Development of a tool combining multiplex PCR and high-throughput sequencing to study the polymorphism of eight *Leptosphaeria maculans* avirulence genes and its application to field surveys in France

Angélique Gautier, Valérie Laval, Marie-Hélène Balesdent

Université Paris-Saclay, INRAE, UR Bioger, Place de l'Agronomie, CS 80022, 91120 Palaiseau Cedex France Corresponding Author email valerie.laval@inrae.fr

Keywords. Survey, virulence, *Brassica napus*, polymorphism, detection, *Leptosphaeria maculans*

Abstract

Leptosphaeria maculans is a pathogen of oilseed rape (*Brassica napus*), mainly controlled by major resistance genes (*Rlm*) in varieties. *Rlm* genes can rapidly be overcome following the selection of virulent isolates with deletions or mutations in the corresponding avirulence genes (*AvrLm*). Reasoned management of *Rlm* genes relies on the detection and monitoring of virulent isolates in populations. Based on previous knowledge of *AvrLm* gene polymorphism, we developed a tool combining multiplex PCR and Illumina sequencing to characterize allelic variants for eight avirulence genes in field *L. maculans* populations.

We tested the method on DNA pools of 71 characterized *L. maculans* isolates and of leaf spots from 32 *L. maculans* isolates. After multiplex-PCR and sequencing with MiSeq technology, reads were mapped on an in-house *AvrLm* sequence database. Data were filtered using thresholds defined from control samples included in each run. Proportions of each allelic variant per gene, including deletions, perfectly correlated with expected ones. The method was applied to around 1300 symptoms (42 pools of mainly 32 leaf spots) from nine fields. The proportion of virulent isolates estimated by sequencing leaf spot pools perfectly correlated with those estimated by pathotyping. In addition, the proportions of allelic variants determined at the national scale also correlated with those previously determined following individual sequencing of eight *AvrLm* genes in a representative collection of isolates. Finally, the method also allowed us to detect still undescribed and rare allelic variants. Altogether, the method appeared suitable for large-scale and regular monitoring of *L. maculans* populations.

31 INTRODUCTION

32 The fight against crop pests is a major issue for food safety in a changing environmental and
33 socio-economic context. Whether chemical or genetic, control measures are more effective
34 as they are adapted to local pest populations. These populations also are constantly evolving
35 in response to the deployment of chemical or genetic protection solutions. For instance,
36 pesticide-resistant individuals, whether fungi or insects, regularly appear in populations
37 following the large-scale use of pesticides (McCulloch & Waters, 2023). The selection of such
38 individuals in populations renders the pesticides ineffective. Similarly, plant resistance genes
39 introduced in crop varieties and effective against fungal isolates possessing matching
40 avirulence genes exert strong selection pressure on pathogenic populations. The selection of
41 virulent isolates in populations rapidly decreases the efficacy of the genetic solution.
42 Monitoring pest populations and their characterization regarding their potential susceptibility
43 to control measures is thus a main issue, not only in choosing an effective management
44 strategy but also in detecting outbreak events and preventing their spread.

45 Monitoring virulent isolates is particularly important for the fungal ascomycete *Leptosphaeria*
46 *maculans*, responsible for the stem canker disease of oilseed rape (*Brassica napus*), for the
47 following reasons; (i) the main control measure deployed to limit the impact of the disease is
48 genetic control, including the use of specific resistance (*Rlm*) genes; (ii) *L. maculans*
49 populations have the potential to rapidly evolve toward virulence, and this has been reported
50 in many instances (Rouxel et al., 2003; Sprague et al., 2006, Zhang et al., 2016); (iii) potentially
51 effective resistance sources are scarce in oilseed rape and the few currently available *Rlm*
52 genes must be carefully managed to avoid new breakdowns (Borhan et al., 2022; Rouxel and
53 Balesdent, 2017). Until recently, the monitoring of *L. maculans* populations consisted of large-

54 scale isolate samplings from diseased *B. napus* plants, followed by inoculation of each isolate
55 on a plant differential set (for instance, Alnajar et al., 2022; Balesdent et al., 2006, 2023;
56 Stachowiak et al., 2006; Zhang et al., 2016). These biological tests are however inadequate for
57 high-throughput analyses and for the detection of rare mutation events in populations. They
58 require adapted facilities and materials (seeds of the plant differential set, control isolates,
59 growth chambers with tight temperature and humidity regulation), and they are time- and
60 labor-intensive and thus, costly. Thus, only a limited number of isolates from a limited number
61 of sites can be analysed each year, and the survey results are obtained six to eight months
62 later. In France, large-scale surveys were done at about 10-year intervals on eight to 12 sites
63 (Balesdent et al., 2023). One full working year is needed to fully characterize populations of
64 around 80 isolates per site. In addition, the size of such samplings limits the precision of the
65 estimates of virulent isolates in a population. For instance, in a sampling of 80 isolates, the
66 detection of five virulent isolates (6.2%) only allows us to say that the actual proportion of
67 virulent isolates ranges between 2% and 14% (Exact CI (inrae.fr)), corresponding to very
68 distinct breakdown situations.

69 Fortunately, *L. maculans* is a fungus for which we have a detailed knowledge of the genetic
70 determinants involved in the recognition of the fungus by the plant, i.e., its avirulence (*AvrLm*)
71 genes. To date, 11 *AvrLm* genes have been cloned (Bohran et al., 2022), and the molecular
72 mechanisms by which these genes are mutated or inactivated to avoid being recognized by
73 the corresponding *Rlm* gene were analysed in detail for a few avirulence genes, such as
74 *AvrLm3* and *AvrLm4-7* (Balesdent et al., 2022). A recent study of the polymorphism of these
75 genes in a French population described contrasting situations, depending on the gene, as
76 regards the level of polymorphism in virulent or avirulent isolates and the type of molecular
77 events responsible for the gain of virulence (Gautier et al., 2023). Some *AvrLm* genes were

78 nearly monomorphic while others were highly polymorphic, and the mutational events in
79 virulent isolates included single nucleotide mutation (SNP), inactivating mutations introducing
80 stop codons, particularly following Reaped-Induced Point mutations (RIP; Fudal et al. 2009;
81 Gladyshev, 2017), deletion of the whole gene including the surrounding region, or the masking
82 of the avirulence function by the presence of another avirulence gene. For instance, the
83 presence of at least some alleles of *AvrLm4-7* is sufficient to mask the *AvrLm3-Rlm3* or the
84 *AvrLm9-Rlm9* interactions (Plissonneau et al., 2017, Ghanbarnia et al., 2018). With this
85 knowledge, one can now envisage developing molecular tools for the survey of virulent
86 isolates at one or more avirulence loci.

87 Some molecular methods were already developed for this purpose, but they are often based
88 on unsatisfactory compromises. For instance, some tools only rely on PCR amplification of the
89 avirulence gene, and the percentage of virulent isolates is deduced from the percentage of
90 isolates that fail to amplify the gene. This approach works well for *AvrLm* genes having gene
91 deletion as the main event in virulent isolates, such as *AvrLm1* (Gout et al., 2007). Even though
92 a few isolates display inactivated versions of *AvrLm1* that can still be amplified in virulent
93 isolates, these situations are rare, with only one such isolate among 89 in the French
94 population (Gautier et al., 2023). The lack of PCR amplification could thus be a good proxy for
95 the frequency of virulent *avrLm1* isolates. A quantitative PCR assay was developed to quantify
96 the proportion of deleted isolates for *AvrLm1*, relative to the ITS (Van de Wouw et al., 2010).
97 In most cases, however, the deletion is not the only mutational event in virulent isolates, and
98 deletions may be rare or even never observed for some *AvrLm* genes (Gautier et al., 2023, Van
99 de Wouw et al., 2023).

100 Thus, *AvrLm* sequencing in individual isolates is the easiest way to access the genotype and to
101 infer its phenotype, provided the correspondence between genotype and phenotype is
102 available. In the situation of epistatic interactions between avirulence genes, such as the
103 masking effect of *AvrLm4-7* on the *AvrLm3*/*Rlm3* and the *AvrLm9*/*Rlm9* interactions, the
104 sequencing of the two genes involved in the epistatic interaction should also give access to
105 the phenotype thanks to the detailed work already published on this model of interaction.

106 To avoid PCR amplification of each Avr gene and subsequent sequencing, different approaches
107 have been proposed. Van de Wouw and Howlett (2014) developed a pyrosequencing method
108 to quantify the proportions of virulent isolates toward *Rlm4*. This quantitative method works
109 on pools of individuals and tags one polymorphic site in *AvrLm4-7* known to be responsible
110 for avirulence towards *Rlm4* (Parlange et al., 2009). However, this method gives sequence
111 information on a small part of the gene only (90 bp) and thus cannot take into account the
112 recently described diversity in this gene (Balesdent et al., 2022). An HRM (High-Resolution
113 Melting analysis) method was also developed to detect sequence polymorphism in *AvrLm4-7*
114 and determine virulence toward *Rlm4* and *Rlm7*. The method worked well in the situation of
115 limited polymorphism of the gene (Carpezat et al., 2014). However, with the recent large-scale
116 use of *Rlm7,* at least in France, and the diversification of sequence polymorphism in virulent
117 isolates (Daverdin et al., 2012; Balesdent et al., 2022), it became more and more complex to
118 convert new HRM profiles into virulent or avirulent alleles.

119 Finally, whole genome sequencing (WGS) of individual isolates gives direct access to the
120 sequences of the full set of *AvrLm* genes in an isolate. For instance, Chen et al. (2020) studied
121 the polymorphism in nine avirulence genes following the WGS of 162 Canadian isolates.
122 Similarly, sequences of avirulence genes in an international collection of 205 *L. maculans*

123 isolates were recently published (Van de Wouw et al., 2023). Although powerful and
124 exhaustive, such an approach cannot be performed on hundreds of isolates every year. In
125 addition, this method necessitates first to isolate and purify each isolate.

126 Here, we developed a tool based on MiSeq sequencing technology, with two levels of
127 multiplexing. First, eight *AvrLm* genes and one control gene (Actin) are amplified together in
128 a multiplex PCR. Second, the multiplex-PCR is done on pools of phoma leaf spots, the typical
129 *L. maculans* symptoms on *B. napus* leaves, to increase the size of the population assessed and
130 to avoid the *L. maculans* isolation step. PCR products of pooled samples are sequenced using
131 MiSeq technology and the proportions of each allele of each avirulence genes in the pooled
132 sample are determined. The tool is named here "MPSeqM tool » (Multiplex on Pools and
133 Sequencing with MiSeq). After the development and validation of the tool on DNA pools of
134 known isolates, we applied the method to samples composed of pools of leaf discs centred on
135 phoma leaf spots from experimental fields. Pools of 32 to 96 different leaf spots (possibly
136 corresponding to as many different individuals, Gout et al., 2006) were tested and the results
137 were compared to phenotyping data of isolates from the same plots and the same sampling
138 date (Balesdent et al., 2023). This approach not only allowed us to define a robust protocol
139 for large-scale monitoring of virulence in the field, but also revealed a diversity of situations
140 between sites, between *AvrLm* genes, and many unknown and rare alleles.

141

142 **MATERIALS AND METHODS**

143 *Isolates*

144 Seventy-one *L. maculans* isolates were chosen, including the reference isolate JN3 (v23.1.3),
145 based on their polymorphism in their avirulence genes (Supplementary Tables S1 and S2;
146 Gautier et al., 2023). The isolate X83.51 was generated following an *in vitro*-cross aiming at
147 accumulating Avr gene deletion in a single isolate. *AvrLm1*, *AvrLm4-7*, *AvrLm11* and *AvrLmS-*
148 *Lep2* are deleted in this isolate. X83.51 was used as a negative control for these genes and in
149 serial DNA dilutions with variable proportions of deleted alleles.

150 *Field samples*

151 The oilseed rape variety Drakkar, devoid of resistance genes, was sown at seven sites in
152 autumn 2019 and at one site in 2020 (Table 1), with two to five plots per site. Individual leaves
153 with typical *L. maculans* leaf spots were collected in autumn (300 per site, all collected from
154 distinct plants). Part of these leaves was previously used to isolate around 80 single-pycnidium
155 isolates per site for pathotyping on a differential plant set (Balesdent et al., 2023). In parallel,
156 leaf disks (one disk per leaf) centred on a single leaf spot were cut using a cork-borer (1 cm
157 diameter). Leaf disks were stored in 24-weel microplates at -20°C until further use
158 (Supplementary Figure S1). Before DNA extraction, leaf spots from a given site were pooled
159 by 24 or 32 in one Falcon tube and freeze-dried for 24 hours. Three to six pools per site were
160 prepared (Table 1).

161 *Plant inoculation for preparation of DNA control*

162 A selection of 32 *L. maculans* isolates with known avirulence alleles was inoculated on
163 cotyledons of *B. napus* cv. Westar (devoid of *Rlm* gene) using standard protocols (Gautier et
164 al., 2023; Supplementary Tables S1 and S2). Fourteen days after inoculation, leaf spots that
165 developed on cotyledons were cut as described above. One spot per isolate was placed in a

166 50 mL Falcon tube, to constitute a pool of 32 leaf spots of known *AvrLm* allele composition.
167 The pool was freeze-dried during 24 hours before DNA extraction.

168 *DNA extraction and manipulation*

169 For each of the 71 *L. maculans* isolates, fungal DNA was extracted independently from 1 mL
170 of crude conidial suspensions (between 10^8 and 10^9 spores mL^{-1}) using the DNeasy 96 Plant
171 Kit (Qiagen) following the manufacturer's instructions (Gautier et al., 2023). Freeze-dried
172 pools of leaf spots, whether from inoculations or from the fields, were crushed with a tungsten
173 carbide bead (3 mm, Qiagen) on a Retsch MM400 mixer mill for 60 seconds at 30 Hz and DNA
174 was extracted using the DNeasy Plant Mini kit (Qiagen). DNA was eluted in 100 µL of elution
175 buffer AE (Qiagen). DNA quality was assessed with NanoDrop™ One (Thermo Scientific™) and
176 DNA concentration was measured with Qubit™ 2.0 fluorimeter (Invitrogen™).

177 *Control DNA samples*

178 Four different controls were included, in triplicate, in each sequencing run: (i), multiplex PCR
179 of JN3 DNA (termed JN3_DNA) used as a positive control; (ii), multiplex PCR of the pool of 32
180 leaf spots from inoculations with known isolates (termed Mix_LSpot); (iii), multiplex PCR of a
181 pool of 71 *L. maculans* DNA (termed Mix_LM); and (iv), multiplex PCR of the 11 points of a
182 dilution series for deletion quantification (termed C_dil_ser). The Mix_LM was prepared by
183 mixing 50ng of DNA of each isolate and used to test the ability to detect all allelic variants in a
184 pool (Supplementary Table S2). The dilution series was prepared by mixing DNA of isolates
185 JN3 and X83.51 in proportions varying from 0 % to 100% (with a 10% increment).

186 *Multiplex PCR and MiSeq sequencing*

187 Primers were designed using the Primer3 website https://primer3.ut.ee/ (Table 2). The PCR
188 reaction contained 40 ng of DNA, 1X Master Mix of Type-it Microsatellite PCR Kit (Qiagen), 4
189 µL Q-Solution, and 0.25 µM of each primer (except *AvrLm2* primers at 0.2 µM and *AvrLmS-*
190 *Lep2* primers at 0.4 µM) in a total volume of 40 µL. PCR reactions were performed using an
191 Applied Biosystems 9700 thermocycler. The amplification parameters consisted of an initial
192 denaturation step at 95˚C for 5 min; then 35 cycles of denaturation at 95˚C for 60 s, annealing
193 at 60˚C for 90 s, and elongation at 72˚C for 60 s; and a final extension step at 72˚C for 10 min.
194 Three independent PCRs were performed for each sample to test the repeatability and
195 robustness of the method. The PCRs were verified on 3% agarose gels and sent to the
196 GenoToul platform (INRAE Toulouse) for Illumina MiSeq sequencing using the chemistry V3
197 (2x250pb).

198 *AvrLm allele database*

199 At the start of the project, the *L. maculans AvrLm* database included the 87 alleles and 63
200 isoforms identified following the sequencing of 11 avirulence genes in a collection of French
201 and reference isolates (Gautier et al., 2023) and available at
202 https://bioinfo.bioger.inrae.fr/portal/data-browser/public/leptosphaeria/avirulence_genes.
203 Additional *AvrLm3* and *AvrLm4-7* alleles described in Balesdent et al. (2022) were added to
204 the database.

205 *Analysis of sequence dataset: pairing, filtering, processing and assigning an allelic*
206 *classification to sequences*

207 First, for each sample, forward (R1) and reverse (R2) reads were paired using the tool Pear -
208 Paired-End read merger (Galaxy Version 0.9.6.2) of Galaxy pipeline (https://vm-galaxy-
209 prod.toulouse.inrae.fr/galaxy_main). Then, the data were processed gene by gene. For each

210 gene, only sequences with the expected primer sequences were retained. A stringent pipeline
211 (Boutigny et al., 2019) was used to filter and process these paired sequences, utilising
212 MOTHUR v.1.36.1 (Schloss et al., 2009) and parameters adapted to each gene (Supplementary
213 Text T1). Sequences with low quality (shorter or including low-quality scores) were removed
214 (Supplementary Text T1). After preprocessing the sequences, sequence files were merged,
215 and unique sequences represented by singletons or 2 reads were removed. Chimeric
216 sequences were removed with the Uchime tool (Edgar et al., 2011) available in MOTHUR, using
217 self as the reference. Uncorrected pairwise distances were determined between the
218 remaining sequences with the Needleman alignment method (Needleman et al., 1970) with
219 match 1.0, mismatch -1.0, gapopen -2.0 and gapextend 1.0 using 1000 iterations. The distance
220 matrix was then clustered with the average neighbor algorithm to assign sequences to
221 operational taxonomic units (OTUs). Using MOTHUR's classifier (Wang et al., 2007), the
222 representative sequence for each OTU was assigned to the allele of the corresponding *AvrLm*
223 gene from the custom *Lmb_avrlm_act genes* reference database using no divergence. An OTU
224 count table was obtained containing OTU number, read count per OTU per sample,
225 representative sequence name, representative sequence, and OTU representative assignation
226 (Figure 1). In a final step, the Galaxy blastn tool was used, to compare of the OTUs sequences
227 retained against *Lmb_avrlm_act genes* references database and distinguish the known alleles
228 with 100% identity in length and sequence homology to new alleles with less than 100%
229 identity in sequence.

230 *Strategy for the selection of relevant OTUs and database cleaning*

231 After the mapping step, a very high number of OTUs were identified (Table 3), most of which
232 correspond to artefacts due to PCR or the MiSeq technology. Three filters were sequentially

233 applied (Figure 1) to keep only reliable OTUs as follows : (i), keep only OTUs present in at least
234 three samples, due to the use of triplicates for each DNA sample; (ii), remove all OTUs for
235 which, whatever the sample analysed, the read numbers were below a threshold. This
236 threshold is named "Maximum read number threshold" and is gene-dependent (Table 3) (iii),
237 remove OTUs that were detected in controls but should not be present. After this OTU
238 selection step, we calculated the mean read number of the three replicates for each OTU and
239 DNA sample. Finally, the resulting count table was simplified by considering an OTU absent in
240 one sample if its proportion in this sample (i.e., read count of the OTU divided by total read
241 counts for all OTUs for this gene) was lower than a threshold ("proportion" threshold), specific
242 for each gene. This percentage threshold is determined based on the expected and observed
243 OTU proportions in controls (Table 3, Figure 1).

244 *Analysis of new AvrLm alleles*

245 When a new OTU was detected, the missing parts of the gene sequence, if any, (i.e., when
246 primers were within the genes) were completed by the corresponding 5' or 3' sequences of
247 the JN3 reference isolate to generate a new putative allele. Introns were removed and the
248 sequence was translated with the Expasy translate tool (https://web.expasy.org/translate/)
249 to generate an extrapolated, putative new isoform (Figure 1).

250 **RESULTS**

251 **Primer design and reproducibility of multiplex PCR and sequencing**

252 Two technical criteria governed primer design ensuring the reliability of the sequencing
253 results: (i) the amplified region must be smaller than 460 bp for correct paired-end subsequent
254 analysis; (ii) all amplified sequences in the multiplex PCR must be of similar length to avoid
255 over-representation of the shortest sequences (maximum of 100 pb difference). The amplified

256 region also must contain a maximum of known polymorphic sites and all codons encoding a
257 cysteine. Based on these constraints, we designed a set of primers fitting these criteria (Table
258 2). These primers amplified part of the genes containing between 66% and 100 % of the known
259 mutations (Table 2, Supplementary Figure S2). A few known variants were not detected, with
260 no or low consequences for the diagnosis of the allele as an avirulent or a virulent one. For
261 instance, four *AvrLm3* variants cannot be detected as mutations are outside the amplified
262 region (Table 2). However, these mutations are either in the last intron or in the peptide signal-
263 coding region, without impact on the mature protein (Supplementary Fig S2-C).

264 Using these primers, the MPSeqM tool was first tested with the concomitant amplification of
265 four, five, or nine genes in a single PCR. Sequencing results were similar whatever the number
266 of genes amplified together (data not shown) and we thus choose the nonaplex amplification
267 in all further experiments. MPSeqM of control DNAs (JN3_DNA, Mix_LSpot, and Mix_Lm)
268 revealed an efficient and reproducible amplification level of all genes (Figure 2 and data not
269 shown). The less amplified gene was *AvrLm6*, followed by *AvrLmS*, still representing more than
270 1250 reads each in each control replicate.

271 **Detection and quantification of deletions**

272 To estimate the proportion of deleted alleles of one *AvrLm* gene in an environmental sample,
273 we used serial dilutions of the DNA of an isolate deleted for this gene with the DNA of an
274 isolate with the gene present. The theoretical proportions of the *AvrLm* gene in the dilution
275 series were compared with the estimated proportions calculated as the total number of reads
276 mapping to this gene over the total number of reads mapping to the *L. maculans* actin gene,
277 which is a single copy and invariant gene. The two values were proportional and allowed us to
278 define standard curves of variable shapes depending on the gene (Figure 3). We approximated

279 the curves to a polynomial equation, whose parameters were used to calculate the proportion
280 of undeleted events and, by difference, the deleted ones, in a pooled sample. For instance,
281 the expected proportion of deleted *AvrLm1* in the Mix_LSpot is 87.5% ([exact confidence
282 interval [71-96]) and the estimated proportion of the deletion calculated with the standard
283 curve equation is 84.5% (Supplementary Table S3).

284 **Threshold determination for OTUs assignation**

285 The use of triplicates for each sample and four types of controls of known allele composition
286 were essential to determine the thresholds used to keep only reliable OTUs and suppress
287 potential PCR or sequencing artefacts. Mainly, the "Maximum number of reads" threshold
288 allowed us to reduce drastically the number of OTUs (1204 OTUs in total, starting from 81024
289 OTUs kept with the criteria of presence in at least three sequencing samples, Table 3). The
290 controls with the reference JN3 isolate or with DNA mixes of known isolates allowed also us
291 to remove all OTUs still unexpectedly found in these controls, which represented 3/4 of the
292 remaining OTUs. Altogether, 300 reliable OTUs (all genes) were kept, among which 206 were
293 not in the database, 74 of which corresponded to highly mutated versions of the genes by RIP
294 (Table 3). Lastly, depending on the gene, OTUs representing less than 0.1 to 0.5% of the total
295 read counts for this gene in a field sample ("proportion" threshold) were considered absent
296 from this sample (Table 3).

297 **Validation of the method on control samples**

298 The previous steps allowed us to select reliable OTUs and to determine the proportion of
299 deleted alleles for *AvrLm1*, *AvrLm11*, and *AvrLm4-7*. Using this cleaned dataset, we compared
300 the proportion of alleles (OTUs) detected in control samples to the expected ones (Figure 4,
301 Supplementary Table S3). In the Mix_Lm control, including 71 isolates, 20 alleles out of 90

302 were not detected, but 19 of them corresponded to a rare occurrence (only one isolate among
303 71). In addition, most of them (13) were RIPped alleles. In the Mix_LSpot control, only 3 out
304 of 47 expected alleles were not detected, all corresponding to RIPped alleles for *AvrLm1* or
305 *AvrLm4-7*. All other alleles, including those present in only one isolate out of 32, were detected
306 (Supplementary Table S3). Overall, observed and expected proportions correlated very well
307 (R^2>0.98). Only one *AvrLm4-7* allele (the deleted allele) was overestimated in the Mix_Lm
308 control, possibly due to the lack of amplification of some RIPped alleles. Apart from this allele,
309 all estimated values were within the confidence interval determined for sample sizes of 32
310 (Mix_LSpot) or 71 (Mix_Lm) individuals (Figure 4).

311 **Analysis of unknown environmental samples: identification of new alleles**

312 The MPSeqM tool, validated on controls, was then used on 42 environmental samples of
313 pooled leaf spots corresponding to 1337 *L. maculans* symptoms from eight French regions
314 (Table 1). This analysis first revealed several still undescribed alleles for most of the *AvrLm*
315 genes studied. Such new alleles were considered as real alleles, not sequencing artefacts,
316 when they combined the following characteristics: (i) they were detected in some samples
317 only, not as a background signal in all samples including controls; (ii) they were consistently
318 detected in the three independent technical replicates of one given environmental sample;
319 (iii) their read numbers were higher than the gene-dependent threshold value determined for
320 each gene; and (iv) they were detected both in one "pool of 32 leaf spots" and in the
321 corresponding "pool of pools", when available, with consistent proportions (Figure 5). Using
322 these criteria, a total of 206 new alleles were detected.

323 **Analysis of unknown environmental samples: gene-by-gene analysis**

324 *AvrLm1*

325 Only one out of the five previously described *AvrLm1* alleles was detected in field samples,
326 allele *AvrLm1_0_A.* In addition, we detected two new alleles, one RIPped allele with 14
327 substitutions and one allele with a stop codon in the middle of the sequence. These two new
328 alleles were considered as virulent, as were deleted isolates whose proportions were
329 estimated using the deletion curve. Altogether, the MPSeqM analysis of environmental
330 samples correlated well with pathotyping; 1.8% of the population was found avirulent on *Rlm1*
331 by pathotyping (Balesdent et al., 2023), while the MPSeqM analysis of populations from the
332 same sites detected 2.14% of such isolates. With both approaches, the site with the highest
333 proportion of avirulent isolates was that of Le Rheu (Supplementary Figure S3). With the
334 MPSeqM tool, it was possible to reproducibly detect low percentages of avirulent isolates in
335 all sites, while, by phenotyping 30 isolates per site, such isolates were not detected (Balesdent
336 et al., 2023).

337 *AvrLm2*

338 No new *AvrLm2* allele was detected in field samples. The *AvrLm2_0_A* was largely prevalent,
339 with proportions ranging from 91% to 100% (average across all pools = 97.3%, Supplementary
340 Figure S4) depending on the pool. The *AvrLm2_1_A* allele was found in all sites at a low
341 frequency (0 to 8% per pool, average across all pools = 2.2%). *AvrLm2_2_A* was found in six
342 sites (10 pools out of 42; between 0 and 6.9%; average across all pools = 0.52%). The other
343 known alleles, including all avirulent ones, were never detected in the pools.

344 *AvrLm3*

345 From the 24 previously described *AvrLm3* alleles, 13 were detected in the environmental
346 samples. Twenty-eight new OTUs were also detected, being either rare (present in one sample
347 only) or more frequent (present in up to 23 samples), allowing us to identify in total 41
348 different OTUs for *AvrLm3*. The new alleles could represent up to 8.5% of the OTUS in a given

349 sample. These new alleles encoded for ten putative new isoforms of the AvrLm3 protein.
350 Based on their alignment with previously described AvrLm3 isoforms, we inferred that six
351 among them potentially correspond to avirulent isoforms and four to virulent ones
352 (Supplementary Figure S5). Despite the allelic diversity detected in field samples, four alleles
353 only represented 96% of the population, with allele *AvrLm3_A1* being by far the most
354 abundant (80.2%), followed by *AvrLm3_C* (11.8%) (Supplementary Figure S6), consistent with
355 previous surveys and sequencing (Gautier et al., 2023).

356 *AvrLm4-7*

357 MPSeqM results for environmental samples displayed 110 new *AvrLm4-7* alleles (Table 3).
358 Most of them (66) corresponded to RIPped alleles with mutations introducing stop codons in
359 the sequence and thus corresponding to virulent alleles, also unmasking *AvrLm3*. Seven
360 sequences displayed one or two bp deletions, also encoding for truncated (virulent) proteins.
361 A few (17) were new alleles of known (virulent or avirulent) isoforms, however, 25 new
362 isoforms were detected. The amino acid (AA) sequence of these new isoforms was compared
363 to previously described ones (Balesdent et al., 2023) and the corresponding phenotypes were
364 more or less easily deduced depending on the AA substitutions. One isoform putatively
365 corresponded to an A4A7 (conferring avirulence on both *Rlm4* and *Rlm7*) isoform, six alleles
366 for a3a7 isoforms (conferring virulence on *Rlm4* and *Rlm7* and on *Rlm3* by masking effect),
367 four for a3A7 (avirulent on *Rlm7* only) isoforms. For the remaining 14 isoforms, it remained
368 impossible to conclude.

369 *AvrLm5-9*

370 Due to primer design, alleles *AvrLm5-9_0_A* and *AvrLm5-9_0_C* (encoding for the same
371 isoform) cannot be distinguished. In field samples, seven new variants were detected. Only

372 three of them corresponded to putative new isoforms of the protein; one displayed one AA
373 change in the signal peptide compared to the AvrLm5-9_0 isoform without any modification
374 of its secretory properties (data not shown). The two others displayed only one AA change
375 each (R63L and K101Q, respectively), whose impact on the phenotype cannot be determined.
376 Among the two alleles encoding avirulent isoforms on *Rlm9*, isoform *AvrLm5-9_4* was never
377 detected in any of the 42 pools, and *AvrLm5-9_1* was found in only nine pools with very low
378 proportions (<0.27%, close to the 0.2% detection limit, 0.06% on average over all pools),
379 confirming the near-fixation of the virulence phenotype toward *Rlm9* in French *L. maculans*
380 populations (Balesdent et al., 2023 Gautier et al., 2023). Similarly, alleles with stop codons in
381 the coding sequence were never detected. Depending on the site, from two to six alleles were
382 detected (Supplementary Figure S7).

383 *AvrLm6*

384 Environmental samples displayed five new OTUs. They were however rare, present in one or
385 two sites only with proportions ranging from 0.6% to 4.6% (Supplementary Figure S8). Each
386 new OTU displayed only one mutation. One was in an intron; three were silent and only one
387 of them introduced a G47C substitution, and thus a new isoform. The impact of this mutation
388 on the phenotype cannot be determined yet. Allele *AvrLm6_0_A* was largely prevalent,
389 ranging from 91.7% to 100% depending on the sample, followed by allele *AvrLm6_1_A* or
390 *AvrLm6_3_A*, two alleles that cannot be discriminated with the current primers (between 0%
391 and 4.1%).

392 *AvrLm11*

393 For *AvrLm11*, two new OTUS displayed a shorter-than-expected size (one or two bp). These
394 gaps were either in the intron or after the stop, thus without impact on the isoform of the

395 protein. In addition, 27 new OTUs were detected in at least one environmental sample,
396 displaying from one to seven mutations along the gene. These new OTUs were often specific
397 to one sampling site (Figure 5) and, depending on the site, from one to 14 new OTUs were
398 present. Most of the mutations (72%) in these OTUs were typical of RIP mutations (C->T and
399 G-> A). In these OTUs however, the number of mutations was limited (from 1 to 7 mutations
400 per *AvrLm11* OTU) compared to what is seen in RIPped alleles for other genes like *AvrLm4-7*
401 (from 1 to 44 mutations, mean=20.2, data not shown). Despite this low number of mutations
402 per allele, these RIP mutations introduced stop codons in seven OTUs at positions 29 or 74 of
403 the protein leading to truncated versions of the protein that thus corresponded to virulent
404 alleles. For 11 OTUs, the mutations were silent, including one mutation in the intron or three
405 mutations after the stop codon, and corresponded to the avirulent AvrLm11_0 isoform. The
406 remaining OTUs corresponded to 13 new isoforms. One of them displayed a mutation in a
407 codon encoding a cysteine, therefore generating a de-structured protein and the probable
408 loss of avirulence. For the other 12 new isoforms, we could not determine yet the impact of
409 these mutations on the interaction phenotype. However, they represented only 0.7% of the
410 detected alleles in average, but could reach up to 7.8% in pools displaying numerous new
411 alleles. Altogether, alleles encoding the AvrLm11_0 isoform were largely prevalent in the
412 population (94.6% in average), while isoforms AvrLm11_1 or AvrLm11_2, previously described
413 in Canada, USA or New Zealand (Chen et al., 2020; Van de Wouw et al., 2023) were never
414 detected. *AvrLm11* deletion was estimated at 3.1% in average, ranging from 0% to 15.4%
415 depending on the sample.

416 *AvrLmS*

417 As already observed following the sequencing of individual isolates (Gautier et al., 2023; Van
418 de Wouw et al., 2023), a high level of polymorphism was detected for *AvrLmS*. Forty-one
419 alleles were detected here, with 27 new alleles compared to previous data (Gautier et al.,
420 2023). The proportions of each *AvrLmS* allele greatly varied between samples including within
421 one given site (Supplementary Figure S9). Two new alleles displayed one mutation introducing
422 a stop codon in the sequence, as previously detected in virulent isolates (Gorse M. and Gautier
423 A., unpublished data; Van de Wouw et al., 2023). These virulent alleles were detected in only
424 one pool from the Boisseau site at low frequency. Nineteen new alleles displayed silent
425 mutations and encoded avirulent isoforms. Three alleles displayed non-synonymous
426 mutations leading to AA substitutions in the signal peptide without modification of the mature
427 (avirulent) protein. Only three alleles displayed non-synonymous mutations (R41Q, S49L and
428 E71G) for which the impact on the phenotype could not be determined. They represented in
429 average 0.56% of the detected alleles (data not shown).

430 **Analysis of unknown environmental samples: comparison between 32 or 96 leaf spots in a**
431 **pool**

432 We compared the proportions of each detected allele in pools of 32 leaf spots or in "pools of
433 pools" containing up to 96 leaf spots. As exemplified for *AvrLm3* for which a high level of alleles
434 were detected, the proportion of each allele estimated from both pool sizes were highly
435 correlated (R^2=0.9996, Supplementary Figure S10). However, when focusing on alleles present
436 at low percentages in the pools (>2%) the correlation was lower (R^2=0.859), due to OTUs that
437 were not detected in pools of 96 leaf spots (detection limit) but also to OTUs detected in pools
438 of 96 but not in pools of 32, probably corresponding to sequencing artefacts (Supplementary
439 Figure S10).

440 **Analysis of unknown environmental samples: comparison between genotyping and**
441 **phenotyping for virulence.**

442 Using leaf samples from the same site and the same sampling date, we compared the
443 percentages of avirulent isolates in *L. maculans* populations as estimated following inoculation
444 tests on a plant differential set (Balesdent et al., 2023) or using the MPSeqM tool (this study).
445 For most avirulence loci, the inference of the phenotype (virulence or avirulence) from the
446 allele sequences or deletion proportions was possible (see above). As regards *AvrLm3*, the
447 situation is more complex since the avirulence phenotype is masked whenever *AvrLm4-7* is
448 present, at least for some *AvrLm4-7* alleles (Balesdent et al., 2023). To determine the
449 proportion of avirulent isolates on *Rlm3*, we multiplied, for each pool, the percentage of
450 *AvrLm3* avirulent alleles by the percentage of *AvrLm4-7* alleles unable to mask the *AvrLm3-*
451 *Rlm3* interaction. Altogether, phenotypic or MPSeqM analyses gave very close estimates of
452 the proportions of avirulent isolates for the interactions investigated, including the complex
453 *AvrLm3-Rlm3* interaction (Figure 6A). For instance, site by site comparison of MPSeqM or
454 phenotypic estimation of virulent isolates on *Rlm7* perfectly correlated (Supplementary Figure
455 11).

456 **Analysis of unknown environmental samples: comparison between Sanger sequencing**
457 **and the MPSeqM tool**

458 We previously estimated *AvrLm* allele frequencies in French *L. maculans* population by Sanger
459 sequencing of full-length avirulence genes in a representative population of 89 isolates
460 collected in 2017 (Gautier et al., 2023). These proportions were compared with those
461 determined by the MPSeqM tool using averaged frequencies of alleles over all 42
462 environmental samples. Proportions of alleles that are undistinguishable with the MiSeq

463 primers (Like *AvrLm3_A_01* and *AvrLm3_A_09*) were first merged, and similarly, proportions
464 of RIPped and deleted alleles for *AvrLm1* were merged. Altogether, 44 distinct alleles were
465 detected for the eight *AvrLm* genes in the 2017 population. The percentages of these alleles
466 estimated with the MPSeqM tool in the 2019 population or by Sanger sequencing of
467 avirulence genes in the 2017 populations perfectly correlated (R^2=0.995, Figure 6B). The
468 difference between the two estimates of a given allele was low (1.8% in average). For only two
469 *AvrLm4-7* alleles, percentages between the two analyses were higher than 5%, *AvrLm4-7_2_A*
470 (an avirulent allele) being lower in 2019 (16.3%) compared to 2017 (22.5%) and the deleted
471 allele being higher in 2019 (45.0%) compared to 2017 (39.3%; Figure 6B). Indeed, these
472 differences may simply reflect the current evolution of French populations toward virulence
473 due to the *Rlm7* selection pressure in France (Balesdent et al., 2022).

474 **Analysis of unknown environmental samples with the** MPSeqM **tool: allelic diversity per**
475 **site**

476 Considering all genes, from 80 to 118 distinct OTUs were detected per site. Relative to the
477 number of leaf spots per site analysed, the site showing the highest diversity (number of
478 distinct OTUs) was that of Frelinghein (Supplementary Figure S12). Depending on the sites,
479 from 40% (Frelinghein) to 53% of the detected OTUs were new OTUs, the site with highest
480 number of new OTUs being that of Coullons, mainly due to the high number of new *AvrLm11*
481 OTUS found at that site (Supplementary Figure S12).

482 **DISCUSSION**

483 In this paper, we describe the design, validation, and use of a molecular tool, termed MPSeqM,
484 for large-scale analysis of polymorphism of eight avirulence genes of *L. maculans* intending to
485 determine the proportion of virulent isolates toward the corresponding *B. napus* resistance
486 genes in field samples. This tool exploits the potentiality of high-throughput sequencing

487 technologies and the in-depth knowledge of *L. maculans AvrLm* polymorphism. It combines
488 four steps: bulk DNA extraction of *L. maculans* symptoms from the field; a multiplex-PCR step
489 amplifying eight *AvrLm* genes and *Actin* as a control, invariant gene; Illumina sequencing of
490 the PCR products, and lastly, blasting reads on an *AvrLm* allelic variant database to finally
491 calculate proportions of each variant in the bulks. The tool was first evaluated on DNA pools
492 of known samples. Whether using pool of DNA from purified and individually extracted
493 isolates or DNA from leaf symptoms developed by defined isolates and extracted in bulks, the
494 MPSeqM tool was found to be versatile, reproducible, quantifiable, and efficient for the
495 detection of small proportions of allelic variants.

496 As regards versatility, the possibility to characterize simultaneously from 4 to 8 avirulence
497 genes with identical results opens the possibility to modify the set of genes analysed
498 depending on the characteristics of the populations to be surveyed. For instance, here,
499 *AvrLm2* is nearly invariant with fixed virulent alleles in the French populations, rendering the
500 *Rlm2* resistance genes useless, as previously described (Balesdent et al., 2023, Gautier et al.,
501 2023). The MPSeqM tool could thus be adapted by replacing *AvrLm2* amplification with other
502 genes of interest, like *AvrLm10* or *AvrLm14*. Alternatively, fewer genes could be analysed
503 simultaneously, to reach deeper sequencing for other genes or to analyse more samples for
504 the same sequencing cost.

505 As regards reproducibility, we showed, using the MPSeqM tool with three independent PCR
506 replicates (different PCR mixes and independent amplifications), that the sequencing results
507 were very reproducible both qualitatively and quantitatively within one sequencing run. In
508 addition, independent runs also gave identical estimates of proportions of each allele (data
509 not shown). The presence of plant material in bulks of leaf spots did not interfere with the
510 MPSeqM results. The quantitative dimension of the tool was clearly established not only when
511 comparing the MPSeqM results with expected proportions of each alleles in bulks of known
512 composition, but also when comparing MPSeqM results from bulks of unknown field
513 symptoms with previously established data, whatever the criteria compared (proportion of
514 virulent isolates per site, or proportions of allelic variants per gene).

515 Three main questions had to be address to validate the tool: (i) what could be the best sample
516 size (number of leaf spots) within each pool for field surveys, (ii) whether gene deletions could
517 be quantified, and (iii) whether new alleles could be detected.

518 MPSeqM results for control DNA with known isolates were used to determine the best
519 number of individual samples in one bulk. In the Mix_LM control containing DNA of 71 isolates,
520 many alleles (>20%) present in one isolate only were not detected. This value dropped to
521 around 6% in the bulk with 32 leaf spots and, more importantly, 100% of alleles encoding non-
522 RIPped sequences were detected. Lack of detection of a RIPped allele present at low
523 frequency is not a problem if the deletions (i.e. lack of amplification) can be accurately
524 detected, because both RIPped and deleted alleles correspond to virulent versions of the
525 gene. Comparison of pools of 32 or 96 leaf spots confirmed a better detection of alleles at low
526 frequencies with pools of 32 leaf spots. Therefore, pools of 32 leaf spots were selected to
527 assess the MPSeqM on unknown samples.

528 Thanks to the use of a dilution series with varying proportions of DNA of two isolates differing
529 by the presence or absence of some avirulence genes, we showed that it was possible to
530 quantify deletions. For that purpose, however, we first had to normalize the data (number of
531 sequences for a given avirulence gene) with the number of reads obtained for a single-copy,
532 conserved gene. Here we used the actin gene for normalization, as it is reliable for species-
533 specific identification of *L. maculans* within the *Leptosphaeria* species complex (Vincenot et
534 al., 2009; Voigt et al., 2005). The normalized read numbers for the deleted genes in the dilution
535 series correlated with the expected proportions, demonstrating again the quantitative
536 dimension of the tool. Using the dilution curves, the proportions of deleted isolates in known
537 bulks were accurately quantified for either high proportions (*AvrLm1*), low proportions
538 (*AvrLm11*), or intermediate proportions (*AvrLm4-7*) of deletion events.

539 Unexpected alleles were detected at low frequencies in bulks with known *AvrLm* allele
540 composition. These sequences corresponded to artefacts generated at the PCR or sequencing
541 steps. Thus, the proportions of such reads in controls were used to establish the thresholds
542 under which MiSeq reads were suppressed. These thresholds were gene-dependent. Once
543 these thresholds were established and applied, the datasets from field samples comprised a
544 reduced number of allelic variants per gene, but still numerous unknown alleles. We defined

545 rules to select reliable alleles among these new sequences, and at this step, the use of
546 independent triplicates was particularly useful. Only unknown sequences consistently present
547 in the three replicates of one field sample and in a few samples only were kept, allowing us to
548 identify 206 new alleles including 74 new RIPped alleles. Depending on the mutations present
549 in these new alleles, the inference of the corresponding phenotype (virulent or avirulent) was
550 more or less easy. RIPped sequences, sequences with frameshift or stop codons were
551 considered virulent. Other variants may display silent mutations or mutations without impact
552 on the mature protein and thus, were easily converted into avirulent or virulent alleles.
553 However, a non-negligible amount of new alleles (9.2% of new alleles detected in pools) could
554 not be converted into a phenotype. These new alleles usually represented low proportions of
555 reads per sample (1.5% on average), thus without a high impact on the survey result. If an
556 unknown allele reaches high proportions in specific sites, it will be necessary either to go back
557 to isolation campaigns at this site, with the aim of recovering and phenotyping isolates with
558 that allelic variant, or to assess the corresponding phenotype following the transformation of
559 a virulent isolate with a synthetic construct displaying this new allele.

560 Although field leaf spot bulks contained 32 leaf spots only, the proportions of reliable alleles
561 in one bulk were frequently lower than 1/32 (3.12%), the expected proportion in case of the
562 presence of this allele in one leaf spot of the bulk only. Two main hypothesis can explain this
563 data. First, despite the care taken to choose individualized leaf spots on the leaves, some leaf
564 spots may have merged into one, or latent and multiple infections may be present around the
565 selected leaf spot without being visible. One leaf spot may thus contain more than one *L.*
566 *maculans* individual. Second, gene or large genomic region duplications in one isolate may
567 occur. Two copies of *AvrLm4-7* or *AvrLm11,* or duplication of their genomic regions were
568 already described (Daverdin et al., 2009; Gautier et al., 2013; Leclair et al., 1996) and the two
569 copies may display polymorphism (Gautier et al., 2023). Whatever the hypothesis, and
570 because the proportions of virulent isolates calculated on the bases of OTUs proportions in
571 bulks or by phenotyping correlated well, we think the percentage of each OTU in one bulk
572 gives a good estimate of its proportion in the natural population.

573 The MPSeqM tool has one technical limitation. Because of the rather short sequence length,
574 only part of the full CDS length is amplified for some *AvrLm* genes, possibility preventing us
575 from detecting mutations in the 5’ or 3’ regions of the genes. Despite this limitation, the

576 MPSeqM tool enabled us to determine frequencies of virulent isolates from field samples
577 highly correlated with those previously determined by pathotyping. This good correlation is
578 probably due to the reasoned choice of the amplified region, including a maximum number of
579 the polymorphic sites previously identified. However, this limitation has to be kept in mind,
580 and regular (for instance, every five years) surveys based on isolation, individual pathotyping,
581 and full-length CDS sequencing will still have to be done in the future, to detect potential new
582 virulent alleles, to enrich the sequence database, and to adapt the primer pairs to current
583 prevalent alleles. Another specific difficulty concerns *AvrLm4-7*. It has been shown previously
584 that some virulent isolates toward *Rlm7* did not amplify the gene with primers located outside
585 to the CDS, but could amplify some internal parts of the gene with a sequence matching
586 avirulent alleles (Daverdin et al., 2012; Gautier et al., 2023). Such virulent isolates,
587 representing 6.7% of the 2017 sampling (Gautier et al., 2023), will be miss-identified as
588 avirulent with the MPSeqM tool. This phenomenon could explain why the MPSeqM tool
589 detects more avirulent isolates toward *Rlm7* than the biological pathotyping (Figure 6).

590 Our MPSeqM field survey revealed highly contrasted levels of polymorphism depending on
591 the avirulence gene. This observation confirms previous data obtained by sequencing
592 avirulence genes in individual isolates, either from France (Gautier et al., 2023) or worldwide
593 (Van de Wouw et al., 2023), with highly polymorphic genes like *AvrLm3*, *AvrLm4-7* or *AvrLmS-*
594 *Lep2*, or low polymorphic ones like *AvrLm2* or *AvrLm6*. Here, by analysing a total of 1337 leaf
595 spots, we not only confirmed these trends but detected a number of new, rare alleles, most
596 of them being site or even plot-specific. This work thus revealed an unexpected diversity of
597 *AvrLm* sequences for at least some genes. Whether this polymorphism is pre-existing
598 everywhere at a basal level, or whether it is locally generated at each generation through
599 meiosis as previously suggested (Daverdin et al., 2012), is still in question. The high number of
600 new RIPped alleles detected here for *AvrLm4-7* (66 new sequences), but not for other genes
601 suggests RIP is still an active mechanism particularly frequent for this gene, generating new
602 alleles every year at meiosis.

603 Finally, the MPSeqM tool gives a detailed and precise characterization of the diversity of a set
604 of *L. maculans* genes in populations present in a define location (geographical location,
605 specific time-point, plant genotype, plant organ). Although the field trials were not designed
606 for statistical analyses of population diversity, we were nevertheless able to highlight

607 contrasting situations between sites, with some sites showing greater diversity but few new
608 alleles, and others being less diverse but richer in new alleles. The MPSeqM tool thus offers
609 new perspectives for additional research questions, such as: do some plant genotypes select
610 some allelic variants along the cropping season? How plant genotypes affect allelic variants?
611 How different are populations present in very close or distant fields? How some resistance
612 management strategies could affect the allelic diversity of avirulence genes? The MPSeqM
613 tool could also be adapted to other genes for which we miss polymorphism information, such
614 as *L. maculans* effector genes expressed at later stages of the infection cycle. Finally, this tool
615 could help bringing extensive field data to determine a "critical percentage of virulent isolates"
616 at which the resistance is in danger of being rapidly overcome, an information currently not
617 available for the *L. maculans* / *B. napus* pathosystem.

618 **Acknowledgments**

619 This work was funded by the CASDAR project C2018-11 (Atipical). We thank all contributors
620 from the CTPS field experiments that provided infected oilseed rape leaves and Julie Noah for
621 her contribution to leaf spot recovery. The "Effectors and Pathogenesis of *L. maculans*" group
622 benefits from the support of Saclay Plant Sciences-SPS (ANR-17-EUR-0007). We thank the GeT
623 Genotoul platform in Toulouse, Midi-Pyrénées for the MiSeq sequencing experiments. We are
624 grateful to the Genotoul Bioinformatics platform in Toulouse, Midi-Pyrénées for providing
625 help in computing and data storage.

626

627 **Data availability statement**

628 Data available on request from the authors

629 **References**

630 Alnajar, D., von Tiedemann, A. & Koopmann, B. (2022) Efficacy of blackleg major resistance
631 genes in *B. napus* in Germany. *Pathogens* 11, 461.

632 Balesdent, M.H., Gautier, A., Plissonneau, C., Le Meur, L., Loiseau, A., Leflon, M. et al. (2022)
633 Twenty years of *Leptosphaeria maculans* population surveys in France suggest

634 pyramiding *Rlm3* and *Rlm7* in rapeseed is a risky resistance management strategy.
635 *Phytopathology* 112, 2126-2137.

636 Balesdent, M.H., Laval, V., Bagot, P., Rouxel, T (2023) Large-scale population survey of
637 *Leptosphaeria maculans* in France highlights both on-going breakdowns and potentially
638 effective resistance genes in oilseed rape. *Pest Management Sciences*
639 https://doi.org/10.1002/ps.7401

640 Balesdent, M.H., Louvard, K., Pinochet, X. & Rouxel, T. (2006) A large-scale survey of races of
641 *Leptosphaeria maculans* occurring on oilseed rape in France. *European Journal of Plant*
642 *Patholo*gy 114, 53-65.

643 Borhan, M.H., Van de Wouw, A.P. & Larkan, N.J. (2022) Molecular interactions between
644 *Leptosphaeria maculans* and *Brassica* species. *Annual Review of Phytopathology* 60,
645 11.1–11.21

646 Boutigny, A.L., Gautier, A., Basler, R., Dauthieux, F., Leite, S., Valade, R. et al. (2019)
647 Metabarcoding targeting the EF1 alpha region to assess *Fusarium* diversity on cereals.
648 PLoS ONE 14(1): e0207988. https://doi.org/10.1371/journal.pone.020798.

649 Carpezat, J., Bothorel, S., Daverdin, G., Balesdent, M.H. & Leflon, M. (2014) Use of high
650 resolution melting analysis to genotype the avirulence *AvrLm4-7* gene of *Leptosphaeria*
651 *maculans*, a fungal pathogen of *Brassica napus. Annals of Applied Biology* 164, 430–444.

652 Chen, Q., Peng, G., Kutcher, R., & Yu, F. (2020) Identification of genome-wide DNA variants
653 and SNP haplotypes associated with avirulence genes in *Leptosphaeria maculans* in
654 Western Canada. *Research Square:* https://doi.org/10.21203/rs.3.rs-24766/v1.

655 Daverdin, G., Rouxel, T., Gout, L., Aubertot, J.N., Fudal, I., Meyer, M. et al. (2012) Genome
656 structure and reproductive behaviour influence the evolutionary potential of a fungal
657 phytopathogen. *PLoS Pathogens* 8, e1003020.

658 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., & Knight, R. (2011) UCHIME improves
659 sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200.
660 https://doi.org/10.1093/bioinformatics/btr381.

661 Fudal, I., Ross, S., Brun, H., Besnard, A.L., Ermel, M., Kuhn, M.L. et al. (2009) Repeat-induced
662 point mutation (RIP) as an alternative mechanism of evolution towards virulence in
663 *Leptosphaeria maculans*. *Molecular Plant-Microbe Interaction* 22, 932-941.

664 Gautier, A., Laval, V., Faure, S., Rouxel, T. & Balesdent, M.H. (2023) Polymorphism of
665 avirulence genes and adaptation to Brassica resistance genes is gene dependent in the
666 phytopathogenic fungus *Leptosphaeria maculans*. *Phytopathology*, 113, 1222-1232.

667 Ghanbarnia, K., Ma, L., Larkan, N.J., Haddadi, P., Fernando, D.W.G. & Borhan, M.H. (2018)
668 *Leptosphaeria maculans AvrLm9*: a new player in the game of hide and seek with
669 *AvrLm4-7*. *Molecular Plant Pathology* 19, 1754-1764.

670 Gladyshev, E. (2017) Repeat-Induced Point Mutation (RIP) and other genome defense
671 mechanisms in fungi. *Microbiology Spectrrum* 5, 10.1128.

672 Gout, L., Eckert, M., Rouxel, T. & Balesdent, M.H. (2006) Genetic variability and distribution of
673 mating-type alleles in field populations of *Leptosphaeria maculans* from France. *Applied*
674 *and Environmental Microbiology* 72, 185-191.

675 Gout, L., Kuhn, M.L., Vincenot, L., Bernard-Samain, S., Cattolico, L., Barbetti, M. et al. (2007)
676 Genome structure impacts molecular evolution at the *AvrLm1* avirulence locus of the
677 plant pathogen *Leptosphaeria maculans*. *Environmental Microbiology* 9, 2978-2992.

678 Leclair, S., Ansan-Melayah, D., Rouxel, T., & Balesdent, M.H. (1996) Meiotic behaviour of the
679 minichromosome in the phytopathogenic ascomycete *Leptosphaeria maculans. Current*
680 *Genetics* 30, 541-548.

681 McCulloch, G. A., & Waters, J. M. (2023) Rapid adaptation in a fast-changing world: emerging
682 insights from insect genomics. *Global Change Biology* 29, 943–954.

683 Needleman, S.B., & Wunsch, C.D. (1970) A general method applicable to the search for
684 similarities in the amino acid sequence of two proteins. *Journal of Molecular Biol*ogy 48,
685 443–453.

686 Parlange, F., Daverdin, G., Fudal, I., Kuhn, M.L., Balesdent, M.H., Blaise, F. et al. (2009)
687 *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity
688 by the *Rlm4* and *Rlm7* resistance genes of oilseed rape, and circumvents *Rlm4*-mediated
689 recognition through a single amino acid change. *Molecular Microbiology* 71, 851-863.

690 Plissonneau, C., Blaise, F., Ollivier, B., Leflon, M., Carpezat, J., Rouxel, T. et al. (2017) Unusual
691 evolutionary mechanisms to escape Effector-Triggered-Immunity in the fungal
692 phytopathogen *Leptosphaeria maculans. Molecular Ecol*ogy 26, 2183-2198.

693 Rouxel, T. & Balesdent, M.H. (2017) Life, death and rebirth of avirulence effectors in a fungal
694 pathogen of Brassica crops, *Leptosphaeria maculans. New Phytologist* 214, 526-532.

695 Rouxel, T., Penaud, A., Pinochet, X., Brun, H., Gout, L., Delourme, R. et al. (2003) A ten-year
696 survey of populations of *Leptosphaeria maculans* in France indicates a rapid adaptation

towards the *Rlm1* resistance gene in oilseed rape. *European Journal of Plant Pathology* 109, 871-881.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B. et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75, 7537–7541.

Sprague, SJ, Balesdent, M-H, Brun, H, Hayden, HL, Marcroft, SJ, Pinochet, X. et al. (2006) Major gene resistance in *Brassica napus* (oilseed rape) is overcome by changes in virulence of populations of *Leptosphaeria maculans* in France and Australia. *European Journal of Plant Pathology* 114, 33–40.

Stachowiak, A., Olechnowicz, J., Jedryczka, M., Rouxel, T., Balesdent, M.H., Happstadius, I., et al. (2006) Frequency of avirulence alleles in field populations of *Leptosphaeria maculans* in Europe. *European Journal of Plant Pathology* 114, 67-75.

Van de Wouw, A.P., Howlett, B.J. (2012) Estimating frequencies of virulent isolates in field populations of a plant pathogenic fungus, *Leptosphaeria maculans*, using high-throughput pyrosequencing. *Journal of Applied Microbiology* 113, 1145–53.

Van de Wouw, A.P., Scanlan, J.L., Al-Mamun, H.A., Balesdent, M.H., Bousset, L., Burketová, L. et al. (2023) A new set of international *Leptosphaeria maculans* isolates as a resource for elucidation of the basis and evolution of blackleg disease on *Brassica napus*. *Plant Pathology*, DOI: 10.1111/ppa.13801.

717 Van de Wouw, A.P., Stonard, J.F., Howlett, B.J., West, J.S., Fitt, B.D.L. & Atkins, S.D. (2010)
718 Determining frequencies of avirulent alleles in airborne *Leptosphaeria maculans*
719 inoculum using quantitative PCR. *Plant Pathology* 59, 809–818.

720 Vincenot, L., Balesdent, M.H., Li, H., Barbetti, M.J., Sivasithamparam, S., Gout, L. et al. (2008)
721 Occurrence of a new subclade of *Leptosphaeria biglobosa* in Western Australia.
722 *Phytopathology* 98, 321-329

723 Voigt, K., Cozijnsen, A.J., Kroymann, J., Pöggeler, S. & Howlett, B.J. (2005) Phylogenetic
724 relationships between members of the crucifer pathogenic *Leptosphaeria maculans*
725 species complex as shown by mating type (MAT1-2), actin, and beta-tubulin sequences.
726 *Molecular Phylogeny and Evolution* 37, 541-57.

727 Wang, Q., Garrity, G.M., Tiedje, J.M., & Cole, J.R. (2007) Naive Bayesian classifier for rapid
728 assignment of rRNA sequences into the new bacterial taxonomy. *Applied and*
729 *Environmental Microbiology* 73, 5261–5267.

730 Zhang, X., Peng, P., Kutcher, R., Balesdent, M.H., Delourme, R. & Fernando, D.W.G. (2016)
731 Breakdown of *Rlm3* resistance in the *Brassica napus–Leptosphaeria maculans*
732 pathosystem in western Canada. *European Journal of Plant Patholo*gy 145, 659-674.

733

734 **Figure legends**

735 **Figure 1. Schematic representation of sequence analyses and dataset cleaning.** Grey boxes
736 correspond to output files at specific stages of the analysis. [a], this threshold ("maximum read
737 number threshold") removes background OTUs due to sequencing errors and is based on
738 unexpected values observed for control DNAs; [b], this threshold ("percentage threshold") is
739 based on the relative proportion of an OTU in an environmental sample and its value is based
740 on relative proportions observed in control DNAs.

741 **Figure 2. Efficiency and reproducibility of MiSeq sequencing following multiplex PCR**
742 **amplification on single isolate DNA (A) or pooled leaf symptoms from 32 known isolates (B).**
743 Replicates (REP) correspond to independent PCR experiments. Values are percentages of
744 reads matching each amplified gene.

745 **Figure 3. Serial dilution curves for *AvrLm* deletion quantification.** A mix of DNA of two isolates
746 differing by the presence or absence of four avirulence genes created deletion proportions
747 ranging from 0% to 100%. The observed proportion of the gene in the mix was estimated with
748 the ratio between the read numbers for this gene in the sample over the Actin read numbers.
749 In blue, *AvrLm1*; orange, *AvrLm4-7*; grey, *AvrLm11*; green, *AvrLmS-Lep2*. Polynomial
750 regression curves and R^2 are displayed for each gene.

751 **Figure 4. Comparison between expected and observed proportions of each allele for eight**
752 ***L. maculans* avirulence genes in DNA pools.** A/ DNA of a pool of 32 leaf spots from known
753 isolates and B/ Pool of DNA of 71 known isolates. The eight *AvrLm* genes were amplified in
754 multiplex PCR with three independent PCR replicates, and then sequenced with MiSeq. Each
755 dot correspond to the mean proportion in the pools of each allele (OTU) over all alleles of
756 every *AvrLm* gene, including deleted alleles, and the colour code on the left refers to the gene.
757 The lower and upper limits correspond to the exact confidence interval around the expected
758 value for a mix of 32 (A) or 71 (B) individuals. Detailed values are in Supplementary Table S3.

759 **Figure 5. An example of detection of new alleles.** Percentages of new OTUs over all *AvrLm11*
760 OTUs are given for the three control samples (left) and the 52 environmental samples from
761 nine sites, separated by the vertical blue lines. Bars in dark orange are those of pools of 32
762 leaf spots that are also in the "pools of pools "of 96 leaf spots (pale orange).

763 **Figure 6. Validation of the MPSeqM tool on environmental samples**. **A**, Comparison between
764 percentages of avirulent isolates in French populations of *Leptosphaeria maculans,* estimated
765 by phenotyping or genotyping. Black, data from phenotyping 690 isolates from 9 sites
766 (Balesdent et al, 2023); grey, data from MultiplexPCR-Miseq analysis of 42 pools of leaf spots
767 from the same sites (including a total of 1337 leaf spots). Values are means and standard
768 errors over all sites. **B**, Comparison between allele frequencies (all genes) estimated by Sanger
769 sequencing of individual isolates (2017 population, Gautier et al, 2013) and by the MPSeqM
770 tool of pools of leaf spots (2019 population)

771

772 **Supplementary Table S1. List and characteristics of *Leptosphaeria maculans* isolates used**
773 **in this study**

774 **Supplementary Table S2. Genotypes of reference *Leptosphaeria maculans* isolates at eight**
775 **avirulence loci**. Isolates in bold are those present in the poolof 32 *Brassica napus* leaf spots
776 used as control.

777 **Supplementary Table S3. List and proportions of *AvrLm* alleles present in the Mix_Lm and**
778 **Mix_LSpot controls and the corresponding estimated proportions with after MiSeq**
779 **sequencing.**

780 **Supplementary Figure S1. Experimental design for environmental samples analyses.** 1,
781 Leaves at reception in the lab, cleaned and air-dried; 2, one leaf spot per leaf is cut and placed
782 in humid conditions for pycnidia production, and 3, mono-pycnidia isolation for 4, inoculation
783 on a plant differential set to determine the phenotype. 5, standardized leaf spot cutting (one
784 per leaf) and long-term storage in 24-well plates. 7, a view of a full plate before storage until
785 8, pool DNA extraction, multiplex PCR and MiSeq sequencing

786 **Supplementary Figure S2. Alignment of known alleles of eight *L. maculans* avirulence genes**
787 **with positions of forward and reverse MiSeq primers.** A, *AvrLm1*; B, *AvrLm2*; C, *AvrLm3*; D,
788 *AvrLm4-7*; E, *AvrLm5-9*; F*, AvrLmS-Lep2*; G*, AvrLm6*; H, *AvrLm11*. The F-primers are located
789 before the start for *AvrLm5-9* and *AvrLm6* and the R-primer is after the stop for *AvrLm11*.

790 **Supplementary Figure S3. Box plot of the frequency per sampling site of avirulent alleles for**
791 ***AvrLm1,* as estimated by MiSeq sequencing in 42 pools of leaf spots.**

792 **Supplementary Figure S4. Proportions of the three *AvrLm2* alleles detected in 42 pools of**
793 **leaf spots from 9 fields**

794 **Supplementary Figure S5 Maximum likelihood tree of all *AvrLm3* isoforms identified here.**
795 Isoforms underlined in green are known avirulent isoforms while those underlined in red are
796 known virulent isoforms**.**

797 **Supplementary Figure S6. Percentages of the 41 *AvrLm3* alleles as determined following**
798 **multiplex-PCR and MiSeq sequencing of 42 environmental samples**

799 **Supplementary Figure S7. Proportions of *AvrLm5-9* alleles detected in 42 pools of leaf spots**
800 **from nine fields**

801 **Supplementary Figure S8. Proportions of *AvrLm6* alleles detected in 42 pools of leaf spots**
802 **from nine fields**

803 **Supplementary Figure S9. Proportions of *AvrLmS-Lep2* alleles detected in 42 pools of leaf**
804 **spots from nine fields**

805 **Supplementary Fig. 10. Relationships between percentages of each *AvrLm3* OTU per site**
806 **estimated in 32 leaf spots or from the "pools of pools" (64 or 96 leaf spots). Each dot**
807 **corresponds to the percentage of one OTU in one site.** A, all data; B, C and D, focus on high
808 percentages (>60%), low percentages (<20%) or very low percentages (<2%). The green circle
809 highlights OTUs that are detected at low frequencies in 24 or 32 leaf spot pools but not in 64
810 or 96 leaf spots. The red circle highlights OTUs that are not detected in pools of 32 leaf spots
811 but found in pools of 96, thus probably corresponding to sequencing artefacts.

812 **Supplementary Figure S11. Correlation between estimates of percentage of avirulent**
813 **isolates on *Rlm7* by pathotyping and with the MPSeqM tool.** Each dot corresponds to one
814 geographical location. Samplings were done at the same place and same time but on different
815 leaves. Pathotyping was described in Balesdent et al. (2023) and was done using around 80
816 isolates per site, while the genotyping was done on 3 to 6 pools of 32 leaf spots per site. The
817 Spearman correlation coefficient is indicated.

818 **Supplementary Figure S12. Relative *AvrLm* allele diversity per site as described with the**
819 **MpSeqM tool.** A, diversity for all genes; B, diversity for *AvrLm11* only. The radar diagrams
820 represent, for each site, the relative proportion of alleles (no. of alleles at this site over total
821 no. of alleles detected in the survey) averaged by the relative proportion of leaf spots analysed
822 at this site. Dark line, overall diversity; dark grey line, diversity for "new" alleles; pale grey line,
823 diversity for site-specific new alleles. Sites as follows: Co, Coullons; Ja, Janville; Th1, Thiverval-
824 Grignon field 1; Th2, Thiverval-Grignon field 2; Boi, Boisseau; Er, Erdre-en-Anjou; Bou, Bouges-
825 le-Château; Fr, Frelinghein; LeR, Le Rheu

826 **Supplementary Text T1. Parameters used for the Mothur pipeline.**

827

828 Table 1 Characteristics of the sites sampled in autumn 2019 and 2020, and number of samples
829 analysed per site

Location	French department	No. of pools of 32 leaf spots	No. of leaf spots (no. of pools) in the "pool of pools"
Boisseau	41	6	96 (3)
Bouges le château	36	5	96 (3) and 64 (2)
Coullons	45	6	96 (3)
Erdre-en-Anjou	49	4	96 (3)
Frelinghien	59	3[a]	88 (3)
Janville	28	4	96 (3)
Le Rheu	35	6	96 (3)
Thiverval Grignon (field 1)	78	4	64 (2)
Thiverval Grignon (field 2)	78	4	64 (2)

830 [a] except for one pool with only 24 leaf spots

831

832 Table 2. Characteristics of the eight *AvrLm* and actin gene regions amplified with MiSeq primers

Gene	Gene size (bp) for JN3	No. of known polymorphic sites[a]	Forward Primer for MiSeq (5'-> 3') [b]	Reverse Primer for MiSeq (5'-> 3') [c]	Amplicon size (bp)[d]	Length of the analysed part of the gene (bp)	No. (%) of known polymorphic sites in the amplicon[e]
AvrLm1	667	3	TCCGTCCAGATAAGCCCTGT	CACGAATGATGCACCAGAGGC	455	399	3 (100%)
AvrLm2	699	5	CCCAGCAACAATTTTCGGAACTCC	TTCAACTGTGTATTGGCCCTGG	433	377	5 (100%)
AvrLm3	539	25	CCAAATCTTGTTTCTATTATTTACCGCAT	GATGATCGCAGTGTACCTCGTCTTT	511	455	21 (84%)
AvrLm4-7	478	13	ACTCGCTCTCTCTATCCCTACA	CAACCACGAGTCCTTGCTTTCA	495	439	13 (100%)
AvrLm5-9	426	6	ACAACCACTCTTCTTCACAGTTT	TTAGAGATAACTAAGCTCCAATCCG	486	430	5 (83.3%)
AvrLm6	581	3	GCACGTTCCAAAGCCCTATTTA	ATCGTCATGGGCTTGGATGTTA	506	450	2 (66.7%)
AvrLm11	365	3	GCTTCATCATGCGTTTCTTGCT	CCACAGCACACTATTCACTTTAGC	504	448	3 (100%)
AvrLmS-Lep2	426	15	GGCCAACTTTATCATGCTCTACCT	GGCAACTGATGGGACCGCTC	487	431	15 (100%)
Actin partial (512F/CA14R)	374	-	ATGTGCAAGGCCGGTTTCG	GCCAGATCTTCTCCATGTCATCCCAGTT	430	374	-

833 [a] From Start to Stop, data from Gautier et al. (2023)
834 [b] all forward primers were tagged in 5' with MiSeq specific sequence CTTTCCCTACACGACGCTCTTCCGATCT
835 [c] all reverse primers were tagged in 5' with MiSeq specific sequence GGAGTTCAGACGTGTGCTCTTCCGATCT
836 [d] including the 5' and 3' tags and for JN3 isolate
837 [e] except RIP mutations; known mutations as described by Gautier et al. (2023).
838

839 Table 3. Sequencing statistics including numbers of OTUs retained at each filtering step

Gene	Mean read no. per sample (min;max)[a]	No. of OTUs in at least 3 samples	Max. read no. threshold[b]	No. of OTUs above the threshold	No. of 0TUs retained after analysis of control DNA[c]	Threshold for the detection of an OTU in a field sample[d]	Final no. of new OTUs	Final no. of new RIPed OTUs[e]	Final no. of new isoforms (full lenght)
AvrLm1	429 (0;6886)	1045	12	50	5	0.10%	2	1	0
AvrLm2	7977 (5971;14799)	18548	25	22	6	0.20%	0	0	0
AvrLm3	4251 (3045;6594)	12383	10	194	46	0.20%	28	0	10
AvrLm4-7	2183 (1242;4377)	10147	10	278	139	0.20%	110	66	25
AvrLm5-9	4398 (3283;8029)	8643	9	198	14	0.20%	7	0	3
AvrLm6	2330 (1523;4686)	4488	9	75	8	0.15%	5	0	1
AvrLm11	3815 (2692;6196)	7656	10	172	36	0.40%	27	7	13
AvrLmS-Lep2	3965 (2896;5498)	18114	8	215	46	0.50%	27	0	7
Total		81024		1204	300		206	74	59

840 [a] Data including control samples (JN3_DNA; Mix_LSpot; Mix_LM) and environmental samples (pools of leaf spots).
841 [b] All OTUs for which read counts are below these values in all samples are removed. These values are gene- and run- dependent
842 [c] based on unexpected OTUs observed in controls and with similar read counts in field samples

843 [d] This "percentage threshold" represents the percentage of read counts for an *AvrLm* OTU over all read counts for this gene in a given sample. All values
844 with OTU read counts below this threshold are replaced by zero in the final dataset.

845 [e] A sequence is considered as RIPped if it displays a few mutations from C→T or G→ A, introducing stop codons in the sequence.

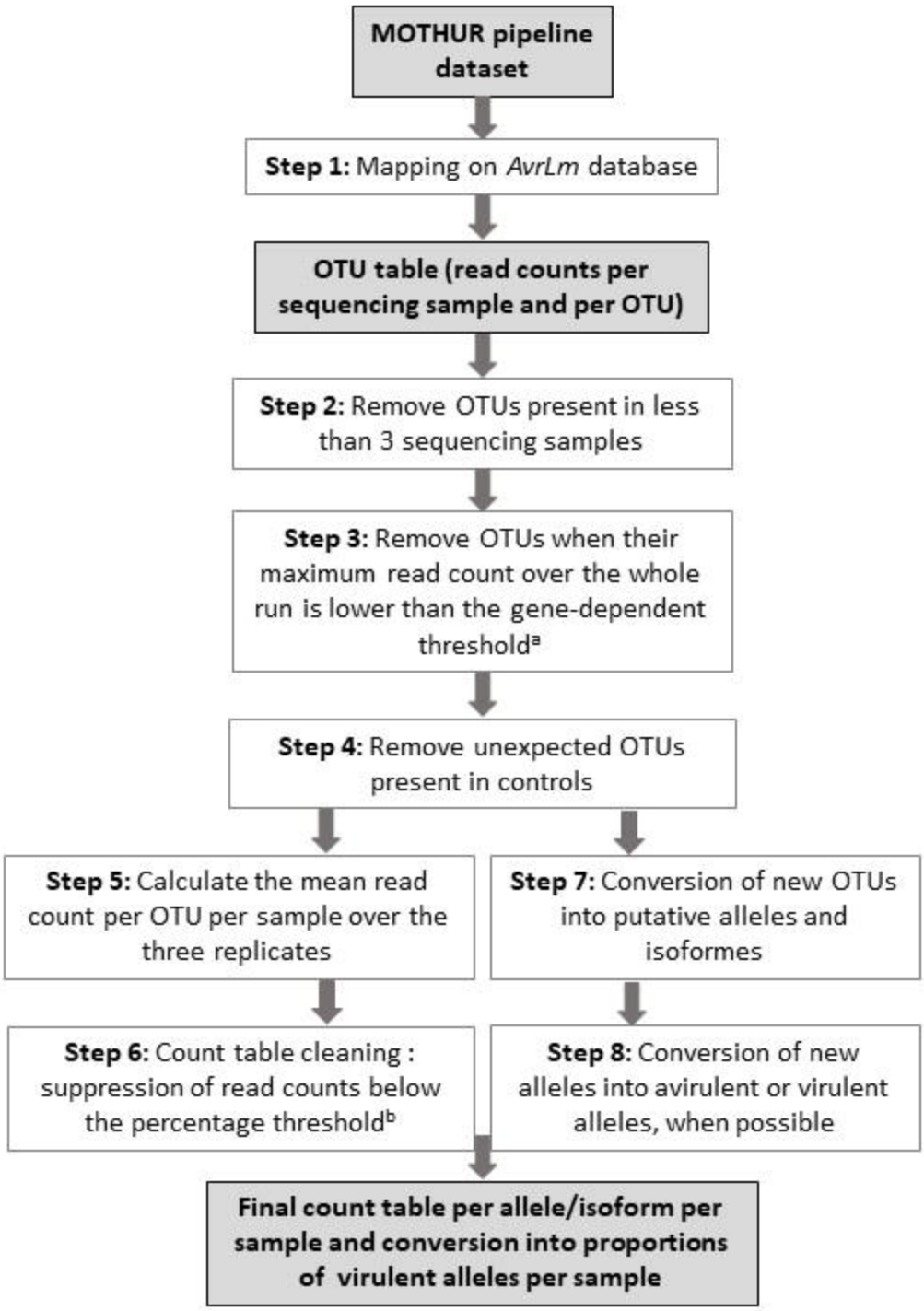

Figure 1. Schematic representation of sequence analyses and dataset cleaning. Grey boxes correspond to output files at specific stages of the analysis. [a], this threshold ("maximum read number threshold") removes background OTUs due to sequencing errors and is based on unexpected values observed for control DNAs; [b], this threshold ("percentage threshold") is based on the relative proportion of an OTU in an environmental sample and its value is based on relative proportions observed in control DNAs.

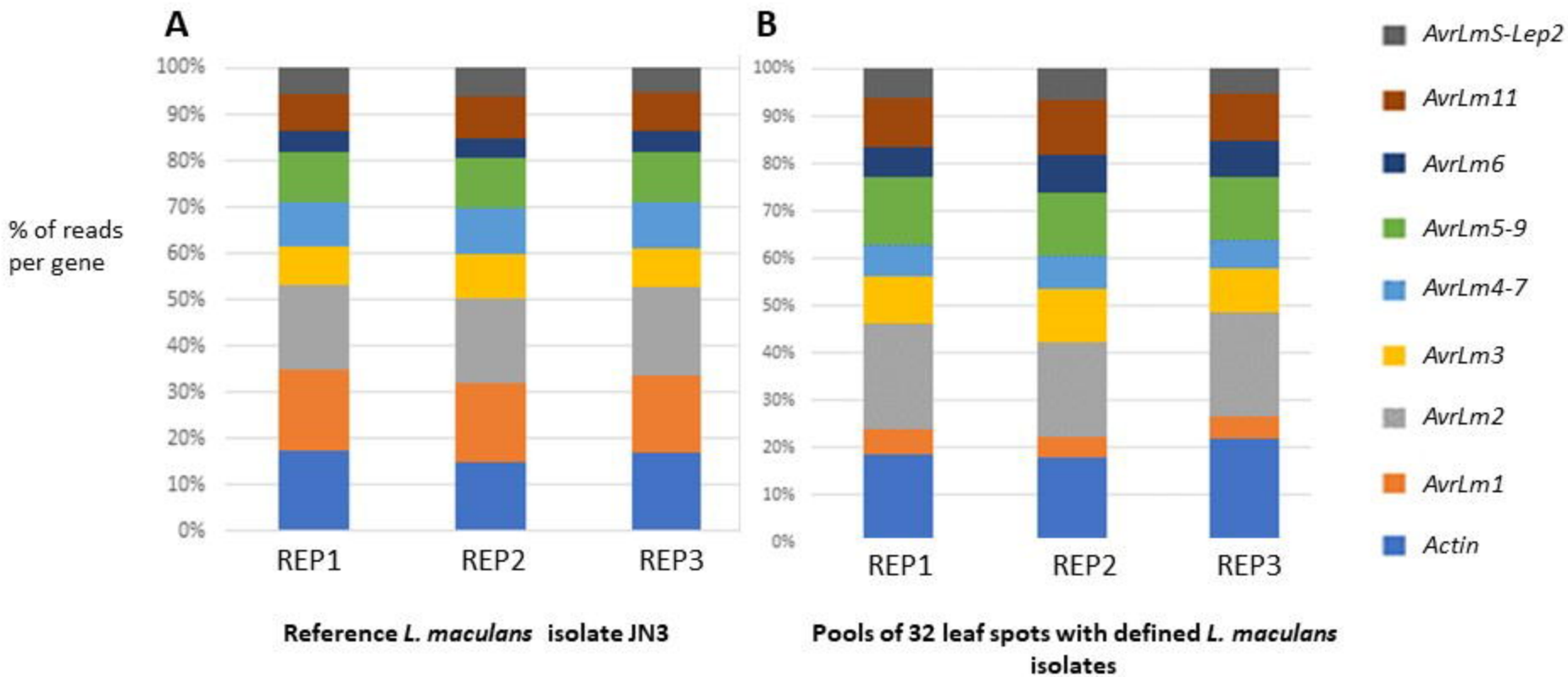

Figure 2. Efficiency and reproducibility of MiSeq sequencing following multiplex PCR amplification on single isolate DNA (A) or pooled leaf symptoms from 32 known isolates (B). Replicates (REP) correspond to independent PCR experiments. Values are percentages of reads matching each amplified gene.

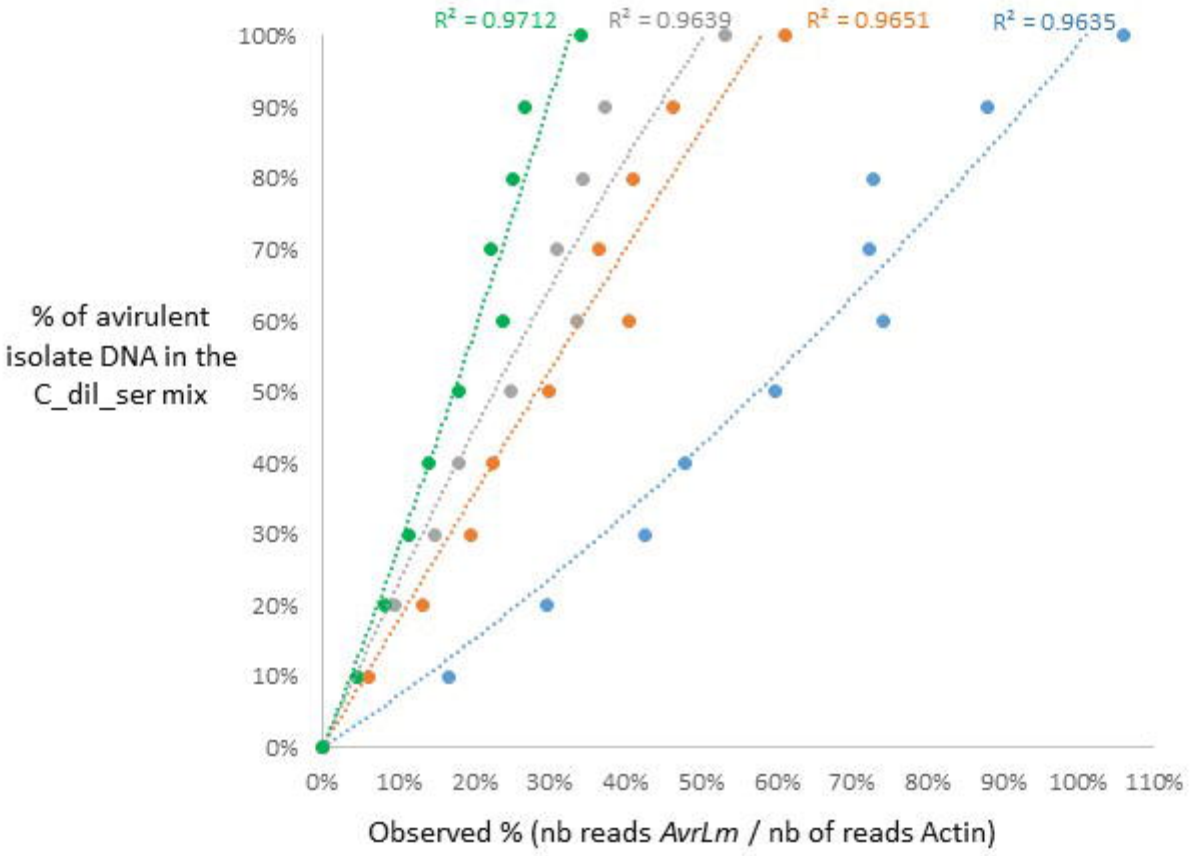

Figure 3. Serial dilution curves for *AvrLm* deletion quantification. A mix of DNA of two isolates differing by the presence or absence of four avirulence genes created deletion proportions ranging from 0% to 100%. The observed proportion of the gene in the mix was estimated with the ratio between the read numbers for this gene in the sample over the Actin read numbers. In blue, *AvrLm1*; orange, *AvrLm4-7*; grey, *AvrLm11*; green, *AvrLmS-Lep2*. Polynomial regression curves and R^2 are displayed for each gene.

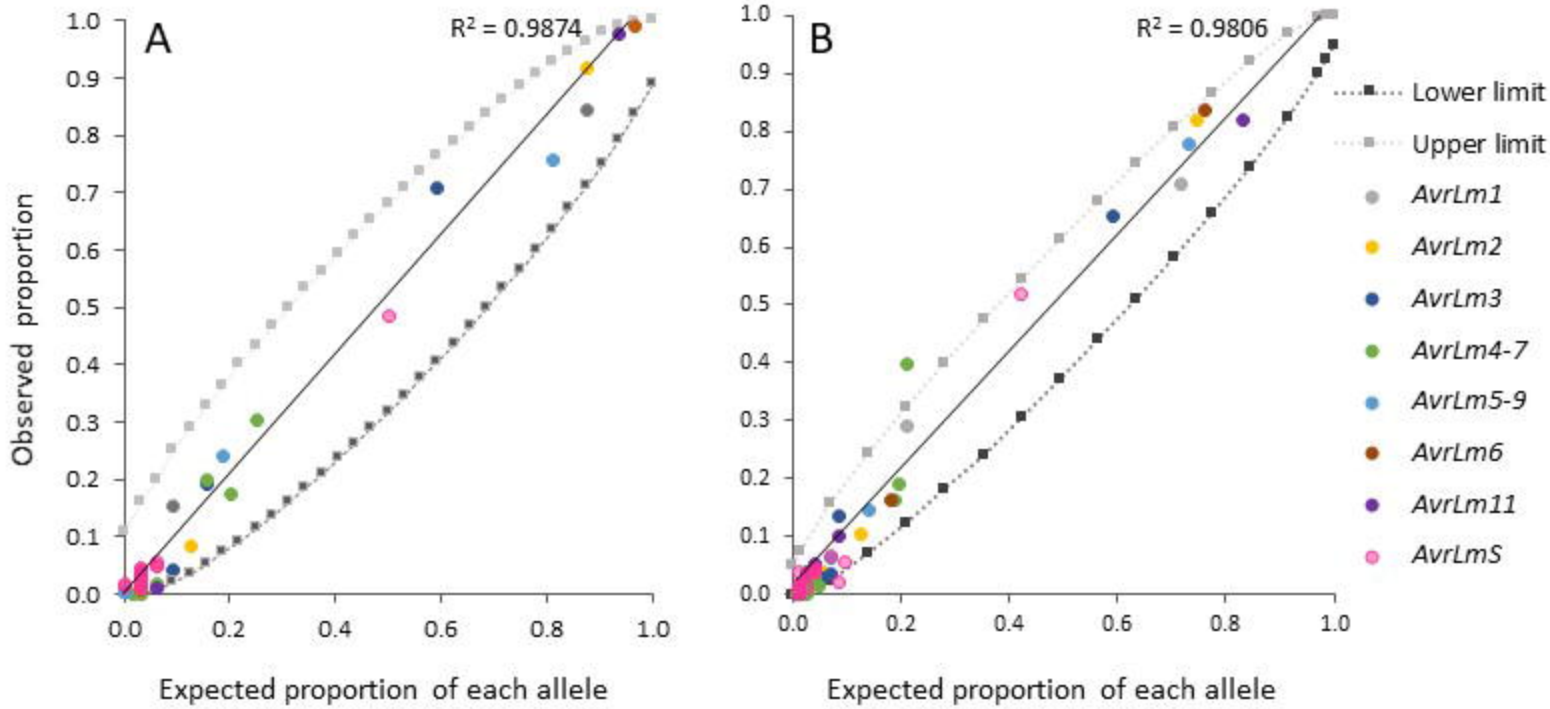

Figure 4. Comparison between expected and observed proportions of each allele for eight *L. maculans* avirulence genes in DNA pools. A/ DNA of a pool of 32 leaf spots from known isolates and B/ Pool of DNA of 71 known isolates. The eight *AvrLm* genes were amplified in multiplex PCR with three independent PCR replicates, and then sequenced with MiSeq. Each dot correspond to the mean proportion in the pools of each allele (OTU) over all alleles of every *AvrLm* gene, including deleted alleles, and the colour code on the left refers to the gene. The lower and upper limits correspond to the exact confidence interval around the expected value for a mix of 32 (A) or 71 (B) individuals. Detailed values are in Supplementary Table S3.

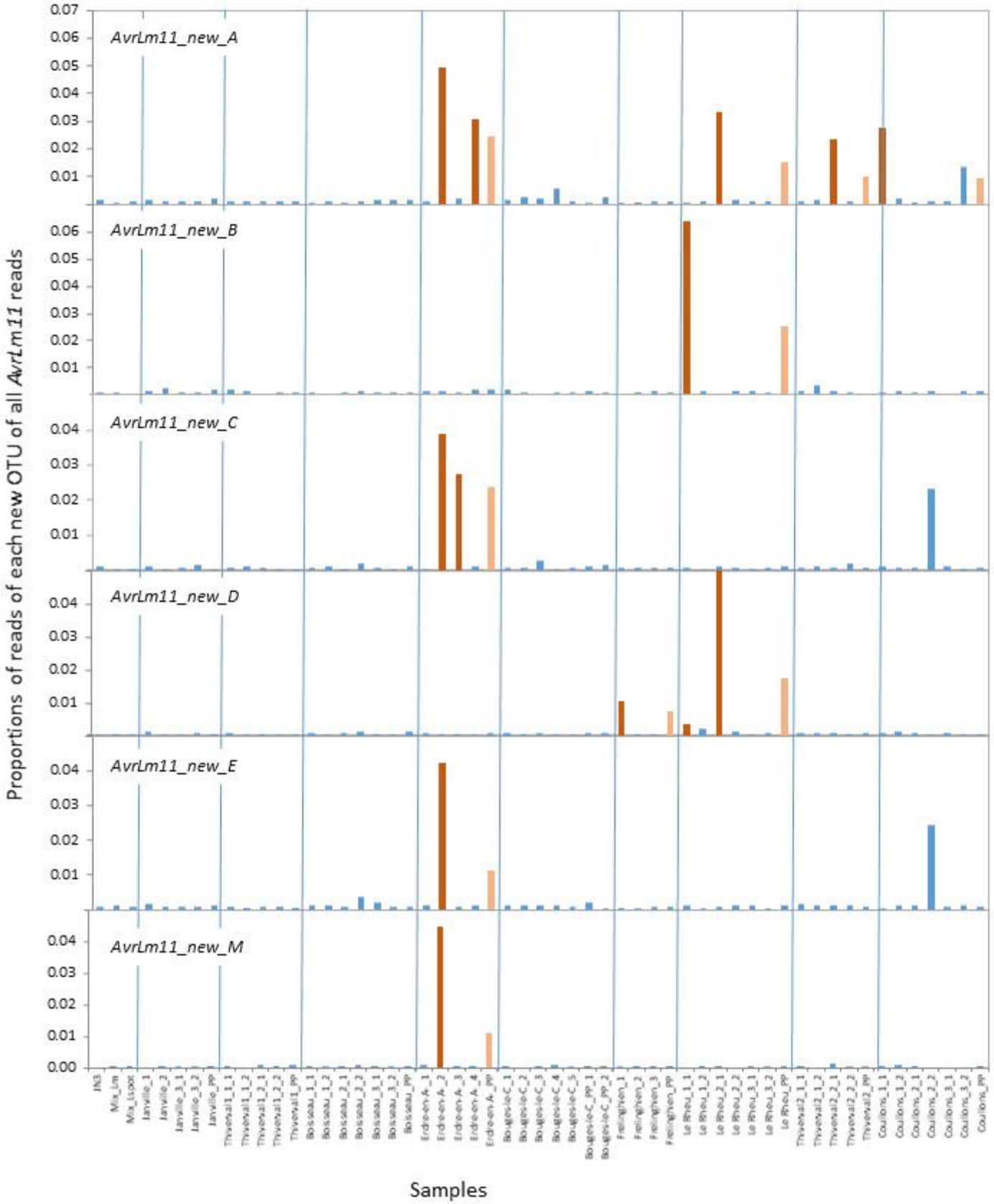

Figure 5. An example of detection of new alleles. Percentages of new OTUs over all *AvrLm11* OTUs are given for the three control samples (left) and the 52 environmental samples from nine sites, separated by the vertical blue lines. Bars in dark orange are those of pools of 32 leaf spots that are also in the "pools of pools "of 96 leaf spots (pale orange).

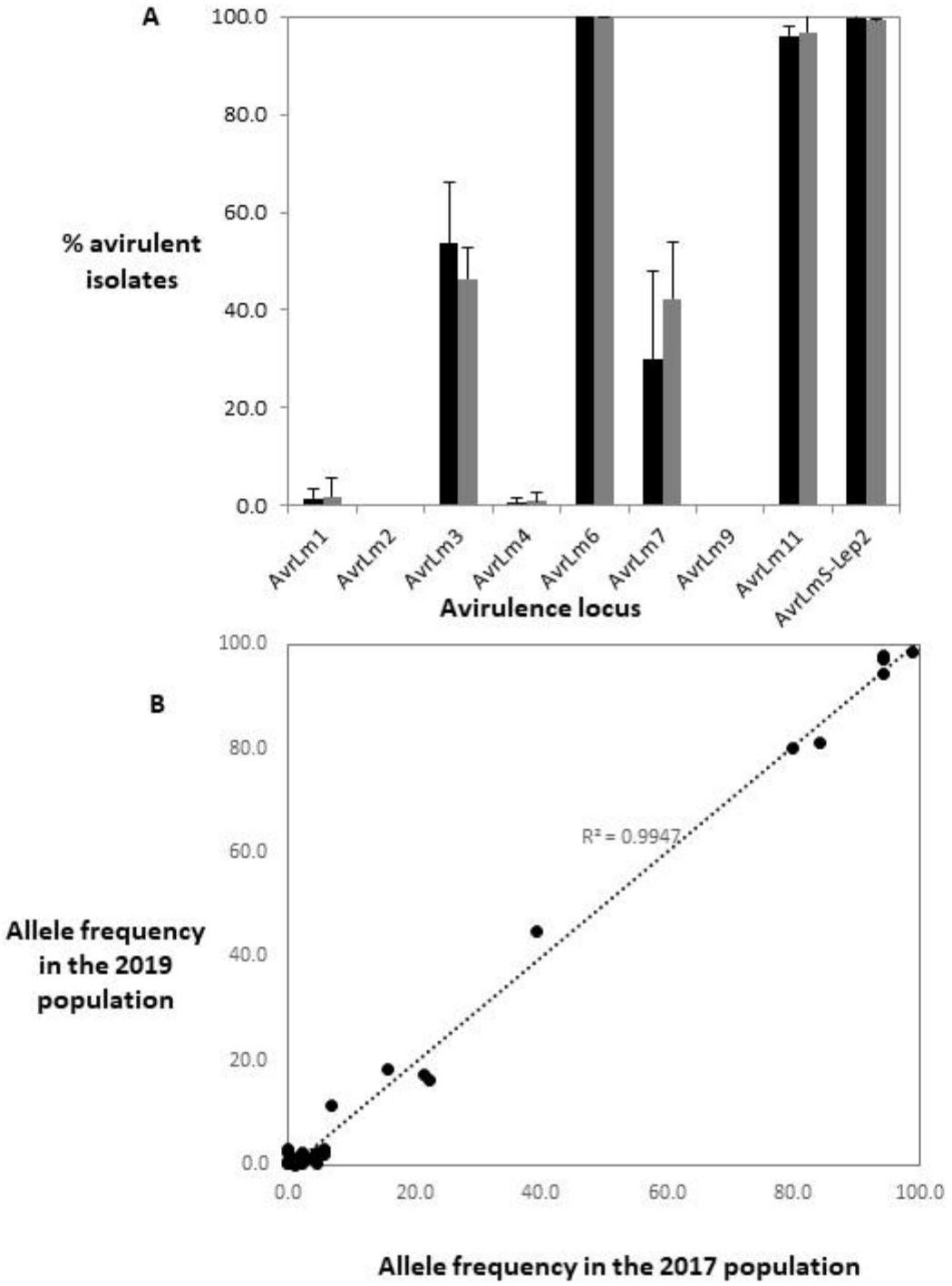

Figure 6. Validation of the MPSeqM tool on environmental samples. **A**, Comparison between percentages of avirulent isolates in French populations of *Leptosphaeria maculans,* estimated by phenotyping or genotyping. Black, data from phenotyping 690 isolates from 9 sites (Balesdent et al, 2023); grey, data from MultiplexPCR-Miseq analysis of 42 pools of leaf spots from the same sites (including a total of 1337 leaf spots). Values are means and standard errors over all sites. **B**, Comparison between allele frequencies (all genes) estimated by Sanger sequencing of individual isolates (2017 population, Gautier et al, 2013) and by the MPSeqM tool of pools of leaf spots (2019 population).