Identifying the genetic architecture of *Plamopara viticola* traits of interest by genome-wide association studies: case study of mating-type locus and effectors genes involved in the breakdown of grapevine Rpv3 resistance

Delmotte[1]*, F., Paineau[1], M., Mazet[1], I. D., Delière[1], L., Fabre[1], F., and Y. Dussert[1]

[1]INRAE, Bordeaux Sciences Agro, SAVE, ISVV, F-33140, Villenave d'Ornon, France

1 Introduction

Downy mildew, caused by *Plasmopara viticola*, is a major disease of grapevines that originates from North America where it infects wild *Vitis* species. Introduced in Bordeaux (France) accidently at the end of the 19th century, the grapevine downy mildew disease rapidly spread on *Vitis vinifera* which is highly sensitive to the pathogen (Fontaine et al. 2021).

Genomic resources for *Pl. viticola* are required to improve our understanding of the life-cycle of the pathogen, together with the genetic mechanisms underlying grape–pathogen interactions. Recent years have seen the publication of various transcriptomes (Mestre et al. 2016) and draft genome sequences based on short-read sequencing (Dussert et al. 2016; Yin et al. 2017; Brilli et al. 2018). Unfortunately, these assemblies have low completeness scores and remain fragmented.

Dussert et al. (2019) have sequenced the genome of *Pl. viticola* with PacBio long reads. They obtained a new 92.94 Mb assembly with high contiguity (359 scaffolds for a N50 of 706.5 kb) due to a better resolution of repeat regions. This assembly presented a high level of gene completeness, recovering 1,592 genes encoding secreted proteins involved in plant–pathogen interactions (Dussert et al. 2019) (Figure 1). *Pl. viticola* had a two-speed genome architecture, with secreted protein-encoding genes preferentially located in gene-sparse, repeat-rich regions and evolving rapidly, as indicated by pairwise dN/dS values. The *Pl. viticola* genome assembly generated here has allowed the development of robust population genomics approaches for investigating phenotypic traits or adaption to biotic and abiotic selective pressures in this species.

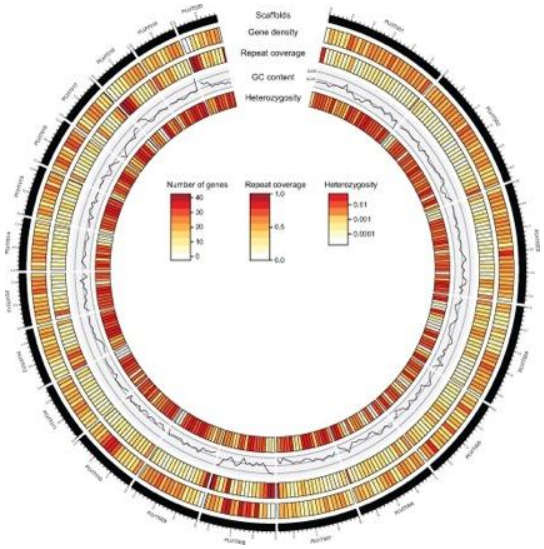

Figure 1: Genomic architecture of *Plasmopara viticola*. Gene density (number of genes per window), repeat coverage (proportion of the window covered by repeated sequences), GC content and the percentage of heterozygous sites are represented for genomic windows of 100 kb for the 20 largest scaffolds of the assembly (32.1 Mb).

The availability of a high-quality reference assembly for the Pl. viticola genome paves the way for robust population genetic studies based on genome-wide diversity (i.e., by resequencing individuals). This incldes population genome scans to detect genes under selection and genome-wide association studies to link genotype and phenotype variations and identify the genomic architecture of pathogen life-history traits of agronomical interest.

2 Identification of the mating-type locus using GWAS.

The life cycle of *Pl. viticola* includes an asexual multiplication phase during the spring and summer and a sexual reproduction event in the fall, generating the thick-walled sexual spores required for overwintering. In *Pl. viticol*a, only two mating types have been observed, and mating can occur only between diploid individuals of different mating types. 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*.

We determined the mating type of 54 diploid individuals of *Pl. viticola* collected across Europe, 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. 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. 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

*Corresponding author: francois.delmotte@inrae.fr

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

structure analyses based on principal component analysis or clustering analysis applied to a dataset filtered for SNPs in close linkage disequilibrium (LD). Conditions were therefore favorable for genetic-phenotype association studies.

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, 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, 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 composed exclusively of tandem repeat arrays covering 327 kb.

We therefore concluded that the mating-type locus was at least 570 kb long. 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 3). We found

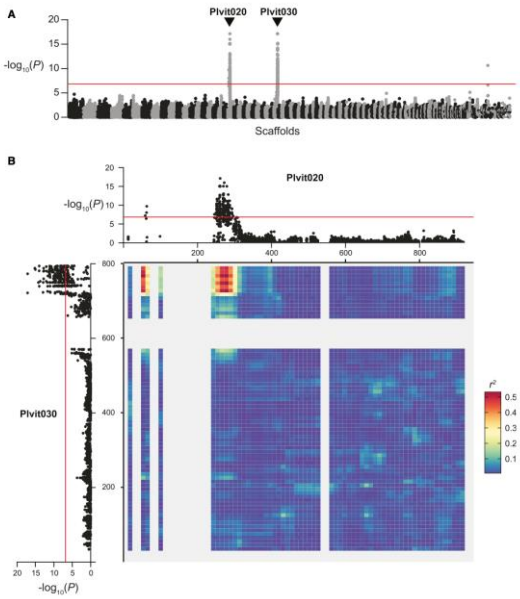

Figure 2: (A) Manhattan plot of the association p values between mating type and SNPs along the Pl. viticola genome. Alternating black and gray blocks of dots mark the limits between scaffolds. The two scaffolds with a significant association signal (Plvit020 and Plvit030) are indicated with arrows. (B) Manhattan plots of association p values between mating type and SNPs for Plvit020 and Plvit030. Linkage disequilibrium between the two scaffolds is represented as a heatmap, with redder colors representing higher r^2 values and therefore lower recombination rates.

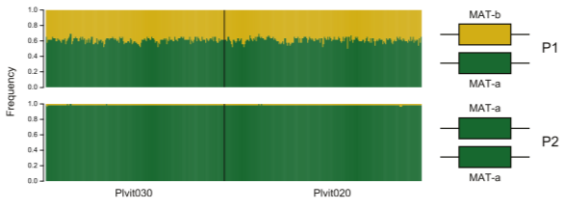

Figure 3: allele frequencies at SNPs associated with the mating-type phenotype in P1 individuals (top) and P2 individuals (bottom). Only results for the reference- based SNP calling approach are shown. Right panel: proposed model for mating-type determination is shown, with homozygosity in P2 individuals 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.

no 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, again consistent with a lack of recombination in this region.

The mating-type region included a total of 40 predicted coding sequences, 26 of which had predicted functions and did not correspond to TEs. 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 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.

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 identification of the mating type locus may guide the development of innovative control methods based on disruption of the sexual cycle of the pathogen. 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.

3 GWAS for the identification of candidate genes involved in the breakdown of grapevine major resistance factors. Several cases of resistance breakdowns have been reported in Europe for the Rpv3.1 resistance factor present in grapevine disease-resistant varieties such as Bianca and Regent (Peressotti et al., 2010; Delmotte et al., 2014; Delmas et al., 2016). The breakdown of resistance observed in the European *Pl. viticola* population could results from a modification of the effector repertoire of the pathogen. Oomycetes indeed secrete effector proteins, which modify host metabolism and defenses for the benefit of the pathogen. RxLR and CRN effectors are the most widely studied class of cytoplasmic effectors of oomycetes. Candidate effectors have been described for *Pl. viticola* (Mestre et al. 2016; Xiang et al. 2017) and several genes encoding RxLR effectors have been

functionally characterized (Xiang et al. 2017; Liu et al. 2018; Ma et al. 2021). To date, no avirulence genes has been described in *Pl. viticola* and the role of effector proteins remains unclear in the adaptation of the pathogen populations to grapevine genetic resistances.

Using the same successful GWAS approach that was used for the identification of the mating-type locus, we have studied the genomic architecture underlying the emergence of *Pl. viticola* Rpv3-resistance breaking strains. We collected 136 *Pl. viticola* isolates in 2018 in Bordeaux, France (INRAE experimental vineyard in Bordeaux). A total of 31 isolates were collected on from the resistant variety 'Regent', 55 isolates from the resistant variety 'Artaban', and 50 isolates were sampled from susceptible *V. vinifera* cv. Cabernet Sauvignon. All the isolates were single spored.

For each of the 136 strains we assessed the area of leaf disc covered by sporulation 5 days after inoculation, on both Cabernet sauvignon and Regent. All the strain were propagated and the DNA was extracted. The genome of each stain was fully sequencing by paired-end illumina sequencing (mean coverage 20X). Using the INRA-Pv221 *Pl. viticola* genome assembly as a reference, reads were mapped, and polymorphisms were detected. This allowed to retain around 1.8 M SNPs across the 359 scaffolds of the genome. First analyses of these phenotypic and genotypic data showed that a region of several kb of the *Pl. viticola* genome is highly associated to the level of sporulation of *Pl. viticola* strains on Regent. Analyses are currently underway to further investigate this promising result and describe the mutations or chromosomal rearrangements associated to the emergence of the Rpv3-resistance breaking strains of *Pl. viticola*.

References

1. Brilli, M., Asquini, E., Moser, M., Bianchedi, P.L., Perazzolli, M. and Si-Ammour, A., 2018. A multi-omics study of the grapevine-downy mildew (Plasmopara viticola) pathosystem unveils a complex protein coding- and noncoding-based arms race during infection. Scientific Reports, 8(1), pp.1-12.
2. 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.
3. Delmotte, F., Mestre, P., Schneider, C., Kassemeyer, H.-H., Kozma, P., Richart-Cervera, S., Rouxel, M., and Delière, L. 2014. Rapid and multiregional adaptation to host partial resistance in a plant pathogenic oomycete: Evidence from European populations of Plasmopara viticola, the causal agent of grapevine downy mildew. *Infect. Genet. Evol.* 27, 500–508.
4. 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.
5. 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 biology and evolution, 11(3), pp.954-969.
6. Dussert, Y., Legrand, L., Mazet, I.D., Couture, C., Piron, M.-C., Serre, R.-F., Bouchez, O., Mestre, P., Toffolatti, S.L., Giraud, T., and Delmotte, F. 2020. Identification of the first oomycete mating-type locus sequence in the grapevine downy mildew pathogen, *Plasmopara viticola*. Curr. Biol., 30 (2020), pp. 3897-3907.e4.
7. Fontaine, M.C., Labbé, F., Dussert, Y., Delière, L., Richart-Cervera, S., Giraud, T. and Delmotte, F., 2021. Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, Plasmopara viticola. Current Biology, 31(10), pp.2155-2166.
8. Ma, T., Chen, S., Liu, J., Fu, P., Wu, W., Song, S., Gao, Y., Ye, W. and Lu, J., 2021. *Plasmopara viticola* effector PvRXLR111 stabilizes VvWRKY40 to promote virulence. Molecular Plant Pathology, 22(2), pp.231-242.
9. 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.
10. 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.
11. Liu, Y., Lan, X., Song, S., Yin, L., Dry, I.B., Qu, J., Xiang, J. and Lu, J., 2018. In planta functional analysis and subcellular localization of the oomycete pathogen Plasmopara viticola candidate RXLR effector repertoire. Frontiers in Plant Science, 9, p.286.
12. Yin, L., An, Y., Qu, J., Li, X., Zhang, Y., Dry, I., Wu, H. and Lu, J., 2017. Genome sequence of Plasmopara viticola and insight into the pathogenic mechanism. Scientific Reports, 7(1), pp.1-12.
13. Xiang, J., Li, X., Yin, L., Liu, Y., Zhang, Y., Qu, J. and Lu, J., 2017. A candidate RxLR effector from Plasmopara viticola can elicit immune responses in Nicotiana benthamiana. BMC plant biology, 17(1), pp.1-14.