Phased grapevine genome sequence assembly of an *Rpv12* carrier for exploration of *Rpv12* associated positional and functional *Plasmopara viticola* resistance genes

Bianca Frommer[1] (Co), Sophia Müllner[2] (Co), Daniela Holtgräwe[1], Prisca Viehöver[1], Bruno Hüttel[3], Reinhard Töpfer[2], Bernd Weisshaar*[1], Eva Zyprian*[2]

*Corresponding authors

[1] Bielefeld University, Genetics and Genomics of Plants, Faculty of Biology & Center for Biotechnology (CeBiTec), Bielefeld, Germany

[2] Julius Kühn-Institute, Institute for Grapevine Breeding Geilweilerhof, Siebeldingen, Germany

[3] Max Planck Institute for Plant Breeding Research, Max Planck-Genome-Centre Cologne, Cologne, Germany

Email addresses:

BF: frommer@CeBiTec.Uni-Bielefeld.DE / ORCID: 0000-0002-5792-0102
SM: sophia.muellner@julius-kuehn.DE / ORCID: 0000-0001-7263-9389
DH: dholtgra@CeBiTec.Uni-Bielefeld.DE / ORCID: 0000-0002-1062-4576
PV: viehoeve@CeBiTec.Uni-Bielefeld.DE / ORCID: 0000-0003-3286-4121
BH: huettel@mpipz.mpg.DE / ORCID: 0000-0001-7165-1714
RT: reinhard.toepfer@julius-kuehn.DE / ORCID: 0000-0003-1569-2495
BW: bernd.weisshaar@Uni-Bielefeld.DE / ORCID: 0000-0002-7635-3473
EZ: eva.zyprian@julius-kuehn.DE / ORCID: 0000-0003-1095-1996

Summary

The downy mildew disease caused by the oomycete *Plasmopara viticola* is a serious threat for grapevine and can cause enormous yield losses in viticulture. The quantitative trait locus *Rpv12,* mediating resistance against *P. viticola*, was originally found in the Asian *Vitis amurensis* species. Here, we present a detailed analysis of this locus and its genes. The *Rpv12*-carrying breeding line Gf.99-03 was long-read sequenced and a high-quality haplotype separated genome sequence assembly including gene annotation was worked out. The defense response against *P. viticola* was investigated through an infection time-course experiment. Differential gene expression analysis revealed approximately 600 up-regulated genes during infection. The *Rpv12* region of both haplotypes of Gf.99-03, one conferring resistance and the other one sensitivity, were structurally and functionally compared to each other as well as to the corresponding alleles of the susceptible grapevine model cultivar PN40024 and revealed great structural differences. Two diverse clusters with important resistance related genes were identified that clearly characterize the *Rpv12* resistance locus. One cluster carries a set of four differentially expressed genes with three *ACCELERATED CELL DEATH 6* related genes. The other cluster carries a set of six *resistance gene analogues* for potential qualitative pathogen resistance. Co-segregating simple sequence repeat markers for *Rpv12* are located in close proximity to the *R*-genes. The *Rpv12* locus and its candidate genes for *P. viticola* resistance provide a precious genetic resource for *P. viticola* resistance breeding.

Significance statement

The detailed characterization of the *Rpv12* locus for *P. viticola* resistance revealed important candidate resistance genes. As the downy mildew disease caused by *P. viticola* is one of the major grapevine diseases, these findings are a cornerstone for further *P. viticola* resistance research and of major impact for breeding of resistant grapevine varieties for more sustainable viticulture.

Keywords (10 in order of importance)

Resistance QTL, fully-phased genome assembly, TrioBinning, *Vitis* spec., time course infection experiment, leaf disc assay, plant defense

Introduction

With the advent of transatlantic migration of people, various plant pathogens have been introduced to Europe during the 19th century. These include the most important diseases affecting viticulture, the causative agents of powdery- and downy mildew: *Erysiphe necator* ((Schwein.) Burr., syn. *Uncinula necator*, anamorph *Oidium tuckeri* Berk.), an obligate biotrophic ascomycete and *Plasmopara viticola* ((Berk. & Curt.) Berl. & de Toni), an obligate biotrophic oomycete, member of the *Peronosporales* (Gessler *et al.*, 2011). Attempts to delimit the damage caused by these pathogens resulted in cross breeding programs that started already in the 19th century (Bavaresco, 2019). This breeding aims to introgress the pathogen resistance traits present in American or Asian *Vitis* species into European grapevine varieties and is enabled by their cross fertility. First results were discouraging because of the prominent fox tones co-inherited from the wild *Vitis* accessions applied as resistance donors (Töpfer *et al.*, 2011, Reynolds, 2022). Later, after several generations of back crosses to the European *V. vinifera* noble varieties, newly bred resistant varieties such as `Regent´ (1967) and `Johanniter´ (1968) were introduced to German viticulture, where they still persist (https://www.vivc.de/) (Statistisches Bundesamt (Destatis), 2016).

However, it became evident that these grapevine varieties carrying only one resistance locus to each downy and powdery mildew may quickly become susceptible to newly emerging pathogen strains (Kast, 2001, Peressotti *et al.*, 2010, Heyman *et al.*, 2021). Genetic analysis during the last decades revealed further resistance loci from several *Vitis* sources that may be combined for durability of the resistance trait. They are commonly tagged by simple sequence repeat (SSR) based markers and applicable in marker assisted selection, used in combination with standardized infection assays in grapevine breeding (Eibach *et al.*, 2007, Consortium, 2016). To date, more than 30 different loci associated with *Plasmopara viticola* resistance have been described (https://www.vivc.de/). The most prominent ones used currently in breeding comprise several alleles of *Rpv3* (e. g. *Rpv3.1* and *Rpv3.3*), *Rpv10* and *Rpv12* (*Rpv* for Resistance *Plasmopara viticola*) (Welter *et al.*, 2007, Di Gaspero *et al.*, 2012, Schwander *et al.*, 2012, Venuti *et al.*, 2013). Most resistance loci used in breeding are

associated with hypersensitive response (HR), a defense mechanism operated through the production of reactive oxygen species (ROS) that often results in necrotic brown lesions (Possamai *et al.*, 2020).

The locus *Rpv3* has been described by Welter *et al.* (Welter *et al.*, 2007) and various alleles of it were found (Di Gaspero *et al.*, 2012). The *Rpv10* locus was identified by Schwander *et al.* in 2012 in the resistant cultivar `Solaris´ and related varieties (Schwander *et al.*, 2012). The locus *Rpv12* was found in 2013 by Venuti *et al*. in the grapevine varieties `Zarja severa´ and `Michurinets´ and further descendants from *V. amurensis* (Venuti *et al.*, 2013). It is delimited by the SSR markers UDV-350 and UDV-370 (Venuti *et al.*, 2013). Within the homologous region of the reference genome sequence (12X.v2) of the grapevine model PN40024 (Jaillon *et al.*, 2007, Canaguier *et al.*, 2017) twelve typical resistance gene analogues (RGAs) of the CNL Type (CC-NBS-LRR, coiled coil, nucleotide binding site and leucine rich repeat encoding genes (see (Han, 2019)) have been identified. However, since the model genotype PN40024 is susceptible to *P. viticola*, this allele does not mediate *P. viticola* resistance. To understand the mechanisms of *P. viticola* resistance mediated by *Rpv12*, this work set out to analyze the genome sequence of a heterozygous *Rvp12* carrying genotype of grapevine by clear separation of the two haplotypes. The haplotypes were analyzed for candidate resistance genes present specifically at the *Rpv12* locus (positional candidates). In addition, a transcriptomic study with a time course experiment after *P. viticola* inoculation was performed to reveal the cellular reactions leading to successful plant defense (functional candidates). This study was accompanied by microscopic characterization.

Results

Pedigree of Gf.99-03

For the trio binning approach employed to generate fully phased haplotype sequences, data of the parental genotypes in addition to those from its offspring are required (Koren *et al.*, 2018). Unfortunately, one or even both of the parental genotypes of well-known *Rpv12* carriers like `Kunbarat´ and `Michurinets´ are unavailable. For this reason, genotype Gf.99-03 (full name GEILWEILERHOF 2014-099-0003) was chosen for analysis as an offspring of the *Rpv12* carrying genotype 65-153-18 and the *Rpv12* non-carrying genotype Gf.43-21 (full name GEILWEILERHOF 2004-043-0021). In order to verify the relationship of the genotypes under investigation, more than 200 SSR-Markers were applied to characterize the genetic profiles of genotypes 65-153-18, Gf.43-21 and Gf.99-03. Out of the markers tested, 154 were segregating (Table S1). Their segregation patterns confirmed the P1/P2 and F1 relationship of the trio of genotypes selected for sequencing (Figure 1). In addition, the relationship of these three genotypes to their ancestors `Blaufraenkisch´, `Calandro´, `Regent´ and `Domina´ was analyzed using 83 markers (Table S2).

The ancestry of Gf.43-21 could not be clarified completely, due to missing of the genotype VRH3082 1-49 (Figure 1). However, `Calandro´ (*Rpv3.1*) as the second ancestor was analyzed. `Calandro´ is a cross product of the `Regent´ (*Rpv3.1*) and `Domina´ grapevine varieties. Since `Regent´ and VRH3082 1-49 introduced the additional loci *Rpv1* and *Rpv3.1*, the absence of *Rpv1* and *Rpv3.1* in the genome assembled F1 *Rpv12* carrier Gf.99-03 was also confirmed by marker analyses.

Plasmopara viticola infection of *Rpv12* carrying and non *Rpv12* carrying genotypes

To characterize the *P. viticola* resistance trait caused by *Rpv12*, the development of sporangiophores and local necrotic tissue reactions indicative of HR were investigated on genotypes Gf.99-03, Gf.43-21, 65-153-18 and `Italia´ at five days post infection (dpi, Figure 2 (a-d)). While `Italia´ as a susceptible control allowed the formation of many

sporangiophores, the parental *Rpv* carriers 65-153-18 (*Rpv12*) and Gf.43-21 (*Rpv1*, *Rpv3.1*) restricted the development to none or only few sporangiophores within 5 dpi, as did Gf.99-03 (*Rpv12*). Gf.99-03 exhibited HR as indicated by small necrotic lesions at the infection site (Figure 2 (d)). Defense responses of GF.99-03 resulted in the formation of ROS as revealed by specific microscopic staining procedures (Figure 2 (e, f); Figure S1). Furthermore, *P. viticola* formed smaller, but conspicuously more haustoria in the resistant genotype Gf.99-03 than in the susceptible control `Italia´ (Figure 2 (g, h)).

To compare the mycelial growth progressing on the leaves, the genotypes Gf.99-03 (*Rpv12*), Gf.43-21 (*Rpv1*, *Rpv3.1*) and 65-153-18 (*Rpv12*) as well as the susceptible variety `Italia´ were investigated after experimental inoculation (Figure 2 (i)). It is evident that the presence of the *Rpv12*-locus in Gf.99-03 (F1) and the paternal genotype 65-153-18, reduces the mycelial growth to half. The maternal genotype Gf.43-21, which carries a combination of two different resistance loci to *P. viticola* shows a stronger reduction of mycelial development as compared to 65-153-18. Gf.99-03 shows reduced mycelial growth compared to 65-153-18 even though it carries only *Rpv12* as known resistance locus against *P. viticola*.

Genomic sequencing data of Gf.99-03 and its parents

The Gf.99-03 long reads consisted of 6,374,217 sequences adding up to ~71 Gbp with an average read length of 11,132 bp and an N50 of 1,565,128 bp. The longest sequence was achieved at 199,614 bp. Assuming a size of around 500 Mbp for each haplotype, the genome was covered at 1n 142 times or around 70-fold for each haplotype.

Around 116 Gbp short Illumina read data (genome coverage 233-fold) with an average read length of 140 bp were used for the maternal genotype Gf.43-21 and ~144 Gbp short Illumina read data (genome coverage 288-fold) with an average read length of 143 bp were used for the paternal genotype 65-153-18.

Gf.99-03 phase-separated genome sequence assembly

The long reads were divided into the Gf.43-21 bin, the 65-153-18 bin and the unassignable read bin (Table S3). The Gf.43-21 bin contained 46.61 % of all bases, the 65-153-18 bin contained 51.14 % of all bases and the unassigned read bin contained 2.25 % of all bases. As advantageous for a successful binning, the parental bins were almost of equal size with each holding about half of the total bases. The unassigned bin comprised only a small amount of bases and contained mostly short reads (average length 3,638 bp).

The parental bins were assembled separately, yielding two haplotype assemblies that were designated Gf9921 for the maternal haplotype derived from Gf.43-21 and Gf9918 for the paternal haplotype derived from 65-153-18 (Table 1). The results show that 90% of the sequence information of both assemblies is represented by less than 240 contigs. The additional about 1,000 contigs in each of the assemblies are mostly relatively short, presumably caused by high content of repetitive sequences.

The contigs of both haplotypes were assigned to two sets of (homologous) pseudochromosomes based on reciprocal best BLAST hits (RBHs) and synteny with other *Vitis* genome sequence assemblies (Table S4). For Gf9921, 237 contigs representing 85.64 % of all bases, and for Gf9918 195 contigs representing 83.10 % of all bases were ordered into pseudochromosomes (Figure S2, Figure S3).

In general, the lengths of the final homologous pseudochromosomes of Gf9921 and Gf9918 were similar (Figure 3a, Table S5). Only the homologs of chr10, chr13 and chr19 differed in their length by more than 3 Mbp. However, the artificial pseudochromosome chrNh with contigs that had no RBH with any of the syntenic genome sequences (see Experimental Procedures) was significantly larger by ~18.5 Mbp for the Gf9918 assembly as compared to Gf9921. Yet, the total size of the Gf9918 assembly is around 33 Mbp larger than that of Gf9921.

To validate the completeness of the haplotype assemblies, the presence of plant core genes was determined with BUSCO (Figure 3b) (Waterhouse *et al.*, 2018). For both haplotypes, a high portion of completeness of core genes was found (97.6 % Gf9921, 97.8 % Gf9918) resulting in a genome wide percentage of 98.5 %. Despite a very low

portion of missing and fragmented plant core genes for the combined assembly (0.9 % fragmented, 0.6 % missing) and for each haplotype, the content of duplicated genes is rather high for the haplotypes (4.7 % Gf9921, 6.2 % Gf9918), compared to the reference PN40024 with 2.1 % duplications.

The contamination analysis of the Gf.99-03 long reads and of the haplotype assemblies revealed no significant contaminations. More than 99 % of all database hits were hits against *Vitis* sequences. The few other hits were mostly hits against the tree *Spondias tuberosa* or against sequences of the genus *Berberidopsis* and the genus *Ampelopsis*.

Phasing status and k-mer completeness

For analysis of the phasing status more than 470 SSR marker candidates were selected from literature. Only ~250 could be assigned to a position on the pseudochromosomes without mismatches in the assay primer sequences. From these ~250 markers, marker products of 203 markers were generated and tested for inheritance using DNA of Gf.99-03 and its parents Gf.43-21 and 65-153-18.

Out of these 203 markers, 154 segregating (more than 1 bp difference) and 22 non-segregating markers were identified for Gf.99-03 (had an appropriate PCR product and PCR product matched to one of the parental haplotypes), 27 had to be discarded (no product or multiple peaks). Every tested segregating marker (154) supported a full phasing (Table S1), e. g. no haplotype switching was identified

The k-mer based phasing assessment analysis with Merqury (Rhie *et al.*, 2020) resulted in a base level quality value (QV) of 35.44 and a k-mer completeness of >97 % for the Gf.99-03 genome assembly (Table S6). The k-mer completeness for the haplotypes was estimated to >99 %. The analyses indicated a higher content of heterozygous sequence (red 1-copy peak) than of homozygous sequence (blue 2-copy peak), almost no missing sequences (read-only peak) and a few artificial duplications (green peak) (Figure S4). For the phasing status, Merqury computed 2,151 haplotype blocks covering 94.63 % of all bases with an N50 of 2.82 Mbp for Gf9921 and 2,092 blocks covering 95.62 % of all bases with an N50 of 2.59 Mbp for Gf9918. The k-mer

switch error rate was <0.03 % per haplotype by only allowing 10 switches per 20 kbp (Table S6). The bob plot showed haplotype-pure contigs (no mixtures between haplotypes, Figure 4).

Gene annotation assisted by RNA-Seq and extraction of RGAs

Annotation of both haplotype assemblies was supported by RNA-Seq data generated from Gf.99-03. Sequencing of RNA extracted from Gf.99-03s' organs leaf, tendril, root and stem and data processing finally resulted in 138,927,184 reads comprising ~12 Gbp (Table S7). The leaf data constitute 53.86 % of the total bases. The annotation of both Gf.99-03 haplotype sequences (see Experimental Procedures) generated a set of 35,783 genes covering 34,713 protein-coding and 1,070 tRNA-coding genes for Gf9921 and a set of 37,145 genes encoding for 36,290 protein-coding and 855 tRNA-coding genes for Gf9918.

Through the RGA annotation pipeline, 2,288 Gf9921 genes and 2,268 Gf9918 genes were classified as RGA (Table 2, Figure S6). Relatively similar numbers of RGAs were identified for both haplotypes over the RGA classes. Most genes were found for the TM-CC class followed by the RLK class (Receptor like kinase).

Comparative RNA-Seq analysis of a *Plasmopara viticola* infection time course experiment challenging grapevine

Based on the phenotypic analyses of the infection process described above, time points and duration for an infection time course were selected. For five time points (0 hpi, 6 hpi, 12 hpi, 24 hpi, 48 hpi) RNA was extracted and subjected to RNA-Seq. On average 6.14 Gbp were available per treated sample and 6.07 Gbp per untreated sample (Table S8). Because of the even amount of sequenced bases, the low content of non-called bases and Q30 values for over 94 % of the base calls for every sample, the data are considered to be of high quality.

Differentially expressed genes

A Principal Component Analysis (PCA) of the count data of the *P. viticola* infection time-course experiment showed time-separated data points in the time course order (x-axis) and condition-separated data points along the y-axis (Figure 5 (a)). The data points of the infection start time at 0 hpi were the only data points that were not separated by condition, but clustered together. In addition, the triplicated data of the time points and conditions clustered together indicating high experimental reproducibility.

A total of 600 to 800 up-regulated DEGs (differentially expressed genes) and 60 to 290 down-regulated DEGs per time point and per haplotype were identified using the likelihood ratio test (LRT, Figure 5 (b)). ~600 up-regulated DEGs were found in the intersection of all time points and ~200 up-regulated genes were shared between 12 hpi, 24 hpi and 48 hpi. For the down-regulated DEGs, time point 6 hpi contained the highest number of genes with ~200.
To investigate the defense response, DEGs were functionally annotated with KAAS (https://www.genome.jp/kegg/kaas/) and the KEGG pathway ko04626 'Plant-pathogen interaction' (https://www.genome.jp/dbget-bin/www_bget?ko04626) was analysed. About 14 proteins of the pathway were hit by ~31 up-regulated DEGs per haplotype and around five proteins of the pathway were hit by six down-regulated DEGs per haplotype (Table S9). The six down-regulated DEGs *Gf9918_03g04061*, *Gf9918_03g04067*, *Gf9918_09g14214*, *Gf9918_10g15495*, *Gf9918_17g27788* and *Gf9918_18g29219* were also found in the set of up-regulated genes (Table S9). Among the up-regulated DEGs were genes encoding proteins associated with (I) production of ROS, (II) implementation of HR of pathogen associated molecular pattern (PAMP)-triggered immunity, (III) defense-related PR1 and FRK1 induction, (IV) induction of programmed cell death and (V) implementation of HR of effector-triggered immunity (Figure S5). The LRR receptor-like serine/threonine-protein kinase elongation factor Tu receptor (*EFR, K13428*)-like gene (*Gf9918_11g18310*) and the resistance to *Pseudomonas syringae* protein 2 (*RPS2, K13459*)-like genes

(*Gf9918_12g19808*, *Gf9918_12g19817*) were uniquely induced in the *Rpv12* carrying haplotype Gf9918.

Design of new markers

The new assembly with detailed information of the *Rpv12*-Locus sequence was exploited for marker design based on four SSR-type repeats localized within the RGA-region. Four primer pairs to amplify markers named GF14-61 to GF14-64 (Table S10) were designed and tested on different genotypes. Every marker primer pair produced a product, but only GF14-63 produced an *Rpv12*-specific allele size of 398 bp. Susceptible genotypes showed a size difference up to 448 bp (Table S11).

Rpv12 locus sequence and investigation of *Rpv12* candidate genes

The *Rpv12* QTL region, delimited by the SSR markers UDV-014 and UDV-370 (Venuti *et al.*, 2013) on chromosome 14, comprised 1.83 Mbp and 75 genes for Gf9921, 1.86 Mbp and 94 genes for Gf9918 and 2.07 Mbp and 146 genes in the corresponding genomic region of PN40024 12X.v2 (Figure 6).

Potential *Rpv12* candidate genes were chosen from the Gf9918 gene models based on the classification as an RGA by the presence of encoded resistance domains (positional candidates), or based on clear differential gene expression between susceptible and resistant haplotypes in the infection time course analysis (functional candidates). As a result, nine RGAs and five DEGs were considered as “narrow” candidates for the causal genes within the *Rpv12* locus. The candidates were analyzed for gene mates with similar protein sequences in the *Rpv12* lacking haplotype Gf9921 and the *P. viticola* susceptible genotype PN40024. Due to this restriction, the DEG *Gf9918_14g22602*, an ethylene-responsive transcription factor, was removed from the list of candidates owing to only two aa exchanges in comparison to Vitvi14g00564.t01 and Gf9921_14g22247, reducing the number of candidate DEGs to four (Table 3).

The four DEGs were localized at the northern end of the *Rpv12* locus in a region that is considerably larger in Gf9918 (417,382 bp) as compared to Gf9921 (158,620 bp) and PN40024 (262,742 bp) (Figure 6 (b)). This region with functional candidates differs significantly among the allelic haplotypes, yet at its borders it is flanked by highly syntenic genes.

The DEG *Gf9918_14g22556* is annotated as protein of unknown function (PUF), but a BLASTP search against the non-redundant database gave the best hit with an ankyrin repeat-containing protein ITN1-like from *V. riparia* (RefSeq XP_034707067.1) with a query coverage of 84 %, an identity of 95 % and an E-value of 4e-55. Nevertheless, the DEG contains only a PGG and TM domain and lacks the actual ankyrin repeat-containing domain. Compared to the *V. riparia* protein with 345 aa (amino acids), the protein sequence is only 115 aa long. Its expression was raised from e. g. 7.63 LFC at 6 hpi to 8.21 LFC after *P. viticola* inoculation (Figure 7 (a)).

The three DEG candidates *Gf9918_14g22558*, *Gf9918_14g22562* and *Gf9918_14g22568* were annotated as 'Similar to ACD6: Protein ACCELERATED CELL DEATH 6' and have a similar protein sequence. In all three proteins an ankyrin repeat-containing domain, a PGG domain and a TM domain were identified. Even if they share almost the same protein sequence with only a few aa substitutions, they differ in their lengths due to extended start sequences of Gf9918_14g22558 and Gf9918_14g22568. Gf9918_14g22558 has a length of 715 aa, Gf9918_14g22562 of 505 aa and Gf9918_14g22568 of 692 aa (Figure S7 (a)). The three *ACD6* DEGs are positioned in around 100 kbp distance to each other. All three *ACD6* genes showed a similar expression time-course profile and had *Gf9921_22214* as gene mate (Figure 7 (b-d)). However, *Gf9921_14g22214* is allelic to *Gf9918_14g22562* with the same length of 505 aa, but with some aa substitutions.

Six of the nine positional candidate RGAs were located at the southern end of the *Rpv12* locus in a region with structural variation similar to the region with functional candidates. It contains co-segregating markers (Figure 6 (c)). Yet, different to the functional candidate region, the positional candidate region (delimited by the markers UDV-345 and UDV-360, see Figure 6) is significantly smaller in Gf9918 (486,318 bp)

compared to Gf9921 (743,407 bp) and to PN40024 (852,604 bp). This region is compressed around the closely located RGAs and also flanked by highly syntenic regions.

The RLP RGA (TM, LRR domain) *Gf9918_14g22614* was annotated as 'probable disease resistance protein At4g27220'. The CN RGA (CC, NB-ARC domain) *Gf9918_14g22616*, which was located close to *Gf9918_14g22614* (534 bp distance), was annotated as 'Similar to At4g10780', a putative R-protein. The NL RGA (NBS, LRR domain) *Gf9918_14g22619* and the CNL RGAs (CC, NB-ARC, LRR domain) *Gf9918_14g22624* and *Gf9918_14g22626* were all annotated as 'Similar to At4g27220' (R-proteins). Besides the RGA domains CC, NB-ARC and LRR, Gf9918_14g22624 encodes two 'Resistant to *P. syringae* 6' domains (PLN03210) (Lu *et al.*, 2020). Gf9918_14g22626 is quite closely related to Gf9918_14g22624 and codes for similar domains. Nevertheless, the domain positions, their length and the aa composition differ (e. g. Figure S7 (b)). The coding frame of Gf9918_14g22626 is lacking 113 aa followed by another lack of 42 aa within a LRR domain in comparison to Gf9918_14g22624.

Additionally, the *Rpv12* locus contained four TM-CC RGAs. Two TM-CCs, *Gf9918_14g22623* inside of the candidate region with positional candidates and *Gf9918_14g22636* southern of this region, were annotated as 'protein of unknown function'. The other TM-CCs were functionally annotated as containing a 'globular-tail (GTD)-binding domain similar to Myosin-binding Protein 1 (MYOB1)' (*Gf9918_14g22546*, *Rpv12* start) and as containing a 'CheQ-like His Kinase A (HisKA) domain similar to ETHYLENE INSENSITIVE 4 (EIN4)' (*Gf9918_14g22588*, *Rpv12* mid).

Looking at the differential expression, significant high counts were found for *Gf9918_14g22624, Gf9918_14g22626* and *Gf9918_14g22588* under the same expression of treated and untreated samples. In contrast *Gf9918_14g22616, Gf9918_14g22619, Gf9918_14g22546* and *Gf9918_14g22623* had no to very low counts. *Gf9918_14g22614* had up to 100 normalised counts and *Gf9918_14g22636* around 300 counts, but both had no significant differential expression. None of the RGAs was differentially expressed.

Furthermore, a gene (*Gf9918_14g22633)* encoding a Serine/threonine-specific protein phosphatase/bis(5-nucleosyl)-tetraphosphatase (PP2A) domain was identified downstream of the CNLs. Compared to homologs of PN40024, Gf9921 and other Gf9918 proteins, a deletion of 13 aa was detected in Gf9918_14g22633. Gf9918_14g22633 belongs the 14-3-3 protein group (Ferl, 1996) which includes serine/threonine protein phosphatases of types PP1 and PP2A (DeLille *et al.*, 2001). This group of proteins affects metabolism, signal transduction, transcription, stress responses, and associated apoptosis (Dougherty and Morrison, 2004). Besides, a central role in immune defense signaling pathways associated with PTI and ETI was identified for 14-3-3 proteins (Bigeard *et al.*, 2015, Cui *et al.*, 2015, Dubrow *et al.*, 2018, Mathe *et al.*, 2019). Based on this, it can be assumed that this gene may play a relevant role in the immune response against *P. viticola,* but is not considered as a candidate gene for *P. viticola* resistance because of its role as a negative regulator.

Discussion

The pedigree of *Rpv12* carrier Gf.99-03

The goal of the work presented here was to characterize the *Rpv12* locus on chromosome 14 which confers resistance to the grapevine pathogen *P. viticola* that causes downy mildew. Since the *Rpv12* locus displays dominance and is often present in heterozygous state, a phase-separated genome sequence was required. The separation of both haplotypes is possible by using the tool TrioCanu (Koren *et al.*, 2018), but requires additional sequence information of both parental genotypes of the genome-sequenced F1. Thus, a complete 'family' or 'trio' with correct descendance is necessary. Unfortunately, the parental genotypes of the well-known *Rpv12* carriers `Kunbarat´ and `Kunleany´ are not available. For this reason, the genotype Gf.99-03 was selected for genome analysis and elucidation of the resistance locus *Rpv12*.

Prior to analysis, its parental genotypes Gf.43-21 and 65-153-18 as well as the grapevine varieties `Blaufraenkisch´, `Calandro´, `Regent´ and `Domina´ were checked for parentage since these lines were included in the pedigree of Gf.99-03 according to breeding records. These lines were tested with 83 genome wide spread SSR markers. The pedigree was completely verified. However, no relationship could be demonstrated between the varieties `Kunbarat´ or `Kunleany´ and the genotype Gf.99-03, due to distant relationships. Based on Koleda *et al.* (Koleda, 1975), the parental genotypes of `Kunbarat´ and 65-153-18 originate from the same pedigree. Regrettably, the interspecific genotypes 28/19 and 4/15, of which 4/15 is expected to be one parent of 65-153-18 (Figure 1), are not available anymore prohibiting any analysis. However, Venuti *et al.* (Venuti *et al.*, 2013) analysed DNA of a *V. amurensis* accession called NT10 in their studies (Venuti *et al.*, 2013), that probably may fit cultivar 28/19. In addition, all genotypes under analysis were checked for the presence of further known and marker-tagged resistance loci. Based on SSR marker analysis, no other characterized *Rpv* locus was detected in Gf.99-03. The pedigree analysis confirmed that the trio Gf.43-21, 65-153-18 and Gf.99-03 displays the expected parent/F1 relationship.

Gf.99-03 is a highly *P. viticola* resistant *Rpv12* carrier

The QTL *Rpv12* mediates resistance to *P. viticola* and is located on chr. 14. It was identified in several independent introgressions of *V. amurensis* and is present in the genomes of diverse varieties like `Michurinets´, `Zarja severa´, as well as in `Kunbarat´, `Petra´, `Lela´, `Kunleany´ and `Mila´ (Venuti *et al.*, 2013). Recently, it was shown that *Rpv12* was inherited by *V. amurensis* to the grapevine variety `Merlot Khorus´ based on introgression maps (Foria *et al.*, 2022).

The effectiveness of *Rpv12* associated resistance properties was demonstrated by Wingerter et al., 2021 (Wingerter *et al.*, 2021). Gf.99-03 responds to *P. viticola* challenge by production of ROS within 24 hpi. Staining of *P. viticola* inoculated leaf discs and image analyses to evaluate mycelial development showed that the *Rpv12* carrying genotype 65-153-18 allowed less than half of the mycelial growth observed in susceptible `Italia´ controls. The inhibition of mycelial development was even more pronounced in its offspring Gf.99-03 (*Rpv12*). Nevertheless, repression of mycelial expansion of *P. viticola* was strongest in the maternal genotype Gf.43-21 that carries two different *P. viticola* resistance loci, namely *Rpv1* and *Rpv3.1*. *Rpv1* is a strong resistance locus from *Muscadinia rotundifolia* (Merdinoglu *et al.*, 2003), and *Rpv3.1* (Welter *et al.*, 2007) is an American *Vitis* spec. locus most likely derived from *V. rupestris* (Rockel *et al.*, 2021). However, marker analyses confirmed that neither *Rpv1* nor *Rpv3.1* were inherited to Gf.99-03.

Although Gf.99-03 carries only *Rpv12* as known resistance locus against *P. viticola*, the effect of this QTL on mycelial growth seems to be more inhibiting than in 65-153-18. It is possible that unrecognized minor resistance factors were inherited to Gf.99-03, which contribute to its stronger defense of *P. viticola*. Some other weak resistance QTLs to *P. viticola* were described in *V. amurensis* `Shuang Hong´, but remain as yet uncharacterized (Fu *et al.*, 2020).

In general, some influence of the genetic background on the efficiency of individual resistance loci is possible (Foria *et al.*, 2018). Complementarily, little is known about the interaction of combined resistance loci directed against different pathogens, such

as *Ren* and *Rpv*-Loci. The genotype Gf.99-03, besides carrying the locus *Rpv12*, additionally contains resistance QTLs directed against powdery mildew (*Erysiphe necator*, *Ren3* and *Ren9*) on chr. 15 (Welter *et al.*, 2007, Zendler *et al.*, 2017, Zendler *et al.*, 2021). It is possible that defense genes within different resistance loci interact and reinforce each other. It is known from the *M. rotundifolia* resistance QTL *Run1/Rpv1* that genes in close proximity act against *P. viticola* and *E. necator* (Feechan *et al.*, 2013).

Overall, the *Rpv12* carrier Gf.99-03 was validated as a highly *P. viticola* resistant genotype. However, it allowed the formation of haustoria in the early phase of pathogen attack (72 hpi), indicating post-invasion and post-haustorial resistance. The *P. viticola* haustoria on Gf.99-03 were more numerous and much smaller as compared to a susceptible genotype, appearing rudimentary or atrophied (Figure 2 (g, h)). The formation of haustoria plays a central role in the infection process of *P. viticola*. It was shown that haustoria formation is stopped in non-host plants and that defense reactions are triggered in resistant grapevine varieties as soon as haustoria are detected (Diez-Navajas *et al.*, 2008). Formation of haustoria may therefore be a prerequisite to initiate defense reactions in the resistant host. In *A. thaliana* the TM-CC proteins RPW8.1 and RPW8.2 mediate broad spectrum resistance to the powdery mildew pathogen *Golovinomyces* spp. (Xiao *et al.*, 2001, Wang *et al.*, 2009, Kim *et al.*, 2014). The RPW8.2 protein is transferred to the extrahaustorial membrane and activates a salicylic acid-dependent defense pathway (Xiao *et al.*, 2003). It leads to ROS accumulation in the vicinity of haustoria presumably to constrain their development (Wang *et al.*, 2009). ROS production was also observed here during *Rpv12* mediated defense, as also in an earlier investigation on the *Rpv3* locus from an American *Vitis* species (Eisenmann *et al.*, 2019, Rockel *et al.*, 2021). A previous study on *P. viticola* defense reaction of the *V. riparia* accession 'Gloire de Montpellier' identified an induced gene family of three *VRP1* CNL genes *VRP1-1, 1-2 and 1-3* located on chr. 10 (Kortekamp *et al.*, 2008), apart from its resistance QTLs *Rpv5* on chr. 9 and *Rpv6* on chr. 12 (Marguerit *et al.*, 2009). Astonishingly, the proteins encoded by *VRP1-1* and *VRP1-2* include an N-terminal domain with similarity to RPW8 (Kortekamp *et al.*, 2008), opening the hypothesis that similar reaction mechanisms may be active in grapevine and in *A. thaliana* during defense against different mildews.

The high number and small size of haustoria formed on the *Rpv12* carrier Gf.99-03 seem to reflect a repeated struggle of the pathogen to parasitize its host cells and gain nutrients. Successful haustorial development is obviously confined early in the resistant host and prevented by the HR (Figure 2 (e, f)), leading to the programmed death of colonized cells. The generation of haplotype genome sequences of Gf.99-03, with a haplotype known to carry *Rpv12* and the other that is known to confer sensitivity, allowed a detailed molecular analysis of the *Rpv12* locus including differential gene expression studies under *P. viticola* challenge to examine this complex defense reaction.

Phased genome assembly and gene annotation

The genome sequencing of Gf.99-03 resulted in the two high-quality phase-separated haplotype assemblies, designated Gf9921 and Gf9918. Successful phase separation was based on the TrioBinning approach that separates the long reads of the F1 into parental read bins prior to assembly (Koren *et al.*, 2018). With testing of different k-mers for read binning, a proper binning was achieved and approximately equal parental read bins were generated for assembly. Based on this well-developed data and even without scaffolding, two haplotypes were generated that display a high average contig length above 400 kbp, a top level contig N50 length of 3.29 Mbp (Gf9921) and 3.92 Mbp (Gf9918) and a high QV of ~35. Moreover, the k-mer completeness for the whole genome assembly was estimated to ~97 % and per haplotype to >99 % indicating a complete genome assembly. The evaluation of haplotype phasing resulted in pure haplotype sequences with a very low switch error rate of less than 0.03 %.

The Gf.99-03 genome assembly thus represents one of the few high-quality and truly phase-separated *Vitis* genome sequence assemblies currently available. Using the RBH approach, contig sequences holding 85.64 % (Gf9921) and 83.10 % (Gf9918) of all bases were successfully assigned to pseudochromosomes. The overall assignment was validated *in silico* with over 450 markers with existing data and of which 176 SSR markers were amplified by PCR (Table S1). For most of the homologues

pseudochromosomes assigned to either one of the two haplotypes, the lengths were comparable and only a few differ significantly in their lengths (chr10, chr13, chr19). However, the Gf9918 haplotype assembly contains around 33 Mbp more than the Gf9921 haplotype. Investigation of the read data for contaminations, showed that the higher base content of Gf9918 is not based on impure genetic material. The sequence hits against none *Vitis* species matched different dicotyledons and close relatives (clade Eudicots, clade Rosids or order *Vitales*). Already Lodhi and Reisch (Lodhi and Reisch, 1995) reported varying genome sizes of 439 to 526 Mbp for Asian *Vitis* species and 411 to 541 Mbp for North American *Vitis* species. Considering the significantly deviant reported *Vitis* genome sizes and the descent of Gf.99-03's parent 65-153-18 from Asian *V. amurensis*, the increased genome size of the paternal haplotype Gf9918 is not unexpected. Despite the higher base content of Gf9918, the gene annotation and RGA annotation of the haplotype assemblies resulted in an even set of gene models and RGAs.

Differentially expressed genes

The analysis of DEGs was based on an infection time-course experiment using *P. viticola* infected leaf disc samples. The number of sequenced-tagged mRNAs across the various infection stages was approximately equal and the number of genes with a count was similar between both haplotypes. PCA analysis indicated a successful infection time-course experiment with the data points well separated according to condition, time and treatment. As expected, only the treated samples of time point 0 hpi when the *P. viticola* infection was initiated, but had no time to spread, clustered closely with the data points of untreated time point 0 hpi. The differential gene expression analysis resulted in a higher amount of up-regulated DEGs than down-regulated DEGs and equal amounts of DEGs for each haplotype and per time point. It is noticeable that the number of low expressed DEGs was around 250 (Gf9921) to 300 (Gf9918) at time point 6 hpi and decreased to a stable number of 60 to 90 DEGs for the other time points.

During the defense response of Gf.99-03 against *P. viticola*, ~31 genes were up-regulated whose proteins putatively play a role in pathways of plant-pathogen interactions, namely in ROS production, during HR induced by PTI (PAMP triggered immunity) or ETI (effector triggered immunity), and in other defense-related reactions. Genes similar to *A*tRPS2 (on chr. 12) and EFR (on chr. 11) were exclusively induced in the *Rpv12* carrying haplotype upon *P. viticola* challenge.

AtRPS2 genes have been described as plant disease resistance genes carrying leucine-rich repeat classes (Bent *et al.*, 1994). More recent publications describe At*RPS2 as* NLR motif encoding gene, which interacts with bacterial effector protein AvrRpt2. After cloning in *Oryza sativa*, it was demonstrated as conferring resistance to fungal and bacterial pathogens like *Magnaporthe oryzae* and *Xanthomonas oryzae* (Qi *et al.*, 2011, Li *et al.*, 2019). In this study the *RPS2* correlating gene Gf9918_12g19808 has 13 predicted isoforms of which some look like NLR and some carry only LRR motifs. Overexpression of *AtRPS2* showed higher accumulation of ROS (Li *et al.*, 2019).

EFR genes (receptor kinases), showing motifs of serine-threonine protein kinases, are described as interacting with bacterial EF-Tu (elongation factor Tu) and activating the basal immunity (Zipfel *et al.*, 2006). Here, it may be involved in the interaction with an oomycete.

The identification of up-regulated genes with a contribution to PTI and to ETI, supports the plant-oomycete interaction model described in Judelson and Ah-Fong (2019), where oomycete infection is detected through PRRs (pattern recognition receptors) in the intercellular space (Judelson and Ah-Fong, 2019). The up-regulated genes of the infection time-course experiment indicate a complex oomycete detection causing a multi-level defense response. As *Rpv12* describes 78.8 % of the resistance phenotype (Venuti *et al.*, 2013), the *Rpv12* locus must contain specific genes that significantly contribute to the *P. viticola* resistance.

Rpv12 locus and potentially resistance related genes

The sequence of the *Rpv12* locus was extracted from both haplotypes of Gf.99-03, namely from the non-*Rpv12* carrying haplotype sequence Gf9921 and the *Rpv12* carrying haplotype sequence Gf9918. Comparison of these sequences and their gene

content was extended to the homologous region of the genome sequence and gene annotation of the susceptible reference model grapevine genotype PN40024 (Jaillon *et al.*, 2007, Canaguier *et al.*, 2017), revealing considerable structural differences and partial hemizygosity. To identify *Rpv12* associated genes, RGAs encoding typical protein domains as well as differentially expressed genes during pathogen challenge were studied on a genome-wide scale.

Resistance research in *Vitis* has so far been focused on CNL/TNL genes, especially in the context of the *Rpv12* locus (Di Gaspero and Cipriani, 2003, Venuti *et al.*, 2013, Chitarrini *et al.*, 2020). CNL and TNL genes represent the largest class of *R*-genes known in the plant kingdom (Dangl and Jones, 2001, Hulbert *et al.*, 2001, Wei *et al.*, 2020) and it is well known that they play an important role in pathogen defense. Nevertheless, there are additional RGA classes known encoding CN, TN, TX, RLP, RLK and TM-CC type proteins as well as transcription factors (Monteiro and Nishimura, 2018, Chattopadhyay *et al.*, 2019). Using the RGA and DEG data, the complete region between the QTL delimiting SSR markers UDV-014 and UDV-370 was analyzed (Venuti *et al.*, 2013). R-Genes are often organized in clusters (Hulbert *et al.*, 2001, Meyers *et al.*, 2003). Two clusters of typically resistance-associated genes were identified in this region. These clusters are structurally highly diverse between the *Rpv12* carrying haplotype Gf9918 and the *Rpv12* non-carrying haplotype Gf9921 as well as compared to the susceptible PN40024 genotype. They differ in sequence size and presence of resistance-associated genes present in Gf9918, but are flanked by highly syntenic gene regions. One cluster, significantly larger in Gf9918 (Figure 6 (b)), includes three copies of *ACD6*-like RGAs that show up as DEGs and encode Ankyrin repeat-containing domains, PGG and TM domains. The three *ACD6*-like genes *Gf9918_14g22558*, *Gf9918_14g22568* and *Gf9918_14g22562* were found within ~221 kbp. In *A. thaliana*, a hyperactive allele of *AtACD6* (*At4g14400*) conferred resistance to various pathogens including microbes and insects (Todesco *et al.*, 2010), making these genes important candidates for *P. viticola* resistance in grapevine. Also, a DEG (*Gf9918_14g22556*) encoding a protein of unknown function containing a PGG and a TM domain was identified.

The locus was further described by co-segregating markers between marker UDV-350 and UDV-370 (Venuti *et al.*, 2013) reducing the size to 1.01 Mbp and 46 genes in

Gf9921, to 0.78 Mbp and 51 genes in Gf9918 and to 1.14 Mbp and 88 genes in PN40024. Here, positional RGA candidates were found. The co-segregating marker GF14-63, that was designed in this study for downstream marker assisted breeding application, was placed between the first three RGAs. The marker UDV-390 is positioned between the other three RGAs.

The southern cluster (Figure 6 (c)) encompasses a diverse set of RGAs, in detail a RLP, a CN, a NL, a TM-CC and two CNL encoding genes. These RGAs, especially CNL and NL, are often described in relation to plant disease resistance. As soon as a pathogen effector is identified, the conformational change of intracellular immune receptors like such RGAs induces a defense reaction that is often linked to HR (Greenberg, 1997, Tameling *et al.*, 2006, Williams *et al.*, 2011, Bentham *et al.*, 2020). CNLs were previously identified as resistance-associated genes in e. g. *Hordeum vulgare*, *Oryza sativa*, *A. thaliana*, *Aegilops tauschii* and *Solanum tuberosum* (Bentham *et al.*, 2020).

The RLP gene *Gf9918_14g22614*, carrying the TM and LRR domains and similar to the NL Gene *Gf9918_14g22619*, is an homolog of the *A. thaliana* gene *At4g27220* as well as of the gene *GbaNA1* described as a resistance mediating gene of *Gossypium barbadense* against *Verticillium dahliae* (Li *et al.*, 2018). It has been shown that *GbaNA1* activates ROS production and ethylene signalling in *A. thaliana (Li et al., 2018).* Besides, *Gf9918_14g22624* and *Gf9918_14g22626*, encoding for proteins with CC, NB-ARC and LRR domains, are orthologous to *At4g27220*. Additionally, they carry the RPS6 (Resistance to *Pseudomonas syringae*) domain which was originally found in proteins of the TNL class (Kim *et al.*, 2009). Furthermore, they show similarities to the protein RAC1 mediating resistance in *A. thaliana* against the oomycete *Albugo candida* (Borhan *et al.*, 2004, Kim *et al.*, 2009). The orthologue *At4g10780* of the CN gene *Gf9918_14g22616* carries an RPS6 domain, while *Gf9918_14g22616* itself only carries an NB-ARC domain. For the TM-CC gene *Gf9918_14g22623,* directly upstream of the CNL gene *Gf9918_14g22624,* no similarities were found.

In addition, three TM-CC RGAs were detected outside of the described two clusters. The TM-CC *Gf9918_14g22546* is located at the northern end of the *Rpv12* locus. This gene was found to encode a protein with a GTD-binding domain similar to that of the

MYOB1 receptor. AtMYOB1 is a transmembrane myosin receptor that serves in vesicle trafficking, interacts with myosins including the variant XI-K and influences plant growth and development (Peremyslov *et al.*, 2013). Myosin XI-K participates in cell wall-mediated penetration resistance against powdery mildew in *A. thaliana* (Yang *et al.*, 2014). The TM domain for membrane attachment together with the GTD-domain that is needed for cargo-binding suggest a molecular function of *Gf9918_14g22546* as a myosin receptor contributing by cell-wall strengthening to support disease resistance.

The second TM-CC encoded by *Gf9918_14g22588* is located in the region between the two clusters and putatively codes for a CheQ-like histidine kinase A (phosphor-acceptor) domain similar to EIN4. AtEIN4 (At3g04580) is known as an ethylene receptor negatively regulating ethylene signaling (Hua and Meyerowitz, 1998) that also displays a serine/threonine kinase activity *in vitro* (Moussatche and Klee, 2004). It is located on the membrane of the endoplasmic reticulum (ER) and the histidine kinase domain is active in signal transduction. Thus, the TM-CC protein encoded by *Gf9918_14g22588* may function in defense-related ethylene signaling.

The last TM-CC *Gf9918_14g22636* is located south of the second cluster and encodes a protein of unknown function.

In conclusion, the *P. viticola* resistance locus *Rpv12* of grapevine mediates post-invasion, post haustorial resistance. The pathogen may form multiple but atrophied haustoria on resistant plants, likely due to the struggle of the oomycete to gain access to plant nutrients, but inhibited by defense reactions of the host cells including the formation of ROS. The defense reaction leads to localized necrosis and HR. This observation agrees with the clearly pathogen-induced expression of three *ACD6*-like genes present in one cluster of the *Rpv12* locus. A second cluster contains several RGAs, whereof *Gf9918_14g22624* and *Gf9918_14g22626* are constitutively expressed. These genes encode CNL domain proteins and may act as primary receptors of pathogen-derived signals (PAMPs) to recognize the attack by *P. viticola*. In addition, further resistance-associated with gene functions were identified in the *Rpv12* region. Considering the whole genomic region, the resistant haplotype shows extensive structural divergence when compared to the susceptible haplotype and to the genome sequence of PN40024, the susceptible *V. vinifera* model genotype.

This situation is reminiscent of a recent local haplotype resolution and characterization of the *M. rotundifolia* `Trayshed´ QTLs *Run1.2* and *Run2.2* directed against *E. necator* (powdery mildew), which revealed quite substantial structural variation of potential disease resistance genes in the QTL regions (Massonnet *et al.*, 2022). Despite this divergence of genomic regions carrying resistance loci in wild American or Asian *Vitis* relatives (or the closely related American *M. rotundifolia*), there seems to remain enough synteny to the noble European *Vitis vinifera* cultivars genomes to allow the introgression of QTL loci to generate naturally resistant grapevines for viticulture. Stacking of several independent loci is required to enhance sustainability of the resistance traits. To this end, tightly QTL-linked molecular tags – ideally linked to the responsible resistance genes- are required.

In this work, a new SSR marker, GF14-63, localized in the center of the polymorphic *Rpv12* region and between the two RGA clusters was developed. This marker should be highly useful in marker assisted resistance breeding of genetically resistant and more sustainable grapevine varieties. Regarding the *Rpv12* mediated resistance pathway, it can be suspected that *P. viticola* penetration is detected inside the cell by one of the two constitutively expressed CNL genes, but that the differentially expressed *ACD6*-like genes contribute to exert the local necrosis, and the resistance cascade is initiated in concert with additionally differentially regulated functions spread over the chromosomes. Other RGAs identified within the *Rpv12* locus may also play an important role.

Experimental Procedures

Plant material

Gf.2014-099-0003 (Gf.99-03, *V*IVC 27131) is a grapevine genotype carrying the *P. viticola* resistance locus *Rpv12*, in addition to *E. necator* resistance loci *Ren3 and Ren*9 (Zendler *et al.*, 2017, Zendler *et al.*, 2021). It is progeny of breeding line Gf.2004-043-0021 (Gf.43-21, *Rpv1*, *Rpv3.1*, *Ren3*, *Ren9, V*IVC 27130) crossed to genotype 65-153-18 (*Rpv12, V*IVC 41129). Plant material of Gf.99-03, Gf.43-21 and 65-153-18 was taken from the germplasm collection of Julius Kühn Institute, Institute for Grapevine Breeding Geilweilerhof (49°12'54.1''N, 8°02'41.3''E). Wooden two eye cuttings were made in 2019, rooted in Jiffy-7® (Jiffy Products International BV, Zwijndrecht, Netherlands) pots and propagated in the greenhouse. While genotype 65-153-18 (4/15 x `Blaufraenkisch´) is an offspring of an Hungarian breeding line based on *V. amurensis* (Koleda, 1975), Gf.43-21 (`Calandro´ x VRH3082 1-49) is product of a long series of crosses, started in 1916 with backcrosses of *Muscadinia rotundifolia* (Olmo, 1986) (https://www.vivc.de/).

gDNA extraction for SSR-Marker-Analysis

Genomic DNA (gDNA) for SSR-Marker analysis of the genotypes Gf.43-21, Gf.99-03, 65-153-18 as well as `Lemberger´ (*V*IVC 1459), `Calandro´ (*V*IVC 21797), `Regent´ (*V*IVC 4572) and `Domina´ (*V*IVC 3632) was isolated. Small leaf pieces (0.25 cm^2) from the first apically inserted leaf were collected, precooled in 1.5 ml reaction tubes (Eppendorf, Hamburg, Germany), lyophilized (Martin Christ Gefriertrocknungsanlagen GmbH) and ground in a paint shaker (SK450, Fast & Fluid Management B. V., Sassenheim, Netherlands) with metal beads (3 mm in diameter, Ditzinger, Brunswick, Germany). DNA was extracted using the NucleoSpin 96 Plant Core Kit (Macherey Nagel) according to the manufacturer's specifications.

SSR-Marker analysis and relationship determination

The sequences of primer pairs flanking simple sequence repeat (SSR-) Markers used in this work are listed in the supplements (Table S1). All primer pairs were synthesized by Metabion International, Planegg, Germany. PCR amplification was done as described (Müllner *et al.*, 2020). A marker set of 203 SSR primer pairs was applied to gDNA of Gf.99-03, Gf.99-21 and 65-153-18 to verify the pedigree and the phasing of genomic sequences.

P. viticola inoculation of leaf discs for a time course of transcriptomic analysis and follow-up by microscopy

Leaves for high molecular genomic DNA and RNA extraction were taken early in the morning to avoid inhibition of purification steps by photosynthetic starch accumulation.

From three clonally replicated plants of Gf.99-03 at least two leaves were taken from the third and fourth apical insertion. They were cleaned in deionized water. Leaf discs of 1.5 cm diameter were punched out by a surface sterilized cork borer. In total, forty leaf discs were tested per plant. They were placed on 1 % agar in H_2O_{dest} (Gustav Essig GmbH & Co. KG, Mannheim) in rectangular light-transmissible incubation plates covered with a lid (Corning 431111, Corning Incorporated, Corning, NY, USA).
20 leaf discs were mock-inoculated with 40 µl H_2O_{dest} and 20 leaf discs were inoculated with 40 µl of a *P. viticola* zoospore-solution. To obtain these zoospores, sporangia (20.000/ml) were diluted in H_2O_{dest} and stirred every 15 min until zoospore release as controlled by bright field microscopy. Zoospores were separated from sporangia with a filtration step using aquarium filter cotton (Dehner Holding GmbH & Co. KG, Rain, Germany) according to a protocol provided by G. Buchholz (AlPlanta, Neustadt WStr.) (Müllner and Zyprian, 2022). Inoculated leaf discs and mock controls were kept at 22 °C and 100 % humidity with a photoperiod of 16 h light and 8 h dark. The ongoing infestation of all samples was checked by fluorescence microscopy. At every time point (0, 6, 12, 24, 48 hpi), four leaf discs per sample were taken, punched again by a smaller

1.3 cm diameter cork borer and shock frozen in liquid nitrogen. They were stored at -70 °C until RNA-Extraction.

Fluorescence microscopy to follow the infestation progress

Experimentally inoculated leaf discs for transcriptomics were studied in parallel on additional leaf pieces by fluorescence microscopy to observe mycelial ingress. At 24 and 48 hpi leaf pieces of every used leaf were bleached in 1 N KOH for at least 2 h at 65 °C. Subsequently, samples were rinsed twice in H_2O_{deion} and 50 µl of Aniline blue staining solution (Aniline blue 0.05 % v/v in 0.067 M K_2HPO_4, (Hood and Shew, 1996)) was applied on top of leaf discs that were positioned upside down. A cover slide was added and the staining reaction allowed to proceed for at least 10 min at room temperature. Samples were cleaned by careful H_2O_{deion} rinsing and studied using a fluorescence microscope (Leica, DM4000B) with GFP-Filter (Leica A, excitation λ=340-380 nm, dichroic mirror at 400 nm, emission of the filter: LP 425).

Detection of reactive oxygen species

Reactive oxygen species were revealed by Nitrotetrazolium blue chloride (Merck KGaA, Darmstadt, Germany) staining according to the protocol of (Kumar *et al.*, 2014).

Assessment of mycelial growth

Mycelial development was followed to study the inhibition of pathogen progress mediated by the *Rpv12*-Locus after 72 hpi by comparison of genotypes Gf.43-21, Gf.99-03 and 65-153-18. The cultivar `Italia´ served as susceptible control. For every genotype under study three leaves were taken and out of every leaf three leaf discs were stained with alkaline aniline blue (see above). Of every leaf disc five pictures of 100x magnification were taken using a fluorescence microscope with GFP-Filter (Leica A). In total 45 images were taken at 72 hpi from every genotype and analyzed using Fiji and R-Studio (Schindelin *et al.*, 2012, Team, 2018). Very bright leaf veins were erased from the images before the following analysis to avoid potential interference

with mycelial evaluation. The images were converted to binary images by “Fiji” (Schindelin *et al.*, 2012). The amount of white pixels (corresponding to mycelium) was set in relation to the complete image (Müllner and Zyprian, 2022) and the area of mycelium was analysed using Welch’s-T-Test and R Software v4.0.3 (Team, 2018) (https://www.R-project.org/).

RNA extraction of Gf.99-03 tissues and the leaf discs from the *P. viticola* inoculation time course experiment

RNA was extracted from leaf discs (at 0 hpi, 6 hpi, 12 hpi, 24 hpi, 48 hpi after experimental resp. mock inoculation), whole leaves, tendrils, roots and stems of Gf.99-03. Samples were ground using mortar and pistil (Jipo, Czech Republic) with liquid nitrogen. To avoid potential contamination by plant phenolics a pinch of PVP-40 (Sigma Aldrich, Merck KGaA, Darmstadt, Germany) was added to the samples. RNA was isolated using the Spectrum™ Plant Total RNA Kit (Sigma Aldrich, Merck KGaA, Darmstadt, Germany) according to the manufacturer’s instructions. These steps were realized at the Institute for Grapevine Breeding Geilweilerhof, Siebeldingen. The material was stored at -70°C until transfer to CeBiTec Bielefeld.

DNA extraction for PacBio Sequencing of Gf.99-03 and for Illumina sequencing of Gf.99-03, Gf.43-21 and 65-153-18

The high molecular weight genomic DNA of the Gf.99-03 genotype was extracted from 2x2 g leaf samples at the Max Planck Society Genome Centre Cologne.
2.5 g fresh leaf samples of Gf.99-03 and the parental genotypes Gf.43-21 and 65-153-18 were transferred into liquid nitrogen and ground in a mortar. Whole-genomic DNA was extracted from young leaves using the CTAB method (Rosso *et al.*, 2003). The samples were purified using QIAGEN Genomic-tip 500/G. The extraction and sequencing was done at the Center for Biotechnology of Bielefeld University.

Library construction and sequencing

PacBio sequencing of Gf.99-03

The Gf.99-03 DNA samples were sequenced on a Sequel I sequencer at the Max Planck Genome Centre Cologne. The PacBio Sequel libraries were prepared and samples sequenced on two runs using sequencing chemistry 3.0, binding kit 3.0 and DNA Polymerase 3.0 on 1Mv3 SMRT cells (Pacific Biosciences).
The subread BAM files were converted to FASTA format with SMRT Link v5.1.0.26412 (PacBio Reference Guide 2018) and reads smaller than 500 bp were removed.

Illumina sequencing Data of Gf.99-03, Gf.43-21 and 65-153-18

The gDNA samples of Gf.99-03, Gf.43-21 and 65-153-18 were sequenced 2x150 bp paired-end on a NextSeq 500 Illumina sequencing machine (Illumina, San Diego, USA) at the Center for Biotechnology of Bielefeld University. Therefore, Illumina NGS libraries were created following the TruSeq DNA Sample Preparation v2 Guide. The genotypes Gf.99-03 and Gf.43-21 were sequenced on 2x4 lanes and the 65-153-18 genotype was sequenced on 1x4 lanes. The Illumina short read data were trimmed with Trimmomatic-v0.39 (Bolger *et al.*, 2014) with the parameter 'LEADING:34 TRAILING:34 SLIDINGWINDOW:4:15 ILLUMINACLIP:2:34:15 MINLEN:90'.

mRNA sequencing of tissue samples and the infection time course experiment

The mRNA tissue samples of tendrils, roots and stems and the mRNA samples of the infection time course experiment were sequenced 2x75 bp single-end on a NextSeq500 instrument at the Center for Biotechnology of Bielefeld University. The RNA-Seq libraries were prepared using 1,000 ng total RNA following the TrueSeq RNA Sample Preparation v2 Guide. The RNA-Seq data were trimmed with Trimmomatic-v0.39 removing reads shorter than 60 bp length. The other parameters were the same as for Illumina data.

mRNA sequencing leaf samples

An RNA sample of Gf.99-03 leaves was sent to CeGaT GmbH Tübingen. The preparation of the RNA-Seq library was done using 1,000 ng of total RNA according to TruSeq Stranded mRNA (Illumina) Kit and sequenced 2x100 bp paired-end on a NovaSeq 6000 Illumina machine. The data were trimmed with Trimmomatic-v0.39 allowing a minimal read length of 80 bp. The other parameters were set as for the genomic short read data.

Phase-separated genome assembly and pseudochromosome construction

The Gf.99-03 long reads were phase separated based on their 24-mer profile with the TrioCanu module of Canu v1.9 (Koren *et al.*, 2017, Koren *et al.*, 2018). As input, the Illumina reads of Gf.99-03s' parental genotypes Gf.43-21 and 65-153-18 together with the long reads of Gf.99-03 were given to TrioCanu and their 24-mer profile computed. According to the k-mer comparison, the long reads were binned into the parental read subsets and one unassigned subset (bins). The usage of 24-mers was based on the assumption that the parental subsets should be of approximately equal in size and the unassigned one as small as possible. In addition, the average and maximal read length per bin were evaluated and compared with different k-mer runs. K-mer sizes varying from 18-mers to 27-mers were tested (data not shown).

To compute the haplotype assemblies of Gf.99-03, named Gf9921 (maternal haplotype) and Gf9918 (paternal haplotype), both parental read-subset were individually assembled with Canu v1.9 with the settings 'corMhapSensitivity=normal', 'correctedErrorRate=0.085' and 'genomeSize=500m'. All steps were performed on a compute cluster. The assembly parameters were chosen through evaluating different assembly runs. The number of sequences, the total assembly length and the repeat length all reported by Canu or determined with QUAST v5.0.2 (Mikheenko *et al.*, 2018) as well as the plant core genes predicted with the program Benchmarking Universal Single-Copy Orthologs (BUSCO) v3.0.2 utilising the pre-release of the database 'eudicots_odb10' comprising 2,326 plant core genes (Waterhouse *et al.*, 2018) were considered. The parameters resulting in haplotype assemblies having the smallest distance in these values were selected. The Gf9921 and Gf9918 haplotype assemblies

were polished two times with *arrow* (gcpp 1.0.0-SL-release-8.0.0; smrtlink_8.0.0.80529) using the haplotype-specific read bin. Around 25 contigs smaller than 10,000 bp were discarded.

Polished contigs were assigned to pseudochromosomes based on reciprocal best BLAST hits (RBHs) with the PN40024 12X.v2 (Jaillon *et al.*, 2007, Canaguier *et al.*, 2017), the *V. riparia* Gloire de Montpellier genome assembly (Girollet *et al.*, 2019) and with the 'Börner' haplotype assemblies (Frommer *et al.*, 2022). Therefore, a gene prediction was performed on Gf9921 and Gf9918 with AUGUSTUS v3.3 (Stanke and Waack, 2003) using the generated PN40024 parameter file from (Frommer *et al.*, 2022), '--UTR=on' and '--allow_hinted_splicesites=atac'. The protein sequences of these gene predictions were used in a BLASTP search against the longest protein sequences per gene of the PN40024 VCost.v3 gene annotation and against the *V. riparia* and 'Börner' longest protein sequences and *vice versa*. Additionally, a BLASTN search of the contigs against the pseudochromosomes of these three genotypes and *vice versa* was performed. RBHs were constructed from the blast hits with maximal E-value of 0.0001 for both directions and ≥80 % coverage and ≥80 % identity for at least one direction. A contig was only assigned through protein RBHs if it had more than 10 RBHs with a pseudochromosome. If two or more pseudochromosomes had RBHs with a contig, the pseudochromosome with the most RBHs was chosen if the RBH amount was higher than 30 %.

For the formation of pseudochromosomes, protein RBHs with PN40024 were given higher preference than the protein RBHs with the other two cultivars and finally the nucleotide RBHs. Contradicting assignments to different pseudochromosomes of the cultivars were solved through choosing the pseudochromosome of the cultivar showing the most protein RBHs.

Based on the position of the RBHs on the pseudochromosome, the contigs were arranged into pseudochromosomes and connected with 100 N's. To further assign remaining contigs, reciprocal BLASTN hits were iteratively computed between the contigs and the pseudochromosomes of the other haplophase (e. g. Gf9921 contigs

against Gf9918 pseudochromosomes). These construction and assignment steps were repeated until no more contigs could be assigned.

To verify the orientation and order of the contigs on the pseudochromosomes, dot plots of 1-to-1 alignments between the homologous pseudochromosomes of the Gf.99-03 haplotypes and dot plots between the haplotypes and the PN40024 12X.v2 pseudochromosomes were generated with DNAdiff v1.3 and mummerplot v3.5 of the MUMmer package v4.0.0beta2 (Marcais *et al.*, 2018). 470 SSR-repeats (Table S1) were identified on the Gf.99-03 and PN0024 genome sequence assemblies with primer search of the EMBOSS v6.6.0.0 package (Rice *et al.*, 2000) allowing 20 % mismatches. These bioinformatically revealed marker positions were depicted on the dot plots. Contig position, orientation and the pseudochromosome assignment was adapted if the dot plots and SSR markers showed a disagreement between the haplotypes and between the haplotypes and PN40024. Furthermore, contigs on which markers of two different chromosomes mapped, were investigated for mis-assemblies. All-versus-all dot plots generated with D-Genies v1.2.0 (Cabanettes and Klopp, 2018) using minimap2 (Li, 2018) and mappings of the haplotype-specific genomic reads computed with minimap2 v2.17 ('-ax map-pb --secondary=no') were examined for breakpoints. If a sequence region within a contig was found where the read coverage was dropping to less than five and if the all-versus-all dot plot and the SSR markers agreed and would assign the parts of the contig sequence to another pseudochromosome, the contig was split and the sequences assigned according to the data, see Methods S2.

A BUSCO analysis was computed on the Gf.99-03 haplotype assemblies at pseudochromosome level with BUSCO v5.1.2 and the 'eudicots_odb10' dataset (Seppey *et al.*, 2019). The pseudochromosome lengths were visualized with cvit v1.2.1 (Cannon and Cannon, 2011). To validate the genetic purity of the Gf.99-03 assembly, the Gf.99-03 long reads as well as the Gf9921 and Gf9918 haplotype assembly were investigated for contaminations through a BLASTN search against the NCBI nucleotide database v4 (parameter '-evalue 0.0001 -max_target_seqs 1 -max_hsps 1').

Examination of haplotype-phasing

19-mer databases were computed for the Illumina read data of Gf.99-03 and for the Illumina read data of the parental genotypes with Merqury v1.3 (Rhie *et al.*, 2020). With the k-mer databases, haplotype specific k-mer sets (hapmers) were computed. The contig sequences of the haplotype assemblies were evaluated setting the parameters ‘num_switch 10’ and ‘short_range 20,000’.

Gene prediction

For both Gf.99-03 haplotypes, an individual repeat library was generated according to the MAKER ‘Advanced Repeat Library Construction Protocol’ (Campbell *et al.*, 2014). In the course of the protocol the tools MITE Hunter v2.2.26 (Han and Wessler, 2010), LTRharvest (Ellinghaus *et al.*, 2008) and LTRdigest (Steinbiss *et al.*, 2009) of the GenomeTools-1.5.9 package (Gremme *et al.*, 2013), ncbi-blast-2.9.0+ (Altschul *et al.*, 1990, Camacho *et al.*, 2009), RepeatModeler-open-1.0.11 (http://www.repeatmasker.org), ProtExcluder1.2 (Campbell *et al.*, 2014) and RepeatMasker version open-4.0.7 (http://www.repeatmasker.org) were used. The programs were run with the same settings as described in the MAKER ‘Advanced Repeat Library Construction Protocol’ protocol.

Additionally, a *de novo* and genome-guided transcriptome assembly were computed from the paired-end RNA-Seq leaf data with Trinity v2.10.0 (Grabherr *et al.*, 2011). First, the trimmed RNA-Seq data were mapped with GMAP/GSNAP version 2020-09-12 (Wu and Watanabe, 2005) on Gf9921 and on Gf9918 using ‘-B 5 --novelsplicing 1’. The trimmed fastq files were used for the Gf.99-03 *de novo* transcriptome assembly and the primary mappings were processed into a Gf9921 and Gf9918 genome-guided transcriptome assembly using the parameter ‘--genome_guided_max_intron 20000’.

Ab initio gene predictions were carried out on each haplotype with GeneID v1.4.5-master-20200916 with the publicly available *V. vinifera* parameter set (vvinifera.param.Jan_12_2007), GlimmerHMM v3.0.4 (Majoros *et al.*, 2004), SNAP v2006-07-28 (Korf, 2004) and BRAKER2 v2.1.5-master_20201122 (Hoff *et al.*, 2016,

Hoff *et al.*, 2019, Bruna *et al.*, 2021). For GlimmerHMM training, a gene prediction training set was generated with PASA v2.4.1 (Haas *et al.*, 2003) and TransDecoder v5.5.0 (https://github.com/TransDecoder/TransDecoder) according to the PASA wiki (https://github.com/PASApipeline/PASApipeline/wiki/). For the generation of PASA alignment assemblies, the Trinity *de novo* and genome-guided transcriptome assemblies, *Vitis* full-length cDNA sequences (NCBI, Nucleotide DB “Vitis” [Organism] AND complete cds[Title], downloaded on 05.06.2020) and *Vitis* EST sequences (NCBI, txid3603[Organism:exp] AND is_est[filter], downloaded on 27.07.2020) were used. The parameters for PASA were the same as in (Frommer *et al.*, 2022). GlimmerHMM was trained with 7,500 (Gf9921) and 5,500 (Gf9918) PASA/TransDecoder genes and a GlimmerHMM gene prediction performed.

SNAP training was performed three times with MAKER v3.01.03 (Holt and Yandell, 2011, Campbell *et al.*, 2014). First, the Trinity *de novo* and genome-guided transcriptome assemblies, the *Vitis* full-length cDNA sequences, *Vitis* protein sequences (NCBI, Protein DB “Vitis”[Organism], downloaded on 05.06.2020), Viridiplantae sequences (UniProt, release 2020_02, "Viridiplantae [33090]" AND reviewed:yes), RepBase monocotyledons repeat library (RepeatMaskerEdition-20181026) (Bao *et al.*, 2015), the generated haplotype specific repeat library and MAKER transposable element sequences were aligned and processed with MAKER using Exonerate v2.4.0 (Slater and Birney, 2005) and genes derived with MAKERs internal algorithm using the parameter ‘max_dna_len=300000, spli_hit=20000’. SNAP was trained with the MAKER gff file and a gene prediction with SNAP was carried out inside of MAKER (input: SNAP hmm file, MAKER gff file). SNAP training and the gene prediction were repeated two times.

Gene prediction with BRAKER2 using the primary leaf RNA-Seq mappings, the *Vitis* protein sequences, protein sequences from OrthoDB v10.1 eudicotyledons (Kriventseva *et al.*, 2019) and the UniProt Viridiplantae protein data was performed in mode ‘--etpmode’ that enables prediction with GeneMark-ETP+ v4.61 (Lomsadze *et al.*, 2005, Lomsadze *et al.*, 2014, Bruna *et al.*, 2020) to generate a training set for AUGUSTUS version master_v3.4.0_20201212 (Stanke *et al.*, 2008). During the BRAKER2 run, the programs DIAMOND v0.9.24.125 (Buchfink *et al.*, 2015), SAMtools v1.9-180-gf9e1caf (Li *et al.*, 2009), SPALN version 2.3.3f (Gotoh, 2008, Iwata and

Gotoh, 2012), ProtHint Version 2.5.0 and BamTools v2.5.1 (Barnett *et al.*, 2011) were called.

All *ab initio* gene predictions and the data of the initial MAKER gff file were combined with MAKER applying EvidenceModeler v1.1.1 (Haas *et al.*, 2008) with the same parameters as above. tRNAs were predicted with tRNAscan-SE v2.0 (Chan *et al.*, 2021) through MAKER. The gene predictions were quality filtered keeping gene models with AED<0.5 and refined with PASA according to the PASA wiki. Parameters remained the same as used in (Frommer *et al.*, 2022). Transcript variants with identical transcript and CDS sequence of a gene were discarded. The refined gene models were functionally annotated with a BLASTP search against the UniProt/SwissProt DB (release 2020_06) and protein domains identified with InterProScan5 v5.48-83.0 (Jones *et al.*, 2014) and PFAM database v33.1. Genes encoding proteins shorter than 50 amino acids carrying no functional annotation were discarded.

Resistance gene prediction

Resistance gene analogs (RGAs) were identified with RGAugury v2.1.7 (Li *et al.*, 2016) using the protein sequences as input. As components of the pipeline, the tools ncoil (Lupas *et al.*, 1991), PfamScan v1.6 (Mistry *et al.*, 2007), InterProScan5 v5.45-80.0 with the databases Gene3D-4.2.0, Pfam-33.1, SMART-7.1 and SUPERFAMILY-1.75 and Phobius-1.01 (Kall *et al.*, 2004) were run to identify relevant protein domains. In short, this pipeline predicts different protein domains with the mentioned tools and databases and then classifies genes into RGA classes based on the identified domains. RGAugurys' BLAST search against its RGAdb to discard protein sequences showing no hit was disabled.

Differential gene expression analysis

The trimmed data were concurrently mapped on the Gf.99-03 haplotype assemblies with HISAT2 v2.2.0 (Kim *et al.*, 2015) disabling softclipping. The mappings were sorted with SAMtools and the mapped reads per gene counted with featureCounts of the package Rsubread v2.4.3 (Liao *et al.*, 2014) utilising the parameters

‘GTF.featureType="exon", GTF.attrType="gene_id", allowMultiOverlap=TRUE, largestOverlap=TRUE, countMultiMappingReads=TRUE’. The count table was filtered for genes with a count sum smaller or equal than one (no expression) and differentially expressed genes (DEGs) determined with the function DESeq of the package DeSeq2 v1.30.0 (Love *et al.*, 2014) setting ‘ref=”untreated”’ and ‘design= ~ condition + time + condition:time’. The count data were transformed with the variance stabilising transformation setting ‘blind=FALSE’. A principal component analysis (PCA) plot was generated with the function plotPCA. Additionally, the count data were investigated for a significant difference over time using the likelihood-ratio test (LRT) with the DESeq function (‘test=”LTR”, reduced = ~ condition + time’). The counts of genes with an adjusted p-value smaller than 0.05 were plotted with ggplot2 v3.3.5 (Wickham, 2009). Using the R library UpSetR v1.4.0 (Lex *et al.*, 2014, Conway *et al.*, 2017), shared DEGs between the time points were visualized and DEGs were functionally annotated against the KEGG database (Kanehisa *et al.*, 2021) with KAAS v2.1 (Moriya *et al.*, 2007) using the amino-acid sequence, eukaryotes as GENES dataset, BLAST and single-directional best hit.

Design of new markers

For design of specific markers within the *Rpv12* locus, the region of the NLR-Cluster of the Gf9918 haplotype was screened for di- tri-, tetra-, penta-, and hexa-repeats using the WebSat online program (Martins *et al.*, 2009). Repeat regions with a maximum of 500 bp flanking sequence were blasted against PN40024 12X.v0, haplotype Gf9918 as well as Gf9921. Primer3Plus (Untergasser *et al.*, 2007) was used for testing PCR analytical suitability focused on primer length between 18-30 bp, melting temperature of 60 °C and GC content between 40-50 %. Resulting, primer pairs were first tested on `Traminer´ and `Kunbarat´ by PCR and Agarose gel electrophoresis and followed by testing for specificity on 18 genotypes (Table S11) by SSR marker analysis using fluorescently labelled forward primers (Table S10, Table S11 & Table S1).

Analysis of *Rpv12* resistance locus

Based on a primer search alignment of the *Rpv12* QTL delimiting marker UDV-014 and UDV-370 and the other *Rpv12* related SSR markers (Venuti *et al.*, 2013), the *Rpv12* locus position was determined on chr14 of Gf9921, Gf9918 and PN40024 12X.v2. As Gf9918 is the genetic source of Gf.99-03s' *Rpv12* locus, significantly DEGs and RGAs of Gf9918 were considered as potential candidate genes for *Rpv12* mediated *P. viticola* resistance. The *Rpv12* locus was visualized with the R package gggenomes (https://github.com/thackl/gggenomes) allowing blast hits with 85 % query coverage and identity.

Orthologues between the *Rpv12* genes of the haplotypes and PN40024 were determined with OrthoFinder v2.3.11 (Emms and Kelly, 2015, Emms and Kelly, 2019) using the longest protein sequences. For the Gf9918 candidate genes, gene mates positioned in the *Rpv12* locus of the susceptible haplotype Gf9921 were identified based on the orthogroups, their position on the locus and their functional annotation. To further validate the uniqueness of the candidates, the protein sequences were used in a BLASTP search against the Gf.99-03 proteins and against the proteins of PN40024 VCost.v3. Only protein sequences with a query coverage and identity above or equal to 70 % were further investigated. Protein alignments were built with CLC Genomics Workbench v21.0.1 (gap open cost 5.0, gap extension cost 1.0, free end gap cost).

Accession numbers

The SMRT sequence data of Gf.99-03 (ERS11455446), the short sequence data of Gf.99-03 (ERS11455447) and of the parental genotypes Gf.43-21 (ERS11455445) and 65-153-18 (ERS11455444) as well as the haplotype sequences of Gf.99-03 (ERS11872527) have been submitted to the DDBJ/EMBL/GenBank databases under project number PRJEB46083. The haplotype assemblies Gf9921 and Gf9921 at chromosome level with structural and functional annotation are available at DOI 10.4119/unibi/2962800.

The RNA-Seq data of Gf.99-03 tissue leaf (ERS11536190), tendril (ERS11536191), root (ERS11536192) and stem (ERS11536193) have been deposited in DDBJ/EMBL/GenBank under project number PRJEB46084. The RNA-Seq data of the *P. viticola* infection experiment on Gf.99-03 leaves (ERS11536194-ERS11536203) have been submitted to the DDBJ/EMBL/GenBank databases under project number PRJEB46085.

Acknowledgements

We would like to thank the members of the Department of Genetics and Genomics of Plants for their support and active exchange.

We thank Abdulkhalek Abdaan for technical assistance, Lucie Gebauer for help with R-Studio and Thomas Proksch for support with Fiji.

This work was funded by the European INTERREG V Upper Rhine Program (“Vitifutur”) and the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG Zy11/11-1).

This work was supported by the BMBF-funded de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI).

Author contributions

SM did the laboratory work including SSR marker analysis, the *P. viticola* infection experiment and the extraction of RNA at the JKI (Julius Kühn-Institute, Institute for Grapevine Breeding Geilweilerhof, Siebeldingen, Germany). SM performed the pedigree analysis, the analysis of the mycelial growth, designed and tested new SSR markers and created figures and tables for these parts. BF, SM and EZ designed the infection time-course experiment. BF did the bioinformatics and analyses concerning the computation of the genome assembly and gene annotation, the evaluation of haplotype phasing, the DEG analysis and the comparisons of the *Rpv12* locus. Bioinformatic analyses were done at Bielefeld University, Faculty of Biology & Center

for Biotechnology (CeBiTec). BF created the figures and tables for the bioinformatic analyses.

DH extracted the DNA for the Illumina sequencing and PV did the library preparation and sequencing of the Illumina data. BH did the DNA extraction, library preparation and sequencing of the PacBio long read data.

DH and BW supervised the work at Bielefeld University and RT and EZ supervised the work at JKI Geilweilerhof.

BF and SM wrote the draft manuscript. EZ and BW edited the manuscript. All authors have read and approved the final manuscript.

Conflict of interest

The authors declare that they have no competing interests.

Data availability statement

All data is publicly available under the accession numbers indicated above or included in the supplemental information of this publication.

Supporting information

Supplemental figures:

Figure S1 Combined staining of inoculated leaf discs of Gf.99-03 with Auramin O *and* Calcofluor white.

Figure S2 All-versus-all dotplots between the pseudochromosomes of the haplotypes and the grapevine reference genotype PN40024.

Figure S3 Mis-assembly of the Gf9921 pseudochromosome chr08.

Figure S4 Spectra-copy number plot of the Gf.99-03 genome assembly.

Figure S5 KEGG pathway 'plant-pathogen interaction'.

Figure S6 Amount of Gf.99-03 genes associated with various RGA classes.

Figure S7 Domain comparisons.

Supplemental tables:

Table S1 SSR-Marker bioinformatic prediction and analysis.

Table S2 SSR marker data to confirm the pedigree of Gf.99-03.

Table S3 Read binning of the Gf.99-03 long reads.

Table S4 Statistics of the chromosome assignment of the Gf.99-03 haplotypes.

Table S5 Pseudochromosome lengths of the Gf.99-03 haplotype assembly.

Table S6 k-mer based phasing assessment with Merqury.

Table S7 RNA-Seq data of Gf.99-03s' tissues.

Table S8 RNA-seq data of Plasmopara viticola infection experiment.

Table S9 Plant-pathogen interaction [PATH:ko04626].

Table S10 Design of new Rpv12 correlated SSR-markers.

Table S11 Testing of new *Rpv12* correlated markers.

Supplemental methods:

Methods S1 - Monitoring necrosis around stomata

Methods S2 - Correction of mis-assemblies

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403-410.

Bao, W., Kojima, K.K. and Kohany, O. (2015) Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*, **6**, 11.

Barnett, D.W., Garrison, E.K., Quinlan, A.R., Stromberg, M.P. and Marth, G.T. (2011) BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics*, **27**, 1691-1692.

Bavaresco, L. (2019) Impact of grapevine breeding for disease resistance on the global wine industry. *Acta Horticulturae*, 7-14.

Bent, A.F., Kunkel, B.N., Dahlbeck, D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J. and Staskawicz, B.J. (1994) RPS2 of Arabidopsis thaliana: a Leucine-Rich Repeat Class of Plant Disease Resistance Genes. *Science*, **265**, 1856-1860.

Bentham, A.R., De la Concepcion, J.C., Mukhi, N., Zdrzalek, R., Draeger, M., Gorenkin, D., Hughes, R.K. and Banfield, M.J. (2020) A molecular roadmap to the plant immune system. *Journal of Biological Chemistry*, **295**, 14916-14935.

Bigeard, J., Colcombet, J. and Hirt, H. (2015) Signaling mechanisms in pattern-triggered immunity (PTI). *Molecular Plant*, **8**, 521-539.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114-2120.

Borhan, M.H., Holub, E.B., Beynon, J.L., Rozwadowski, K. and Rimmer, S.R. (2004) The arabidopsis TIR-NB-LRR gene RAC1 confers resistance to Albugo candida (white rust) and is dependent on EDS1 but not PAD4. *Molecular Plant-Microbe Interactions*, **17**, 711-719.

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

Bruna, T., Lomsadze, A. and Borodovsky, M. (2020) GeneMark-EP+: eukaryotic gene prediction with self-training in the space of genes and proteins. *NAR Genomics and Bioinformatics*, **2**, lqaa026.

Buchfink, B., Xie, C. and Huson, D.H. (2015) Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, **12**, 59-60.

Cabanettes, F. and Klopp, C. (2018) D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ*, **6**, e4958.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 421.

Campbell, M.S., Law, M., Holt, C., Stein, J.C., Moghe, G.D., Hufnagel, D.E., Lei, J., Achawanantakun, R., Jiao, D., Lawrence, C.J., Ware, D., Shiu, S.H., Childs, K.L., Sun, Y., Jiang, N. and Yandell, M. (2014) MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations. *Plant Physiology*, **164**, 513-524.

Canaguier, A., Grimplet, J., Di Gaspero, G., Scalabrin, S., Duchêne, E., Choisne, N., Mohellibi, N., Guichard, C., Rombauts, S., Le Clainche, I., Bérard, A., Chauveau, A., Bounon, R., Rustenholz, C., Morgante, M., Le Paslier, M.C., Brunel, D. and Adam-Blondon, A.F. (2017) A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3). *Genomics Data*, **14**, 56-62.

Cannon, E.K. and Cannon, S.B. (2011) Chromosome visualization tool: a whole genome viewer. *International Journal of Plant Genomics*, **2011**, 373875.

Chan, P.P., Lin, B.Y., Mak, A.J. and Lowe, T.M. (2021) tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Research*, **49**, 9077-9096.

Chattopadhyay, A., Purohit, J., Tiwari, K.K. and Deshmukh, R. (2019) Targeting Transcription Factors for Plant Disease Resistance:Shifting Paradigm. *Current Science*, **117**.

Chitarrini, G., Riccadonna, S., Zulini, L., Vecchione, A., Stefanini, M., Larger, S., Pindo, M., Cestaro, A., Franceschi, P., Magris, G., Foria, S., Morgante, M., Di Gaspero, G. and Vrhovsek, U. (2020) Two-omics data revealed commonalities and differences between Rpv12- and Rpv3-mediated resistance in grapevine. *Scientific Reports*, **10**, 12193.

Consortium, R.E.X. (2016) Combining Selective Pressures to Enhance the Durability of Disease Resistance Genes. *Frontiers in Plant Science*, **7**, 1916.

Conway, J.R., Lex, A. and Gehlenborg, N. (2017) UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*, **33**, 2938-2940.

Cui, H., Tsuda, K. and Parker, J.E. (2015) Effector-triggered immunity: from pathogen perception to robust defense. *Annual Review of Plant Biology*, **66**, 487-511.

Dangl, J.L. and Jones, J.D. (2001) Plant pathogens and integrated defence responses to infection. *Nature*, **411**, 826-833.

DeLille, J.M., Sehnke, P.C. and Ferl, R.J. (2001) The Arabidopsis 14-3-3 Family of Signaling Regulators. *Plant Physiology*, **126**, 35-38.

Di Gaspero, G. and Cipriani, G. (2003) Nucleotide binding site/leucine-rich repeats, Pto-like and receptor-like kinases related to disease resistance in grapevine. *Mol Genet Genomics*, **269**, 612-623.

Di Gaspero, G., Copetti, D., Coleman, C., Castellarin, S.D., Eibach, R., Kozma, P., Lacombe, T., Gambetta, G., Zvyagin, A., Cindric, P., Kovacs, L., Morgante, M. and Testolin, R. (2012) Selective sweep at the Rpv3 locus during grapevine breeding for downy mildew resistance. *Theoretical and Applied Genetics*, **124**, 277-286.

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

Dougherty, M.K. and Morrison, D.K. (2004) Unlocking the code of 14-3-3. *Journal of Cell Science*, **117**, 1875-1884.

Dubrow, Z., Sunitha, S., Kim, J.G., Aakre, C.D., Girija, A.M., Sobol, G., Teper, D., Chen, Y.C., Ozbaki-Yagan, N., Vance, H., Sessa, G. and Mudgett, M.B. (2018) Tomato 14-3-3 Proteins Are Required for Xv3 Disease Resistance and Interact with a Subset of Xanthomonas euvesicatoria Effectors. *Molecular Plant Microbe Interactions*, **31**, 1301-1311.

Eibach, R., Zyprian, E., Welter, L. and Töpfer, R. (2007) The use of molecular markers for pyramiding resistance genes in grapevine breeding. *Vitis*, **46**, 120-124.

Eisenmann, B., Czemmel, S., Ziegler, T., Buchholz, G., Kortekamp, A., Trapp, O., Rausch, T., Dry, I. and Bogs, J. (2019) Rpv3-1 mediated resistance to grapevine downy mildew is associated with specific host transcriptional responses and the accumulation of stilbenes. *BMC Plant Biology*, **19**, 343.

Ellinghaus, D., Kurtz, S. and Willhoeft, U. (2008) LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics*, **9**, 18.

Emms, D.M. and Kelly, S. (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, **16**, 157.

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

Feechan, A., Anderson, C., Torregrosa, L., Jermakow, A., Mestre, P., Wiedemann-Merdinoglu, S., Merdinoglu, D., Walker, A.R., Cadle-Davidson, L., Reisch, B., Aubourg, S., Bentahar, N., Shrestha, B., Bouquet, A., Adam-Blondon, A.F., Thomas,

M.R. and Dry, I.B. (2013) Genetic dissection of a TIR-NB-LRR locus from the wild North American grapevine species Muscadinia rotundifolia identifies paralogous genes conferring resistance to major fungal and oomycete pathogens in cultivated grapevine. *The Plant Journal*, **76**, 661-674.

Ferl, R.J. (1996) 14-3-3 Proteins and Signal Transduction. *Annual Review of Plant Physiology and Plant Molecular Biology*, **47**, 49-73.

Foria, S., Magris, G., Jurman, I., Schwope, R., De Candido, M., De Luca, E., Ivanisevic, D., Morgante, M. and Di Gaspero, G. (2022) Extent of wild-to-crop interspecific introgression in grapevine (Vitis vinifera) as a consequence of resistance breeding and implications for the crop species definition. *Horticultural Research*, **9**, uhab010.

Foria, S., Magris, G., Morgante, M., Di Gaspero, G. and Debener, T. (2018) The genetic background modulates the intensity of Rpv3-dependent downy mildew resistance in grapevine. *Plant Breeding*, **137**, 220-228.

Frommer, B., Hausmann, L., Holtgräwe, D., Viehöver, P., Hüttel, B., Reinhardt, R., Töpfer, R. and Weisshaar, B. (2022) A fully phased interspecific grapevine rootstock genome sequence representing V. riparia and V. cinerea and allele-aware annotation of the phylloxera resistance locus Rdv1. *bioRxiv*, posted 2022-2007-2008.

Fu, P., Wu, W., Lai, G., Li, R., Peng, Y., Yang, B., Wang, B., Yin, L., Qu, J., Song, S. and Lu, J. (2020) Identifying Plasmopara viticola resistance Loci in grapevine (Vitis amurensis) via genotyping-by-sequencing-based QTL mapping. *Plant Physiology and Biochemistry*, **154**, 75-84.

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

Girollet, N., Rubio, B., Lopez-Roques, C., Valiere, S., Ollat, N. and Bert, P.F. (2019) De novo phased assembly of the Vitis riparia grape genome. *Scientific Data*, **6**, 127.

Gotoh, O. (2008) A space-efficient and accurate method for mapping and aligning cDNA sequences onto genomic sequence. *Nucleic Acids Research*, **36**, 2630-2638.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. and Regev, A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**, 644-652.

Greenberg, J.T. (1997) Programmed Cell Death in Plant-Pathogen Interactions. *Annual Review of Plant Physiology and Plant Molecular Biology*, **48**, 525-545.

Gremme, G., Steinbiss, S. and Kurtz, S. (2013) GenomeTools: a comprehensive software library for efficient processing of structured genome annotations. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **10**, 645-656.

Haas, B.J., Delcher, A.L., Mount, S.M., Wortman, J.R., Smith Jr, R.K., Hannick, L.I., Maiti, R., Catherine M. Ronning, Rusch, D.B., Town, C.D., Salzberg, S.L. and White, O. (2003) Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Research*, **31**, 5654-5666.

Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R. and Wortman, J.R. (2008) Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology*, **9**, R7.

Han, G.Z. (2019) Origin and evolution of the plant immune system. *New Phytologist*, **222**, 70-83.

Han, Y. and Wessler, S.R. (2010) MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Research*, **38**, e199.

Heyman, L., Höfle, R., Kicherer, A., Trapp, O., Ait Barka, E., Töpfer, R. and Höfte, M. (2021) The Durability of Quantitative Host Resistance and Variability in Pathogen

Virulence in the Interaction Between European Grapevine Cultivars and Plasmopara viticola. *Frontiers in Agronomy*, **3**, 684023.

Hoff, K.J., Lange, S., Lomsadze, A., Borodovsky, M. and Stanke, M. (2016) BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*, **32**, 767-769.

Hoff, K.J., Lomsadze, A., Borodovsky, M. and Stanke, M. (2019) Whole-Genome Annotation with BRAKER. *Methods in Molecular Biology*, **1962**, 65-95.

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

Hood, M.E. and Shew, H.D. (1996) Applications of KOH-aniline blue fluorescence in the study of plant-fungal interactions. *Phytopathology*, **86**, 704-708.

Hua, J. and Meyerowitz, E.M. (1998) Ethylene Responses Are Negatively Regulated by a Receptor Gene Family in Arabidopsis thaliana. *Cell*, **94**, 261-271.

Hulbert, S.H., Webb, C.A., Smith, S.M. and Sun, Q. (2001) Resistance gene complexes: evolution and utilization. *Annual Review of Phytopathology*, **39**, 285-312.

Iwata, H. and Gotoh, O. (2012) Benchmarking spliced alignment programs including Spaln2, an extended version of Spaln that incorporates additional species-specific features. *Nucleic Acids Research*, **40**, e161.

Jaillon, O., Aury, J.M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Hugueney, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyere, C., Billault, A., Segurens, B., Gouyvenoux, M., Ugarte, E., Cattonaro, F., Anthouard, V., Vico, V., Del Fabbro, C., Alaux, M., Di Gaspero, G., Dumas, V., Felice, N., Paillard, S., Juman, I., Moroldo, M., Scalabrin, S., Canaguier, A., Le Clainche, I., Malacrida, G., Durand, E., Pesole, G., Laucou, V., Chatelet, P., Merdinoglu, D., Delledonne, M., Pezzotti, M., Lecharny, A., Scarpelli, C., Artiguenave, F., Pe, M.E., Valle, G., Morgante, M., Caboche, M., Adam-Blondon, A.F., Weissenbach, J., Quetier, F., Wincker, P. and French-Italian Public Consortium for Grapevine Genome, C. (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, **449**, 463-467.

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

Judelson, H.S. and Ah-Fong, A.M.V. (2019) Exchanges at the Plant-Oomycete Interface That Influence Disease. *Plant Physiology*, **179**, 1198-1211.

Kall, L., Krogh, A. and Sonnhammer, E.L. (2004) A combined transmembrane topology and signal peptide prediction method. *Journal of Molecular Biology*, **338**, 1027-1036.

Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. and Tanabe, M. (2021) KEGG: integrating viruses and cellular organisms. *Nucleic Acids Research*, **49**, D545-D551.

Kast, W.K. (2001) Inter-isolate variation of virulence of *Plasmopara viticola* on resistant vine varieties. *IOBC/WPRS Bulletin*, **24**, 45-49.

Kim, D., Langmead, B. and Salzberg, S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, **12**, 357-360.

Kim, H., O'Connell, R., Maekawa-Yoshikawa, M., Uemura, T., Neumann, U. and Schulze-Lefert, P. (2014) The powdery mildew resistance protein RPW8.2 is carried on VAMP721/722 vesicles to the extrahaustorial membrane of haustorial complexes. *The Plant Journal*, **79**, 835-847.

Kim, S.H., Kwon, S.I., Saha, D., Anyanwu, N.C. and Gassmann, W. (2009) Resistance to the Pseudomonas syringae effector HopA1 is governed by the TIR-NBS-LRR protein RPS6 and is enhanced by mutations in SRFR1. *Plant Physiology*, **150**, 1723-1732.

Koleda, I. (1975) Ergebnisse von Kreuzungen zwischen Vitis amurensis und Vitis vinifera in der Züchtung frostwiderstandsfähiger Reben. *Vitis*, **14**, 1-5.

Koren, S., Rhie, A., Walenz, B.P., Dilthey, A.T., Bickhart, D.M., Kingan, S.B., Hiendleder, S., Williams, J.L., Smith, T.P.L. and Phillippy, A.M. (2018) De novo assembly of haplotype-resolved genomes with trio binning. *Nature Biotechnology*, **36**, 1174-1182.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H. and Phillippy, A.M. (2017) Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, **27**, 722-736.

Korf, I. (2004) Gene finding in novel genomes. *BMC Bioinformatics*, **5**, 59.

Kortekamp, A., Welter, L., Vogt, S., Knoll, A., Schwander, F., Töpfer, R. and Zyprian, E. (2008) Identification, isolation and characterization of a CC-NBS-LRR candidate disease resistance gene family in grapevine. *Molecular Breeding*, **22**, 421-432.

Kriventseva, E.V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simao, F.A. and Zdobnov, E.M. (2019) OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Research*, **47**, D807-D811.

Kumar, D., Yusuf, M.A., Singh, P., Sardar, M. and Sarin, N.B. (2014) Histochemical detection of superoxide and H2O2 accumulation in Brassica juncea seedlings. *bio-protocol*, **4**, e1108.

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R. and Pfister, H. (2014) UpSet: Visualization of Intersecting Sets. *IEEE Trans Vis Comput Graph*, **20**, 1983-1992.

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.

Li, N.Y., Zhou, L., Zhang, D.D., Klosterman, S.J., Li, T.G., Gui, Y.J., Kong, Z.Q., Ma, X.F., Short, D.P.G., Zhang, W.Q., Li, J.J., Subbarao, K.V., Chen, J.Y. and Dai, X.F. (2018) Heterologous Expression of the Cotton NBS-LRR Gene GbaNA1 Enhances Verticillium Wilt Resistance in Arabidopsis. *Frontiers in Plant Science*, **9**, 119.

Li, P., Quan, X., Jia, G., Xiao, J., Cloutier, S. and You, F.M. (2016) RGAugury: a pipeline for genome-wide prediction of resistance gene analogs (RGAs) in plants. *BMC Genomics*, **17**, 852.

Li, Z., Huang, J., Wang, Z., Meng, F., Zhang, S., Wu, X., Zhang, Z. and Gao, Z. (2019) Overexpression of Arabidopsis Nucleotide-Binding and Leucine-Rich Repeat Genes RPS2 and RPM1(D505V) Confers Broad-Spectrum Disease Resistance in Rice. *Frontiers in Plant Science*, **10**, 417.

Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923-930.

Lodhi, M.A. and Reisch, B.I. (1995) Nuclear DNA content of Vitis species, cultivars, and other genera of the Vitaceae. *Theoretical and Applied Genetics*, **90**, 11-16.

Lomsadze, A., Burns, P.D. and Borodovsky, M. (2014) Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Research*, **2014**, e119.

Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y.O. and Borodovsky, M. (2005) Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Research*, **33**, 6494-6506.

Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Marchler, G.H., Song, J.S., Thanki, N., Yamashita, R.A., Yang,

M., Zhang, D., Zheng, C., Lanczycki, C.J. and Marchler-Bauer, A. (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Research*, **48**, D265-D268.

Lupas, A., Van Dyke, M. and Stock, J. (1991) Predicting coiled coils from protein sequences. *Science*, **252**, 1162-1164.

Majoros, W.H., Pertea, M. and Salzberg, S.L. (2004) TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics*, **20**, 2878-2879.

Marcais, G., Delcher, A.L., Phillippy, A.M., Coston, R., Salzberg, S.L. and Zimin, A. (2018) MUMmer4: A fast and versatile genome alignment system. *PLoS Computational Biology*, **14**, e1005944.

Marguerit, E., Boury, C., Manicki, A., Donnart, M., Butterlin, G., Nemorin, A., Wiedemann-Merdinoglu, S., Merdinoglu, D., Ollat, N. and Decroocq, S. (2009) Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. *Theoretical and Applied Genetics*, **118**, 1261-1278.

Martins, W.S., Lucas, D.C., Neves, K.F. and Bertioli, D.J. (2009) WebSat--a web software for microsatellite marker development. *Bioinformation*, **3**, 282-283.

Massonnet, M., Vondras, A.M., Cochetel, N., Riaz, S., Pap, D., Minio, A., Figueroa-Balderas, R., Walker, M.A. and Cantu, D. (2022) Haplotype-resolved powdery mildew resistance loci reveal the impact of heterozygous structural variation on NLR genes in Muscadinia rotundifolia. *G3 Genes Genomes Genetics*.

Mathe, C., Garda, T., Freytag, C. and M, M.H. (2019) The Role of Serine-Threonine Protein Phosphatase PP2A in Plant Oxidative Stress Signaling-Facts and Hypotheses. *International Journal of Molecular Sciences*, **20**.

Merdinoglu, D., Wiedeman-Merdinoglu, S., Coste, P., Dumas, V., Haetty, S., Butterlin, G. and Greif, C. (2003) Genetic Analysis of Downy Mildew Resistance Derived from Muscadinia Rotundifolia. *Acta Horticulturae*, **603**, 451-456.

Meyers, B.C., Kozik, A., Griego, A., Kuang, H. and Michelmore, R.W. (2003) Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *The Plant Cell*, **15**, 809-834.

Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D. and Gurevich, A. (2018) Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*, **34**, i142-i150.

Mistry, J., Bateman, A. and Finn, R.D. (2007) Predicting active site residue annotations in the Pfam database. *BMC Bioinformatics*, **8**, 298.

Monteiro, F. and Nishimura, M.T. (2018) Structural, Functional, and Genomic Diversity of Plant NLR Proteins: An Evolved Resource for Rational Engineering of Plant Immunity. *Annual Review of Phytopathology*, **56**, 243-267.

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

Moussatche, P. and Klee, H.J. (2004) Autophosphorylation activity of the Arabidopsis ethylene receptor multigene family. *Journal of Biological Chemistry*, **279**, 48734-48741.

Müllner, S., Röckel, F., Maul, E. and Zyprian, E. (2020) 'Kunbarat' and 'Kunleany' – full not half-siblings. *Vitis*, **59**, 91-92.

Müllner, S. and Zyprian, E. (2022) The effect of homozygosity of locus Rpv12 on downy mildew resistance of grapevine leaves. *Vitis*, **61**, 71-76.

Olmo, H.P. (1986) The potential role of (vinifera x rotundifolia) hybrids in grape variety improvement. *Experientia*, **42**, 921-926.

Peremyslov, V.V., Morgun, E.A., Kurth, E.G., Makarova, K.S., Koonin, E.V. and Dolja, V.V. (2013) Identification of myosin XI receptors in Arabidopsis defines a distinct class of transport vesicles. *The Plant Cell*, **25**, 3022-3038.

Peressotti, E., Wiedemann-Merdinoglu, S., Delmotte, F., Bellin, D., Di Gaspero, G., Testolin, R., Merdinoglu, D. and Mestre, P. (2010) Breakdown of resistance to

grapevine downy mildew upon limited deployment of a resistant variety. *BMC Plant Biology*, **10**, 147.

Possamai, T., Migliaro, D., Gardiman, M., Velasco, R. and De Nardi, B. (2020) Rpv Mediated Defense Responses in Grapevine Offspring Resistant to Plasmopara viticola. *Plants*, **9**.

Qi, Y., Tsuda, K., Nguyen le, V., Wang, X., Lin, J., Murphy, A.S., Glazebrook, J., Thordal-Christensen, H. and Katagiri, F. (2011) Physical association of Arabidopsis hypersensitive induced reaction proteins (HIRs) with the immune receptor RPS2. *Journal of Biological Chemistry*, **286**, 31297-31307.

Reynolds, A.G. (2022) Viticultural and vineyard management practices and their effects on grape and wine quality. In *Managing Wine Quality (Second Edition)*.

Rhie, A., Walenz, B.P., Koren, S. and Phillippy, A.M. (2020) Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biology*, **21**, 245.

Rice, P., Longden, I. and Bleasby, A. (2000) EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics*, **16**, 276-277.

Rockel, F., Trapp, O., Zyprian, E., Hausmann, L., Migliaro, D., Vezzulli, S., Topfer, R. and Maul, E. (2021) A 'Regent' pedigree update: ancestors, offspring and their confirmed resistance loci. *Vitis*, **60**, 189-193.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B. (2003) An *Arabidopsis thaliana* T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics. *Plant Molecular Biology*, **53**, 247-259.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P. and Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. *Nature Methods*, **9**, 676-682.

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

Seppey, M., Manni, M. and Zdobnov, E.M. (2019) BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods Mol Biol*, **1962**, 227-245.

Slater, G.S. and Birney, E. (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, **6**, 31.

Stanke, M., Diekhans, M., Baertsch, R. and Haussler, D. (2008) Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*, **24**, 637-644.

Stanke, M. and Waack, S. (2003) Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*, **19**, ii215–ii225.

Statistisches Bundesamt (Destatis) (2016) Grunderhebung der Rebflächen 2015. In *Fachserie 3 Reihe 3.1.5*. Wiesbaden.

Steinbiss, S., Willhoeft, U., Gremme, G. and Kurtz, S. (2009) Fine-grained annotation and classification of de novo predicted LTR retrotransposons. *Nucleic Acids Research*, **37**, 7002-7013.

Tameling, W.I., Vossen, J.H., Albrecht, M., Lengauer, T., Berden, J.A., Haring, M.A., Cornelissen, B.J. and Takken, F.L. (2006) Mutations in the NB-ARC domain of I-2 that impair ATP hydrolysis cause autoactivation. *Plant Physiology*, **140**, 1233-1245.

Team, R.C. (2018) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Todesco, M., Balasubramanian, S., Hu, T.T., Traw, M.B., Horton, M., Epple, P., Kuhns, C., Sureshkumar, S., Schwartz, C., Lanz, C., Laitinen, R.A., Huang, Y., Chory, J., Lipka, V., Borevitz, J.O., Dangl, J.L., Bergelson, J., Nordborg, M. and Weigel, D. (2010) Natural allelic variation underlying a major fitness trade-off in Arabidopsis thaliana. *Nature*, **465**, 632-636.

Töpfer, R., Hausmann, L., Harst, M., Maul, E., Zyprian, E. and Eibach, R. (2011) New horizons for grapevine breeding. In *Fruit, vegetable and cereal science and biotechnology,* : Global Science Books, pp. 79-100.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J.A. (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research*, **35**, W71-74.

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

Wang, W., Wen, Y., Berkey, R. and Xiao, S. (2009) Specific targeting of the Arabidopsis resistance protein RPW8.2 to the interfacial membrane encasing the fungal Haustorium renders broad-spectrum resistance to powdery mildew. *The Plant Cell*, **21**, 2898-2913.

Waterhouse, R.M., Seppey, M., Simao, F.A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E.V. and Zdobnov, E.M. (2018) BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Molecular Biology and Evolution*, **35**, 543-548.

Wei, H., Liu, J., Guo, Q., Pan, L., Chai, S., Cheng, Y., Ruan, M., Ye, Q., Wang, R., Yao, Z., Zhou, G. and Wan, H. (2020) Genomic Organization and Comparative Phylogenic Analysis of NBS-LRR Resistance Gene Family in Solanum pimpinellifolium and Arabidopsis thaliana. *Evolutionary Bioinformatics Online*, **16**, 1176934320911055.

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

Wickham, H. (2009) *ggplot2 - Elegant Graphics for Data Analysis*: Springer.

Williams, S.J., Sornaraj, P., deCourcy-Ireland, E., Menz, R.I., Kobe, B., Ellis, J.G., Dodds, P.N. and Anderson, P.A. (2011) An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Molecular Plant-Microbe Interactions*, **24**, 897-906.

Wingerter, C., Eisenmann, B., Weber, P., Dry, I. and 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.

Wu, T.D. and Watanabe, C.K. (2005) GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics*, **21**, 1859-1875.

Xiao, S., Charoenwattana, P., Holcombe, L. and Turner, J.G. (2003) The Arabidopsis genes RPW8.1 and RPW8.2 confer induced resistance to powdery mildew diseases in tobacco. *Molecular Plant-Microbe Interactions*, **16**, 289-294.

Xiao, S., Ellwood, S., Calis, O., Patrick, E., Li, T., Coleman, M. and Turner, J.G. (2001) Broad-spectrum mildew resistance in Arabidopsis thaliana mediated by RPW8. *Science*, **291**, 118-120.

Yang, L., Qin, L., Liu, G., Peremyslov, V.V., Dolja, V.V. and Wei, Y. (2014) Myosins XI modulate host cellular responses and penetration resistance to fungal pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 13996-14001.

Zendler, D., Schneider, P., Töpfer, R. and Zyprian, E. (2017) Fine mapping of Ren3 reveals two loci mediating hypersensitive response against Erysiphe necator in grapevine. *Euphytica*, **213**.

Zendler, D., Topfer, R. and Zyprian, E. (2021) Confirmation and Fine Mapping of the Resistance Locus Ren9 from the Grapevine Cultivar 'Regent'. *Plants*, **10**.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D., Boller, T. and Felix, G. (2006) Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. *Cell*, **125**, 749-760.

Tables

Table 1 Statistics of the Gf.99-03 assembly.

Gf9921 denotes the maternal haplotype assembly and Gf9918 the paternal haplotype assembly, both derived from phase-separated reads of Gf.99-03.

	Gf9921	**Gf9918**	**Gf.99-03 assembly**
contigs	1,245	1,253	2,498
size [bp]	511,095,651	544,110,372	1,055,206,023
largest contig [bp]	14,810,751	18,073,517	18,073,517
N50 length [bp]	3,294,480	3,920,956	3,763,673
N50 contigs	42	38	79
N90 length [bp]	200,578	203,624	201,530
N90 contigs	239	235	474

Table 2 List of bioinformatically identified potential genes encoding domain structures associated with RGA classes.

The table shows the number of genes identified for each RGA class. 'Gf.99-03 assembly' contains the total number of RGAs. NBS = Nucleotide Binding Site; CNL = coiled-coil, nucleotide-binding site and leucine-rich repeat protein; TNL= TIR, nucleotide-binding site and leucine-rich repeat protein; CN= coiled-coil and nucleotide-binding site; TN=TIR and nucleotide-binding site; NL= nucleotide-binding site and leucine-rich repeat protein; TX= TIR-unknown site/domain; RLP= Receptor like protein; RLK= Receptor like kinase; TM-CC= Transmembrane and coiled coil.

	Gf9921	**Gf9918**	**Gf.99-03 assembly**
NBS	62	59	121
CNL	106	159	265
TNL	20	25	45
CN	35	30	65
TN	0	6	6
NL	110	102	212
TX	12	33	45
others	6	3	9
RLP	210	231	441
RLK	748	719	1,467
TM-CC	979	901	1,880
total	**2,288**	**2,268**	**4,556**

Table 3 *Rpv12* locus specific candidate genes of Gf9918. The designation of the *Rpv12* candidate gene, the putative function (put. function) based on prediction analysis as described (gene prediction), the detected encoded domains as well as the adjusted p-value (adj. p-value) and the cluster of the locus are described. Significant adjusted p-values (p-value<0.05) and RGA classes are written in bold. For domains the abbreviations mean the following: NBS = Nucleotide Binding Site; CNL = coiled-coil, nucleotide-binding site and leucine-rich repeat protein; TNL= TIR, nucleotide-binding site and leucine-rich repeat protein; CN= coiled-coil and nucleotide-binding site; TN=TIR and nucleotide-binding site; NL= nucleotide-binding site and leucine-rich repeat protein; TX= TIR-unknown site/domain; RLP= Receptor like protein; RLK= Receptor like kinase; TM-CC= Transmembrane and coiled coil.

***Rpv12* candidate gene**	**Put. function**	**Encoded Domains**	**adj. p-value**	**cluster**
Gf9918_14g22546	GTD-binding domain similar to MYOB1: Myosin-binding Protein 1	**TM-CC**	0.9745	-
Gf9918_14g22556	Protein of unknown function	PGG, TM	**0.007672**	1
Gf9918_14g22558	Similar to ACD6: Protein ACCELERATED CELL DEATH 6 (*Arabidopsis thaliana*)	Ankyrin repeat-containing domain, PGG, TM	**1.278e-08**	1
Gf9918_14g22562	Similar to ACD6	Ankyrin repeat-containing domain, PGG, TM	**7.634e-09**	1
Gf9918_14g22568	Similar to ACD6	Ankyrin repeat-containing domain, PGG, TM	**5.891e-08**	1
Gf9918_14g22588	CheQ-like His Kinase A (phosphor-acceptor) domain similar to EIN4	**TM-CC**	0.8976	-
Gf9918_14g22614	Similar to At4g27220: Probable disease resistance protein At4g27220 (*A. thaliana*)	TM, LRR (**RLP RGA**)	0.854	2
Gf9918_14g22616	Similar to At4g10780: Putative disease resistance protein At4g10780 (*A. thaliana*)	NB-ARC (**CN RGA**)	0.4264	2
Gf9918_14g22619	Similar to At4g27220	NBS, LRR (**NL RGA**), Hydrolase	NA	2
Gf9918_14g22623	Protein of unknown function	**TM-CC**	0.7058	2

Gf9918_14g22624	Similar to At4g27220	NB-ARC, LRR (**CNL RGA**), Hydrolase	0.7973	2
Gf9918_14g22626	Similar to At4g27220	NB-ARC, LRR (**CNL RGA**), Hydrolase	0.7808	2
Gf9918_14g22636	Protein of unknown function	**TM-CC**	0.419	-

Figure Legends

Figure 1 Pedigree of the heterozygous *Rpv12*-carrier GF.99-03 chosen for phased genome sequence assembly. This pedigree shows the confirmed relationship between the sequenced genotypes 65-153-18, Gf.43-21 and Gf.99-03. Genotypes in green were available while genotypes in grey were not available. Identified resistance-associated loci are given below the genotype.

Figure 2 *Plasmopara viticola* infection of *Rpv12* carrying and non *Rpv12* carrying genotypes. **(a-d)** Leaf disc assay of different genotypes 5 dpi after artificial inoculation with *P. viticola*. **(a)** Leaf discs from genotype `Italia´ covered by sporangiophores. **(b)** 65-153-18 (*Rpv12*) shows strongly reduced sporangiophore development and HR visible as necrotic spots. **(c)** Genotype Gf.43-21 (*Rpv1*, *Rpv3.1*) exhibits no sporangiophores but HR visible as necrotic spots. **(d)** Gf.99-03 (*Rpv12*) has very little sporangiophores and HR is visible as necrotic spots. **(e & f)** Nitrotetrazolium blue chloride staining for detection of ROS. *P. viticola* infected leaf discs of `Italia´ (susceptible) and Gf.99-03 (*Rpv12*) were incubated for 24 h and were stained with Nitrotetrazolium blue chloride. **(e)** The susceptible genotype `Italia´ shows no ROS. **(f)** Genotype Gf.99-03 (*Rpv12*) produced ROS that led to visible brown lesions. **(g & h)** Comparison of haustoria formation. Fluorescence microscopic images of *P. viticola* infestation on the genotypes `Italia´ and Gf.99-03 at 72 hpi. Mycelium was stained with alkaline aniline blue. Arrows are pointing on haustoria. **(g)** `Italia´ (susceptible) shows less haustoria and a widely spread mycelium. **(h)** Gf.99-03 (*Rpv12*) shows many small haustoria with reduced width. **(i)** Comparison of mycelial growth. Mycelial density was evaluated on leaves of Gf.99-03 and its parental genotypes Gf.43-21 and 65-153-18 including `Italia´ as susceptible control. The results are visualized as box-plots with standard deviation (Welch's t-test). Relevant resistance loci are given above the error bar. The median is marked as a thick black line within the box while the whiskers subscript the minimum and maximum of all data. The data of the variety `Italia´ as reference was already published (Müllner and Zyprian, 2022).

Figure 3 Pseudochromosome lengths and BUSCO analysis of the Gf.99-03 haplotypes and of PN40024. (a) Pseudochromosome lengths of both Gf.99-03 haplotypes and of PN40024 12X.v2. The lengths and values are shown in magenta for Gf9921, in turquoise for Gf9918, and in grey for PN40024. chrUn, chr00 and chrNh represent artificial contig collections. The reference sequence PN40024 12X.v2 contains a pseudochromosome chr00 for sequences not assigned to chromosomes, this collection has been created as well and is referred to here as "chrUn" for Gf9921 and Gf9918. chrNh collects contigs that can not be assigned to chromosomal locations based on similarity to established *Vitis* genome sequences. The scale on the y-axis and the lengths of chrNh are interrupted at 48 Mbp. **(b)** Number of plant core genes for Gf.99-03 and PN40024.2,326 plant core genes were searched with BUSCO. Note that the bar graph is truncated at the left and focuses on only duplicated, fragmented and missing BUSCOs. The track Gf.99-03 combines the summed up results of the single haplotypes.

Figure 4 Bob plot of the Gf.99-03 contigs. Contigs are displayed as circles. Magenta circles represent Gf9921 contigs and turquoise circles Gf9918 contigs. The circle size depends on the total number of k-mers assigned to a given contig. The axes indicate the number of k-mers found in the hap-mer sets (y-axis Gf.43-21 hap-mers, x-axis 65-153-18 hap-mers; the analysis tool Mercury refers to haplotype-specific k-mers as hap-mers).

Figure 5 PCA plot and differentially expressed genes of the *Plasmopara viticola* infection time course experiment. (a) Principal Component Analysis (PCA) plot of the *P. viticola* infection experiment. Untreated samples (blue colored dots) are control samples that were not inoculated with *P. viticola* zoospores, but they were taken at the given time point. Treated samples (red colored dots) represent the inoculated samples. **(b) Differentially expressed genes across time points.** The upset plot (Lex *et al.*, 2014, Conway *et al.*, 2017) shows the intersection of DEGs between the time points for the up-regulated DEGs (upper part) and down-regulated DEGs (lower part) per haplotype. Data of Gf9921 are colored in magenta and data of Gf9918 in turquoise. The most left bar plot displays the total amount of up-regulated DEGs per time point and per haplotype and the most right bar plot the total amount of down-regulated DEGs per time point and per haplotype. Intersections are displayed as connecting lines with dots. The values on the y-axis represent the amount of shared DEGs.

Figure 6 Synteny plot of the *Rpv12* locus. Connecting lines are based on blastp protein hits with a query coverage and identity higher than 85 % and an E-value of 0.0001. To facilitate comparison, the *Rpv12* locus of PN40024 12X.v2 is displayed twice. **(a)** Synteny plot of the *Rpv12* locus between marker UDV-014 and UDV-370. The two regions with candidate genes were underlain with a blue square. **(b)** Zoom in of the northern gene candidate region comprising functional candidates (marker GF14-21 to GF14-31). The gene mate *Gf9921_14g22214* for the *ACD6* DEGs is displayed with its name. **(c)** Zoom in of the southern gene candidate regions with positional RGA candidates (marker UDV-345 to UDV-360).

Figure 7 Expression over time of the differentially expressed *Rpv12* candidates. The error bars represent the standard deviation. **(a)** Expression course of *Gf9918_14g22556*. **(b)** Expression course of *Gf9918_14g22558* and its gene mate *Gf9921_14g22214*. **(c)** Expression course of *Gf9918_14g22562* and its gene mate *Gf9921_14g22214*. **(d)** Expression course of *Gf9918_14g22568* and its gene mate *Gf9921_14g22214*.

Figure 1

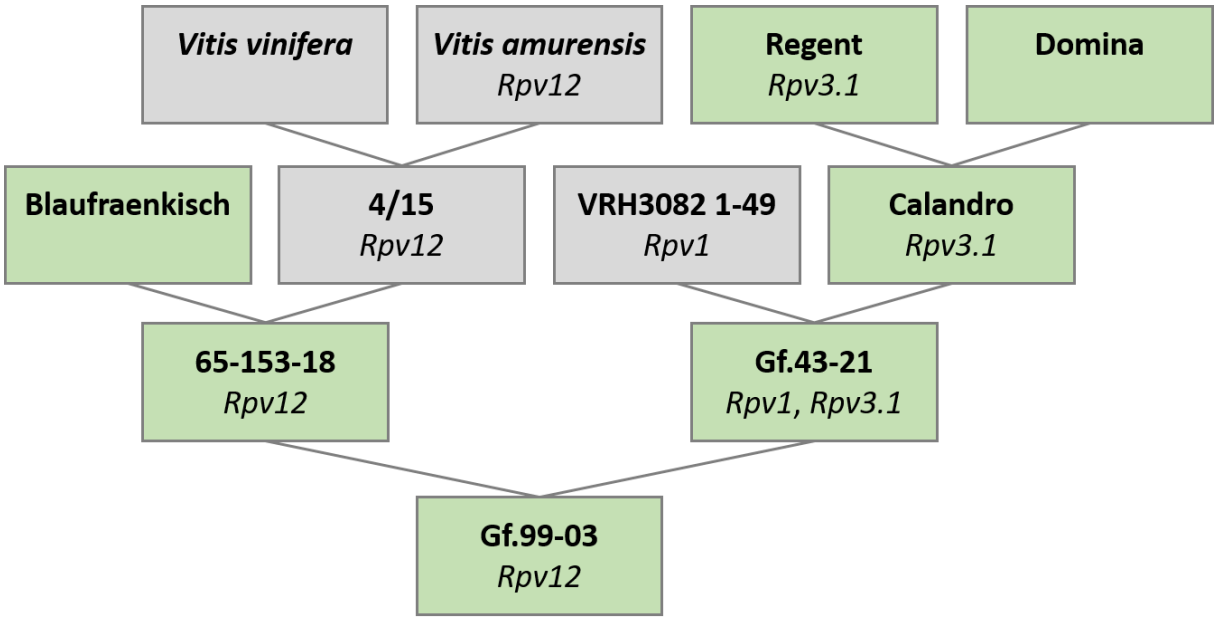

Vitis vinifera
Vitis amurensis
Rpv12
Regent
Rpv3.1
Domina
Blaufraenkisch
4/15
Rpv12
VRH3082 1-49
Rpv1
Calandro
Rpv3.1
65-153-18
Rpv12
Gf.43-21
Rpv1, Rpv3.1
Gf.99-03
Rpv12

Figure 2

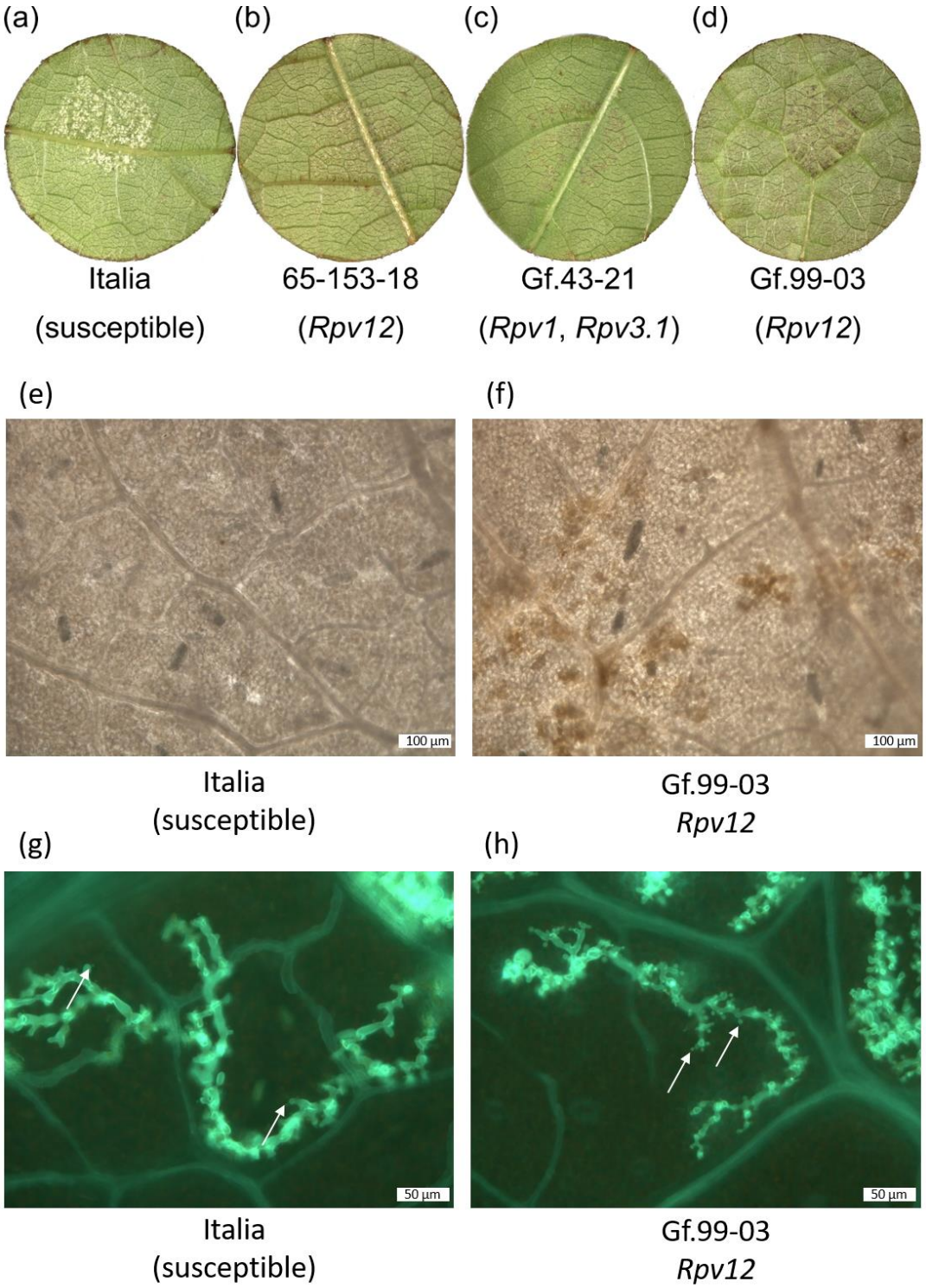

(a)
(b)
(c)
(d)
Italia
(susceptible)
65-153-18
(*Rpv12*)
Gf.43-21
(*Rpv1*, *Rpv3.1*)
Gf.99-03
(*Rpv12*)
(e)
(f)
100 µm
100 µm
Italia
(susceptible)
Gf.99-03
Rpv12
(g)
(h)
50 µm
50 µm
Italia
(susceptible)
Gf.99-03
Rpv12

Figure 2 contd

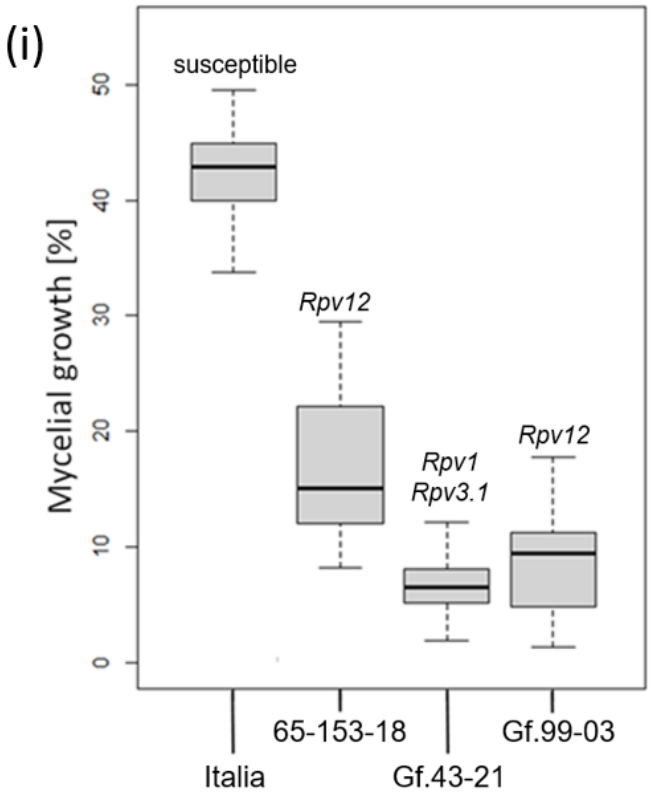


Figure 3

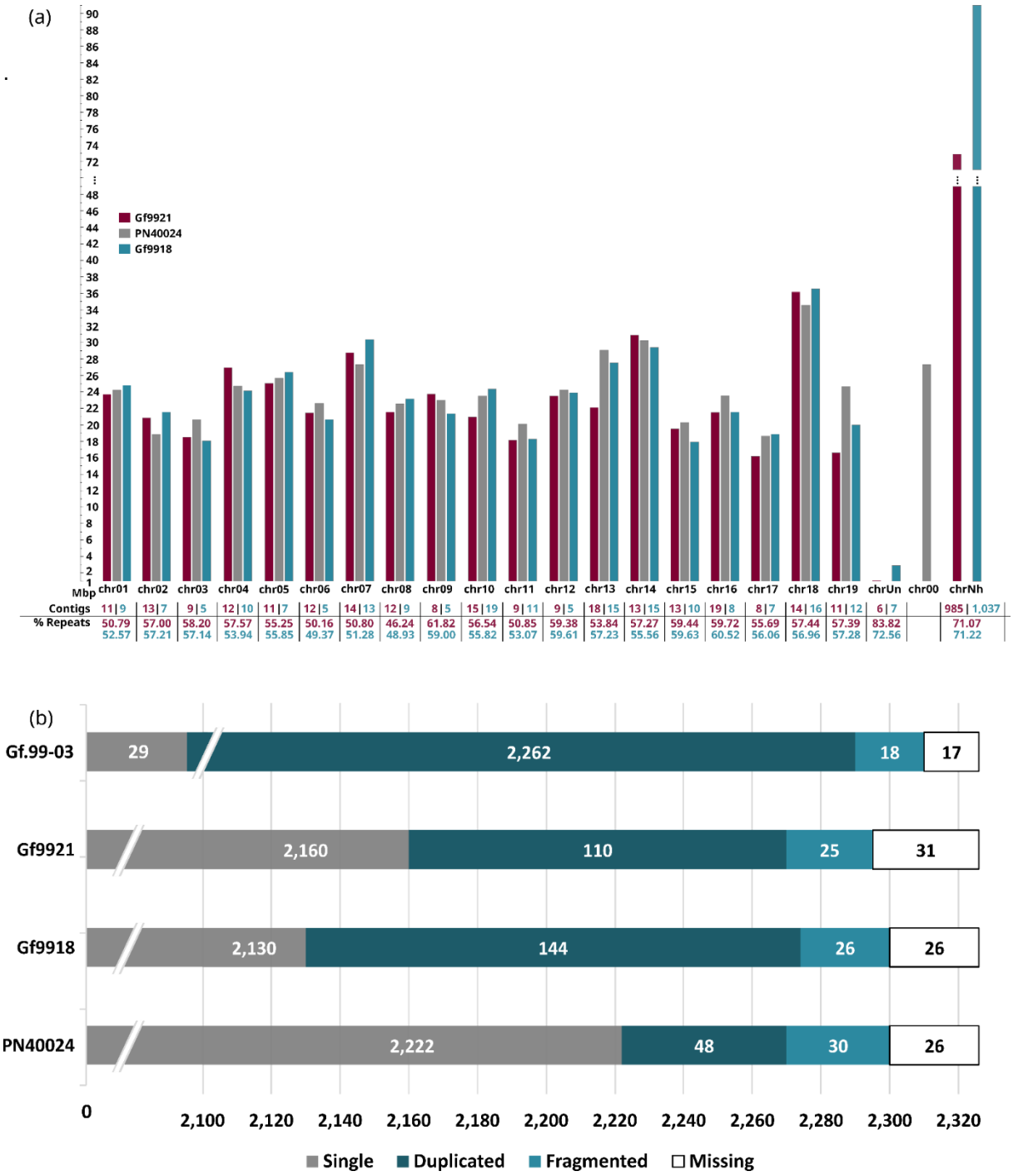

(a)
Gf9921
PN40024
Gf9918
Mbp
chr01 chr02 chr03 chr04 chr05 chr06 chr07 chr08 chr09 chr10 chr11 chr12 chr13 chr14 chr15 chr16 chr17 chr18 chr19 chrUn chr00 chrNh
Contigs
11|9 13|7 9|5 12|10 11|7 12|5 14|13 12|9 8|5 15|19 9|11 9|5 18|15 13|15 13|10 19|8 8|7 14|16 11|12 6|7 985|1,037
% Repeats
50.79 57.00 58.20 57.57 55.25 50.16 50.80 46.24 61.82 56.54 50.85 59.38 53.84 57.27 59.44 59.72 55.69 57.44 57.39 83.82 71.07
52.57 57.21 57.14 53.94 55.85 49.37 51.28 48.93 59.00 55.82 53.07 59.61 57.23 55.56 59.63 60.52 56.06 56.96 57.28 72.56 71.22
(b)
Gf.99-03
29
2,262
18
17
Gf9921
2,160
110
25
31
Gf9918
2,130
144
26
26
PN40024
2,222
48
30
26
0
2,100 2,120 2,140 2,160 2,180 2,200 2,220 2,240 2,260 2,280 2,300 2,320
Single
Duplicated
Fragmented
Missing

Figure 4

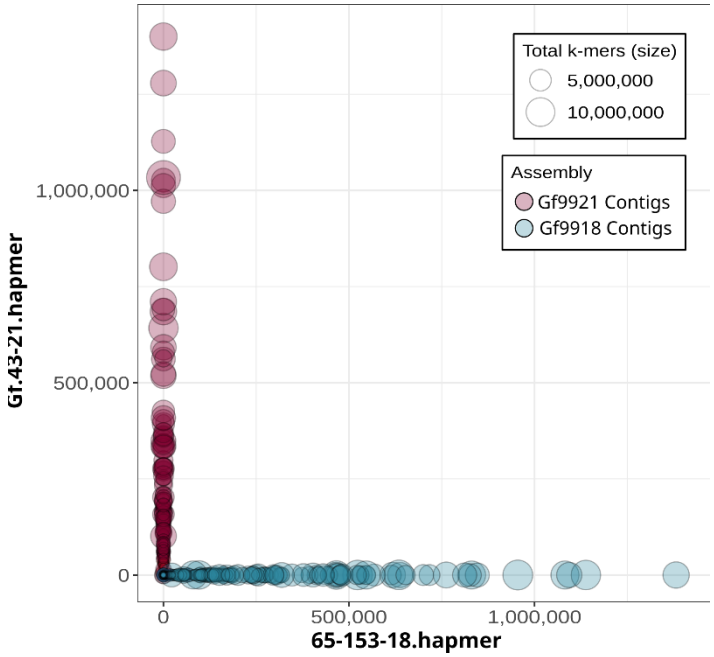

Total k-mers (size)
5,000,000
10,000,000
Assembly
Gf9921 Contigs
Gf9918 Contigs
1,000,000
500,000
0
Gt.43-21.hapmer
0
500,000
1,000,000
65-153-18.hapmer

Figure 5

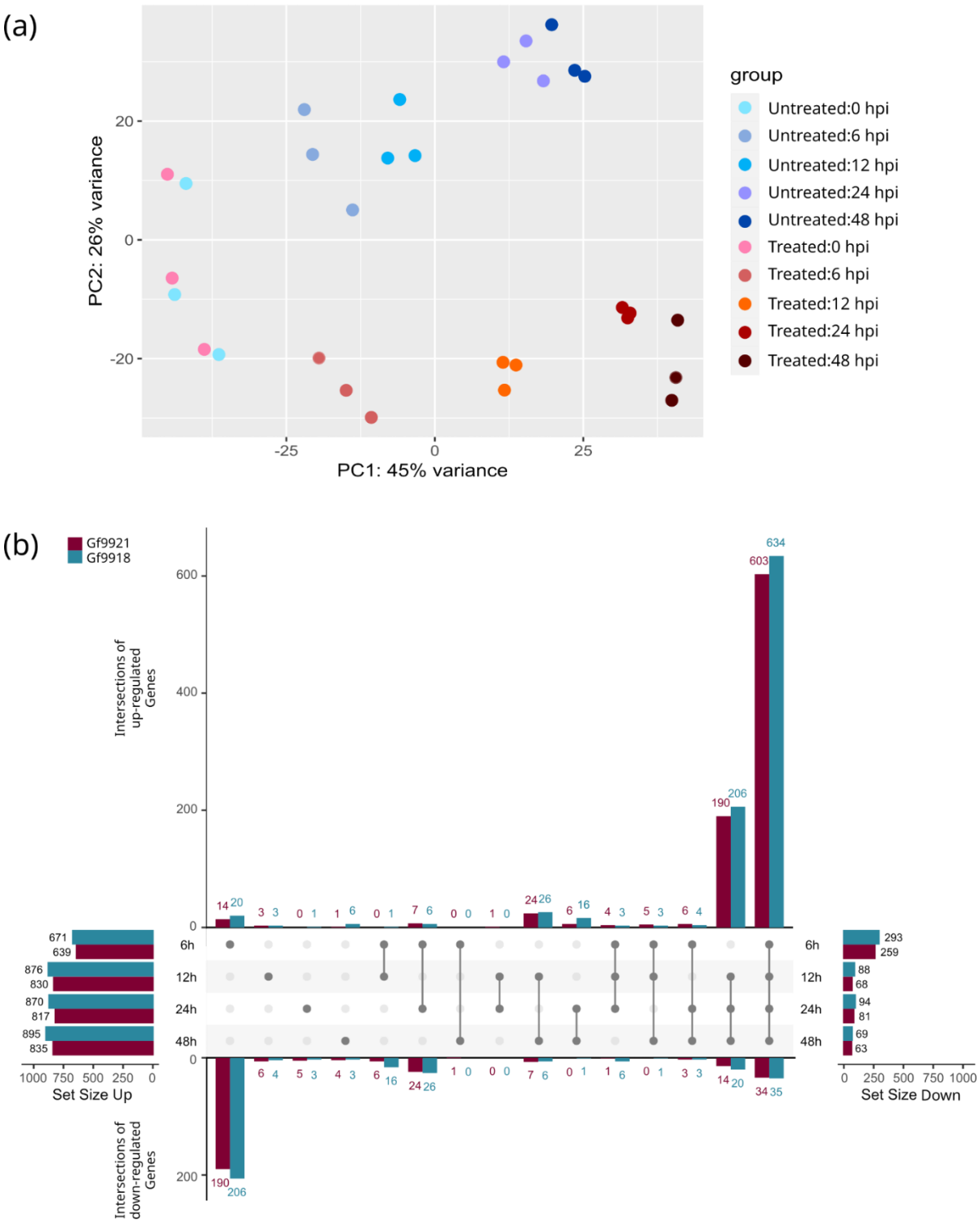

(a)
group
Untreated:0 hpi
Untreated:6 hpi
Untreated:12 hpi
Untreated:24 hpi
Untreated:48 hpi
Treated:0 hpi
Treated:6 hpi
Treated:12 hpi
Treated:24 hpi
Treated:48 hpi
PC2: 26% variance
PC1: 45% variance
(b)
Gf9921
Gf9918
Intersections of up-regulated Genes
Intersections of down-regulated Genes
Set Size Up
Set Size Down
6h
12h
24h
48h

Figure 6

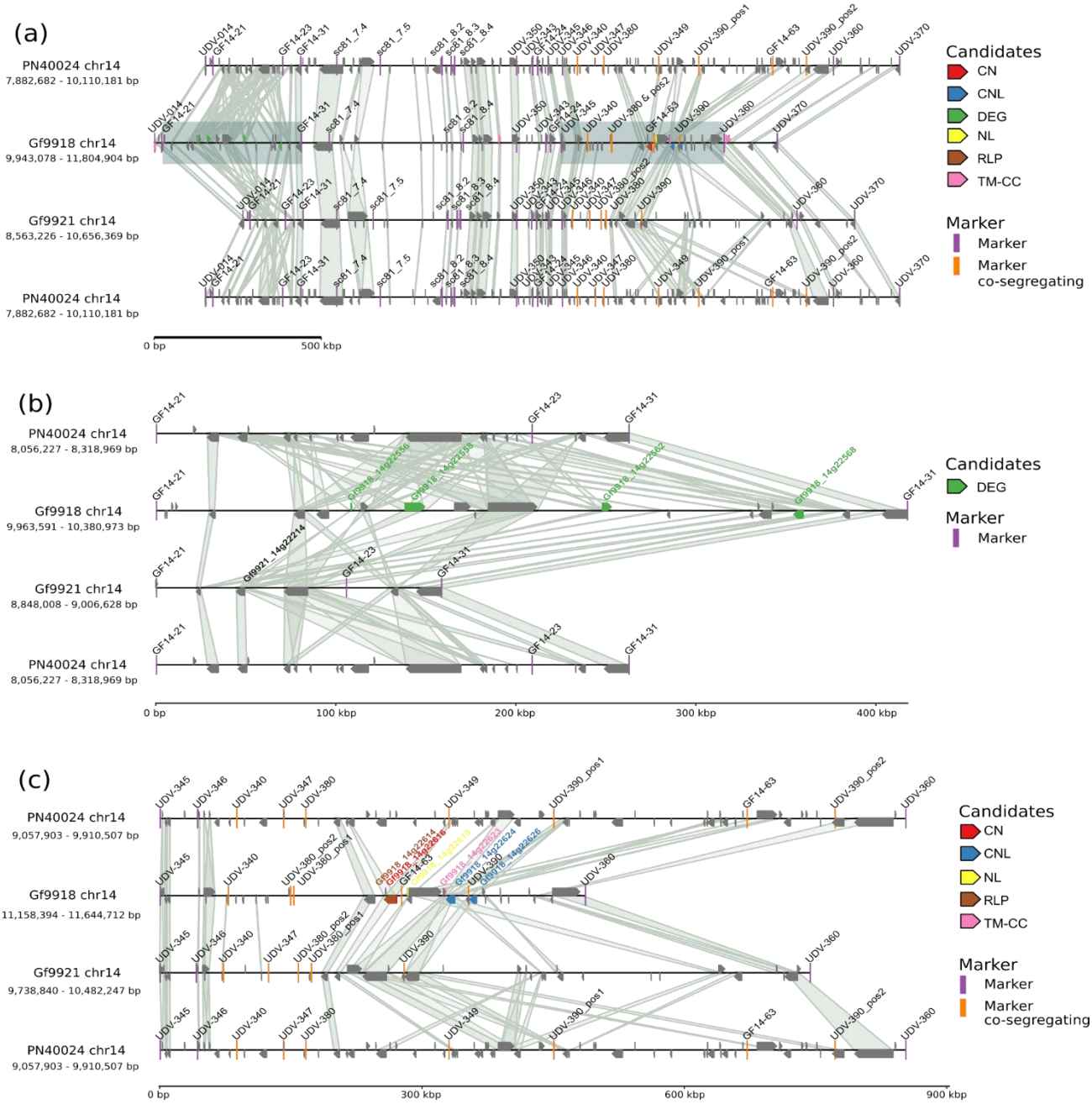
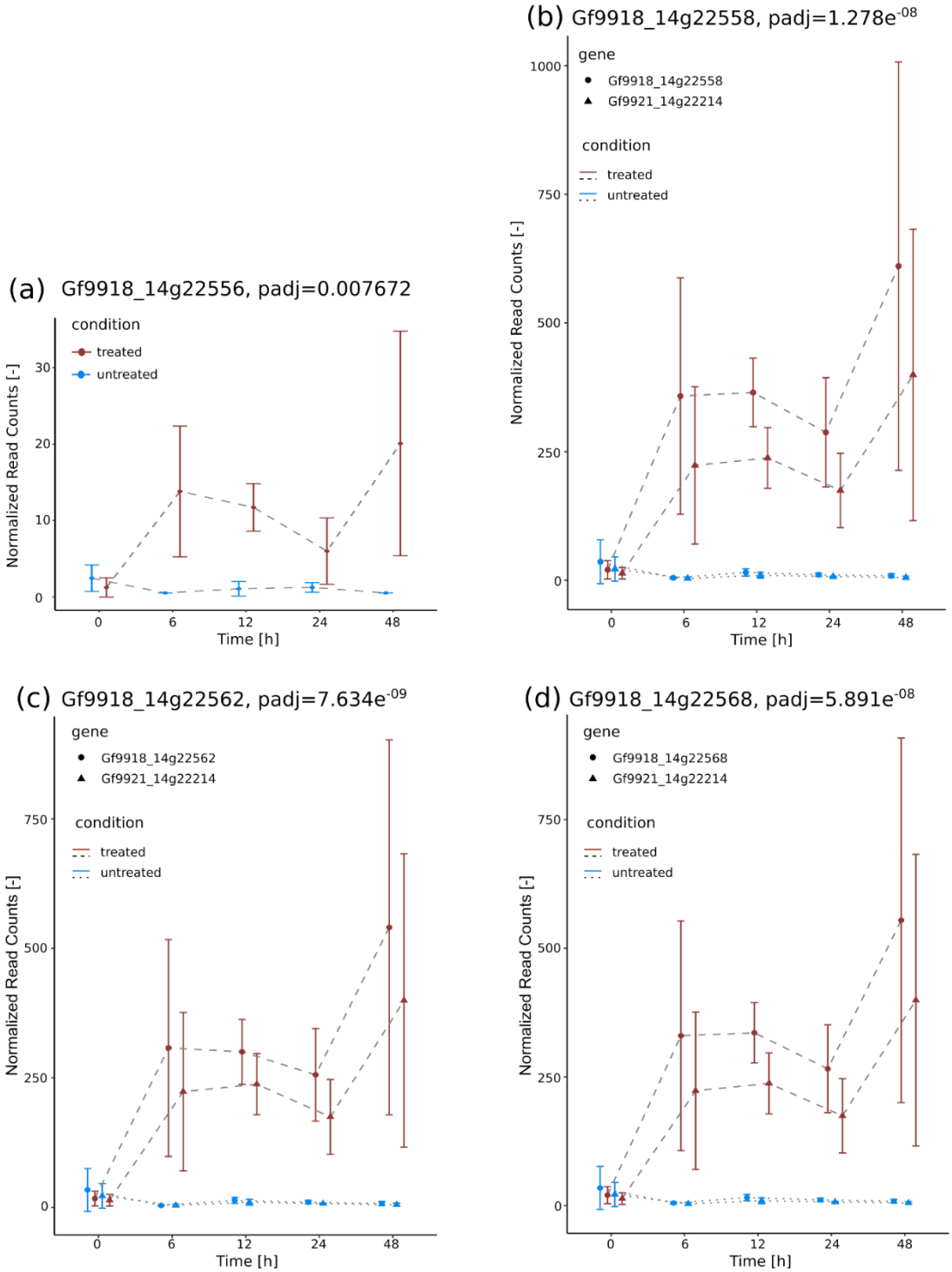

(a)
PN40024 chr14
7,882,682 - 10,110,181 bp
Gf9918 chr14
9,943,078 - 11,804,904 bp
Gf9921 chr14
8,563,226 - 10,656,369 bp
PN40024 chr14
7,882,682 - 10,110,181 bp
0 bp
500 kbp
Candidates
CN
CNL
DEG
NL
RLP
TM-CC
Marker
Marker
Marker co-segregating
(b)
PN40024 chr14
8,056,227 - 8,318,969 bp
Gf9918 chr14
9,963,591 - 10,380,973 bp
Gf9921 chr14
8,848,008 - 9,006,628 bp
PN40024 chr14
8,056,227 - 8,318,969 bp
0 bp
100 kbp
200 kbp
300 kbp
400 kbp
Candidates
DEG
Marker
Marker
(c)
PN40024 chr14
9,057,903 - 9,910,507 bp
Gf9918 chr14
11,158,394 - 11,644,712 bp
Gf9921 chr14
9,738,840 - 10,482,247 bp
PN40024 chr14
9,057,903 - 9,910,507 bp
0 bp
300 kbp
600 kbp
900 kbp
Candidates
CN
CNL
NL
RLP
TM-CC
Marker
Marker
Marker co-segregating

Figure 7


(a) Gf9918_14g22556, padj=0.007672
condition
treated
untreated
Normalized Read Counts [-]
Time [h]
(b) Gf9918_14g22558, padj=1.278e-08
gene
Gf9918_14g22558
Gf9921_14g22214
condition
treated
untreated
Normalized Read Counts [-]
Time [h]
(c) Gf9918_14g22562, padj=7.634e-09
gene
Gf9918_14g22562
Gf9921_14g22214
condition
treated
untreated
Normalized Read Counts [-]
Time [h]
(d) Gf9918_14g22568, padj=5.891e-08
gene
Gf9918_14g22568
Gf9921_14g22214
condition
treated
untreated
Normalized Read Counts [-]
Time [h]