service, San Joaquin Valley Agricultural Center, Parlier, CA 93648, United States

³Genome Center, University of California Davis, Davis, CA 95616, United States

*Corresponding author: Department of Viticulture and Enology, University of California Davis, Davis, CA 95616, United States. Email: dacantu@ucdavis.edu

Previous genetic mapping studies indicate that multiple haplotypes of the *Pierce's disease* (PD) Resistance 1 (*PdR1*) locus occur in *Vitis arizonica* and its hybrids. To characterize sequence diversity at this locus, we assembled chromosome-scale diploid genomes for four PD-resistant (PD-R) accessions: b43-17 (*PdR1a*⁺/*PdR1b*⁺), the backcross 07744-094 (*PdR1c*⁺/*PdR1*⁺), b46-43 (*PdR1e*⁺/*PdR1f*⁺), and b42-26 (*PdR1*⁺/*PdR1*⁺), which displays quantitative PD resistance not associated with *PdR1*. Haplotype resolution of *PdR1a*, *PdR1b*, *PdR1c*, and *PdR1e* revealed substantial variation in intergenic repeat content and gene composition between *PdR1* haplotypes and their alternatives at the *PdR1* locus not associated with PD resistance phenotype (*PdR1*⁺), as well as among *PdR1* haplotypes, demonstrating extensive sequence diversity at the *PdR1* locus. Sequence graph analysis uncovered substantial structural divergence concentrated in approximately one quarter of the locus, together with smaller-scale variation across haplotypes. This analysis identified *PdR1*-specific graph nodes, showing that *PdR1a* and *PdR1b* share most of their *PdR1*-specific features, whereas *PdR1c* contains the highest number of private nodes, followed by *PdR1e*. Integration of sequence graph features with gene expression data further refined a set of defense-related candidate genes within *PdR1c*. Together, these results identify candidate genes for functional validation and indicate that multiple resistance determinants co-localized within the *PdR1* locus may contribute to PD resistance, highlighting opportunities for targeted genetic improvement strategies.

Keywords: *Vitis arizonica*; Pierce's disease; *Xylella fastidiosa*; disease resistance; haplotype-resolved genomes; sequence graph

Introduction

Pierce's disease (PD) is caused by the bacterium *Xylella fastidiosa* ssp. *fastidiosa* (hereafter *Xf*), which colonizes grapevine xylem. In California, PD remained limited until the late 1990s when the introduction of the non-native glassy-winged sharpshooter (*Homalodisca vitripennis*) in the southern part of the state led to a rapid rise in disease incidence (Rapicavoli et al. 2018). These insects, along with the native blue-green sharpshooter (*Graphocephala atropunctata*) and spittlebug species (family Cercopidae), are principal vectors of *Xf*. All *Vitis vinifera* cultivars are susceptible to PD, whereas some wild grape species from the southern US and northern Mexico, including *V. arizonica* and its hybrids, are PD-resistant.

In grapes, *Xf* colonization triggers the formation of balloon-shaped tyloses in the xylem, which help block pathogen spread and limit air embolisms (Castro et al. 2021). While this defense response is effective in PD-resistant grapes, tylose formation is spatially and temporally misaligned with *Xf* colonization in PD-susceptible cultivars, leading to extensive vessel occlusion (Sun et al. 2013). Coupled with systemic *Xf* spread through bacterial biofilms, excess tylose development sharply reduces stem hydraulic conductivity in susceptible grapevines (Rapicavoli et al. 2018). This results in leaf scorching, berry desiccation, irregular periderm development, abnormal leaf abscission, and vine death (Rapicavoli et al. 2018).

PD management mostly relies on the use of insecticides to control vectors and the removal of infected plants (Haviland et al. 2021). Despite these practices, annual production loss is estimated at \$56.1 million per year in California (Tumber et al. 2014). In addition, growing application restrictions and emerging pesticide resistance have made this approach unsustainable (Haviland et al. 2021). Therefore, the long-term solution is the development of grape cultivars with strong and durable PD resistance combined with desirable agronomic and enological traits. Such resistance can be achieved through the introgression of functionally diverse PD resistance-associated genes or alleles into *V. vinifera* (Michelmore et al. 2013).

One resistance (R) locus located on chromosome 14, PD Resistance 1 (*PdR1*), was identified in grapes through genetic mapping (Krivanek et al. 2006; Riaz et al. 2006). Using marker-based maps and tightly linked molecular markers, putative *PdR1* haplotypes were distinguished among *V. arizonica* accessions and their hybrids based on marker size polymorphisms rather than sequence data (Riaz et al. 2008, 2018). The presence of multiple marker-defined haplotypes at the same locus (*PdR1a-f*) suggests that PD resistance may involve different genes or allelic variants within *PdR1* (Table 1). The diploid genome of *V. arizonica* b40-14, which carries *PdR1c* and *d* (*PdR1c*⁺/*PdR1d*⁺), together with a

Received on 22 December 2025; accepted on 23 April 2026

© The Author(s) 2026. Published by Oxford University Press on behalf of The Genetics Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

genome-wide association and gene expression profiling, enabled the identification of four candidate R genes encoding extracellular receptors: two leucine-rich repeat (LRR) receptor-like proteins (RLPs), one LRR receptor-like kinase (LRR-RLK), and one lysin motif (LysM)-RLK (Morales-Cruz et al. 2023). However, limited haplotype information prevented confirmation of the locus phasing and whether the studied *PdR1* haplotype corresponded to the c or d haplotype of b40-14.

PdR1b has already been deployed in the University of California Davis grape breeding program and underlies the five PD-resistant cultivars released in 2019 (<https://fps.ucdavis.edu/newsarticle.cfm?newsid=57>). These cultivars have generated significant industry interest, motivating efforts to characterize additional *PdR1* haplotypes that could further strengthen resistance and improve its durability in future breeding. In this study, to better understand the range of resistance encoded at this locus, we assess sequence diversity among five *PdR1* haplotypes and its impact on genes potentially involved in *Xf* resistance. We sequenced, assembled, and annotated diploid chromosome-scale genomes for four PD-resistant accessions: b43-17 (*PdR1a**/*PdR1b**), the back-cross 07744-094 (*PdR1c**/*PdR1*⁻), b46-43 (*PdR1e**/*PdR1f**), and b42-26 (*PdR1*⁻/*PdR1*⁻), which displays quantitative PD resistance but not associated with *PdR1* (Table 1) (Riaz et al. 2018). After localizing the *PdR1* haplotypes using marker data and phasing them, we compared the gene content potentially associated with *Xf* resistance across five *PdR1* haplotypes and three alternative haplotypes. *PdR1*⁻ will be used to describe the alternative haplotype at the *PdR1* locus not associated with PD-resistance phenotype. A sequence graph built from both *PdR1* and *PdR1*⁻ haplotypes allowed us to identify variants specific to the *PdR1* haplotypes, compare these variants across haplotypes, and determine which ones altered the coding sequences of defense-related genes. Finally, we show how the information of the sequence graph, coupled with gene expression, can help to narrow down candidate genes among the refined region of *PdR1c* (Morales-Cruz et al. 2023).

Materials and methods

Plant material

Young leaves from the accessions b43-17 (*PdR1a**/*PdR1b**), 07744-094 (*PdR1c**/*PdR1*⁻), b46-43 (*PdR1e**/*PdR1f**), and b42-26 (*PdR1*⁻/*PdR1*⁻) were collected, immediately flash frozen, and ground in liquid nitrogen. All four accessions were field-grown plants maintained at the Department of Viticulture and Enology, University of California, Davis, CA, at the time of collection. Accessions b43-17, 07744-094, b46-43, and b42-26 were planted in 2003, 2008, 2014, and 2015, respectively.

DNA extraction, library preparation, and sequencing

High-molecular-weight genomic DNA was isolated from the young leaves of the four accessions like in Chin et al. (Chin et al. 2016). DNA concentration was evaluated using the DNA High Sensitivity Kit on a Qubit 2.0 Fluorometer (Life Technologies, CA, USA), and DNA purity with a Nanodrop 2000 spectrophotometer (Thermo Scientific, IL, USA). Assessment of the DNA integrity was done using the FEMTO Pulse system (Agilent, CA, USA). Preparation of the HiFi libraries was performed using the SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, CA, USA) following the manufacturer's protocol. Libraries were size-selected using 3.1 X v/v of 35% Ampure PB beads (Pacific Biosciences, CA, USA), and sequenced on a PacBio Sequel II platform (DNA Technology Core Facility, University of California,

Davis, CA, USA). Sequencing resulted in 54, 54, 57, and 40X of coverage of a 500-Mbp haplotype for b43-17 (*PdR1a**/*PdR1b**), 07744-094 (*PdR1c**/*PdR1*⁻), b46-43 (*PdR1e**/*PdR1f**), and b42-26 (*PdR1*⁻/*PdR1*⁻), respectively (Supplementary Table 1).

Genome assembly

HiFi reads were assembled using hifiasm (Cheng et al. 2021) with the option “-n 13” for b43-17 (*PdR1a**/*PdR1b**), 07744-094 (*PdR1c**/*PdR1*⁻), and b42-26 (*PdR1*⁻/*PdR1*⁻), and “-n 25” for b46-43 (*PdR1e**/*PdR1f**) (Supplementary Table 2). Diploid chromosome-scale pseudomolecules were scaffolded from contigs using HaploSync v1.0 (Minio et al. 2022) and a consensus grape synteny map (Cochetel et al. 2025) (Supplementary Table 3). Telomere repeat units were searched as described in Massonnet et al. (Massonnet et al. 2025) using TIDK v.0.2.63 (<https://github.com/tolk/tidk/telomeric-identifier>). Gene space completeness of the genome assemblies was evaluated using BUSCO v.5.4.7 with the library embryophyte_odb10 (Waterhouse et al. 2018).

Genome annotation

Repeat and gene annotations were performed as in Massonnet et al. (Massonnet et al. 2025). Repeat annotation was made using RepeatMasker v.4.1.7-p1 (Smit et al. 2013) and a grape repeat library (Minio et al. 2019). PN40024 V5.1 predicted protein-coding gene sequences (https://grapedia.org/t2t_annotation/), along with the high-quality transcripts obtained from the Iso-Seq full-length reads of *V. arizonica* b40-14 (Cochetel et al. 2023) and *V. vinifera* cv. Cabernet Sauvignon (Minio et al. 2019), were used as transcriptomic evidence for the gene annotation. For each accession, a set of high-quality gene models was created using the PASA pipeline v2.5.3 (<https://github.com/PASAPipeline/PASAPipeline>), GMAP v.2024-11-20 (Wu and Watanabe 2005), pblat v.2.5.1 (Wang and Kong 2019), and the transcriptomic evidence aforementioned. The set of high-quality gene models was then used as input for the *ab initio* predictors Augustus v.3.5.0 (Stanke et al. 2006) and GeneMark v.3.68.lic (Lomsadze et al. 2005). Consensus models from the gene models produced by the *ab initio* predictions and PASA were made using EvidenceModeler v.2.1.0 (Haas et al. 2008). Gene models encoding protein sequences with a length less than 50 amino acids, or in-frame stop codons, or lacking a start methionine or a stop codon, were removed.

Functional annotation

Predicted proteins of the four genomes were aligned onto the proteins of *Arabidopsis thaliana* (Araport11_genes.201606.pep.fasta; <https://www.arabidopsis.org/download/index.jsp>) using BLASTP v.2.12.0+ (Camacho et al. 2009). Homologous proteins were identified by: first, keeping the alignments with an identity greater than 30%, and both reference:query and query:reference coverages between 75 and 125%; second, selecting the alignment with the highest product of identity by query coverage by reference coverage for each grape protein. The same method was used to determine homologous proteins in *V. vinifera* PN40024 genome V1 annotation and V5.1 annotations (Jaillon et al. 2007; Shi et al. 2023).

To identify protein domains, hmmsearch from HMMER v.3.3.1 (Eddy 2011) was used with the Pfam-A Hidden Markov Models (HMM) database (El-Gebali et al. 2019) (downloaded on March 12th, 2025). Protein domains with an independent E-value less than 1.0, and at least 50% of the HMM, were selected (Supplementary Tables 4–7).

Transmembrane helices were predicted with TMHMM2 v2.0c (Krogh et al. 2001).

PdR1 region localization

The five PdR1 haplotypes and the three PdR1⁻ haplotypes were located by *in silico* amplification of two PdR1-associated markers, VMCNg2b7.2 and UDV095 (Riaz et al. 2008), on the genome assemblies using *dispr* (<https://github.com/douglasgscfield/dispr>).

Haplotyping

To identify PdR1c in 07744-094 (PdR1c⁺/PdR1⁻) and PdR1b in b43-17 (PdR1a⁺/PdR1b⁺), we used short DNA-seq reads from *V. arizonica* b40-14 (PdR1c⁺/PdR1d⁺) and five backcrossed cultivars carrying PdR1b: Ambulo blanc, Caminante blanc, Camminare noir, Errante noir, and Paseante noir (Morales-Cruz et al. 2023), and a coverage analysis of unambiguous alignments. Adapter sequence removal and quality read selection was performed using Trimmomatic v.0.36 (Bolger et al. 2014) with the following parameters: "LEADING:3 TRAILING:3 SLIDINGWINDOW:10:20 MINLEN:36 CROP:150". High-quality reads of *V. arizonica* b40-14 and the five backcrossed cultivars were aligned on 07744-094 and b43-17 genomes, respectively, using BWA v0.7.17 (Li and Durbin 2009) with the parameter "-a" to report all multiple alignments. Unambiguous alignments (i.e. with no mismatches) were retained with *bamtools* filter v.2.5.1 (Barnett et al. 2011) and the option "NM:0". Base coverage was obtained using *genomcov* from BEDTools v2.29.1 (Quinlan 2014). Coverage in repetitive elements was removed using BEDTools intersect. Median coverage per window of 10-kbp was computed by BEDTools map. Median coverage per 10-kbp window was normalized by dividing by the sequencing coverage of the accessions. SNPs and INDELs between the two haplotypes of b43-17 were called using *NUCmer* from *MUMmer* v.4.0.0 (Marçais et al. 2018) with the "--mum" option and *show-snps* with the parameters "-Clr".

Sequence graph construction

Sequence graph of the five PdR1 haplotypes and the three PdR1⁻ alternative haplotypes was built using the *nf-core/pangenome* pipeline (Heumos et al. 2024) with *NextFlow* v.25.04.6 (Di Tommaso et al. 2017) and the parameters "-r 1.1.2 -profile singularity". 2D representation was generated using *ODGI* v0.9.3 (Guarracino et al. 2022).

Alignment of PD resistance-associated kmers

PD resistance-associated kmers from *V. arizonica* were retrieved from Morales-Cruz et al. (Morales-Cruz et al. 2023). They were mapped on the genome of 07744-094 (PdR1c⁺/PdR1⁻) using *BLASTN* v.2.12.0+ (Camacho et al. 2009) and a word size of 8 bp. Full-length and perfect alignments (i.e. 100% coverage and identity) were retained.

Gene expression analysis

RNA-seq reads of three PD-resistant 07744 lines were retrieved from NCBI BioProject PRJNA956994 (Morales-Cruz et al. 2023). High-quality reads were obtained using Trimmomatic v.0.36 (Bolger et al. 2014) with the following parameters: "LEADING:3 TRAILING:3 SLIDINGWINDOW:10:20 MINLEN:36" (Supplementary Table 11). Transcript abundance was assessed with *Salmon* v.1.5.1 (Patro et al. 2017). A transcriptome index file was built using the protein-coding sequences of 07744-094 and its genome as decoy, using the parameters "-k 31 --keepDuplicates". Transcript abundance was quantified with the parameters "--gcBias --seqBias --validateMappings". Gene expression in transcript per million (TPM) was extracted from quantification files with the R package *tximport* v.1.30.0 (Soneson et al. 2016). Differential gene expression analysis was performed using *DESeq2* v.1.42.1 (Love et al. 2014).

Results

Diploid chromosome-scale genome assemblies

To characterize the sequence diversity among the PdR1 haplotypes, we sequenced the genomes of three PD-resistant *V. arizonica* hybrids: *V. arizonica* x *V. candicans* hybrid b43-17 (PdR1a⁺/PdR1b⁺), *V. arizonica glabrous* x *V. monticola* hybrid b46-43 (PdR1e⁺/PdR1f⁺), and *V. arizonica* x *V. girdiana* hybrid b42-26 (PdR1⁻/PdR1⁻) which has quantitative PD resistance not associated with PdR1 (Table 1) (Riaz et al. 2018), using long-read DNA sequencing (PacBio HiFi reads; Supplementary Table 1). In addition, we sequenced the genome of the backcross line 07744-094 to obtain the PdR1c haplotype from *V. arizonica* b40-14 (PdR1c⁺/PdR1d⁺), because in the available b40-14 genome, the two PdR1 haplotypes are fragmented and not haplotype-phased (Supplementary Fig. 1) (Morales-Cruz et al. 2023).

Draft diploid genomes were composed of an average of 192 ± 43.4 contigs (Supplementary Table 2). After pseudomolecule reconstruction, each set of 19 chromosomes spanned 491.9 ± 67.8 Mbp, with unanchored contigs accounting for only 1.7 ± 0.8% of the assembly (16.7 ± 8.1 Mbp) (Table 2; Supplementary Table 3). On average, 99.2 ± 0.2% of expected single-copy orthologs (BUSCOs) were detected in each set of 19 chromosomes, indicating a high completeness. Furthermore, we detected telomeric repeats at both ends of 30, 28, 24, and 16 chromosomes for 07744-094, b42-26, b46-43, and b43-17, respectively, and at one end of 8, 10, 11, and 21 chromosomes, respectively, out of 38. These results indicate that the four haplotype-resolved assemblies are highly contiguous and complete.

PdR1 localization and haplotyping

The boundaries of the PdR1 and PdR1⁻ haplotypes were located through *in silico* amplification of the PdR1-linked markers VMCNg2b7.2 and UDV095 (Riaz et al. 2008) (Supplementary Table 8). Three additional markers (Ch14-70, Ch14-77, and Ch14-02) were used to confirm their localization (Riaz et al. 2018). Amplification revealed amplicon-size variation between accessions for all markers. Variation within accessions, reflecting differences between the two homologous chromosomes, was also observed for at least one marker in each accession, except b46-43, whose two haplotypes produced identical amplicon sizes for all markers (Supplementary Table 8). The size of the PdR1 region varied across the four grape accessions, providing further evidence of haplotype diversity (Table 3; Supplementary Table 9). Substantial differences in locus size between haplotypes were found in 07744-094 (PdR1c⁺/PdR1⁻) (1,103.9 and 842.8 kbp for Haplotype 1 and 2, respectively)

Table 1. List of the PD-resistant grape accessions whose genome was sequenced in this study.

Accession	PdR1 composition	Species	Reference
b43-17	PdR1a ⁺ / PdR1b ⁺	<i>V. arizonica</i> x <i>V. candicans</i> hybrid	Krivanek et al. 2006 Riaz et al. 2008
07744-094	PdR1c ⁺ / PdR1 ⁻	(<i>V. rupestris</i> Wichita refuge x <i>V. arizonica glabrous</i> b40-14) x <i>V. vinifera</i> cv. Airen	Riaz et al. 2018
b46-43	PdR1e ⁺ / PdR1f ⁺	<i>V. arizonica glabrous</i> x <i>V. monticola</i> hybrid	Riaz et al. 2018
b42-26	PdR1 ⁻ /PdR1 ⁻	<i>V. arizonica</i> x <i>V. girdiana</i> hybrid	Riaz et al. 2018

Table 2. Statistics of the four genome assemblies.

Accession	b43-17			07744-094			b46-43			b42-26		
PdR1 composition	PdR1a ⁺ /PdR1b ⁺			PdR1c ⁺ /PdR1 ⁻			PdR1e ⁺ /PdR1f ⁺			PdR1 ⁻ /PdR1 ⁻		
Haplotype	Hap1	Hap2	Unp	Hap1	Hap2	Unp	Hap1	Hap2	Unp	Hap1	Hap2	Unp
Assembly length (Mbp)	502.1	488.3	15.4	484.6	491.8	13.7	501.1	489.8	9.4	493.6	484.1	28.1
Number of sequences	19	19	81	19	19	70	19	19	56	19	19	93
Telomeres	29	24	0	33	35	0	29	30	0	33	33	0
Number of genes (x1,000)	39.2	38.3	1.0	37.5	36.1	3.3	32.1	31.4	0.8	33.1	32.4	2.1
Repeats (%)	56.3	56.0	79.9	56.5	56.0	76.8	57.3	56.6	69.7	56.1	55.8	78.9
BUSCOs (%)	99.2	98.9	1.1	99.4	99.5	0.4	99.1	99.1	0.1	99.1	99.1	1.8

Abbreviations: Hap, Haplotype; Unp, Unplaced contigs.

Table 3. Features of the PdR1 and PdR1⁻ haplotypes.

Accession	b43-17		07744-094		b46-43	b42-26	
PdR1 composition	PdR1a ⁺ /PdR1b ⁺		PdR1c ⁺ /PdR1 ⁻		PdR1e ⁺ /PdR1f ⁺	PdR1 ⁻ /PdR1 ⁻	
Haplotype	Hap1 PdR1b	Hap2 PdR1a	Hap1 PdR1 ⁻	Hap2 PdR1c	Hap1 PdR1e	Hap1 PdR1 ⁻	Hap2 PdR1 ⁻
Length (kbp)	889.4	887.7	1,103.9	842.8	877.5	969.7	848.1
Intergenic repeat (kbp)	303.4	304.0	452.4	291.4	282.4	371.3	318.5
Gene loci	118	118	125	110	96	134	115
Defense-related genes	42	42	47	40	38	45	40
LRR-RLP	18	18	24	16	15	21	17
Peroxidase	1	1	1	2	1	1	1
MPL1-like protein	2	2	1	1	2	2	1
Calcium-binding EF-hand family protein gene	1	1	1	1	0	1	1

Defense-related genes with different numbers across the haplotypes are depicted in this table. Number of defense-related genes composing each haplotype can be retrieved in [Supplementary Table 11](#). Gene content of each haplotype and their defense-related category and subcategory is detailed in [Supplementary Tables 12–18](#). Abbreviations: Hap, Haplotype; LRR-RLP, leucine-rich repeat receptor-like proteins; MPL1, *Myzus persicae*-induced lipase 1

and b42-26 (PdR1⁻/PdR1⁻) (969.7 and 848.1 kbp on Haplotype 1 and 2, respectively). In contrast, the two PdR1 haplotypes of b43-17 (PdR1a⁺/PdR1b⁺) differed by only 1.7 kbp (889.4 and 887.7 kbp on Haplotype 1 and 2, respectively), and the two haplotypes of b46-43 (PdR1e⁺/PdR1f⁺) were identical except for a single SNP located in a repetitive region at 472,149 bp after VMCNg2b7.2. Because our aim is to investigate the sequence diversity among PdR1 haplotypes, only Haplotype 1 of b46-43 was retained for the subsequent analyses and named PdR1e.

Because 07744-094 (PdR1c⁺/PdR1⁻) is a backcross to *V. vinifera*, one haplotype derives from *V. arizonica* (carrying PdR1c) and the other from *V. vinifera* (PdR1⁻). Accordingly, we used short DNA-seq reads of b40-14 (PdR1c⁺/PdR1d⁺) to distinguish PdR1c from its homologous *V. vinifera* PdR1⁻ haplotype. Normalized median base coverage (per 10 kbp) was evaluated on both haplotypes of 07744-094 using only reads that aligned perfectly (i.e., without any mismatch). This strict criterion was applied to prevent calculating base coverage from reads carrying SNPs. Short DNA-seq reads of b40-14 covered the entire region at 1.1 ± 0.3x on the Haplotype 2 of chromosome 14, whereas little to no coverage was observed on Haplotype 1 ([Fig. 1a](#)). This indicates that PdR1c is located on Haplotype 2 of 07744-094 genome, while Haplotype 1 corresponds to the *V. vinifera* PdR1⁻ haplotype. Interestingly, coverage on Haplotype 2 between markers Ch14-77 and Ch14-02 (26.8–27.01 Mbp) averaged 1.5 ± 0.2x, suggesting that this region is quite homozygous between PdR1c and PdR1d.

Conversely, we assessed the normalized median base coverage of five PD-resistant BC₄ and BC₅ cultivars carrying only PdR1b ([Morales-Cruz et al. 2023](#)) to distinguish PdR1b from PdR1a in b43-17 ([Fig. 1b](#); [Supplementary Figs. 2 and 3](#)). The average coverage across the five PdR1b⁺ cultivars was 0.4 ± 0.1x between markers

VMCNg2b7.2 and Ch14-02 on both haplotypes of the b43-17 genome, suggesting that the region is quite homozygous in PdR1a and PdR1b. In the 3' portion of the locus, between the markers Ch14-02 and UDV095, Haplotype 1 showed little to no coverage, whereas 0.7 ± 0.1x coverage was found on Haplotype 2. This suggests that the five PD-resistant cultivars likely possess this part of the locus. In addition, a short region between markers VMCNg2b7.2 and Ch14-70 (from 25.2 to 25.21 Mbp) was uncovered in Haplotype 2, while a coverage peak (~0.8x) was observed at 26.17 Mbp on Haplotype 1. The localized differences in coverage between haplotypes could reflect either recombination or assembly-related haplotype switches. Remarkably, small polymorphisms (< 50 bp) between the two haplotypes of b43-17 were mostly located in the same two regions ([Fig. 1b](#)). Furthermore, no sign of haplotype switch in these two regions could be observed on the alignment of the HiFi reads of b43-17 on the genome ([Supplementary Figs. 4 and 5](#)). Altogether, these results indicate that a recombination event occurred in the 3' of the locus during the breeding of the five PD-resistant cultivars, and that PdR1a and PdR1b are located on Haplotype 2 and Haplotype 1 of chromosome 14 of b43-17, respectively.

PdR1 repeat and gene content

To determine the origin of PdR1 haplotype size variation, we annotated the repeat and gene content of the four genomes. Annotation of repetitive elements showed that the total length of intergenic repeats in PdR1 haplotypes ranged from 452.4 kb in the *V. vinifera* PdR1⁻ to 303.4 kb in PdR1b of b43-17 ([Table 3](#), [Supplementary Table 10](#)). These data indicate that PdR1 haplotype size differences are primarily driven by variation in intergenic repeat content.

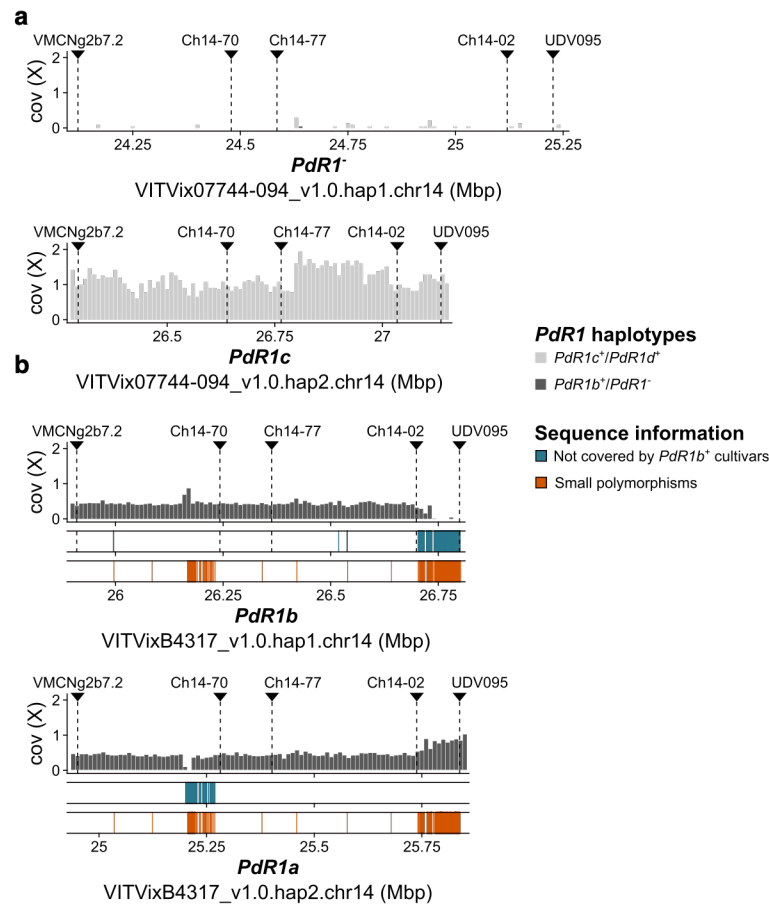


Fig. 1. PdR1 haplotyping of 07744-094 (PdR1c⁺/PdR1⁻) and b43-17 (PdR1a⁺/PdR1b⁺). a) Normalized median DNA-seq coverage per 10 kbp of b40-14 (PdR1c⁺/PdR1d⁺) on the diploid genome of 07744-094 (PdR1c⁺/PdR1⁻). b) Average of the normalized median DNA-seq coverage per 10 kbp of five PdR1b⁺ cultivars on the genome of b43-17 (PdR1a⁺/PdR1b⁺). Bases not covered by any PdR1b⁺ cultivar, and small polymorphisms between the two haplotypes of b43-17 (PdR1a⁺/PdR1b⁺) are depicted for each haplotype. Chromosomal position of the PdR1-associated genetic markers is indicated by black triangles and dashed lines.

The number of annotated gene loci ranged from 96 in PdR1e to 134 in the PdR1⁻ of b42-26 Haplotype 1. Based on protein domain composition and homology to *A. thaliana* proteins, we identified genes potentially involved in plant responses to biotic stress in each haplotype (Supplementary Tables 11–18). Defense-related genes within PdR1 haplotypes included those encoding extracellular receptors and transcription factors, as well as genes involved in reactive oxygen species (ROS) production and scavenging, cell wall modification, programmed cell death, ethylene- and jasmonic acid-mediated signaling, and genes with putative roles in biotic stress responses of unknown mechanism (Supplementary Table 11). The number of defense-related genes ranged from 38 in PdR1e to 47 in the *V. vinifera* PdR1⁻ (Table 3; Supplementary Table 11).

Differences in defense-related gene content among PdR1 haplotypes were mainly driven by variation in genes encoding LRR-RLPs between markers Ch14-77 and Ch14-02, which ranged from 16 to 24 genes. In addition, PdR1c contained two peroxidase-coding genes,

compared with a single gene in the other haplotypes. Two genes encoding homologs of *Myzus persicae*-induced lipase 1 (MPL1) from *A. thaliana* were present in PdR1a, PdR1b, PdR1e, and the PdR1⁻ Haplotype 1 of b42-26, whereas only one MPL1 homolog was detected in the remaining haplotypes. In *A. thaliana*, MPL1 restricts *Fusarium graminearum* infection by limiting jasmonic acid accumulation (Alam et al. 2022). A gene encoding a calcium-binding EF-hand family protein was present in all haplotypes except PdR1e.

Altogether, variation in intergenic repeat and gene composition, particularly among defense-related genes, provides clear evidence of sequence diversity between PdR1 and PdR1⁻ haplotypes, as well as among the PdR1 haplotypes themselves.

PdR1 sequence graph

To further evaluate sequence diversity among the PdR1 haplotypes, we constructed a sequence graph including both PdR1 and PdR1⁻ haplotypes. This graph comprised 92,807 nodes spanning

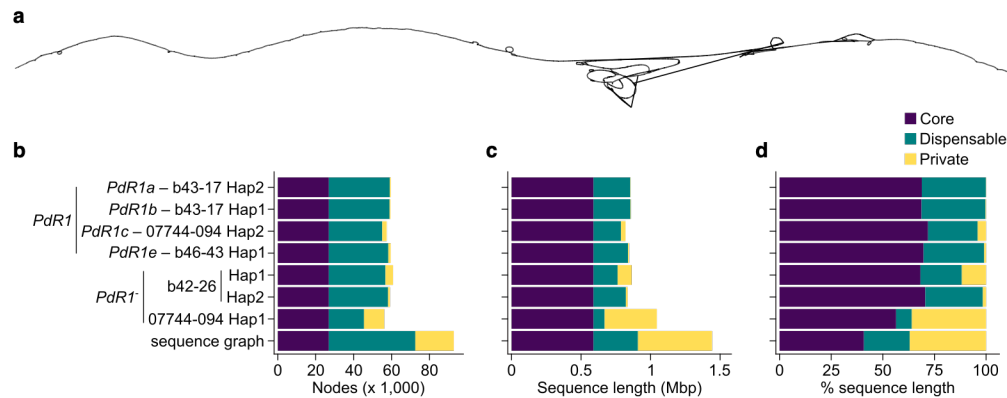


Fig. 2. Sequence graph of the PdR1 and PdR1⁻ haplotypes. a) 2D visualization of the sequence graph built using four PdR1 haplotypes: PdR1a and PdR1b of b43-17, PdR1c from 07744-094 (PdR1c⁺/PdR1⁻), PdR1e of b46-43, and three PdR1⁻ haplotypes: the two haplotypes of b42-26 (PdR1⁻/PdR1⁻), and the *V. vinifera* haplotype of 07744-094. Number of nodes (b), cumulative sequence length (c), and percentage of sequence length (d), made of core, dispensable, and private nodes in each haplotype path. Panels b–d share the same legend. Abbreviations: Hap, Haplotype.

1,442,276 bp. An extensive and complex bubble located approximately three-quarters along the graph indicated divergent paths consistent with large structural variation in this region (Fig. 2a). Several smaller bubbles elsewhere in the graph reflected finer-scale polymorphisms, indicating substantial structural divergence concentrated in a specific region of the locus, together with additional localized variation among haplotypes.

To characterize this variation, we classified graph nodes into three categories: core nodes (shared across all seven haplotypes), dispensable nodes (present in two to six haplotypes), and private nodes (specific to a single haplotype) (Supplementary Fig. 6). Nearly half of the nodes were dispensable (49.1%), whereas core and private nodes accounted for 29.0 and 21.9%, respectively (Fig. 2b). Private nodes were most abundant in the *V. vinifera* PdR1⁻ haplotype (10,711 nodes), and, after core nodes, represented the largest contribution to node content, followed by nodes shared among *V. arizonica* and its hybrids (Supplementary Fig. 4). This pattern indicates marked sequence divergence between *V. vinifera* and *V. arizonica*-derived haplotypes in this region. Among PdR1 haplotypes, PdR1c contained the greatest number of private nodes (2,334), followed by PdR1e (1,253), and the two haplotypes of b43-17 (476 and 550). Although fewer than dispensable nodes, private nodes accounted for a greater proportion of total graph length (37% vs 22.2%) (Fig. 2c and d), indicating that private nodes tend to be longer than shared nodes.

Comparison of PdR1 and PdR1⁻ haplotype paths identified 8,562 PdR1-specific nodes, defined as nodes present in at least one PdR1 haplotype and absent from all PdR1⁻ haplotypes, representing 9.2% of all graph nodes. All four PdR1 haplotype paths contained PdR1-specific nodes (Fig. 3a), ranging from 3,139 in PdR1e to 4,295 in PdR1b. The two b43-17 haplotypes contained the highest numbers of PdR1-specific nodes (3,479 and 3,505 in PdR1a and PdR1b, respectively). Most PdR1-specific nodes were shared among multiple haplotype paths, particularly between the two b43-17 haplotypes and PdR1e. In contrast, PdR1c exhibited a distinct pattern, with 71.8% of its PdR1-specific nodes being private, accounting for the largest cumulative private-node length (33.1 kbp) (Fig. 3b).

Among PdR1-specific nodes, 212 overlapped protein-coding sequences (CDSs) of defense-related genes, corresponding to 2.5% of all PdR1-specific nodes. These nodes were detected across all four

PdR1 haplotype paths, with the highest numbers in PdR1b and PdR1c (90 and 89, respectively), followed by PdR1a (82) and PdR1e (58) (Fig. 3c). Overall, these nodes overlapped approximately half of the defense-related gene content of each PdR1 haplotype, ranging from 52.4 to 53.7%. They affected all defense-related gene categories, albeit in different proportions among haplotype paths (Supplementary Table 19). Only eight nodes in defense-related CDSs were shared across all four PdR1 haplotypes, six of which overlapped genes involved in cell wall modification (Fig. 3c, Supplementary Table 19).

Most PdR1-specific, defense-related nodes were private to PdR1c (86.5%), whereas nodes in the two b43-17 haplotypes were largely shared between each other and with PdR1e. In PdR1c, private nodes predominantly overlapped CDSs of LRR-RLP genes (45 nodes) and genes involved in ROS production (16 nodes) (Supplementary Table 19). PdR1a and PdR1b shared nodes overlapping genes encoding extracellular receptors (LRR-RLPs (14) and LRR-RLK (2)), the metal-nicotianamine transporter YSL6-like protein (9), and cell wall modification genes (7). In contrast, PdR1e shared 24 PdR1-specific, defense-related nodes with one or two other PdR1 haplotypes, mostly with the b43-17 haplotypes, and contained 26 private nodes, which overlapped genes associated with cell death and genes with putative roles in biotic stress responses of unknown mechanism.

In summary, the sequence graph analysis revealed that each PdR1 haplotype path contains PdR1-specific nodes, but with distinct distributions, lengths, and functional enrichments, affecting different categories of defense-related genes.

PdR1c candidate genes

Because the assembly of 07744-094 (PdR1c⁺/PdR1⁻) enabled unambiguous phasing of PdR1c and the sequence graph revealed PdR1-specific features of the region, we investigated the PdR1c gene content to identify candidate genes associated with PD resistance. We previously refined the boundaries of the PdR1 region in *V. arizonica* b40-14 between the markers Ch14-77 and Ch14-02 (Morales-Cruz et al. 2023). In PdR1c, these markers delimited a region spanning 26,764,799–27,034,556 bp on chromosome 14 of Haplotype 2 of 07744-094. This refined region encompassed 39 genes, including 18 defense-related genes encoding a chloride channel protein,

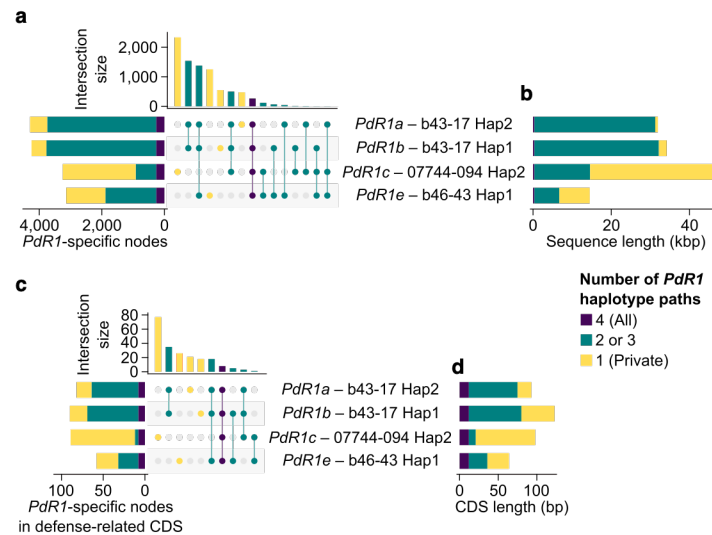


Fig. 3. Nodes specific to the paths of the PdR1 haplotypes in the sequence graph. a) Number of PdR1-specific nodes (i.e. found only in the paths of the PdR1 haplotypes) in each PdR1 haplotype path, and their distribution among the paths. b) Cumulative sequence length of the PdR1-specific nodes in each haplotype path. Number of PdR1-specific nodes overlapping the protein-coding sequence (CDS) of the genes potentially involved in biotic stress response, their repartition among the paths (c), and their cumulative length (d). Panels a–d share the same legend. Abbreviations: Hap, Haplotype.

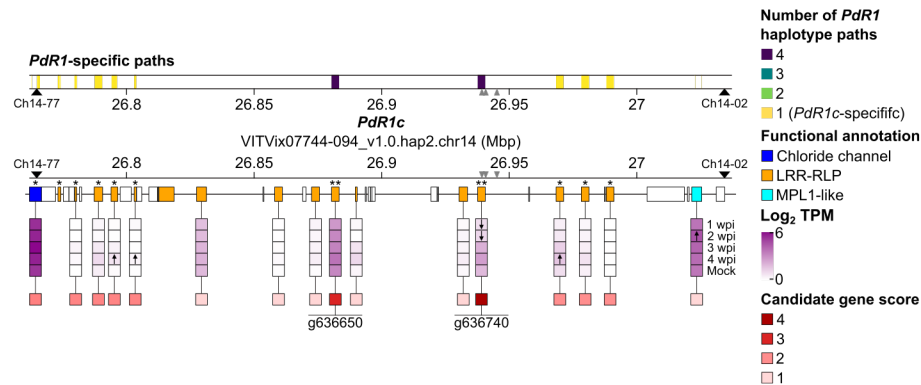


Fig. 4. Narrowing candidate PdR1c genes using sequence graph information and gene expression. PdR1-specific paths in the defense-related genes of the refined PdR1c region. Gene content of the PdR1c locus with the defense-related genes colored according to their functional annotation. Genes with a protein sequence specific to PdR1c are indicated with *, genes with a protein sequence specific to all PdR1 haplotypes are depicted with **. Expression of the defense-related genes in stems of 3 PD-resistant 07744 lines at 1, 2, 3, 4 weeks post Xf inoculation (wpi) and stems mock-inoculated (Mock). Gene expression is represented as the log₂-transformed mean Transcripts per Million (TPM) for the genes expressed in at least one condition (TPM > 0). Genes more highly and lower expressed are indicated by † and ‡, respectively. Score of the defense-related genes of the refined PdR1c region as candidate genes. Scoring was made based on the position of the genes regarding the PD resistance-associated kmers, the specificity of their protein sequence to all PdR1 haplotypes and PdR1c, and gene expression. Black triangles represent the position of the PdR1-associated genetic markers. Gray triangles indicate the positions of the PD resistance-associated kmers.

16 LRR-RLPs, and an MPL1-like protein (Fig. 4, Supplementary Table 14). Sequence graph analysis identified 709 PdR1-specific nodes within this interval, 97% of which were private to PdR1c (Supplementary Table 20).

To assess the functional impact of PdR1-specific variation, we compared CDS paths of defense-related genes in PdR1c with those of the three other PdR1 and three PdR1⁻ haplotypes (Fig. 4).

PdR1-specific CDS paths were detected in 12 defense-related genes, encoding the chloride channel protein, 10 LRR-RLPs, and the MPL1-like protein. All PdR1-specific CDS paths were unique to PdR1c, except for two monoexonic LRR-RLP genes (g636650 and g636740) (Supplementary Table 21). Comparison of CDS paths and predicted protein sequences showed that nine genes encoded PdR1c-specific proteins, including the chloride channel protein

gene and eight LRR-RLP genes, whereas g636650 and g636740 encoded identical protein sequences across all *PdR1* haplotypes. Notably, g636740 overlapped PD resistance-associated kmers, previously identified using 167 *V. arizonica* accessions and hybrids, providing independent support for its association with PD resistance (Morales-Cruz et al. 2023). Forty-six kmers aligned perfectly (Supplementary Table 22), spanning 26,939,281–26,945,206 bp within the Ch14-77–Ch14-02 interval, which contains only the CDS of g636740.

To further narrow down candidate genes, we profiled the expression of defense-related genes in *Xf*-inoculated stems from three PD-resistant 07744 backcross lines, including 07744-094, at 1, 2, 3, and 4 weeks post-inoculation, as well as in mock-inoculated controls. Sixteen of the 18 defense-related genes were expressed (TPM > 0) in at least one condition (Fig. 4). These included the chloride channel protein, 14 LRR-RLPs, and the MPL1-like gene. Only four genes showed differential expression in response to *Xf*, comprising four LRR-RLP genes and the MPL1-like gene.

Based on these results, we scored defense-related genes within the refined *PdR1c* region as candidate PD resistance genes using three criteria: (i) proximity to PD resistance-associated kmers, (ii) protein sequence specificity to all *PdR1* haplotypes and/or *PdR1c*, and (iii) constitutive expression and/or transcriptional response to *Xf* infection (Fig. 4; Supplementary Table 23). The LRR-RLP gene g636740 had the highest score (4), followed by the LRR-RLP gene g636650 (3). Eight genes encoding the chloride channel protein and 7 LRR-RLPs had a score of 2, while 6 genes coding for the MPL1-like and 5 LRR-RLPs had a score of 1. Accordingly, LRR-RLP genes g636740 and g636650 are the strongest *PdR1c* candidates.

Discussion

Based on bacterial titer after inoculation with *Xf*, *V. arizonica* b40-14 have similar level of PD resistance to the two *V. arizonica* hybrids b43-17 and b46-43 (Huerta-Acosta et al. 2023; Morales-Cruz et al. 2023). However, the presence of multiple marker-defined haplotypes at *PdR1* in *V. arizonica* and its hybrids (*PdR1a-f*) suggests that PD resistance may involve different genes or allelic variants within the locus. To answer this hypothesis, we evaluated the sequence diversity among the five *PdR1* haplotypes (*PdR1a,b,c,e,f*) by first assembling and haplotyping the diploid genome of b43-17 (*PdR1a⁺/PdR1b⁺*), the backcross 07744-094 (*PdR1c⁺/PdR1⁻*), and b46-43 (*PdR1e⁺/PdR1f⁺*), and an accession which PD resistance is not associated with *PdR1*, b42-26 (*PdR1⁻/PdR1⁻*) (Table 1) (Riaz et al. 2018). Genome assembly using HiFi reads allowed to generate all *PdR1* and *PdR1⁻* haplotypes in a single contig, which is an improvement for *PdR1c* compared to two fragmented loci in the genome assembly of *V. arizonica* b40-14 (*PdR1c⁺/PdR1d⁺*) (Supplementary Fig. 1). Localization of the *PdR1* region through the *in silico* amplification of *PdR1*-associated genetic markers showed differences in amplicon size and locus length between accessions and within each accession, except in b46-43, where both haplotypes were identical, with the exception of 1 SNP. Differences were also found in the gene and repeat contents of the *PdR1* haplotypes, especially in the number of LRR-RLP genes. All these differences provide clear evidence of sequence diversity between *PdR1* and *PdR1⁻* haplotypes, as well as among the *PdR1* haplotypes themselves.

Sequence diversity was further investigated through the construction of a sequence graph including 7 haplotypes. These consisted of 4 *PdR1* haplotypes: *PdR1a*, *PdR1b*, *PdR1c*, and *PdR1e*, and 3 *PdR1⁻* haplotypes: one from *V. vinifera* and two from the

PD-resistant accession b42-26. The 2D representation of the sequence graph (Fig. 2a) showed an extensive structural divergence located approximately three-quarters along the graph, together with additional localized structural variations among haplotypes. Categorization of the graph nodes as core, dispensable, and private determined a substantial sequence divergence between *V. vinifera* and *V. arizonica*-derived haplotypes in the locus (Fig. 2b–d). Interspecific sequence diversity has also been observed in other grape disease resistance loci, such as the powdery mildew-resistance loci *Run1.2* and *Run2.2* (Cochetel et al. 2021), *Ren6* and *Ren7* (Massonnet et al. 2025), and the downy mildew-resistance loci *Rpv3* (Wilkerson et al. 2025). Using the inter- and intra-specific sequence diversity between the 4 *PdR1* haplotypes and the *V. vinifera PdR1⁻* haplotype of 07744-094 and the two haplotypes of b42-26, we determined the nodes specific to the *PdR1* haplotypes (i.e. found in at least one *PdR1* haplotype but absent in all *PdR1⁻* haplotypes). Investigating the *PdR1*-specific nodes showed a diversity in sequence between the four *PdR1* haplotypes (Fig. 3), while examining their overlap with defense-related genes showed that they affected different defense-related gene categories. *PdR1a* and *PdR1b* were found to share most of their *PdR1*-specific features, whereas *PdR1c* contained the highest number of private nodes, followed by *PdR1e*. Normalized median DNA-seq coverage and location of small polymorphisms in the five *PdR1b⁺* cultivars compared to b43-17's *PdR1a* and *PdR1b* indicated a high level of homozygosity between the two haplotypes for nearly three-quarters of the locus (Fig. 1b). This corroborates the great number of *PdR1*-specific features shared between *PdR1a* and *PdR1b*. On the other hand, *PdR1a* and *PdR1b* also shared a great sequence specificity compared to *PdR1c* and *PdR1e*. Shared nodes by *PdR1a* and *PdR1b* (1,540) and their respective private nodes (476 and 550, respectively) represent a total of 2,016 and 2,090 nodes, respectively (Fig. 3a), which is slightly lower than the number of private nodes in *PdR1c* (2,334) but higher than *PdR1e* (1,253). The three *PdR1⁺* accessions (b40-14, b43-17, and b46-43) were collected from three different collection sites in Texas and Mexico (Riaz et al. 2018), which could explain the sequence diversity among their *PdR1* haplotypes.

Finally, we demonstrated the utility of the sequence graph to narrow down candidate genes in the refined region of *PdR1c*, especially when coupled with gene expression. Scoring of the expressed defense-related genes within the refined *PdR1c* region based on proximity to PD resistance-associated kmers, protein sequence specificity, and gene expression, pointed two strong *PdR1c* candidates: the LRR-RLP genes g636740 and g636650 (Fig. 4). Because these two LRR-RLP genes are present in all the four *PdR1* haplotypes, this suggests that the same R gene(s) are potentially involved in PD resistance in all the three *PdR1⁺* accessions. To answer this, it would be necessary to profile their gene expression in petioles or stems at both the constitutive level and in response to *Xf*. Monitoring the gene expression of the defense-related genes among the 3 other *PdR1* haplotypes would also help identify any additional candidate R gene(s). Two LRR-RLP genes were also proposed as candidate genes in *V. arizonica* b40-14 (Morales-Cruz et al. 2023). These genes correspond to g636600 and g636650 in the 07744-094 genome. Based on the results of our sequence graph, g636650 is a very strong candidate as its protein sequence is identical for all the *PdR1* haplotypes and absent in all *PdR1⁻* haplotypes. However, the protein sequence of g636600 was also found present in b42-26, decreasing its score as a candidate gene. Gene expression profiling of *Xf*-infected tissues of b42-26 could help to determine whether

g636600 is still a strong candidate. Regarding g636740, the gene was absent in the genome of *V. arizonica* b40-14. In the b40-14 genome, the refined PdR1 region was made of 2 and 4 contigs on Haplotype 1 and 2, respectively (Morales-Cruz et al. 2023)(Supplementary Fig. 1). The fragmentation of the locus region and the absence of haplotyping information for the genome assembly are likely at the origin of the absence of g636740 in the genome of *V. arizonica* b40-14.

Future work will focus on the functional validation of the PdR1c candidate genes, the incorporation of additional PdR1 haplotypes in the sequence graph, and the profiling of the gene expression of the defense-related genes in several PdR1 haplotypes, at both the constitutive level and in response to Xf. This would help to determine whether the PD resistance associated with the four PdR1 haplotypes is based on the expression of the same gene(s). As a protocol for stable protoplast-based gene editing has been developed for *V. arizonica* (Tricoli and Debernardi 2024), stable gene editing of the PdR1c candidate genes in *V. arizonica* b40-14 (PdR1c+/PdR1d+) could be an option. However, the lack of sequence information about PdR1d might complicate the process. Stable expression of the candidate genes in *V. vinifera* could be another option to assess the role of these genes in PD resistance. On the other hand, using transient transformation methods, such as *Agrobacterium*-mediated expression (Zhang et al. 2024), virus-induced gene silencing (Yang et al. 2022), and exogenous dsRNA-based RNAi (Marciano et al. 2021), would likely be more challenging to evaluate the function of the candidate genes as Xf is a xylem-restricted pathogen.

From a breeding perspective, the comparison of the PdR1 and PdR1⁺ haplotypes through the sequence graph could greatly help the development of precise markers, either specific to all PdR1 haplotypes or specific to each single haplotype. Such markers could accelerate the screening in traditional breeding programs or the incorporation of the genes responsible for PD resistance in elite cultivars using gene editing (Greenwood et al. 2023).

Data availability

HiFi reads are accessible through NCBI under the BioProject PRJNA1380937. Genome sequences and their annotation files are available at Zenodo doi: 10.5281/zenodo.17925270. [Supplemental Material](https://doi.org/10.25387/g3.31864987) available at figshare: <https://doi.org/10.25387/g3.31864987>.

Acknowledgments

We thank the UC Davis DNA Technologies Core for sequencing assistance.

Funding

This work was funded by the grant #22-0555-000-SA from the Pierce's Disease/Glassy-Winged Sharpshooter Board of the California Department of Food and Agriculture, and partially funded by the Ray Rossi Endowment.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Author contributions

Conceptualization: D.C. Formal analysis: M.M., M.Z., N.C. Investigation: M.M., M.Z., N.C., R.F.-B. Resources: S.R.

Writing-original draft preparation: M.M., D.C. Writing-review and editing: M.M., D.C. Visualization: M.M. Project administration: D.C. Funding acquisition: D.C., S.R. All authors read and agreed to the published version of the manuscript.

Literature cited

- Alam ST, et al. 2022. Opposing effects of MYZUS PERSICAE-INDUCED LIPASE 1 and jasmonic acid influence the outcome of *Arabidopsis thaliana*-fusarium graminearum interaction. *Mol Plant Pathol.* 23:1141–1153. <https://doi.org/10.1111/mpp.13216>.
- Barnett DW, et al. 2011. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics.* 27:1691–1692. <https://doi.org/10.1093/bioinformatics/btr174>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Camacho C, et al. 2009. BLAST+: architecture and applications. *BMC Bioinformatics.* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
- Castro C, DiSalvo B, Roper MC. 2021. *Xylella fastidiosa*: a reemerging plant pathogen that threatens crops globally. *PLoS Pathog.* 17: e1009813. <https://doi.org/10.1371/journal.ppat.1009813>.
- Cheng H et al. 2021. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods.* 18: 170–175. <https://doi.org/10.1038/s41592-020-01056-5>.
- Chin C-S, et al. 2016. Phased diploid genome assembly with single molecule real-time sequencing. *Nat Methods.* 13:1050–1054. <https://doi.org/10.1038/nmeth.4035>.
- Cochetel N, et al. 2021. Diploid chromosome-scale assembly of the *Muscadinia rotundifolia* genome supports chromosome fusion and disease resistance gene expansion during Vitis and *Muscadinia* divergence. *G3 (Bethesda).* 11:jkab033. <https://doi.org/10.1093/g3journal/jkab033>.
- Cochetel N, et al. 2023. A super-pangenome of the North American wild grape species. *Genome Biol.* 24:290. <https://doi.org/10.1186/s13059-023-03133-2>.
- Cochetel N, et al. 2025. Phased epigenomics and methylation inheritance in a historical *Vitis vinifera* hybrid. *Genome Biol.* 26:392. <https://doi.org/10.1186/s13059-025-03858-2>.
- Di Tommaso P, et al. 2017. Nextflow enables reproducible computational workflows. *Nat Biotechnol.* 35:316–319. <https://doi.org/10.1038/nbt.3820>.
- Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput Biol.* 7:e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
- El-Gebali S, et al. 2019. The Pfam protein families database in 2019. *Nucleic Acids Res.* 47:D427–D432. <https://doi.org/10.1093/nar/gky995>.
- Greenwood JR, Zhang X, Rathjen JP. 2023. Precision genome editing of crops for improved disease resistance. *Curr Biol.* 33:R650–R657. <https://doi.org/10.1016/j.cub.2023.04.058>.
- Guarracino A, et al. 2022. ODGI: understanding pangenome graphs. *Bioinformatics.* 38:3319–3326. <https://doi.org/10.1093/bioinformatics/btac308>.
- Haas BJ, et al. 2008. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the program to assemble spliced alignments. *Genome Biol.* 9:R7. <https://doi.org/10.1186/gb-2008-9-1-r7>.
- Haviland DR, Stone-Smith B, Gonzalez M. 2021. Control of Pierce's disease through areawide management of glassy-winged sharpshooter (Hemiptera: Cicadellidae) and roguing of infected grapevines. *J Integr Pest Manag.* 12:1–10. <https://doi.org/10.1093/jipm/pmab008>.

- Heumos S, et al. 2024. Cluster-efficient pangenome graph construction with nf-core/pangenome. *Bioinformatics*. 40:btac609. <https://doi.org/10.1093/bioinformatics/btac609>.
- Huerta-Acosta KG, Riaz S, Tenschler A, Walker MA. 2023. Genetic characterization of Pierce's disease resistance in a *Vitis arizonica*/monticola wild grapevine. *Am J Enol Vitic*. 74:0740003. <https://doi.org/10.5344/ajev.2022.22021>.
- Jaillon O, et al. 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*. 449: 463–467. <https://doi.org/10.1038/nature06148>.
- Krivanek AF, Riaz S, Walker MA. 2006. Identification and molecular mapping of PdR1, a primary resistance gene to Pierce's disease in *Vitis*. *Theor Appl Genet*. 112:1125–1131. <https://doi.org/10.1007/s00122-006-0214-5>.
- Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes1. *J Mol Biol*. 305:567–580. <https://doi.org/10.1006/jmbi.2000.4315>.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Lomsadze A, Ter-Hovhannisyan V, Chernoff YO, Borodovsky M. 2005. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res*. 33:6494–6506. <https://doi.org/10.1093/nar/gki937>.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Marçais G, et al. 2018. MUMmer4: a fast and versatile genome alignment system. *PLoS Comput Biol*. 14:e1005944. <https://doi.org/10.1371/journal.pcbi.1005944>.
- Marciano D, et al. 2021. RNAi of a putative grapevine susceptibility gene as a possible downy mildew control strategy. *Front Plant Sci*. 12:667319. <https://doi.org/10.3389/fpls.2021.667319>.
- Massonnet M, et al. 2025. Dissection of the *Ren6* and *Ren7* powdery mildew resistance loci in *Vitis piasezkii* DVIT2027 using phased parental-progeny genomes and intraspecific locus graph reconstruction. *G3 (Bethesda)*. 15:jkaf250. <https://doi.org/10.1093/g3journal/jkaf250>.
- Micheltmore RW, Christopoulou M, Caldwell KS. 2013. Impacts of resistance gene genetics, function, and evolution on a durable future. *Annu Rev Phytopathol*. 51:291–319. <https://doi.org/10.1146/annurev-phyto-082712-102334>.
- Minio A, et al. 2019. Iso-seq allows genome-independent transcriptome profiling of grape berry development. *G3 (Bethesda)*. 9: 755–767. <https://doi.org/10.1534/g3.118.201008>.
- Minio A, et al. 2022. Assembly of complete diploid-phased chromosomes from draft genome sequences. *G3 (Bethesda)*. 12:jkac143. <https://doi.org/10.1093/g3journal/jkac143>.
- Morales-Cruz A, et al. 2023. Multigenic resistance to *Xylella fastidiosa* in wild grapes (*Vitis* spp.) and its implications within a changing climate. *Commun Biol*. 6:580. <https://doi.org/10.1038/s42003-023-04938-4>.
- Patro R, et al. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 14:417–419. <https://doi.org/10.1038/nmeth.4197>.
- Quinlan AR. 2014. BEDTools: the Swiss-army tool for genome feature analysis. *Curr Protoc Bioinformatics*. 47:11.12.1–11.12.34. <https://doi.org/10.1002/0471250953.bi1112s47>.
- Rapicavoli J, et al. 2018. *Xylella fastidiosa*: an examination of a re-emerging plant pathogen. *Mol Plant Pathol*. 19:786–800. <https://doi.org/10.1111/mpp.12585>.
- Riaz S, et al. 2008. Fine-scale genetic mapping of two Pierce's disease resistance loci and a major segregation distortion region on chromosome 14 of grape. *Theor Appl Genet*. 117:671–681. <https://doi.org/10.1007/s00122-008-0802-7>.
- Riaz S, Huerta-Acosta K, Tenschler AC, Walker MA. 2018. Genetic characterization of *Vitis* germplasm collected from the southwestern US and Mexico to expedite Pierce's disease-resistance breeding. *Theor Appl Genet*. 131:1589–1602. <https://doi.org/10.1007/s00122-018-3100-z>.
- Riaz S, Krivanek AF, Xu K, Walker MA. 2006. Refined mapping of the Pierce's disease resistance locus, PdR1, and sex on an extended genetic map of *Vitis rupestris* x *V. arizonica*. *Theor Appl Genet*. 113:1317–1329. <https://doi.org/10.1007/s00122-006-0385-0>.
- Shi X, et al. 2023. The complete reference genome for grapevine (*Vitis vinifera* L.) genetics and breeding. *Hortic Res*. 10:uhad061. <https://doi.org/10.1093/hr/uhad061>.
- Smit A, Hubley R, Green P. 2013. <http://www.repeatmasker.org>.
- Soneson C, Love MI, Robinson MD. 2016. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res*. 4:1521. <https://doi.org/10.12688/f1000research.7563.2>.
- Stanke M, et al. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res*. 34:W435–W439. <https://doi.org/10.1093/nar/gkl200>.
- Sun Q, Sun Y, Walker MA, Labavitch JM. 2013. Vascular occlusions in grapevines with Pierce's disease make disease symptom development Worse1[OA]. *Plant Physiol*. 161:1529–1541. <https://doi.org/10.1104/pp.112.208157>.
- Tricoli DM, Debernardi JM. 2024. An efficient protoplast-based genome editing protocol for *Vitis* species. *Hortic Res*. 11:uhad266. <https://doi.org/10.1093/hr/uhad266>.
- Tumber K, Alston J, Fuller K. 2014. Pierce's disease costs California \$104 million per year. *CalAg*. 68:20–29. <https://doi.org/10.3733/ca.v068n01p20>.
- Wang M, Kong L. 2019. Pblat: a multithread blat algorithm speeding up aligning sequences to genomes. *BMC Bioinformatics*. 20:28. <https://doi.org/10.1186/s12859-019-2597-8>.
- Waterhouse RM, et al. 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol*. 35: 543–548. <https://doi.org/10.1093/molbev/msx319>.
- Wilkerson D, et al. 2025. Comparative genomics of Rpv3, a multiallelic downy mildew resistance locus in grapevine (*Vitis* sp.). *OENO One*. 59. <https://doi.org/10.20870/oeno-one.2025.59.1.8244>.
- Wu TD, Watanabe CK. 2005. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics*. 21:1859–1875. <https://doi.org/10.1093/bioinformatics/bti310>.
- Yang B, et al. 2022. VvANR silencing promotes expression of VvANS and accumulation of anthocyanin in grape berries. *Protoplasma*. 259:743–753. <https://doi.org/10.1007/s00709-021-01698-y>.
- Zhang Q, et al. 2024. Ectopic and transient expression of VvDIR4 gene in Arabidopsis and grapes enhances resistance to anthracnose via affecting hormone signaling pathways and lignin production. *BMC Genomics*. 25:895. <https://doi.org/10.1186/s12864-024-10830-0>.

Editor: S. Smith

CHAPTER 4

Dual genome editing strategy development to validate Pierce's disease candidate resistance genes

Abstract

Although candidate genes associated with PD resistance have been identified, experimental validation is required before they can be employed for targeted genome engineering. Here, complementary gain- and loss-of-function approaches were applied to begin the functional validation of four PD candidate genes encoding different classes of cell-surface immune receptors. For gain-of-function, overexpression constructs were generated and introduced into the susceptible *Vitis vinifera* Thompson Seedless through *Agrobacterium*-mediated transformation. Transformation resulted in the regeneration of ten putatively transgenic plants for two candidate genes, while constructs carrying the other two genes consistently produced necrotic embryos following transformation. For loss-of-function studies, four single-guide RNAs (sgRNAs) were designed for each candidate gene and assembled into expression constructs for CRISPR/Cas9 editing of protoplasts of the PD resistant *V. arizonica*. Editing efficiency was assessed at an early stage using DNA extracted from putatively edited regenerated microcalli, and varied substantially among sgRNAs, ranging from 1% to nearly 99% depending on the target site. Highly efficient guides predominantly generated large deletions resulting from the simultaneous activity of two closely spaced sgRNAs, while small indels typically associated with Cas9-mediated repair were also observed. Poorly performing guides were characterized mainly by substitutions. Mutation profiles were strongly guide-dependent, highlighting the importance of sgRNA validation before plant regeneration. Together, these results establish a dual experimental platform for functional validation of PD candidate genes in grapevine. The putatively transgenic and edited lines generated in this study provide valuable resources for future disease assays aimed at determining the contribution of these receptors to PD resistance.

Introduction

Pierce's disease (PD) of grapevine is a devastating disease caused by the bacterial pathogen *Xylella fastidiosa* (*Xf*) (Castro *et al.*, 2021). This bacterium is spread plant-to-plant by xylem-feeding insects. Once

the bacterial population establishes itself in the vascular tissues of the plants, nutrient and water uptake are disrupted. This results in leaf scorch, desiccated canes, and shriveled clusters, and consequently vine death (Rapicavoli *et al.*, 2018). Current control strategies consist of pesticides application to control the insect population and removal of infected vines. However, these approaches are not sustainable or effective and genetic resistance is considered the most sustainable way to address this issue.

Cultivars of the European grapevine (*Vitis vinifera* L.), commonly used for wine, table grape, and raisin production, are susceptible to PD. Natural sources of resistance have been identified in several wild *Vitis* spp., especially *V. arizonica* and its natural hybrids (Huerta-Acosta *et al.*, 2023; Riaz *et al.*, 2020). Previous studies identified a major locus associated with PD resistance in *V. arizonica*, called Pierce's disease resistance 1 (*PdRI*) locus, that was used in breeding programs for trait introgression in susceptible cultivars (Krivanek *et al.*, 2006; Walker *et al.*, 2014). The underlying candidate resistance genes, however, were not resolved until later. In Chapter 2, genome-wide association analysis combined with gene expression identified four candidate PD resistance genes: two encoding leucine-rich repeat receptor-like proteins (LRR-RLPs), one encoding an LRR receptor-like kinase (RLK), and one encoding a lysin motif (LysM) RLK (Morales-Cruz *et al.*, 2023). In Chapter 3, sequence-graph approaches combined with gene expression narrowed these to the two LRR-RLP genes as the strongest candidates for PD resistance (Massonnet *et al.*, 2026).

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas genome editing has made it possible to test candidate resistance genes directly (El-Mounadi *et al.*, 2020), and recent advances allow precise insertion of DNA at a defined locus, or knock-in (KI), to speed deployment of resistance genes into elite cultivars (Muzaffar *et al.*, 2026). Experimental validation of these candidates is a necessary step toward establishing their contribution to PD resistance.

To initiate the effort of validating all candidate genes identified in Chapters 2 and 3, an effort that will extend beyond the scope of this dissertation, I implemented a complementary strategy combining gain-of-function and loss-of-function approaches. For gain-of-function studies, candidate genes were overexpressed under the control of the constitutive CaMV35S promoter and introduced into somatic embryos of *V.*

vinifera cv. Thompson Seedless (TS) via *Agrobacterium*-mediated transformation. For loss-of-function studies, single guide RNAs (sgRNAs) targeting each gene of interest (GOI) were designed and assembled in pairs into expression constructs in a customized vector and combined with Cas9 for genome editing of *V. arizonica* protoplasts (Tricoli & Debernardi, 2024), followed by plant regeneration. Together, these complementary strategies establish a framework to experimentally validate resistance genes previously identified within the *PdRI* locus.

Materials and methods

Construct assembly for overexpression in *V. vinifera* Thompson Seedless

The binary vector pGWB402 (Plasmid #74796; Addgene), which carries the CaMV35S promoter and kanamycin resistance within the T-DNA region, was selected for stable expression of candidate resistance genes in *Vitis vinifera* Thompson Seedless (TS). Constructs were designed using SnapGene v7.2.1. For three out of four genes, their coding sequences were cloned into pGWB402 by GENEWIZ by removing the Gateway cassette between the single-cutters XbaI and SacI and inserting the GOI between the CaMV35S promoter and the NOS terminator. Each construct was analyzed to confirm the correct reading frame relative to the 35S promoter and absence of premature stop codons before and after synthesis.

The coding sequence of one of the two LRR-RLP genes could not be synthesized in pGWB402, and it was delivered in pUC19. Cloning was done in house using InFusion® cloning. First, the gene of interest was PCR-amplified using PrimeSTAR Max polymerase (Catalog No. 638945; In-Fusion® Snap Assembly Starter Bundle; Takara Bio) and gene-specific primers (**Table 1**). The PCR reaction was carried out as follows: 30 cycles of 98°C for 10 s, 57°C for 5 s, and 72°C for 50 s. The amplicon was purified using the NucleoSpin Gel and PCR Cleanup Kit. After purification, the concentration was measured at Qubit and 1 µL of the purified product was used to run a gel electrophoresis to confirm correct size. 2 µL were digested with HindIII for further confirmation.

The pGWB402 backbone (11,685 bp) was digested with the restriction enzymes XbaI and SacI to create a linearized vector. The 9,936 bp linear fragment was gel-purified using the Zymoclean Gel DNA Recovery

Kit (Catalog No. D4001; Zymo Research) and concentration was measured at Qubit. Purification of the fragment was confirmed on a 1% TAE gel electrophoresis.

The gene insert and linearized vector were assembled using the In-Fusion® Snap Assembly Master Mix. A linearized pUC19 + 2kb insert was used as a positive control while the linearized pGWB402 was used as a negative control. The cloning reaction was incubated at 50°C for 15 min and chilled on ice prior to transformation into Stellar competent cells. The transformation reaction as well as the controls were plated on selection. Three transformed colonies were picked and screened by colony PCR using primers reported in **Table 1**. Plasmid DNA was extracted from one of the transformed colonies and sequenced.

Agrobacterium tumefaciens engineering with overexpression constructs

For each construct, at least 1µg of purified plasmid DNA was used for the transformation of disarmed *Agrobacterium tumefaciens* (At) EHA105 pCH32 via heat shock. Spectinomycin was used as a selective marker for transformant colonies screening. One colony per construct was screened by colony PCR and amplicon sequencing to verify correct insertion and directionality of the construct. Primer sequences and expected amplicon sizes are listed in **Table 2** and **Table 3**, respectively. After sequencing, one transformed *Agrobacterium* clone harboring the correct plasmid for each gene of interest was used for transformation of somatic embryos of *V. vinifera* Thompson Seedless grapevines.

Table 2. Primer names and sequences used to check for correct insertion of the constructs carrying g238150, g238250 and g238290.

Primer name	Primer sequence (5' – 3')	Gene
CaMV35S	GACGTAAGGGATGACGCACAAT	CaMV35S promoter
M13rev	AGCGGATAACAATTTACACAGG	NOS terminator
RLP50_R1	TAAGCGTAGTTCCAAGAGTGAAGAA	g238150 (LRR-RLP)
RLP50_F2	TCCCTTGTCTATGGCTAAGATTACG	
RLP50_R2	TCCATGTTGTCTTCTTTCCCATTG	
LYSM_R1	GGATTAATTACTGGGATGCTGGTTG	g238250 (LysM-RLK)
LYSM_F2	TAATAAGGCTGTTAGGGACTTGCTT	
LYSM_R2	CTTGATTTTCTCTGGAAGCCAACT	
RLK_R1	CTTAAACCGGTTATTCCTCAAGCTC	g238290 (LRR-RLK)
RLK_F2	GAGATGAACCTGGGTTTGTGACAT	
RLK_R2	GTCACCTCAACTTGTCTATCCTATGC	

Transformation of Thompson Seedless somatic embryos and plant regeneration

Embryogenic cultures of *V. vinifera* cv. Thompson Seedless (TS) were obtained from immature anthers. Initiation, maintenance, transformation and regeneration of TS embryogenic cultures was performed at the UC Davis Plant Transformation Facility (UCDPTF) following the protocol described in (Agüero *et al.*, 2006). Briefly, overnight *Agrobacterium* cultures were diluted at an OD600 of 0.1-0.2 and supplemented with 200 μ M of acetosyringone prior to transformation. TS embryogenic calli were placed on sterile Whatman filter paper in an empty 100x20 mm petri dish. The *Agrobacterium* culture was poured over the plant tissue, and excess bacterial culture was allowed to evaporate before the calli were transferred onto fresh co-cultivation medium. After 48 h in the dark, callus tissue was sub-divided into small clumps and transferred onto selective media containing kanamycin for recovery of transformed tissue. Cells were subcultured on fresh medium every two weeks. Developing secondary embryos were then transferred onto germination medium and subsequently onto regeneration media for plant regeneration. Putatively transformed plants regenerated from secondary embryos were transferred to soil.

sgRNA design and off-target evaluation

Candidate genes coding sequences were analyzed using TMHMM (Krogh *et al.*, 2001) to predict the location of extracellular, transmembrane and intracellular domains. For genes encoding LRR-RLPs, sgRNAs were designed for the LRR motif only. For receptor-like kinases (LysM-RLK and LRR-RLK), sgRNAs were distributed across both extra- and intracellular domains. Guide design was performed using the command-line version of CRISPOR v5.01 (Concordet & Haeussler, 2018) with default parameters. The *V. arizonica* b40-14 v2.0 chromosome-scale genome assembly was used as the reference and the CDS of each candidate gene was used as query. For each gene, CRISPOR generated a summary table containing guide sequences, predicted editing efficiencies, specificity scores, and numbers of predicted off-targets, as well as an off-target table reporting mismatch position and genomic locations. Guide selection was refined manually based on the following criteria. Guides were initially filtered using CFD (cutting frequency

determination) and MIT (Massachusetts Institute of Technology) specificity scores. Guides with CFD value > 90 and MITscore > 40 were retained. Remaining guides were further filtered using Doench2016 efficiency scores, and guides with predicted efficiency > 40 were retained. Guides containing sequences associated with reduced editing efficiency (Graf *et al.*, 2019) were excluded. Filtered sgRNAs that fell outside of the targeted domains were discarded as well. Potential off-target sites were manually inspected. The guides that had a few predicted off-targets that showed mismatches located near the PAM-distal region were retained.

sgRNA expression cassette assembly

To generate the constructs for the CRISPR/Cas9 editing in *V. arizonica*, sgRNAs were cloned in pairs in a custom vector (JD969) developed by our collaborators at the UCDPTF. Multiple sgRNAs are post-transcriptionally processed through the tRNA approach (Xie *et al.*, 2015). Each cassette contained two sgRNA separated by a tRNA sequence. The two resulting plasmids per gene were co-transfected with a plasmid expressing Cas9 to target all four sites simultaneously.

sgRNAs cassettes were amplified using the primers indicated in **Supplementary Table 1** by PCR using AmpliTaq Gold™ DNA polymerase (Thermo Fisher Scientific) and 1 µg of DNA template in a 25 µL reaction volume. Amplification was performed as follows: 95°C x 10 min, 95°C x 15 sec, 60°C x 15 sec, 72°C x 30 sec and 72°C x 5 min. PCR products were then purified using KAPA HyperPure Beads (Roche). Purified PCR products (gRNA cassettes) were cloned into JD969 using Golden Gate assembly. The reaction was performed using T4 ligase, ligase buffer and BsaI-HF v2® enzyme (New England Biolabs) and a 1:1 vector to PCR product ratio for a final reaction volume of 15 µL. The ligation reaction was performed using the following conditions: 37°C x 20 sec, 25 cycles (37°C x 3 min, 16°C x 4 min), 50°C x 5 min, 80°C x 5 min. The assembled plasmids were transformed into chemically competent DH5α *E. coli* cells, and the reaction was plated on LB supplemented with ampicillin and incubated overnight at 37°C. Individual colonies for each plasmid were picked, grown in liquid culture and plasmid DNA was extracted. Plasmid constructs were screened by restriction digestion using HindIII and PstI. The empty vector was added to

the reaction and used as positive control. Digested products were analyzed by agarose gel electrophoresis. Plasmid showing the expected restriction pattern were quantified using Qubit and submitted for sequencing.

V. arizonica protoplasts transfection

V. arizonica protoplasts generation, transformation, culture and whole plant regeneration were performed at the UCDPTF following the protocol described in (Tricoli & Debernardi, 2024) with some modifications. For the PEG-mediated transfection of *V. arizonica* protoplasts, validated JD969 plasmids carrying paired sgRNAs expression cassettes were combined with Cas9 protein (CASPROT-50UG; Sigma) and introduced into *V. arizonica* protoplasts by PEG-mediated transfection.

Amplicon sequencing to detect CRISPR editing activity

To estimate editing efficiency before plant regeneration, genomic DNA was extracted from putatively edited protoplast-derived calli encapsulated in alginate beads. Individual encapsulated protoplasts regenerate into small, isolated callus colonies, allowing for an early assessment of CRISPR/Cas9 editing events. Detection of editing events was assessed by amplicon sequencing following the protocol described in (Zhang, 2022). Briefly, primers flanking each sgRNA target site were designed to produce amplicons in the range of 280-290 bp (**Supplementary Table 2**). If two target sites were close to each other, a single pair of primers was designed to amplify both, otherwise, individual primer pairs were designed for each target. Sequencing was performed using the services provided by MGH CCIB DNA Core (https://dnacore.mgh.harvard.edu/new-cgi-bin/site/pages/crispr_sequencing_main.jsp).

Editing was evaluated for seven independent alginate beads selected according to tissue availability and developmental stage. The number of regenerated callus colonies analyzed varied among beads and target genes.

Sequencing reads were analyzed using the command-line version of CRISPResso2 v2.3.3 (Clement *et al.*, 2019) with the CRISPRessoBatch option and default parameters. Reads were demultiplexed, quality filtered and aligned to the reference amplicon sequence. Samples containing fewer than 200 reads were excluded from downstream analysis. Editing efficiency was calculated as the percentage of modified reads over the

Figure 1. Strategy to create the overexpression constructs for *Agrobacterium* transformation. A) On the left, the map of the pGWB402 binary vector (11,685 bp). The Gateway cassette (1,749 bp) is removed by restriction digestion with single-cutters XbaI and SacI, both highlighted in bold. On the right, one of the genes of interest (g238150) cloned into the linearized pGWB402. B) Shows the InFusion cloning strategy. pGWB402 is linearized by restriction digestion with XbaI and SacI (top). The gene of interest (g238180) is PCR-amplified from pUC19 (bottom). A 15 bp overlap between the fragment and the vector allows for seamless and directional fusion of the two DNA fragments. Figures generated using SnapGene software (<http://www.snapgene.com/products/snapgene/>) (GSL Biotech LLC).

For one of the LRR-RLP genes, cloning in pGWB402 using restriction enzyme-based assembly was not successful and the construct was generated using InFusion cloning, a highly efficient and directional cloning method (Irwin et al., 2012) (**Figure 1B**).

For this, pGWB402 was linearized and purified to obtain a 9,936 bp fragment (**Figure 2A**), and g238180 was PCR-amplified directly from the pUC19 template, yielding the expected 3,018 bp fragment (**Figure 2B**). The purified linearized vector and the PCR-amplified insert were used for InFusion cloning. HindIII digestion of the GOI produced the expected fragments of 2,344 bp and 674 bp, confirming its integrity (**Figure 2C**). The transformation reaction was plated in two aliquots, resulting in an overall efficiency of $\sim 9 \times 10^3$ CFU/ μ g DNA (**Figure 2E**). The positive control (pUC19 + 2kb insert) produced confluent growth (**Figure 2F**), while the negative control (linear pGWB402) yielded only two colonies, indicating low background transformation (**Figure 2G**). Three colonies were screened by colony PCR (**Figure 2D**, **Table 1**), and all of them yielded the expected overlapping amplicons of 1,686 bp and 1,792 bp. The two PCR products covered the full insert and both junction regions, including portions of the 35S and the terminator. This confirmed correct orientation and integration of the insert. Amplicon sequencing further confirmed 100% sequence identity to the expected design.

Overall, four overexpression constructs were generated and introduced into *Agrobacterium tumefaciens* to produce engineered strains for plant transformation.

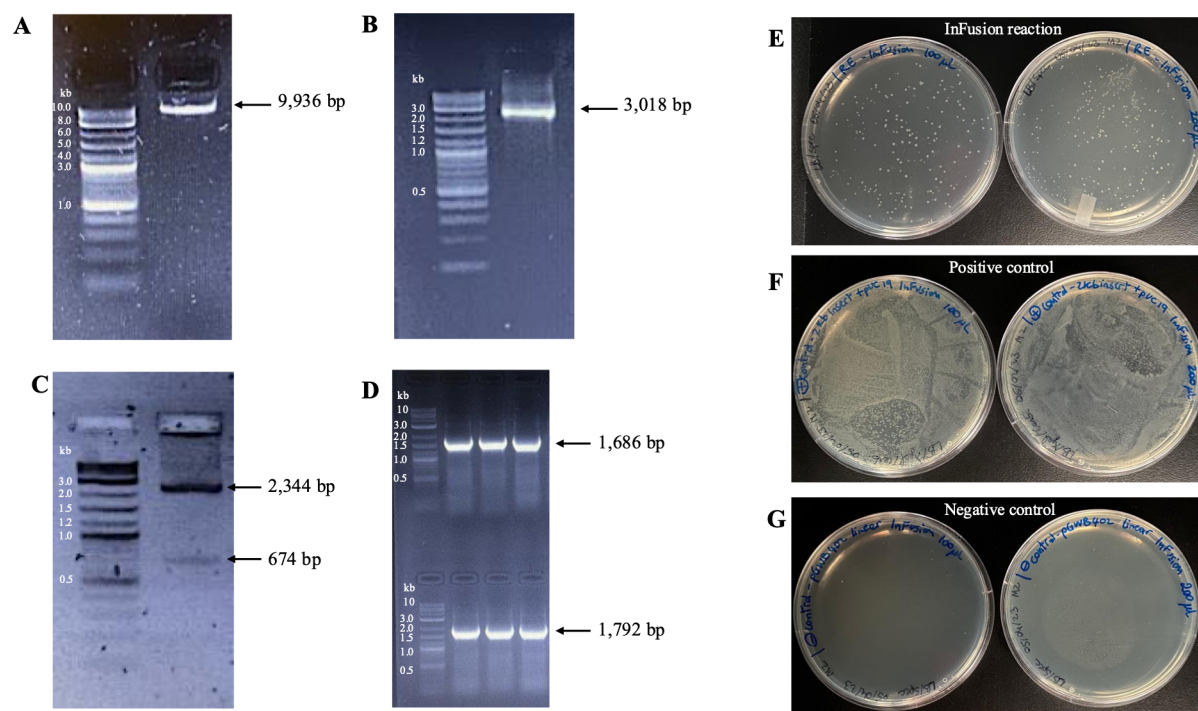


Figure 2. InFusion cloning strategy to clone g238180 in pGWB402. A) Represents the purified and linearized 9,936 bp pGWB402 while B) represents the purified PCR-amplified g238180 (3,018 bp). C) shows the two fragments produced by restriction digestion with HindIII of g238180. D) Shows the results of the colony PCR for three clones after transformation. Two fragments of 1,686 and 1,792 bp are obtained using overlapping primers covering the full insert including the promoter and terminator junctions. For all of them, a 1 kb ladder was used on a 1% TAE agarose gel. E) Shows the two aliquots of the transformation plated on LB (Luria Bertani) and selection and colonies growing on both plates. F) Shows confluent growth for the InFusion reaction positive control (pUC19 + 2kb insert) while G) shows the growth of InFusion reaction negative control (linearized pGWB402).

Table 1. Primer sequences used in this study for the subcloning of g238180 from pUC19 and to verify correct integration and orientation.

Name	Sequence	Purpose
RE_FOR	CACGGGGGACTCTAGAATGGCTACTACCAAT	

RE_REV	GATCGGGGAAATTCGAGCTCTCAAGTATGATG	Gene PCR-amplification from pUC19
80_R1	AGTTTGGAGTTCTGCCACTTAATTG	Correct gene integration and orientation
80_F2	GAGTGGTTTGGGAAGTTAGACTTGG	

Molecular validation of transformed *Agrobacterium* strains

The four overexpression constructs were successfully introduced into *Agrobacterium tumefaciens* to generate recombinant strains for transformation of Thompson Seedless somatic embryos. Construct integrity and orientation for g238180 were validated after InFusion cloning. For g238150, g238250 and g238290, construct integrity and orientation were confirmed separately using overlapping primer pairs spanning the CaMV35S-GOI and GOI-NOS junctions (**Figure 3A, Table 2**). For these constructs, colony PCR produced the expected amplification patterns predicted from the *in silico* construct designs (**Figure 3B, Table 3**). Amplicon sequencing further confirmed correct insertion and orientation of each GOI, absence of mutations, and 100% sequence identity to the expected construct sequence. One sequence-verified *Agrobacterium* clone harboring each overexpression construct was selected for transformation of Thompson Seedless somatic embryos.

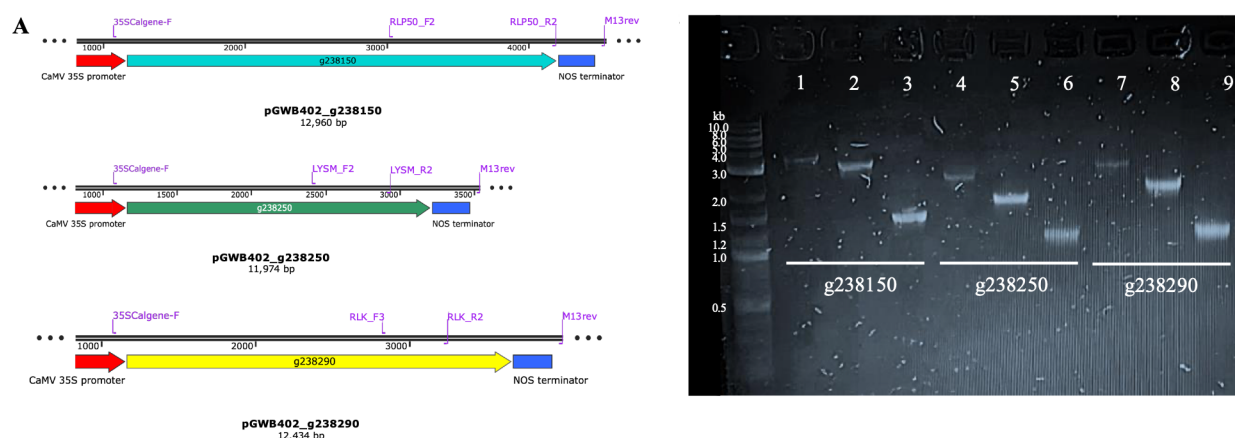


Figure 3. Molecular validation of recombinant *Agrobacterium* strains. A) Schematic representation of three out of four overexpression constructs showing the position of the primers used to verify construct integrity and orientation. Primers were designed to amplify overlapping regions spanning the entire length

of the gene, including promoter and terminator. The genes of interest are color-coded as follows: g238150 (teal), g238250 (green) and g238290 (yellow). The CaMV35S promoter is shown in red, the NOS terminator in blue and primer-binding sites in purple. Construct schematics were generated using SnapGene software (GSL Biotech LLC). B) Colony PCR products obtained from three independent *Agrobacterium* transformants for each construct. Lanes 1–3 correspond to g238150, lanes 4–6 to g238250, and lanes 7–9 to g238290. Expected amplicon sizes for each primer pair are reported in Table 1. PCR products were resolved on a 1% agarose gel in 1X TAE buffer. A 1kb DNA ladder was included as the molecular size marker, and the corresponding fragment sizes are indicated.

Table 3. Pair of primers used to confirm orientation and integrity of the genes of interest in the binary vector pGWB402 for overexpression in *V. vinifera* cv. Thompson Seedless. LRR-RLP: leucine-rich repeat receptor-like protein; LysM-RLK: lysin-motif receptor-like kinase; LRR-RLK: leucine-rich repeat receptor-like kinase.

GOI	Gene family	Primer pair	Amplicon size (bp)	Lane on gel electrophoresis
g238150	LRR-RLP	35S/M13rev	3,460	1
		35S/RLP50_R2	3,116	2
		RLP50_F2/M13rev	1,515	3
g238250	LysM-RLK	35S/M13rev	2,464	4
		35S/LYSM_R2	1,862	5
		LYSM_F2/M13rev	1,129	6
g238290	LRR-RLK	35S/M13rev	2,924	7
		35S/RLK_R2	2,175	8
		RLK_F3/M13rev	1,174	9

Overexpression of candidate genes in Thompson Seedless results in different regeneration outcomes

Distinct regeneration phenotypes were observed among the four overexpression constructs following *Agrobacterium*-mediated transformation of Thompson Seedless somatic embryos. Constructs carrying g238150 (LRR-RLP) and g238250 (LysM-RLK) repeatedly produced necrotic embryos that failed to regenerate to whole plants (**Figure 4, B and C**) over six rounds of transformation. In contrast, constructs

carrying g238290 (LRR-RLK) and g238180 (LRR-RLP) consistently produced healthy secondary embryos that continued to proliferate on selection medium (**Figure 4A**). For these two constructs, stable secondary embryo development was observed approximately six months after transformation. Subsequent transfer of these secondary embryos to germination media resulted in the regeneration of ten putatively transgenic TS plants, five for each construct (**Figure 4D**).

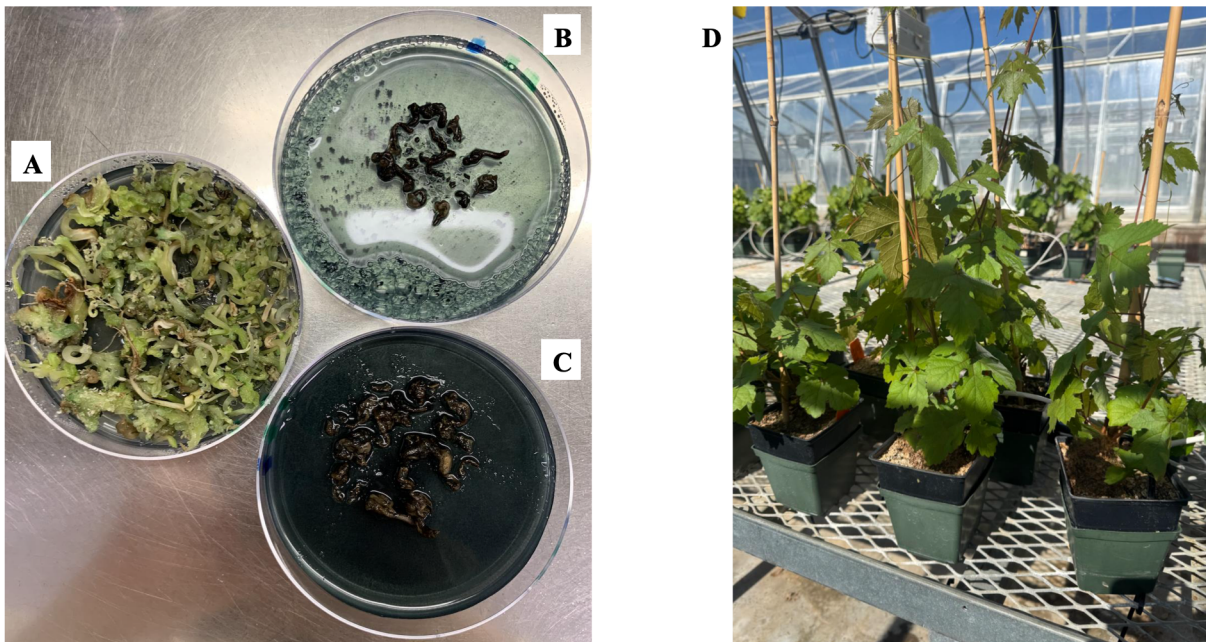


Figure 4. Regeneration outcome following *Agrobacterium*-mediated transformation of Thompson Seedless somatic embryos. A) Healthy secondary embryos regenerated on kanamycin selection medium following transformation with the g238290 (LRR-RLK) overexpression construct. (B, C) Representative necrotic embryos observed after transformation with g238150 (LRR-RLP) and g238250 (LysM-RLK) overexpression constructs, respectively. D) Putatively transgenic Thompson Seedless plants regenerated from secondary embryos transformed with g238290 (LRR-RLK) and g238180 (LRR-RLP) constructs.

Selection of target sequences for candidate genes targeted mutagenesis in *V. arizonica*

To generate putatively edited *V. arizonica* b40-14 plants four target sites were selected within each candidate gene to induce mutations that are expected to disrupt gene function. For the two genes coding for LRR-RLPs (g238150 and g238180), all four guides were designed within the extracellular LLR motif. For g238250 (LysM-RLK) and g238290 (LRR-RLK), two guides were designed to target the extracellular domain (LRR; LysM), and two the kinase domain (**Figure 5A; Table 4**).

CRISPOR identified a list of candidate guides for each gene, but the selected sgRNAs were manually refined to maximize Cas9 cutting efficiency while minimizing off-target edits. Initial guide prediction identified 229 to 326 candidate sgRNAs per gene. After applying the filtering criteria, the number of candidate guides was reduced to 13 for g238150, 14 for g238180, 59 for g238250 and 101 for g238290 (**Table 5**). Manual inspection of predicted off-targets further reduced the list of sgRNAs. Overall, four sgRNA per each gene were selected based on a combination of high predicted specificity and editing efficiency, low number of predicted off-targets, off-target genomic location and mismatch position relative to the PAM sequence. The sgRNAs that satisfied these criteria were used to build the vectors for *V. arizonica* protoplast transfection.

Assembly and validation of eight sgRNA constructs for genome editing

Eight sgRNA plasmids were successfully generated to enable simultaneous targeting of the four sites at each gene. To increase multiplex editing efficiency, the four sgRNAs were divided into two plasmids using the tRNA expression system (Xie *et al.*, 2015), resulting in two sgRNA cassettes per gene (**Figure 5B**).

Following cloning, the plasmids were transformed into *E. coli* and two colonies per construct were screened by restriction digestion with HindIII and PstI (**Figure 5C**). All constructs produced the expected digestion pattern predicted in silico, with fragments of 2,252 bp and 746 bp. The RLK_1-2 construct was the only one that produced three fragments (2,252, 628, and 118 bp) due to the presence of an additional HindIII site in one of the sgRNA sequences. Sequencing of the validated plasmids confirmed correct assembly of all sgRNA cassettes and 100% sequence identity to the original in silico design. These results confirmed the successful assembly of all eight sgRNA expression constructs, which were used for *V. arizonica* protoplast transfection.

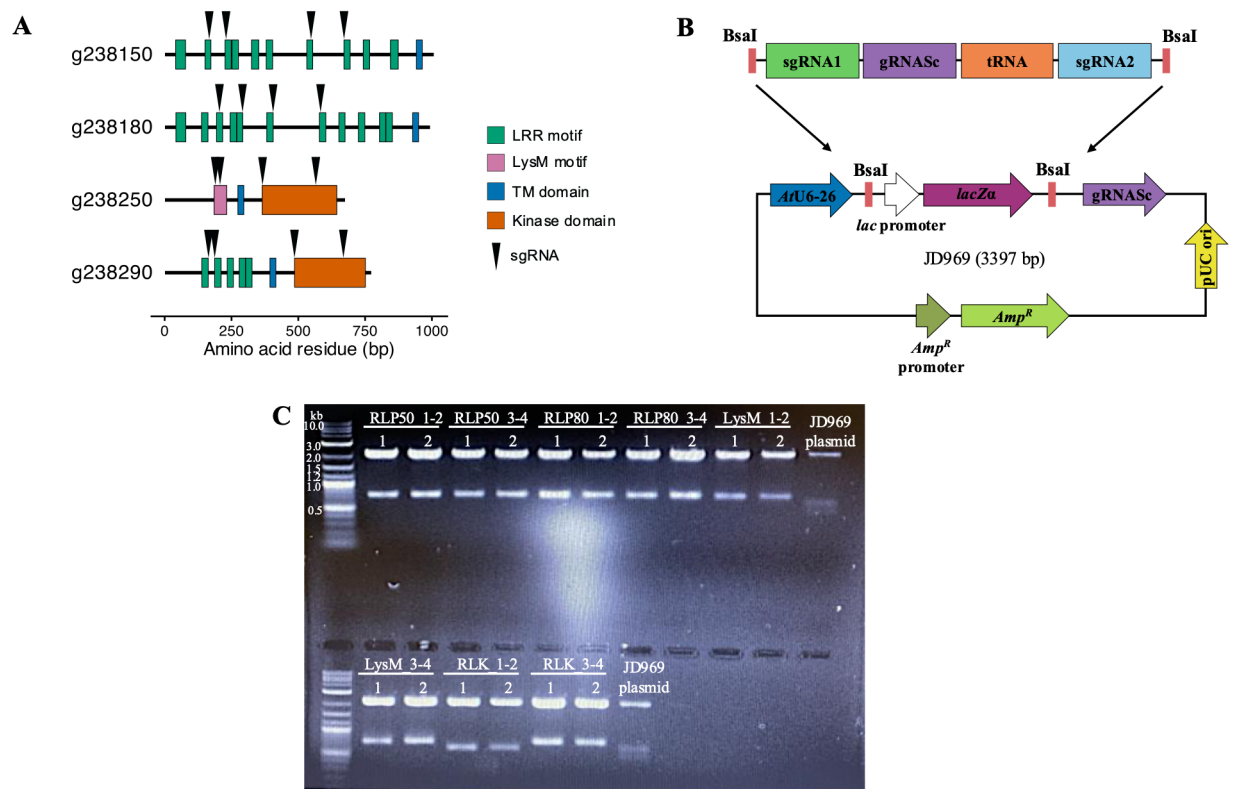


Figure 5. Design and validation of CRISPR/Cas9 constructs for targeted *V. arizonica* editing. A) Schematic representation of the four candidate genes showing the predicted functional domains and the position of the four sgRNAs designed for each gene. Functional domains are color-coded as follows: leucine-rich repeat (LRR; green), lysin motif (LysM; pink), transmembrane domain (TM; blue), and kinase domain (orange). Black arrows indicate the sgRNA target sites. B) Schematic overview of the multiplex sgRNA cloning strategy. Two sgRNAs were assembled into a single expression cassette consisting of sgRNA1–sgRNA scaffold–tRNA–sgRNA2 and inserted into the custom vector JD969 through Golden Gate cloning using the BsaI restriction sites (red rectangles). Two expression constructs were generated for each candidate gene and then combined with purified Cas9 to assemble CRISPR/Cas9 ribonucleoprotein (RNP) complexes *in vitro* for PEG-mediated transfection of *V. arizonica* protoplasts. C) Restriction analysis of the eight sgRNA expression constructs following cloning into JD969. Two independent *E. coli* colonies were screened for each construct by HindIII/PstI digestion. Construct names are indicated above each lane, and

colonies are labelled as 1 and 2. The empty JD969 vector was included as a control. DNA fragments were resolved on a 1% agarose gel in 1X TAE buffer using a 1 kb DNA ladder.

Table 4. List of sgRNAs designed in this study to target the PD candidate resistance genes g238150 (LRR-RLP), g238180 (LRR-RLP), g238250 (LysM-RLK) and g238290 (LRR-RLK). Underlined nucleotides indicate the PAM motif in the sgRNA sequence.

Target gene	sgRNA	Sequence (5'- 3')	GC content (%)
g238150	RLP50_g1	CAACCAGTATGCTAGTAGTCC <u>GG</u>	52.2
	RLP50_g2	TTTCAAATAATCAGCTGAGC <u>GGG</u>	43.5
	RLP50_g3	AATGAAGCGACGGATGACTAT <u>GG</u>	47.8
	RLP50_g4	ATTTGTAGAAAATGCCGTTGT <u>TGG</u>	39.1
g238180	RLP80_g1	GATTGACTCCTTCTAAATTA <u>AGG</u>	34.8
	RLP80_g2	AAGTAATGAAAATTTAGCGAT <u>TGG</u>	30.4
	RLP80_g3	GACAGGTTTCCTAATGAATA <u>AGG</u>	39.1
	RLP80_g4	GGAGCAGAGAACCATTGAAAT <u>TGG</u>	47.8
g238250	LYSM_g1	TGGCATTGCTGAATCGCTGG <u>AGG</u>	56.5
	LYSM_g2	CACCCCGCATAAGATTGTAGT <u>TGG</u>	52.2
	LYSM_g3	AACCCGAGCATCCAACCATG <u>GGG</u>	60.9
	LYSM_g4	GGGTTGGGGAGATTATACCT <u>TGG</u>	56.5
g238290	RLK_g1	ACCAAGGTAGATACCTTCGGT <u>TGG</u>	52.2
	RLK_g2	GCAGCAAAGCTTGGGACACA <u>AGG</u>	56.5
	RLK_g3	TGACAGTGTCCCCGATTGGAT <u>TGG</u>	56.5
	RLK_g4	CTCGTGGGCCCTATAGTAACT <u>TGG</u>	56.5

Table 5. Total number of sgRNA sequences generated by CRISPOR per each gene and filters applied to reduce the number of sequences to use for the study. The numbers in the table represent the number of sgRNA retained in each filtering step.

Gene	Initial number of sgRNAs	Filters applied			
		CFD	MITscore	Doench2016	Graf2019
g238150	326	142	17	14	13
g238180	320	124	20	18	14
g238250	229	183	77	60	59
g238290	233	164	162	117	101

Regeneration of putatively edited plants from *V. arizonica* edited protoplasts

Sequence-validated sgRNA plasmids were co-transfected with Cas9 into isolated *V. arizonica* protoplasts using PEG-mediated delivery to generate multiplex CRISPR/Cas9 editing complexes.

Following protoplast encapsulation and culture, regenerated embryos became visible approximately three months after protoplast isolation (**Figure 6A**). Developing embryos were transferred to larger containers where they continued to regenerate into whole plants (**Figure 6B**). Protoplasts that generated micro calli were used for molecular characterization of CRISPR/Cas9-induced editing events.

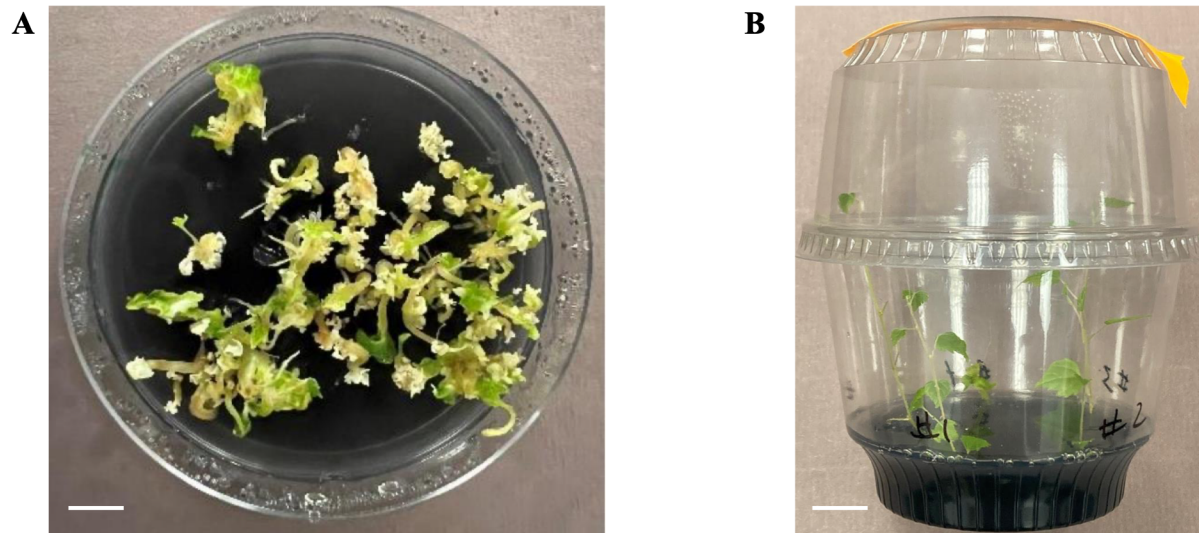


Figure 6. Regeneration of *V. arizonica* plants from edited protoplasts. A) Developing embryos derived from PEG-transfected *V. arizonica* protoplasts encapsulated in alginate beads. B) Individual regenerated embryos transferred to a plastic dome for continued growth and development into whole plants. The identification numbers (#1, #2, etc) on the dome correspond to the alginate bead from which the embryo originated, allowing regenerated plants to be traced back to their source. Scale = 1 cm.

Early assessment of editing efficiency and mutation patterns in edited *V. arizonica* calli

Only a subset of encapsulated protoplasts regenerated micro calli suitable for CRISPR editing analysis. Seven beads were available, and the number of events recovered differed among targeted genes (**Table 6**).

Five beads (#21, #16, #26, #72, #1) originated from co-transfection with sgRNAs targeting individual candidate genes, whereas two beads (#99 and #105) were generated following co-transfection with all the sgRNA constructs. A total of 152 PCR amplicons were analyzed with CRISPResso2. After quality filtering, 60% of the samples containing more than 200 reads were retained for downstream analysis. Across these samples, an average of 97% of reads aligned to the corresponding reference amplicons.

Editing efficiency varied substantially among target sites (**Table 6**). RLK_g1-2 showed the highest editing efficiency, with approximately 99% in both beads #99 and #105. RLP80_g4 exhibited medium to high efficiency across three independent beads (#16, #26 and #72), ranging from approximately 46% to 93%. Guides targeting g238150 showed mixed results. RLP50_g1-2 had an editing efficiency of 53% in bead #21, intermediate efficiency in bead #99 (31%) and very low (3%) in bead #105. RLP50_g3-4 showed consistent low efficiency (< 1%) across all beads. Similarly, for bead #1, one of the guides targeting g238250 (LYSM_g4), showed efficiency of ~1% across all analyzed events.

Highly efficient guides generated stable edited genotypes. Cas9 activity at RLK_g1-2 consistently produced biallelic heterozygous deletions in both beads #99 and #105. For RLP80_g4 more variable allelic profiles were observed, including beads showing biallelic heterozygous edits as well as beads with more than two edited alleles. In two out of three beads targeting RLP50_g1-2, a monoallelic mutation was observed. No edited genotypes were observed for RLP50_g3-4 or LYSM_g4, as mutations at these sites were mainly substitutions outside of the quantification window (3 bp upstream from the PAM).

Across all edited alleles, deletions represented the predominant mutation class (82%), followed by substitutions (10.1%), insertions (7.4%) and combined mutation events (0.5%) (**Figure 7A**). Mutation profiles differed among sgRNAs (**Figure 7B**). Editing at RLK_g1-2 and RLP50_g1-2 predominantly generated deletions, whereas editing at RLP80_g4 produced a combination of deletions and insertions, with deletions being the most frequent mutations. RLP50_g3-4 and LYSM_g4 were characterized almost exclusively by substitution events.

Mutation length also differed according to the target site (**Figure 7C**). Editing at RLK_g1-2 consistently generated large deletions of 71 and 73 nucleotides, while RLP50_g1-2 produced 31 bp deletions. On the

other hand, RLP80_g4 generated a broad spectrum of short indels, mainly of 1-, 3-, 5, 6- and 8-bp. Representative nucleotide composition generated by CRISPResso2 illustrates the distinct editing pattern observed at each target site for each bead (**Figure 7D**).

Overall, efficient editing was obtained for one of the sgRNAs targeting g238180, intermediate editing was observed at the first two guides targeting g238150, and little to no editing was observed for RLP50_g3-4 and the guide targeting g238250. High efficiency was also observed in the two beads co-transfected with sgRNAs targeting multiple genes at the site targeting g238290.

Table 6. Summary of editing outcomes across regenerated *V. arizonica* calli. Editing outcomes determined by CRISPResso2 analysis of regenerated *V. arizonica* calli following PEG-mediated transfection. For each target gene and regenerated bead, the table reports the number of events screened and retained after quality filtering (in parentheses), the sgRNA, editing efficiency (mean $\pm$ SE when multiple events were available), the allelic profile and the corresponding genotype description. Editing efficiency was calculated as the percentage of modified reads and the total number of reads that aligned to the reference amplicon. Genotype descriptions were assigned based on the allelic profile detected in each regenerated event.

Target gene	Bead ID	Events (analyzed)	sgRNA	Editing efficiency (%)	Allelic profile	Genotype description
g238150	21	11 (9)	RLP50_g1-2	53.2 $\pm$ 1.9	-31/wt	Monoallelic
		11 (6)	RLP50_g3-4	0.8 $\pm$ 0.1	wt	-
g238180	16	6 (2)	RLP80_g4	46.3 $\pm$ 6.6	wt/+1/-6/-8	Putative cell fusion
	26	10 (3)	RLP80_g4	93.2 $\pm$ 2.1	wt/+1/-3/-5	Putative cell fusion
	72	3 (1)	RLP80_g4	88.2	+1/-8	Biallelic, heterozygous
g238250	1	31 (29)	LYSM_g4	1.7 $\pm$ 0.1	wt	-
		9 (7)	RLP50_g1-2	31.0 $\pm$ 5.1	wt/-31	Monoallelic
		9 (4)	RLP50_g3-4	1.1 $\pm$ 0.4	wt	-
All	99	9 (5)	RLP80_g4	30.0 $\pm$ 2.1	wt/-1/+1	Putative cell fusion
		9 (5)	RLK_g1-2	99.9 $\pm$ 0.1	-71/-73	Biallelic, heterozygous

105	11 (8)	RLP50_g1-2	2.7 ± 0.4	wt	-
	11 (6)	RLP50_g3-4	0.9 ± 0.1	wt	-
	11 (2)	RLP80_g4	93.3 ± 1.8	-1/-3	Biallelic, heterozygous
	11 (1)	RLK_g1-2	99.9	-71/-73	Biallelic, heterozygous

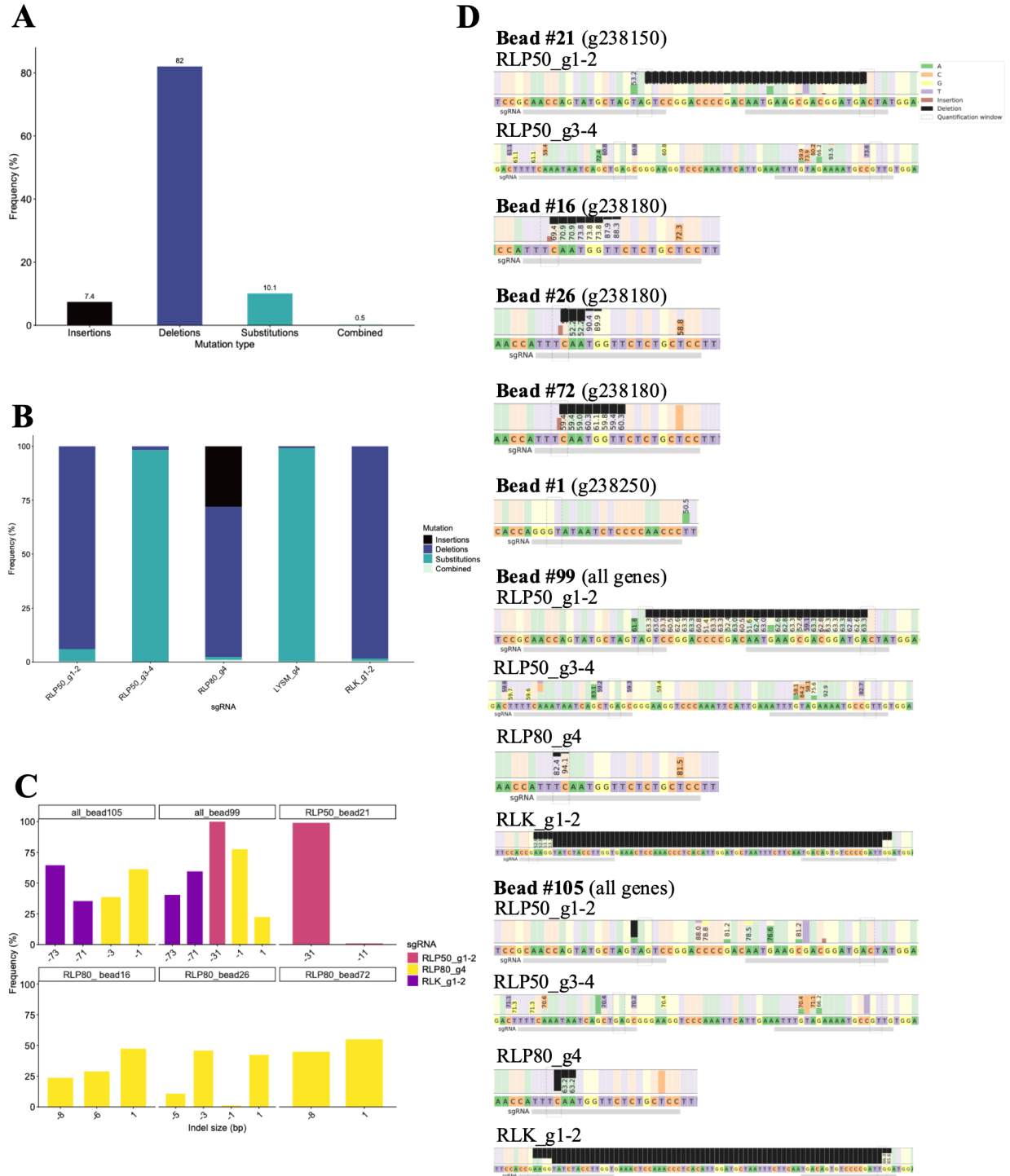


Figure 7. CRISPR/Cas9 editing outcomes and mutation profiles in *V. arizonica* regenerated calli. A) Overall frequency of mutation types detected across all edited samples following CRISPR/Cas9 editing. B) Frequency of mutation types separated per sgRNA. Panels A and B share the same color legend. C) Frequency of mutation length generated by each sgRNA across the analyzed beads. Negative values

(numbers with a – sign on the x axis) indicate deletions, while positive values indicate insertions. Only sgRNA reporting insertions or deletions are shown. D) Representative CRISPResso2 nucleotide composition plots illustrating editing outcomes for each bead and sgRNA. The reference amplicon sequence is displayed at the bottom of each plot, with the sgRNA target site represented by a grey bar below the reference sequence. The dashed box represents the CRISPResso2 quantification window (3bp upstream the PAM site). Colored bars represent the proportion of sequencing reads containing each nucleotide at every position (A = green; C = orange; G = yellow; T = purple). Black bars indicate deletions. Brown bars between nucleotides represent insertions.

Discussion

Cell-surface receptors are promising targets for disease resistance breeding (Gao et al., 2025), and in this study we established an experimental platform to functionally validate four candidates PD resistance genes encoding for cell-surface receptors.

Despite challenges related to transgene removal and chimerism, *Agrobacterium*-mediated transformation of embryogenic tissue is a widely adopted method to generate transgenic grapes (Nakajima *et al.*, 2017) and to validate genes putatively involved in disease resistance (Yan *et al.*, 2023). To investigate gene function, the candidate genes were overexpressed under the control of the CaMV35S promoter, a successful strategy adopted for RLKs functional analysis in *Arabidopsis* (Gou *et al.*, 2010). Although constitutive expression does not reproduce the native spatial or temporal regulation of immune receptors, it represents an effective strategy for an initial assessment of gene function. This approach has been widely adopted for exploratory functional studies and has revealed the contribution of numerous genes to disease resistance and other agronomic traits (Gambino *et al.*, 2005; Su *et al.*, 2018; Zok *et al.*, 2010).

All four candidate genes were successfully assembled into overexpression vectors, validated, and introduced in *Agrobacterium*. The same binary vector, transformation protocol and explant source were used for all the candidates, however, only two out of the four constructs ultimately regenerated putative transgenic TS plants. While constructs carrying g238290 (LRR-RLK) and g238180 (LRR-RLP)

consistently produced secondary embryos that regenerated into putatively transgenic plants, six transformation attempts with the constructs carrying g238250 (LysM-RLK) and g238150 (LRR-RLP) were unsuccessful and resulted in repeated embryo necrosis on antibiotic selection. These results suggest that the constitutive overexpression of these two receptors might negatively affect embryogenic tissue viability, and that correct temporal and spatial expression might be needed, as reported in another study (François *et al.*, 2008). Unsuccessful regeneration of only two out of four genes may indicate that constitutive expression of specific receptors can potentially interfere with developmental processes required for embryo development. Another possible explanation is that the overaccumulation of immune receptors can trigger autoimmunity, excessive ROS production and consequently programmed cell death (Hailemariam *et al.*, 2024; R. Wang *et al.*, 2022).

To achieve candidate genes targeted mutagenesis, CRISPR/Cas9 genome editing was used to knock-out the four candidate genes in the resistant *V. arizonica* background. Despite challenges related to low regeneration efficiency from edited protoplasts in *Vitis*, we were able to regenerate *V. arizonica* putatively edited plantlets by applying a protocol developed by the UCDPTF that works for wild *Vitis* species (Tricoli & Debernardi, 2024).

Both editing efficiency and mutation spectrum were evaluated and the results showed that Cas9 activity is strongly guide-dependent. Due to difficulties related to regeneration of whole plants from protoplasts after transfection, only a subset of beads was analyzed, depending on their development.

For each candidate gene, four sgRNAs were designed. A recent study in *Arabidopsis* revealed that using multiple sgRNAs increases gene targeting efficiency and their combination can compensate for the presence of less efficient sgRNAs (Li *et al.*, 2024). In our study, even if the four sgRNAs designed for each candidate gene were co-transfected during PEG-mediated transfection, edits were not observed for all the target sites. Two plausible reasons are either the distance between target sites or defective primer design around the target sites. For example, editing at LYSM_g4, one of the guides targeting g238250, produced almost exclusively substitutions and therefore had very low efficiency despite favorable computational predictions. RLP50_g3-4, two of the four guides targeting g238150, displayed very low efficiency, but this result was

compensated by good efficiency observed at RLP50_g1-2 that produced a monoallelic deletion of 31 nucleotides. Edits at RLP80_g4, one of the guides targeting g238180, showed more than two alleles, which either suggests a fusion of wild type and edited protoplasts during PEG-mediated transfection (Scintilla *et al.*, 2022) or a chimeric event. A much larger number of events needs to be screened to confirm this hypothesis.

Two beads transfected with all the sgRNAs targeting all four candidate genes were included to evaluate whether simultaneous delivery of multiple expression cassettes could induce editing at multiple target loci following a single transfection (Ma *et al.*, 2015). In these beads, editing at RLK_g1-2, two guides targeting g238290, resulted in very high efficiency and consistently generated biallelic heterozygous mutation, a desirable outcome for genome editing in woody plants (X. Wang *et al.*, 2018). However, because g238290 was only evaluated in the multiplex co-transfection experiment and not in independently transfected beads, it remains unclear if the observed efficiency reflects the real sgRNAs performance or the influence of the multiplex delivery strategy. In contrast, RLP50_g1-2 exhibited variable editing efficiency between the two beads, highlighting the variability of CRISPR-editing in individual transformations and reinforcing the importance to evaluate many independent events to accurately estimate guide performance.

Although computational tools have improved guide selection, prediction scores do not always correlate with experimental editing efficiencies in (Doench *et al.*, 2016; Hassan *et al.*, 2021). In addition, factors such as sgRNA GC content, target sequence, and chromatin state are known to affect cleavage efficiency (Jung *et al.*, 2024). Taken together, these results are consistent with previous observations that experimental validation remains fundamental when selecting sgRNA for plant genome editing.

The mutation profile observed agrees with previous reports in plants showing that repair of Cas9-induced double-strand breaks (DSBs) predominantly occurs through the non-homologous end joining (NHEJ) pathway, an error-prone repair mechanism that generates deletions and, less commonly, insertions or substitutions (Yang *et al.*, 2017).

Although deletions predominated, the mutation spectrum was highly guide dependent. In all the target sites that presented mutations, paired sgRNAs (RLK_g1-2, RLP50_g1-2) consistently generated large deletions,

while editing at RLP80_g4 primarily produced short insertions and deletions. Simultaneous cleavage by two nearby sgRNA is expected to remove the whole DNA fragment if both DSBs happen before repair, generating large deletions. In contrast, a single Cas9-induced break is generally repaired through local NHEJ, resulting in short indels around the cleavage site that frequently generate frameshift mutations that introduce premature stop codons, ultimately producing truncated proteins (Vu *et al.*, 2020; X. Wang *et al.*, 2018). The guides for which mostly substitutions were generated were not accompanied by insertions or deletions and therefore are unlikely to represent canonical CRISPR/Cas9 editing events. Instead, they could be naturally occurring sequence variations or PCR sequencing errors. This is consistent with the very low editing efficiency estimated for these guides.

In this study, we estimated sgRNA editing efficiency, mutation type and size, and allelic composition of individual edited events. However, these analyses were performed on DNA extracted from micro calli rather than fully regenerated plants. This work provides an important early assessment of editing outcomes and guide performance, but the regenerated plants will require independent molecular characterization to confirm editing status.

Conclusions

Overall, in this study we generated ten putatively transgenic TS lines and established an early genome editing screening platform in *V. arizonica*, revealing substantial variation in editing efficiency among sgRNAs. Putatively edited *V. arizonica* plants are currently being regenerated. Pathogen bioassays will be conducted on both transgenic TS and edited *V. arizonica* lines to determine whether these receptors contribute to PD resistance. Future studies should also investigate the combined expression of multiple candidate receptors, as plant immune receptors often function as interconnected signaling networks and phenotype in single loss-of-function mutants might not display because of the redundancy of structurally similar genes (C. Gao *et al.*, 2025; Lv *et al.*, 2016). In addition, expressing candidate genes under their native promoters may provide a more biologically relevant assessment of their function by preserving their natural spatial and temporal expression.

For targeted mutagenesis studies, further optimization of *V. arizonica* protoplast regeneration remains one of the main bottlenecks (Tricoli & Debernardi, 2024) and needs improvement before it can be applied to edit vegetatively propagated crops. Because regeneration efficiency is low and CRISPR editing is highly variable, screening a large number of events is essential to accurately evaluate genome editing outcomes and identifying desirable mutations. Repeating the multiplex editing experiment with a greater number of regenerated events will also provide a more rigorous evaluation of simultaneous editing of multiple candidate genes, an approach that is increasingly recognized as an important strategy for dissecting complex traits controlled by multiple loci (Lin *et al.*, 2026). This study provides the first experimental platform to validate the biological function of PD-associated candidate genes and represents an important step toward experimentally testing their contribution to PD resistance in grapevine.

References

- Agüero, C., Meredith, C., & Dandekar, A. (2006). Genetic transformation of *Vitis vinifera* L. cvs Thompson Seedless and Chardonnay with the pear PGP and GFP encoding genes. *Vitis - Journal of Grapevine Research*, 45.
- Castro, C., DiSalvo, B., & Roper, M. C. (2021). *Xylella fastidiosa*: A reemerging plant pathogen that threatens crops globally. *PLOS Pathogens*, 17(9), e1009813. <https://doi.org/10.1371/journal.ppat.1009813>
- Clement, K., Rees, H., Canver, M. C., Gehrke, J. M., Farouni, R., Hsu, J. Y., Cole, M. A., Liu, D. R., Joung, J. K., Bauer, D. E., & Pinello, L. (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nature Biotechnology*, 37(3), 224–226. <https://doi.org/10.1038/s41587-019-0032-3>
- Concordet, J.-P., & Haeussler, M. (2018). CRISPOR: Intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research*, 46(W1), W242–W245. <https://doi.org/10.1093/nar/gky354>
- Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H. W., Listgarten, J., & Root, D. E. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology*, 34(2), 184–191. <https://doi.org/10.1038/nbt.3437>
- El-Mounadi, K., Morales-Floriano, M. L., & Garcia-Ruiz, H. (2020). Principles, Applications, and Biosafety of Plant Genome Editing Using CRISPR-Cas9. *Frontiers in Plant Science*, 11. <https://www.frontiersin.org/articles/10.3389/fpls.2020.00056>
- François, J., Lallemand, M., Fleurat-Lessard, P., Laquitaine, L., Delrot, S., Coutos-Thévenot, P., & Gomès, E. (2008). Overexpression of the VvLTP1 gene interferes with somatic embryo development in grapevine. *Functional Plant Biology: FPB*, 35(5), 394–402. <https://doi.org/10.1071/FP07303>

- Gambino, G., Gribaudo, I., Leopold, S., Scharl, A., & Laimer, M. (2005). Molecular characterization of grapevine plants transformed with GFLV resistance genes: I. *Plant Cell Reports*, 24(11), 655–662. <https://doi.org/10.1007/s00299-005-0006-4>
- Gao, C., Chen, J., Sun, Z., & Zhang, H. (2025). Emerging Roles of Receptor-Like Proteins in Plant Immunity: Crosstalk, Signalling Networks and Prospects for Disease Resistance Breeding. *Molecular Plant Pathology*, 26(11), e70167. <https://doi.org/10.1111/mpp.70167>
- Gou, X., He, K., Yang, H., Yuan, T., Lin, H., Clouse, S. D., & Li, J. (2010). Genome-wide cloning and sequence analysis of leucine-rich repeat receptor-like protein kinase genes in *Arabidopsis thaliana*. *BMC Genomics*, 11, 19. <https://doi.org/10.1186/1471-2164-11-19>
- Graf, R., Li, X., Chu, V. T., & Rajewsky, K. (2019). sgRNA Sequence Motifs Blocking Efficient CRISPR/Cas9-Mediated Gene Editing. *Cell Reports*, 26(5), 1098-1103.e3. <https://doi.org/10.1016/j.celrep.2019.01.024>
- Hailemariam, S., Liao, C.-J., & Mengiste, T. (2024). Receptor-like cytoplasmic kinases: Orchestrating plant cellular communication. *Trends in Plant Science*, 29(10), 1113–1130. <https://doi.org/10.1016/j.tplants.2024.04.006>
- Hassan, M. M., Zhang, Y., Yuan, G., De, K., Chen, J.-G., Muchero, W., Tuskan, G. A., Qi, Y., & Yang, X. (2021). Construct design for CRISPR/Cas-based genome editing in plants. *Trends in Plant Science*, 26(11), 1133–1152. <https://doi.org/10.1016/j.tplants.2021.06.015>
- Huerta-Acosta, K. G., Riaz, S., Tenschler, A., & Walker, M. A. (2023). Genetic Characterization of Pierce's Disease Resistance in a *Vitis arizonica*/monticola Wild Grapevine. *American Journal of Enology and Viticulture*, 74(1). <https://doi.org/10.5344/ajev.2022.22021>
- Irwin, C. R., Farmer, A., & Willer, D. O. (2012). In-Fusion® Cloning with Vaccinia Virus DNA Polymerase. In S. N. Isaacs (Ed.), *Vaccinia Virus and Poxvirology: Methods and Protocols* (pp. 23–35). Humana Press. https://doi.org/10.1007/978-1-61779-876-4_2
- Jung, W. J., Park, S.-J., Cha, S., & Kim, K. (2024). Factors affecting the cleavage efficiency of the CRISPR-Cas9 system. *Animal Cells and Systems*, 28(1), 75–83. <https://doi.org/10.1080/19768354.2024.2322054>
- Krivanek, A. F., Riaz, S., & Walker, M. A. (2006). Identification and molecular mapping of PdR1, a primary resistance gene to Pierce's disease in *Vitis*. *Theoretical and Applied Genetics*, 112(6), 1125–1131. <https://doi.org/10.1007/s00122-006-0214-5>
- Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology*, 305(3), 567–580. <https://doi.org/10.1006/jmbi.2000.4315>
- Li, J., Kong, D., Ke, Y., Zeng, W., & Miki, D. (2024). Application of multiple sgRNAs boosts efficiency of CRISPR/Cas9-mediated gene targeting in *Arabidopsis*. *BMC Biology*, 22, 6. <https://doi.org/10.1186/s12915-024-01810-7>
- Lin, J., Hazaisi, H., Guan, Y., & Bai, M. (2026). Multiplex gene editing drives revolution in crop breeding: Overlaid editing of multiple genes and customization of complex traits. *Advanced Biotechnology*, 4(1), 5. <https://doi.org/10.1007/s44307-026-00099-7>
- Lv, Y., Yang, N., Wu, J., Liu, Z., Pan, L., Lv, S., & Wang, G. (2016). New insights into receptor-like protein functions in *Arabidopsis*. *Plant Signaling & Behavior*, 11(7), e1197469. <https://doi.org/10.1080/15592324.2016.1197469>
- Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., Wang, B., Yang, Z., Li, H., Lin, Y., Xie, Y., Shen, R., Chen, S., Wang, Z., Chen, Y., Guo, J., Chen, L., Zhao, X., Dong, Z., & Liu, Y.-G. (2015). A

- Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Molecular Plant*, 8(8), 1274–1284.
<https://doi.org/10.1016/j.molp.2015.04.007>
- Massonnet, M., Zaccheo, M., Cochetel, N., Figueroa-Balderas, R., Riaz, S., & Cantu, D. (2026). Haplotype graph analysis of PdR1 uncovers resistance diversity to Pierce's disease in *Vitis arizonica* and its hybrids. *G3 Genes|Genomes|Genetics*, jkag115.
<https://doi.org/10.1093/g3journal/jkag115>
- Morales-Cruz, A., Aguirre-Liguori, J., Massonnet, M., Minio, A., Zaccheo, M., Cochetel, N., Walker, A., Riaz, S., Zhou, Y., Cantu, D., & Gaut, B. S. (2023). Multigenic resistance to *Xylella fastidiosa* in wild grapes (*Vitis* spp.) and its implications within a changing climate. *Communications Biology*, 6(1), 1–15. <https://doi.org/10.1038/s42003-023-04938-4>
- Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., Tabata, R., Kawai, T., Tanaka, K., Niwa, Y., Watanabe, Y., Nakamura, K., Kimura, T., & Ishiguro, S. (2007). Improved Gateway Binary Vectors: High-Performance Vectors for Creation of Fusion Constructs in Transgenic Analysis of Plants. *Bioscience, Biotechnology, and Biochemistry*, 71(8), 2095–2100.
<https://doi.org/10.1271/bbb.70216>
- Nakajima, I., Ban, Y., Azuma, A., Onoue, N., Moriguchi, T., Yamamoto, T., Toki, S., & Endo, M. (2017). CRISPR/Cas9-mediated targeted mutagenesis in grape. *PLOS ONE*, 12(5), e0177966.
<https://doi.org/10.1371/journal.pone.0177966>
- Rapicavoli, J., Ingel, B., Blanco-Ulate, B., Cantu, D., & Roper, C. (2018). *Xylella fastidiosa*: An examination of a re-emerging plant pathogen: *Xylella fastidiosa*. *Molecular Plant Pathology*, 19(4), 786–800. <https://doi.org/10.1111/mpp.12585>
- Riaz, S., Tenschler, A. C., Heinitz, C. C., Huerta-Acosta, K. G., & Walker, M. A. (2020). Genetic analysis reveals an east-west divide within North American *Vitis* species that mirrors their resistance to Pierce's disease. *PLOS ONE*, 15(12), e0243445. <https://doi.org/10.1371/journal.pone.0243445>
- Scintilla, S., Salvagnin, U., Giacomelli, L., Zeilmaker, T., Malnoy, M. A., Rouppe van der Voort, J., & Moser, C. (2022). Regeneration of non-chimeric plants from DNA-free edited grapevine protoplasts. *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.1078931>
- Su, H., Jiao, Y.-T., Wang, F.-F., Liu, Y.-E., Niu, W.-L., Liu, G.-T., & Xu, Y. (2018). Overexpression of VpPR10.1 by an efficient transformation method enhances downy mildew resistance in *V. vinifera*. *Plant Cell Reports*, 37(5), 819–832. <https://doi.org/10.1007/s00299-018-2271-z>
- Tricoli, D. M., & Debernardi, J. M. (2024). An efficient protoplast-based genome editing protocol for *Vitis* species. *Horticulture Research*, 11(1), uhad266. <https://doi.org/10.1093/hr/uhad266>
- Vu, T. V., Sivankalyani, V., Kim, E.-J., Doan, D. T. H., Tran, M. T., Kim, J., Sung, Y. W., Park, M., Kang, Y. J., & Kim, J.-Y. (2020). Highly efficient homology-directed repair using CRISPR/Cpf1-geminiviral replicon in tomato. *Plant Biotechnology Journal*, 18(10), 2133–2143.
<https://doi.org/10.1111/pbi.13373>
- Walker, M. A., Riaz, S., & Tenschler, A. (2014). OPTIMIZING THE BREEDING OF PIERCE'S DISEASE RESISTANT WINEGRAPES WITH MARKER-ASSISTED SELECTION. *Acta Horticulturae*, (1046), 139–143. <https://doi.org/10.17660/ActaHortic.2014.1046.17>
- Wang, R., Li, C., Li, Q., Ai, Y., Huang, Z., Sun, X., Zhou, J., Zhou, Y., & Liang, Y. (2022). Tomato receptor-like cytosolic kinase RIPK confers broad-spectrum disease resistance without yield penalties. *Horticulture Research*, 9, uhac207. <https://doi.org/10.1093/hr/uhac207>

- Wang, X., Tu, M., Wang, D., Liu, J., Li, Y., Li, Z., Wang, Y., & Wang, X. (2018). CRISPR/Cas9-mediated efficient targeted mutagenesis in grape in the first generation. *Plant Biotechnology Journal*, 16(4), 844–855. <https://doi.org/10.1111/pbi.12832>
- Xie, K., Minkenberg, B., & Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proceedings of the National Academy of Sciences*, 112(11), 3570–3575. <https://doi.org/10.1073/pnas.1420294112>
- Yan, C., Yang, N., Li, R., Wang, X., Xu, Y., Zhang, C., Wang, X., & Wang, Y. (2023). Alfin-like transcription factor VqAL4 regulates a stilbene synthase to enhance powdery mildew resistance in grapevine. *Molecular Plant Pathology*, 24(2), 123–141. <https://doi.org/10.1111/mpp.13280>
- Yang, H., Wu, J.-J., Tang, T., Liu, K.-D., & Dai, C. (2017). CRISPR/Cas9-mediated genome editing efficiently creates specific mutations at multiple loci using one sgRNA in *Brassica napus*. *Scientific Reports*, 7(1), 7489. <https://doi.org/10.1038/s41598-017-07871-9>
- Zhang, J. (2022). Check CRISPR Editing Events in Transgenic Wheat with Next-Generation Sequencing. In S. H. Wani & A. Kumar (Eds.), *Genomics of Cereal Crops* (pp. 95–106). Springer US. https://doi.org/10.1007/978-1-0716-2533-0_4
- Zok, A., Oláh, R., Hideg, É., Horváth, V. G., Kós, P. B., Majer, P., Váradi, Gy., & Szegedi, E. (2010). Effect of *Medicago sativa* ferritin gene on stress tolerance in transgenic grapevine. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 100(3), 339–344. <https://doi.org/10.1007/s11240-009-9641-8>

Supplementary data

Supplementary Table 1. Primers used for sgRNA expression cassette cloning. Forward (Fw) and reverse (Rev) primer sequences contain the BsaI tail (underlined), the sgRNA (in bold) and either part of the sgRNA scaffold (Fw, lower case) or the tRNA (Rev, lower case) sequence for cloning purposes.

Gene	Name of the construct	Sequence (5' – 3')
g238150	RLP50_1-2	Fw: <u>GATGGTCTCTATTG</u> CAACCAGTATGCTAGTAGTCggttttaga
		gctagaaatagc
	RLP50_3-4	Rev: <u>GATGGTCTCTAAAC</u> GCTCAGCTGATTATTTGAAAtgcacc
		agccgggaatc
g238180	RLP50_1-2	Fw: <u>GATGGTCTCTATTG</u> AATGAAGCGACGGATGACTAggttttaga
		gctagaaatagc
	RLP80_1-2	Rev: <u>GATGGTCTCTAAAC</u> CAACGGCATTCTTCTACAAATtgcacc
		agccgggaatc
g238180	RLP80_1-2	Fw: <u>GATGGTCTCTATTG</u> GATTGACTCCTTCTAAATTAgttttagag
		ctagaaatagc
	RLP80_1-2	Rev: <u>GATGGTCTCTAAAC</u> TCGCTAAATTTTCATTACTTtgcacca
		gccgggaatc

	RLP80_3-4	Fw: <u>GATGGTCTCTATTGGACAGGTTTCCTAATGAATA</u> gttttaga gctagaaatagc Rev: <u>GATGGTCTCTAAACTTTCAATGGTTCTCTGCTCC</u> tgacc agccgggaatc
g238250	LysM_1-2	Fw: <u>GATGGTCTCTATTGTGGCATTGCTGAATCGCTGG</u> gttttaga gctagaaatagc Rev: <u>GATGGTCTCTAAACCTACAATCTTATGCGGGGTG</u> tgacc agccgggaatc
	LysM_3-4	Fw: <u>GATGGTCTCTATTGAACCCGAGCATCCAACCATG</u> gttttaga gctagaaatagc Rev: <u>GATGGTCTCTAAACGGGTATAATCTCCCCAACCC</u> tgacc agccgggaatc
g238290	RLK_1-2	Fw: <u>GATGGTCTCTATTGACCAAGGTAGATACCTTCGG</u> gttttaga gctagaaatagc Rev: <u>GATGGTCTCTAAACTGTGTCCCAAGCTTTGCTGC</u> tgac cagccgggaatc
	RLK_2-3	Fw: <u>GATGGTCTCTATTGTGACAGTGTCCCCGATTGGA</u> gttttaga gctagaaatagc Rev: <u>GATGGTCTCTAAACGTTACTATAGGGCCCACGAG</u> tgacc agccgggaatc

Supplementary Table 2. Primers flanking the targeted sites. Primers were designed so that the CRISPR editing site is located within the first 100 bp of the amplicon.

GOI	Targeted site	Primer sequence	Amplicon size (bp)
g238150	RLP50_g1-g2	Fw:TCCTCTGTCACGGAAGCGCAAGGCCACC TCGAGTCATC Rev:TTTAGCCTCCCCACCGACATGACTTATCT CGCCGCCAA	122
	RLP50_g3-g4	Fw:TCCTCTGTCACGGAAGCGTGGTTTTGGG AGTTAGACTTGC Rev:TTTAGCCTCCCCACCGACGTGGGAAAGA ACCGTGAAAGC	142
g238180	RLP80_g1	Fw:TCCTCTGTCACGGAAGCGACCCTGATGA GTCAAGCCAG Rev:TTTAGCCTCCCCACCGACAGGTACACAA GATTCCTCAGCTG	289
	RLP80_g2	Fw:TCCTCTGTCACGGAAGCGTCTGGGTGCA ACAACCTGTAGT	293

		Rev:TTTAGCCTCCCCACCGACTGGGCTTCTG TTAGAACACCT	
	RLP80_g3	Fw:TCCTCTGTCACGGAAGCGTTGGTTGCGTT AGATATCTCTGAGA Rev:TTTAGCCTCCCCACCGACTGGTGTCTGA CATTGAGCA	280
	RLP80_g4	Fw:TCCTCTGTCACGGAAGCGAACTGGACTT GGGCTATAATCAA Rev:TTTAGCCTCCCCACCGACTCCAGAGTCA CTAAACCAATCAGT	286
	LYSM_g1-g2	Fw:TCCTCTGTCACGGAAGCGTCAGCCTTTGT TGATCCCAA Rev:TTTAGCCTCCCCACCGACAGGAACCGTA TCACCTTCACT	292
g238250	LYSM_g3	Fw:TCCTCTGTCACGGAAGCGGTCTCGTTTGA GGGATCCCA Rev:TTTAGCCTCCCCACCGACTCAAGCAAGT CCCTAACAGCC	283
	LYSM_g4	Fw:TCCTCTGTCACGGAAGCGTGAAGATGAT GCTGAAGACTCCC Rev:TTTAGCCTCCCCACCGACATTCTCACCCA ATGCGCTGT	280
	RLK_g1-g2	Fw:TCCTCTGTCACGGAAGCGGGGACCACTT CCTGATAAAATCC Rev:TTTAGCCTCCCCACCGACCATTGTGAGA CAGAGCAACATCC	281
g238290	RLK_g3	Fw:TCCTCTGTCACGGAAGCGTCTTCCACAG GGATTTACACAG Rev:TTTAGCCTCCCCACCGACTGAAAGCAAA TCCAGGCGAAG	282
	RLK_g4	Fw:TCCTCTGTCACGGAAGCGAGGCAAAGAA GGAAGGAGGC Rev:TTTAGCCTCCCCACCGACCTCAAATGAG GGACGGGTAGT	284

CHAPTER 5

Knock-out of *VvMLO3,7,11* by RNP-mediated genome editing leads to reduced susceptibility to grape powdery mildew

Abstract

Powdery mildew (PM) is an economically important fungal disease of grapevine. *Mildew Resistance Locus O* (*MLO*) genes are conserved susceptibility factors that have been targeted to reduce PM susceptibility in several plant species. In this study, we used CRISPR/Cas9-mediated ribonucleoprotein (RNP) editing to simultaneously target *VvMLO3*, *VvMLO7*, and *VvMLO11* in *Vitis vinifera* Thompson Seedless (TS). Following RNP-mediated transformation of grapevine protoplasts, 170 regenerated plants were screened, of which approximately 15% carried edits in at least one target gene. Eight edited lines were selected for detailed characterization and carried diverse combination of mutations across the three *MLO* loci. Disease susceptibility was evaluated and four edited lines showed significantly reduced susceptibility compared with the non-edited TS control. Lines with extensive homozygous or biallelic mutations generally showed stronger reduction in susceptibility. Lines carrying mutations at both sites of *VvMLO11* showed reduced susceptibility, suggesting a potential contribution of this gene to disease resistance. Visual disease scores were positively correlated with fungal biomass, although genotype-specific differences were observed. Overall, our results demonstrate the potential of multiplex RNP-mediated editing of *MLO* genes to reduce PM susceptibility in grapevine and provide a framework for further dissecting the individual and combined contribution of *MLO* genes to disease resistance.

Introduction

Cultivated grapevine (*Vitis vinifera* L.) is highly susceptible to PM, a widespread fungal disease caused by the obligate biotroph *Erysiphe necator* (Feechan *et al.*, 2011). This pathogen infects all green tissues of the plant, severely reducing both fruit quality and overall yield. Characteristic symptoms include a white, powdery fungal growth on leaves and berries, associated with fruit shriveling and cracking. Vineyards

worldwide rely heavily on frequent and costly fungicide applications to control PM (Sambucci *et al.*, 2019). While effective, the use of chemicals raises serious economic and environmental concerns, underlining the need for sustainable strategies to develop disease-resistant grape cultivars.

Traditionally, breeding for resistance has focused on introducing dominant resistance (R) genes into elite grapevine backgrounds. However, this process is laborious and time-consuming for clonally propagated crops like grapes. An alternative and increasingly promising approach involves targeting host susceptibility (S) genes. In this case, resistance can be generated by loss-of-function at the S genes in the host, a mutation that can be achieved in elite cultivars using CRISPR-Cas9 technology (Kieu *et al.*, 2021). S gene-mediated resistance is generally more durable than R gene-based resistance because the pathogen must overcome dependence on a host S factor (Dong & Ronald, 2019). S genes have been identified in multiple crops, and research has shown that their disruption through genome editing leads to resistance to bacterial, viral and fungal pathogens (Zaidi *et al.*, 2018).

Among the best-characterized S gene families are the *MLO* genes. These genes are conserved across monocots and dicots, and their inactivation can confer PM resistance in several crops, including wheat (Wang *et al.*, 2014), rice (Elliott *et al.*, 2002), apple (Pessina *et al.*, 2014), *Arabidopsis* (Consonni *et al.*, 2006), tomato (Bai *et al.*, 2008), pea (Pavan *et al.*, 2011), tobacco (Appiano *et al.*, 2015), and pepper (Zheng *et al.*, 2013). Phylogenetic analyses have revealed that MLO proteins involved in PM susceptibility cluster in clades IV and V, sharing conserved sequence features such as a C-terminal tetrapeptide motif (Acevedo-Garcia *et al.*, 2014; Devoto *et al.*, 2003; Kusch & Panstruga, 2017).

In grapevine, several *MLO* homologs have been implicated in PM susceptibility (Feechan *et al.*, 2008; Winterhagen *et al.*, 2008; **Table 1**). Pessina *et al.* (2016) demonstrated that simultaneous knockdown of *VvMLO6*, *VvMLO7* and *VvMLO11* reduced PM symptoms. More recent studies demonstrated that CRISPR/Cas9-mediated mutagenesis of *VvMLO11* (called *VvMLO3* in Wan *et al.*, 2020) and a double mutant *VvMLO6/VvMLO7* (Moffa *et al.*, 2024) show enhanced resistance to PM. Transcriptomic analysis further supported the role of *VvMLO3*, which is upregulated in PM-infected *Vitis pseudoreticulata* leaves (Weng *et al.*, 2014). Despite these advances, the simultaneous knockout of *VvMLO3*, *VvMLO7*, and

VvMLO11 has not yet been explored. This gap limits our understanding of the combined contribution of these genes to disease susceptibility.

Genome editing approaches based on CRISPR-Cas9 have been extensively used to deliver genome-editing components into protoplasts. The DNA-free editing method is particularly appealing for vegetatively propagated crops like grapevine, where transgene removal by segregation can only be achieved by outcrossing, which can lead to unwanted changes in the genetic background of the plant (Scintilla *et al.*, 2022).

In this study, we applied an efficient protoplast-based, RNP-mediated genome editing protocol (Tricoli & Debernardi, 2024) to generate *V. vinifera* cv. Thompson Seedless grapevines with loss-of-function mutations in three *VvMLO* genes (*VvMLO3*, *VvMLO7*, and *VvMLO11*). Then, we evaluated the phenotype of the edited lines for resistance to PM by visual and molecular assessment of pathogen biomass accumulation after infection. This work lays the foundation for future use of S gene editing as an alternative approach to achieve durable, non-transgenic disease resistance in grapevine.

Materials and methods

Phylogenetic analysis of MLO sequences

Protein sequences of *MLO* genes were obtained from Appiano *et al.* (2015), using the GenBank accession numbers provided therein. For *Vitis vinifera* (*Vv*), the sequences of *VvMLO7* and *VvMLO11* were retrieved from the Grapedia database (<https://grapedia.org/genes/>) using the latest PN40024 v5.1 T2T annotation release (Shi *et al.*, 2023) and translated into protein sequences. The *VvMLO3* sequence was retrieved from Genoscope (<http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/>, gene ID: GSVIVT01025160001). The scientific community currently uses nomenclature from (Winterhagen *et al.*, 2008) and (Feechan *et al.*, 2008) and we opted for Winterhagen's nomenclature (**Table 1**).

Table 1. Cross-reference of *MLO* gene nomenclature used in this and previous studies. Genoscope IDs identify the grapevine gene models corresponding to the three *MLO* genes targeted in this study. The

nomenclature reported by each reference is provided to facilitate comparison of gene identifiers across studies, as different *MLO* naming conventions have been used in the literature. The F in parenthesis refers to the Feechan nomenclature.

Target gene	GenoscopeID	Reference	Nomenclature in reference
<i>VvMLO7</i>	GSVIVT01016304001	Winterhagen <i>et al.</i> , 2008	<i>VvMLO7</i>
		Feechan <i>et al.</i> , 2008	<i>VvMLO17</i>
		Malnoy <i>et al.</i> , 2016	<i>VvMLO7</i>
		Shi <i>et al.</i> , 2023	<i>VvMLO7</i> (Vitvi05_01chr13g07790)
<i>VvMLO11</i>	GSVIVT01025653001	Winterhagen <i>et al.</i> , 2008	<i>VvMLO11</i>
		Feechan <i>et al.</i> , 2008	<i>VvMLO3</i>
		Wan <i>et al.</i> , 2020	<i>VvMLO3</i> (F)
		Shi <i>et al.</i> , 2023	<i>VvMLO11</i> (Vitvi05_01chr08g11640)
<i>VvMLO3</i>	GSVIVT01025160001	Winterhagen <i>et al.</i> , 2008	—
		Feechan <i>et al.</i> , 2008	—
		Shi <i>et al.</i> , 2023	<i>VvMLO3</i> (Vitvi05_chr06g04700)

The PN40024 v5.1 CDS sequences were used as queries in BLAST searches against both haplotypes of the TS genome on Grapegnomics (Cochetel & Cantu, 2024), and the corresponding sequences were translated into proteins.

Multiple sequence alignment of the protein sequences was performed using MAFFT v7.511 (Katoh & Standley, 2013). The alignment was filtered using Gblocks v0.91b (Castresana, 2000). The best-fit model of protein evolution was determined using ModelTest-NG v0.1.7 (Darriba *et al.*, 2020). Phylogenetic reconstruction was carried out using RAXML-NG v0.9.0 (Kozlov *et al.*, 2019) with the filtered alignment, and 100 bootstrap replicates for support estimation. Phylogenetic tree was drawn using Figtree v1.4.4 (Rambaut, 2018).

RNP-mediated protoplast transfection and screening of edited lines

Thompson Seedless protoplasts generation, transfection, culture and whole plant regeneration were performed at the UC Davis Plant Transformation Facility following the protocol described in (Tricoli & Debernardi, 2024). Single guide RNAs (sgRNAs) targeting *VvMLO7* and *VvMLO11* were previously published (Malnoy *et al.*, 2016; Wan *et al.*, 2020), while sgRNAs for *VvMLO3* were designed in this study to target the third exon of the gene. Two independent sites per gene were targeted to increase the chances of getting edits and maximize knockout efficiency (Malnoy *et al.*, 2016) (**Table 2**). TS protoplasts were isolated and transfected with a mixture of the three pairs of sgRNAs targeting the *MLO* genes. DNA extraction was performed on young leaves of fully regenerated plants and amplicon sequencing was carried out as described in (Zhang, 2022). Amplicons were submitted for sequencing using the CRISPR sequencing service at MGH (https://dnacore.mgh.harvard.edu/new-cgi-bin/site/pages/crispr_sequencing_main.jsp). Reads were analyzed using the CRISPR edits analysis tool Geneious Prime, 2023.2.1 (<https://www.geneious.com>) by aligning the reads to the reference amplicon. The edited plants were then used for infection assays.

Table 2. List of sgRNAs used in this study to target the three *VvMLO* genes. Underlined nucleotides indicate the PAM motif in the sgRNA sequence.

Target gene	Target site	sgRNA sequence (5' – 3')	Reference
<i>VvMLO3</i>	1	AATATGCATACCCAATAACG <u>CGG</u>	This study
	2	TTCTGATTATTGCAGTTCCA <u>AGG</u>	
<i>VvMLO7</i>	1	TATGCATTTCTGAGAGTGTT <u>GGG</u>	Malnoy <i>et al.</i> , 2016
	2	TTTTAAGAACACATGCTCTG <u>AGG</u>	
<i>VvMLO11</i>	1	GGAGCAAACACCGACTTGGG <u>CGG</u>	Called <i>VvMLO3</i> (F) in Wan <i>et al.</i> , 2020
	2	GTGGTCATCCGCATATCAAT <u>GGG</u>	

E. necator inoculation and disease severity assessment

Phenotypic evaluations were performed at 1, 5, and 10 days post inoculation (dpi), followed by molecular analysis of fungal biomass and correlation between phenotypic and molecular measurements. To do this, the edited lines and the non-edited control (TS) were propagated to obtain four biological replicates per

line. The *E. necator* strain UCD-TS1 was multiplied as described in (Bonarota *et al.*, 2025), resuspended in 0.001% Tween20 solution, and used as inoculum to infect the *MLO*-edited lines and the non-edited control under controlled environment conditions. Plants were maintained in two growth chambers, with one chamber used for pathogen inoculation and the other for mock-inoculation. Mock-inoculated plants were sprayed with a 0.001% Tween20 solution. Inoculation and disease severity assessment were performed as described in (Amrine *et al.*, 2015). Three leaves were collected from each biological replicate at 1, 5 and 10 dpi from both the mock- and pathogen-inoculated plants. Two leaves were immediately frozen in liquid nitrogen and used for DNA extraction and subsequent quantitative PCR (qPCR) analysis. The third leaf was used for Coomassie Brilliant Blue R-250 staining (Riaz *et al.*, 2013), followed by evaluation at the microscope and visual scoring of infection rate.

Total plant and fungal DNA was extracted from leaf samples collected at 10 dpi and used as input for qPCR to measure *E. necator* biomass accumulation in infected tissues. Four biological replicates were analyzed for each line, with three technical replicates per biological replicate for the qPCR. Fungal biomass was quantified by amplifying the *E. necator* β -tubulin (TUB) gene, using the *V. vinifera* actin (ACT) gene as a reference (Jones *et al.*, 2014). Primers for the TUB were obtained from (Jones *et al.*, 2014) while ACT primers were designed in this study and had an efficiency of 97.7% (ACT_F: 5'-TGTGCTGGATTCTGGTGATGGT-3' and ACT_R: 5'-GTCACGTCCAGCAAGGTCAAGA-3'). Relative DNA accumulation was calculated using the formula $2^{-(En\ TUB\ Ct - Vvin\ ACT\ Ct)}$. To assess differences in fungal biomass accumulation among genotypes, a one-way ANOVA was performed on ΔCt values. Pairwise comparisons were conducted using Tukey's HSD test with adjusted p-values, and group differences were summarized using a compact letter display (CLD). The correlation between fungal biomass accumulation and visual disease scores was evaluated using Spearman's rank correlation coefficient. All statistical analysis was performed in R v4.5.1.

Results

Phylogenetic analysis reveals distinct evolutionary relationship among the three edited *VvMLO* genes

To investigate the evolutionary relationships among the three *MLO* genes targeted for editing in this study, a phylogenetic analysis of *VvMLO* and representative *MLO* protein sequences from other plant species was performed. Consistent with previous reports, *VvMLO* family members were distributed across multiple clades, with the three target genes clustering within clades IV and V (**Figure 1**).

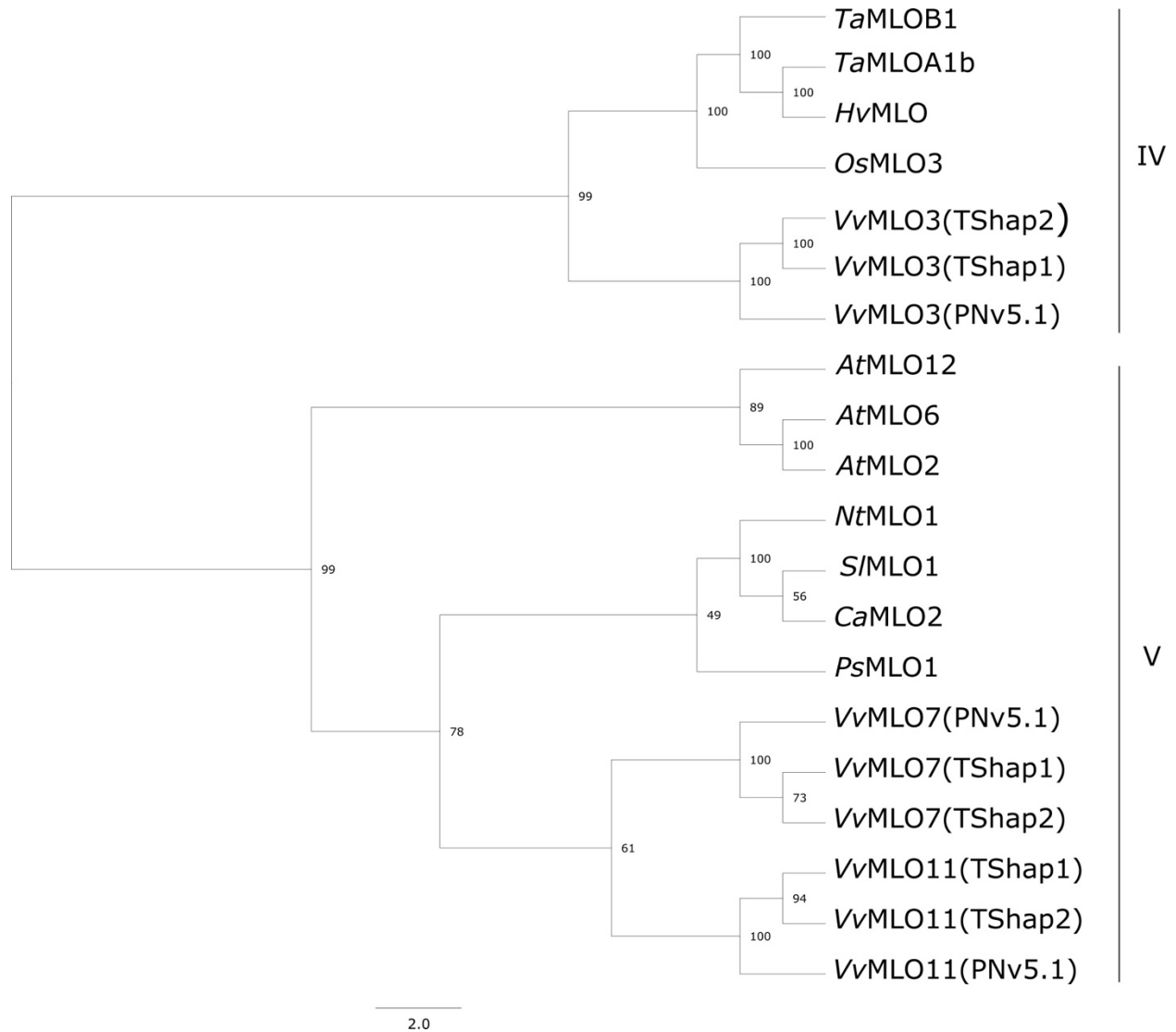


Figure 1. Phylogenetic relationship of *VvMLO3*, *7* and *11* among PM susceptibility-associated *MLO* homologs. A maximum likelihood (ML) phylogenetic tree was constructed using *MLO* protein sequences from monocots and dicots species known to be associated with PM susceptibility. Branch length represents sequence divergence, and number at nodes indicate bootstrap support values from 100 replicates. Species

included in the analysis were wheat (*Ta*), barley (*Hv*), rice (*Os*), *Arabidopsis* (*At*), tobacco (*Nt*), tomato (*Sl*), pepper (*Ca*), pea (*Ps*) and grapevine (*Vv*). For grapevine MLOs, sequences of PN40024v5.1 (PNv5.1) and both haplotypes of *V. vinifera* cv. Thompson Seedless (TS), which is the genetic background of the edited lines, were included.

To determine whether *Vv*MLO3 corresponded to any of the grape genes previously reported to cluster within clade IV, known to be confined to monocot MLO homologs (Devoto *et al.*, 2003), its sequence was compared with *Vv*MLO14 (Feechan *et al.*, 2008) and *Vv*MLO1 (Winterhagen *et al.*, 2008) (**Figure 2**). Although the sequence alignment showed approximately 90% query coverage, sequence identity was relatively low (~40%), indicating that *Vv*MLO3 is distinct from both previously reported grape MLO sequences. Together, these results place *Vv*MLO3 within a phylogenetic group comprising monocot MLO homologs and distinguish it from the other grape *MLO* genes examined in this study.

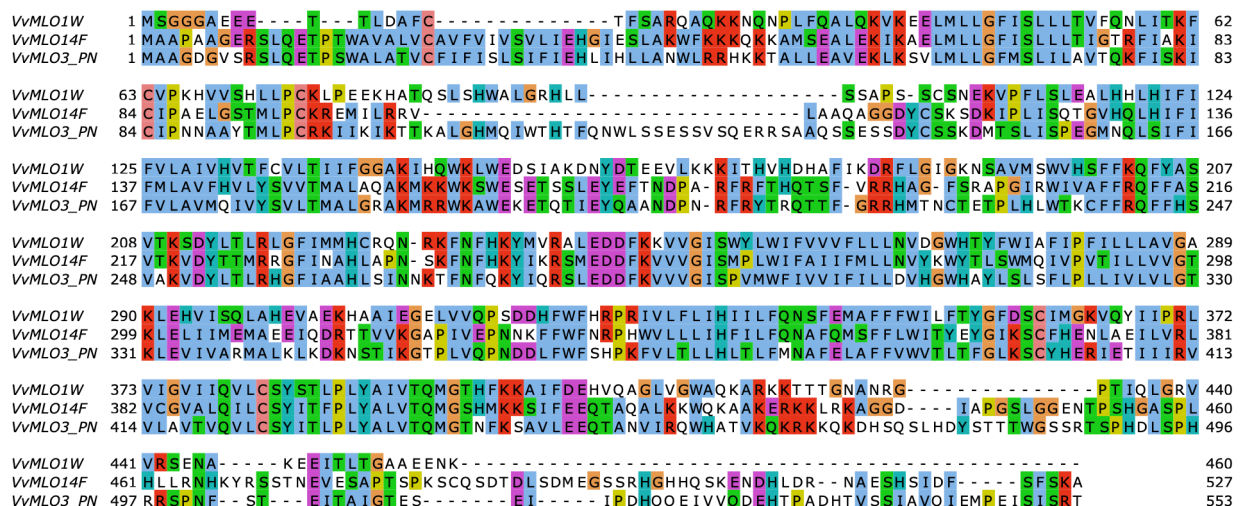


Figure 2. Multiple sequence alignment shows sequence divergence among *Vv*MLOs clustering in clade IV. Protein sequence alignment of *Vv*MLO1 (Winterhagen *et al.*, 2008) and *Vv*MLO14 (Feechan *et al.*, 2008), previously reported to cluster within clade IV and the PN40024 v5.1 *Vv*MLO3 protein sequence. ClustalX color scheme was applied based on residue conservation of 50%. Multiple sequence alignment was visualized using Jalview (Waterhouse *et al.*, 2009).

CRISPR analysis reveals distinct editing patterns across the three *VvMLO* target genes

To characterize the outcomes of CRISPR editing across the three *VvMLO* target genes, regenerated plants were screened for edits at the targeted sites and subsequently characterized based on their editing profiles. Following protoplast transfection with sgRNAs targeting the three *VvMLO* genes, 170 TS regenerated plants were screened for CRISPR editing. Of these plants, approximately 15% displayed edits in at least one of the targeted *VvMLO* genes. The edited plants were grouped into 16 independent events and screened for Cas9 integration by PCR. All the tested lines were negative for Cas9 integration, suggesting transient Cas9 expression. Among the 16 edited events, eight were selected for more detailed characterization (**Table 3**). Line 24-1 displayed large deletions and edits at all three *VvMLO* genes, as well as line 32-1. In contrast, line 45-1 exhibited minimal editing, with edits detected at one site in *VvMLO7* and one site in *VvMLO11*. Lines 40-1, 39-2, and 39-3 exhibited the same editing pattern and may represent clonal events. Lines 39-1 and 41-1 had distinct overall genotypes, although both lacked edits at site 1 of *VvMLO3* and *VvMLO7*. Across the edited lines, the three *VvMLO* genes displayed distinct editing patterns. In *VvMLO7*, edits predominantly consisted of T insertions (+T), which were typically homozygous at site 1 and heterozygous at site 2. In *VvMLO3*, three main types of edits were observed at similar frequencies, including single- and double-base deletions (-A and -TT) and T insertions (+T). *VvMLO11* exhibited the most complex editing pattern with a homozygous T insertion (+T) at site 1 and multiple deletions (-TCA), along with occasional insertions, at site 2. Overall, these results demonstrate successful editing at all three *VvMLO* target genes that led to the generation of a diverse set of *MLO*-edited lines differing in the number, type, and allelic state of mutations across the targeted loci.

Table 3. Summary of editing outcomes across the eight *MLO*-edited lines. For each edited line, the variant frequency (%), allelic profile and genotype description at each *MLO* gene and site (1 and 2) is reported. Variant frequency was calculated as the ratio of the number of reads supporting the variant and the total number of reads. The + and – sign of the allelic profile indicate insertion and deletion, respectively.

Genotype description was assigned based on the allelic profile detected at each edited gene/site (Hom = homozygous; Het = heterozygous; WT = wild type).

Site	Edited line	Target gene (<i>Vv</i>)	Variant frequency (%)		Allelic profile		Genotype description	
			1	2	1	2	1	2
24-1		<i>MLO3</i>	90	80	+A/-74	+T/-74	Hom	Hom
		<i>MLO7</i>	99	46.7	+T	-T/-GC	Hom	Biallelic
		<i>MLO11</i>	100	95.3	+T	+T	Hom	Hom
32-1		<i>MLO3</i>	98	46	-A	+T	Hom	Het
		<i>MLO7</i>	99.7	46.7	+T	+T	Hom	Het
		<i>MLO11</i>	59.7	44.7	+T/-TT	+A/A>T	Biallelic	Biallelic
39-1		<i>MLO3</i>	42	0	-A	-	Het	WT
		<i>MLO7</i>	51.7	0	+T	-	Het	WT
		<i>MLO11</i>	100	98.5	-T	+A	Hom	Hom
39-2		<i>MLO3</i>	98.6	95.8	TA>AT	-TT	Hom	Hom
		<i>MLO7</i>	99.7	50.3	+T	+T	Hom	Het
		<i>MLO11</i>	100	55.2	+T	-TCA/+T	Hom	Biallelic
39-3		<i>MLO3</i>	95.5	95.5	TA>AT	-TT	Hom	Hom
		<i>MLO7</i>	99.5	46.3	+T	+T	Hom	Het
		<i>MLO11</i>	100	50.1	+T	-TCA/+T	Hom	Biallelic
40-1		<i>MLO3</i>	90.9	88.9	TA>AT	-TT	Hom	Hom
		<i>MLO7</i>	100	54.5	+T	+T	Hom	Het
		<i>MLO11</i>	100	48.5	+T	-TCA/+T	Hom	Biallelic
41-1		<i>MLO3</i>	47.4	0	-A	-	Het	WT
		<i>MLO7</i>	46.5	0	+T	-	Het	WT
		<i>MLO11</i>	42	99.8	+T	+AA	Het	Hom
45-1		<i>MLO3</i>	0	0	-	-	WT	WT
		<i>MLO7</i>	45.7	0	+T	-	Het	WT
		<i>MLO11</i>	0	99.4	-	+A	WT	Hom

MLO-edited lines show genotype-dependent differences in PM susceptibility

To determine whether the modifications in the *MLO*-edited lines were associated with changes in PM susceptibility, disease progression was evaluated using both visual disease scoring and molecular quantification of *E. necator* accumulation in infected tissues. The eight *MLO*-edited lines and the susceptible non-edited control TS were inoculated with the *E. necator* strain UCD-TS1.

Disease progression through visual assessment of disease symptoms was evaluated in all the genotypes at 1, 5, and 10 dpi (**Figure 3A**). At 1 dpi, no clear differences in susceptibility scores were observed among genotypes, with all the lines showing uniform low levels of disease symptoms. By 5 dpi, disease symptoms became established, and variability among genotypes was more apparent. Although statistically significant differences were detected at this time point, the differences in susceptibility were modest, with substantial overlap among genotypes. At 10 dpi, differences in susceptibility among genotypes were more pronounced, allowing clearer discrimination between the susceptible control and the edited lines. TS and the minimally edited line 45-1 exhibited high susceptibility, characterized by dense mycelial growth and abundant conidiation. Similarly, lines 39-2, 39-3, and 40-1 showed susceptibility levels comparable to TS. In contrast, lines 41-1 and 32-1 showed significantly reduced susceptibility relative to TS, with limited disease progression characterized by the development of secondary hyphae and minimal conidiation. Lines 24-1 and 39-1 also exhibited significantly reduced susceptibility compared to TS but had intermediate phenotypes relative to the least susceptible lines. In these two lines, fungal colonization was moderate, with extended mycelial growth and limited conidiophore formation (**Figure 4**).

Because differences among genotypes were most evident at 10 dpi, pathogen accumulation was quantified at this time point. Molecular analysis was restricted to the four edited lines that showed reduced susceptibility relative to TS in the phenotypic assessment (24-1, 32-1, 39-1, and 41-1), with TS included as susceptible reference (**Figure 3B**). Analysis of variance performed on ΔC_t values revealed a marginal effect of genotype on fungal biomass accumulation (ANOVA, $F = 3.37$, $p = 0.06$). Although the genotype effect was only marginally significant, pairwise differences were explored using Tukey's HSD, which maintains the family-wise error rate regardless of the overall F-test results; these comparisons are interpreted as exploratory given the modest sample size. Pairwise comparisons were subsequently performed to evaluate differences among genotypes. Line 24-1 showed significantly lower fungal biomass accumulation than line 39-1 (adjusted $p < 0.05$). None of the edited lines differed significantly from TS, but line 24-1 showed the lowest pathogen accumulation among the analyzed genotypes. In contrast, line 39-1 accumulated the highest fungal biomass among the analyzed lines. Lines 32-1 and 41-1 accumulated intermediate levels of

fungal biomass comparable to TS, with 41-1 showing slightly lower pathogen accumulation than the control.

To assess the relationship between disease symptoms and pathogen accumulation, fungal biomass was correlated with visual disease scores at 10 dpi. A significant positive correlation was observed between the two measurements (Spearman's $\rho = 0.5$; $p < 0.05$; **Figure 3C**), indicating that genotypes with higher visual disease scores generally also exhibited higher *E. necator* biomass.

Together, these results demonstrated genotype-dependent differences in PM susceptibility among the *MLO*-edited lines, with four lines showing reduced disease symptoms relative to the control. The molecular assessment further showed variation in fungal biomass among these lines with the relationship between visual disease score and fungal biomass generally following a positive trend, but varying among individual genotypes.

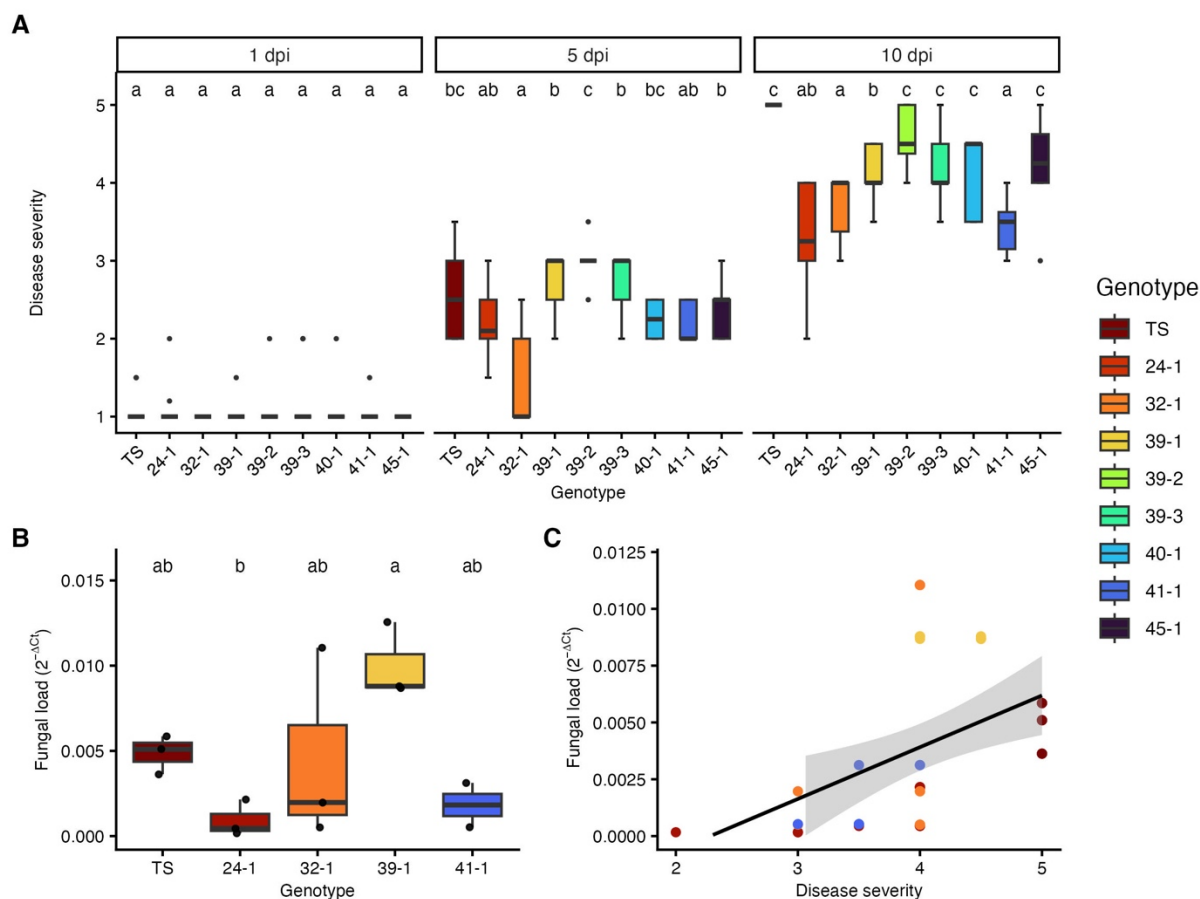


Figure 3. *MLO*-edited genotypes differ in PM susceptibility and fungal biomass accumulation. A) Disease susceptibility scores of *MLO*-edited grape genotypes and the susceptible non-edited control (TS) at 1, 5, and 10 days post inoculation (dpi) with *E. necator*. Disease severity was scored on a scale from 1 to 5, with higher scores indicating higher susceptibility. Boxplots show the distribution of visual susceptibility scores for each genotype at each time point; the center line represents the median, the box represents the interquartile range (IQR), and whiskers extend to 1.5 x IQR. Letters indicate statistically significant differences among genotypes within each time point according to Tukey HSD post-hoc test (adjusted $p < 0.05$). (B) Fungal biomass accumulation at 10 dpi in TS and the four selected *MLO*-edited genotypes, quantified by qPCR and expressed as relative fungal DNA accumulation ($2^{-\Delta Ct}$). Boxplots show the distribution of biological replicates for each genotype; letters indicate statistically significant differences among genotypes ($p < 0.05$). (C) Correlation between visual susceptibility score and fungal biomass at 10

dpi. Points represent individual biological replicates and are colored according to genotype. The solid line represents the fitted linear regression, with the shaded area indicating the 95% confidence interval. Spearman's rank correlation coefficient was used to assess the association between visual susceptibility scores and fungal biomass. The same genotype color scheme is used across all the panels.

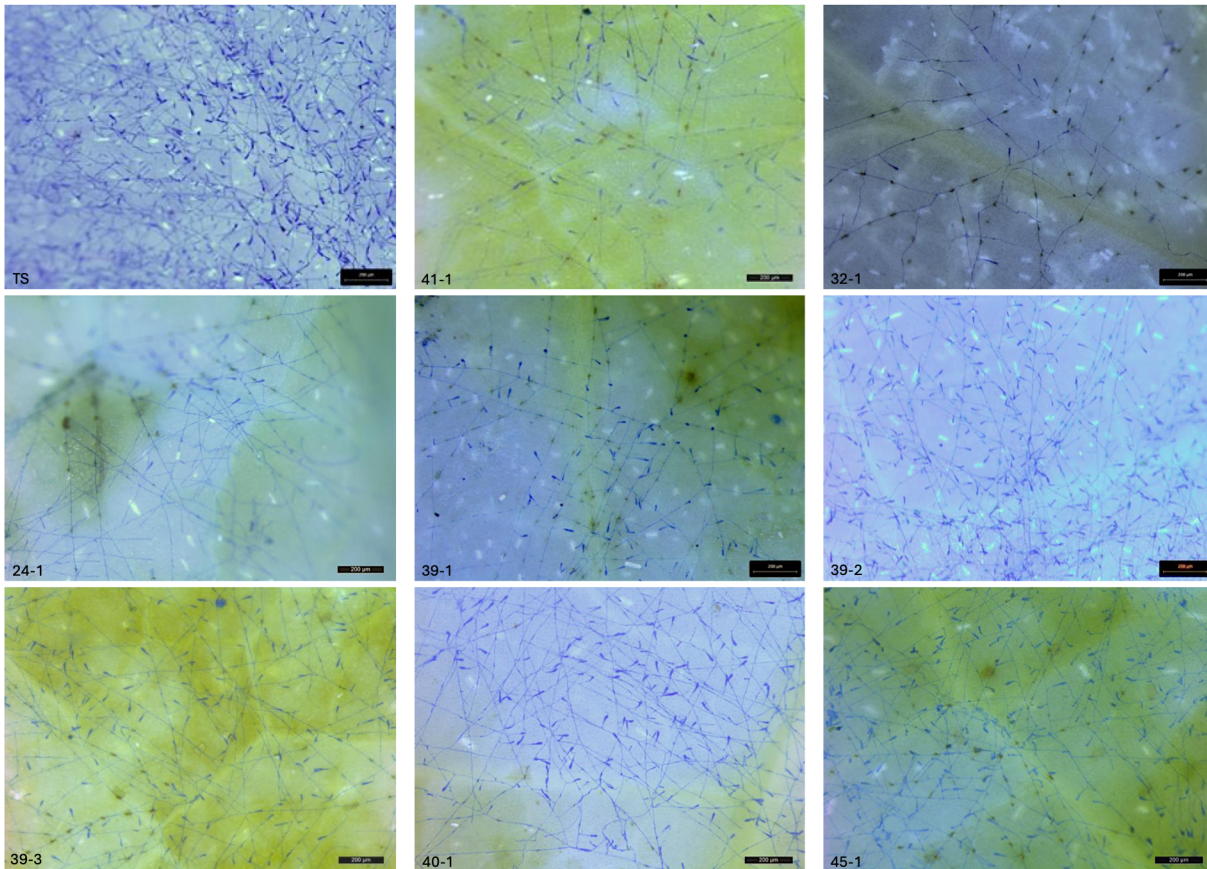


Figure 4. *MLO*-edited lines exhibit differences in *E. necator* development on leaf surface at 10 dpi.

Micrographs show *E. necator* development on leaves of the susceptible non-edited control (TS) and *MLO*-edited genotypes at 10 dpi following Coomassie Brilliant Blue staining. Conidiophores (blue club-shaped structures) and hyphae (blue filamentous structures) are visible on the leaf surface. Micrographs are arranged according to the extent of fungal development, from minimal hyphal growth in lines 41-1 and 32-1, to intermediate growth with several conidiophores in lines 24-1 and 39-1, and extensive mycelia coverage

and conidiation in lines 39-2, 39-3, 40-1, and 45-1. The susceptible non-edited control (TS) is shown in the first panel on the left. Scale bars = 0.2 mm.

Discussion

Editing of multiple loci to enhance disease resistance has become increasingly feasible with advances in CRISPR/Cas technologies (Li *et al.*, 2018; Singer *et al.*, 2024). In particular, the use of Cas9-RNP complexes in grapes protoplasts provides an effective approach for generating edited plants without stable integration of the editing construct (Malnoy *et al.*, 2016). Reduced PM susceptibility through knock-down or editing of multiple *MLO* genes has previously been reported in grapevine (Gouthu & Deluc, 2026; Moffa *et al.*, 2025; Pessina *et al.*, 2016). To our knowledge, this is the first study to simultaneously target *VvMLO3*, *VvMLO7*, and *VvMLO11* using RNP-mediated editing in grape protoplasts. The obtained edited lines carried diverse allelic combinations and exhibited different levels of susceptibility to *E. necator*.

Previous studies of grapevine *MLO* genes associated with PM susceptibility have primarily focused on genes belonging to clade V. Interestingly, *VvMLO3*, one of the genes targeted in this study, clustered within clade IV, which has traditionally been associated with monocot *MLO* homologs (Devoto *et al.*, 2003). *VvMLO3* has also previously been reported to be upregulated during PM infection (Weng *et al.*, 2014). In our study, lines carrying mutations in *VvMLO3* exhibited reduced susceptibility to the pathogen, providing evidence for a role of *VvMLO3* in the grape response to PM despite its unusual phylogenetic placement. This observation suggests that the canonical distinction between monocot- and dicot-associated *MLO* clades might not fully predict the susceptibility-associated function of individual grape *MLO* genes. This might be related to the early evolutionary diversification of the *MLO* family, which is thought to have preceded the divergence of monocots and dicots (Devoto *et al.*, 2003).

Among the three targeted genes, *VvMLO11* emerged as a potential contributor to the reduced susceptibility phenotype. All four lines showing reduced susceptibility relative to the control (24-1, 32-1, 39-1 and 41-1) carried mutations at both targeted sites of *VvMLO11*. Interestingly, line 41-1 showed one of the strongest

reductions in disease symptoms despite retaining wild type sequences at site 2 of both *VvMLO3* and *VvMLO7*. In this line, *VvMLO11* was the only gene for which both sites were edited.

Previous studies reported different conclusions regarding the contribution of *VvMLO11* to PM susceptibility. Pessina *et al.*, (2016) found that *VvMLO11* has a limited contribution to PM resistance when silenced in combination with other two *VvMLO* genes, whereas Wan *et al.*, (2020) reported enhanced resistance in *VvMLO11*-edited lines. Our results are more consistent with the latter observation; however, it is important to notice that different genetic background and editing strategy were used. Pessina *et al.* (2016) used the cultivar Brachetto and reduced *MLO* expression through RNA interference, whereas Wan *et al.* (2020) and the present study used Thomson Seedless and complete knock-out. These differences make direct comparison difficult and makes us hypothesize that the contribution of *VvMLO11* may depend on genetic background and/or the extent of gene disruption. Nevertheless, *VvMLO11* was edited together with *VvMLO3* and *VvMLO7* in this study and its individual contribution cannot be separated from those of the other targeted genes.

Although extensive homozygous or biallelic mutations were frequently associated with reduced susceptibility, the relationship between the extent of editing and disease phenotype was not linear. Lines 32-1 and 24-1, which carried mutations at both sites of all three targeted genes, exhibited strong reduction in susceptibility, while line 45-1, which contained fewer edits, remained highly susceptible and was comparable to TS. However, lines 39-2, 39-3 and 40-1 demonstrated that the number of edited sites alone cannot explain the observed phenotypes since their susceptibility remained comparable to the control. Notably, the mutation at *VvMLO3* site 1 in these lines was a nucleotide substitution rather than an insertion or a deletion, and this type of mutation may have a different impact on gene or protein function compared to the indels identified in the other lines. Similarly, homozygous or biallelic mutations do not necessarily result in complete loss of *MLO* function, as the effect of an edit depends on its type and location within the gene (Acevedo-Garcia *et al.*, 2017). The differences in susceptibility among edited lines may therefore reflect the type of mutation generated and potential functional redundancy among the targeted *MLO* genes. Other *MLO* genes may partially compensate for the loss of one gene, while mutations at multiple loci are

required to achieve a stronger reduction in susceptibility (Acevedo-Garcia *et al.*, 2014; Berman *et al.*, 2026). While the present results suggest that extensive editing increases the likelihood of reduced susceptibility, they also indicate that the relationship between genotype and disease phenotype is complex and cannot be predicted solely from the number of edited loci.

The positive correlation between the visual disease score and fungal biomass further supports an overall relationship between disease development and pathogen accumulation, although this relationship was not consistent across all genotypes. For example, line 39-1 exhibited significantly reduced visual susceptibility relative to TS while accumulating the highest fungal biomass among the genotypes included in the molecular analysis. Visual score primarily reflects the extent of visible fungal development and symptoms progression on the leaf surface, whereas the molecular assessment provides a measure of relative fungal DNA accumulation in the sampled tissue. The discrepancy observed in line 39-1 highlights the importance of using both phenotypic and molecular measurements when evaluating disease response in edited plants (Schnippenkoetter *et al.*, 2021). Additional experiments will be required to determine the individual contributions of *VvMLO3*, *VvMLO7* and *VvMLO11* to the reduced-susceptibility phenotype and to establish whether the mutations identified in these genes result in loss or alteration of *MLO* function.

Conclusions

The generation of *MLO*-edited grapevine lines with reduced susceptibility to *E. necator* demonstrates the potential of multiple CRISPR/RNP-mediated editing for improving PM resistance. Future studies should further investigate the molecular effects of the identified mutations through transcriptomic analysis. Generating independent *VvMLO3*, *VvMLO7* and *VvMLO11* mutants, as well as combinations of these and other *MLO* genes in the same genetic background, will help determine the individual and combined contribution to disease susceptibility. These approaches will also help identify the most effective *MLO* combination to achieve durable PM resistance.

References

- Acevedo-Garcia, J., Kusch, S., & Panstruga, R. (2014). Magical mystery tour: MLO proteins in plant immunity and beyond. *New Phytologist*, 204(2), 273–281. <https://doi.org/10.1111/nph.12889>
- Acevedo-Garcia, J., Spencer, D., Thierion, H., Reinstädler, A., Hammond-Kosack, K., Phillips, A. L., & Panstruga, R. (2017). Mlo-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach. *Plant Biotechnology Journal*, 15(3), 367–378. <https://doi.org/10.1111/pbi.12631>
- Amrine, K. C. H., Blanco-Ulate, B., Riaz, S., Pap, D., Jones, L., Figueroa-Balderas, R., Walker, M. A., & Cantu, D. (2015). Comparative transcriptomics of Central Asian *Vitis vinifera* accessions reveals distinct defense strategies against powdery mildew. *Horticulture Research*, 2, 15037. <https://doi.org/10.1038/hortres.2015.37>
- Appiano, M., Catalano, D., Santillán Martínez, M., Lotti, C., Zheng, Z., Visser, R. G. F., Ricciardi, L., Bai, Y., & Pavan, S. (2015). Monocot and dicot MLO powdery mildew susceptibility factors are functionally conserved in spite of the evolution of class-specific molecular features. *BMC Plant Biology*, 15(1), Article 1. <https://doi.org/10.1186/s12870-015-0639-6>
- Bai, Y., Pavan, S., Zheng, Z., Zappel, N. F., Reinstädler, A., Lotti, C., De Giovanni, C., Ricciardi, L., Lindhout, P., Visser, R., Theres, K., & Panstruga, R. (2008). Naturally occurring broad-spectrum powdery mildew resistance in a Central American tomato accession is caused by loss of mlo function. *Molecular Plant-Microbe Interactions: MPMI*, 21(1), 30–39. <https://doi.org/10.1094/MPMI-21-1-0030>
- Berman, A., Zylberberg, I., Mayrose, I., & Shani, E. (2026). Navigating genetic redundancy in plant genomes: Insights for research and breeding. *Trends in Plant Science*, 31(7), 961–973. <https://doi.org/10.1016/j.tplants.2026.02.004>
- Bonarota, M.-S., Garcia, J. F., Massonnet, M., Zaccheo, M., Figueroa-Balderas, R., Cochetel, N., & Cantu, D. (2025). Dual Single-Nucleus Gene Expression Atlas of Grapevine and Erysiphe necator During Early Powdery Mildew Infection. *Molecular Plant-Microbe Interactions*, 38(6), 1016–1030. <https://doi.org/10.1094/MPMI-08-25-0099-R>
- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17(4), 540–552. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>
- Cochetel, N., & Cantu, D. (2024, January). *Grapegenomics.com: A web portal with genomic data and analysis tools for wild and cultivated grapevines* (Version 2.0). Zenodo. <https://doi.org/10.5281/zenodo.10570120>
- Consonni, C., Humphry, M. E., Hartmann, H. A., Livaja, M., Durner, J., Westphal, L., Vogel, J., Lipka, V., Kemmerling, B., Schulze-Lefert, P., Somerville, S. C., & Panstruga, R. (2006). Conserved requirement for a plant host cell protein in powdery mildew pathogenesis. *Nature Genetics*, 38(6), 716–720. <https://doi.org/10.1038/ng1806>
- Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Molecular Biology and Evolution*, 37(1), 291–294. <https://doi.org/10.1093/molbev/msz189>

- Devoto, A., Hartmann, H. A., Piffanelli, P., Elliott, C., Simmons, C., Taramino, G., Goh, C.-S., Cohen, F. E., Emerson, B. C., Schulze-Lefert, P., & Panstruga, R. (2003). Molecular Phylogeny and Evolution of the Plant-Specific Seven-Transmembrane MLO Family. *Journal of Molecular Evolution*, 56(1), 77–88. <https://doi.org/10.1007/s00239-002-2382-5>
- Dong, O. X., & Ronald, P. C. (2019). Genetic Engineering for Disease Resistance in Plants: Recent Progress and Future Perspectives[OPEN]. *Plant Physiology*, 180(1), 26–38. <https://doi.org/10.1104/pp.18.01224>
- Elliott, C., Zhou, F., Spielmeyer, W., Panstruga, R., & Schulze-Lefert, P. (2002). Functional Conservation of Wheat and Rice Mlo Orthologs in Defense Modulation to the Powdery Mildew Fungus. *Molecular Plant-Microbe Interactions*®, 15(10), 1069–1077. <https://doi.org/10.1094/MPMI.2002.15.10.1069>
- Feechan, A., Jermakow, A. M., Torregrosa, L., Panstruga, R., & Dry, I. B. (2008). Identification of grapevine MLO gene candidates involved in susceptibility to powdery mildew. *Functional Plant Biology*, 35(12), 1255. <https://doi.org/10.1071/FP08173>
- Feechan, A., Kabbara, S., & Dry, I. B. (2011). Mechanisms of powdery mildew resistance in the Vitaceae family. *Molecular Plant Pathology*, 12(3), 263–274. <https://doi.org/10.1111/j.1364-3703.2010.00668.x>
- Gouthu, S., & Deluc, L. G. (2026). CRISPR-Cas9 editing of grapevine clade V-related mlo genes revealed new insights into the contribution of each mlo gene to grapevine powdery mildew resistance (*Erysiphe necator*) and plant development. *bioRxiv*. <https://doi.org/10.64898/2026.01.29.702653>
- Jones, L., Riaz, S., Morales-Cruz, A., Amrine, K. C., McGuire, B., Gubler, W. D., Walker, M. A., & Cantu, D. (2014). Adaptive genomic structural variation in the grape powdery mildew pathogen, *Erysiphe necator*. *BMC Genomics*, 15(1), 1081. <https://doi.org/10.1186/1471-2164-15-1081>
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kieu, N. P., Lenman, M., Wang, E. S., Petersen, B. L., & Andreasson, E. (2021). Mutations introduced in susceptibility genes through CRISPR/Cas9 genome editing confer increased late blight resistance in potatoes. *Scientific Reports*, 11(1), 4487. <https://doi.org/10.1038/s41598-021-83972-w>
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAXML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics (Oxford, England)*, 35(21), 4453–4455. <https://doi.org/10.1093/bioinformatics/btz305>
- Kusch, S., & Panstruga, R. (2017). mlo-Based Resistance: An Apparently Universal “Weapon” to Defeat Powdery Mildew Disease. *Molecular Plant-Microbe Interactions*®, 30(3), 179–189. <https://doi.org/10.1094/MPMI-12-16-0255-CR>
- Li, C., Hao, M., Wang, W., Wang, H., Chen, F., Chu, W., Zhang, B., Mei, D., Cheng, H., & Hu, Q. (2018). An Efficient CRISPR/Cas9 Platform for Rapidly Generating Simultaneous Mutagenesis of Multiple Gene Homoeologs in Allotetraploid Oilseed Rape. *Frontiers in Plant Science*, 9. <https://doi.org/10.3389/fpls.2018.00442>

- Malnoy, M., Viola, R., Jung, M.-H., Koo, O.-J., Kim, S., Kim, J.-S., Velasco, R., & Nagamangala Kanchiswamy, C. (2016). DNA-Free Genetically Edited Grapevine and Apple Protoplast Using CRISPR/Cas9 Ribonucleoproteins. *Frontiers in Plant Science*, 7. <https://doi.org/10.3389/fpls.2016.01904>
- Moffa, L., Mannino, G., Bevilacqua, I., Gambino, G., Perrone, I., Pagliarani, C., Berteà, C. M., Spada, A., Narduzzo, A., Zizzamia, E., Velasco, R., Chitarra, W., & Nerva, L. (2024). CRISPR/Cas9-driven double modification of grapevine MLO6-7 imparts powdery mildew resistance, while editing of NPR3 augments powdery and downy mildew tolerance. *The Plant Journal*, n/a(n/a). <https://doi.org/10.1111/tpj.17204>
- Moffa, L., Mannino, G., Bevilacqua, I., Gambino, G., Perrone, I., Pagliarani, C., Berteà, C. M., Spada, A., Narduzzo, A., Zizzamia, E., Velasco, R., Chitarra, W., & Nerva, L. (2025). CRISPR/Cas9-driven double modification of grapevine MLO6-7 imparts powdery mildew resistance, while editing of NPR3 augments powdery and downy mildew tolerance. *The Plant Journal: For Cell and Molecular Biology*, 122(2), e17204. <https://doi.org/10.1111/tpj.17204>
- Pavan, S., Schiavulli, A., Appiano, M., Marcotrigiano, A. R., Cillo, F., Visser, R. G. F., Bai, Y., Lotti, C., & Ricciardi, L. (2011). Pea powdery mildew er1 resistance is associated to loss-of-function mutations at a MLO homologous locus. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, 123(8), 1425–1431. <https://doi.org/10.1007/s00122-011-1677-6>
- Pessina, S., Lenzi, L., Perazzolli, M., Campa, M., Dalla Costa, L., Urso, S., Valè, G., Salamini, F., Velasco, R., & Malnoy, M. (2016). Knockdown of MLO genes reduces susceptibility to powdery mildew in grapevine. *Horticulture Research*, 3, 16016. <https://doi.org/10.1038/hortres.2016.16>
- Pessina, S., Pavan, S., Catalano, D., Gallotta, A., Visser, R. G., Bai, Y., Malnoy, M., & Schouten, H. J. (2014). Characterization of the MLO gene family in Rosaceae and gene expression analysis in *Malus domestica*. *BMC Genomics*, 15(1), 618. <https://doi.org/10.1186/1471-2164-15-618>
- Riaz, S., Braun, U., Lejkina, I., Gubler, W. D., & Walker, M. A. (2013). Report of a new grape powdery mildew morphotype with branched conidiophores. *Plant Pathology & Quarantine*, 3(1), 19–27. <https://doi.org/10.5943/ppq/3/1/4>
- Sambucci, O., Alston, J. M., Fuller, K. B., & Lusk, J. (2019). The Pecuniary and Nonpecuniary Costs of Powdery Mildew and the Potential Value of Resistant Grape Varieties in California. *American Journal of Enology and Viticulture*, 70(2), 177–187. <https://doi.org/10.5344/ajev.2018.18032>
- Schnippenkoetter, W., Hoque, M., Maher, R., Van de Wouw, A., Hands, P., Rolland, V., Barrett, L., & Sprague, S. (2021). Comparison of non-subjective relative fungal biomass measurements to quantify the *Leptosphaeria maculans*—*Brassica napus* interaction. *Plant Methods*, 17(1), 122. <https://doi.org/10.1186/s13007-021-00822-6>
- Scintilla, S., Salvagnin, U., Giacomelli, L., Zeilmaker, T., Malnoy, M. A., Rouppe van der Voort, J., & Moser, C. (2022). Regeneration of non-chimeric plants from DNA-free edited grapevine protoplasts. *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.1078931>
- Shi, X., Cao, S., Wang, X., Huang, S., Wang, Y., Liu, Z., Liu, W., Leng, X., Peng, Y., Wang, N., Wang, Y., Ma, Z., Xu, X., Zhang, F., Xue, H., Zhong, H., Wang, Y., Zhang, K., Velt, A., ... Zhou, Y. (2023). The

complete reference genome for grapevine (*Vitis vinifera* L.) genetics and breeding. *Horticulture Research*, 10(5), uhad061. <https://doi.org/10.1093/hr/uhad061>

Singer, S. D., Mukhtar, M. M., Subedi, U., Poudel, H., Chen, G., Foroud, N., & Chatterton, S. (2024). CRISPR/Cas-Mediated Gene Editing in Plant Immunity and Its Potential for the Future Development of Fungal, Oomycete, and Bacterial Pathogen-Resistant Pulse Crops. *Plant, Cell & Environment*, n/a(n/a). <https://doi.org/10.1111/pce.15174>

Tricoli, D. M., & Debernardi, J. M. (2024). An efficient protoplast-based genome editing protocol for *Vitis* species. *Horticulture Research*, 11(1), uhad266. <https://doi.org/10.1093/hr/uhad266>

Wan, D.-Y., Guo, Y., Cheng, Y., Hu, Y., Xiao, S., Wang, Y., & Wen, Y.-Q. (2020). CRISPR/Cas9-mediated mutagenesis of VvMLO3 results in enhanced resistance to powdery mildew in grapevine (*Vitis vinifera*). *Horticulture Research*, 7(1), 1–14. <https://doi.org/10.1038/s41438-020-0339-8>

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., & Qiu, J.-L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology*, 32(9), 947–951. <https://doi.org/10.1038/nbt.2969>

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9), 1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>

Weng, K., Li, Z.-Q., Liu, R.-Q., Wang, L., Wang, Y.-J., & Xu, Y. (2014). Transcriptome of Erysiphe necator-infected *Vitis pseudoreticulata* leaves provides insight into grapevine resistance to powdery mildew. *Horticulture Research*, 1, 14049. <https://doi.org/10.1038/hortres.2014.49>

Winterhagen, P., Howard, S. F., Qiu, W., & Kovács, L. G. (2008). Transcriptional Up-Regulation of Grapevine *MLO* Genes in Response to Powdery Mildew Infection. *American Journal of Enology and Viticulture*, 59(2), 159–168. <https://doi.org/10.5344/ajev.2008.59.2.159>

Zaidi, S. S.-A., Mukhtar, M. S., & Mansoor, S. (2018). Genome Editing: Targeting Susceptibility Genes for Plant Disease Resistance. *Trends in Biotechnology*, 36(9), 898–906. <https://doi.org/10.1016/j.tibtech.2018.04.005>

Zhang, J. (2022). Check CRISPR Editing Events in Transgenic Wheat with Next-Generation Sequencing. In S. H. Wani & A. Kumar (Eds.), *Genomics of Cereal Crops* (pp. 95–106). Springer US. https://doi.org/10.1007/978-1-0716-2533-0_4

Zheng, Z., Nonomura, T., Appiano, M., Pavan, S., Matsuda, Y., Toyoda, H., Wolters, A.-M. A., Visser, R. G. F., & Bai, Y. (2013). Loss of function in *Mlo* orthologs reduces susceptibility of pepper and tomato to powdery mildew disease caused by *Leveillula taurica*. *PloS One*, 8(7), e70723. <https://doi.org/10.1371/journal.pone.0070723>