Article

Current Biology

Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*

Highlights

- Identification of the first oomycete mating-type locus using a GWAS approach
- This 570-kb repeat-rich non-recombining region showed two highly divergent alleles
- One mating type was heterozygous and the other homozygous, as in XY sex chromosomes

Authors

Yann Dussert, Ludovic Legrand, Isabelle D. Mazet, ..., Silvia Laura Toffolatti, Tatiana Giraud, François Delmotte

Correspondence

dussert.yann@gmail.com (Y.D.), francois.delmotte@inrae.fr (F.D.)

In Brief

Mating-type genes have been widely studied in plants and fungi but have yet to be identified in oomycetes. Dussert et al. utilize a genome-wide association approach to identify the mating-type locus of the oomycete plant pathogen *Plasmopara viticola*. The mating-type locus region displays similarity to heteromorphic sex chromosome systems.

Dussert et al., 2020, Current Biology *30*, 1–11
October 19, 2020
https://doi.org/10.1016/j.cub.2020.07.057

CellPress

Please cite this article in press as: Dussert et al., Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.07.057

Current Biology

CellPress

Article

Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*

Yann Dussert,[1,7,8,*] Ludovic Legrand,[2] Isabelle D. Mazet,[1] Carole Couture,[1] Marie-Christine Piron,[3] Rémy-Félix Serre,[4] Olivier Bouchez,[4] Pere Mestre,[3] Silvia Laura Toffolatti,[5] Tatiana Giraud,[6] and François Delmotte[1,*]

[1]SAVE, INRAE, Bordeaux Sciences Agro, Université de Bordeaux, F-33140 Villenave d'Ornon, France
[2]LIPM, INRAE, Université de Toulouse, CNRS, Castanet-Tolosan, France
[3]SVQV, INRAE, Université de Strasbourg, F-68000 Colmar, France
[4]INRAE, US 1426 GeT-PlaGe, Genotoul, Castanet-Tolosan, France
[5]Dipartimento di Scienze Agrarie e Ambientali, Università degli Studi di Milano, Milano, Italy
[6]Ecologie Systematique et Evolution, CNRS, AgroParisTech, Universite Paris-Saclay, 91400 Orsay, France
[7]Present address: School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
[8]Lead Contact
*Correspondence: dussert.yann@gmail.com (Y.D.), francois.delmotte@inrae.fr (F.D.)
https://doi.org/10.1016/j.cub.2020.07.057

SUMMARY

Mating types are self-incompatibility systems that promote outcrossing in plants, fungi, and oomycetes. Mating-type genes have been widely studied in plants and fungi but have yet to be identified in oomycetes, eukaryotic organisms closely related to brown algae that cause many destructive animal and plant diseases. We identified the mating-type locus of *Plasmopara viticola*, the oomycete responsible for grapevine downy mildew, one of the most damaging grapevine diseases worldwide. Using a genome-wide association approach, we identified a 570-kb repeat-rich non-recombining region controlling mating types, with two highly divergent alleles. We showed that one mating type was homozygous, whereas the other was heterozygous at this locus. The mating-type locus encompassed 40 genes, including one encoding a putative hormone receptor. Functional studies will, however, be required to validate the function of these genes and find the actual determinants of mating type. Our findings have fundamental implications for our understanding of the evolution of mating types, as they reveal a unique determinism involving an asymmetry of heterozygosity, as in sex chromosomes and unlike other mating-type systems. This identification of the mating-type locus in such an economically important crop pathogen also has applied implications, as outcrossing facilitates rapid evolution and resistance to harsh environmental conditions.

INTRODUCTION

Mating systems control the degree of outcrossing in natural populations, thus affecting adaptability [1–5]. Indeed, outcrossing promotes gene flow and, therefore, the rapid spread of beneficial alleles, the generation of new allelic combinations, and the purging of deleterious alleles. Various genetic systems have evolved across the tree of life, which enforce outcrossing, such as separate sexes or mating types [6]. The sex of many plants and animals is determined by sex chromosomes [6]. Mating types, preventing mating between individuals carrying the same alleles, have evolved independently in many lineages of the tree of life, including fungi, ciliates, green algae, and oomycetes [7–9]. Plants possess a molecular system called self-incompatibility, which enforces outcrossing and is analogous to mating type [10]. Given the fundamental importance of mating types in life cycles and evolution, their molecular determinism has been extensively studied, particularly in plants, fungi, and ciliates. Mating types are controlled by a number of different mechanisms, even within these groups [7, 8, 10].

Mating type systems enforcing outcrossing at the diploid stage, involving only two mating types, have been described in oomycetes on the basis of cross incompatibilities [11]. However, no genomic sequence for a mating-type locus has yet been identified in this lineage. Oomycetes are diploid eukaryotic organisms closely related to diatoms and brown algae [12]. This group includes a number of animal and plant pathogens causing significant environmental and economic damage. Oomycetes are responsible for damaging diseases, such as saprolegniosis, a lethal disease affecting wild and farmed fish [13]; sudden oak death; downy mildew in several crops (e.g., *Bremia lactucae*, responsible for lettuce downy mildew, and *Plasmopara halstedii*, responsible for sunflower downy mildew), and the infamous potato blight caused by *Phytophthora infestans*, responsible for the Irish potato famine in the 1840s.

Mating types are of the utmost importance in these organisms, as they control gamete compatibility, and sexual reproduction generates thick-walled spores called oospores that can resist harsh conditions and survive for several years. Outcrossing also produces recombinant genotypes, which can mediate

Please cite this article in press as: Dussert et al., Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.07.057

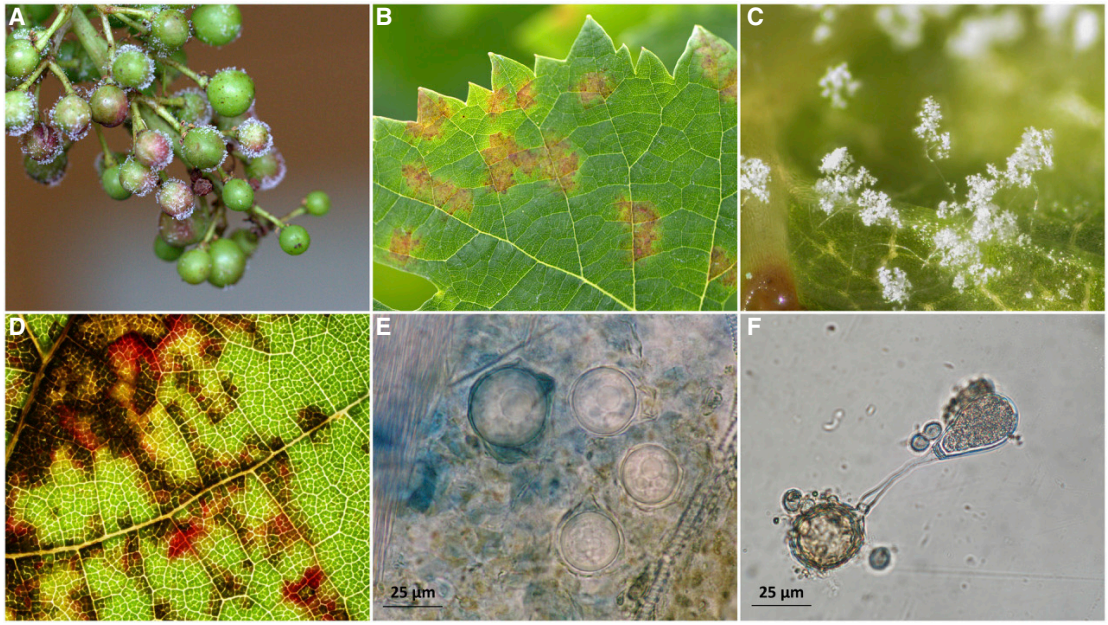

Figure 1. Symptoms and Asexual and Sexual Structures of the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*
(A) Diseased young grape berries (asexual reproduction).
(B) Typical pale yellow lesions (oil spots) on a diseased grape leaf at spring (asexual reproduction).
(C) Sporangiophores and sporangia (asexual structures) emerging through the stomata of diseased grape organs.
(D) Typical patchwork of angular yellow to brown red spots on the upper surface of a diseased leaf at fall, when sexual reproduction occurs.
(E) Oospores formed on leaf disks after sexual reproduction between two individuals of opposite mating types.
(F) Oospore (left) germinating to produce a macrosporangium (right) containing 4–6 asexual zoospores.

faster adaptation to control measures. For example, invasive populations of *Ph. infestans* were not able to reproduce sexually for a long period, due to the lack of one mating type, but the recent introduction of the missing mating type has resulted in higher frequencies of recombinant genotypes in some areas [14] and the emergence of an aggressive lineage [15]. In oomycetes, the mating-type locus has been located on genetic linkage maps in *Phytophthora* species [16–19] and *Bremia lactucae* [20]. A recent study [21] identified several contigs associated with mating types on the *Ph. capsici* reference genome but could not precisely locate the locus controlling the phenotype. The only known mating-type factors are purified hormones in *Phytophthora spp* [22–24]. However, no genes determining mating type have yet been identified in oomycetes.

Plasmopara viticola, a pathogen causing grapevine downy mildew (Figure 1), one of the most devastating grapevine diseases worldwide, is an oomycete. *Plasmopara viticola* was introduced into European vineyards from North America in the 1870s [25] and rapidly spread, invading vineyards all over the world. The life cycle of *Pl. viticola* includes an asexual multiplication phase during the spring and summer (Figures 1A–1C) and a sexual reproduction event in the fall (Figure 1D), generating the thick-walled sexual spores (oospores; Figures 1E and 1F) required for overwintering [26]. In *Pl. viticola*, only two mating types have been observed, and mating can occur only between diploid individuals of different mating types [27]. This results in high rates of outcrossing, a key element explaining rapid adaptation to fungicides [28, 29] and to resistant cultivars [30, 31] in this species. We identified the first oomycete mating-type locus sequence, with an approach combining phenotypic analysis (mating types determined by crosses) with genome-wide single-nucleotide polymorphisms (SNPs) in *Pl. viticola*.

RESULTS

We determined the mating type of 54 diploid individuals of *Pl. viticola* collected across Europe (Table S1), in experimental crosses against six testers, three for each of the two mating types, arbitrarily called P1 and P2. The P1 mating type was inferred for an individual if it produced oospores when inoculated on grapevine leaves with P2 testers, but not with P1 testers, and conversely for P2. We identified 26 individuals of the P1 mating type and 28 of the P2 mating type (Table S2). We sequenced the genomes of these 54 diploid individuals with short-read technology and mapped the reads onto a recent high-quality reference sequence obtained by long-read sequencing with high coverage of the INRA-PV221 individual [32]. Our mating-type phenotyping revealed that this reference INRA-PV221 individual had the P2 mating type. The reference genome covers around 81% of the estimated 115-Mb genome size of *Pl. viticola* and exhibits a high genome-wide heterozygosity of 0.8%. We retained 2.011 million SNPs after filtering. No genetic subdivision associated with mating types was detected in population structure analyses based on principal-component analysis or clustering analysis applied to a dataset filtered for SNPs in close linkage disequilibrium (LD) (Figure S1). Conditions were therefore favorable for genetic-phenotype association studies.

Please cite this article in press as: Dussert et al., Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.07.057

CellPress

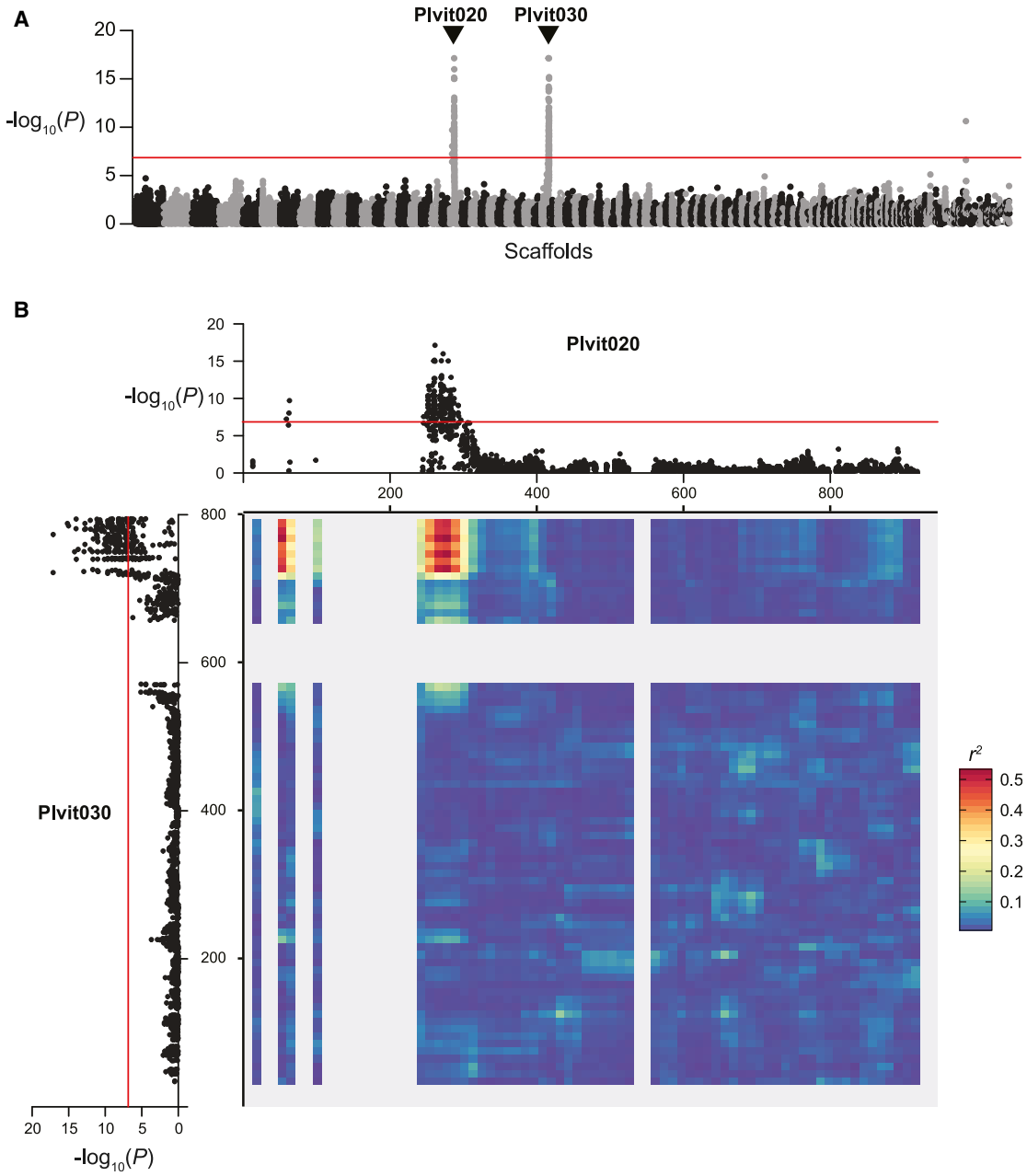

Figure 2. Genome-wide Association Analysis for Identifying Mating-type Regions in *Plasmopara viticola*

(A) Manhattan plot of the negative $\log_{10}$-transformed association p values between mating type and SNPs along the *Pl. viticola* genome. Alternating black and gray blocks of dots mark the limits between scaffolds. The two scaffolds with a significant association signal (Plvit020 and Plvit030) are indicated with arrows. Only 5% of the SNPs with a $p > 0.1$ are represented, to keep the number of plotted points manageable.

(B) Manhattan plots of the negative $\log_{10}$-transformed association p values between mating type and SNPs for Plvit020 and Plvit030. The significance threshold for the association analysis, computed with 10,000 permutations, is represented as a red line in both panels. Linkage disequilibrium between the two scaffolds is represented as a heatmap (80 bins), with redder colors representing higher r^2 values (higher linkage disequilibrium) and therefore lower recombination rates. See also Figures S1–S4.

Using a genome-wide association approach, we identified two genomic regions with significant signals of association with the mating-type phenotype, located at the edges of the scaffolds Plvit020 and Plvit030 (Figure 2). SNPs at these two scaffold extremities were in very strong LD (Figure 2B), indicating that they were located in the same genomic region and that this region probably lacked recombination. The rate of LD decay was much lower in this region than in the rest of the genome (Figure S2), providing further support for the hypothesis of a lack of recombination. The incomplete assembly of this locus in the reference genome was probably due to its high repeat content. Indeed, we observed two large regions

Please cite this article in press as: Dussert et al., Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.07.057

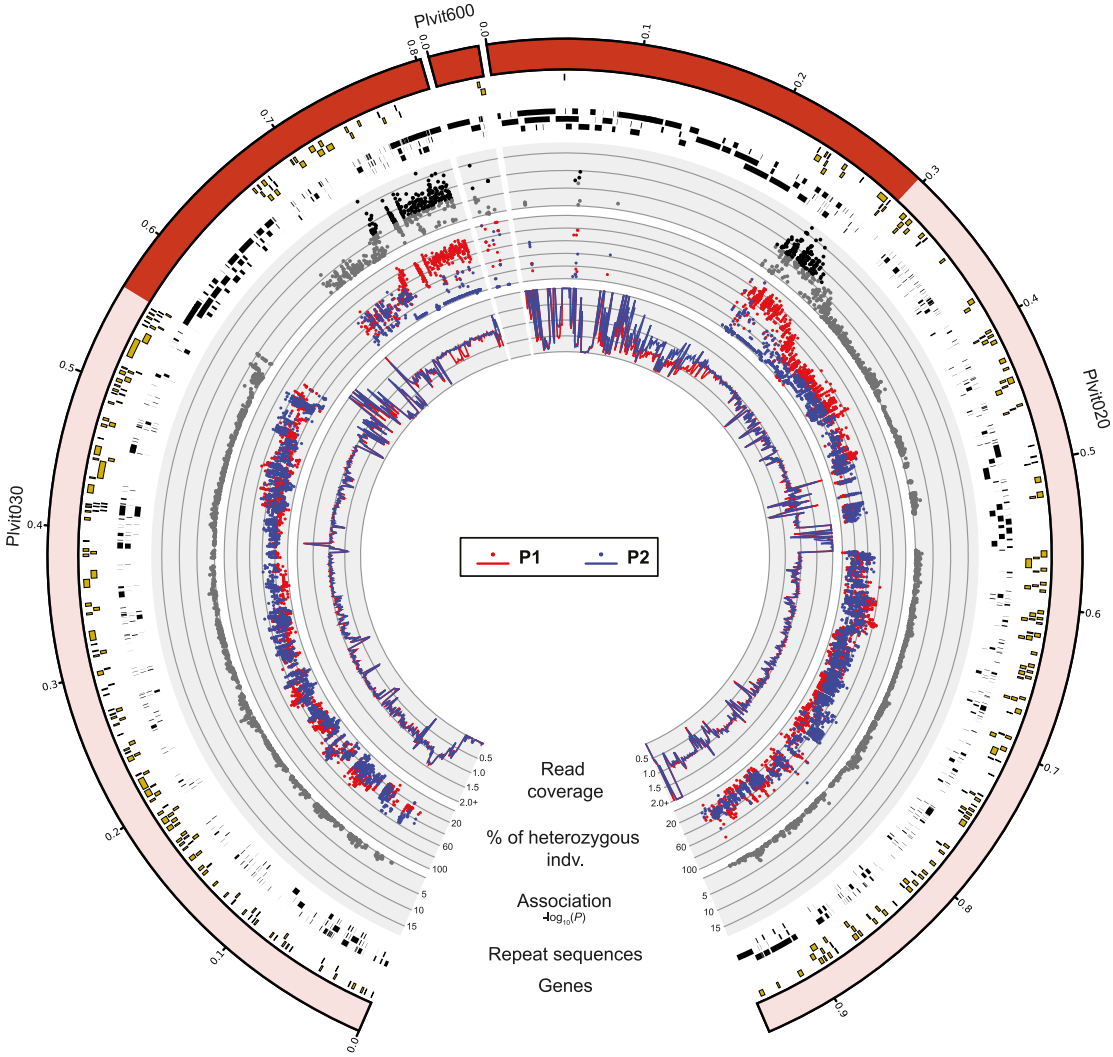

Figure 3. Genomic Regions Associated with Mating Type in *Plasmopara viticola*

The regions significantly associated with the mating-type phenotype are shown in dark red within the scaffold ideograms (outer track). Annotated genes are represented as yellow rectangles and repeated sequences as black rectangles (see Figure S3 for details about the repeat content in the mating-type locus region). Significant negative $\log_{10}$-transformed p values from the association analysis are represented by black dots, and other values are represented by gray dots. The percentage of heterozygous individuals for each SNP is shown for individuals of mating types P1 by red dots and P2 by blue dots. The mean normalized read coverage (for adjacent 1-kb windows) is shown by red lines for P1 individuals and blue lines for P2 individuals. The data were obtained with the reference-based SNP calling approach for the scaffolds Plvit020 and Plvit030 and the reference-free SNP calling approach for Plvit600 (STAR Methods). See also Figure S2 and Table S3.

composed exclusively of tandem repeat arrays covering 327 kb (Figure S3).

To identify potential missing sequences between the two scaffolds, we used a reference-free SNP calling method based on the analysis of the de Bruijn graph built from raw sequencing reads [33]. We detected short genomic sequences carrying SNPs with significant signals of association with the mating-type phenotype that were missing from the reference assembly (STAR Methods). We reassembled the PacBio reads corresponding to these short sequences and obtained a new contig, Plvit600 (32,220 bp), with SNPs significantly associated with the mating-type phenotype in strong linkage disequilibrium with SNPs located at the edges of Plvit020 and Plvit030 and therefore probably located between these two scaffolds (Figure S4). Plvit600 also included a large region entirely composed of tandem repeat arrays (Figure S3). We therefore concluded that the mating-type locus was at least 570 kb long (Figure 3). The mating-type locus might actually be larger given the difficulty of assembling and mapping reads in such a repeat-rich genomic region, but our use of a long-read reference assembly and the addition of a reference-free SNP calling method probably mitigated this issue.

P2 individuals were homozygous for the reference allele (designated MAT-a) at the mating-type locus, whereas P1 individuals were heterozygous, carrying the MAT-a reference allele and a second allele, called MAT-b (Figure 4). We found no

Please cite this article in press as: Dussert et al., Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.07.057

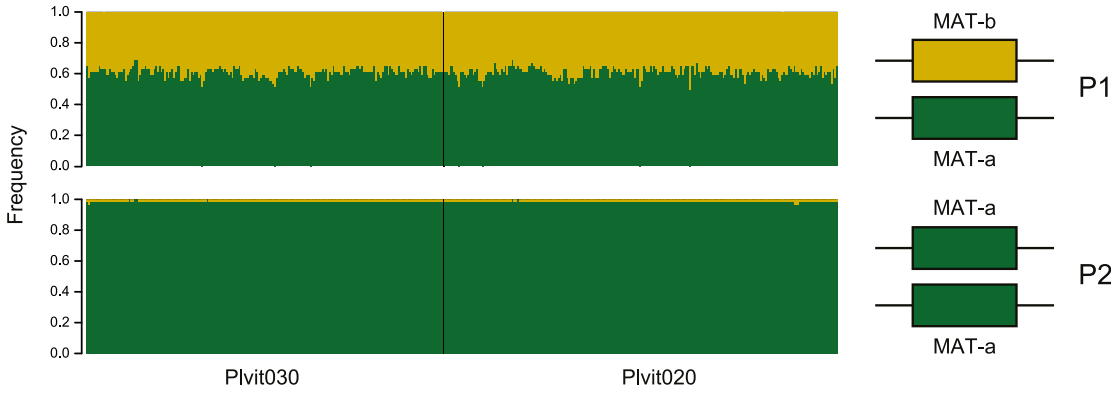

Figure 4. Mating Type Is Determined by a Two-Allele System in *Plasmopara viticola*

Left panel: allele frequencies at SNPs associated with the mating-type phenotype in P1 individuals (top) and P2 individuals (bottom). Only results for the reference-based SNP calling approach are shown. Right panel: proposed model for mating-type determination is shown, with homozygosity in P2 individuals and heterozygosity in P1 individuals. For both panels, the reference allele (i.e., the allele found in the reference genome) is shown in green and the alternative allele is shown in yellow.

homozygous individual for the alternative MAT-b allele. This suggested that the MAT-b allele is dominant. The two alleles were highly differentiated along the 570 kb, as shown by the large difference of heterozygosity levels (Figure 3), again consistent with a lack of recombination in this region. Unfortunately, with only short-read data and given the richness in repeats in the mating-type locus region, we could not assess the degree of rearrangements between the MAT-a and MAT-b alleles. It is, however, likely that structural variation exists in such a non-recombining and repeat-rich region.

The mating-type region included a total of 40 predicted coding sequences (Table S3), 26 of which had predicted functions and did not correspond to TEs (Table 1). Based on the predicted functions of these genes, the most promising candidate for involvement in mating-type determination was a gene encoding a transmembrane protein (PVIT_0010929.T1) with a sterol-sensing domain and lipid transport activity. This protein might act as a hormone receptor, and hormones have been identified as mating-type factors during initiation of sexual reproduction in *Phytophthora spp* [22–24]. Another two genes (PVIT_0010925.T1 and PVIT_0010927.T1) encoded proteins with lipid-binding domains that could potentially interact with mating hormones. We also found a putative mitogen-activated protein (MAP) kinase gene (PVIT_0027771.T1) in the identified mating-type region. MAP kinases play an important role in the mating pheromone response pathway in fungi [34], and genes encoding kinases are present at the mating-type loci of some fungi [35]. We also found genes encoding a protein with an HRCD (helicase and RNaseD C-terminal) (PVIT_0010933.T1) domain and a RecQ-mediated genome instability protein 2 (PVIT_0008617.T1). Proteins of this type can interact with DNA and are often associated with RecQ helicases and, thus, DNA repair. These proteins may also be involved in recombination suppression or gene regulation. This is potentially also the case for a Cdc48-like protein (PVIT_0027772.T1) encoded by a gene present at the mating-type locus. Cdc48 can interact with protein complexes on chromatin and is involved in gene regulation and cell cycle progression. Finally, we also found four genes (PVIT_0008616.T1, PVIT_0010916.T1, PVIT_0010917.T1, and PVIT_0010924.T1) encoding proteins with ubiquitin-like or ubiquitin-interacting domains, which could potentially be involved in addressing proteins to the proteasome and, thus, in protein degradation.

Using a reciprocal best-hit approach with whole proteomes, we located the regions orthologous to the *Pl. viticola* mating-type region in the available genomes of other oomycetes. The genes orthologous to putative *Pl. viticola* mating-type genes were found on three scaffolds for *Ph. infestans* and *Ph. capsici*: one for *Br. lactucae* and two for *Pl. halstedii* (Figure 5; Table S4). Orthologs of the genes from the Plvit600 contig were found in these four species close to the other candidate gene orthologs, confirming that this contig most probably belonged to the mating-type locus in *Pl. viticola*. A previous study in *Ph. capsici* based on genotyping by sequencing (GBS) (sequencing only a fraction of the genomes) identified SNPs associated with mating types in five different scaffolds [21]. We found that none of these scaffolds corresponded to the regions orthologous to the *Pl. viticola* mating-type locus. Association or functional studies are required to check whether the regions orthologous to the *Pl. viticola* mating-type locus also control mating type in other oomycete species.

DISCUSSION

We report here the first identification of an oomycete mating-type locus sequence, in the genome of *Pl. viticola*. The strong signal of association with mating type constitutes compelling evidence that the identified region is involved in controlling mating types. Functional studies deleting this region would, however, be required to provide definitive evidence for its role in mating-type determinism and identify the precise mating-type genes, but such functional studies are currently not possible in *Pl. viticola*. The *Pl. viticola* mating-type locus was found in a 570-kb non-recombining, repeat-rich region of the genome encompassing 40 genes. The region was heterozygous (MAT-a/MAT-b) in the P1 mating type and homozygous (MAT-a/MAT-a) in the P2 mating type, indicating dominance of the MAT-b allele. These findings are consistent with previous findings for marker segregation

Please cite this article in press as: Dussert et al., Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.07.057

Table 1. Putative Function of Genes in the Mating-type Locus of *Plasmopara viticola*

Protein ID	Scaffold	Length	Putative Function	Non-synonymous Mutations[a]
PVIT_0008600.T1	Plvit020	795	transferase activity, transferring acyl groups	
PVIT_0008605.T1	Plvit020	752	transferase activity, transferring acyl groups	
PVIT_0008608.T1	Plvit020	713	transferase activity, transferring acyl groups	I234T, Y239C
PVIT_0008616.T1	Plvit020	1,375	similar to peroxin-1, involved in protein targeting to peroxisome; contains a SEP and an UBX domains	P179S, D392G, E540D, R763K
PVIT_0008617.T1	Plvit020	160	similar to RecQ-mediated genome instability protein 2	
PVIT_0008618.T1	Plvit020	363	disulfide isomerase, involved in cell redox homeostasis and protein folding in the endoplasmic reticulum	
PVIT_0008619.T1	Plvit020	83	molybdopterin synthase sulfur carrier subunit	
PVIT_0008620.T1	Plvit020	303	similar to peroxin-13	V79G, L233S, E237D
PVIT_0010914.T1	Plvit030	419	glycosyltransferase GlcNAc	
PVIT_0010915.T1	Plvit030	270	protein N-terminal amidase, similar to NTA1	
PVIT_0010916.T1	Plvit030	218	similar to sentrine-/SUMO-specific protease 8 (SENP8); contains an Ulp1 catalytic domain	
PVIT_0010917.T1	Plvit030	449	similar to 26S proteasome regulatory subunit N13; contains a ubiquitin interacting motif	
PVIT_0010918.T1	Plvit030	765	cation efflux protein/zinc transporter	
PVIT_0010919.T1	Plvit030	137	dynein light-chain roadblock type	
PVIT_0010922.T1	Plvit030	203	similar to protein N-lysine methyltransferase METTL21A	
PVIT_0010924.T1	Plvit030	388	thiol-dependent, ubiquitin-specific protease activity	
PVIT_0010925.T1	Plvit030	1,071	FYVE-finger-containing Rab5 effector protein Rabenosyn-5-related; contains a START-like domain	
PVIT_0010926.T1	Plvit030	993	calcium binding protein	
PVIT_0010927.T1	Plvit030	1,230	bactericidal permeability-increasing protein; lipid and protein binding	
PVIT_0010928.T1	Plvit030	633	similar to DNA polymerase eta; contains a DNA polymerase, Y-family, little finger domain; DNA repair	H163Y

(*Continued on next page*)

Please cite this article in press as: Dussert et al., Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.07.057

Table 1. *Continued*

Protein ID	Scaffold	Length	Putative Function	Non-synonymous Mutations[a]
PVIT_0010929.T1	PlvitO30	1,479	transmembrane protein with sterol-sensing domain; involved in lipid transport, translational initiation, cell division	I297M, L402M, E1211D
PVIT_0010931.T1	PlvitO30	417	similar to serine-/arginine-rich splicing factor	D229G, L284R, S294G, S303G, T306S, G310S, G325V
PVIT_0010932.T1	PlvitO30	390	ADP-ribosylation-factor-like-2-binding protein domain	G3D, L5F, S52T, H159Q, K219E, premature start
PVIT_0010933.T1	PlvitO30	479	contains an HRDC domain	D130N, K176R, L165I, V249I, T250S
PVIT_0027771.T1	PlvitO600	444	mitogen-activated protein kinase	
PVIT_0027772.T1	PlvitO600	842	similar to transitional endoplasmic reticulum ATPase CDC48	

See Table S3 for the detailed functional annotation for all genes in the region.
[a]Non-synonymous mutations caused by alternate alleles (i.e., present in the MAT-b allele) for SNPs associated with the mating-type phenotype

analyses in *Phytophthora* species and *Br. lactucae* [16, 19–21, 36]. This genetic determinism of mating type is unique in the tree of life and resembles determinism based on sex chromosomes in plants and animals [6], with one sex typically being homozygous and the other heterozygous. By contrast, fungal mating types are determined at the haploid stage, with mating occurring between gametes carrying different mating-type alleles and many species harboring multiple alleles. Diploid or dikaryotic fungal individuals are necessarily heterozygous at the mating-type locus in heterothallic species and can thus undergo selfing [37]. Mating types in fungi therefore do not prevent diploid selfing; they only prevent same-clone mating [7], although mating types in oomycetes and plants do prevent diploid selfing.

As expected for non-recombining regions under long-term balancing selection, we found the two *Pl. viticola* mating-type alleles to be highly divergent. The mating-type locus in *Ph. infestans* is also heteromorphic, with highly differentiated alleles, hemizygous fragments, and genomic translocations, and it is associated with deleterious recessive alleles [18, 36, 38]. The high degree of divergence between alleles, the number of rearrangements, the high repeat content, and the sheltering of deleterious alleles by permanent heterozygosity are typical of sex and mating-type chromosomes in plants, animals, and fungi [39–42].

The identification of regions orthologous to the *Pl. viticola* mating-type locus in other oomycete species paves the way for comparative genomic studies. Further association studies are, however, required to test whether these regions also control mating-type determination in other oomycetes. Only in *Ph. capsici* were enough genomic data available for association studies with mating type: a previous study identified scaffolds with different allelic frequencies between mating types [21], but these scaffolds did not include any of the genes orthologous to the putative *Pl. viticola* mating-type genes. Mating-type genes may have moved into different genomic backgrounds during the *Ph. capsici* and *Pl. viticola* divergence, as is often observed for animal sex-determining genes, with a high rate of sex chromosome turnover [43]. Alternatively, the mating-type determinism itself may have changed, as it occurred in fungi: mating type is determined by pheromone and pheromone receptor genes in basidiomycetes, although it is determined by transcription factors controlling the expression of these genes in ascomycetes [44]. *Ph. capsici* and *Pl. viticola* are very distantly related [45], leaving time for genetic changes in mating-type locus or mating-type determinism. Additional association studies in other oomycetes species and functional validation are consequently required for drawing inferences on mating-type gene evolution in oomycetes.

In conclusion, our study identifies the localization of the mating-type genomic sequence in the oomycete pathogen *Pl. viticola* responsible for the devastating downy mildew grapevine disease and elucidates the determinism of mating type (homozygous versus heterozygous). We identified possible mating-type-determining genes, including one encoding a putative hormone receptor. This may guide the development of innovative control methods based on disruption of the sexual cycle of the pathogen. The pathogen indeed overwinters as oospores, which play a major role during grapevine downy mildew epidemics [46]. Furthermore, the identification of the mating-type locus in such an economically important crop pathogen may improve our understanding of pathogen adaptation, as sex and outcrossing promote rapid evolution and sexual structures confer resistance to harsh conditions.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact

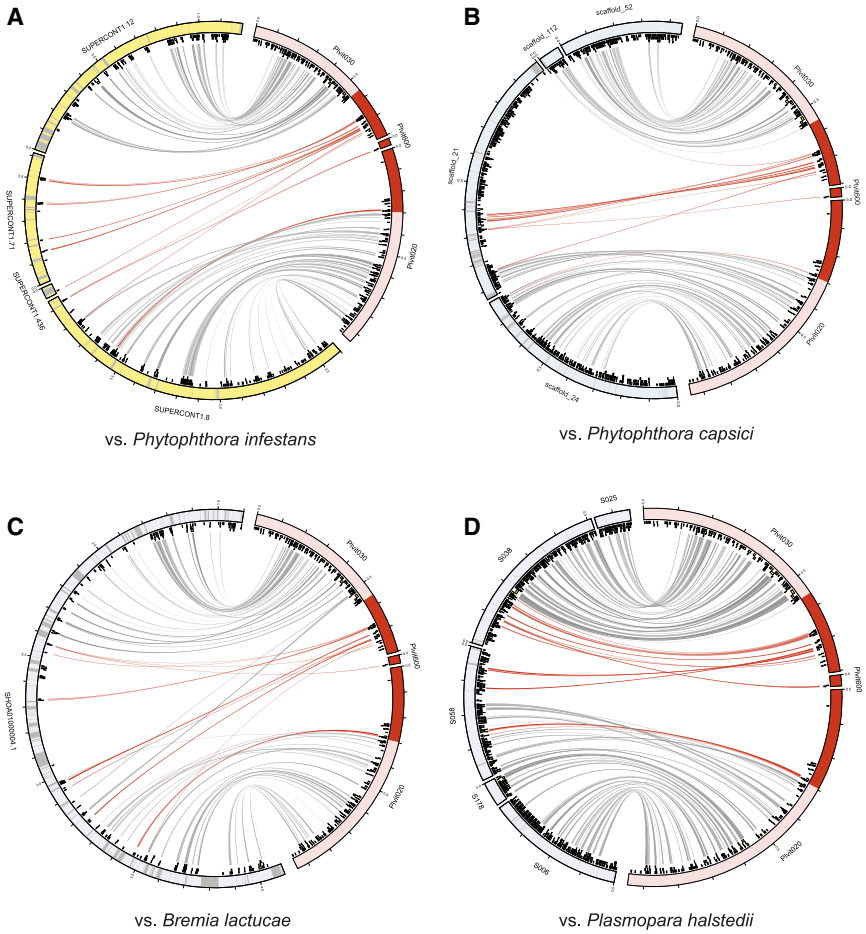

Figure 5. Orthologs of the *Plasmopara viticola* Mating-type Locus Genes in the Genomes of *Phytophthora infestans*, *Phytophthora capsici*, *Bremia lactucae*, and *Plasmopara halstedii*

Orthology relationships among protein-coding sequences between *Pl. viticola* and (A) *Ph. infestans*, (B) *Ph. capsici*, (C) *Br. lactucae*, and (D) *Pl. halstedii* were determined with reciprocal best hits and are represented by links between ideograms. The regions significantly associated with the mating-type phenotype in *Pl. viticola* and the links corresponding to genes in these regions are shown in dark red. Predicted genes are represented by yellow boxes, and the grayed-out regions in ideograms represent gaps in the assemblies. See Table S4 for the corresponding genomic coordinates in the mating-type region.

- Materials Availability
- Data and Code Availability

- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - DNA extraction and whole-genome resequencing
  - Mating-type determination
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Read mapping and single-nucleotide polymorphism (SNP) calling
  - Genetic diversity and association of SNPs with mating types
  - Identification of genomic regions associated with mating type and absent from the reference genome
  - Comparative genomics with *Phytophthora infestans*, *Phytophthora capsici*, *Bremia lactucae* and *Plasmopara halstedii*

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2020.07.057.

ACKNOWLEDGMENTS

We thank Vincent Thomas for helping with the production of material and Pal Kozma (University of Pécs, Hungary), Mauro Jermini (Agroscope, Switzerland), Sara Legler (UCSC, Italy), Hervé Steva (Bordeaux, France), and Atanas Atanassov (JGC, Bulgaria) for providing *Pl. viticola* isolates. We are grateful to Jérôme Gouzy (INRAE, Toulouse) for his help and suggestions concerning data analysis. We thank the Genotoul bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo Genotoul) for providing help and computing resources. This work was performed in collaboration with the GeT core facility, Toulouse, France (http://get.genotoul.fr). This study was supported by the European Commission (INNOVINE, FP7-KBBE-311775) and the French National Research Agency (GANDALF project, ANR-12-ADAP-0009; EFFECTOORES project, ANR-13-ADAP-0003; Investments for the Future Program in the Cluster of Excellence COTE, ANR-10-LABX-45; Investments for the Future Program France Génomique national infrastructure, ANR-10-INBS-09). T.G. acknowledges receipt of the ERC advanced grant EvolSexChrom no. 832252.

AUTHOR CONTRIBUTIONS

Conceptualization, F.D.; Methodology, Y.D. and F.D.; Formal Analysis, Y.D. and L.L.; Investigation, I.D.M., C.C., M.-C.P., R.-F.S., O.B., P.M., F.D., and S.L.T.; Writing – Original Draft, Y.D. and T.G.; Writing – Review & Editing, Y.D., L.L., I.D.M., C.C., M.-C.P., R.-F.S., O.B., P.M., S.L.T., T.G., and F.D.; Supervision, F.D. and T.G.; Funding Acquisition, P.M. and F.D.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 24, 2020
Revised: June 2, 2020
Accepted: July 16, 2020
Published: August 13, 2020

REFERENCES

1. Charlesworth, D., and Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Syst. *18*, 237–268.
2. Charlesworth, D., Morgan, M.T., and Charlesworth, B. (1990). Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multi-locus system with no linkage. Evolution *44*, 1469–1489.
3. Lande, R., and Schemske, D.W. (1985). The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution *39*, 24–40.
4. Igic, B., Lande, R., and Kohn, J.R. (2008). Loss of self-incompatibility and its evolutionary consequences. Int. J. Plant Sci. *169*, 93–104.
5. Lande, R., and Porcher, E. (2015). Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. Genetics *200*, 891–906.
6. Beukeboom, L.W., and Perrin, N. (2014). The Evolution of Sex Determination (Oxford University).
7. Billiard, S., López-Villavicencio, M., Hood, M.E., and Giraud, T. (2012). Sex, outcrossing and mating types: unsolved questions in fungi and beyond. J. Evol. Biol. *25*, 1020–1038.
8. Orias, E., Singh, D.P., and Meyer, E. (2017). Genetics and epigenetics of mating type determination in *Paramecium* and *Tetrahymena*. Annu. Rev. Microbiol. *71*, 133–156.
9. Sekimoto, H. (2017). Sexual reproduction and sex determination in green algae. J. Plant Res. *130*, 423–431.
10. Fujii, S., Kubo, K., and Takayama, S. (2016). Non-self- and self-recognition models in plant self-incompatibility. Nat. Plants *2*, 16130.
11. Judelson, H.S. (2009). Sexual reproduction in oomycetes: biology, diversity and contributions to fitness. In Oomycete Genetics and Genomics: Diversity, Interactions and Research Tools, K. Lamour, and S. Kamoun, eds. (John Wiley & Sons), pp. 121–138.
12. Lévesque, C.A. (2011). Fifty years of oomycetes—from consolidation to evolutionary and genomic exploration. Fungal Divers. *50*, 35.
13. van West, P. (2006). *Saprolegnia parasitica*, an oomycete pathogen with a fishy appetite: new challenges for an old problem. Mycologist *20*, 99–104.
14. Brurberg, M.B., Elameen, A., Le, V.H., Naerstad, R., Hermansen, A., Lehtinen, A., Hannukkala, A., Nielsen, B., Hansen, J., Andersson, B., and Yuen, J. (2011). Genetic analysis of *Phytophthora infestans* populations in the Nordic European countries reveals high genetic variability. Fungal Biol. *115*, 335–342.
15. Gavino, P.D., Smart, C.D., Sandrock, R.W., Miller, J.S., Hamm, P.B., Lee, T.Y., Davis, R.M., and Fry, W.E. (2000). Implications of sexual reproduction for *Phytophthora infestans* in the United States: generation of an aggressive lineage. Plant Dis. *84*, 731–735.
16. Fabritius, A.L., and Judelson, H.S. (1997). Mating-type loci segregate aberrantly in *Phytophthora infestans* but normally in *Phytophthora parasitica*: implications for models of mating-type determination. Curr. Genet. *32*, 60–65.
17. Judelson, H.S., Spielman, L.J., and Shattock, R.C. (1995). Genetic mapping and non-Mendelian segregation of mating type loci in the oomycete, *Phytophthora infestans*. Genetics *141*, 503–512.
18. van der Lee, T., Testa, A., Robold, A., van 't Klooster, J., and Govers, F. (2004). High-density genetic linkage maps of *Phytophthora infestans* reveal trisomic progeny and chromosomal rearrangements. Genetics *167*, 1643–1661.
19. Lamour, K.H., Mudge, J., Gobena, D., Hurtado-Gonzales, O.P., Schmutz, J., Kuo, A., Miller, N.A., Rice, B.J., Raffaele, S., Cano, L.M., et al. (2012). Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. Mol. Plant Microbe Interact. *25*, 1350–1360.
20. Sicard, D., Legg, E., Brown, S., Babu, N.K., Ochoa, O., Sudarshana, P., and Michelmore, R.W. (2003). A genetic map of the lettuce downy mildew pathogen, *Bremia lactucae*, constructed from molecular markers and avirulence genes. Fungal Genet. Biol. *39*, 16–30.
21. Carlson, M.O., Gazave, E., Gore, M.A., and Smart, C.D. (2017). Temporal genetic dynamics of an experimental, biparental field population of *Phytophthora capsici*. Front. Genet. *8*, 26.
22. Harutyunyan, S.R., Zhao, Z., Hartog, Td., Bouwmeester, K., Minnaard, A.J., Feringa, B.L., and Govers, F. (2008). Biologically active *Phytophthora* mating hormone prepared by catalytic asymmetric total synthesis. Proc. Natl. Acad. Sci. USA *105*, 8507–8512.
23. Ojika, M., Molli, S.D., Kanazawa, H., Yajima, A., Toda, K., Nukada, T., Mao, H., Murata, R., Asano, T., Qi, J., and Sakagami, Y. (2011). The second *Phytophthora* mating hormone defines interspecies biosynthetic cross-talk. Nat. Chem. Biol. *7*, 591–593.
24. Qi, J., Asano, T., Jinno, M., Matsui, K., Atsumi, K., Sakagami, Y., and Ojika, M. (2005). Characterization of a Phytophthora mating hormone. Science *309*, 1828.
25. Fontaine, M.C., Austerlitz, F., Giraud, T., Labbé, F., Papura, D., Richard-Cervera, S., and Delmotte, F. (2013). Genetic signature of a range expansion and leap-frog event after the recent invasion of Europe by the grapevine downy mildew pathogen *Plasmopara viticola*. Mol. Ecol. *22*, 2771–2786.
26. Vercesi, A., Toffolatti, S.L., Zocchi, G., Guglielmann, R., and Ironi, L. (2010). A new approach to modelling the dynamics of oospore germination in *Plasmopara viticola*. Eur. J. Plant Pathol. *128*, 113–126.
27. Wong, F.P., Burr, H.N., and Wilcox, W.F. (2001). Heterothallism in *Plasmopara viticola*. Plant Pathol. *50*, 427–432.
28. Delmas, C.E.L., Dussert, Y., Delière, L., Couture, C., Mazet, I.D., Richart Cervera, S., and Delmotte, F. (2017). Soft selective sweeps in fungicide resistance evolution: recurrent mutations without fitness costs in grapevine downy mildew. Mol. Ecol. *26*, 1936–1951.
29. Toffolatti, S.L., Prandato, M., Serrati, L., Sierotzki, H., Gisi, U., and Vercesi, A. (2011). Evolution of QoI resistance in *Plasmopara viticola* oospores. Eur. J. Plant Pathol. *129*, 331–338.
30. 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 Biol. *10*, 147.
31. Delmas, C.E.L., Fabre, F., Jolivet, J., Mazet, I.D., Richart Cervera, S., Delière, L., and Delmotte, F. (2016). Adaptation of a plant pathogen to partial host resistance: selection for greater aggressiveness in grapevine downy mildew. Evol. Appl. *9*, 709–725.
32. Dussert, Y., Mazet, I.D., Couture, C., Gouzy, J., Piron, M.-C., Kuchly, C., Bouchez, O., Rispe, C., Mestre, P., and Delmotte, F. (2019). A high-quality grapevine downy mildew genome assembly reveals rapidly evolving and lineage-specific putative host adaptation genes. Genome Biol. Evol. *11*, 954–969.
33. Peterlongo, P., Riou, C., Drezen, E., and Lemaitre, C. (2017). DiscoSnp++: de novo detection of small variants from raw unassembled read set(s). bioRxiv. https://doi.org/10.1101/209965.
34. Zhao, X., Mehrabi, R., and Xu, J.R. (2007). Mitogen-activated protein kinase pathways and fungal pathogenesis. Eukaryot. Cell *6*, 1701–1714.
35. Karos, M., Chang, Y.C., McClelland, C.M., Clarke, D.L., Fu, J., Wickes, B.L., and Kwon-Chung, K.J. (2000). Mapping of the Cryptococcus neoformans MATalpha locus: presence of mating type-specific mitogen-activated protein kinase cascade homologs. J. Bacteriol. *182*, 6222–6227.
36. Judelson, H.S. (1996). Chromosomal heteromorphism linked to the mating type locus of the oomycete *Phytophthora infestans*. Mol. Gen. Genet. *252*, 155–161.
37. Giraud, T., Yockteng, R., López-Villavicencio, M., Refrégier, G., and Hood, M.E. (2008). Mating system of the anther smut fungus *Microbotryum violaceum*: selfing under heterothallism. Eukaryot. Cell *7*, 765–775.
38. Randall, T.A., Ah Fong, A., and Judelson, H.S. (2003). Chromosomal heteromorphism and an apparent translocation detected using a BAC contig

spanning the mating type locus of *Phytophthora infestans*. Fungal Genet. Biol. *38*, 75–84.

39. Bachtrog, D. (2005). Sex chromosome evolution: molecular aspects of Y-chromosome degeneration in *Drosophila*. Genome Res. *15*, 1393–1401.
40. Charlesworth, D. (2015). Plant contributions to our understanding of sex chromosome evolution. New Phytol. *208*, 52–65.
41. Badouin, H., Hood, M.E., Gouzy, J., Aguileta, G., Siguenza, S., Perlin, M.H., Cuomo, C.A., Fairhead, C., Branca, A., and Giraud, T. (2015). Chaos of rearrangements in the mating-type chromosomes of the anther-smut fungus *Microbotryum lychnidis-dioicae*. Genetics *200*, 1275–1284.
42. Graves, J.A.M. (2006). Sex chromosome specialization and degeneration in mammals. Cell *124*, 901–914.
43. Jeffries, D.L., Lavanchy, G., Sermier, R., Sredl, M.J., Miura, I., Borzée, A., Barrow, L.N., Canestrelli, D., Crochet, P.A., Dufresnes, C., et al. (2018). A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. Nat. Commun. *9*, 4088.
44. Heitman, J., Kronstad, J.W., Taylor, J.W., and Casselton, L.A. (2007). Sex in Fungi: Molecular Determination and Evolutionary Implications (ASM).
45. Bourret, T.B., Choudhury, R.A., Mehl, H.K., Blomquist, C.L., McRoberts, N., and Rizzo, D.M. (2018). Multiple origins of downy mildews and mito-nuclear discordance within the paraphyletic genus *Phytophthora*. PLoS ONE *13*, e0192502.
46. Gobbin, D., Jermini, M., Loskill, B., Pertot, I., Raynal, M., and Gessler, C. (2005). Importance of secondary inoculum of *Plasmopara viticola* to epidemics of grapevine downy mildew. Plant Pathol. *54*, 522–534.
47. Haas, B.J., Kamoun, S., Zody, M.C., Jiang, R.H.Y., Handsaker, R.E., Cano, L.M., Grabherr, M., Kodira, C.D., Raffaele, S., Torto-Alalibo, T., et al. (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature *461*, 393–398.
48. Fletcher, K., Gil, J., Bertier, L.D., Kenefick, A., Wood, K.J., Zhang, L., Reyes-Chin-Wo, S., Cavanaugh, K., Tsuchida, C., Wong, J., and Michelmore, R. (2019). Genomic signatures of heterokaryosis in the oomycete pathogen *Bremia lactucae*. Nat. Commun. *10*, 2645.
49. Pecrix, Y., Buendia, L., Penouilh-Suzette, C., Maréchaux, M., Legrand, L., Bouchez, O., Rengel, D., Gouzy, J., Cottret, L., Vear, F., and Godiard, L. (2019). Sunflower resistance to multiple downy mildew pathotypes revealed by recognition of conserved effectors of the oomycete *Plasmopara halstedii*. Plant J. *97*, 730–748.
50. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.
51. Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. *22*, 568–576.
52. Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics *27*, 2987–2993.
53. Cingolani, P., Platts, A., Wang, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin) *6*, 80–92.
54. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.
55. Zhang, C., Dong, S.-S., Xu, J.-Y., He, W.-M., and Yang, T.-L. (2019). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Bioinformatics *35*, 1786–1788.
56. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.
57. Luu, K., Bazin, E., and Blum, M.G.B. (2017). *pcadapt*: an R package to perform genome scans for selection based on principal component analysis. Mol. Ecol. Resour. *17*, 67–77.
58. R Core Team (2017). R: A language and environment for statistical computing (R Foundation for Statistical Computing). http://www.R-project.org/.
59. Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., and François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. Genetics *196*, 973–983.
60. Frichot, E., and François, O. (2015). LEA: an R package for landscape and ecological association studies. Methods Ecol. Evol. *6*, 925–929.
61. Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y., and Buckler, E.S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics *23*, 2633–2635.
62. 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.
63. Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004). Versatile and open software for comparing large genomes. Genome Biol. *5*, R12.
64. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
65. Sedlazeck, F.J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., von Haeseler, A., and Schatz, M.C. (2018). Accurate detection of complex structural variations using single-molecule sequencing. Nat. Methods *15*, 461–468.
66. Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., and Earl, A.M. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE *9*, e112963.
67. Quesneville, H., Bergman, C.M., Andrieu, O., Autard, D., Nouaud, D., Ashburner, M., and Anxolabehere, D. (2005). Combined evidence annotation of transposable elements in genome sequences. PLoS Comput. Biol. *1*, 166–175.
68. Flutre, T., Duprat, E., Feuillet, C., and Quesneville, H. (2011). Considering transposable element diversification in *de novo* annotation approaches. PLoS ONE *6*, e16526.
69. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.
70. 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.
71. Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics *30*, 1236–1240.
72. Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talón, M., Dopazo, J., and Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. *36*, 3420–3435.
73. Kohany, O., Gentles, A.J., Hankus, L., and Jurka, J. (2006). Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics *7*, 474.
74. Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. *19*, 1639–1645.

75. Sung, C.R.-T.M., and Clerjeau, M. (1988). Techniques for formation, maturation, and germination of *Plasmopara viticola* oospores under controlled conditions. Plant Dis. *72*, 938–941.

76. Privé, F., Aschard, H., Ziyatdinov, A., and Blum, M.G.B. (2018). Efficient analysis of large-scale genome-wide data with two R packages: bigstatsr and bigsnpr. Bioinformatics *34*, 2781–2787.

77. Dussert, Y., Gouzy, J., Richart-Cervera, S., Mazet, I.D., Delière, L., Couture, C., Legrand, L., Piron, M.-C., Mestre, P., and Delmotte, F. (2016). Draft genome sequence of *Plasmopara viticola*, the grapevine downy mildew pathogen. Genome Announc. *4*, e009877-16.

78. Mestre, P., Carrere, S., Gouzy, J., Piron, M.-C., Tourvieille de Labrouhe, D., Vincourt, P., Delmotte, F., and Godiard, L. (2016). Comparative analysis of expressed CRN and RXLR effectors from two *Plasmopara* species causing grapevine and sunflower downy mildew. Plant Pathol. *65*, 767–781.

Please cite this article in press as: Dussert et al., Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.07.057

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
CTAB	Sigma-Aldrich	CAT#H6269-250 g
Chloroform-Isoamyl alcohol	Sigma-Aldrich	CAT#C0549-1PJ
RNase A	QIAGEN	CAT#19101
Sodium Acetate (3 M), pH 5.5, RNase-free	Invitrogen	CAT#AM9740
TE Buffer	Invitrogen	CAT#12090015
Tris (1 M), pH 7.0, RNase-free	Invitrogen	CAT#AM9851
UltraPure 0.5M EDTA, pH 8.0	Invitrogen	CAT#15575020
Sodium dodecyl sulfate	Sigma-Aldrich	CAT#436143-25G
Agar	SETEXAM	CAT#FAL2
Critical Commercial Assays		
DNeasy Plant Mini Kit	QIAGEN	CAT#69104
Deposited Data		
Plasmopara viticola INRA-PV221 reference genome & sequencing reads	[32]	PRJNA329579
Plasmopara viticola sequencing reads	This paper	PRJNA644583
Phytophthora infestans reference genome	[47]	GCA_000142945.1
Phytophthora capsici reference genome	[19]	GCA_000325885.1
Bremia lactucae reference genome	[48]	GCA_004359215.1
Plasmopara halstedii reference genome	[49]	GCA_003724065.1
Experimental Models: Organisms/Strains		
Plasmopara viticola	See Table S1	See Table S1
Software and Algorithms		
Glint v1.0.rc12.826:833	N/A	https://forge-dga.jouy.inra.fr/projects/glint
SAMtools 1.3.1	[50]	http://www.htslib.org/download/
VARSCAN 2.4.3	[51]	http://varscan.sourceforge.net/
BCFtools 1.1-60-g3d5d3d9	[52]	http://www.htslib.org/download/
snpEff 4.3	[53]	http://snpeff.sourceforge.net/
VCFtools	[54]	https://vcftools.github.io/
PopLDdecay 3.28	[55]	https://github.com/BGI-shenzhen/PopLDdecay
PLINK 1.07	[56]	http://zzz.bwh.harvard.edu/plink/
PCAdapt R package	[57]	https://cran.r-project.org/web/packages/pcadapt/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
R	[58]	https://www.R-project.org/
sNMF	[59]	N/A
LEA R package	[60]	https://www.bioconductor.org/packages/release/bioc/html/LEA.html
TASSEL 5.2.41	[61]	https://www.maizegenetics.net/tassel
discoSnp++	[33]	https://github.com/GATB/DiscoSnp
BLAST+ 2.4.0	[62]	ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/
SMARTdenovo	N/A	https://github.com/ruanjue/smartdenovo
MUMmer 3.23	[63]	http://mummer.sourceforge.net/
BWA 0.7.12	[64]	https://github.com/lh3/bwa
NGMLR	[65]	https://github.com/philres/ngmlr
Pilon 1.22	[66]	https://github.com/broadinstitute/pilon
REPET 2.2	[67, 68]	https://urgi.versailles.inra.fr/Tools/REPET
STAR 2.4	[69]	https://github.com/alexdobin/STAR
AUGUSTUS 3.1	[70]	https://github.com/Gaius-Augustus/Augustus
InterProScan 5	[71]	https://www.ebi.ac.uk/interpro/download/
TransposonPSI	N/A	http://transposonpsi.sourceforge.net
Blast2GO 4.1.9	[72]	https://www.blast2go.com/blast2go-pro/download-b2g
CENSOR 4.2.29	[73]	https://www.girinst.org/censor/
Circos	[74]	http://circos.ca/software/download/circos/

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Yann Dussert (dussert.yann@gmail.com).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

Raw reads for each individual have been deposited in GenBank: PRJNA644583 and analysis files have been deposited in Dataverse: http://doi.org/10.15454/ILQ12T.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Grapevine leaves infected with *Pl. viticola* were collected across Europe (Table S1). Diploid oomycete individuals were isolated from diseased leaves by dilution to isolate single spores, as follows: isolates of *Pl. viticola* were first propagated on detached leaves of *Vitis*

Please cite this article in press as: Dussert et al., Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.07.057

vinifera cv. Cabernet-Sauvignon. Sporangia were collected and suspended in water at a concentration of 10^4 sporangia/mL (estimated using Malassez cells) and serial dilutions were performed to obtain a concentration of 0.1 sporangia/10 μL. We applied 10 μL of this inoculum on 18-mm-diameter leaf disks, excised using a cork borer from V. *vinifera* cv. Cabernet-Sauvignon leaves, which were placed on filter paper soaked with deionized water in Petri dishes. After seven days in a growth chamber at 22°C (12 h light/12 h dark photoperiod), we retained a single sporulating disk per isolate, which was, thus, derived from a single diploid sporangium.

METHOD DETAILS

DNA extraction and whole-genome resequencing

Detached leaves from *V. vinifera* cv. Cabernet-Sauvignon or Muscat Ottonel plants were sterilized with bleach, washed with sterilized water and infected with fresh *Pl. viticola* inoculum. Leaves were placed in a growth chamber for 4-7 days at 21-22°C (12 h/12 h or 16 h/8 h light/dark photoperiod). Sporangiophores and sporangia were then collected in distilled water. The suspension was centrifuged at maximum speed, the supernatant was removed and the pellet was stored at −80°C. DNA was extracted with a modified CTAB procedure [32] or with the QIAGEN DNeasy Plant Mini Kit.

Genomic DNA was sequenced by Beckman Coulter Genomics (Grenoble, France) or at the GeT-PlaGe facility (Toulouse, France), on an Illumina HiSeq 2000 sequencer (2x100 bp paired-end reads) or an Illumina HiSeq 3000 sequencer (2x150 bp), except for the PV221 individual, for which sequencing reads were already available [32] (SRA accession numbers: SRX1970160, SRX1970161, SRX1970162). Read lengths and mean genome coverage are provided in Table S1.

Mating-type determination

Six *Pl. viticola* individuals (Tables S1 and S2) were chosen as mating-type testers: three individuals of each mating type. These testers were first crossed with each other in all possible pairings, for separation of the six individuals into two opposite mating-type groups, P1 and P2 (three individuals each). We then crossed 54 individuals with these six testers to determine their mating type. In all experiments, 18-mm-diameter leaf disks, excised from *V. vinifera* cv. Cabernet-Sauvignon leaves, were placed on 20 g/L agar in Petri dishes and co-inoculated with 15 μL droplets containing $4x10^4$ sporangia/mL from each individual (three droplets per disk, six replicate disks per pairing). After four to five days in a growth chamber at 22°C (12 h light/12 h dark photoperiod), disks were examined for sporulation, to check for successful inoculation. The leaf disks were then transferred to a growth chamber at 12°C (12 h light/12 h dark photoperiod) for about three weeks [75]. Each leaf disk was then observed with a binocular microscope, to check for the presence of oospores, indicative of successful mating.

QUANTIFICATION AND STATISTICAL ANALYSIS

Read mapping and single-nucleotide polymorphism (SNP) calling

Reads from each of the 54 sequenced individuals were mapped onto the *Pl. viticola* PacBio reference genome [32] with the Glint aligner v1.0.rc12.826:833, allowing a maximum of 15% mismatches, to handle genomic regions with high heterozygosity (other parameters: mappe mode,–no-lc-filtering–best-score–lrmin 80), and filtering out reads with a mapping quality below 3 and reads that were not properly paired.

A pileup file was generated from the resulting alignment files with SAMtools 1.3.1 [50] mpileup without probabilistic alignment quality computation (-B parameter). SNPs and short insertions and deletions (indels) were then called with the multi-sample procedure (mpileup2snp and mpileup2indel) of VARSCAN 2.4.3 [51] (–min-coverage 10–min-reads2 5–min-avg-qual 15–min-var-freq 0.2–min-freq-for-hom 0.75–p value 0.001–strand-filter 1). For each individual, sites with a read coverage more than 1.5 standard deviations on either side of the mean value for genome coverage were discarded with BCFtools 1.1-60-g3d5d3d9 [52]. SNPs located within 2 bp of a detected indel were also filtered out. Finally, only sites with less than 10% missing data were retained. After filtering, the dataset contained 2,011 million SNPs. Variant annotation was carried out using snpEff 4.3 [53] (parameters:–no-downstream –no-upstream).

Genetic diversity and association of SNPs with mating types

We used SAMtools to assess the normalized read coverage for 1 kb windows for each individual. Using only SNPs with a minor allele frequency (MAF) below 0.1, the percentage of heterozygous individuals for each site along the genome was calculated separately for P1 and P2 individuals with VCFtools [54]. Linkage disequilibrium between sites, as assessed with the r^2 statistic, was also computed with VCFtools. Finally, linkage disequilibrium decay along scaffolds was analyzed with PopLDdecay 3.28 [55].

SNP thinning was performed by clumping [76] with PLINK 1.07 [56], with MAF as the statistic of importance and a LD threshold (r^2) of 0.2. Population genetic structure was investigated by principal component analysis (PCA) with PCAdapt [57] in R [58]. We also investigated population clustering by computing ancestry coefficients for individuals (Q-matrix) with the non-negative matrix factorization algorithm sNMF [59] implemented in the LEA package [60] in R. We set the number of groups *K* between 1 and 10, with 50 runs for each value of *K*. The optimal number of groups was chosen using the minimum cross-entropy criterion.

For identification of the genomic region controlling mating types, we analyzed the association between genotype and mating type (genome-wide association approach) with TASSEL 5.2.41 [61] on SNPs with no missing data and a MAF above 0.1, using a GLM

model with the Q-matrix estimated by sNMF. *P*-values were corrected with 10,000 permutations (threshold: 0.05). We filtered out the small number of association signals found in short scaffolds, corresponding to isolated SNPs in repeat regions.

Identification of genomic regions associated with mating type and absent from the reference genome

The discovery of SNPs associated with mating type in the edges of two scaffolds suggested that the mating-type locus had been only partially assembled in the reference assembly. We therefore used discoSnp++ [33], a reference-free method for identifying SNPs. Briefly, SNPs are detected from bubbles found in a de Bruijn graph built from raw sequencing reads, and short contigs are reconstructed around the SNPs. This analysis was carried out with a subset of 30 individuals (Table S1), as computing resource issues arose if the analysis was extended to a larger number of individuals. We used a *k*-mer size of 31, the smart branching strategy (-b 1), a maximum of five SNPs in unique bubbles (-P 5), and polymorphisms were extended with left and right contigs (-T). Only sites with less than 10% missing data and a MAF below 0.1 were retained. Contigs were mapped onto the reference genome with the VCF_creator module of discoSnp++ and an association analysis was carried out with TASSEL, as described above, to check that our results were robust to the SNP calling method.

Contigs from discoSnp++ including SNPs significantly associated with mating type were aligned with the self-corrected PacBio reads used for reference genome assembly [32] with BLASTN+ [62], retaining hits displaying at least 95% identity and 95% query coverage. PacBio reads corresponding to these hits were assembled with SMARTdenovo (available from https://github.com/ruanjue/smartdenovo) using default parameters. Assembled contigs were aligned against the reference scaffolds including SNPs with a significant association signal, with NUCmer from the MUMmer 3.23 package [63] using default parameters, and only regions absent from the reference assembly were retained. Available paired-end reads and 3 kb mate-pair reads [77] and the self-corrected PacBio reads were aligned against the reference genome and the new mating-type contigs, with BWA 0.7.12 [64] for short reads and NGMLR [65] for long reads. Aligned reads were used to polish the new mating-type contigs with Pilon 1.22 [66].

After sequence polishing, contigs from the discoSNP analysis were again mapped, but this time to the reference assembly and the newly assembled contigs, and a final association analysis was carried out. Newly assembled contigs with no SNP significantly associated with the matin type were discarded. For the remaining contigs, repeat sequences were annotated with TEannot from REPET 2.2 [67, 68], using the *Pl. viticola* repeat library [32]. Available RNA-seq data [32, 78] were aligned against the reference assembly and the new contigs with STAR 2.4 [69], and were used as hints for gene prediction in the new contigs with AUGUSTUS 3.1 [70], together with the training files produced by the previous annotation [32]. Proteins for which functional domains of transposable elements were detected with InterProScan 5 [71] or for which hits were obtained with TransposonPSI (http://transposonpsi.sourceforge.net; 30% coverage, e-value < $1e^{-10}$) were filtered out. The remaining proteins were aligned against the NCBI nr database with BLASTP+ (-evalue $1e^{-5}$ -max_hsp 20 -num_alignment 20), and functionally annotated with Blast2GO 4.1.9 [72] using the BLASTP+ and InterProScan results for gene ontology (GO) term mapping.

For all candidate genes within the mating-type locus (i.e., from the scaffolds in the reference assembly and the new contigs), we further checked whether coding sequences corresponded to transposable elements: protein sequences for those genes were masked with CENSOR 4.2.29 [73] using RepBase 23.08, which allowed us to compute for each protein a percentage of coverage from known transposable elements.

Comparative genomics with *Phytophthora infestans*, *Phytophthora capsici*, *Bremia lactucae* and *Plasmopara halstedii*

We identified the orthologs of the mating-type candidate genes in the annotated proteomes of *Ph. infestans* [47], *Ph. capsici* [19], *Br. lactucae* [48] and *Pl. halstedii* [49], with the reciprocal best hits (RBH) method using BLASTP+ (-evalue $1e^{-6}$ -qcov_hsp_perc 50 –use_sw_tback, min. identity 40%). Results were visualized with Circos [74].

Current Biology, Volume *30*

Supplemental Information

Identification of the First Oomycete Mating-type Locus Sequence in the Grapevine Downy Mildew Pathogen, *Plasmopara viticola*

Yann Dussert, Ludovic Legrand, Isabelle D. Mazet, Carole Couture, Marie-Christine Piron, Rémy-Félix Serre, Olivier Bouchez, Pere Mestre, Silvia Laura Toffolatti, Tatiana Giraud, and François Delmotte

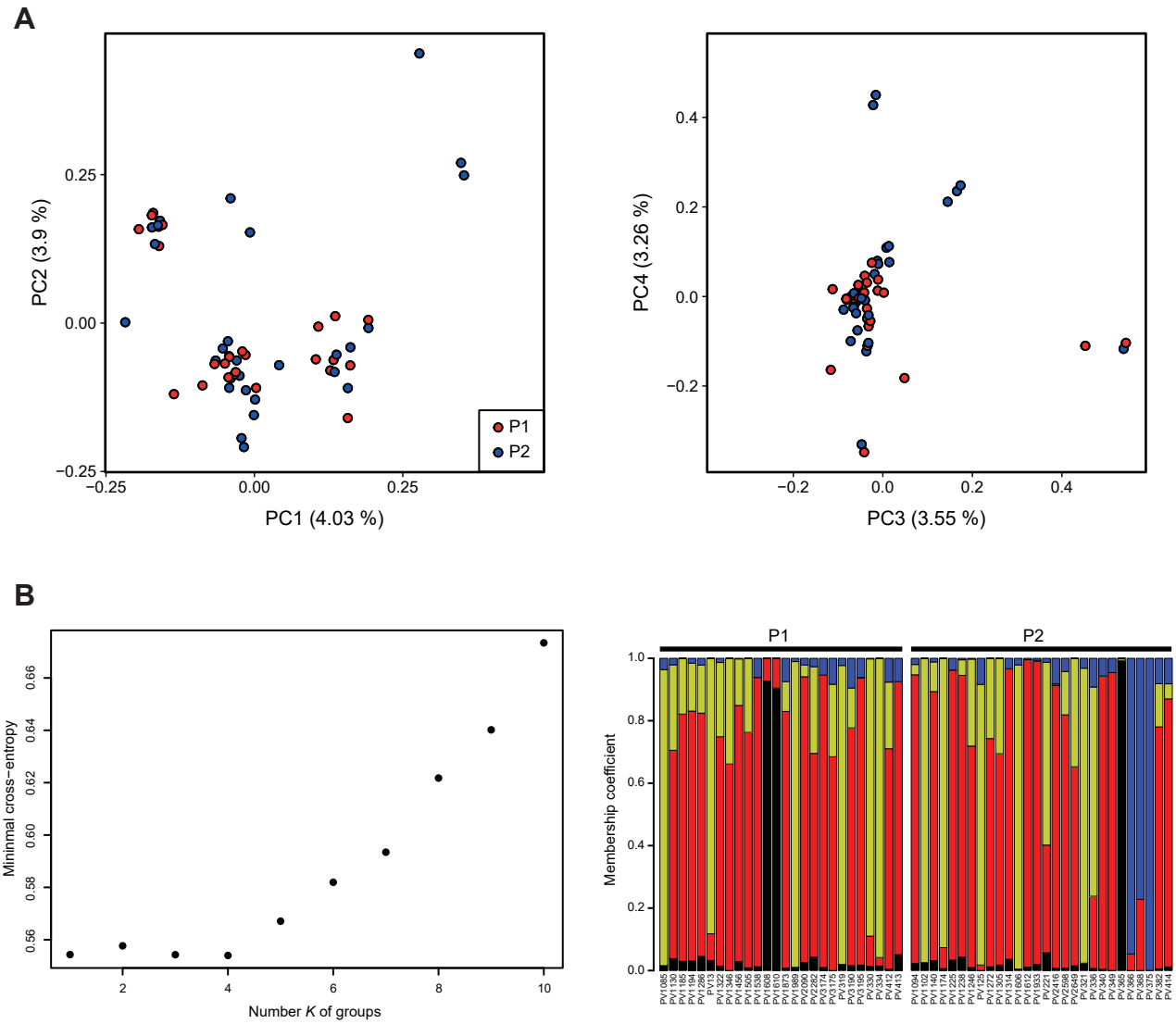

Figure S1 Population structure analysis in *Plasmopara viticola*, based on a LD-thinned dataset of 35,516 SNPs. Related to Figure 2.

A: Principal component analysis. Individuals with the P1 mating type are represented in red and individuals with the P2 mating type in blue. The percentage of variation explained by each principal component is indicated in brackets.

B: Population clustering by the non-negative matrix factorization algorithm sNMF. The values of minimum cross-entropy for each number *K* of groups are represented in the left panel. The membership coefficients for each individual for the run with the lower (i.e. better) minimum cross-entropy at *K* = 4 are represented in the right panel. Individuals are represented by vertical lines, divided into four colored sections corresponding to the individuals' membership coefficient in each group. See Table S1 for the information on individuals

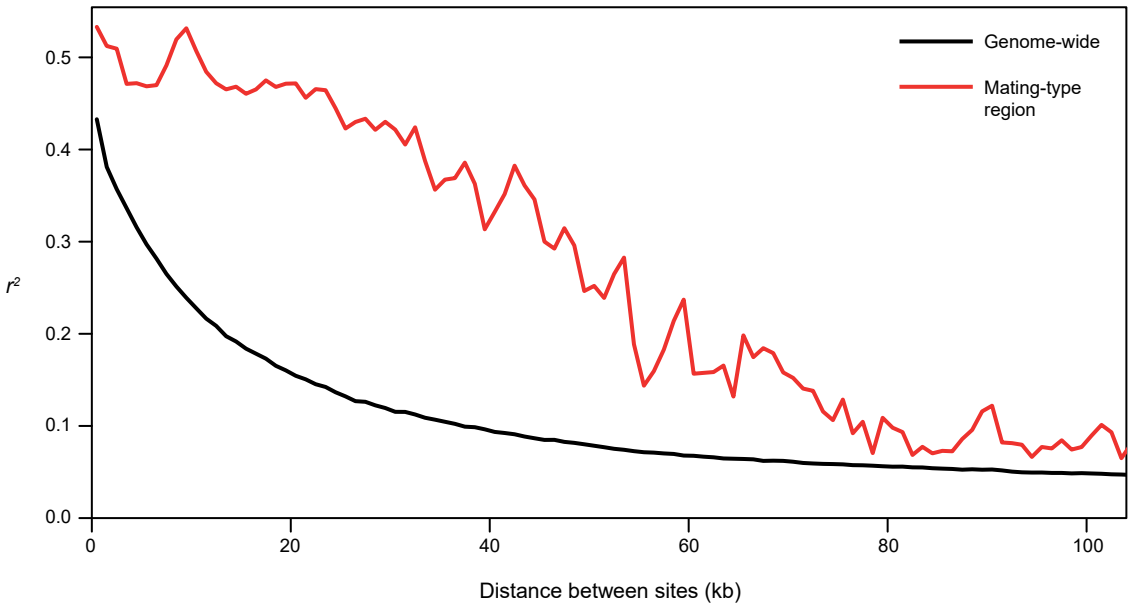

Figure S2 Linkage disequilibrium (LD) decay in *Plasmopara viticola*. Related to Figure 2; Figure 3.
The LD between SNPs, measured by r^2, is represented as a function of physical distance between SNPs, with r^2 values averaged for windows of 1000 bp. The LD decay curve is represented in black for the whole genome, and in red only for the region associated with the mating-type phenotype in the Plvit020 and Plvit030 scaffolds. This shows that the LD is much higher in the mating-type region compared to the rest of the genome, suggesting a lack of recombination in this region.

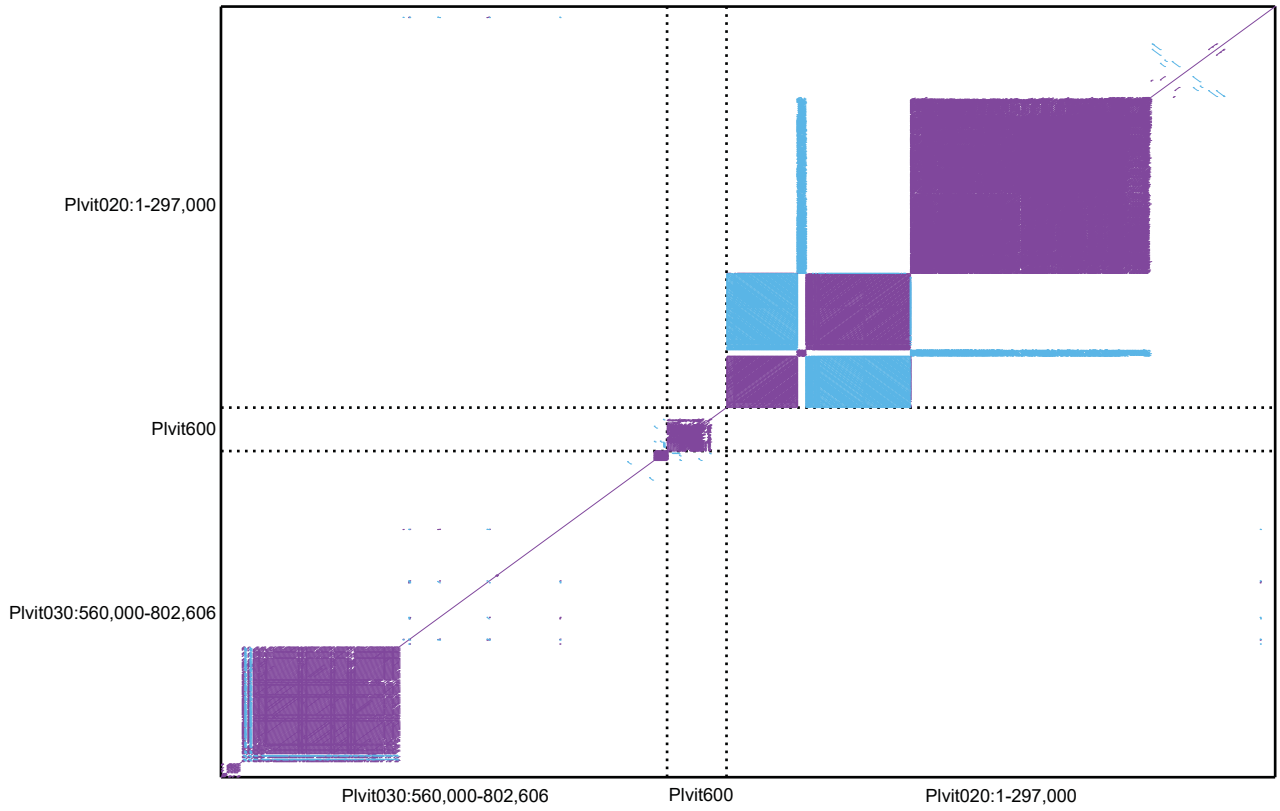

Figure S3 Tandem repeat arrays in the region associated with the mating type phenotype in *Plasmopara viticola*. Related to Figure 2; Figure 3.
All-versus-all alignments for sequences associated with the mating type are represented as a dot plot, with forward matches in purple and reverse matches in blue. Results are shown for the two scaffolds found in the reference genome (Plvit020 and Plvit030) and a third scaffold (Plvit600) assembled using data from the reference-free SNP calling method (see main text). See also Figure 3 for a representation of the repeat content and other features along the scaffolds containing the mating-type region.

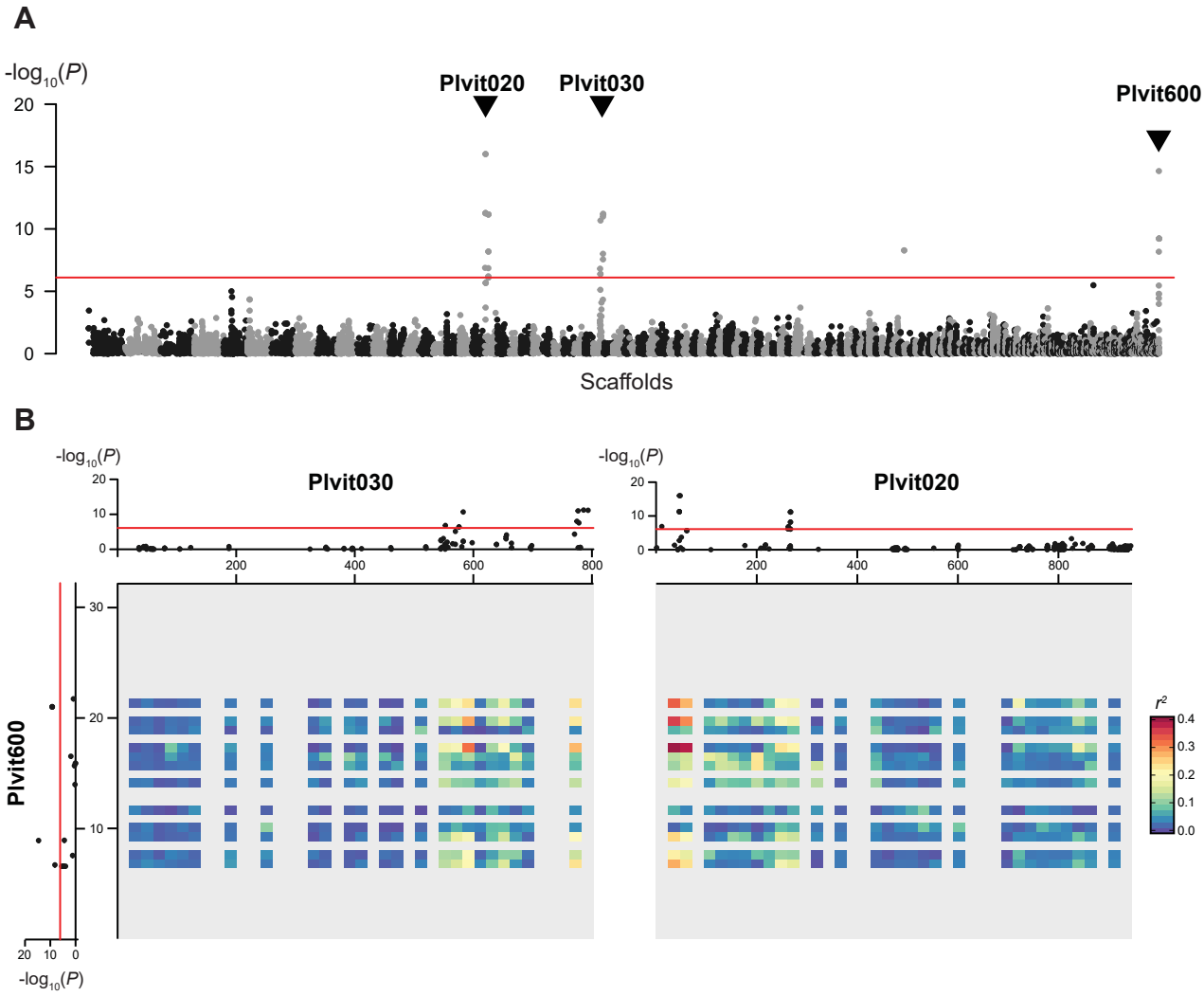

Figure S4 Genome-wide association analysis for identifying mating-type regions in *Plasmopara viticola*, based on a reference-free SNP calling method. Related to Figure 2.

A: Manhattan plot of the negative $\log_{10}$-transformed association *P*-values between mating type and SNPs along the *Pl. viticola* genome. Alternating black and grey blocks of dots mark the limits between scaffolds. Scaffolds with a significant association signal (Plvit020, Plvit030 and Plvit600) are indicated with arrows. Plvit600 was assembled using data from the reference-free SNP calling method (see main text). B: Manhattan plots of the negative $\log_{10}$-transformed association *P*-values between mating type and SNPs for Plvit020, Plvit030 and Plvit600, and linkage disequilibrium between Plvit600 and the two other scaffolds, represented as a heatmap (40 bins). The significance threshold for the association analysis, computed with 10,000 permutations, is represented as a red line in both panels.

ID	Country	Location	Latitude (decimal degrees)	Longitude (decimal degrees)	Read length (bp)	Mean coverage	Strains used for the reference-free SNP analysis
PV13	France	Latresne	44.79	-0.50	100	57.1	
PV125	Hungary	Pècs	46.08	18.23	100	107.3	
PV221[a]	France	Blanquefort	44.92	-0.62	100	100.7	X
PV319	France	Côte d'Or	NA	NA	100	93.7	X
PV321	Germany	Kröv	49.99	7.09	100	144.1	X
PV330	Germany	Pfaffenweiler	47.94	7.76	100	60.3	
PV334	Germany	Ehrenkirchen	47.91	7.75	100	79.2	X
PV336	Hungary	Eger	47.91	20.38	100	47.0	X
PV340	Hungary	Tolcsva	48.28	21.44	150	23.0	
PV349	Switzerland	Cugnasco	46.18	8.92	150	19.4	X
PV365	Switzerland	Cugnasco	46.18	8.92	100	45.0	
PV366	Switzerland	Cugnasco	46.18	8.92	150	26.3	
PV368	Switzerland	Cugnasco	46.18	8.92	150	19.3	X
PV375	Switzerland	Cugnasco	46.18	8.92	100	149.7	
PV382	Switzerland	Cugnasco	46.18	8.92	150	33.8	
PV412	Switzerland	Cugnasco	46.18	8.92	150	37.5	
PV413[a]	Switzerland	Cugnasco	46.18	8.92	150	23.5	X
PV414	Switzerland	Cugnasco	46.18	8.92	150	27.8	
PV1085	Switzerland	Chexbres	46.48	6.77	100	126.8	X
PV1094	Switzerland	Chexbres	46.48	6.77	150	24.0	X
PV1102	Switzerland	Nyon	46.40	6.23	100	115.2	
PV1130	Switzerland	Leytron	46.19	7.22	150	32.7	
PV1140	Switzerland	Leytron	46.19	7.22	150	23.3	X
PV1174	Germany	Ihringen	48.05	7.62	100	168.0	X
PV1185	Germany	Ihringen	48.05	7.62	150	30.0	
PV1194	Germany	Ihringen	48.05	7.63	150	28.8	X
PV1225	Germany	Ihringen	48.05	7.63	150	18.7	X
PV1238[a]	Germany	Pfaffenweiler	47.94	7.76	150	23.0	X
PV1246	Germany	Ehrenkirchen	47.93	7.73	150	36.4	
PV1272	Germany	Ihringen	48.05	7.62	150	32.3	X
PV1286	Germany	Ehrenkirchen	47.93	7.73	150	27.5	X
PV1305	Germany	Ebringen	47.96	7.78	150	33.8	
PV1314	Germany	Ebringen	47.96	7.78	150	17.9	
PV1322	Germany	Ehrenkirchen	47.93	7.73	150	29.6	
PV1346	Germany	Pfaffenweiler	47.93	7.75	150	34.8	
PV1400[a]	Germany	Pfaffenweiler	47.94	7.75	-[b]	-[b]	
PV1456	Germany	Pfaffenweiler	47.94	7.76	150	28.4	X
PV1505	Germany	Pfaffenweiler	47.94	7.76	150	30.9	
PV1538[a]	Czech Republic	NA	NA	NA	150	21.7	X
PV1606	Switzerland	Leytron	46.19	7.22	100	109.7	
PV1608	Switzerland	Pully	46.51	6.67	150	25.7	X
PV1610[a]	Switzerland	Pully	46.51	6.67	150	27.2	X
PV1612	Switzerland	Pully	46.51	6.67	150	22.4	X
PV1873	Switzerland	Cugnasco	46.18	8.92	150	28.7	X
PV1933	Switzerland	Pully	46.51	6.67	150	25.3	X
PV1989	France	Pouilly le Monial	45.96	4.65	100	84.4	
PV2090	France	Arles	43.68	4.63	150	23.9	X
PV2282	Spain	Sant Sadurni D'Anoia	41.43	1.79	150	34.1	X
PV2416	Italy	Pianello Val Tidone	44.93	9.38	150	27.0	
PV2598	Spain	Agoncillo	42.47	-2.29	150	31.1	X
PV2649	France	Nogent-l'Abesse	49.24	4.16	150	39.7	X
PV3174	Italy	Piateda	46.16	9.94	150	15.5	X
PV3175	Italy	Canevino	44.94	9.28	150	38.9	
PV3190	Georgia	NA	NA	NA	150	34.0	
PV3195	Bulgaria	Vidin	43.99	22.87	150	22.1	X

Table S1 *Plasmopara viticola* individuals used for the mating-type association study. Related to STAR Methods.

The country and location (town or region) of collection are indicated, with their GPS coordinates, as well as the mean read coverage for genome sequencing, and the strains used for the reference-free SNP analysis for the de novo SNP detection with discoSNP++.

[a]: tester individuals used for mating-type phenotyping

[b]: resequencing was not successful for this individual

	P1 testers			P2 testers			
	PV413	**PV1538**	**PV1610**	**PV1238**	**PV221**	**PV1400**	Mating type
PV13				■	■	■	P1
PV125	■						P2
PV221	■	■					**P2**
PV319					■		P1
PV321	■	■					P2
PV330				■	■	■	P1
PV334					■		P1
PV336			■				P2
PV340	■	■					P2
PV349		■					P2
PV365			■				P2
PV366	■		■				P2
PV368	■						P2
PV375			■				P2
PV382	■	■					P2
PV412				■			P1
PV413				■	■		**P1**
PV414	■						P2
PV1085					■		P1
PV1094	■	■	■				P2
PV1102	■	■	■				P2
PV1130				■			P1
PV1140	■	■	■				P2
PV1174	■						P2
PV1185				■			P1
PV1194				■		■	P1
PV1225			■				P2
PV1238	■	■					**P2**
PV1246	■	■	■				P2
PV1272	■	■					P2
PV1286				■	■	■	P1
PV1305			■				P2
PV1314		■					P2
PV1322				■			P1
PV1346				■	■		P1
PV1400	■	■					**P2**
PV1456				■			P1
PV1505				■	■		P1
PV1538				■	■		**P1**
PV1606		■					P2
PV1608					■		P1
PV1610				■	■		**P1**
PV1612	■	■					P2
PV1873				■	■		P1
PV1933		■	■				P2
PV1989				■		■	P1
PV2090				■			P1
PV2282				■		■	P1
PV2416		■					P2
PV2598	■	■	■				P2
PV2649		■					P2
PV3174				■	■		P1
PV3175					■		P1
PV3190				■			P1
PV3195				■	■		P1

Table S2 Mating-type phenotyping for *Plasmopara viticola* individuals. Related to STAR Methods.

All individuals (rows) have been crossed with the six testers (columns). Successful crosses (i.e. with oospores) are indicated by black cells.

Individuals successfully mating with P2 testers were determined as having the P1 mating type, and vice versa. Individual IDs in bold designate tester individuals.

Protein ID	Length	Coverage TE (%)	Panther definition	GO terms	InterPro domains	KEGG definition	KEGG pathway	Non-syn mutations
PVIT_0008596.T1	135	0.00						
PVIT_0008599.T1	101	99.01		nucleic acid binding				
PVIT_0008600.T1	795	16.60		transferase activity, transferring acyl groups; transferase activity				
PVIT_0008601.T1	328	38.41						
PVIT_0008603.T1	193	79.27						
PVIT_0008604.T1	358	80.45	FAMILY NOT NAMED	transferase activity, transferring acyl groups; transferase activity	Endonuclease/exonuclease/phosphatase			
PVIT_0008605.T1	752	25.66		transferase activity, transferring acyl groups; transferase activity				
PVIT_0008608.T1	713	0.00		transferase activity, transferring acyl groups; transferase activity				Ile234Thr, Tyr239Cys
PVIT_0008609.T1	462	82.47	FAMILY NOT NAMED	DNA binding; transferase activity, transferring acyl groups; transferase activity	Endonuclease/exonuclease/phosphatase			
PVIT_0008615.T1	91	95.60		nucleic acid binding; zinc ion binding; metal ion binding; DNA integration				
PVIT_0008616.T1	1375	0.00	AAA-FAMILY ATPASE	protein targeting to peroxisome; peroxisome organization; peroxisomal membrane; peroxisome; protein binding; ATP binding; ATPase activity, coupled	ATPase, AAA-type, conserved site; UBX domain; AAA+ ATPase domain; ATPase, AAA-type, core; SEP domain; Peroxisome biogenesis factor 1, N-terminal, psi beta-barrel fold; Peroxisome biogenesis factor 1; P-loop containing nucleoside triphosphate hydrolase	PEX1; peroxin-1	Peroxisome	Pro179Ser, Asp392Gly, Glu540Asp, Arg763Lys
PVIT_0008617.T1	160	0.00	RECQ-MEDIATED GENOME INSTABILITY PROTEIN 2			RMI2; RecQ-mediated genome instability protein 2	Fanconi anemia pathway	
PVIT_0008618.T1	363	0.00	PROTEIN DISULFIDE ISOMERASE	integral component of membrane; isomerase activity; cell redox homeostasis; protein folding; endoplasmic reticulum; protein disulfide isomerase activity; response to endoplasmic reticulum stress	Thioredoxin, conserved site; Disulphide isomerase; Thioredoxin-like fold; Thioredoxin domain	TXNDC5, ERP46; thioredoxin domain-containing protein 5	Protein processing in endoplasmic reticulum	
PVIT_0008619.T1	83	0.00	FAMILY NOT NAMED	Mo-molybdopterin cofactor biosynthetic process; cell part	Sulfur carrier ThiS/MoaD-like; Molybdopterin converting factor, subunit 1; Beta-grasp domain superfamily; Molybdopterin synthase/thiamin biosynthesis sulphur carrier, beta-grasp	MOCS2A, CNXG; molybdopterin synthase sulfur carrier subunit	Sulfur relay system	
PVIT_0008620.T1	303	50.50	PEROXISOMAL MEMBRANE PROTEIN PEX13	membrane; integral component of membrane		PEX13; peroxin-13	Peroxisome	Val79Gly, Leu233Pro, Leu233Phe, Glu237Asp
PVIT_0010914.T1	419	0.00	SUBFAMILY NOT NAMED	integral component of membrane; transferase activity, transferring glycosyl groups	Glycosyltransferase, GlcNAc			
PVIT_0010915.T1	270	0.00	FAMILY NOT NAMED	hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds; nitrogen compound metabolic process	Carbon-nitrogen hydrolase	NTA1; protein N-terminal amidase [EC:3.5.1.-]		
PVIT_0010916.T1	218	92.66	SENTRIN/SUMO-SPECIFIC PROTEASE	proteolysis; cysteine-type peptidase activity	Ulp1 protease family, C-terminal catalytic domain	SENP8, NEDP1, DEN1; sentrin-specific protease 8 [EC:3.4.22.68]		
PVIT_0010917.T1	449	0.00		DNA binding; cytoplasm; nucleus	Ubiquitin interacting motif; Proteasomal ubiquitin receptor Rpn13/ADRM1	RPN13; 26S proteasome regulatory subunit N13	Proteasome;Epstein-Barr virus infection	
PVIT_0010918.T1	765	0.00	CATION EFFLUX PROTEIN/ ZINC TRANSPORTER	zinc ion transmembrane transporter activity; membrane; integral component of membrane; cation transmembrane transporter activity; cation transmembrane transport; Golgi apparatus; regulation of sequestering of zinc ion; cation transport; response to zinc ion; transmembrane transport	Cation efflux protein; Cation efflux transmembrane domain superfamily	SLC30A5_7, ZNT5_7, MTP, MSC2; solute carrier family 30 (zinc transporter), member 5/7		
PVIT_0010919.T1	137	0.00	DYNEIN LIGHT CHAIN ROADBLOCK-TYPE 2	microtubule-based movement; cytoplasmic dynein complex	Roadblock/LAMTOR2 domain	DYNLRB, DNCL2; dynein light chain roadblock-type		
PVIT_0010920.T1	570	0.00						
PVIT_0010921.T1	84	0.00						
PVIT_0010922.T1	203	0.00	HEPATOCELLULAR CARCINOMA-ASSOCIATED ANTIGEN	methyltransferase activity; transferase activity; methylation	Lysine methyltransferase	METTL21A; protein N-lysine methyltransferase METTL21A [EC:2.1.1.-]		
PVIT_0010923.T1	265	0.00						
PVIT_0010924.T1	388	0.00		thiol-dependent ubiquitin-specific protease activity; proteolysis; protein binding	PDZ domain; Ankyrin repeat-containing domain; Ankyrin repeat			
PVIT_0010925.T1	1071	10.92	FYVE-FINGER-CONTAINING RAB5 EFFECTOR PROTEIN RABENOSYN-5-RELATED	protein binding; metal ion binding	FYVE zinc finger; WW domain; Zinc finger, FYVE-related; Zinc finger, FYVE/PHD-type; Zinc finger, RING/FYVE/PHD-type; START-like domain superfamily			

PVIT_0010926.T1	993	27.39		calcium ion binding	EF-Hand 1, calcium-binding site; EF-hand domain; EF-hand domain pair			
PVIT_0010927.T1	1230	0.00	BACTERICIDAL PERMEABILITY-INCREASING BPI PROTEIN-RELATED	lipid binding; protein binding	PDZ domain; Bactericidal permeability-increasing protein, alpha/beta domain superfamily			
PVIT_0010928.T1	633	0.00	DNA REPAIR POLYMERASE UMUC / TRANSFERASE FAMILY MEMBER	damaged DNA binding; transferase activity; DNA repair	UmuC domain; DNA polymerase, Y-family, little finger domain	POLH; DNA polymerase eta [EC:2.7.7.7]	Platinum drug resistance;Fanconi anemia pathway	His163Tyr
PVIT_0010929.T1	1479	0.00	PATCHED-RELATED	lipid transporter activity; integral component of membrane; cell division; translational initiation; lipid transport; translation initiation factor activity	Sterol-sensing domain; Protein patched/dispatched	NPC1; Niemann-Pick C1 protein	Lysosome;Cholesterol metabolism	Ile297Met, Leu402Met, Glu1211Asp
PVIT_0010931.T1	417	68.35	SERINE/ARGININE RICH SPLICING FACTOR	nucleic acid binding; nucleotide binding	RNA recognition motif domain; Nucleotide-binding alpha-beta plait domain superfamily			Asp229Gly, Leu284Arg, Ser294Gly, Ser303Gly, Thr306Ser, Gly310Ser, Gly325Val
PVIT_0010932.T1	390	77.69	PHOSPHODIESTERASE HL		ADP-ribosylation factor-like 2-binding protein, domain			Gly3Asp, Leu5Phe, Ser52Thr, His159Gln, Lys219Glu, premature start
PVIT_0010933.T1	479	33.19		nucleic acid binding; organic cyclic compound binding; nucleotide binding; cellular metabolic process; heterocyclic compound binding; intracellular; catalytic activity	HRDC domain; HRDC-like superfamily			Asp130Asn, Lys176Arg, Leu165Ile, Val249Ile, Thr250Ser
PVIT_0010934.T1	151	86.75		nucleic acid binding; DNA integration				
PVIT_0010935.T1	161	60.87		nucleic acid binding; DNA binding; zinc ion binding; metal ion binding; DNA integration				
PVIT_0010940.T1	76	75.00						Gly82Cys
PVIT_0027771.T1	444	0.00	MITOGEN-ACTIVATED PROTEIN KINASE	protein kinase activity; protein phosphorylation; nucleotide binding; ATP binding; phosphorylation	Protein kinase, ATP binding site; Protein kinase domain; Protein kinase-like domain superfamily			
PVIT_0027772.T1	842	0.00	SPERMATOGENESIS-ASSOCIATED PROTEIN 5	ATP binding	ATPase, AAA-type, conserved site; AAA+ ATPase domain; P-loop containing nucleoside triphosphate hydrolase	VCP, CDC48; transitional endoplasmic reticulum ATPase	Protein processing in endoplasmic reticulum;Legionellosis	

Table S3 Functional annotation of all candidate genes in the mating-type locus of *Plasmopara viticola* . Related to Figure 3; Table 1.
Length: length of the protein sequence. Coverage TE: % of the protein sequence covered by a known transposable element sequence from the RepBase database.
Non-syn mutations: Non-synonymous mutations caused by alternate alleles (i.e. present in the MAT-b allele) for SNPs associated with the mating-type phenotype.

Plasmopara viticola		*Phytophthora infestans*		*Phytophthora capsici*		*Bremia lactucae*		*Plasmopara halstedii*	
Gene	Genomic coordinates	Gene	Genomic coordinates	Gene	Genomic coordinates	Gene	Genomic coordinates	Gene	Genomic coordinates
PVIT_0008615.T1	Plvit020 : 268538-268813	-	-	-	-	-	-	S037g16879	S037 : 514445-514675
PVIT_0008616.T1	Plvit020 : 284031-290118	PITG_06660T0	supercont1.8 : 3555056-3559199	-	-	TDH71904.1	SHOA01000004.1 : 3465462-3469694	S058g22075	S058 : 253289-257441
PVIT_0008617.T1	Plvit020 : 290201-290843	PITG_06658T0	supercont1.8 : 3553223-3553655	113615	scaffold_24 : 780134-780497	-	-	S058g22076	S058 : 257449-258211
PVIT_0008618.T1	Plvit020 : 290981-292072	PITG_06657T0	supercont1.8 : 3552016-3553107	507114	scaffold_24 : 778821-779936	TDH71905.1	SHOA01000004.1 : 3471347-3472438	S058g22077	S058 : 258213-259360
PVIT_0008619.T1	Plvit020 : 292484-292735	-	-	507113	scaffold_24 : 767579-767839	TDH72680.1	SHOA01000004.1 : 3312387-3312638	S058g22080	S058 : 270145-270396
PVIT_0008620.T1	Plvit020 : 292778-293875	PITG_13927T0	supercont1.30 : 1209885-1210997			TDH72681.1	SHOA01000004.1 : 3311222-3312707	S058g22081	S058 : 270418-271480
PVIT_0010914.T1	Plvit030 : 664127-665563	PITG_19514T0	supercont1.113 : 42045-43651	131769	scaffold_112 : 15063-16082	TDH72556.1	SHOA01000004.1 : 2470640-2472250	S038g16979	S038 : 178602-180207
PVIT_0010915.T1	Plvit030 : 666787-667975	-	-	546543	scaffold_21 : 676424-677113	TDH72560.1	SHOA01000004.1 : 2432641-2433906	S038g16969	S038 : 159428-160444
PVIT_0010916.T1	Plvit030 : 668239-668997	PITG_18649T0	supercont1.71 : 380428-381093	527673	scaffold_21 : 675378-676040	TDH72113.1	SHOA01000004.1 : 2674193-2674840	S038g16968	S038 : 158665-159334
PVIT_0010917.T1	Plvit030 : 669103-671428	PITG_18650T0	supercont1.71 : 381201-383324	-	-	TDH72114.1	SHOA01000004.1 : 2672938-2674102	S038g16967	S038 : 157440-158663
PVIT_0010918.T1	Plvit030 : 671952-674249	PITG_18652T0	supercont1.71 : 386886-388831	-	-	TDH72111.1	SHOA01000004.1 : 2668398-2670467	S038g16964	S038 : 151191-153602
PVIT_0010919.T1	Plvit030 : 674409-674822	PITG_18659T0	supercont1.71 : 509731-510138	566521	scaffold_21 : 636972-637277	-	-	S038g16963	S038 : 150556-150984
PVIT_0010920.T1	Plvit030 : 674890-677139	PITG_18658T0	supercont1.71 : 508361-509646	112024	scaffold_21 : 637351-638642	-	-	S038g16962	S038 : 148198-150509
PVIT_0010922.T1	Plvit030 : 678416-679027	PITG_18657T0	supercont1.71 : 505463-506107	16685	scaffold_21 : 639868-641344	TDH72357.1	SHOA01000004.1 : 3035742-3036617	S038g16961	S038 : 147238-147907
PVIT_0010924.T1	Plvit030 : 698965-700465	PITG_18655T0	supercont1.71 : 499107-500480	566524	scaffold_21 : 646246-647536	TDH71914.1	SHOA01000004.1 : 3038525-3042607	-	-
PVIT_0010925.T1	Plvit030 : 700575-703790	PITG_18656T0	supercont1.71 : 500643-504444	527667	scaffold_21 : 642758-646105	TDH72358.1	SHOA01000004.1 : 3038038-3041208	S038g16945	S038 : 117114-120572
PVIT_0010926.T1	Plvit030 : 711559-714862	PITG_18644T0	supercont1.71 : 200175-203646	534233	scaffold_21 : 628164-630255	-	-	-	-
PVIT_0010927.T1	Plvit030 : 715181-719081	PITG_18641T0	supercont1.71 : 143372-147280	-	-	-	-	S058g22006	S058 : 116854-120892
PVIT_0010928.T1	Plvit030 : 719171-721298	PITG_18642T0	supercont1.71 : 147405-149483	111820	scaffold_21 : 622078-623521	TDH72635.1	SHOA01000004.1 : 3068903-3071123	S058g21970	S058 : 58220-61000
PVIT_0010929.T1	Plvit030 : 722330-726769	PITG_21540T0	supercont1.436 : 41521-43864	506722	scaffold_21 : 616552-620039	TDH72218.1	SHOA01000004.1 : 3229267-3233712	S058g21968	S058 : 52165-56527
PVIT_0010931.T1	Plvit030 : 739521-740979	PITG_06695T0	supercont1.8 : 3784993-3786396	506740	scaffold_21 : 839844-841191	TDH72050.1	SHOA01000004.1 : 3516872-3518268	S038g16929	S038 : 79475-80755
PVIT_0010932.T1	Plvit030 : 740980-742486	PITG_06694T0	supercont1.8 : 3783665-3784865	112033	scaffold_21 : 841340-842455	-	-	S038g16928	S038 : 77507-79384
PVIT_0010933.T1	Plvit030 : 749107-750546	PITG_06702T0	supercont1.8 : 3915861-3917400	527674	scaffold_21 : 678227-679333	TDH72557.1	SHOA01000004.1 : 2434271-2435659	S038g16937	S038 : 98939-100506
PVIT_0027771.T1	Plvit600 : 27269-28831	PITG_06701T0	supercont1.8 : 3914139-3915718	527675	scaffold_21 : 679468-681021	TDH72558.1	SHOA01000004.1 : 2435764-2437367	S038g16935	S038 : 94334-95948
PVIT_0027772.T1	Plvit600 : 28900-31757	PITG_06700T0	supercont1.8 : 3912033-3914071	-	-	-	-	S038g16936	S038 : 95950-98826

Table S4 Orthologs of the *Plasmopara viticola* mating-type genes in the genomes of *Phytophthora infestans* , *Phytophthora capsici* , *Bremia lactucae* and *Plasmopara halstedii.* Related to Figure 5.

Genomic coordinates are indicated as scaffold_ID : start_position-end_position, with 1-based coordinates. Only *Pl. viticola* candidate proteins with a reciprocal best hit in at least one species are represented. For *Ph. capsici* , transcript IDs, and not full gene names, are indicated.