A point mutation and large deletion at the candidate avirulence locus AvrMlp7 in the poplar rust fungus correlate with poplar RMlp7 resistance breakdown

Clémentine Louet, Méline Saubin, Axelle Andrieux, Antoine Persoons, Mathilde Gorse, Jérémy Pétrowski, Bénédicte Fabre, Stéphane de Mita, Sébastien Duplessis, Pascal Frey, et al.

DR. SEBASTIEN DUPLESSIS (Orcid ID : 0000-0002-2072-2989)

DR. FABIEN HALKETT (Orcid ID : 0000-0001-8856-0501)

Article type : Special Issue

A point mutation and large deletion at the candidate avirulence locus *AvrMlp7* in the poplar rust fungus correlate with poplar RMlp7 resistance breakdown

Authors:

Clémentine LOUET*, Méline SAUBIN*, Axelle ANDRIEUX*, Antoine PERSOONS*, Mathilde GORSE*†, Jérémy PETROWSKI*, Bénédicte FABRE*, Stéphane DE MITA*‡, Sébastien DUPLESSIS*, Pascal FREY* and Fabien HALKETT*

Affiliations:

*Université de Lorraine, INRAE, IAM, F-54000 Nancy, France

†Present address: INRAE, AgroParisTech, Université Paris-Saclay, BIOGER, F-78850 Thiverval-Grignon, France

‡Present address: INRAE, Cirad, IRD, Montpellier SupAgro, Université de Montpellier, PHIM, F-34000 Montpellier, France

Corresponding author:

Fabien Halkett

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.16294

INRAE Centre Grand-Est – Nancy, UMR 1136 Interactions Arbres-Microorganismes, F-54280, Champenoux, France

E-mail: fabien.halkett@inrae.fr

Running title: Two genome alterations explain virulence

ABSTRACT

The deployment of plant varieties carrying resistance genes (*R*) exerts strong selection pressure on pathogen populations. Rapidly evolving avirulence genes (*Avr*) allow pathogens to escape *R*-mediated plant immunity through a variety of mechanisms, leading to virulence. The poplar rust fungus *Melampsora larici-populina* is a damaging pathogen of poplars in Europe. It underwent a major adaptive event in 1994, with the breakdown of the poplar *RMlp7* resistance gene. Population genomics studies identified a locus in the genome of *M. larici-populina* that likely corresponds to the candidate avirulence gene *AvrMlp7*. Here, to further characterize this effector, we used a population genetics approach on a comprehensive set of 281 individuals recovered throughout a 28-year period encompassing the resistance breakdown event. Using two dedicated molecular tools, genotyping at the candidate locus highlighted two different alterations of a predominant allele found mainly before the resistance breakdown: a non-synonymous mutation and a complete deletion of this locus. This results in six diploid genotypes: three genotypes related to the avirulent phenotype and three related to the virulent phenotype. The temporal survey of the candidate locus revealed that both alterations were found in association during the resistance breakdown event. They pre-existed before the breakdown in a heterozygous state with the predominant allele cited above. Altogether, these results suggest that the association of both alterations at the candidate locus *AvrMlp7* drove the poplar rust adaptation to *RMlp7*-mediated immunity. This study demonstrates for the first time a case of adaptation from standing genetic variation in rust fungi during a qualitative resistance breakdown.

Keywords

molecular evolution, gene-for-gene interaction, fungal pathogen, effector

1. INTRODUCTION

Fungal plant pathogens are a major threat to food security and biodiversity worldwide (Savary et al. 2019). Therefore, it is essential to gain more insight into the evolutionary history of plant pathogens to better control these harmful microorganisms by knowing the limits of their adaptive potential (Zhan et al., 2015).

To colonize and exploit host plants, fungal pathogens express a specific set of genes encoding secreted proteins (Petre et al., 2020). Among these, pathogenicity factors (or effector proteins) modulate host defense responses. To control diseases, modern agriculture widely uses plants with qualitative resistances relying on major resistance genes (*R*) (Stukenbrock & McDonald, 2008). Typically, R proteins recognize effector proteins and trigger host defense responses (Kourelis & Van Der Hoorn, 2018). The genes encoding such recognized effectors are termed avirulence (*Avr*) genes, according to the gene-for-gene model (Flor, 1971) and the avirulent (Avr) allele is dominant over virulent (avr) allele(s) (Flor, 1956).

R–Avr interactions usually result in an all-or-nothing infection outcome and thus tend to cause strong selection pressure on pathogen populations, ultimately resulting in resistance breakdown (Möller & Stukenbrock, 2017; Scholthof, 2006). *R* genes deployed in genetically uniform monocultures are the main risk factor causing pathogen evolution towards resistance breakdown (McDonald & Linde, 2002). Such breakdowns may result from different types of alteration in *Avr* genes. The two most common alterations are non-synonymous SNPs (single nucleotide polymorphisms) mutating *Avr* genes, and partial or complete deletion of *Avr* genes causing their inactivation (Fouché et al., 2018; Frantzeskakis et al., 2020; Rouxel & Balesdent, 2017; Sánchez-Vallet et al., 2018). A single pathogen species can evolve *via* multiple routes to become virulent on a given host. For instance, in *Leptosphaeria maculans*, virulence towards the same *R*-mediated resistance can be conferred by different alterations at the same locus that coexist in pathogen populations (Daverdin et al., 2012; Rouxel & Balesdent, 2017).

Fundamental questions about the dynamics of alterations during a resistance breakdown remain. Two evolutionary scenarios can be considered. On the one hand, adaptation is driven by a *de novo* mutation: virulent individuals emerge rapidly when a new allele appears. On the other hand, adaptation results from standing genetic variation (Barrett & Schluter, 2008): all alleles already exist in the population, but must be combined properly to produce an adapted phenotype. A way to disentangle these two scenarios is to examine the temporal evolution of the alterations leading to resistance breakdown. Here, we tackled this issue in the poplar rust fungus, a plant pathogen subjected to enduring selection pressure exerted by the deployment of *R* genes in the environment (Xhaard et al., 2011).

The poplar rust fungus *Melampsora larici-populina* (Basidiomycota, Pucciniales) is a biotrophic pathogen. The poplar rust life cycle consists of clonal multiplication on poplars alternating with annual sexual

reproduction on larch (Duplessis et al., 2021). Commercial poplar cultivation is particularly susceptible to poplar rust because of the intensive and monoclonal poplar plantations carrying selected qualitative resistances (Gérard et al., 2006) that are overcome systematically by the pathogen (Pinon & Frey, 2005). The most damaging resistance breakdown occurred in 1994 when the rust resistance gene *RMlp7* (carried by, e.g., *Populus × interamericana* 'Beaupré') was overcome, leading to the invasion in Western Europe of virulent 7 (a7) individuals of *M. larici-populina* (Persoons et al., 2017; Xhaard et al., 2011). Only four years elapsed between detection of the first virulent a7 individuals in Belgium in 1994 and the spread of these individuals throughout France, causing damaging epidemics (Pinon & Frey, 2005; Pinon et al., 1998). Although planting of 'Beaupré' decreased, and indeed totally ceased by 2005, this cultivar still predominated in the landscape for around 20 years, maintaining a strong selection pressure on *M. larici-populina* populations. As a consequence, the genetic structure and distribution of poplar rust populations in France was strongly impacted (Persoons et al., 2017; Xhaard et al., 2011). Three genetic groups were identified in France: (*i*) an ancestral group called 'Fossil', containing only avirulent 7 (A7) individuals, which declined from 1994 and vanished in 1998; (*ii*) a contemporaneous genetic group called 'Cultivated', containing almost exclusively virulent a7 individuals, which emerged at the time of the resistance breakdown (1994), and (*iii*) a third genetic group called 'Wild', containing exclusively avirulent A7 individuals that predominates mostly in south-eastern France where pathogen populations are found mostly on wild poplars and are not affected by commercial poplar plantations (Xhaard et al., 2011; Persoons et al., 2017).

Population genomic analyses identified a unique genomic region under positive selection on chromosome 15 of the *M. larici-populina* 98AG31 reference genome (Persoons et al., 2021; Duplessis et al., 2011). This region harbors a candidate gene that may be associated with *RMlp7*-mediated recognition (Persoons et al., 2021) and thus likely corresponds to the avirulence gene *AvrMlp7*. Three SNPs targeting this gene are associated with the virulence phenotype a7 (Fig. 1), including a non-synonymous SNP at position 2.233.369 (leading to a serine to glycine amino acid substitution at position 81, hereafter abbreviated to $S^{81}G$). Because the A7 and a7 alleles were found predominantly in avirulent A7 or virulent a7 populations, respectively (Persoons et al., 2021), we denominated them as Avr (avirulent) and avr (virulent) alleles. Moreover, 11% of virulent individuals appeared to have no sequence coverage at the candidate locus, which would indicate that this region is deleted in comparison to the 98AG31 reference genome (Persoons et al., 2021). This points to a second putative alteration of the avirulent allele.

The aim of this study was to further elucidate whether evolution of the candidate gene is related to resistance breakdown, and which scenario holds for the evolution of allelic frequencies. For this purpose, we used a historical collection of 281 *M. larici-populina* individuals encompassing the resistance

breakdown event (i.e. individuals sampled between 1989 and 2016). We thus expanded to a greater extent the sampling used in Persoons et al. (2021), which focused on only four populations representative of the three genetic groups. In the following, we (*i*) investigated the occurrence of the two alterations at the candidate locus, (*ii*) assessed genotypes and tested the link with phenotypes, and (*iii*) examined the temporal evolution of allele frequencies in relation to the neutral population structure.

2. MATERIALS AND METHODS

2.1. *Melampsora larici-populina* individuals

Sampling comprised 354 individuals collected from 1989 to 2016, mainly in the Northeast part of France and Belgium (where the resistance breakdown occurred). These individuals were selected from a historical collection maintained at INRAE Grand-Est Nancy, France. This sampling included most individuals from the previous study of Persoons et al. (2017) and extended the time-period covered by the analysis (Table 1; Table S1, Supporting Information).

Individuals were phenotyped for their ability to overcome *RMlp7*-mediated recognition (Section 2.3) and were genotyped with 22 microsatellite loci (Section 2.4). A total of 73 individuals were excluded from further analyses either because they had missing data, they were clonemates or they appeared to be mixtures of genotypes. Of the remaining 281 individuals, 111 were original for this study.

2.2. Investigation and delineation of the deletion

The genomes of 66 individuals had previously been sequenced at high depth (80x on average) using Illumina technology, and mapped on the *M. larici-populina* 98AG31 reference genome version 2 (Persoons et al., 2021).

To detect structural variants and identify the boundaries of the deletion, we used GROM (Smith et al., 2017), which operates on paired-end short reads. As recommended by the authors, we aligned reads using the BWA (v.0.7.17) MEM command with default settings. We ran GROM (v.1.0.3) with the -M option (duplicate read filtering). We extracted from GROM's output all structural variants (insertions, deletions or inversions) falling into or overlapping the region located between 2.1 and 2.3 Mbp on chromosome 15. For each variant, we recorded whether these events overlapped with at least one coding exon of any gene, and/or the SNP_2233369, and we computed Pearson's correlation coefficient between the virulence status of each individual and whether it bore this variant.

To examine further the occurrence of the deletion, we computed the sequencing depth at the position of the SNP_2233369 on chromosome 15. In order to correct for differences in sequencing depth among individuals, we calculated the average depth with the surrounding region (2.2–2.3 Mbp). Variation in

relative sequencing depth was partitioned using *K*-means clustering as implemented in the R package stats (Cran, v.3.5.2).

2.3. Phenotype assessment

Qualitative virulence phenotypes (i.e. pathotypes) of all individuals were confirmed by inoculation of a poplar cultivar carrying the *RMlp7* resistance gene (*P. x interamericana* 'Beaupré') and of a universal cultivar susceptible to all poplar rust individuals (*P. × euramericana* 'Robusta'), as a positive control, as described in Barrès et al. (2008). Infected 'Robusta' leaf discs of each individual were frozen and stored at –20°C until genomic DNA isolation.

2.4. Genomic DNA isolation and genotyping

Genomic DNA was extracted from infected poplar leaf discs using the DNeasy96 DNA Plant Kit (Qiagen) as described in Maupetit et al. (2018). Genotyping was performed using 22 microsatellite markers (Xhaard et al., 2011) for genotype purity and to avoid potential clones. Genotypes were assigned to the three genetic groups 'Fossil', 'Cultivated' and 'Wild' using the software TESS (Chen et al., 2007) as described in Persoons et al. (2017). Individuals studied in Persoons et al. (2017) were used as controls (Table S1, Supporting Information). A total of 37 individuals displaying mixed membership were removed from further analyses. For population genetic analyses, individuals were pooled according to sampling location and year, resulting in 14 population samples, with at least eight individuals per sample (mean sample size of 17 individuals).

2.5. Analysis of genotype at the candidate locus

2.5.1. Copy number variation:

To detect deletion of the candidate locus, a TaqMan Copy Number Assay was performed using the software PRIMEREXPRESS (v.3.0.1, Applied Biosystems) to compare the copy number of the gene of interest against a reference gene with known copy number. Ultra-specific qPCR primers for the candidate gene (forward: 5'-CACTAACCCTAACCTAACCTAT-3'; reverse: 5'-TTAAGTGTTGAGTTGCCTAAC-3'; Fig. 1) and for the α-Tubulin gene (JGI protein ID no.73137; forward: 5'-CACTAACCCTAACCTAACCTAT-3'; reverse: 5'-TTAAGTGTTGAGTTGCCTAAC-3) were designed to amplify products of 73 bp and 64 bp, respectively. Two TaqMan fluorogenic probes with different fluorescent dye reporters were used: the first for the candidate gene (5'-FAM-CAACCACCTCGAACCCGCTTCCA-3') and the second for the α-Tubulin gene (5'-VIC-CCCGCTCGCTTCTTACGCCCC-3'). Multiplex assays were performed in technical triplicate in 384-well plates, along with negative controls in each plate, with a final volume of 13 µL containing the following reaction

mixture: 2.7 µL of nuclease-free water, 6.5 µL of 2X TaqMan Master Mix (Applied Biosystems), 0.13 µL of each primer at 30 µM, 0.13 µL of each probe at 10 µM and 3 µL of uncalibrated genomic DNA. Quantitative PCR runs were performed on a QuantStudio 6 Flex machine (Life Technologies) with the following cycle parameters: 10 min at 95°C, followed by 50 cycles of 15 s at 90°C and 1 min at 60°C. Assuming a signal intensity proportional to the quantity of gene copies, we considered the number of gene copies as the relative difference between the cycle threshold (Ct) of the candidate gene compared with the Ct of the α-*Tubulin* gene ($2^{-\Delta Ct}$ calculation, Livak & Schmittgen, 2001), present in four copies in the *M. larici-populina* genome (Zhao et al., 2014). Using *K*-means clustering as implemented in the R package STATS (Cran, v.3.5.2), individuals were assigned to three categories depending on the number of gene copies detected: two, one, or none. Attribution to one of these categories was checked using error bars representative of the standard deviation from the mean. If error bars overlapped between categories, individuals were re-tested or excluded from subsequent analyses.

2.5.2. Allelic discrimination at SNP_2233369:

To determine the genotype at SNP_2233369, TaqMan SNP Genotyping Assays were performed using primers and probes designed using PRIMEREXPRESS (v.3.0.1, Applied Biosystems). Two ultra-specific qPCR primers were designed to amplify a 130-bp product encompassing SNP_2233369 of the candidate gene (forward: 5'-ACGTGCTTTTGTGGAGCAGTT-3'; reverse: 5'-TCAGTATGCGTCCTGACTCGTAA-3'). Two fluorogenic TaqMan minor groove binder (MGB) probes were designed, with different fluorescent dye reporters. The FAM-labelled probeAvrMlp7 (5'-FAM-CTGGTGCCACCGTG-MGB-3') and the VIC-labelled probeavrMlp7 (5'-VIC-CTGGTGCCACTGTG-MGB-3') were designed to target specifically the avirulent and virulent alleles of the candidate gene, respectively (Fig. 1). Multiplex assays were performed in technical triplicate in 384-well plates, along with negative controls in each plate, with a final volume of 5 µL containing the following reaction mixture: 2.5 µL of 2X TaqMan Master Mix (Applied Biosystems), 0.25 µL of 20X Assay Working Stock (containing primers and both probes) and 2.25 µL of uncalibrated genomic DNA. Quantitative PCR runs were performed on a QuantStudio 6 Flex machine (Life Technologies) with the following cycle parameters: 30 s at 60°C, 10 min at 95°C, followed by 40 cycles of 15 s at 90°C and 1 min at 60°C and 30 s at 60°C to finish. Fluorescence data were obtained using the QuantStudio Real-Time PCR Software (v.1.3). The average of FAM and VIC fluorescences were normalized by the ROX basal fluorescence present in the 2X TaqMan Master Mix and compared with each other. Using *K*-means clustering as described above, individuals were assigned to four categories: homozygous Avr/Avr, heterozygous Avr/avr, homozygous avr/avr and undetermined. Attribution to one of these categories was

checked using error bars representative of the standard deviation from the mean. If error bars overlapped between categories, individuals were re-tested or excluded from subsequent analyses.

2.6. Analysis of genetic diversity at the candidate locus and neutral loci

The genetic diversity of temporal samples was analysed using the nonneutral candidate locus. Hardy-Weinberg equilibrium was tested using Fisher's exact tests as implemented in the R package GENEPOP (Rousset et al., 2020) on all individuals pooled into 14 populations according to the collection year and regardless of the initial location. To estimate the *P*-value for Fisher's exact test, a Markov chain algorithm was implemented using the following parameters: dememorization, 1000; number of batches, 200; iterations per batch, 5000. Using neutral loci (22 microsatellite loci), we also analysed basic population genetics statistics using the R package HIERFSTAT (Goudet & Jombart, 2015) and the population structure with a discriminant analysis of principal components (DAPC, Jombart et al., 2010) as implemented in the R package ADEGENET (Jombart, 2008; Jombart & Ahmed, 2011). In a second step, we grouped individuals according to their genotype at the candidate locus to examine further the link between neutral genetic variation and the locus under selection. These analyses are presented as supplementary information (Annex S1, Supporting Information).

2.7. Statistical analyses

A Fisher's exact test was used to test the null hypothesis of independence between the genotype at the candidate locus and the phenotype towards *RMlp7*-mediated resistance. This test was performed using the fisher.test R function with a two-sided alternative hypothesis; a *P*-value < 0.001 was considered significant. To evaluate the evolution of allele and genotype frequencies over years, a linear regression was performed using the software R; a *P*-value < 0.05 was considered significant.

3. RESULTS

3.1 A deletion occurs at the candidate locus

With GROM, we detected 339 structural variants in the 2.1–2.3 Mbp region of chromosome 15 using data from Persoons et al. (2021). Among these 339 variants, 202 affected the coding sequence of at least one gene, including 121 deletion or inversion variants affecting SNP_2233369. No insertions were found close to this SNP. Among the 121 variants spanning SNP_2233369, there were 52 large inversions (>1 Mbp and spanning the whole region) unlikely to have a functional effect. Among the remaining 69 deletions, only one displayed a strong positive correlation with virulence a7 phenotype (r=0.25). This deletion had a length of 30 kb ranging from position 2.230.644 to 2.261.286 (Fig. 1).

To investigate the occurrence of this deletion at the candidate gene, we examined the variation in sequencing depth at this locus on genomic sequences from 66 whole-genome sequenced individuals. Strong differences were observed (Table S2, Supporting Information). In particular, four individuals showed a sequencing depth lower than 1X both at the SNP_2233369 position and over the flanking region, suggesting the presence of a large deletion estimated between position 2.23 and 2.26 Mbp on chromosome 15 of the 98AG31 reference genome (Fig. 2A; Table S2, Supporting Information). Furthermore, these four individuals (two from 1994 and two from 1998) originated from the 'Cultivated' group and were phenotypically virulent toward the *RMlp7*-mediated resistance. However, the large heterogeneity of sequencing depth at the locus (from 0 to 153.4X) could be indicative of a deletion affecting only one chromosome (i.e. a deletion allele at the heterozygous state). To correct for overall differences in sequencing depth, we computed the average sequencing depth along a larger genomic region ranging from 2.2 to 2.3 Mbp for all other individuals. As illustrated in Fig. 2A, the weighted sequencing depth at the candidate gene *AvrMlp7* clustered in three patterns (for details see Table S2, Supporting Information): a reference pattern exhibiting a regular sequencing depth in comparison with flanking regions, as in the *M. larici-populina* 98AG31 reference genome (pattern A); a halved sequencing depth in comparison to flanking regions (pattern B); and a lack of reads in this region (pattern C). Altogether, these results identified one unique deletion at the candidate locus affecting either one or both chromosomes.

3.2. The candidate locus displays six possible genotypes

To determine the genotype at the candidate locus, two TaqMan discrimination assays were performed on 281 individuals, after validation on 66 whole-genome sequenced individuals (Table S2, Supporting Information). In a first step, we determined the occurrence of the deletion resulting in the null (Ø) allele. Following the relative signal intensity ($2^{-\Delta Ct}$ ratio), the copy number variation assay indicated that clustering in three categories (*K*max = 3) was the best partition (Fig. 2B). In addition, the three categories identified by the *K*-means partition perfectly matched the grouping into the three sequencing depths described above. Therefore, pattern A corresponds to individuals with two copies of the gene, pattern B corresponds to individuals with one copy of the gene and a null allele, and pattern C corresponds to individuals with two null alleles (denoted as Ø/Ø). In a second step, allelic discrimination at the SNP_2233369 locus was performed to determine the specific alleles at the one or two copies of the gene present: only the avirulent allele, only the virulent allele, or both alleles (i.e. heterozygous individuals) (Fig. S1, Supporting Information). By combining the results of the two assays, six genotypes were identified at the *AvrMlp7* locus: the homozygous genotypes Avr/Avr, avr/avr and Ø/Ø and the

heterozygous genotypes Avr/avr, Avr/Ø and avr/Ø. No significant departure from Hardy-Weinberg proportions was observed (*P*-values ranging from 0.11 to 1; Table S3, Supporting Information). To conclude, the three alleles can be found in random association at the genotype level.

3.3. Genotypes at the candidate locus are linked to virulence phenotypes

To evaluate whether the candidate avirulence gene is involved in *RMlp7*-mediated recognition, we compared the genotypic states at the candidate locus with the qualitative virulence phenotype for the 281 *M. larici-populina* individuals (Table 2; Table S1, Supporting Information). Overall, three genotypes (Avr/Avr, Avr/avr and Avr/Ø) are associated with avirulent A7 individuals and the other three genotypes (avr/avr, avr/Ø and Ø/Ø) are associated with virulent a7 individuals (Table 2), suggesting dominance of the avirulent allele and a gene-for-gene relationship scenario. Only five individuals (out of 281), i.e. 1.8% of analyzed samples, presented a discrepancy between the genotype and the expected phenotype according to the gene-for-gene relationship. The match between genotypes and expected virulence phenotypes was highly significant according to a Fisher's exact test; the probability of observing this result by chance was $P < 0.001$ (*P*-value = 0.0005). In conclusion, we observed a tight link between the genotype at the candidate avirulence locus and the virulence phenotype.

3.4. Both the virulent and the null alleles expanded rapidly at the time of resistance breakdown

To study the temporal evolution of the polymorphism at the candidate locus during the RMlp7 resistance breakdown event, we calculated allelic frequencies according to years and genetic groups (Fig. 3). In the 'Fossil' group, including only avirulent A7 individuals, we observed a predominance of the avirulent allele (75%) with no significant difference in frequency over time (*P*-value = 0.610). Notably, both the virulent and the null alleles were identified in the 'Fossil' group (8% and 17%, respectively), already five years before the breakdown of *RMlp7*-mediated resistance. In the 'Wild' group containing only avirulent A7 individuals, we also observed the predominance of the avirulent allele (69%). Conversely, in the 'Cultivated' group, emerging at the time of the resistance breakdown in 1994, the avirulent allele had almost disappeared (0.01%) in favor of the virulent (48%) and the null (51%) alleles. In addition, the frequency of the null allele increased significantly (*P*-value = 0.019) from 17% in 1994 to 82% in 2016, whereas the frequency of the virulent allele decreased significantly during this period (*P*-value = 0.012). Therefore, the allelic evolution of the candidate locus coincided with the change in neutral population structure of *M. larici-populina* (Annex S1, Supporting Information). To conclude, the RMlp7 resistance breakdown event is associated with a concomitant increase in frequency of both the virulent and the null alleles, and the null allele has taken over in recent years.

4. DISCUSSION

Our study has helped elucidate the complexity of a double molecular alteration in a diploid organism, the poplar rust fungus, *M. larici-populina*. We based our analyses on a comprehensive temporal sampling that was made possible thanks to a large historical collection covering 28 years of contemporaneous evolution encompassing the RMlp7 breakdown event. Three alleles were detected in a specific candidate for the avirulence gene *AvrMlp7*, the avirulent allele that predominated initially, and the virulent and null alleles that burst onto the scene at the time of resistance breakdown. The evolution of this candidate gene is related to the emergence of the virulent phenotype a7 and the change in genetic structure in poplar rust populations.

4.1. The candidate gene seems to follow a gene-for-gene relationship

Through our screening of 281 individuals of *M. larici-populina*, we found an almost perfect concordance between genotypes at the candidate locus and phenotypes towards the *RMlp7*-mediated resistance. Three genotypes (Avr/Avr, Avr/avr and Avr/Ø) are associated with avirulence and the three others (avr/avr, avr/Ø and Ø/Ø) with virulence. Assuming a gene-for-gene relationship (Flor, 1971), our results are consistent with the avirulent allele being dominant over other alterations conferring virulence, namely the virulent and the null alleles. However, we found five individuals out of the 281 studied that did not follow this phenotype-genotype match, suggesting the existence of other alterations not yet identified. Our analysis of candidate gene polymorphisms is based on the genomic variation identified from a limited number of genomes (Persoons et al. 2021). By enlarging the screening of individuals, we broadened the genomic background, hence revealing mismatches that would be indicative of other putative alterations occurring at a much-reduced frequency. A given avirulence gene can be altered in different ways (Rouxel & Balesdent, 2017). These additional genomic alterations may impact pathogen fitness more heavily, which could explain their reduced frequency, as observed for viruses (Fabre et al. 2012). Conversely, the two alterations identified here would be the least costly, and were therefore retained throughout the evolution of poplar rust populations during resistance breakdown.

Until now, only a handful of *Avr* genes have been characterized in rust fungi (Chen et al., 2017; Petre et al., 2014; Salcedo et al., 2017; Upadhyaya et al., 2021; Wu et al., 2020). The obligate biotrophy nature of rust fungi represents a challenge, which explains the limited number of functional studies (Lorrain et al., 2019). Some authors have managed to bypass these limitations with indirect characterization using progenies and genetic mapping, e.g. in the wheat stem rust *Puccinia graminis* f. sp. *tritici* (Zambino et al., 2000), in the pine rust *Cronartium quercuum* f. sp. *fusiforme* (Kubisiak et al., 2011) and, most recently, in

the wheat yellow rust *Puccinia striiformis* f. sp. *tritici* (Wang et al., 2018). Our study proved that population genetics can be an efficient method with which to identify and investigate a putative avirulence gene in diploid pathogens. A formal demonstration of avirulence function will require functional approaches targeting the avirulent allele.

4.2. Virulence towards RMlp7 results from two non-exclusive molecular mechanisms

Here, we identified two non-exclusive alterations at the candidate locus *AvrMlp7* within diploid genotypes. Generally, an allele resulting from non-synonymous SNP(s) presents an altered version of the avirulence Avr protein, allowing escape from recognition mediated by the corresponding *R* protein (Ve et al., 2013; Wang et al., 2007). Nevertheless, some virulent pathogens are characterized mainly by knockout alterations such as the complete loss of the avirulence gene. In the wheat stem rust fungus *P. graminis* f. sp. *tritici*, the genes *AvrSr35* and *AvrSr50* contain large insertions of 57 kbp and 26 kbp, respectively, that resulted from transposon-mediated disruptions (Chen et al., 2017; Salcedo et al., 2017). Similarly, a deletion of a large genomic region containing an avirulence gene was shown to be responsible for a gain in virulence in the wheat pathogen *Zymoseptoria tritici* (Hartmann et al., 2017) and the oilseed rape pathogen *L. maculans* (Gout et al., 2007). In our study, sequencing depth analyses highlighted that the two breakpoints of the deletion encompassing the candidate gene *AvrMlp7* exhibited the same motif, suggesting a single deletion event, which then spread widely within poplar rust populations.

In haploids, any alteration of the avirulence gene (e.g. a non-synonymous mutation (avr) or gene inactivation caused by a deletion (Ø)) determines the virulence phenotype in a straightforward relationship (Fig. 4A, B). For diploid pathogens, the relationship is more complex as alleles conferring virulence are usually recessive (Flor, 1956). In this case, the occurrence of an Avr allele (even in a heterozygous state) is sufficient to trigger plant resistance (Fig. 4C). It was thus already shown that a diploid pathogen can evolve through several alteration events leading to virulence towards an *R* gene. High levels of polymorphism in avirulence genes have also been reported in other rust fungi, as in the flax rust fungus *Melampsora lini,* where proteins AvrL567 and AvrM exhibit multiple allelic variants or a copy number variation with different alleles, directly impacting pathogenicity (Catanzariti et al., 2006; Dodds et al., 2004). Here, we demonstrated that any association of alleles can confer the virulence phenotype (Fig. 4D).

4.3. A point mutation and a large deletion were combined during resistance breakdown

Our study revealed the presence of the two alterations at least five years before RMlp7 resistance breakdown, but, once this resistance was overcome, their frequency increased rapidly. Therefore,

adaptation of *M. larici-populina* to *RMlp7*-mediated resistance resulted from selection from standing genetic variation (Barrett & Schluter, 2008). However, at least five years elapsed before the establishment of a favorable combination leading to virulence towards the *RMlp7*-mediated resistance. The dikaryotic state of rust fungi impedes the evolution of virulence, because of the dominance of the avirulent allele. Alleles conferring virulence are subjected mostly to drift as long as they are in the heterozygous state. In this respect, a recent modelling study on the dynamics of resistance breakdown highlighted that the diploid state leads to much more unpredictable evolutionary trajectories, with a delayed time for fixation of the virulence compared with haploids (Saubin et al., 2021). This phenomenon is further marked in pathogens such as *M. larici-populina* that require an alternate host to complete their life cycle. Indeed, obligate migration to resistant and susceptible hosts after sexual reproduction on the alternate host causes high rates of mortality of avirulent individuals, especially when the resistant host predominates in the landscape. As long as virulent allele(s) frequency is low, no genotypes missing the avirulent allele are produced, hence no phenotypically virulent individual exists. By contrast, once the resistance is overcome, the virulent allele(s) spread massively in the landscape (over both resistant and susceptible hosts) and the avirulent allele collapses rapidly, even in individuals living on susceptible hosts (Saubin et al., 2021). Such modeling work is in accordance with our observations. Here, the scenario is complicated by the occurrence of the virulent and null alleles, conferring virulence. The fact that they are not mutually exclusive likely fostered the pace of this breakdown event and contributed to the rapid emergence of the 'Cultivated' group. According to this scenario, the 'Cultivated' population might have been founded from a single event, as previously suggested (Persoons et al., 2017). Replacement of the 'Fossil' population (described in Persoons et al., 2017) coincided with the replacement of the avirulent allele by the two alleles conferring virulence.

4.4. Does the null allele confer a fitness advantage?

Whereas the two alleles conferring virulence exploded at the time of resistance breakdown, their relative frequency was not stable through time afterwards. We observed a significant increase in the frequency of one allele leading to the predominance of the null allele in contemporary populations. This variation can be neutral (resulting from genetic drift), but may also suggest a consequence on fitness for individuals carrying either allele. Genetic modifications leading to the acquisition of a new virulence may be accompanied by unfavorable effects (Leach et al., 2001). Indeed, the avirulence genes are involved primarily in various functions necessary for infection (Khan et al., 2018). Any alteration at the avirulence locus can thus translate into a reduction in fitness, the so-called fitness cost of virulence (van der Plank, 1963). No such cost has been reported so far in poplar rust; virulent individuals were found to display the

same level of disease-associated traits as avirulent individuals (Maupetit et al., 2021). However, this latter study did not consider the two types of alterations conferring virulence. The deletion may provide an advantage over point mutation mechanisms because of the absence of cost associated with the synthesis of the mutated protein. Examining the fitness advantage or disadvantage of each allele in more detail will require a dedicated study, monitoring several traits associated with completion of the infection process and life cycle.

5. Conclusion

Efforts dedicated to developing new analysis strategies for avirulence genes have focused mostly on haploid organisms (Bakkeren & Szabo, 2020). Here, we proposed a method combining population genetics and molecular approaches to investigate the candidate avirulence gene *AvrMlp7* in *M. larici-populina*, a diploid fungus responsible for poplar rust disease. We identified two alterations at the candidate locus: a single nucleotide mutation and a complete deletion. These two alleles co-occurred within pathogen populations and associated randomly within individuals. To our knowledge, this is the first demonstration that two alterations conferring virulence can combine within a diploid individual. Their association increased the chance of observing virulent phenotypes and was probably responsible for the breakdown of *RMlp7*-mediated resistance in 1994.

ACKNOWLEDGMENTS

We thank Jean Pinon for creating the historical poplar rust collection at INRAE Nancy in the 1980s. We warmly thank Fanny Hartmann for constructive comments on a previous version of the manuscript, and Benjamin Petre for proofreading an early version of manuscript, a thousand blessings to him. We also thank Tatiana Giraud and four anonymous reviewers for detailed comments and suggestions which helped improve the manuscript. Helen Rothnie corrected the English language. This work was supported by grants from the French National Research Agency (ANR-18-CE32-0001, Clonix2D project; ANR-17-CE32-0010, WABSARF project; ANR-11-LABX-0002-01, Cluster of Excellence ARBRE). Clémentine Louet was supported by a PhD fellowship from the Région Lorraine and the French National Research Agency (ANR-18-CE32-0001, Clonix2D project). Méline Saubin was supported by a PhD fellowship from INRAE and the French National Research Agency (ANR-18-CE32-0001, Clonix2D project). Antoine Persoons was supported by a PostDoc fellowship from the Region Grand-Est and the French National Research Agency (ANR-11-LABX-0002-01, Lab of Excellence ARBRE).

REFERENCES

Bakkeren G, Szabo LJ (2020) Progress on molecular genetics and manipulation of rust fungi. *Phytopathology*, **110**, 532–543

Barrès B, Halkett F, Dutech C *et al.* (2008) Genetic structure of the poplar rust fungus *Melampsora larici-populina*: evidence for isolation by distance in Europe and recent founder effects overseas. *Infection, Genetics and Evolution*, **8**, 577–587.

Barrett LG, Thrall, PH, Burdon JJ, Linde CC (2008). Life history determines genetic structure and evolutionary potential of host-parasite interactions. *Trends in Ecology and Evolution*, **23**, 678–685.

Catanzariti AM, Dodds PN, Lawrence GJ, Ayliffe MA, Ellis, JG (2006) Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *The Plant Cell*, **18**, 243–256.

Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes*, **7**, 747–756.

Chen J, Upadhyaya NM, Ortiz D *et al.* (2017) Loss of *AvrSr50* by somatic exchange in stem rust leads to virulence for *Sr50* resistance in wheat *Science*, **358**, 1607–1610.

Daverdin G, Rouxel T, Gout L *et al.* (2012) Genome structure and reproductive behaviour influence the evolutionary potential of a fungal phytopathogen. *PLoS Pathogens*, **8**, e1003020.

Dodds PN, Lawrence GJ, Catanzariti AM, Ayliffe MA, Ellis JG (2004) The *Melampsora lini AvrL567* avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *The Plant Cell*, **16**, 755–768.

Duplessis S, Lorrain C, Petre B *et al.* (2021) Host adaptation and virulence in heteroecious rust fungi. *Annual Review of Phytopathology*, **59**, 171-1720.

Flor HH (1956) The complementary genic systems in flax and flax rust. *Advances in Genetics*, **8**, 29–54.

Flor HH (1971) Current status of the gene-for-gene concept. *Annual Review of Phytopathology*, **9**, 275–296.

Fouché S, Plissonneau C, Croll D (2018) The birth and death of effectors in rapidly evolving filamentous pathogen genomes. *Current Opinion in Microbiology*, **46**, 34–42.

Frantzeskakis L, Di Pietro A, Rep M, *et al.* (2020) Rapid evolution in plant–microbe interactions – a

molecular genomics perspective. *New Phytologist*, **225**, 1134–1142.

Gérard PR, Husson C, Pinon J, Frey P (2006) Comparison of genetic and virulence diversity of *Melampsora larici-populina* populations on wild and cultivated poplar and influence of the alternate host. *Phytopathology*, **96**, 1027–1036.

Goudet J, Jombart T (2015) Hierfstat: estimation and tests of hierarchical F-statistics.

Gout L, Kuhn ML, Vincenot L *et al.* (2007) Genome structure impacts molecular evolution at the *AvrLm1* avirulence locus of the plant pathogen *Leptosphaeria maculans. Environmental Microbiology*, **9**, 2978–2992.

Hacquard S, Petre B, Frey P *et al.* (2011) The poplar-poplar rust interaction: insights from genomics and transcriptomics. *Journal of Pathogens*, **2011**.

Hartmann FE, Sánchez-Vallet A, McDonald BA, Croll D (2017) A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. *The ISME Journal*, **11**, 1189–1204.

Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jombart T, Ahmed I (2011) adegenet 13-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, **27**, 3070–3071.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.

Khan M, Seto D, Subramaniam R, Desveaux D (2018) Oh, the places they'll go! A survey of phytopathogen effectors and their host targets. *The Plant Journal*, **93**, 651–663.

Kourelis J, Van Der Hoorn RAL (2018) Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *The Plant Cell*, **30**, 285–299.

Kubisiak TL, Anderson CL, Amerson HV *et al.* (2011) A genomic map enriched for markers linked to *Avr1* in *Cronartium quercuum* fsp *fusiforme. Fungal Genetics and Biology*, **48**, 266–274.

Leach JE, Vera Cruz CM, Bai J, Leung H (2001) Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Annual Review of Phytopathology*, **39**, 187–224.

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods*, **25**, 402–408.

Lorrain C, Gonçalves dos Santos KC, Germain H, Hecker A, Duplessis S (2019) Advances in understanding obligate biotrophy in rust fungi. *New Phytologist*, **222**, 1190–1206.

Maupetit A, Larbat R, Pernaci M *et al.* (2018) Defense compounds rather than nutrient availability shape aggressiveness trait variation along a leaf maturity gradient in a biotrophic plant pathogen. *Frontiers in Plant Science*, **9**, 1396.

Maupetit A, Fabre B, Pétrowski J *et al.* (2021) Evolution of morphological but not aggressiveness-related traits following a major resistance breakdown in the poplar rust fungus, *Melampsora larici-populina*. *Evolutionary Applications*, **14**, 513–523.

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, **40**, 349–379.

Möller M, Stukenbrock EH (2017) Evolution and genome architecture in fungal plant pathogens. *Nature Reviews Microbiology*, **15**, 756–771.

Persoons A, Hayden KJ, Fabre B *et al.* (2017) The escalatory Red Queen: population extinction and replacement following arms race dynamics in poplar rust. *Molecular Ecology*, **26**, 1902–1918.

Persoons A, Maupetit A, Louet C *et al.* (2021) Genomic signatures of a major adaptive event in the pathogenic fungus *Melampsora larici-populina. BioRxiv* 2021.04.09.439223.

Petre B, Joly DL, Duplessis, S (2014) Effector proteins of rust fungi. *Frontiers in Plant Science*, **5**, 416.

Petre B, Lorrain C, Stukenbrock EH, Duplessis S (2020) Host-specialized transcriptome of plant-associated organisms. *Current Opinion in Plant Biology*, **56**, 81–88.

Pinon J, Frey P (2005) Interactions between poplar clones and *Melampsora* populations and their implications for breeding for durable resistance. In: *Rust diseases of willow and poplar* (eds Pei MH, McCraken AR), pp. 139–154. CAB International, Wallingford, UK.

Pinon J, Frey P, Husson C, Schipfer A (1998) Poplar rust (*Melampsora larici-populina*): the development of E4 pathotypes in France since 1994. In: *Proceedings of the First IUFRO Rusts of Forest Trees Conference*, Saariselkä, Finland (eds Jalkanen R, Crane PE, Walla JA, Aalto T), pp. 57–64. Finnish Forest Research Institute, Research Paper 712, Saarijïarvi.

Rousset F, Lopez J, Belkhir K (2020) genepop: population genetic data analysis using genepop.

Rouxel T, Balesdent MH (2017) Life, death and rebirth of avirulence effectors in a fungal pathogen of

Brassica crops, Leptosphaeria maculans. New Phytologist, **214**, 526–532.

Salcedo A, Rutter W, Wang S *et al.* (2017) Variation in the *AvrSr35* gene determines *Sr35* resistance against wheat stem rust race Ug99. *Science*, **358**, 1604–1606.

Sánchez-Vallet A, Fouché S, Fudal I *et al.* (2018) The genome biology of effector gene evolution in filamentous plant pathogens. *Annual Review of Phytopathology*, **56**, 21-210.

Saubin M, De Mita S, Zhu X, Sudret B, Halkett F (2021) Impact of ploidy and pathogen life cycle on resistance durability. *BioRxiv* 2021.05.28.446112.

Savary S, Willocquet L, Pethybridge SJ *et al.* (2019) The global burden of pathogens and pests on major food crops. *Nature Ecology and Evolution*, **3**, 430–439.

Scholthof KBG (2006) The disease triangle: pathogens, the environment and society. *Nature Reviews Microbiology*, **5**, 152–156.

Smith SD, Kawash JK, Grigoriev A (2017) Lightning-fast genome variant detection with GROM. *GigaScience*, **6**, 1–7.

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, **46**, 75–100.

Upadhyaya NM, Mago R, Panwar V *et al.* (2021) Genomics accelerated isolation of a new stem rust avirulence gene-wheat resistance gene pair. *Nature Plants*, **7**, 1220–1228.

van der Plank JE (1963). *Plant diseases: Epidemics and control*. Academic Press, New York, 349 p.

Ve T, Williams SJ, Catanzariti AM *et al.* (2013) Structures of the flax-rust effector AvrM reveal insights into the molecular basis of plant-cell entry and effector-triggered immunity. *PNAS*, **110**, 17594–17599.

Wang CIA, Guncar G, Forwood JK *et al.* (2007) Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity. *The Plant Cell*, **19**, 2898–2912.

Wang L, Zheng D, Zuo S *et al.* (2018) Inheritance and linkage of virulence genes in chinese predominant race CYR32 of the wheat stripe rust pathogen *Puccinia striiformis* f sp *tritici. Frontiers in Plant Science*, **9**, 120.

Wu JQ, Dong C, Song L, Park RF (2020) Long-read–based de novo genome assembly and comparative genomics of the wheat leaf rust pathogen *Puccinia triticina* identifies candidates for three avirulence

genes. *Frontiers in Genetics*, **11**, 521.

Xhaard C, Fabre B, Andrieux A *et al.* (2011) The genetic structure of the plant pathogenic fungus *Melampsora larici-populina* on its wild host is extensively impacted by host domestication. *Molecular Ecology*, **20**, 2739–2755.

Xu X (2012) Super-races are not likely to dominate a fungal population within a life time of a perennial crop plantation of cultivar mixtures: a simulation study. *BMC Ecology*, **12**, 1–10.

Zambino PJ, Kubelik AR, Szabo LJ (2000) Gene action and linkage of avirulence genes to DNA markers in the rust fungus *Puccinia graminis. Phytopathology*, **90**, 819–826.

Zhan J, Thrall PH, Papaïx J, Xie L, Burdon JJ (2015) Playing on a pathogen's weakness: using evolution to guide sustainable plant disease control strategies. *Annual Review of Phytopathology*, **53**, 19–43.

Zhao Z, Liu H, Luo Y *et al.* (2014) Molecular evolution and functional divergence of tubulin superfamily in the fungal tree of life. *Scientific Reports*, **4**, 6746.

DATA ACCESSIBILITY

Virulence phenotype towards the *RMlp7-mediated* resistance and genotype at the *AvrMlp7* locus as well as microsatellite data from the 281 individuals are available on Dryad repository: https://doi.org/10.5061/dryad.h70rxwdhr.

AUTHOR CONTRIBUTIONS

C.L., M.S., P.F. and F.H. conceived and designed the study. C.L. and A.A. designed and performed molecular experiments. M.S., J.P. and B.F. performed individual cultures and virulence type analyses. C.L., M.S., M.G. and F.H. performed the genotyping. S.D.M. and F.H. performed the alignment. C.L., M.S., S.D.M., A.P. and F.H. analyzed the data. C.L. and F.H. prepared the manuscript. All authors revised and approved the manuscript.

TABLE AND FIGURE LEGENDS

Table 1 *Melampsora larici-populina* individuals used in the study, with their phenotype towards the *RMlp7*-mediated resistance and their assignment to one of three genetic groups (Fossil, Cultivated, Wild).

Table 2 Genotypes identified at the candidate locus *AvrMlp7* in relation to phenotypes towards the *RMlp7* resistance gene.

Fig. 1 Genomic organization in the region containing the candidate gene *AvrMlp7*, based on the *Melampsora larici-populina* reference genome 98AG31. Synonymous single-nucleotide polymorphisms (SNPs) and non-synonymous SNPs are indicated by grey and red bars, respectively. Displayed SNPs occurred in at least 10% of genomes (i.e. minor allele frequency above 0.1). Genes and the candidate gene are indicated in grey and pink, respectively. The exact limits of the deletion are indicated by the grey dotted box. The vertical red dotted line highlights the non-synonymous SNP at position 2.233.369 on chromosome 15 ($C^{241}T$) leading to a $G^{81}S$ amino acid substitution in the predicted protein. Relative positions of primers and probes for the allelic discrimination at the SNP_2233369 (1) and testing the copy number variation (2) are shown in the enlargement of the candidate gene. The third exon nucleotide sequences of the candidate gene and the probe locations (ProbeavrMlp7 and ProbeAvrMlp7) used for allelic discrimination at SNP_2233369 are indicated for both virulent and avirulent alleles.

Fig. 2 Patterns of sequencing depth in the region containing the candidate locus *AvrMlp7* and amplification of the candidate gene in eight *Melampsora larici-populina* individuals. (A) Raw sequencing depth from eight individuals after mapping of the sequencing reads on chromosome 15 of the 98AG31 reference genome. The region impacted by the deletion was shortened to better illustrate the three distinct patterns of sequencing depth among individuals: pattern A with a regular sequencing depth in comparison with flanking regions, pattern B with a decreased sequencing depth in comparison with flanking regions and pattern C with a total lack of sequencing depth in comparison with flanking regions. The vertical red dotted lines encompass the exact position of the candidate gene and the horizontal red line represents the average depth between flanking regions of the candidate locus (2.2 to 2.3 Mbp) along chromosome 15. (B) Mean normalized amplification of the candidate gene from the same eight individuals. The colors represent the individual assignment by *K*-means clustering (*K*max = 3). Error bars indicate the standard deviation of the mean.

Fig. 3 Cumulative frequencies of alleles at the candidate locus *AvrMlp7* in *Melampsora larici-populina* individuals according to genetic group and sampling years. Numbers at the top indicate the number of individuals assayed. To balance sample sizes, the years 1990 and 1997 in Fossil group gather individuals

collected in 1989–1990 and 1996–1997–1998, respectively and the year 1997 in Cultivated group gather individuals collected in 1996–1997. Avr: avirulent allele; avr: virulent allele; Ø: null allele.

Fig. 4 Difference in molecular determinism of R-mediated resistance between haploid and diploid pathogens. For haploid pathogens, the outcome of the interaction between the pathogen's avirulence gene and the resistance gene (*R*) is straightforward: (A) the effector protein encoded by the avirulent allele (Avr) is recognized, directly or indirectly, by the *R* gene product, which triggers host resistance; (B) the virulent allele, e.g. a non-synonymous mutation (avr) or a deletion (Ø) of the avirulent allele leads to host susceptibility because of a lack of recognition by the *R* gene product. For diploid pathogens the interaction is complicated by the dominance of the avirulent allele: (C) a single haploid copy of the avirulent allele is sufficient for recognition by the *R* gene product and initiation of host resistance; (D) only the association of two recessive virulent alleles can lead to host susceptibility.

Table 1 *Melampsora larici-populina* individuals used in the study, with their phenotype towards the *RMlp7*-mediated resistance and their assignment to one of three genetic groups (Fossil, Cultivated, Wild).

	Phenotype[†]			Genetic group			
Year	A7	a7	No. of isolates	Fossil	Cultivated	Wild	Hybrid[‡]
1989	10	-	10	8	-	-	2
1990	5	-	5	5	-	-	-
1991	13	-	13	13	-	-	-
1992	24	-	24	23	-	-	1
1993	19	-	19	19	-	-	-
1994	11	28	39	9	29	1	-
1995	8	22	30	8	22	-	-
1996	3	2	5	1	2	1	1
1997	7	6	13	6	6	1	-
1998	2	22	24	1	21	1	1
1999	3	8	11	-	9	-	2
2001	1	12	13	-	11	2	-
2008	20	30	50	-	30	18	2
2016	2	23	25	-	25	-	-
Total	128	153	281	93	155	24	9

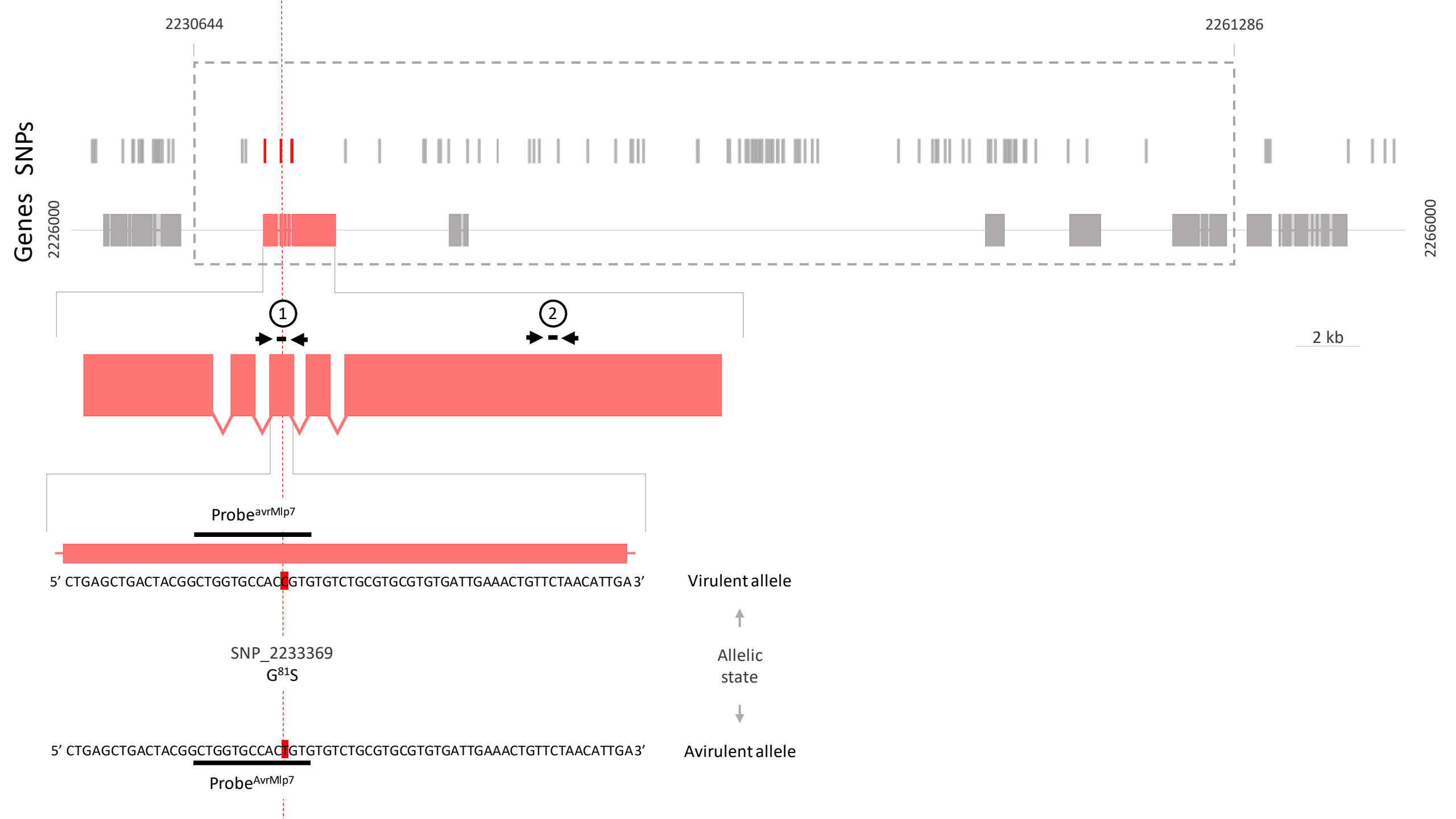
[†]A7, avirulent 7 phenotype towards *RMlp7*; a7, virulent 7 phenotype towards *RMlp7.*
[‡]Hydrids were considered as missing data for subsequent analyses.

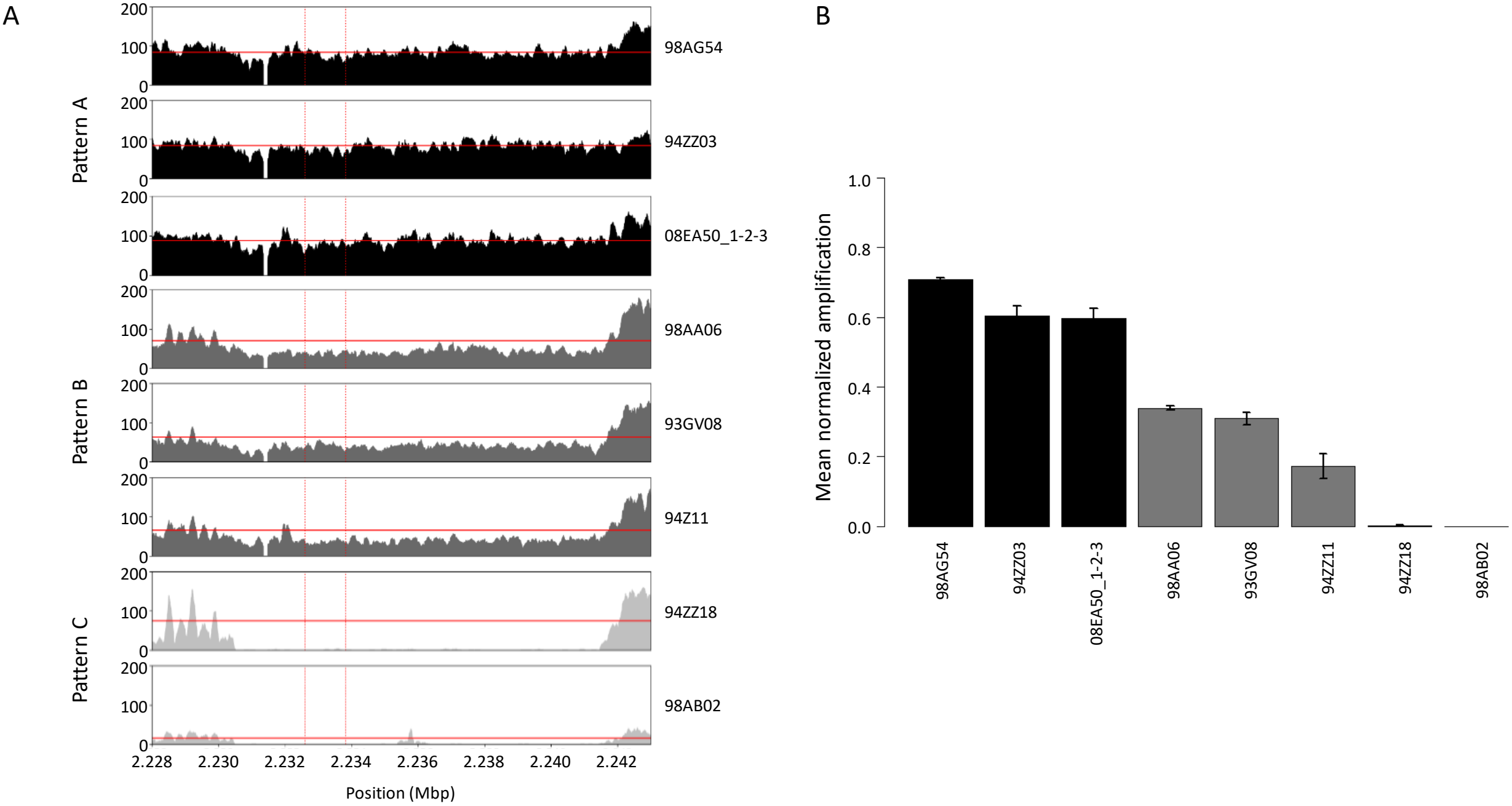
Table 2 Genotypes identified at the candidate locus *AvrMlp7* in relation to phenotypes towards the *RMlp7* resistance gene.

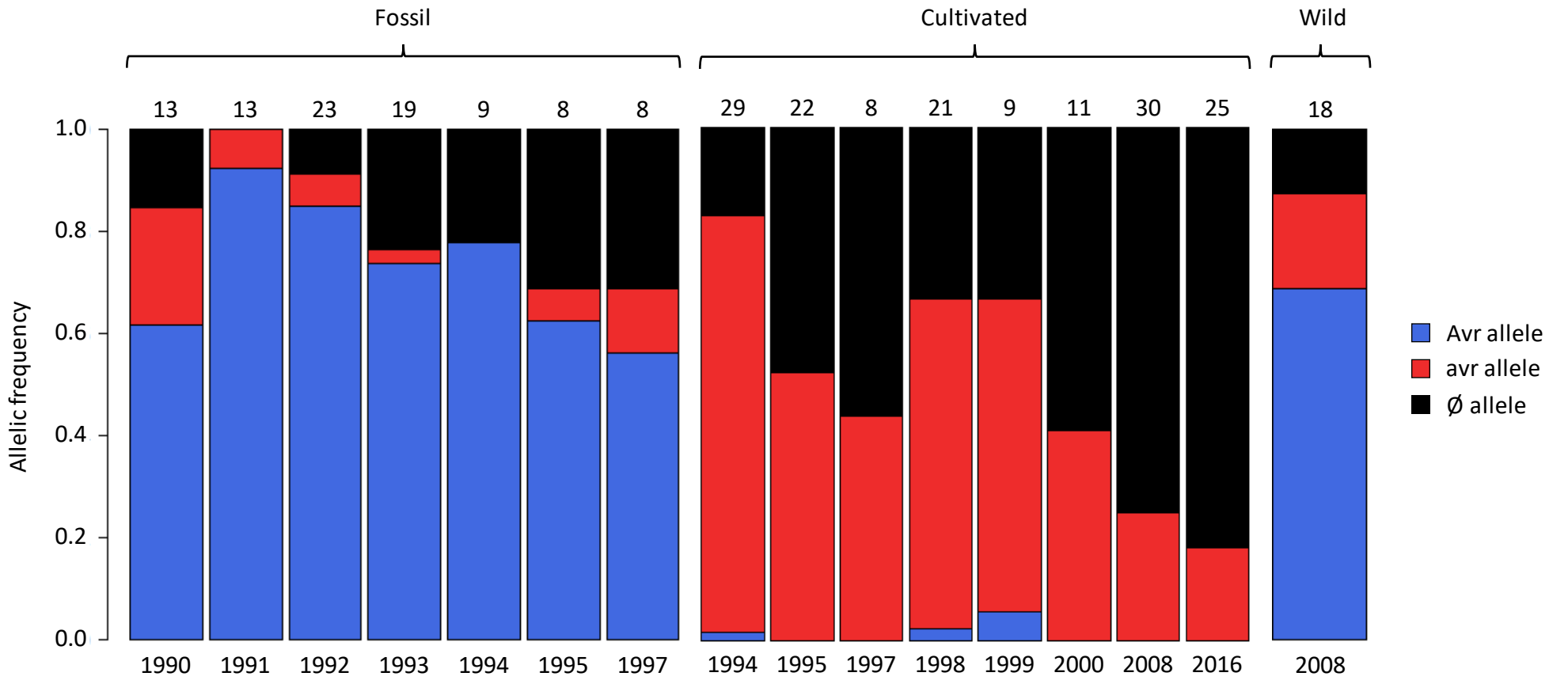
Phenotype† \ Genotype‡	Avr / Avr	Avr / avr	Avr / Ø	avr / avr	avr / Ø	Ø / Ø	Total
A7	64	27	33	-	1	3	128
a7	-	1	-	44	58	50	153
Total	64	28	33	44	59	53	281

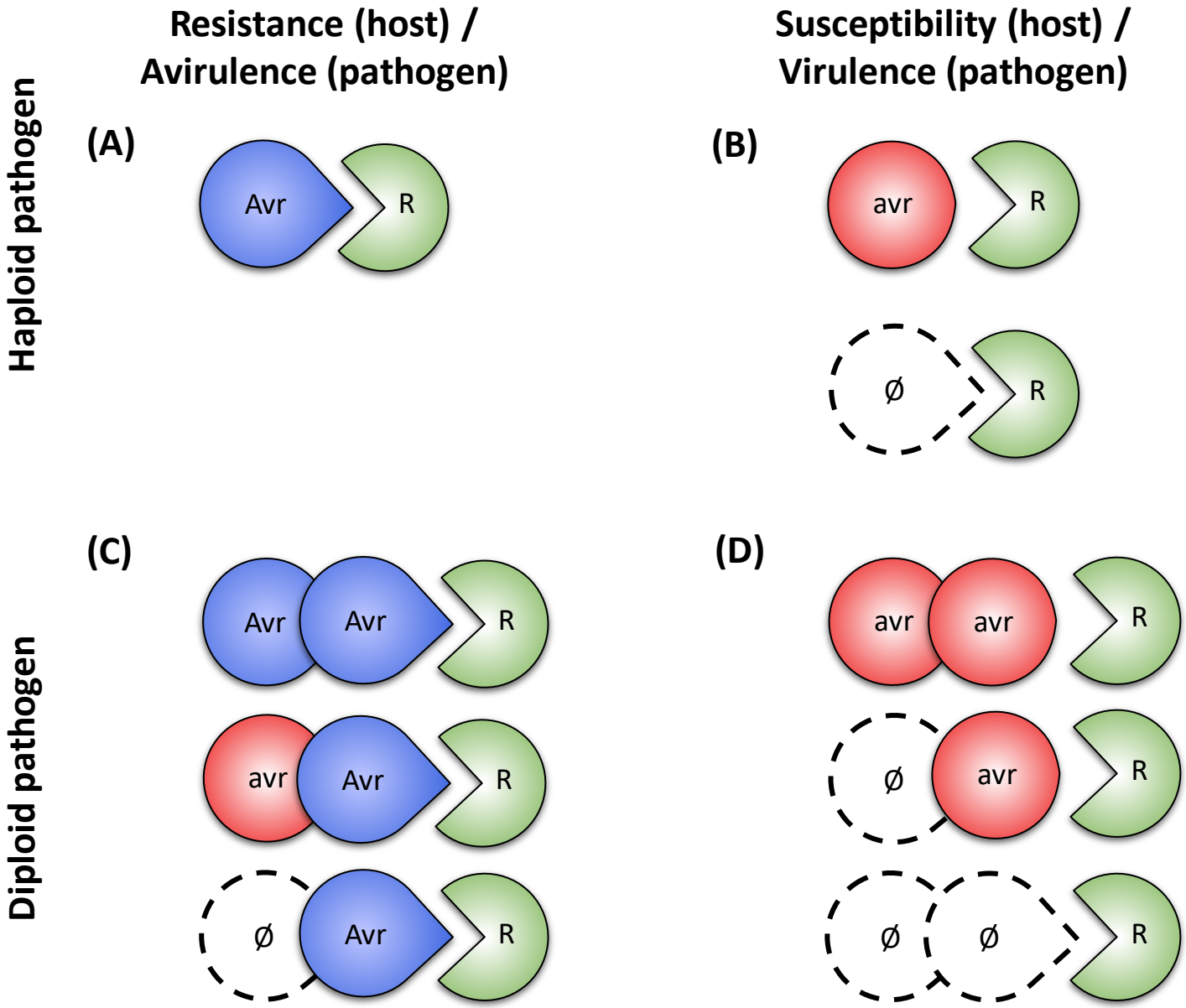
†A7, avirulent 7 phenotype towards *RMlp7*; a7, virulent 7 phenotype towards *RMlp7.*

‡ Genotype at the *AvrMlp7* candidate locus: Avr: avirulent allele; avr: virulent allele; Ø: null allele.


2230644
2261286
Genes
SNPs
2226000
2266000
1
2
2 kb
Probe^avrMlp7
5’ CTGAGCTGACTACGGCTGGTGCCACCGTGTGTCTGCGTGCGTGTGATTGAAACTGTTCTAACATTGA 3’
Virulent allele
SNP_2233369
G81S
Allelic
state
5’ CTGAGCTGACTACGGCTGGTGCCACTGTGTGTCTGCGTGCGTGTGATTGAAACTGTTCTAACATTGA 3’
Avirulent allele
Probe^AvrMlp7


A
Pattern A
Pattern B
Pattern C
200
100
0
98AG54
94ZZ03
08EA50_1-2-3
98AA06
93GV08
94Z11
94ZZ18
98AB02
2.228
2.230
2.232
2.234
2.236
2.238
2.240
2.242
Position (Mbp)
B
Mean normalized amplification
1.0
0.8
0.6
0.4
0.2
0.0
98AG54
94ZZ03
08EA50_1-2-3
98AA06
93GV08
94ZZ11
94ZZ18
98AB02


Fossil
Cultivated
Wild
13 13 23 19 9 8 8
29 22 8 21 9 11 30 25
18
1.0
0.8
0.6
0.4
0.2
0.0
Allelic frequency
1990 1991 1992 1993 1994 1995 1997
1994 1995 1997 1998 1999 2000 2008 2016
2008
Avr allele
avr allele
Ø allele


Resistance (host) /
Avirulence (pathogen)
Susceptibility (host) /
Virulence (pathogen)
Haploid pathogen
Diploid pathogen
(A)
(B)
(C)
(D)
Avr
R
avr
Ø