Complete genomes of grapevine downy mildew reveal effector cluster evolution driven by complex structural variations

Lianzhu Zhou,[1] Shaowei Cui[1], Hao Zhang,[1] Fanfang Kong,[1] Qi Wang,[2] Zhongyue Wang,[1] Yongfeng Zhou,[3] Shidong Li,[1] Fei Du,[4] Xiaoqing Huang[1]*, Yongqiang Liu[1]*

[1]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China

[2]Department of Plant Pathology, College of Plant Protection, China Agricultural University, Beijing, China

[3]National Key Laboratory of Tropical Crop Breeding, Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, China

[4]Department of Plant Pathology, College of Plant Protection, Yunnan Agricultural University, Kunming, China

*Corresponding authors: XQ Huang, E-mail: huangxiaoqing0718@126.com; YQ Liu, E-mail: liuyongqiang@caas.cn

Abstract

Plasmopara viticola, the causal agent of grapevine downy mildew, exhibits substantial intraspecific variation in pathogenicity and genetic diversity, yet the genomic features underlying this variation remain incompletely characterized. Here, we sequenced and assembled two *P. viticola* isolates, PvH (from *Vitis vinifera*) and PvS (from *V. amurensis*), using PacBio HiFi sequencing, and performed comparative genomic analysis. Two complete genome assemblies (17 chromosomes) of *P. viticola* (PvH: 115.3 Mb; PvS: 113.0 Mb) were generated and revealed that nearly 90% of the putative effectors exist as local duplicated gene clusters. Comparative genomics uncovered distinct intraspecific expansion, deletion and diversification of putative effectors driven by local segmental, tandem, and proximal duplication events in *P. viticola*. Specifically, PvH exhibited a ~1.4-fold increase in CRNs (PvH: 237; PvS: 183; PV221: 169) and harbored 35 strain-specific CRNs. These differential effectors were predominantly clustered in complex structural variation hotspots (SVs, duplication and inversion). Notably, 104 putative effectors—including 21 RxLRs, 59 CRNs, and 24 CAZymes—located within inversion regions. Together, our results highlight a highly dynamic genome architecture in *P. viticola*, in which SV and local gene duplication are closely associated with effector diversification. This study provides a genome-resolved comparative framework for understanding intraspecific genomic diversity in *P. viticola* and establishes a foundation for future population-level and functional investigations.

Key words: *Plasmopara viticola*, complete genome assembly, structural variations, effector expansion, adaptive evolution

Introduction

Downy mildew is one of the most significant diseases affecting cultivated grapevines (*Vitis vinifera* L.) globally (Gessler et al., 2011). The pathogen responsible for this disease is *Plasmopara viticola*, an obligate biotrophic and heterothallic diploid oomycete. *P. viticola* demonstrates a strong adaptive potential, not only through the rapid development of fungicide resistance (Delmas et al., 2017), but also through the repeated breakdown of grapevine resistance loci, including *Rpv3.1*, *Rpv10*, *Rpv12*, and more recently *Rpv1* (Peressotti et al., 2010; Delmotte et al., 2014; Wingerter et al., 2021; Paineau et al., 2022; Martínez et al., 2025). Recent studies have revealed that the five formae speciales (f. sp.) of *P. viticola* are found on wild *Vitis* species across North America (Rouxel et al., 2013; 2014), and that populations invading vineyards worldwide all originate from only one of the five native North American lineages (*P. viticola* f. sp. *aestivalis*) (Fontaine et al., 2021). However, *P. viticola* populations within China exhibit high levels of genetic and genotypic diversity (Yin et al., 2014; Zhang et al., 2017), with most of this genetic variation occurring among isolates on different *Vitis* species (Yin et al., 2014). Notably, our recent studies revealed significant differences in pathogenicity and sporangia size between *P. viticola* isolates from *V. vinifera* (a widely cultivated Eurasian species) and *V. amurensis* (a wild *Vitis* species endemic to East Asia) in China (Huang 2020). All these observations suggest that although the *P. viticola* populations within China belong to the same species-corresponding to the *P. viticola* f. sp. *aestivalis* lineage (Fontaine et al., 2021)—this pathogen has likely undergone differential adaptation to *V. vinifera* and *V. amurensis* in China. However, the genetic basis underlying this differential adaptation remains largely unclear.

High-quality genome analysis of pathogens is essential for uncovering the molecular basis of their adaptive evolution. However, the obligate biotrophy of *P. viticola* complicates the acquisition of large quantities of purified hyphae, which greatly restricts research on its genome and genetic evolution. Although several assemblies of the *P. viticola* genome have been released using Illumina short-read sequencing and PacBio long-read sequencing technologies (Brilli et al., 2018, Dussert et al., 2016,

Dussert et al., 2019, Yin et al., 2017), including recently released pseudo-diploid assembly PV221 (Paineau et al., 2024) and the chromosome-level diploid genome Pv1419, (Dvořák et al., 2025a), these assemblies remain incomplete. This deficiency is particularly pronounced in complex regions characterized by highly repetitive genomic content. These regions often serve as hotspots for genomic rearrangements, facilitating gene innovation and rapid gain-of-function adaptations (Long et al., 2024; Wang et al., 2024). For instance, genome segmental duplications can influence genome structure and drive the formation of gene cluster—a group of genes that are located in close physical proximity on the genome and typically share similar or functionally related roles—which tend to collapse in fragmented genome assemblies (Chen et al., 2018; Li et al., 2021; Müller et al., 2019). Long-read sequencing technology, particularly high-fidelity (HiFi) sequencing using circular consensus sequencing (CCS), significantly enhances the continuity of genome assembly. The CCS algorithm combines subreads from a single polymerase read to generate high-quality consensus reads. This approach minimizes errors associated with overlapping and error correction, especially in repeat regions, thereby enabling the generation of high-quality genome assemblies (Li et al., 2021; Shi et al., 2023).

Oomycetes usually deploy several classes of effectors to manipulate host immunity. Among them, RxLR effectors (containing the conserved RxLR motif) are translocated into host cells and can suppress immune signaling or alter host cellular processes. CRN (Crinkler) effectors represent another class of intracellular effectors that often induce cell death or tissue collapse in host plants. In addition, pathogens also encode a large repertoire of CAZymes (Carbohydrate-Active enZymes), which function in modifying or degrading plant cell wall polysaccharides, thereby facilitating host tissue colonization and nutrient acquisition (Lo Presti and Kahmann, 2017). In the host-pathogen arms race, effectors rapidly evolve to evade plant immunity. Genomic studies have shown that effector genes in filamentous pathogens are often clustered in genomic regions rich in repetitive sequences, accelerating their evolution through duplication events (Dong et al., 2015, Müller et al., 2019, Wang et al., 2017). For instance, in *Phytophthora betacei*, *Phytophthora megakarya*, and *Phytophthora palmivora*, whole-

genome duplication (WGD) events and transposon invasions have resulted in a larger effector repertoire compared to other *Phytophthora* species (Ayala-Usma et al., 2021, Morales-Cruz et al., 2020). Similarly, chromosome segmental duplications and inversions have been shown to promote the rapid emergence of new genes that contribute to host adaptation (Chen et al., 2018, de Jonge et al., 2013). Similar patterns have been reported in the fungal pathogens *Blumeria graminis* f. sp. *tritici* and *Zymoseptoria tritici*, where effector gene duplication has generated new allelic variants that accelerate pathogen adaptation and evolution (Hartmann and Croll, 2017; Müller et al., 2019). In *P. viticola*, numerous predicted effector genes have been identified based on genomic and transcriptomic analyses (Brilli et al., 2018; Dussert et al., 2019; Yin et al., 2015; Xiang et al., 2021; Xiang et al., 2022; Yin et al., 2019; Yin et al., 2022). However, only a limited number of effectors have been functionally characterized, and even fewer have been shown to interact directly with grapevine resistance loci (Dvořák et al., 2025a, Dvořák et al., 2025b, Paineau et al., 2024). Consequently, although nucleotide polymorphisms among effector candidates have been reported (Xiang et al., 2021), the mechanisms driving effector diversification and host-specific recognition remain unresolved, partly due to the lack of complete and comparative genome assemblies until recently.

The objectives of this study are to: (1) construct high-quality genome assemblies of *P. viticola* isolates from Eurasian *V. vinifera* (PvH) and wild *V. amurensis* (PvS) using PacBio HiFi sequencing technology; (2) describe the overall genome architecture and characterize the number and distribution patterns of effectors on the chromosomal landscape; (3) perform comparative genomic analyses to document patterns of genomic and effector variation among *P. viticola* isolates; and (4) investigate genomic features, including gene duplication and structural variation, that are associated with effector diversification and turnover. Collectively, this work provides a genome-resolved comparative framework for understanding intraspecific genomic diversity in *P. viticola* and establishes a foundation for future population-level and functional studies aimed at dissecting the genetic determinants of pathogenicity variation.

Results

Complete genome assemblies and annotations of two *P. viticola* isolates

The genomes of two *P. viticola* isolates from the Eurasian cultivated species *V. vinifera* cv. 'Red Globe' (PvH) and the East Asian wild species *V. amurensis* cv. 'Shuanghong' (PvS) were sequenced and were assembled into 17 chromosomes using 7.9 Gb and 9.6 Gb HiFi reads from the PacBio Sequel II System, respectively (Table 1, Fig. 1). The average sequencing depth of these two genomes is appropriately 68× and 83×, respectively. The high quality of the sequencing data supported the generation of complete chromosome-level assemblies (Table S1). The final genome of PvH was assembled into 17 contigs corresponding to 17 chromosomes, 14 of which are complete and gapless telomere-to-telomere (T2T) chromosome sequences that harbor oomycete-specific telomeric repeats (C3TA3) at both extremities, underlining the complete assembly of chromosome ends. The remaining three contigs (Chr5, Chr14 and Chr17) contained a single telomere, two (Chr14 and Chr17) of which are missing ~10 Kb sequences in the telomeric region at the other end of the chromosome and the other (Chr5) terminating in a repetitive rDNA sequence (Fig. 1; Table S2). Chromosome assembly length ranged from 4.3 Mb (Chr17) to 14.9 Mb (Chr2), resulting in a total length of 115.3 Mb with a contig N_{50} size of 7.0 Mb (Table 1 and Table S2). The final genome of PvS was assembled into 18 contigs corresponding to 17 chromosomes. Among them, 15 chromosomes are complete and gapless T2T chromosomes; Chr12 contains a single gap that spans the nucleolar organizer region (NOR) between two contigs; and the remaining one contig (Chr5) terminates with a repetitive rDNA sequence. Chromosome assembly length ranged from 4.2 Mb (Chr17) to 14.5 Mb (Chr2), resulting in a total length of 113.0 Mb with a contig N_{50} size of 6.3 Mb (Table 1 and Table S2). The GC content of the genome assemblies for the two *P. viticola* isolates (PvH and PvS) was 44.8%, with chromosome-specific values ranging from 44.1% to 46.1% for PvH and from 44.2% to 46.0% for PvS (Table S2). Putative centromeric regions were identified on all chromosomes in the PvH and PvS assemblies

(Fig. 1; Table S2), which were absent from the previously released genome versions.

A high proportion of repeats was annotated in both assemblies, with 64.3 Mb (55.8%) in PvH and 62.2 Mb (55.1%) in PvS (Table 1). Of these repeats, nearly 23% were annotated as transposable elements (TEs), including approximately 20% retrotransposons and 2.5% DNA transposons (Fig. 2A; Table S3). The proportion of repeats in the new assemblies is approximately 15% to 20% higher than in previous versions (e.g., JL-7-2 = 25.6%; PV221 = 37.3%), suggesting that complex repeat regions have been more comprehensively assembled in these updated genomes (Table S4). The heterozygosity rate for PvH and PvS is 0.892% and 0.837%, respectively, which is slightly higher than the PV221 heterozygosity rate of 0.795% (Table S4). In the PvH assembly, a total of 14,472 protein-coding genes were predicted, while the PvS assembly predicted 14,431 protein-coding genes. Approximately 64% of all annotated gene models were supported by homologies with known protein databases or functional domains (Table 1).

The accuracy and completeness of the new assemblies were validated in multiple aspects. High BUSCO scores indicated strong completeness, with 96.5% (165 of 171) and 100% (100 of 100) complete single-copy orthologs detected at the alveolate (n = 171) and stramenopiles (n = 100) levels, respectively (Table S5). Additionally, the assemblies demonstrated high mapping rates and coverages for both HiFi reads (mapping rate > 98%, coverage > 99%) and Illumina reads (mapping rate > 97%, coverage > 99%), along with a high degree of synteny with the recently released pseudo-diploid assembly PV221, and the chromosome-level diploid genome Pv1419, as well as with other high-quality oomycete genomes (Table S5; Figs. S1A, B; S2C). Importantly, the genome sizes of PvH and PvS closely matched the estimated genome size of 115 Mb, previously determined by Feulgen staining (Voglmayr and Greilhuber, 1998), further supporting the conclusion that the genomes assembled in this study were highly complete and accurate (Table S4).

Effector gene clusters on chromosome landscape

Oomycete pathogens secreted effectors to facilitate successful host colonization

and growth. In the PvH and PvS genomes, 1012 and 1026 putative effector genes, including encoding candidate cytoplasmic RxLR and CRN effectors and apoplastic effectors, such as CAZymes, Elicitins, CAPs, NLPs, Transglutaminase, and PAN/Apple were identified, respectively (Table S6). For a comparison, we reannotated the previous genome *P. viticola* INRA-PV221 (PV221) (Dussert et al., 2019), using the same method and parameters as in this study. A near-identical number of apoplastic effectors (except for CAZymes) but different cytoplasmic effectors were found among these three assemblies (PvH, PvS and PV221). The number of putative RxLR effectors in PvH (405) is lower than those in PvS (419) and PV221 (445), whereas PvH (379) has more CRN effectors than those in PvS (240) and PV221 (219) (Table S6).

We then analyzed the distribution of putative effector genes on each chromosome and found that the majority of effector genes were clustered on specific chromosome regions rather than randomly distributed (Fig. 2B, C; Table S7). For example, most of the RxLRs were clustered on Chr7, Chr16, Chr15, Chr8 and Chr14, while the CRNs were concentrated on Chr3, Chr4, Chr10 and Chr14, and the CAZymes on chromosomes Chr1, Chr4, Chr6 and Chr11, which may play a major role in the expansion and evolution of effectors. In addition, about twice as many CRNs were found on Chr4, Chr10, and Chr14 in PvH as in PvS (Table S7), suggesting that expansion of the CRN effector gene family may have occurred in PvH.

We observed that genomic regions enriched in effector clusters exhibited higher TE density compared to the genomic background (Fig. 2C), suggesting that TEs may contribute to local genomic plasticity. To further assess the association between TEs and effector diversification, we quantified the distances between effector genes and annotated TEs. CRN effectors showed significant enrichment of TEs overlapping or located within 2 kb, relative to non-effector genes (Fig. S3). These patterns are consistent with TE-associated genome remodeling reported in other oomycetes (Dong et al., 2015; Dussert et al., 2019), indicating that TEs may facilitate effector turnover and expansion within *P. viticola*.

Comparative genomics analysis reveals the diversification of pathogenicity-related

genes among *P. viticola* isolates

A total of 21,437 orthogroups were identified by clustering the predicted gene models from 18 oomycete species with OrthoFinder (Emms and Kelly, 2019). A phylogenetic tree was constructed for 18 species, using 206 single-copy genes (Fig. S4). Phylogenetic analysis confirms that PvH, PvS, and PV221 form a genetically cohesive cluster. However, PvH displays a slight degree of separation (Fig. S4), indicating potential intraspecific variation within this taxon. The clustering results of the predictive gene models for the three *P. viticola* assemblies yielded 11,144 (88.0%) common gene families, 1,395 (11.0%) dispensable gene families, and 127 (1.0%) unique gene families, respectively (Fig. 3A). Among them, PvS and PV221 shared the highest number of gene families (676); while PV221 (52) and PvH (49) had almost the same number of private gene families, which was about twice as many as the number of private gene families in PvS (26) (Fig. 3A). Functional enrichment analyses of those private genes across the three genomes revealed distinct GO terms (Fig. 3B). Private genes in PvH and PvS were associated with fewer GO terms, primarily enriched for transport and localization processes. In contrast, PV221 private genes were enriched in a broader range of functional categories, including protease and hydrolase activities as well as hydrolysis, metabolism, and transport functions (Fig. 3B).

The analysis of effector gene family clustering among the three *P. viticola* assemblies indicates significant differences in expansion and evolutionary patterns, particularly in the context of the PvH assembly. Notably, PvH contains more core and dispensable effector genes than the other isolates (Fig. 3C, Table S6), with a substantial expansion of CRN effectors (344 in PvH, compared to 240 in PvS and 211 in PV221). When considering isolate-specific effectors, PvH and PV221 each carry a similar number of private effector genes (44 and 47, respectively), whereas PvS contains none (Fig. 3C, Table S6). Furthermore, PvH displays a larger set of isolate-specific CRN effectors (35), while PV221 has more isolate-specific CAZymes (8) (Fig. 3C, Table S6). These patterns highlight intraspecific variation in effector repertoires across isolates.

Phylogenetic analyses of the first six core CRN gene families across the three *P.*

viticola assemblies reveal distinct patterns of expansion, particularly in PvH and PvS. The most pronounced expansions are observed in PvH (Fig. 3D), as illustrated by the largest CRN gene family (CRN-family1), which contains 37 and 36 CRN genes in PvH and PvS, respectively, while PV221 has only one CRN in this family (Fig. 3E). This data underscores the evolutionary dynamics within these genomes, with PvH exhibiting the most significant expansions in CRN gene families.

Local gene duplications facilitate the expansion of pathogenicity-related genes, especially effectors

To explore the possible causes of gene family expansion (i.e., an increase in the total copy number) in the *P. viticola* genomes, the assemblies of these three *P. viticola* isolates and two closely related downy mildew species with high-quality genome assemblies (*B. lactucae* and *P. effusa*) were selected for further analysis. Compared to other downy mildew species, the genomes of *P. viticola* are characterized by a relatively large genome size and a remarkably huge number of genes (Fig. 4A). A total of 8999 (62.8%), 9944 (66.0%), 9044 (62.5%), 5799 (59.4%), and 5338 (55.2%) duplicated genes were identified in PvS, PV221, PvH, *B. lactucae*, and *P. effusa*, respectively (Table S8). In all three *P. viticola* genomes, approximately 7%~10% of the genes were the result of duplicated gene blocks (DGB, containing at least five collinear genes), 10%~12% of the genes were from tandem duplication (TD, two adjacent similar genes on the same chromosome), around 10% were due to proximal duplication (PD, two similar genes with intervals of less than 20 other genes on the same chromosome), and more than one-third of the total genes were from dispersed duplication (DSD, two similar genes with intervals of more than 20 other genes) and single-copy (SL), respectively (Fig. 4B). Compared to *B. lactucae* and *P. effusa*, the genes in *P. viticola* (PvS, PV221 and PvH) are characterized by a relatively higher proportion of genes derived from DGB, TD, and PD, and a relatively lower proportion of DSD and SL genes (Fig. 4B). Smaller Ks values and higher Ka/Ks ratios were observed for gene pairs derived from DGB, TD and PD in *P. viticola* and *P. effusa* (Figs. 4C; S5A, C, D), indicating that the gene duplications via DGB, TD and PD represent recent processes

and are under stronger positive selection. Furthermore, we observe that the overall expression levels of these fast-evolving genes (under positive selection, Ka/Ks > 1) are significantly lower than those of slow-evolving genes (Fig. S5B). The lower expression levels of fast-evolving genes suggest that these duplicates may be in the early stages of functional divergence. Collectively, these patterns imply that recent gene duplication events contributed to gene expansion, particularly effector genes in *P. viticola*.

To comprehend the impact of these gene duplication events on the evolution of *P. viticola* isolates, a screen for significantly expanded gene families (SEGFs) was carried out. Eventually, 54 gene families were identified in PvH, 25 in PvS, and 101 in PV221. Further, we discovered that the majority of the SEGFs in the three *P. viticola* genomes are isolate-specific. Specifically, only four SEGFs are shared between PvH and PvS or between PvH and PV221, and there are no SEGFs common to all three genomes (Fig. 4D). Although the number of significantly expanded gene families (SEGFs) differed between PvH (54) and PvS (28), the number of expansion genes (SEGs) within these families was nearly identical in PvH (289 genes) and PvS (284 genes). However, the number of SEGs in PV221 (459 genes) was significantly higher than that in both PvH and PvS (Fig. 4E). A significant proportion of SEGs (> 60%) in all three *P. viticola* assemblies were attributed to duplicated genes generated by recent DGB, TD and PD (Fig. 4E), indicating newly generated genes through DGB, TD and PD are the important sources of gene family expansion. In PvH, DGB-SEGFs were primarily enriched in the biological processes of 'chromatin remodeling/organization', 'metabolic', and 'protein-DNA complex organization', whereas the majority of the TD-SEGFs and PD-SEGFs featured unknown biological processes (Fig. 4F). For the overlapping genes of duplicated and SEGFs in PvS, most of them featured unknown biological functions, except the TD-SEGFs, which mainly related to metabolic and proteolysis processes (Fig. 4F). For the TD-SEGFs with known biological processes in PV221, genes primarily involved in 'localization', 'proteolysis', and 'catabolic processes' were found, while genes of PD-SEGFs related to 'transporter' activity were most numerous (Fig. 4F). These uniquely expanded gene families caused by recent local gene duplications (DGB, TD, PD) may contribute to adaptive evolution of *P. viticola* isolates, especially

311 those SEGFs involved in pathogenicity-related terms.

312 Interestingly, we found 95% of the RxLRs and CRNs and nearly 90% of the
313 CAZymes existed as duplications and distributed nonrandomly in gene-sparse regions
314 on the chromosome in the three *P. viticola* assemblies (Table S9). Among those above
315 duplicated effectors, the TD pattern predominated, followed by PD, while DGB played
316 an important role in the duplication of CRN effectors (Fig. S6). Using the RxLR and
317 CRN effectors in PvH as examples, we observe distinct patterns in the distribution of
318 those effectors. For CRN effectors, gene clusters on Chr10 as well as in parts of the
319 gene clusters on Chr3 mainly existed as TD, while the remaining large CRN gene
320 clusters on Chr3, Chr4, and Chr14 are predominantly populated by DGB and PD genes
321 (Fig. 5A). Conversely, for RxLR effectors, gene clusters on Chr7, Chr16, Chr15, Chr8
322 and Chr14 existed as TD and PD; notably, RxLR gene clusters on chromosome 15
323 mainly present as DGB (Fig. 5B). The above results suggest that localized gene
324 duplications (DGB, TD and PD) drive the formation and expansion of effector gene
325 clusters.

326 **Genomic structural variations contribute to candidate effector diversification**

327 Genome alignment between PvH with PvS and PV221 (relatively complete, from
328 cultivated *V. vinifera* cultivars of Bordeaux vineyards in France) revealed the presence
329 of several large SVs among three *P. viticola* assemblies. A total of 100 inversions (PvH:
330 4.4 Mb, PvS: 4.7 Mb), 294 translocations (PvH: 5.2 Mb, PvS: 5.3 Mb), and 2205
331 duplications (PvH: 11.2 Mb; PvS: 6.2 Mb) were identified between the PvH and PvS.
332 A total of 84 inversions (PvH: 1.7 Mb, PV221: 1.3 Mb), 192 translocations (PvH: 2.9
333 Mb, PV221: 2.7 Mb), and 1949 duplications in PvH (23.6 Mb) and 493 duplications in
334 PV221 (3.0 Mb) were identified between PvH and PV221 (Table S10). Compared to
335 the PvS and PV221 genomes, the PvH genome contained a higher proportion of
336 duplications, with approximately twice the amount found in PvS (11.2 Mb versus 6.2
337 Mb) and eight times that of PV221 (23.6 Mb versus 3.0 Mb). Moreover, a greater
338 number of inversions and translocations were identified between the PvH and PvS
339 genomes compared to those observed between PvH and PV221 genomes (Fig. 6A;

Table S10). The distribution of structural variants was non-random across the chromosome but concentrated on specific regions of chromosomes. Notably, two large inversions were observed on Chr2 and Chr4 between PvH and PvS, which were absent between PvH and PV221 (Fig. 6A). Thus, even though PvH, PvS and PV221 belong to the same species and have nearly identical genome architecture (Fig. S1), there were also several SVs among genomes, especially chromosome segmental duplications and inversions.

To investigate the impact of SVs on *P. viticola* adaptive evolution, the SVs between the PvH and PvS assemblies were annotated, resulting in the identification of 1,259 genes located within SV-associated regions, including 761 in inversions, 81 in translocations, and 417 in segmental duplications (Table S11). GO enrichment analysis showed that these genes are enriched in functional categories associated with 'hydrolase activity', 'transporter', and 'catalytic' functions, as well as 'catabolism/metabolism' processes (Fig. S7). Among these SV-associated genes, 93 secreted proteins and 223 putative effectors (30 RxLRs, 157 CRNs, and 36 CAZymes) were identified (Table S11). Notably, CRN effectors were significantly enriched in inversion regions (59/379 CRNs; odds ratio = 3.52, $p = 4.25 \times 10^{-14}$), whereas RxLRs and CAZymes did not show significant enrichment ($p = 0.56$ and $p = 0.059$, respectively), indicating a pronounced impact of SVs on CRN family dynamics (Table S12). Further analyses revealed non-random chromosomal distributions of these effector genes: secreted proteins, RXLRs, and CAZymes in inversion regions were mainly located on Chr2, Chr6, and Chr8, whereas CRNs in inversion regions were concentrated on Chr4 and Chr14, and CRNs in segmental duplication regions were concentrated on Chr4, Chr10, and Chr14 (Table S11). For example, a ~1 Mb inversion on Chr2 harbors four CRNs, nine CAZymes, and ten secreted proteins shared by PvH and PvS (Fig. 6B). Furthermore, within the large segmental duplication region on Chr10 (Fig. 6C) and within inverted blocks nested in duplicated regions on Chr4 and Chr14 (Fig. S8), we observed a PvH-specific duplication of CRN effectors, which is consistent with the previously observed pattern of CRN gene family expansion (Fig. 3E). Moreover, most of these SV-associated genes exhibited differential expression patterns or levels between the two isolates (Fig. 6D;

Fig. S9). These findings suggest that structural variation, particularly chromosomal inversions and segmental duplications, contributes to CRN effector turnover and diversification within *P. viticola*.

Furthermore, we identified the previously characterized *AvrRpv3.1*, *AvrRpv10*, and *AvrRpv12* loci in both the PvH and PvS assemblies, which correspond to grapevine resistance interactions mediated by Rpv3.1, Rpv10, and Rpv12 loci, respectively (Paineau et al., 2024; Dvořák et al., 2025a). At the *AvrRpv3.1* locus, we detected the same effector-gene deletion in PvS that was previously reported for the virulent haplotype (Paineau *et al.*, 2024), whereas multiple novel effector variants were present in PvH (Fig. 6E). Similar patterns were observed at the *AvrRpv12* (Fig. 6E) and *AvrRpv10* (Fig. S10A) loci: several known effector genes were absent or showed substantial divergence in PvS, while multiple novel effector variants were detected in both PvH and PvS genomes. Gene expression profiling further revealed that effector genes deleted or mutated in PvS exhibited no detectable expression, whereas the corresponding genes were normally expressed in PvH (Fig. S10B). Together, these observations provide strong evidence for gene loss at these *Avr* loci in PvS. Notably, a substantial enrichment of TEs was observed surrounding or within these deletion sites and newly generated effector variants (Fig. 6E), further demonstrating the role of TEs in the diversification and evolution of oomycete effectors. These findings further demonstrated the role of SVs in facilitating the frequent loss, duplication, mutation, and evolution of genes, and thus the evolution of organisms.

Discussion

Although multiple *P. viticola* lineages have been reported in North America (Rouxel et al., 2014; Fontaine et al., 2021), all isolates analyzed in this study belong to *P. viticola sensu stricto*—the lineage that was introduced to Europe and spread worldwide (Fontaine et al., 2021). Here, two complete genome assemblies of *P. viticola* from *V. vinifera* (PvH) and *V. amurensis* (PvS) were generated, respectively (Table 1). We preliminarily characterized the genome architecture of *P. viticola*, including the putative locations of centromeres, telomeres, and repeat-rich regions, and identified the number and distribution patterns (clustering) of effectors on the chromosomal landscape. Comparative analyses revealed extensive PAV and CNV differences among candidate effectors, many of which occur within duplicated and structurally variable regions. These observations provide a reference framework to explore effector diversity within *P. viticola*, while recognizing that population-level data will be required to determine the evolutionary forces underlying these patterns.

Complete genome assembly and annotation of *P. viticola*

The assembly of the two *P. viticola* genomes (PvH: 115.3 Mb, PvS: 113.0 Mb) into 17 chromosomes represents a significant improvement over previous assemblies using short-read Illumina or long-read PacBio sequences (Table S4) (Brilli et al., 2018, Dussert et al., 2019, Dvořák et al., 2025a, Paineau et al., 2024, Yin et al., 2017). The newly generated genomes have an equivalent genome size (~115 Mb) to that determined by Feulgen staining previously (Voglmayr and Greilhuber, 1998) and exhibit high-level synteny with the reference genome (PV221) and the diploid genome reassembly (PV221Primary, PV221Haplotigs), as well as the chromosome-level diploid genome Pv1419 of *P. viticola* (Figs. S1A, B) (Dussert et al., 2019, Dvořák et al., 2025a, Paineau et al., 2024). Notably, this assembly is 20 Mb larger than the previous one obtained by long-read sequencing (PV221, Dussert et al., 2019), and the majority of the additional 20 Mb consists of repetitive sequences, including TEs and some duplicated chromosomal segments (Fig. 6A; Table S4), suggesting that repeat regions are more completely assembled in these new releases. These assemblies will

serve as important resources for comparative genomics, functional genomics, and population genetics, as well as references for other oomycete genome assemblies.

Local gene duplication as a major source of effector number variation in *P. viticola*

To characterize effector diversity within *P. viticola sensu stricto*, we compared the numbers and chromosomal distributions of predicted effector genes across high-quality assemblies. Candidate effectors were non-randomly distributed and predominantly clustered in specific chromosomal regions (Fig. 2C). Notably, CRN effectors showed substantial quantitative variation across isolates (PvH: 379; PvS: 240; PV221: 219), whereas PV221 encoded more RxLR effectors (PvH: 405; PvS: 419; PV221: 445). Gene-family clustering further revealed isolate-specific effector complements, including 37 and 23 putative cytoplasmic effectors uniquely detected in PvH (2 RxLRs, 35 CRNs) and PV221 (15 RxLRs, 8 CRNs), respectively, and a higher number of unique CAZymes in PV221 (24) relative to PvH and PvS (Table S6). PvH also exhibited an expansion of core CRN genes (PvH: 237; PvS: 183; PV221: 169), consistent with a strain possessing a particularly large CRN repertoire. Take together, these data indicate that *P. viticola* harbors a highly dynamic effector gene content. This interpretation aligns with recent diploid-resolved assemblies and population studies showing structural variation, deletions, and extended haplotypes at effector loci (e.g. AvrRpv3.1, AvrRpv12, Rpv10) (Dvorák et al., 2025a; Paineau et al., 2024). Such patterns resemble those observed in other fungi and oomycetes, where CNV and PAV contribute to effector repertoire plasticity (Bourras et al., 2015; Hartmann and Croll, 2017; Müller et al., 2019). However, effector CNV/PAV may arise through multiple mechanisms—including recombination suppression, TE-associated instability, high heterozygosity, or neutral gene turnover—and distinguishing these processes will require population-level analyses.

Effector genes typically evolve through duplication-based mechanisms (Dutheil et al., 2016; Müller et al., 2019; Yilmaz et al., 2024). Consistent with this, we found that nearly 85% of the secretome, over 95% of CRN and RxLR effectors, and approximately 90% of CAZymes occur as duplicated or highly similar sequences (Fig. S6). Most duplicated effectors were associated with local duplication events, including DGB, TD,

and PD, which together accounted for more than 60% of duplicated copies (Table S9). Local duplications were also the predominant contributors to effector gene clusters and a major source of CNV across isolates (Fig. 5). Similar duplication-driven effector expansions have been documented in *B. graminis f. sp. tritici* (Müller et al., 2019), *P. sojae* (Zhang et al., 2024), *P. infestans* (Matson et al., 2022), and *P. agathidicida* (Cox et al., 2022), suggesting that duplication represents a conserved mechanism for expanding repertoires of secreted proteins in filamentous plant pathogens.

While gene duplication has been proposed to facilitate dosage effects or functional diversification in several pathosystems (Ali et al., 2017; Morales-Cruz et al., 2020), our study does not include functional assays or infection phenotyping; therefore, we refrain from interpreting copy number differences as direct evidence of altered virulence. Moreover, *P. viticola* exhibits high heterozygosity (Paineau et al., 2024), and our estimates for PvH (0.892%) and PvS (0.837%) indicate substantial intra-genome polymorphism that may contribute to haplotype-specific effector content. Accordingly, observed PAV/CNV patterns could result from haplotype divergence, neutral gene loss/gain, or demographic history rather than host-associated adaptation. While the effector differences described here are compatible with host-associated diversification scenarios, additional data—particularly phase-resolved haplotypes, broader isolate sampling, and functional validation—will be required to discriminate among alternative evolutionary explanations.

Genomic structural variations and effector repertoire plasticity

SVs are increasingly recognized as a major source of genomic plasticity in filamentous plant pathogens, providing a dynamic context for gene turnover, copy number changes, and sequence diversification (Chen et al., 2018; de Jonge et al., 2013). Comparative analyses among *P. viticola* isolates PvH, PvS, and PV221 revealed that—despite broadly conserved chromosome-scale architectures—SVs including inversions, translocations, and segmental duplications have accumulated between lineages (Fig. 6A). PvH in particular exhibited an elevated number of large segmental duplications, highlighting a more structurally dynamic genome compared to the reference isolates. These differences may reflect isolate-specific evolutionary histories, but population-level sampling is required. We note that SVs reported here were inferred

computationally from genome assemblies and comparative analyses, and their precise breakpoints and configurations remain to be experimentally validated.

More than 1,200 predicted genes were embedded within SV-associated regions, and these genes were enriched for molecular functions such as hydrolase activity, transporter activity, and metabolic processes (Table S11; Fig. S7). This enrichment suggests that SVs could affect pathways beyond effector biology, although functional effects remain untested. Notably, 223 predicted effector genes were located within SV regions, including core CRN genes, many of which resided in duplicated segments in PvH and PvS. For example, chromosome 10 contained extensive segmental duplications in PvH and PvS relative to PV221, and these segments contained clusters of locally duplicated CRN genes (Fig. 6C). PvH-specific CRN genes were further enriched in repetitive and inverted genomic blocks (Fig. S8). Genes within SV regions also displayed expression pattern and expression-level differences between PvH and PvS (Figs. 6D, S9), although the consequences of such expression differences require functional investigation. We acknowledge that the transcriptomic analyses in this study were primarily used to support gene annotation and to examine overall expression patterns, and were not designed to comprehensively resolve stage-specific effector expression or coordinated host–pathogen transcriptional dynamics. Future studies integrating finer temporal sampling, additional infection stages, and simultaneous host–pathogen transcriptome profiling will be required to fully elucidate the regulatory dynamics and functional relevance of candidate effectors. Together, these observations indicate that SVs provide structural contexts in which effector loci are frequently duplicated, rearranged, or lost. Similar associations between effector clusters and dynamic genomic compartments have been described in other plant pathogens (Dong et al., 2015; Müller et al., 2019; Lofgren et al., 2021; Zhang et al., 2024). However, the presence of effectors within SV regions does not alone establish a causal role for SVs in virulence or host adaptation. Alternative explanations—including neutral gene turnover, background heterozygosity, TE-mediated instability, or recombination suppression—remain possible and require population-scale data to disentangle.

Analyses of Avr loci (AvrRpv3.1, AvrRpv10, AvrRpv12) provide a clear example

of how structural variation may impact effector repertoires. PvS displayed gene deletions or divergent haplotypes at these loci, while PvH retained multiple candidate effectors (Figs. 6E; S10A). The absence of expression for deleted or mutated PvS alleles suggests functional loss, paralleling observations in other *P. viticola* isolates where large deletions at Avr loci correlate with virulence on resistant cultivars (Paineau et al., 2024; Dvořák et al., 2025a, b). Additionally, we observed TE enrichment within and adjacent to effector genes and Avr loci, consistent with the proposed role of TEs in mediating genomic instability and effector diversification in other pathogens (Nakamoto et al., 2023; Abraham Leen et al., 2024; Sampaio and Croll, 2025). While these findings suggest mechanistic links between SVs, TEs, and effector turnover, determining their functional relevance for pathogenicity will require experimental assays and expanded sampling.

Taken together, our results support an emerging view—reinforced by Paineau et al. (2024) and Dvořák et al. (2025a)—that *P. viticola* possesses a highly dynamic effector repertoire shaped by structural variation, gene duplication, and haplotype divergence. The chromosome-scale assemblies and comparative analyses presented here lay a foundation for exploring the genomic bases of virulence and host interactions. However, given that only three isolates were examined, population-level evolutionary conclusions cannot be drawn. To address this limitation, we have initiated systematic sampling of downy mildew isolates from distinct geographical regions and host backgrounds in China, which will support future population genomics, genotype–phenotype association analyses, and functional validation of effector variants.

Conclusion

Our study assembled two complete *P. viticola* genomes from the Eurasian *V. vinifera* and the wild *V. amurensis*. Comparative genomics revealed that candidate effector genes in *P. viticola* predominantly exist as local gene clusters and exhibit strain-specific expansion, deletion and diversification. Notably, many of these effector variations are enriched in genomic regions characterized by SVs, suggesting that SV-rich genomic compartments are associated with effector turnover and diversification.

Together, these findings highlight a highly dynamic genome architecture that may facilitate the generation and maintenance of effector diversity within this species. While the observed patterns are consistent with a role for gene duplication and SV in shaping effector repertoires, functional validation and population-scale analyses will be required to determine their precise contributions to pathogenicity and host adaptation. Our genome resources provide a valuable foundation for future studies aimed at disentangling the evolutionary and functional significance of effector diversification in *P. viticola* and advancing our understanding of plant–pathogen co-evolution.

Materials and Methods

Sample preparation, DNA extraction, and PacBio HiFi sequencing

The *P. viticola* isolates Hs2101 (designated '*P. viticola* PvH' in the genome assembly) and Sl2124 (designated '*P. viticola* PvS') were collected from different grapevine species in China. Hs2101 was isolated from the cultivated *V. vinifera* cv. 'Red Globe' in Qingxu County, Shanxi Province, while Sl2124 was obtained from the wild grape species *V. amurensis* cv. 'Shuanghong' in Liuhe County, Jilin Province. Both isolates were purified by inoculating single-sporangiophore onto leaf discs from their respective host plants (Zhang et al., 2017). Fresh sporangiophores and sporangia on the surface of infected leaves from each isolate were collected, stored at -20°C, and used for DNA extraction with the SP Fungal DNA kit (Omega, Norcross, GA) following the manufacturer's instructions (Huang et al., 2023). DNA quality was verified using NanoDrop spectrophotometry, the Qubit fluorometer, and agarose gel electrophoresis. The DNA samples were fragmented into 10-15 kb segments to construct HiFi libraries, which were then sequenced using the PacBio Sequel II platform. Library preparation and sequencing were performed by Shanghai Personal Biotechnology Co., Ltd. (http://www.personalbio.cn/). The DNA from both isolates was also sent to Beijing Thermo Lily Biotechnology Co., Ltd. (http://www.sinobiocore.com/) for sequencing on an Illumina NovaSeq 6000 platform (150 bp paired-end reads).

RNA extraction and sequencing

Leaves of *V. vinifera* cv. 'Red Globe' and *V. amurensis* cv. 'Shuanghong' were inoculated with *P. viticola* isolates PvH and PvS, respectively. Inoculated leaves were sampled at three time points: 24, 72 and 120 h post-inoculation (hpi). For the 120 hpi samples, only fresh sporangiophores and sporangia were collected. Each sample was immediately snap-frozen in liquid nitrogen and stored at -80°C until RNA extraction. Total RNA was extracted from each sample using TRIzol reagent (Vazyme). RNA purity was assessed using a NanoDrop spectrophotometer (Thermo Scientific) and

further verified by 2% agarose gel electrophoresis, and the RNA integrity was assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Qualified total RNA was submitted to Shanghai Personal Biotechnology Co., Ltd. (http://www.personalbio.cn) for library preparation and sequencing. Sequencing was performed on an Illumina NovaSeq 6000 platform with a paired-end 150 bp (PE150) strategy.

Genome assembly and optimization

Approximately 7.9 Gb and 9.6 Gb HiFi reads of PvH and PvS were generated from the PacBio Sequel II System and CCS module, respectively. In addition, approximately 5.1 Gb and 2.7 Gb of Illumina short reads were generated for PvH and PvS, respectively, using the Illumina NovaSeq 6000 platform. The PacBio HiFi data provided sufficient coverage and high sequencing accuracy to support telomere-to-telomere genome assembly. Illumina short-read datasets used for transcriptome analysis and assembly validation exhibited high base quality, with consistently high Q30 ratios. Comprehensive sequencing quality control metrics are summarized in Table S1. De novo assembly of these HiFi reads was performed using hifiasm v0.13 with default parameters (Cheng et al., 2021); each primary contig GFA file was converted to FASTA format using the awk command. A 136 Mb assembly in 613 contigs for PvH and a 150 Mb assembly in 707 contigs for PvS were generated by hifiasm, respectively. These primary assemblies were decontaminated according to the following two criteria: Firstly, contigs with a coverage rd (read depth) value less than 10 were filtered; as the majority of contigs had a coverage rd value between 30 and 40, based on the hifiasm assembly result GFA file, the remaining contigs were further filtered according to the result of the BLASTn (Altschul et al., 1990) hit in the NCBI nucleotide (nt) database, and contigs with a top hit of a non-oomycete accession were filtered. After filtering, the PvH genome size was 119 Mb in 64 contigs, and the PvS genome size was 121 Mb in 61 contigs. Both genomes were 96.5% complete with 3.5% duplication, as calculated with BUSCO v5.0 (Simão et al., 2015) using alveolata_odb10 datasets.

The telomere sequences of each end of the chromosome of two *P. viticola* primary

genomes were detected by searching telomeric satellite sequences (5'-CCCTAAA-3') using FindTelomeres.py (https://github.com/JanaSperschneider/FindTelomeres). 12 contigs were complete telomere-to-telomere (T2T) chromosomes in the PvH primary assembly and 11 contigs in the PvS primary assembly, respectively. The contiguity of genomic assembly was improved by using related species-based reference genome-assisted assembly technology. Synteny of two cleaned *P. viticola* hifiasm assemblies with the genomes of *P. effusa* (Fletcher et al., 2022), *B. lactucae* (Fletcher et al., 2019), and *P. sojae* (Zhang et al., 2024), which have higher assembly integrity in the current oomycete genomes, was analyzed using MCScanX (Wang et al., 2012). This approach is justified by the high collinearity observed among Peronosporaceae genomes (Fletcher et al., 2022), which allows robust inference of contig order and chromosomal structure. The first 21 and 28 major contigs in PvH and PvS, respectively, are collinear with the genome of *P. effusa*, as well as of *B. lactucae* and *P. sojae*. Among them, the intact T2T contigs in both *P. viticola* genomes were found to be collinear with *P. effusa* T2T chromosomes one-to-one. In the remaining contigs, multiple PvH and PvS contigs were found to be collinear with other *P. effusa* chromosomes. HiFi reads were realigned to these contigs with minimap2 (Li, 2018) and assembled independently again with hifiasm v0.13 (Cheng et al., 2021), yielding two and four complete T2T contigs in PvH and PvS, respectively. Ultimately, the final PvH and PvS assemblies comprised 17 and 18 contigs, respectively, corresponding to 17 chromosomes and showing one-to-one correspondence with the linkage maps of *P. viticola* (Fig. S2A, B) established by Dvořák et al. (2025b). Accordingly, we adopted the chromosome numbering and nomenclature proposed by Dvořák et al. (2025b) to maintain consistency across studies and to facilitate direct comparison of synteny and locus coordinates. In PvH, 14 contigs represent complete and gapless T2T chromosomes, each carrying the oomycete-specific telomeric repeat (C3TA3) at both ends. The remaining three contigs each contain only a single telomere: two (Chr15 and Chr17) lack approximately 8~10 kb of sequence at the opposite telomeric end, and one (Chr5) terminates within a highly repetitive rDNA region that could not be fully resolved. In PvS, 15 chromosomes are assembled as complete, gapless T2T sequences. Chr15 contains a single gap of 100 Ns

spanning the nucleolar organizer region (NOR) between two contigs, and Chr5 terminates within a repetitive rDNA region. All unplaced contigs in both PvH and PvS were fragments that corresponded to rDNA arrays or mitochondrial genome sequences, and contained very minimal coding sequences (Fig. S1C, D). Specifically, for PvH, the total size of unplaced contigs is ~1,703 kb (accounting for only 1.43% of the total PvH assembly size). Gene prediction identified eight protein-coding genes, three annotated as reverse transcriptase, all non-effectors. For PvS, the unplaced contigs have a total size of ~1,838 kb (1.52% of the total PvS assembly size). Gene prediction yielded 17 protein-coding genes, none with known functional annotations and all non-effectors. Consequently, these short contigs were excluded from subsequent analysis.

Centromere identification

Putative centromeric regions were identified by the presence (BLASTn alignment length >500 bp) of a *P. sojae* copia-like transposon proposed as a distinctive feature of oomycete centromeres (Fang et al. 2020). We then calculated repeat density and gene density in 50 kb sliding windows (step size = 10 kb) across each chromosome. Regions showing a local maximum in repeat content together with a local minimum in gene density were considered candidate centromeric intervals. Finally, these candidate regions were confirmed via synteny with the *P. effusa* centromeres described by Fletcher et al. (2022), and the regions supported by repeat/gene density patterns and synteny conservation were retained. The final coordinates for all putative centromeric regions are listed in Table S2.

Heterozygosity estimation

Illumina reads for each isolate were used to generate k-mer frequency histograms (k = 21) with Jellyfish, and heterozygosity was estimated using GenomeScope1 (http://genomescope.org/).

Genome annotations

Repeat elements were annotated by a combination of homology-based and de novo approaches using RepeatModeler v2.0.1 (Flynn et al., 2020) and RepeatMasker v4.0.6 (http://repeatmasker.org). Annotation of protein-coding genes was conducted with a

combination of ab-initio prediction, transcriptome-based prediction, and homology-based prediction methods using the MAKER2 (Holt and Yandell, 2011) annotation pipeline. In brief, RNA-seq reads were aligned against the genome using HISAT2 (Kim et al., 2019) and then assembled de novo into transcripts with Trinity (Grabherr et al., 2011), using a genome-guided approach. Gene models for species were trained using BRAKER2 (Brůna et al., 2021) based on RNA-seq bam files, and ab-initio gene predictions on a repeats-masked genome were conducted using AUGUSTUS v.3.0.3 (Stanke et al., 2006) based on the trained models. Further, MAKER2 (Holt and Yandell, 2011) uses the assembled transcripts and proteomes of five other oomycete species (*B. lactucae*, *P. effusa*, *P. viticola*, *P. sojae*, and *P. ramorum*) as evidence to generate the final non-redundant gene sets. A reannotation of the *P. viticola* INRA_PV221 assembly (hereafter called PV221, Dussert et al., 2019) was carried out with the same parameters and input data for comparison purposes. The function of protein-coding genes was annotated by searching their protein sequence against the SwissProt, UniProt, and KOG/NOG databases using BLASTP (Altschul et al., 1997), PANNZER2 (Törönen et al., 2018), and eggNOG-mapper v2 (Cantalapiedra et al., 2021), respectively. Protein domains were identified by searching against the InterPro v.32.0 and Pfam databases, using InterProScan v5.59 (Jones et al., 2014) and HMMER3 (Eddy, 1998), respectively. The Gene Ontology term for each gene was obtained from the corresponding InterPro or pannzer entry. KEGG metabolic pathway annotation was performed using KAAS (Moriya et al., 2007).

Assessment of genome assembly quality

To evaluate the completeness of the genome assembly, the HiFi reads were remapped to the assembly using minimap2 (Li, 2018), and Illumina short reads were mapped to the assembly using BWA-mem v0.7.17 (Li, 2013). The mapping rates and the genome coverages were calculated using samtools v0.1.19 (Danecek et al., 2021). BUSCO v5.0 (Simão et al., 2015) software was used to assess the completeness of gene regions using alveolate_odb10 (n = 171) and stramenopiles_odb (n = 100) datasets in Benchmarking Universal Single-Copy Orthologs (BUSCO). The assembly was found to contain 96.5%

(163 of 171) and 100% (100 of 100) complete single-copy orthologs at the alveolate and stramenopiles levels, respectively.

Candidate effector characterization

Proteins that possess a signal peptide structure and lack a transmembrane structural domain are classified as secreted proteins. The prediction of protein signal peptides was conducted using SignalP v3 (Dyrløv Bendtsen et al., 2004). We used SignalP v3 for signal peptide prediction, as it exhibits superior sensitivity for detecting oomycete signal peptides (Sperschneider et al., 2015) and maintains methodological consistency with recent Peronosporaceae effector annotation studies (Fletcher et al., 2022; Zhang et al., 2024). Subsequently, protein transmembrane structural domains were predicted using the DeepTMHMM (Hallgren et al., 2022). Proteins with a positive prediction from SignalP, an HMM S probability value $\geqslant$ 0.9, an NN Ymax score of $\geqslant$ 0.5, an NN D score of $\geqslant$ 0.5, and lacking TM domains were considered as secreted proteins. Then, the predicted secreted proteins were submitted to ApoplastP (Sperschneider et al., 2018) to determine whether they are localized to the plant apoplast.

Candidate RxLR and CRN effectors were identified using a conservative, multi-step strategy that integrates motif-based, profile-based, and homology-based approaches. RxLR candidates were predicted following the framework of McGowan and Fitzpatrick (2017), incorporating the Win, Regex, HMM, and BLAST methods. In parallel, RxLR effectors were independently predicted using the effectR v1.0.2 package (Tabima and Grünwald, 2019). Only non-redundant proteins supported by at least one of these approaches and fulfilling the secretome criteria were retained as putative RxLR effectors. CRN effectors were predicted using a combination of regular expression searches and hidden Markov model (HMM) profiling. Proteins identified by the initial motif search were aligned using MAFFT (Katoh and Standley, 2013), and an HMM profile was constructed using HMMER3 (Eddy, 1998). This CRN-specific HMM was subsequently searched against the predicted proteomes, and proteins with detectable similarity (bit score > 0) were retained as putative CRNs. All predicted CRNs were required to encode a signal peptide and lack transmembrane domains. To minimize

false positives, all predicted candidate effectors were subjected to post-processing steps, including redundancy removal, confirmation of secretion signals, and exclusion of proteins with predicted transmembrane regions. Throughout the manuscript, RxLRs and CRNs are therefore referred to as candidate or putative effectors. Functional annotation of carbohydrate-active enzymes (CAZymes) was performed using the dbCAN2 meta-server, integrating three independent computational pipelines: HMMER (E-Value < 1e-15, coverage > 0.35), DIAMOND (E-Value < 1e-102), and CGCFinder (Distance <= 2, signature genes = CAZyme+TC). To ensure high-confidence functional assignment, a consensus-based strategy was adopted, where CAZymes were filtered and retained only if predicted by at least two of the three aforementioned tools (Zhang et al., 2018). While CAZymes can contribute to host colonization and pathogenicity through cell wall modification (Lo Presti and Kahmann, 2017), they are not considered canonical effectors in oomycetes. Accordingly, CAZymes are treated here as secreted pathogenicity-related proteins rather than effectors sensu stricto, and no direct functional role in virulence is inferred in the absence of experimental validation. Additional secreted protein families, including elicitins (PF00964), necrosis- and ethylene-inducing proteins (NLPs; PF05630), CAP proteins (PF00188), transglutaminases (PF16683), and PcF phytotoxins (PF09461), were annotated based on Pfam domain searches.

Distribution of transposable element (TE) density and their spatial relationship with effectors

Genome-wide TE density was calculated in 10-kb sliding windows with the BEDTools v2.29 software (Quinlan and Hall, 2010) coverage function. The physical proximity between effector genes and TEs was assessed using TEGRiP (Meguerditchian et al., 2021). The designation of genes as TE-associated was determined by the presence of overlapping or adjacent genes within a 2-kb window of a TE. The enrichment of TE proximity in effector genes relative to non-effectors was assessed using Fisher's exact test. Statistical analyses were conducted in R v 4.4.1.

Whole-genome synteny and structural variation analysis

The PvS and PV221 assemblies were aligned to the PvH assembly using Mummer v4.0.0 (Marçais et al., 2018) with parameter settings "--maxmatch -c 1000 -l 50". The alignment results were initially filtered using the "delta-filter -m" option to retain many-to-many alignment blocks for synteny analyses. To facilitate chromosome-level comparisons with PvH and PvS, the more fragmented PV221 assembly was reconstructed into pseudochromosomes. This was achieved by ordering and orienting contigs based on synteny with the PvH reference genome using SyRI (Goel et al., 2019) with the "chroder" parameter. This reconstruction facilitated the visualization of structural variations. SVs were identified using SyRI (Goel et al., 2019) based on whole-genome alignments, including inversions, translocations, duplications, and insertions/deletions larger than 50 bp. Default SyRI filtering criteria were applied, which require SVs to be supported by collinear alignment blocks with consistent orientation and sufficient alignment length. SVs overlapping assembly gaps or low-complexity regions were excluded from downstream analyses.

For quantitative analyses, alignments were further filtered using *delta-filter -1* to retain one-to-one syntenic blocks. Unaligned regions extracted using *show-diff* were classified as presence/absence variations (PAVs). Functional annotation of SV-associated genes was performed using ANNOVAR (Wang et al., 2010). Only high-confidence SVs detected between PvH and PvS were used for downstream analyses, including effector enrichment tests. Enrichment of candidate effector genes within inversion regions was assessed using Fisher's exact test, with all annotated genes as background. Circos plots were generated using Circos v0.69 (Krzywinski et al., 2009), and sliding window coverage analyses were performed using BEDTools v2.29 (Quinlan and Hall, 2010).

Gene family and phylogenetic analyses

The protein sequences of 20 selected oomycete species were subjected to gene family clustering analysis using OrthoFinder (Emms and Kelly, 2019). The conserved single-copy orthologous genes were used for phylogenetic tree construction via the maximum likelihood method using FastTree (Price et al., 2009). The resulting

780 evolutionary trees were visualized using the iTOL (Letunic and Bork, 2019).
781 Divergence times were estimated using the mcmctree in the PAML package (Yang,
782 2007). The expansion and contraction of gene families were quantified using CAFÉ
783 v5.0 (De Bie et al., 2006).

784 **Gene duplication analysis**

785 Gene duplications refer to specific molecular events that produce additional copies
786 of a single gene. We performed an all-by-all blastp (1e-5) alignment with the proteins
787 coded by the predicted genes within each assembly for three *P. viticola* (PvH, PvS and
788 PV221) and their relatively downy mildew species *B. lactucae* and *P. effusa*. Then,
789 different modes of gene duplication, including duplicated gene blocks (DGB,
790 containing at least five collinear genes), tandem duplication (TD, two adjacent similar
791 genes on the same chromosome), proximal duplication (PD, two similar genes with
792 intervals of less than 20 other genes on the same chromosome), dispersed duplication
793 (DSD, two similar genes with intervals of more than 20 other genes), and single-copy
794 (SL) genes were classified using the duplicate_gene_classifier program in MCScanX
795 (Wang et al., 2012). The Ks and Ka/Ks values for each type of duplicated gene pair
796 were calculated based on the YN method implemented in the PAML program (Yang,
797 2007). The Ks and Ka/Ks distribution were plotted and displayed using R v 4.4.1.

798 **Differential expression analysis**

799 Raw RNA-seq reads were quality-filtered using fastp v0.23.2 (Chen et al., 2018)
800 with default parameters. Clean reads were aligned to the PvH reference genome using
801 hisat2 (Kim et al., 2019). Gene-level read counts were quantified using featureCounts.
802 The resulting count matrix was imported into DESeq2 in R for normalization and
803 differential expression testing. Differentially expressed genes (DEGs) were identified
804 using the following criteria: $|\log_2(FC)| \geq 2$, FDR < 0.05. Effectors and secreted protein
805 annotations were integrated with the DEG list to identify differentially expressed
806 effector candidates. Expression heatmaps were generated using Chiplot
807 (https://www.chiplot.online/).

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2023YFD1401405), the China Agricultural Research System (CARS-29), the Expert Workstation Project in Yunnan Province (202305AF150129), and Central Public-interest Scientific Institution Basal Research Fund (82025X131).

Authors' contributions

Y-Q Liu, X-Q Huang, Z-Y Wang, Q Wang, H Zhang and L-Z Zhou conceived and designed the experiments. L-Z Zhou, S-W Cui and F-F Kong carried out experiments and performed the bioinformatic analyses. L-Z Zhou and X-Q Huang wrote the initial manuscript. Y-Q Liu, X-Q Huang, Y-F Zhou, F Du and S-D Li revised and contributed to the rewriting of sections of the manuscript.

Data availability

The newly assembled *Plasmopara viticola* (PvH and PvS) genomes, along with all raw sequencing data—including HiFi reads, Illumina DNA reads, and RNAseq data—of both PvH and PvS have been submitted to NCBI with project no. PRJNA1145503. Gene and effector annotation of *P. viticola* PvH and PvS assemblies are available on Zenodo [https://doi.org/10.5281/zenodo.17873764].

Competing interests

The authors declare that they have no competing interests.

References

Abraham Leen N, Oggenfuss U, Croll D (2024) Population-level transposable element expression dynamics influence trait evolution in a fungal crop pathogen. mBio **15:** e02840-02823

Ali SS, Shao J, Lary DJ, Kronmiller B, Shen D, Strem MD, Amoako-Attah I, Akrofi AY, Begoude BA, Ten Hoopen GM, Coulibaly K, Kebe BI, Melnick RL, Guiltinan MJ, Tyler BM, Meinhardt LW, Bailey BA (2017) *Phytophthora megakarya* and *P. palmivora*, closely related causal agents of cacao black pod rot, underwent increases in genome sizes and gene numbers by different mechanisms. Genome Biol Evol **9:** 536-557

Abraham Leen N, Oggenfuss U, Croll D. 2024. Population-level transposable element expression dynamics influence trait evolution in a fungal crop pathogen. *mBio* **15**(3): e02840-02823.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403-410

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389-3402

Ayala-Usma DA, Cárdenas M, Guyot R, Mares MCD, Bernal A, Muñoz AR, Restrepo S (2021) A whole genome duplication drives the genome evolution of *Phytophthora betacei*, a closely related species to *Phytophthora infestans*. BMC Genomics **22:** 795

Beckerson WC, Rodríguez de la Vega RC, Hartmann FE, Duhamel M, Giraud T, Perlin MH (2019) Cause and Effectors: Whole-Genome Comparisons Reveal Shared but Rapidly Evolving Effector Sets among Host-Specific Plant-Castrating Fungi. mBio **10**

Bourras S, McNally KE, Ben-David R, Parlange F, Roffler S, Praz CR, Oberhaensli S, Menardo F, Stirnweis D, Frenkel Z, Schaefer LK, Flückiger S, Treier G, Herren G, Korol AB, Wicker T, Keller B (2015) Multiple Avirulence Loci and Allele-Specific Effector Recognition Control the Pm3 Race-Specific Resistance of Wheat to Powdery Mildew. Plant Cell **27:** 2991-3012

Brilli M, Asquini E, Moser M, Bianchedi PL, Perazzolli M, Si-Ammour A (2018) A multi-omics study of the grapevine-downy mildew (*Plasmopara viticola*) pathosystem unveils a complex protein coding- and noncoding-based arms race during infection. Scientific Reports **8:** 757

Brůna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M (2021) BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. NAR Genomics and Bioinformatics **3:** lqaa108

Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J (2021) eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. Mol Biol Evol **38:** 5825-5829

Chen JY, Liu C, Gui YJ, Si KW, Zhang DD, Wang J, Short DPG, Huang JQ, Li NY, Liang Y, Zhang WQ, Yang L, Ma XF, Li TG, Zhou L, Wang BL, Bao YM, Subbarao KV, Zhang GY, Dai XF (2018) Comparative genomics reveals cotton-specific virulence factors in flexible genomic regions in *Verticillium dahliae* and evidence of horizontal gene transfer from *Fusarium*. New Phytol **217:** 756-770

Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics **34:** i884-i890

Cheng H, Concepcion GT, Feng X, Zhang H, Li H (2021) Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nature Methods **18:** 170-175

Cox MP, Guo Y, Winter DJ, Sen D, Cauldron NC, Shiller J, Bradley EL, Ganley AR, Gerth ML, Lacey RF, McDougal RL, Panda P, Williams NM, Grunwald NJ, Mesarich CH, Bradshaw RE (2022) Chromosome-level assembly of the *Phytophthora agathidicida* genome reveals adaptation in effector gene families. Front Microbiol **13:** 1038444

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H (2021) Twelve years of SAMtools and BCFtools. Gigascience **10**

De Bie T, Cristianini N, Demuth JP, Hahn MW (2006) CAFE: a computational tool for the study of gene family evolution. Bioinformatics **22:** 1269-1271

de Jonge R, Bolton MD, Kombrink A, van den Berg GC, Yadeta KA, Thomma BP (2013) Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Res **23:** 1271-1282

Delmas CEL, Dussert Y, Delière L, Couture C, Mazet ID, Richart Cervera S, Delmotte F (2017) Soft selective sweeps in fungicide resistance evolution: recurrent mutations without fitness costs in grapevine downy mildew. Molecular Ecology **26:** 1936-1951

Delmas CEL, Fabre F, Jolivet J, Mazet ID, Richart Cervera S, 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

Delmotte F, Mestre P, Schneider C, Kassemeyer H-H, Kozma P, Richart-Cervera S, Rouxel M, Delière L (2014) Rapid and multiregional adaptation to host partial resistance in a plant pathogenic oomycete: Evidence from European populations of Plasmopara viticola, the causal agent of grapevine downy mildew. Infection, Genetics and Evolution **27:** 500-508

Dong S, Raffaele S, Kamoun S (2015) The two-speed genomes of filamentous pathogens: waltz with plants. Curr Opin Genet Dev **35:** 57-65

Dussert Y, Gouzy J, Richart-Cervera S, Mazet ID, Delière L, Couture C, Legrand L, Piron MC, Mestre P, Delmotte F (2016) Draft Genome Sequence of Plasmopara viticola, the Grapevine Downy Mildew Pathogen. Genome Announc **4**

Dussert Y, Mazet ID, Couture C, Gouzy J, Piron MC, 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 Biol Evol **11:** 954-969

Dutheil JY, Mannhaupt G, Schweizer G, C MKS, Münsterkötter M, Güldener U, Schirawski J, Kahmann R (2016) A Tale of Genome Compartmentalization: The Evolution of Virulence Clusters in Smut Fungi. Genome Biol Evol **8:** 681-704

Dvořák E, Dumartinet T, Mazet ID, Chataigner A, Paineau M, Cantù D, Mestre P, Foulongne-Oriol M, Delmotte F (2025a) Parallel adaptation and admixture drive the evolution of virulence in the grapevine downy mildew pathogen. bioRxiv**:** 2025.2005.2018.654733

Dvořák E, Mazet ID, Couture C, Delmotte F, Foulongne-Oriol M (2025b) Recombination landscape and karyotypic variations revealed by linkage mapping in the grapevine downy mildew pathogen *Plasmopara viticola*. G3 (Bethesda) **15**

Dyrløv Bendtsen J, Nielsen H, von Heijne G, Brunak S (2004) Improved Prediction of Signal Peptides: SignalP 3.0. Journal of Molecular Biology **340:** 783-795

Eddy SR (1998) Profile hidden Markov models. Bioinformatics **14:** 755-763

Emms DM, Kelly S (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biology **20:** 238

Fang Y, Coelho MA, Shu H, Schotanus K, Thimmappa BC, Yadav V, Chen H, Malc EP, Wang J, Mieczkowski PA, Kronmiller B, Tyler BM, Sanyal K, Dong S, Nowrousian M, Heitman J (2020) Long transposon-rich centromeres in an oomycete reveal divergence of centromere features in Stramenopila-Alveolata-Rhizaria lineages. PLoS Genet **16:** e1008646

Fletcher K, Gil J, Bertier LD, Kenefick A, Wood KJ, Zhang L, Reyes-Chin-Wo S, Cavanaugh K, Tsuchida C, Wong J, Michelmore R (2019) Genomic signatures of heterokaryosis in the oomycete pathogen *Bremia lactucae*. Nature Communications **10:** 2645

Fletcher K, Shin OH, Clark KJ, Feng C, Putman AI, Correll JC, Klosterman SJ, Van Deynze A, Michelmore RW (2022) Ancestral Chromosomes for Family Peronosporaceae Inferred from a Telomere-to-Telomere Genome Assembly of *Peronospora effusa*. Mol Plant Microbe Interact **35:** 450-463

Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF (2020) RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A **117:** 9451-9457

Gessler C, Pertot I, Perazzolli M (2011) *Plasmopara viticola*: a review of knowledge on downy mildew of grapevine and effective disease management. Phytopathologia Mediterranea **50:** 3-44

Goel M, Schneeberger K (2022) plotsr: visualizing structural similarities and rearrangements between multiple genomes. Bioinformatics **38:** 2922-2926

Goel M, Sun H, Jiao WB, Schneeberger K (2019) SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. Genome Biol **20:** 277

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol **29:** 644-652

Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE, Cano LM, Grabherr M, Kodira CD, Raffaele S, Torto-Alalibo T, Bozkurt TO, Ah-Fong AMV, Alvarado L, Anderson VL, Armstrong MR, Avrova A, Baxter L, Beynon J, Boevink PC, Bollmann SR, Bos JIB, Bulone V, Cai G, Cakir C, Carrington JC, Chawner M, Conti L, Costanzo S, Ewan R, Fahlgren N, Fischbach MA, Fugelstad J, Gilroy EM, Gnerre S, Green PJ, Grenville-Briggs LJ, Griffith J, Grünwald NJ, Horn K, Horner NR, Hu C-H, Huitema E, Jeong D-H, Jones AME, Jones JDG, Jones RW, Karlsson EK, Kunjeti SG, Lamour K, Liu Z, Ma L, MacLean D, Chibucos MC, McDonald H, McWalters J, Meijer HJG, Morgan W, Morris PF, Munro CA, O'Neill K, Ospina-Giraldo M, Pinzón A, Pritchard L, Ramsahoye B, Ren Q, Restrepo S, Roy S, Sadanandom A, Savidor A, Schornack S, Schwartz DC, Schumann UD, Schwessinger B, Seyer L, Sharpe T, Silvar C, Song J, Studholme DJ, Sykes S, Thines M, van de Vondervoort PJI, Phuntumart V, Wawra S, Weide R, Win J, Young C, Zhou S, Fry W, Meyers BC, van West P, Ristaino J, Govers F, Birch PRJ, Whisson SC, Judelson HS, Nusbaum C (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature **461:** 393-398

Hallgren J, Tsirigos K, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A, Winther O (2022) DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *In*. bioRxiv

Hartmann FE, Croll D (2017) Distinct Trajectories of Massive Recent Gene Gains and Losses in Populations of a Microbial Eukaryotic Pathogen. Molecular Biology and Evolution **34:** 2808-2822

Holt C, Yandell M (2011) MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics **12:** 491

Huang X (2020) Genomics diversity and phylogenetic analysis of *Plasmopara viticola* of Chinese wild and Eurasian *Vitis* species. Chinese Academy of Agricultural Sciences. DOI:10.27630/d.cnki.gznky. 2020.000963

Huang X, Wang X, Zhou L, Kong F, Liu Y, Wang Z, Zhang H (2023) TaqMan-MGB PCR Method for Rapid Detection of QoI Fungicide Resistance in Chinese Populations of *Plasmopara viticola*. Plant Disease **107:** 3007-3013

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S (2014) InterProScan 5: genome-scale protein function classification. Bioinformatics **30:** 1236-1240

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol **30:** 772-780

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nature Biotechnology **37:** 907-915

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res **19:** 1639-1645

Letunic I, Bork P (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res **47:** W256-w259

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv **1303**

Li H (2018) Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics **34:** 3094-3100

Li K, Jiang W, Hui Y, Kong M, Feng LY, Gao LZ, Li P, Lu S (2021) Gapless indica rice genome reveals synergistic contributions of active transposable elements and segmental duplications to rice genome evolution. Mol Plant **14:** 1745-1756

Lin L, Sun T, Guo J, Lin L, Chen M, Wang Z, Bao J, Norvienyeku J, Zhang D, Han Y, Lu G, Rensing C, Zheng H, Zhong Z, Wang Z (2024) Transposable elements impact the population divergence of rice blast fungus *Magnaporthe oryzae*. mBio **15:** e0008624

Lo Presti L, Kahmann R (2017) How filamentous plant pathogen effectors are translocated to host cells. Current Opinion in Plant Biology **38:** 19-24

Lofgren LA, Nguyen NH, Vilgalys R, Ruytinx J, Liao H-L, Branco S, Kuo A, LaButti K, Lipzen A, Andreopoulos W, Pangilinan J, Riley R, Hundley H, Na H, Barry K, Grigoriev IV, Stajich JE, Kennedy PG (2021) Comparative genomics reveals dynamic genome evolution in host specialist ectomycorrhizal fungi. New Phytologist **230:** 774-792

Long Q, Cao S, Huang G, Wang X, Liu Z, Liu W, Wang Y, Xiao H, Peng Y, Zhou Y (2024) Population comparative genomics discovers gene gain and loss during grapevine domestication. Plant Physiol **195:** 1401-1413

Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A (2018) MUMmer4: A fast and versatile genome alignment system. PLoS Comput Biol **14:** e1005944

Martínez JR, Miclot A-S, Dvořák E, Mazet ID, Couture C, Delière L, Fabre F, Hoarau I, Yobrégat O, Foulongne-Oriol M, Delmotte F (2025) A multivirulent *Plasmopara viticola* strain from Cilaos on Réunion Island breaks down Rpv1, Rpv3.1 and Rpv10 mediated resistance of grapevine. bioRxiv**:** 2025.2009.2002.673614

Matson MEH, Liang Q, Lonardi S, Judelson HS (2022) Karyotype variation, spontaneous genome rearrangements affecting chemical insensitivity, and expression level polymorphisms in the plant pathogen *Phytophthora infestans* revealed using its first chromosome-scale assembly. PLoS Pathog **18:** e1010869

McGowan J, Fitzpatrick David A (2017) Genomic, Network, and Phylogenetic Analysis of the Oomycete Effector Arsenal. mSphere **2:** 10.1128/msphere.00408-00417

Meguerditchian C, Ergun A, Decroocq V, Lefebvre M, Bui QT (2021) Pipeline to detect the positional and directional relationship between transposable elements and adjacent genes in host genome. Peer Community Journal, **1:** e48

Menardo F, Praz CR, Wicker T, Keller B (2017) Rapid turnover of effectors in grass powdery mildew (*Blumeria graminis*). BMC Evol Biol **17:** 223

Morales-Cruz A, Ali SS, Minio A, Figueroa-Balderas R, García JF, Kasuga T, Puig AS, Marelli JP, Bailey BA, Cantu D (2020) Independent Whole-Genome Duplications Define the Architecture of the Genomes of the Devastating West African Cacao Black Pod Pathogen *Phytophthora megakarya* and Its Close Relative *Phytophthora palmivora*. G3 (Bethesda) **10:** 2241-2255

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res **35:** W182-185

Müller MC, Praz CR, Sotiropoulos AG, Menardo F, Kunz L, Schudel S, Oberhänsli S, Poretti M, Wehrli A, Bourras S, Keller B, Wicker T (2019) A chromosome-scale genome assembly reveals a highly dynamic effector repertoire of wheat powdery mildew. New Phytologist **221:** 2176-2189

Nakamoto AA, Joubert PM, Krasileva KV (2023) Intraspecific Variation of Transposable Elements Reveals Differences in the Evolutionary History of Fungal Phytopathogen Pathotypes. Genome Biology and Evolution **15:** evad206

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 (2024) Multiple deletions of candidate effector genes lead to the breakdown of partial grapevine resistance to downy mildew. New Phytologist **243:** 1490-1505

Peressotti E, Wiedemann-Merdinoglu S, Delmotte F, Bellin D, Di Gaspero G, Testolin R, Merdinoglu D, Mestre P (2010) Breakdown of resistance to grapevine downy mildew upon limited deployment of a resistant variety. BMC Plant Biology **10:** 147

Praz CR, Bourras S, Zeng F, Sánchez-Martín J, Menardo F, Xue M, Yang L, Roffler S, Böni R, Herren G, McNally KE, Ben-David R, Parlange F, Oberhaensli S, Flückiger S, Schäfer LK, Wicker T, Yu D, Keller B (2017) AvrPm2 encodes an RNase-like avirulence effector which is conserved in the two different specialized forms of wheat and rye powdery mildew fungus. New Phytol **213:** 1301-1314

Price MN, Dehal PS, Arkin AP (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol **26:** 1641-1650

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics **26:** 841-842

Rouxel M, Mestre P, Baudoin A, Carisse O, Delière L, Ellis MA, Gadoury D, Lu J,

Nita M, Richard-Cervera S, Schilder A, Wise A, Delmotte F (2014) Geographic Distribution of Cryptic Species of *Plasmopara viticola* Causing Downy Mildew on Wild and Cultivated Grape in Eastern North America. Phytopathology® **104:** 692-701

Rouxel M, Mestre P, Comont G, Lehman BL, Schilder A, Delmotte F (2013) Phylogenetic and experimental evidence for host-specialized cryptic species in a biotrophic oomycete. New Phytol **197:** 251-263

Sampaio AM, Croll D (2025) Transposable elements create distinct genomic niches for effector evolution among *Magnaporthe oryzae* lineages. BMC Biology **23:** 282

Sánchez-Vallet A, Fouché S, Fudal I, Hartmann FE, Soyer JL, Tellier A, Croll D (2018) The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens. Annu Rev Phytopathol **56:** 21-40

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, Avia K, Holtgräwe D, Grimplet J, Matus JT, Ware D, Wu X, Wang H, Liu C, Fang Y, Rustenholz C, Cheng Z, Xiao H, Zhou Y (2023) The complete reference genome for grapevine (*Vitis vinifera* L.) genetics and breeding. Horticulture Research **10**

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics **31:** 3210-3212

Sperschneider J, Dodds PN, Singh KB, Taylor JM (2018) ApoplastP: prediction of effectors and plant proteins in the apoplast using machine learning. New Phytol **217:** 1764-1778

Stanke M, Schöffmann O, Morgenstern B, Waack S (2006) Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinformatics **7:** 62

Tabima JF, Grünwald NJ (2019) effectR: An Expandable R Package to Predict Candidate RxLR and CRN Effectors in Oomycetes Using Motif Searches. Mol Plant Microbe Interact **32:** 1067-1076

Törönen P, Medlar A, Holm L (2018) PANNZER2: a rapid functional annotation web server. Nucleic Acids Research **46:** W84-W88

Voglmayr H, Greilhuber J (1998) Genome size determination in peronosporales (Oomycota) by Feulgen image analysis. Fungal Genet Biol **25:** 181-195

Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research 38: e164-e164

Wang Q, Jiang C, Wang C, Chen C, Xu JR, Liu H (2017) Characterization of the Two-Speed Subgenomes of *Fusarium graminearum* Reveals the Fast-Speed Subgenome Specialized for Adaption and Infection. Front Plant Sci **8:** 140

Wang X, Liu Z, Zhang F, Xiao H, Cao S, Xue H, Liu W, Su Y, Liu Z, Zhong H, Zhang F, Ahmad B, Long Q, Zhang Y, Liu Y, Gan Y, Hou T, Jin Z, Wu X, Liu G, Wang Y, Peng Y, Zhou Y (2024) Integrative genomics reveals the

polygenic basis of seedlessness in grapevine. Curr Biol **34:** 3763-3777.e3765

Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH (2012) MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res **40:** e49

Wingerter 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

Xiang G, Yin X, Niu W, Chen T, Liu R, Shang B, Fu Q, Liu G, Ma H, Xu Y (2021) Characterization of CRN-Like Genes From *Plasmopara viticola*: Searching for the Most Virulent Ones. Front Microbiol **12:** 632047

Xiang J, Cheng J, Wei L, Li M, Wu J (2022) Functional analysis of the Nep1-like proteins from *Plasmopara viticola*. Plant Signal Behav **17:** 2000791

Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol **24:** 1586-1591

Yilmaz F, Karageorgiou C, Kim K, Pajic P, Scheer K, Beck CR, Torregrossa AM, Lee C, Gokcumen O, Audano PA, Austine-Orimoloye O, Beck CR, Eichler EE, Hallast P, Harvey WT, Hastie AR, Hoekzema K, Hunt S, Korbel JO, Kordosky J, Lee C, Lewis AP, Marschall T, Munson KM, Pang A, Yilmaz F (2024) Reconstruction of the human amylase locus reveals ancient duplications seeding modern-day variation. Science **386:** eadn0609

Yin L, An Y, Qu J, Li X, Zhang Y, Dry I, Wu H, Lu J (2017) Genome sequence of *Plasmopara viticola* and insight into the pathogenic mechanism. Scientific Reports **7:** 46553

Yin L, Li X, Xiang J, Qu J, Zhang Y, Dry IB, Lu J (2015) Characterization of the secretome of *Plasmopara viticola* by de novo transcriptome analysis. Physiological and Molecular Plant Pathology **91:** 1-10

Yin X, Fu Q, Shang B, Wang Y, Liu R, Chen T, Xiang G, Dou M, Liu G, Xu Y (2022) An RxLR effector from *Plasmopara viticola* suppresses plant immunity in grapevine by targeting and stabilizing VpBPA1. Plant J **112:** 104-114

Yin X, Shang B, Dou M, Liu R, Chen T, Xiang G, Li Y, Liu G, Xu Y (2019) The Nuclear-Localized RxLR Effector PvAvh74 From *Plasmopara viticola* Induces Cell Death and Immunity Responses in *Nicotiana benthamiana*. Front Microbiol **10:** 1531

Zhang H, Kong F, Wang X, Liang L, Schoen CD, Feng J, Wang Z (2017) Tetra-primer ARMS PCR for rapid detection and characterisation of *Plasmopara viticola* phenotypes resistant to carboxylic acid amide fungicides. Pest Manag Sci **73:** 1655-1660

Zhang Z, Zhang X, Tian Y, Wang L, Cao J, Feng H, Li K, Wang Y, Dong S, Ye W, Wang Y (2024) Complete telomere-to-telomere genomes uncover virulence evolution conferred by chromosome fusion in oomycete plant pathogens. Nature Communications **15:** 4624

1178 # Table

1179 **Table 1** Summary statistics of assembly and annotation of two *P. viticola* genomes

Sequencing	PvH	PvS
Raw bases of PacBio HiFi (Gb)	7.9	9.6
Raw base of RNAseq-Illumina (Gb)	6.8	7.8
Assembly		
Genome size (Mb)	115.3	113.0
Number of chromosomes	17	17
Number of contigs	17	18
Contig N_{50} (bp)	7,025,164	6,267,842
Longest contig (bp)	14,934,798	14,512,239
Shortest contig (bp)	4,336,716	4,235,849
Number of gaps	0	1
Number of Ns	0	100
GC content of the genome (%)	44.8	44.8
Annotation		
Repeat elements (%)	55.8	55.1
Number of predicted genes	14,473	14,341
Complete BUSCOs (%)	96.5	96.5
Genes annotated by UniProt	9,276 (64.1%)	9,223 (64.3%)
Genes annotated by SwissProt	7,396 (51.1%)	7,426 (51.8%)
Genes annotated by Interproscan	8,172 (56.5%)	8,172 (57.0%)
Genes annotated by eggNOG	8,456 (58.4%)	8,288 (57.8%)
Genes annotated by Pfam	7,127 (49.2%)	7,143 (49.8%)
Genes annotated by GO	8,134 (56.2%)	8,144 (56.8%)
Genes annotated by KEGG	3,612 (25.0%)	3,638 (25.4%)

1180

1181

Figure legends

Fig. 1 Architecture of the 17 chromosomes of *P. viticola* isolates PvH. **a** Scaled chromosomes and ticks showing lengths. **b** Dark dots at the end of each chromosome were telomeres, orange boxes indicate putative positions of centromeric regions, and blue boxes represent rDNA arrays. **c** densities of protein-coding genes. **d** densities of and repetitive sequence. **e** densities of GC content. **f** distribution of LTR retrotransposon (green represents Copia, red represents Gypsy). **g** distribution of predicted RxLR (blue) and CRN (red) effectors. **h** distribution of carbohydrate-active enzymes (CAZymes). **i** distribution of NLPs (orange), CAPs (green), elicitins (yellow), and transglutaminase (blue). **j** Collinearity between the chromosomes.

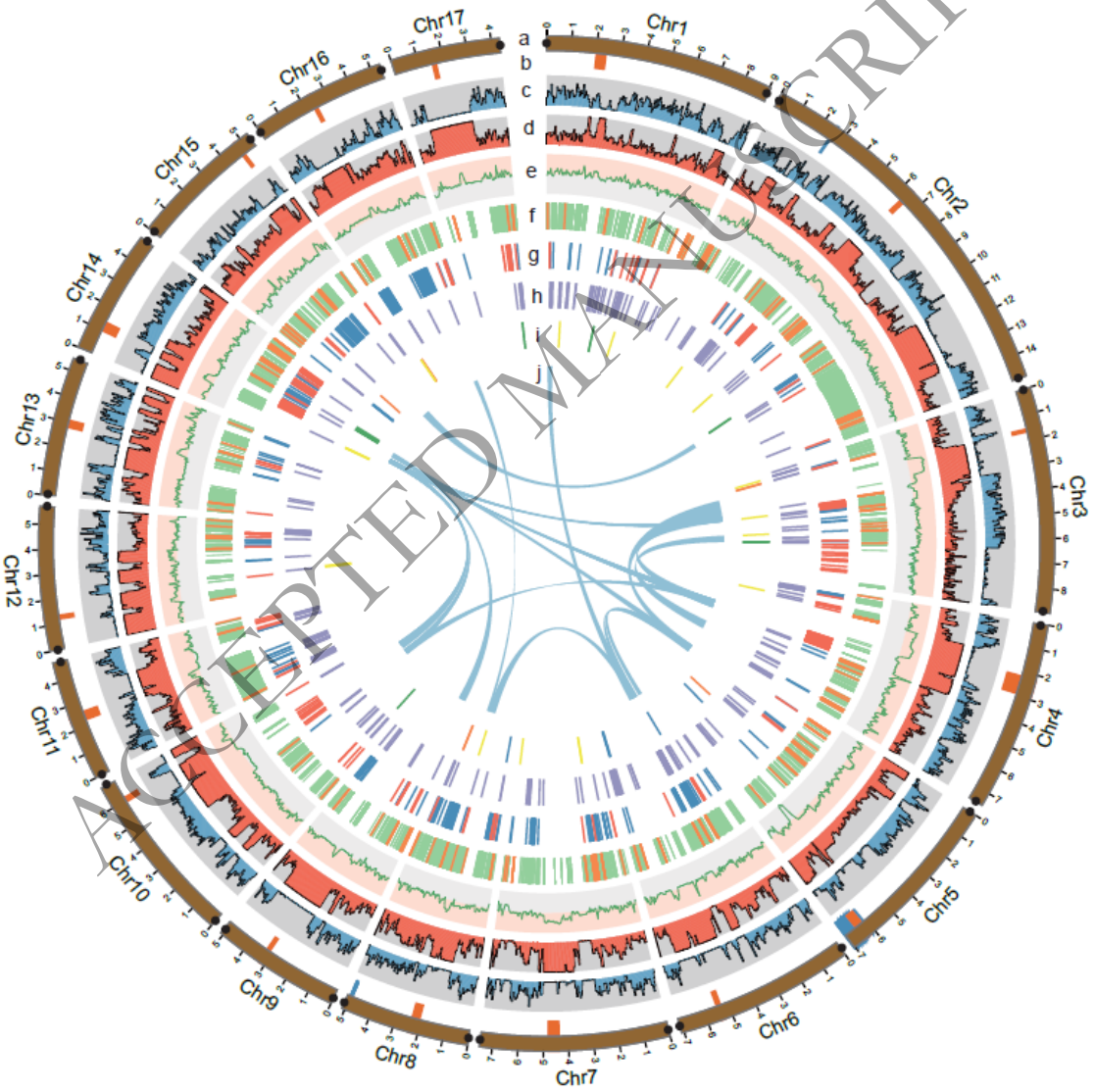

Fig. 2 Annotation of repeats and effector gene clusters in the *P. viticola* assemblies (PvH, PvS). **(A)** Types and percentages of different repeats detected in the *P. viticola* genome (PvH). **(B)** Number and distribution of effectors across chromosomes in PvH and PvS genomes. Frequency of different types of gene families among three *P. viticola* genomes. **(C)** Distribution patterns of effectors across chromosomes in the PvH and PvS genomes. The uppermost and nethermost density plots illustrate the distribution density of TEs along the chromosome.

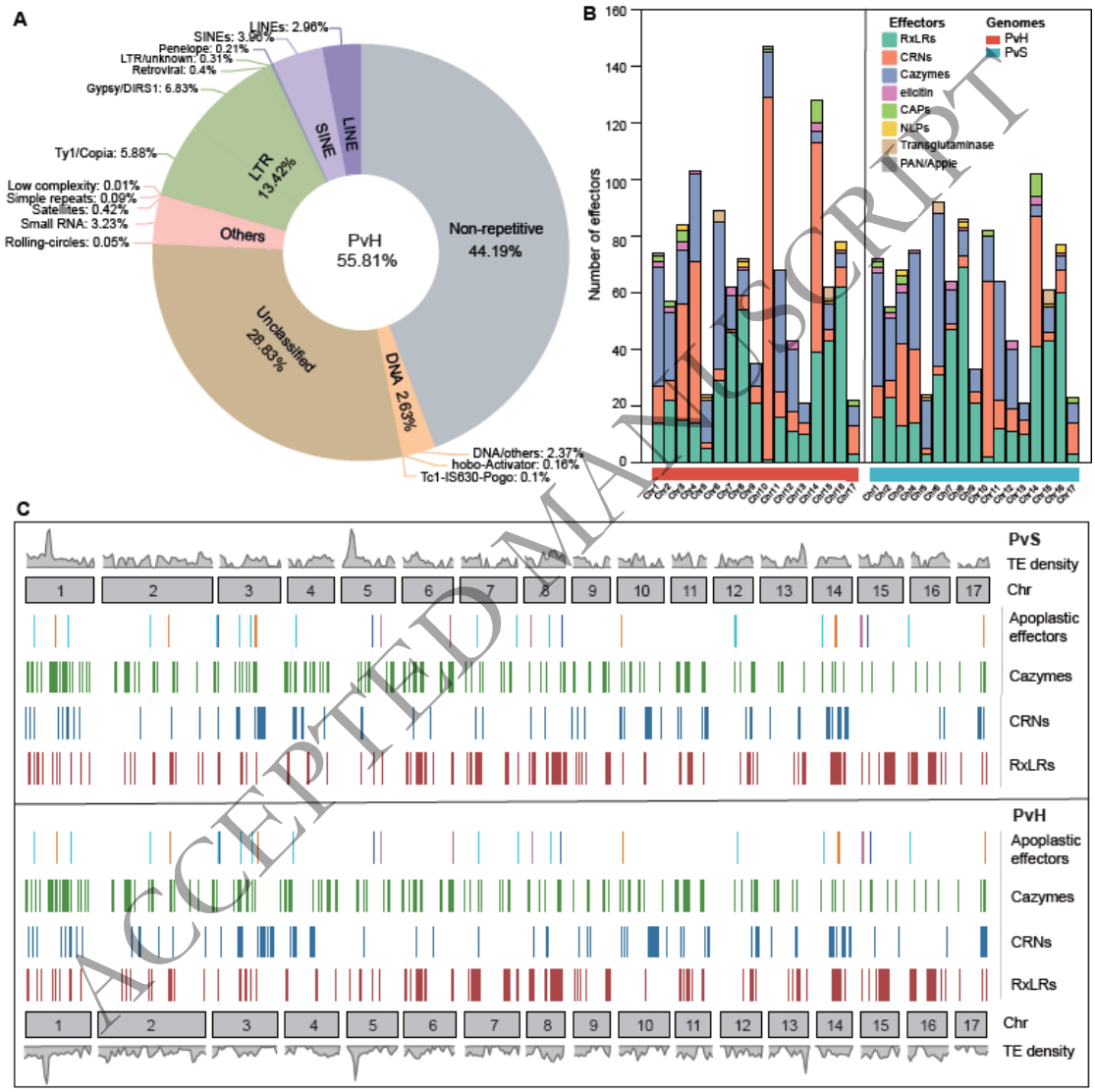

Fig. 3 Gene synteny and evolution among the *P. viticola* assemblies (PvH, PvS and PV221). **(A)** Venn diagram showing the core, dispensable, and private gene families among three *P. viticola* genomes. **(B)** Functional enrichment analysis of private genes in PvH, PvS and PV221, respectively, the enriched GO terms with a corrected *P* value < 0.05 are presented. **(C)** Summary of core, dispensable and private effector genes in PvH, PvS and PV221 assemblies. **(D)** Numbers of genes in the top six CRN families per genome. **(E)** Phylogenetic analysis of the largest CRN family across three *P. viticola* genomes (PvH: blue, PvS: yellow, PV221: purple). PvH and PvS showed clear branch-specific expansion patterns.

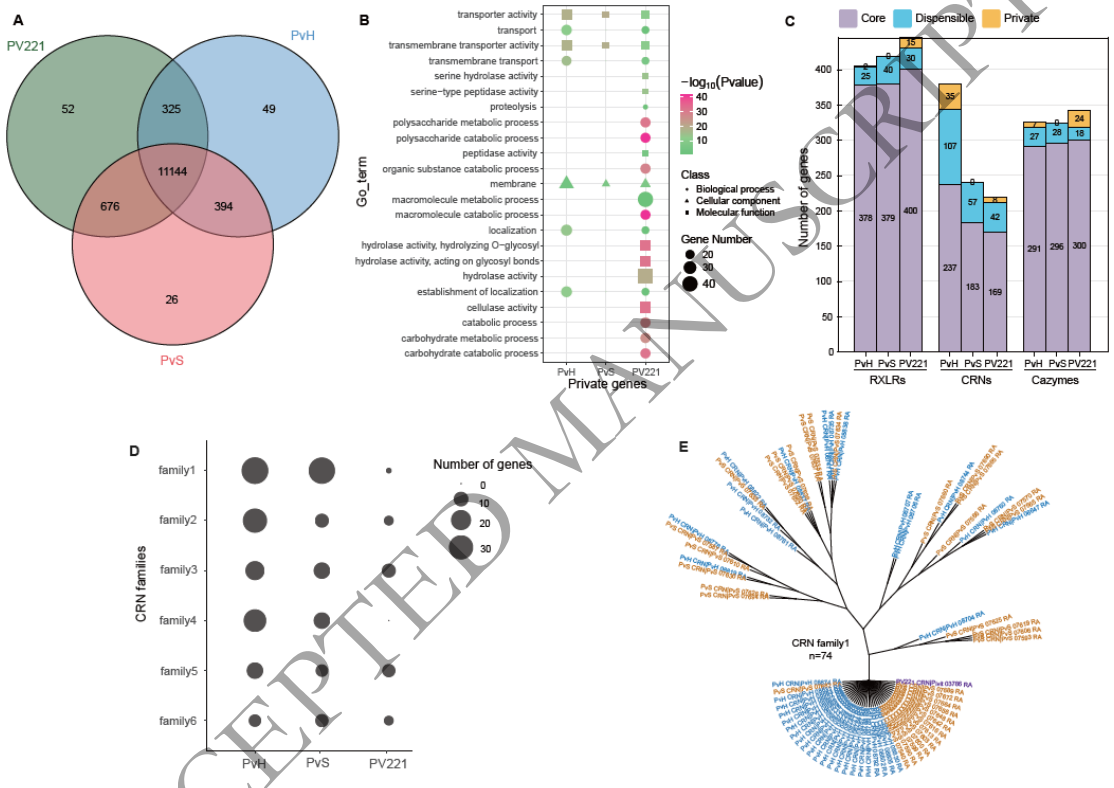

Fig. 4 Local gene duplications contribute to specific gene family expansions in *P. viticola*. **(A)** Schematic cladogram, assembly size, and number of predicted genes are shown for PvS (blue), PV221 (green), PvH (red), and closely related downy mildew species (gray), *B. lactucae* and *P. effusa*. **(B)** Gene duplication categories and percentages in PvS, PV221, PvH, Brlac (*B. lactucae*), and Peeff (*P. effusa*). DGB: duplicated gene blocks, containing at least five collinear genes; TD: tandem duplication, two adjacent similar genes on the same chromosome; PD: proximal duplication, two similar genes with intervals of less than 20 other genes on the same chromosome; DSD: dispersed duplication, two similar genes with intervals of more than 20 other genes; SL: singleton. **(C)** Ka/Ks distribution of gene pairs from DGB, TD, and PD. **(D)** The Venn diagram shows the overlap between significantly expanded gene families in three *P. viticola* genomes. **(E)** The proportions of modes of duplication (DGB, TD, PD, and DSD) and singleton (SL) genes in significantly expanded gene families (SEGFs). SEGs: genes in significantly expanded gene families. **(F)** Functional enrichment analysis of genes overlapping between significantly expanded gene families (SEGFs) and recent gene duplications (DGB, TD and PD). '*P* adjust' is the Benjamini-Hochberg false discovery rate (FDR) adjusted *P* value. The enriched GO terms with a corrected *P* value < 0.01 are presented.

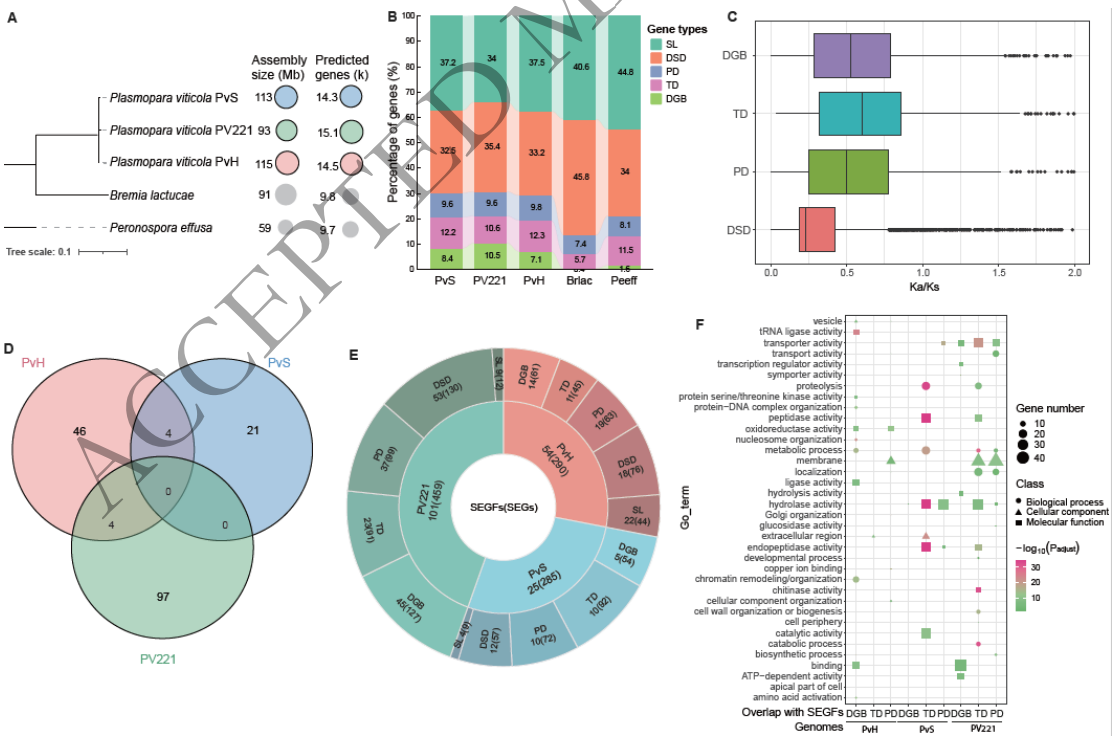

Fig. 5 Local gene duplications contribute to the formation of effector gene clusters. **(A)** Distribution of duplicated (triangles) and single-copy (squares) genes within CRN effector gene clusters. **(B)** Distribution of duplicated (triangles) and single-copy (squares) genes within RxLR effector gene clusters. Gene duplications are classified by type: blue triangles, duplicated gene blocks (DGB, containing at least five collinear genes); red triangles, tandem duplication (TD, two adjacent similar genes on the same chromosome); green triangles, proximal duplication (PD, two similar genes with intervals of less than 20 other genes on the same chromosome); and purple triangles, dispersed duplication (DSD, two similar genes with intervals of more than 20 other genes). Gray squares represent single-copy genes. Numbers indicate the number of duplicated genes in each cluster. Chromosome bars are colored according to gene density, with yellow indicating low density and blue indicating high density.

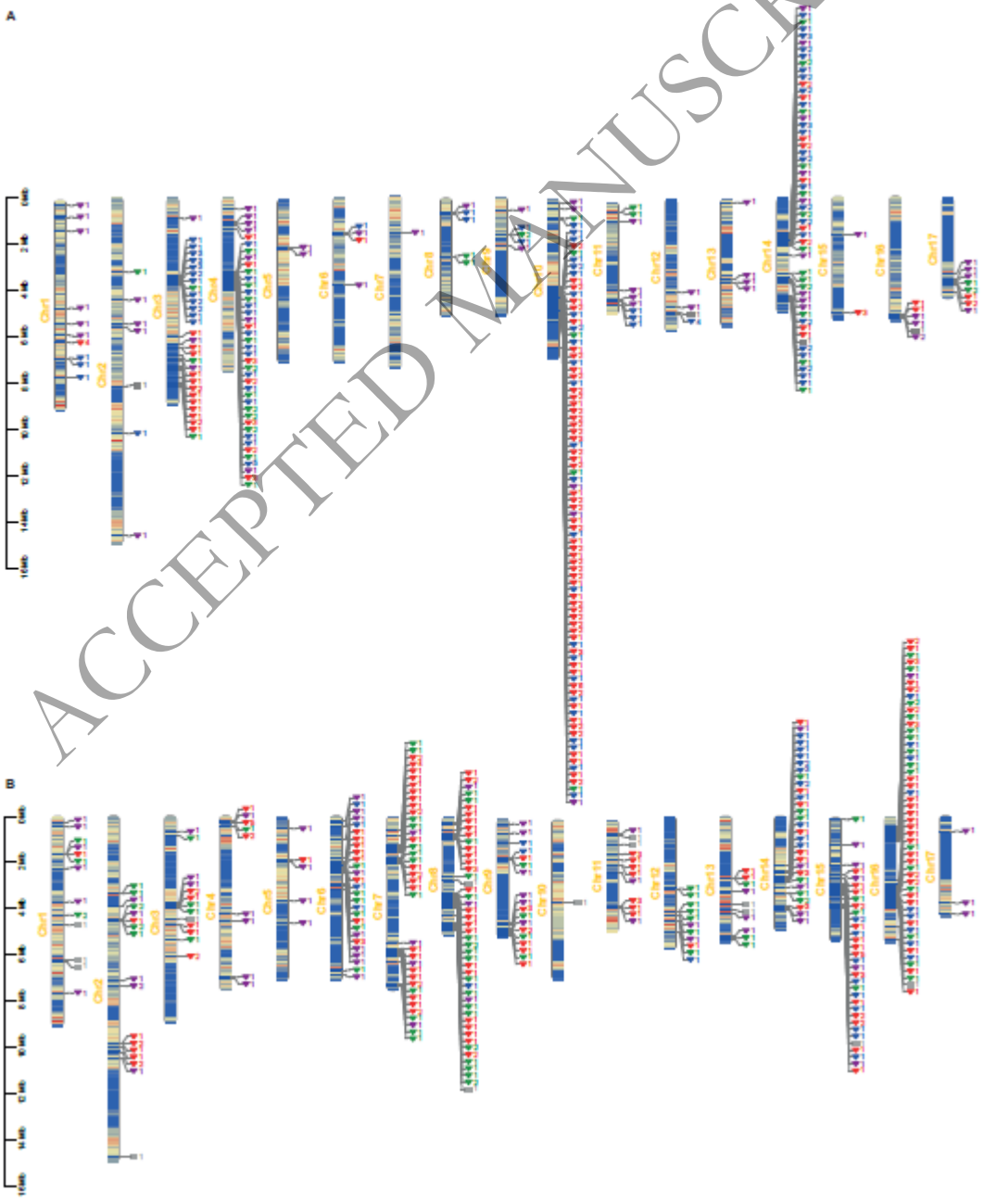

Fig. 6 Detection and functional annotation of structural variations between the three *P. viticola* assemblies (PvH, PvS and PV221). **(A)** Structural variations between complete genome PvH, PvS, and fragment assembly PV221 of *P. viticola*. **(B)** The annotation of the inversion region in Chr2 between PvH and PvS. Red lines represent CRN effectors, and green lines represent CAZymes. **(C)** Multiple local duplicated CRN genes enriched in chromosomal segmental duplication regions in PvH and PvS compared to PV221. **(D)** Integrated analysis of genes located in the structural variation regions and the differentially expressed genes in PvH and PvS. The blue bars indicate the numbers of genes uniquely identified or those overlapping through this approach. INV: Inversion; DUPS: Duplications; TRANS: Translocations; DEGs: Differentially expressed genes. **(E)** Identification and synteny analysis of the *AvrRpv3.1* and *AvrRpv12* loci in PvH and PvS genomes. Genes annotated above PvH represent those identified only present in the avirulent haplotype (Dvořák et al., 2025a). Genes are linked in a gray line when they are present in both genomes. Genes present only in one g enome are colored in red (PvH) or pink (PvS). Transposable element (TE) distribution is shown above the PvH locus and below the PvS locus.