Sexual reproduction contributes to the evolution of resistance breaking isolates of the spinach pathogen *Peronospora effusa*

Petros Skiadas[1,2,*], Joël Klein[2,*], Thomas Quiroz Monnens[2], Joyce Elberse[2], Ronnie de Jonge[2], Guido Van den Ackerveken[2,**], Michael F. Seidl[1,**,#]

[1]Theoretical Biology and Bioinformatics, Utrecht University, Padualaan 8 3584 CH, Utrecht, The Netherlands

[2]Plant-Microbe Interactions, Utrecht University, Padualaan 8 3584 CH, Utrecht, The Netherlands

[*]these authors contributed equally

[**]these authors also contributed equally

[#]**corresponding author:** Padualaan 8, Room N611, 3584 CH Utrecht, the Netherlands, +31 30 253 3694, m.f.seidl@uu.nl

Keywords: *Peronospora effusa*, resistance breaking, oomycete, mitochondrial genome, structural rearrangements, phylogeny, asexual reproduction, sexual reproduction, inverted repeats

Abstract

Peronospora effusa causes downy mildew, the economically most important disease of cultivated spinach worldwide. To date, 19 *P. effusa* races have been denominated based on their capacity to break spinach resistances, but their genetic diversity and the evolutionary processes that contribute to race emergence are unknown. Here, we performed the first systematic analysis of *P. effusa* races showing that those emerge by both asexual and sexual reproduction. Specifically, we studied the diversity of 26 *P. effusa* isolates from 16 denominated races based on mitochondrial and nuclear comparative genomics. Mitochondrial genomes based on long-read sequencing coupled with diversity assessment based on short-read sequencing uncovered two mitochondrial haplogroups, each with distinct genome organization. Nuclear genome-wide comparisons of the 26 isolates revealed that ten isolates from six races could clearly be divided into three asexually evolving groups, in concordance with their mitochondrial phylogeny. The remaining isolates showed signals of reticulated evolution and discordance between nuclear and mitochondrial phylogenies, suggesting that these evolved through sexual reproduction. Increased understanding of this pathogen's reproductive modes will provide the framework for future studies into the molecular mechanisms underlying race emergence and into the *P. effusa*-spinach interaction, thus assisting in sustainable production of spinach through knowledge-driven resistance breeding.

Significance statement

Many microbial plant pathogens depend on the successful colonization of their hosts to complete their life cycle, thereby damaging food crops worldwide. The most effective way of disease control is to deploy genetic disease resistances. However, the extensive use of resistant crop varieties exerts strong selective pressure on microbial plant pathogens to adapt in order to escape resistance. Through yet unknown mechanisms, the spinach pathogen *Peronospora effusa* can rapidly break the resistance of newly introduced varieties, often within a single growing season. Thus, there is an urgent need to better understand the mechanisms driving adaptation in *P. effusa*. This information will lead the way to knowledge-driven resistance breeding. Here, we capture for the first time the genetic variation of 26 *P. effusa*, 16 of which can break a different combination of host resistances. We demonstrate that *P. effusa* isolates evolve by both asexual and sexual reproduction, and thereby provide the framework to study the molecular mechanisms of the interactions between *P. effusa* and spinach.

Introduction

Many microbial plant pathogens need to successfully colonize their hosts to complete their life cycle. Colonization leads to extensive damage and disease on the infected host plants (McMullan *et al.*, 2015). To mitigate these effects, plants have evolved defence mechanisms to resist pathogens, while in turn pathogens evolve to overcome the host immune system to successfully establish an infection (Cook *et al.*, 2015). In natural populations, host and pathogen populations are slowly coevolving and maintain diversity of resistance and virulence alleles (Barrett *et al.*, 2009; Möller and Stukenbrock, 2017). In the last decades, modern agricultural practices have introduced resistant host cultivars that are often deployed in monocultures (Miller *et al.*, 2020). This practice exerts strong evolutionary pressure on pathogen populations to overcome the introduced host resistance, and consequently pathogens are engaging in rapid evolutionary arms races (Möller and Stukenbrock, 2017; Mohd-Assaad *et al.*, 2019).

The oomycete downy mildew *Peronospora effusa* (*Pe*) is an obligate biotrophic pathogen of spinach, and economically the most important disease of cultivated spinach worldwide (Lyon *et al.*, 2016; Ribera *et al.*, 2020). Severe downy mildew infection can rapidly destroy entire spinach fields, and even minor disease symptoms may require removing infected leaves before harvest, thus significantly decreasing profitability (Correll *et al.*, 2011). In conventional spinach production, downy mildew is managed by the deployment of resistant cultivars and/or fungicide applications for short-term control (Koike *et al.*, 1992). However, spinach production nowadays is shifting to organic practices, and consequently development and deployment of resistant cultivars are the most practical management tool (Kandel *et al.*, 2020). Modern farming practices of high-density, year-round spinach production are an ideal environment for downy mildew, providing ‘green bridges’ that sustain pathogen survival (Lyon *et al.*, 2016). Due to the added pressure from newly developed resistant spinach cultivars, *P. effusa* evolves rapidly and repeatedly overcomes host resistance, resulting in new epidemics and severe economic losses (Ribera *et al.*, 2020).

Newly identified *P. effusa* isolates are assigned to different races based on their ability to infect a defined set of differential spinach lines developed by the International Seed Federation (Correll *et al.*, 2015). These spinach differentials contain one or more resistance (*R*) loci that are

76 known to effectively contain *P. effusa* races (Feng *et al.*, 2018). New races are determined by
77 the International Working Group on *Peronospora*, which annually evaluates newly identified *P.*
78 *effusa* isolates for their capacity to overcome previously known host resistances. Downy mildew
79 of spinach was first identified and reported in Britain in 1824 (Ribera *et al.*, 2020). More than a
80 century later, race 2 was described in 1958. Before 1990, only three races of the pathogen were
81 known, and the disease could be well controlled (Koike *et al.*, 1992). Likely driven by the
82 extensive deployment of resistance loci (*R*-loci) in commercial spinach varieties, after 1990 the
83 number of identified races increased tremendously, and 16 additional races have been
84 discovered within the last 20 years (Supplementary Figure 1) (Lyon *et al.*, 2016; Feng *et al.*,
85 2018; Klein *et al.*, 2020).

86 Like many oomycetes, *P. effusa* is a heterozygous diploid organism. Its genome sequence
87 is relatively small compared with other oomycetes (58.6 Mb compared with for instance
88 *Phytophthora infestans* which is 240 Mb (Haas *et al.*, 2009)), and it is organised in 17
89 chromosomes with 53.7% repeat content (Fletcher *et al.*, 2021). *P. effusa* can reproduce both
90 asexually and sexually. During the spinach growing season, *P. effusa* reproduces asexually on
91 infected plant tissues, releasing millions of asexual spores (Kandel *et al.*, 2020). These spores
92 are produced on leaves and other infected parts of the plant on sporangiophores that are visible
93 with the naked eye. By the end of the growing season, *P. effusa* will also reproduce sexually,
94 which happens when two *P. effusa* isolates of the opposite mating type (P1 and P2) co-occur in
95 the same infected spinach tissue. Sexual reproduction results in new combinations of parental
96 chromosomes, and this process has been thought to be a powerful driver of the emergence of
97 new *P. effusa* races (Feng *et al.*, 2018). However, *P. effusa* field isolates from various locations
98 in the USA displayed limited genotypic diversity suggesting that these isolates are primarily the
99 result of asexual reproduction. Nevertheless, genotypic diversity in historical isolates is likely
100 influenced by sexual recombination (Lyon *et al.*, 2016). The complete genome assembly of the
101 *P. effusa* isolate UA202013 has been compared with closely related isolates from races 12, 13,
102 and 14 with limited variation being observed (Fletcher *et al.*, 2021). Consequently, we still know
103 surprisingly little about genome-wide diversity between most of *P. effusa* isolates and races, as
104 well as evolutionary processes that contribute to the emergence of new *P. effusa* races.

105 Mitochondrial genomes are ideal markers to study oomycete taxonomy and evolution as the
106 mitochondrial genome has a five to ten times higher evolutionary rate compared to the nuclear
107 genome (Choi *et al.*, 2011; Yuan *et al.*, 2017; Bourret *et al.*, 2018). Previously, the mitochondrial
108 genes *cox2* and *nad1* have been used to shed light on the relationship between different *P.*
109 *effusa* isolates (Choi *et al.*, 2011), which uncovered two distinct mitochondrial haplotype groups
110 that could be linked to their geographical origin (Group I for Asia and Oceania and Group II for
111 American, European, and Japan) (Choi *et al.*, 2011). However, the mitochondrial phylogeny of
112 denominated *P. effusa* races is currently unknown. Here, by using long-read sequencing data,
113 we generated mitochondrial assemblies for four *P. effusa* isolates, including *Pe1* the first
114 officially denominated *P. effusa* race which was isolated in the United States (Lyon *et al.*, 2016).
115 These four isolates differ in their mitochondrial genome structure, which correlates with their
116 separation into two haplogroups. Combined with whole-genome short-read sequencing data,
117 we developed both mitochondrial and nuclear phylogenies of 26 *P. effusa* isolates. Comparisons
118 of both phylogenies revealed discordance and signatures of shared genetic material, which
119 suggest that pervasive sexual recombination contributed to the evolution of resistance-breaking
120 *P. effusa* isolates.

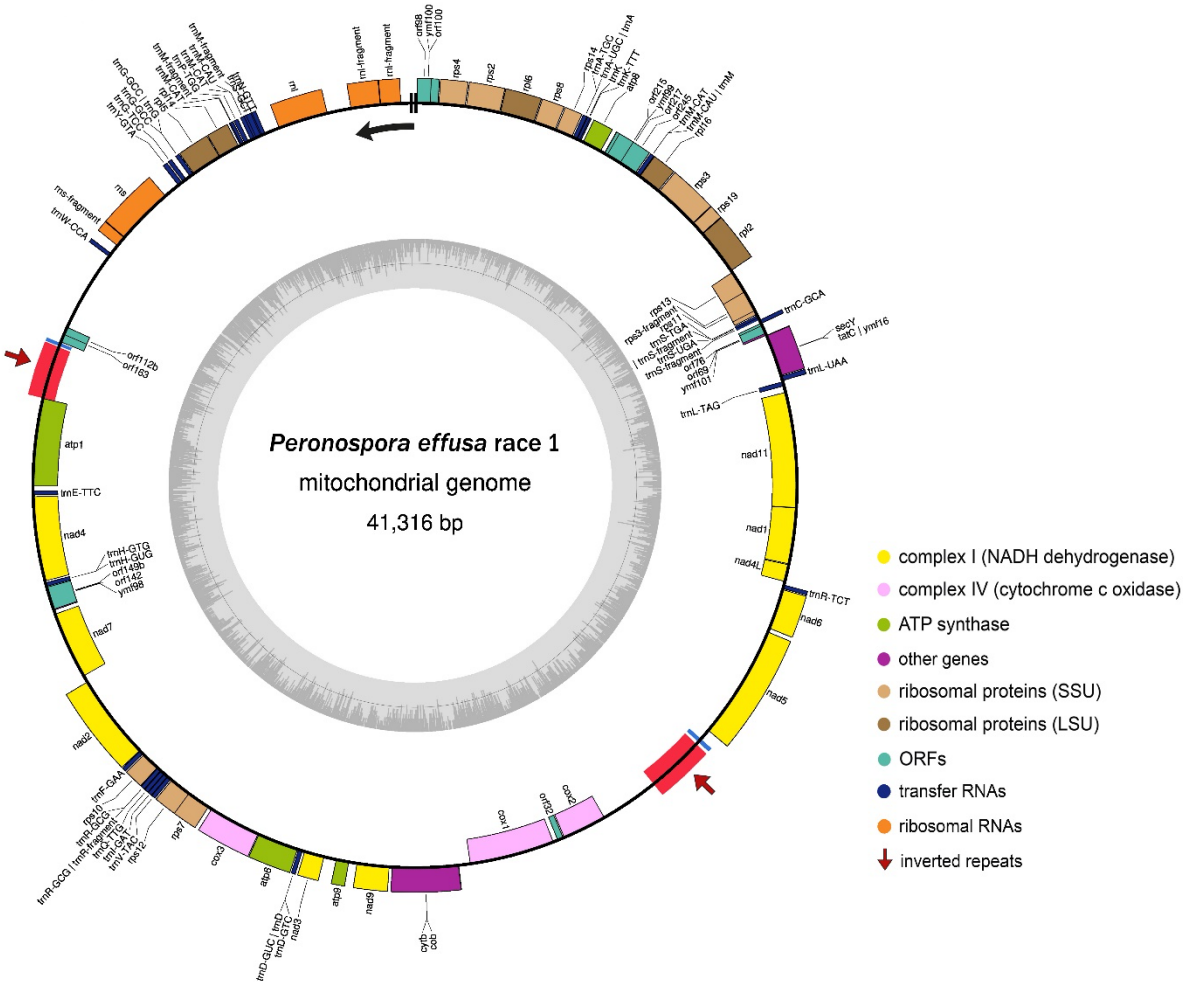
121 **Results**

122 ***De novo P. effusa* mitochondrial genome assembly**

123 To study the mitochondrial phylogeny of *P. effusa* isolates, we first sought to reconstruct the
124 complete mitochondrial genome of *Pe1*, the first denominated *P. effusa* race. We performed
125 whole-genome long-read sequencing using Oxford Nanopore sequencing technology on total
126 genomic DNA obtained from *Pe1* spores (Supplementary Table 1B). To facilitate the assembly
127 of the *Pe1* mitochondrial genome sequence, we recovered 19,627 reads based on sequence
128 similarity searches to a collection of 26 publicly available oomycete mitochondrial genome
129 sequences (Supplementary Table 1B; Supplementary Table 2). The sequencing reads were
130 error-corrected and assembled using Canu, yielding one linear contig of 49,964 bp long. We
131 manually curated this assembly to reconstruct a circular contig, which we further corrected for
132 sequencing errors with high-quality Illumina short reads (Supplementary Table 1B). The
133 resulting complete mitochondrial genome sequence of *Pe1* is a 41,316 bp circular molecule with
134 22.8% GC content, which is similar to mitochondrial genome assemblies previously reported for
135 other oomycetes (Yuan *et al.*, 2017; Fletcher *et al.*, 2018). Based on sequence homology
136 searches, we identified 42 protein-coding genes encoding 19 respiratory chain proteins, 16
137 ribosomal proteins, two protein transport proteins, and five hypothetical proteins also present in
138 other oomycetes; of these, 12 were annotated in the negative strand and 30 in the positive
139 strand (Figure 1, Supplementary Table 3). We observed that the majority of genes are
140 conserved across the 26 selected mitochondrial reference sequences. Notably, we also
141 identified three predicted genes (open reading frames; ORFs) without homologs in any of the
142 references including *R13* and *R14*, which is comparable to the five ORFs without homology
143 previously identified in *P. effusa* isolates *R13* and *R14* (Fletcher *et al.*, 2018) and to the six to
144 twelve genus-specific ORFs found in different *Phytophthora* species (Yuan *et al.*, 2017; Cai and
145 Scofield, 2020).

146 Mitochondrial genomes are known to contain inverted repeats (Achaz et al., 2003; Voineagu
147 et al., 2008), and thus we used blastn to self-align the genome and identified two 100% identical
148 832 bp-long inverted repeats in *Pe1* (Figure 1). Inverted repeats are challenging to resolve
149 correctly in genome assemblies when sequencing reads are not long enough to span the

150 repetitive sequence (Wang *et al.*, 2018). To corroborate the genome organization of the *Pe1*
151 mitochondrial assembly, we mapped both short- and long-read sequencing data of *Pe1* to the
152 assembly, which did not reveal any discordant alignments (Supplementary Figure 2), suggesting
153 that the overall structure of the *Pe1* assembly is correct.



154

155 **Figure 1. Circular representation of the mitochondrial genome assembly of *Peronospora effusa***
156 **race 1 (*Pe1*).** Protein-coding genes, tRNA, rRNA, and other open-reading frames (ORFs) are shown
157 along the outer ring (positive strand is the outside of the ring and the negative strand is the inside). The
158 two inverted repeats are highlighted by red boxes and with arrows. The inner ring depicts the GC content.
159 The start and end of the linear representations (Figure 2B, Figure 3B) of the circular genome assembly is
160 indicated with two black lines at the start of the black arrow, which indicates the direction.

161 **Structural rearrangement in the *Peronospora effusa race 1* mitochondrial genome**

162 Structural rearrangements in mitochondrial genomes of oomycetes are common (Yuan *et al.*,
163 2017). We sought to investigate the mitochondrial genome structures in other Peronosporaceae

164 and compare it to *Pe1*. To this end, we used the publicly annotated mitochondrial genome
165 assemblies of 15 closely related species, which belong to the Peronosporaceae family, together
166 with the *Pe1* and *R13* assembly. We first reconstructed the phylogenetic relationship between
167 the 17 Peronosporaceae based on the mitochondrial genes *cox2* and *nad1*, which are two highly
168 conserved genes that have been previously used as taxonomic markers to study the
169 relationships of oomycetes including *P. effusa* (Choi *et al.*, 2011). The phylogenetic tree based
170 on concatenated *cox2* and *nad1* alignments recovers the known relationship between
171 Peronosporaceae by clearly separating the different genera (Figure 2A).

172 Guided by their phylogenetic relationship, we reconstructed gene-based alignments between
173 these 17 mitochondrial genomes. We observed six independent large-scale structural
174 rearrangements across the 17 Peronosporaceae mitochondrial genomes, with *Pe1* having a
175 unique structure that is different from *R13*; *R13* is collinear with *R14* and the publicly available
176 assembly of the closest relative, the tobacco pathogen *Peronospora tabacina* (Fletcher *et al.*,
177 2018). All the observed rearrangements can be summarized by displaying the differences in just
178 ten species (Figure 2B); three of the observed structural rearrangements (*i*, *ii*, and *iv*) are
179 conserved across multiple species, while three are unique to the species in which they were
180 identified (*iii*, *v*, and *vi*). Two rearrangements are small (2.5 Kb) and involve only two genes (*ii*,
181 *iii*), while the remaining four are larger, 8-16 Kb, and involve 11-15 genes. We found no evidence
182 of the same region being rearranged more than once, although the regions of the
183 rearrangements *iv* and *vi* fully overlap with those of *i* and *ii*. Thus, our data suggest that structural
184 rearrangements are common in Peronosporaceae mitochondrial genomes, yet the organisation
185 of mitochondrial genomes is typically conserved in closely related species. Like in *P. effusa*, we
186 observed inverted repeats in many Peronosporaceae mitochondrial genomes. Their localization
187 and size vary considerably between species, and they are often absent in *Phytophthora* spp.
188 (Figure 2C), implying that they can be gained or lost multiple times independently during
189 evolution. Notably, inverted repeats directly flank two of the unique rearrangements (*v* and *vi*),
190 suggesting that these structures recombine to result in inversions in Peronosporaceae (Yuan *et*
191 *al.*, 2017).

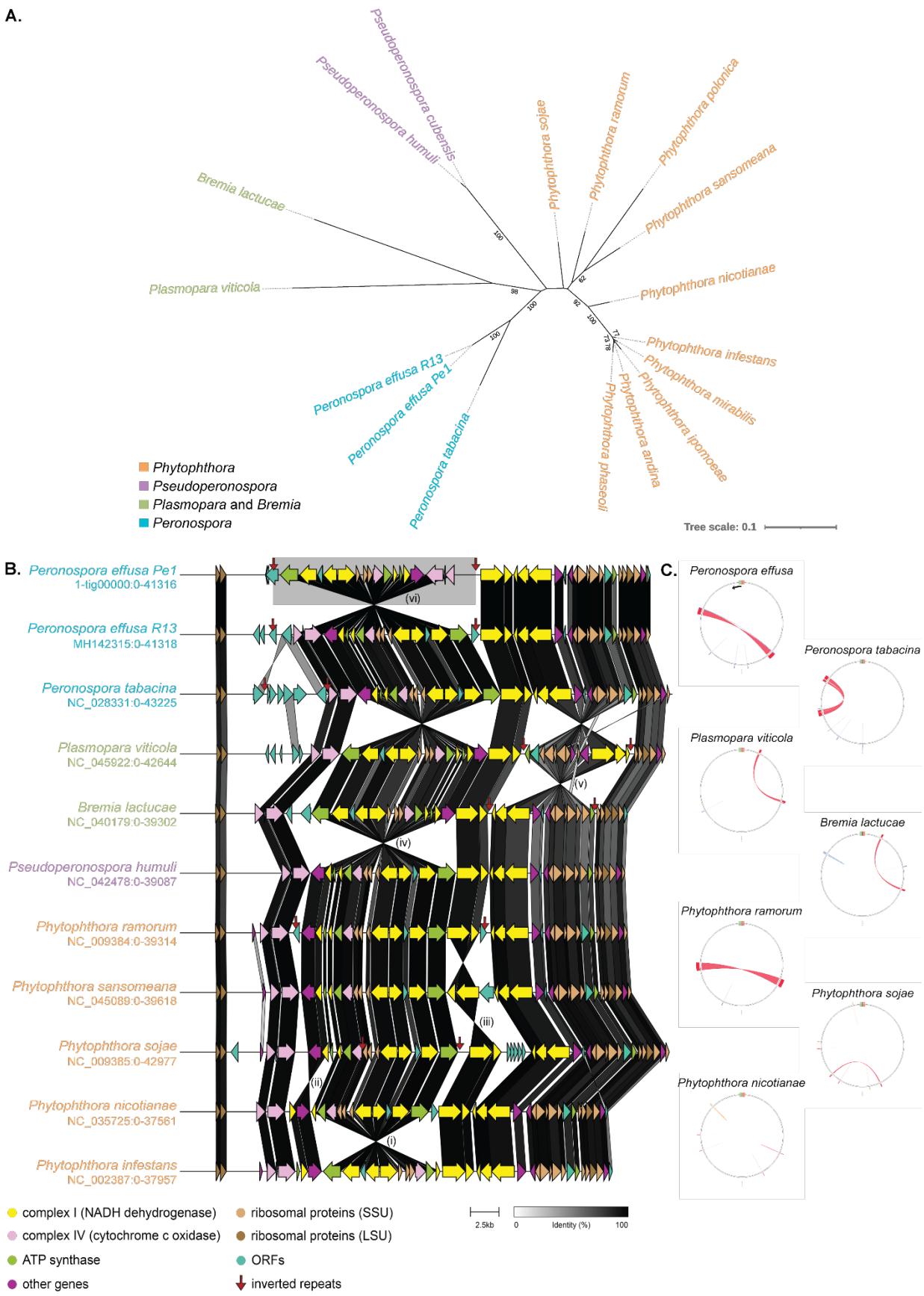


Figure 2. Mitochondrial genome assemblies of ten Peronosporaceae species reveal structural rearrangements that are often flanked by inverted repeats. A) Unrooted maximum-likelihood phylogeny of 17 different Peronosporaceae mitochondria (16 species) based on the sequences of the

196 mitochondrial *cox2* and *nad1* genes. **B)** Linear representation of mitochondrial gene order is shown for
197 ten Peronosporaceae, representative of the 16 species in the phylogenetic tree. Genes are coloured
198 based on their functions and they are connected by ribbons shaded based on the percentage of protein
199 identity (from white to black). The relative positions of the inverted repeats are shown by red arrows and
200 the structural rearrangements are numbered (*i-vi*). **C)** Comparison of the inverted repeats in
201 Peronosporaceae. The ribbons display the alignment scores, reflecting the sequence similarity
202 (percentage): red (>75%), orange (50-75%), and blue (25%). The respective start-end of the linear
203 representation of the circular genomes is indicated with the green/orange stripes, with the arrow indicating
204 the direction.

205 **Mitochondrial genome phylogeny separates *Peronospora effusa* into two distinct groups**

206 Phylogenetic analysis of the global *P. effusa* population based on the mitochondrial genes *cox2*
207 and *nad1* has previously uncovered two distinct mitochondrial haplogroups that can be linked to
208 the geographic origin of the individual isolates (Choi *et al.*, 2011). To study the phylogenetic
209 relationship between the denominated *P. effusa* races, we obtained spore material for 24 *P.*
210 *effusa* isolates. The pathogenicity of the isolates is determined based on the current standard
211 set of international spinach differentials (Supplementary Figure 1). To discover the evolutionary
212 relationship of the 24 *P. effusa* isolates, we generated on average 58.8 million Illumina paired-
213 end sequencing reads, which were subsequently mapped to the *Pe1* mitochondrial assembly to
214 determine sequence variants (Supplementary Table 1A).

215 We initially focused on the mitochondrial specific genes *cox2* and *nad1* as this enabled us to
216 study our 24 isolates together with the previously analysed 33 isolates from the global population
217 (Choi *et al.*, 2011) as well as *R13* and *R14* (Fletcher *et al.*, 2018). The alignment as well as
218 phylogenetic analyses of these genes reveal that *P. effusa* isolates can be divided into two
219 distinct mitochondrial haplogroups (Supplementary Figure 3), thereby corroborating previous
220 results by Choi and colleagues (Choi *et al.*, 2011). Based on Choi and colleagues, European
221 and North American isolates belong to Haplogroup II, while Asian and Australian isolates belong
222 to Haplogroup I (Choi *et al.*, 2011). However, we observed that multiple isolates denominated
223 to Haplogroup I were isolated in the USA (*Pe5*, *Pe4*, *Pe16*, *US-13a*), and thus we could not

224 establish the same correlation between the mitochondrial haplogroups and the geographical
225 location of the *P. effusa* isolates.

226 To further support the separation into two haplogroups, we performed phylogenetic analyses
227 based on the whole mitochondrial genomes of the 26 isolates (16 denominated races, eight
228 pathotypes, as well as *R13* and *R14*). We could identify 85 variant sites (76 SNPs, 9 indels)
229 across the mitochondrial genome in the 26 isolates based on the comparison to the *Pe1*
230 reference mitochondrial assembly (Supplementary Table 1A). Of those, 48 variants were
231 identified in protein-coding regions (two in ORFs that lack similarity to known protein-coding
232 genes). Based on these 85 variants, we constructed a maximum-likelihood phylogeny, which
233 similar to *cox2* and *nad1* supports a separation into two distinct haplogroups (separated by 46
234 of the 85 variants, and 100% bootstrap support); isolates within Haplogroup I differ by 21 variant
235 sites and within Haplogroup II by 18 variant sites (Figure 3A).

236 Based on the discovery of the novel structure of the *Pe1* mitochondrial genome (Figure 2B),
237 we sought to further investigate the *P. effusa* mitochondrial genome organization in the two
238 haplogroups. To this end, we selected three additional *P. effusa* isolates for Nanopore long-read
239 sequencing: *Pe4* and *Pe6* from Haplogroup I and *Pe11* from Haplogroup II (Supplementary
240 Table 1B). We recovered ~153,943 reads with an average length of 6.8 Kb based on sequence
241 similarity searches to a collection of 26 publicly available oomycete mitochondrial genome
242 sequences (Supplementary Table 2). We assembled the long reads with Canu and manually
243 corrected the assemblies into circular genome sequences of 41.4 Kb on average
244 (Supplementary Figure 4). Like in *Pe1*, these assemblies encode 43 protein-coding genes and
245 two 100% identical, 832 bp long inverted repeats. Despite these similarities, the gene-based
246 alignment of the four *P. effusa* mitochondrial genomes revealed a distinct genome structure for
247 each haplogroup that is characterized by an inversion at the inverted repeats (Figure 3B). An
248 ORF predicted on one edge of this structural rearrangement in *Pe1* and *Pe11* is missing from
249 *Pe4* and *Pe6*. Notably, the same 16 Kb long structural rearrangement that we previously
250 described between *Pe1* and *P. tabacina* (Figure 2B) also differentiates the two *P. effusa*
251 haplogroups, suggesting that the structural rearrangement was introduced only in the *P. effusa*
252 Haplogroup II.

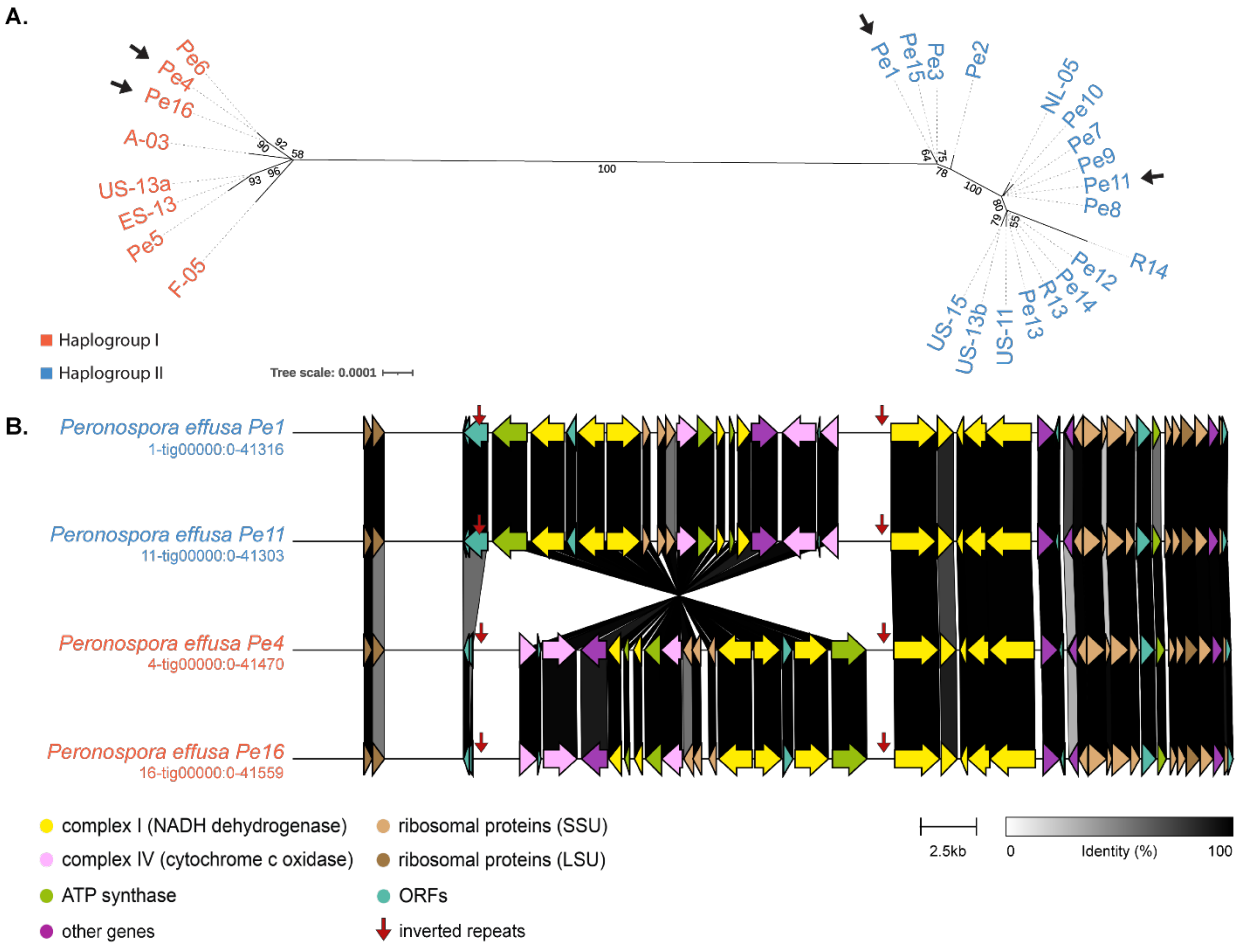


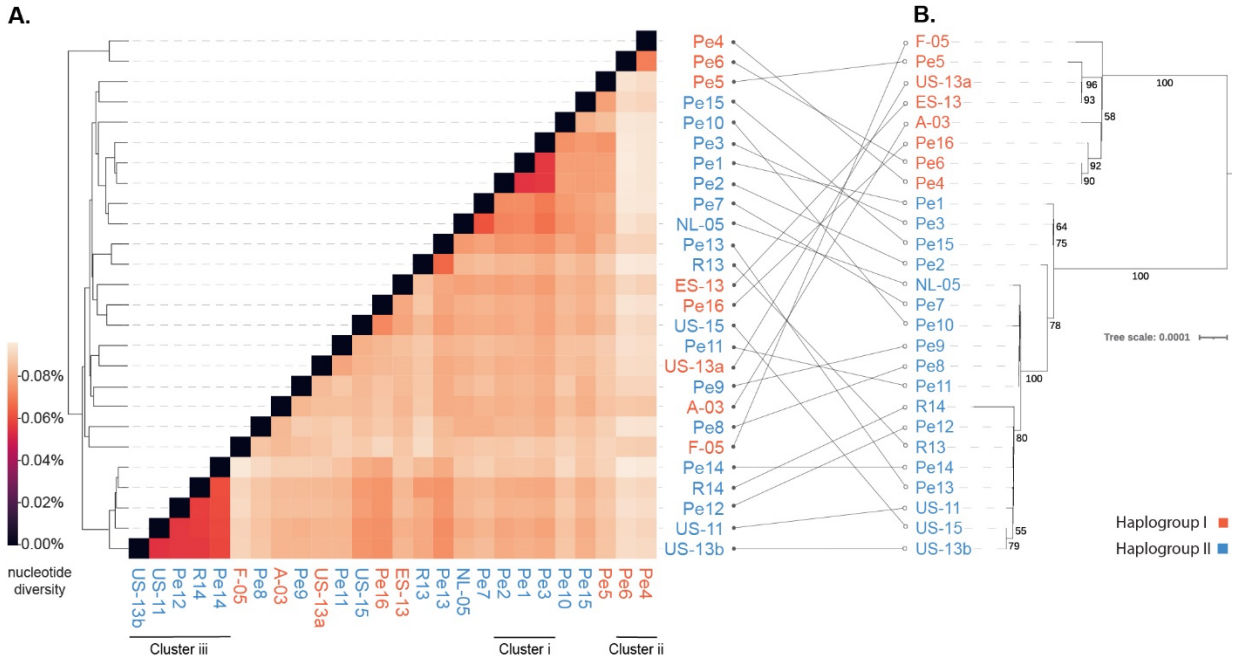
Figure 3. Comparison of mitochondrial genomes between *Peronospora effusa (Pe)* isolates reveals two distinct groups. A) Unrooted maximum-likelihood phylogeny based on the complete mitochondrial genome sequences of 26 *P. effusa* isolates clearly reveals two haplogroups (I and II). The black arrows point to the four *P. effusa* isolates, which were selected for Nanopore long-read sequencing, mitochondrial genome assembly, and gene annotation (*Pe1*, *Pe4*, *Pe6*, and *Pe11*). **B)** Mitochondrial gene orders of the selected four *P. effusa* isolates. Genes are coloured based on their functions and they are connected by ribbons coloured based on the percentage of protein identity (from white to black). The relative positions of the inverted repeats are shown by the red arrows.

262 **Nuclear and mitochondrial genome phylogenies of *Peronospora effusa* isolates are**
263 **discordant**

264 *P. effusa* reproduces both sexually and asexually (Kandel *et al.*, 2020), and since the
265 mitochondria are typically inherited from only one parental strain, mitochondrial variation alone
266 cannot fully reveal the relationship and diversity of *P. effusa* isolates (Bourret *et al.*, 2018). To
267 further investigate the relationship between *P. effusa* isolates, we performed in-depth
268 comparisons of mitochondrial and nuclear genome variation of the 26 *P. effusa* isolates used in
269 this study. Using Illumina short-read data, we performed variant calling with GATK against the
270 public nuclear genome of *Pe1* (Klein *et al.*, 2020). In total, we identified 314,276 multiallelic
271 variant sites that can be separated in 280,750 SNPs and 35,677 indels. On average, 35.8% of
272 the variants per isolate are homozygous and 64.2% are heterozygous. Apart from *Pe1*, on
273 average 1.53% of the variants per isolate are unique, 1.75% are shared by only two, and 19.67%
274 are shared between all isolates (Supplementary Figure 5; Supplementary Table 1A). To uncover
275 the evolutionary relationship of the 26 *P. effusa* isolates, we used 260,616 biallelic SNPs to
276 determine the pairwise nuclear genome nucleotide diversity for each isolate, as a percentage of
277 difference over the entire genome. The average nucleotide diversity (nd) between all isolates
278 was 0.083% (0.008% standard deviation), with the closest related isolates being *US-11* and *US-*
279 *13b* (0.055% nd), and the most distant isolates being *Pe6* and *Pe14* (0.097% nd). Based on
280 these observations, we assigned three genomic clusters of closely related isolates that have
281 been denominated to different races: Cluster i contains *Pe1*, *Pe2*, and *Pe3* (0.056% average
282 nd); Cluster ii contains *Pe4* and *Pe6* (0.071% nd); and Cluster iii contains *Pe12*, *Pe14*, *R14*, *US-*
283 *11*, and *US-13b* (0.058% average nd). We observed high similarity between additional isolates,
284 for example *R13* and *Pe13* (0.066% nd), which were not assigned into separate clusters
285 because they are designated to the same race (Figure 4A).

286 To further explore the relationships between the isolates, we directly compared their similarity
287 based on the nuclear variation with that identified from the mitochondrial phylogeny (Figure 4B).
288 We observed that the ten isolates that cluster based on genome-wide nuclear similarity
289 (>0.073% nd) also clearly cluster based on the mitochondrial phylogeny and belong to the same
290 mitochondrial haplogroup (Figure 4B). For example, all isolates assigned to Cluster i (*Pe1*, *Pe2*,

291 and *Pe3*) belong to Haplogroup II, while all the isolates assigned to Cluster ii (*Pe4* and *Pe6*)
292 belong to Haplogroup I. By contrast, we observed clear discordance between nuclear clustering
293 and the mitochondrial phylogeny for the 16 remaining isolates (Figure 4B). For example, isolates
294 with high genome-wide nucleotide similarity such as *US-15* and *Pe16* (0.073% nd), *Pe5* and
295 *Pe15* (0.078% nd), or *Pe11* and *US-13a* (0.079% nd) belong to different mitochondrial
296 haplogroups.



297

298 **Figure 4. Nuclear and mitochondrial variation discordance in *Peronospora effusa (Pe)*. A)**
299 Hierarchically clustered heatmap of pairwise nuclear genome nucleotide diversity. Three distinct clusters
300 formed by related isolates belonging to different races are highlighted. **B)** From the right-to the left,
301 maximum-likelihood phylogeny based on the mitochondrial genome sequences of the 26 *P. effusa*
302 isolates, rooted on the beet downy mildew (*Peronospora farinosa* f. sp. *betae, ES-15*). Concordance
303 between this phylogeny and the hierarchical clustering shown in A) is highlighted by connecting lines
304 between isolates.

305 **Mitochondrial and nuclear variation suggests that sexual recombination contributes to**
306 **the evolution of resistance-breaking *Peronospora effusa* isolates**

307 To explain the discordance between nuclear clustering and mitochondrial phylogeny of *P.*
308 *effusa*, we further explored the genetic relatedness between *P. effusa* isolates. Based on the
309 previously generated diversity matrix (260,616 biallelic SNPs), we explored the genetic

310 differences by principal component analysis (PCA); the first principal component explains
311 14.14% of the observed variation and the second 13.22%, while the third principal component
312 only explains 7.01%. When visualizing the first two principal components, we observed the same
313 three distinct clusters that were previously recovered based on the genome-wide nuclear
314 diversity (Figure 5A). The remaining isolates are localized in the middle of the PCA, sharing
315 genetic material but did not clearly cluster in a distinct group.

316 To better visualise the complex phylogenetic relationship of *P. effusa* isolates, we performed
317 neighbor-net phylogenetic network analyses (260,616 biallelic SNPs), which is well-suited to
318 illustrate the complex evolutionary relationships among different isolates in the presence of
319 phylogenetic conflicts caused by, for instance, sexual recombination, hybridization, or horizontal
320 gene transfer (Galtier *et al.*, 2009). The neighbor-net phylogenetic network shows clear evidence
321 of phylogenetic conflicts (Figure 5B), for instance between *Pe13* and *Pe5* that localized close to
322 the centre of the network. *P. effusa* is known to reproduce sexually (Kandel *et al.*, 2020), and
323 thus, the most likely explanation for phylogenetic conflicts observed in our data is sexual
324 recombination, which is further supported by the PHI test for recombination (260,616 biallelic
325 SNPs, $P < 0.0001$) (Bruen *et al.*, 2006). In contrast, the above defined three *P. effusa* nuclear
326 clusters are connected to the main network with relatively long branches, suggesting that they
327 accumulated a substantial number of cluster-specific variants (e.g., *Pe4* and *Pe6* uniquely share
328 5,029 (3.5%) variants), and thus, at least for some time, these isolates likely evolved without
329 recombination with isolates outside of the respective cluster.

330 To further examine the exchange of genetic material between these isolates, we analysed
331 the variation of *P. effusa* isolates using STRUCTURE, a population structure inference approach
332 that uses Bayesian clustering based on allele frequencies to assign isolates to a pre-defined
333 number of populations (K) (Linck and Battey, 2019). We performed STRUCTURE analysis with
334 two, three, and four predefined sub-populations (K=2, K=3, K=4) (Figure 5C); any analysis with
335 K>4 had identical results with K=4. When assuming two sub-populations (K=2), isolates
336 belonging to both Cluster i and Cluster ii, and belonging to both haplogroups are present in a
337 single sub-population (purple), thus their variation is not fully explained. When the isolates are
338 divided into three and four sub-populations (K=3, K=4), the same three distinct clusters (i, ii, iii)

339 that were previously identified by phylogenetic analysis were also assigned into three distinct
340 sub-populations (purple, yellow, green) (Figure 4 and 5). This division in these three distinct sub-
341 populations further confirms that the isolates in each cluster did likely reproduce asexually. In
342 the case of K=3, all isolates belonging to the mitochondrial Haplogroup I are either a part of
343 nuclear Cluster ii (yellow) or a recombination of nuclear Cluster ii with other clusters. For K=4,
344 isolates that previously shared genetic material are now assigned to a distinct sub-population
345 (red), which contains isolates belonging to both mitochondrial haplogroups. The presence of
346 both mitochondrial haplogroups suggests that the isolates assigned to the red sub-population
347 but have different virulence spectrum, do not have a common origin, but have likely evolved
348 after one or more sexual recombination events between Cluster i and ii isolates. For example,
349 for K=3 we observed that *Pe5* contains genetic information that is partially shared with Cluster i
350 and ii (purple and yellow). We similarly observed genetic combinations of Clusters i, ii, and iii for
351 some of the isolates (Figure 5C). For example, for K=3, *Pe13* is represented as a mix of the
352 three defined sub-populations. This pattern could have emerged from multiple sexual
353 recombination events involving different isolates of all three defined genetic sub-populations.
354 We also observed a large variation in the number of unique variants for each isolate
355 (Supplementary Figure 5). For *F-05*, for example, we observed 3,630 (2.79%) unique variants,
356 which suggests that this isolate likely evolved asexually after the possible sexual recombination
357 event. In contrast, we only observed 910 (0.73%) unique variants in *Pe11*, which suggests more
358 recent sexual recombination events.

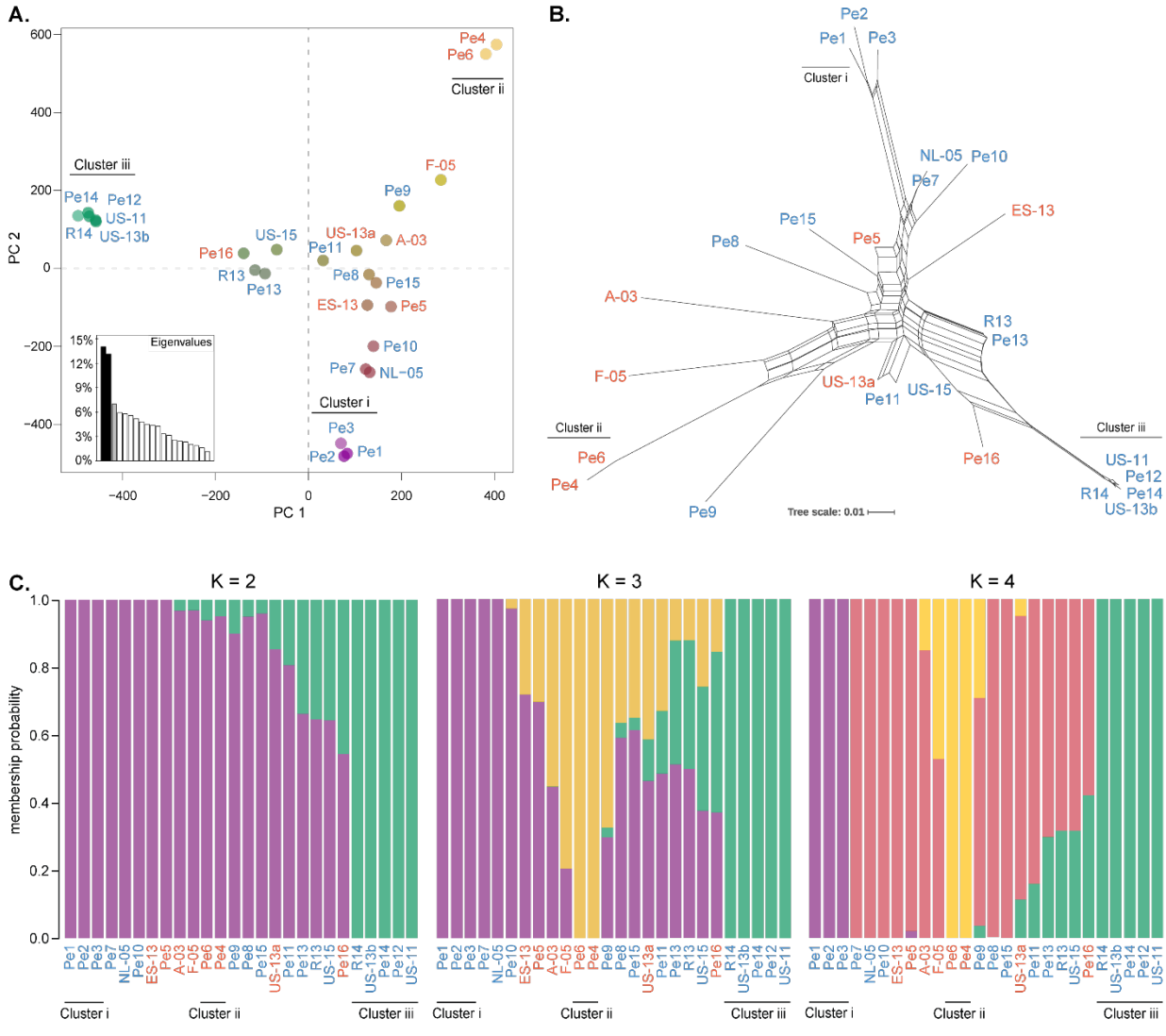


Figure 5. *Peronospora effusa* evolves by both asexual propagation and sexual recombination. Visualisations are constructed based on a distance matrix of genome-wide nucleotide differences, containing 260,616 biallelic sites for A and B and 169,016, minor allele frequency-filtered biallelic sites for C. The isolates are coloured based on their mitochondrial haplogroup (Haplogroup I in orange and Haplogroup II in blue), and the previously identified three distinct clusters of closely related isolates belonging to different races are highlighted and coloured similarly between A and C (Figure 4A). **A)** Principal component analysis; the axes depict the first two principal components that collectively explain 27.36% of the observed variation. **B)** Neighbor-net phylogenetic network; the branch lengths are proportional to the calculated number of substitutions per site. The parallel edges connecting different isolates indicate conflicting phylogenetic signals. **C)** STRUCTURE analysis for two, three, and four genetic sub-populations (K), without any prior assumptions for the grouping of isolates. Vertical bars represent *P. effusa* isolates, and the proportion of the coloured segments indicates the probability for the isolate to belong to *K* genetic sub-populations.

373 **Discussion**

374 In the last two decades, new races of the spinach downy mildew *Peronospora effusa* (*Pe*) have
375 been rapidly emerging, and thus far 19 *P. effusa* races have been denominated; the two latest
376 ones have been added in 2021 (Correll and Smilde, 2021). However, we know little about the
377 processes contributing to the diversity and evolution of different *P. effusa* races. To better
378 understand the molecular differences between *P. effusa* races, we here described the
379 mitochondrial and nuclear relationships between 16 denominated races and eight pathotypes of
380 *P. effusa*; for the most recent races (*Pe* 17-19) no genomic data was yet available. Based on
381 mitochondrial data, *P. effusa* isolates can be divided into two clear groups based on the
382 sequence and structure. Based on nuclear data, we observed that a subset of isolates could be
383 described by three distinct groups, which are in concordance with their mitochondrial phylogeny,
384 in turn, suggesting that these isolates likely evolved through asexual reproduction. The
385 remaining 16 isolates display discordance between nuclear and mitochondrial clustering as well
386 as signs of shared genetic material, which uncovers that also sexual reproduction contributed
387 to the emergence of *P. effusa* races.

388 Newly found *P. effusa* isolates are assigned to races based on their ability to infect a defined
389 set of differential spinach lines (Supplementary Figure 1). The phenotyping and maintenance of
390 *P. effusa* races, and the development of differential spinach lines is handled by the International
391 Seed Federation (Correll *et al.*, 2015) (Supplementary Table 4). The 24 isolates sequenced in
392 this study have been carefully maintained and screened based on their phenotypes since their
393 date of isolation, to avoid mixtures of multiple races (Feng et al., 2014). As it has been shown
394 in other obligate biotrophic plant pathogens, what is described as an isolate could be a
395 population of genetically distinct isolates with identical phenotypes, rather than a single
396 genetically distinct isolate (Barsoum *et al.*, 2020). However, based on our genome sequencing
397 data, we are unable to find any indication of mixed genotypes in a single isolate. Based on our
398 variant calling analysis against the mitochondrial and nuclear genome of *Pe1*, we did not
399 observe any multiallelic sites for individual isolates, i.e., more than one allele for mitochondria
400 and more than two alleles for nuclear genomes. In the same analysis of nuclear genome
401 variants, we did not observe a significant deviation in allele frequency from the expected 50:50

402 of a diploid organism. Similarly, comparisons of isolates with the same phenotype (*Pe13* with
403 *R13*, and *Pe14* with *R14*) shows that they have almost identical genotypes (0.066% and 0.061%
404 nucleotide diversity, respectively), despite being isolated and sequenced independently
405 (Fletcher *et al.*, 2018). Thus, we conclude that our data does not suggest that any of our isolates
406 is in fact a mixture of different genotypes.

407 Inverted repeats are known to form cruciform DNA structures that can cause chromosomal
408 instabilities, leading to the formation of structural rearrangements (Achaz *et al.*, 2003; Voineagu
409 *et al.*, 2008). Here, we observed inverted repeats to be associated with mitochondrial genome
410 rearrangements in *P. effusa* as well as throughout related Peronosporaceae. Inverted repeats
411 are genomic regions that are particularly challenging to assemble correctly when sequencing
412 reads are not long enough to span these repeats (Wang *et al.*, 2018). Inverted repeats in the *P.*
413 *effusa* mitochondrial genome are 832 bp in size, making it impossible for short-read data to
414 faithfully resolve mitochondrial chromosome structures. The *R13* and *R14* mitochondrial
415 genomes display the structure of Haplogroup I isolates, even though these isolates are clearly
416 assigned to Haplogroup II based on mitochondrial sequence variation. This discrepancy could
417 be due to the short-read sequencing data used to reconstruct the *R13* and *R14* mitochondrial
418 sequences (Fletcher *et al.*, 2018). By contrast, we utilised long-read sequencing data to
419 assemble mitochondrial genomes. The long reads span the inverted repeats, as corroborated
420 by uniform read coverage throughout the mitochondrial sequence, which suggests that our
421 mitochondrial assemblies of *Pe1*, *Pe4*, *Pe11*, and *Pe16* are accurate. The structural
422 rearrangement that was initially observed in *Pe1* separates the two mitochondrial haplogroups
423 of *P. effusa*. This rearrangement is not present in any queried oomycete, suggesting it is of
424 relatively recent origin. Similarly, we can assume that the unique structural rearrangement in
425 *Plasmopara viticola* is also formed by a relatively recent event. The DNA of these two recent
426 rearrangements are flanked by inverted repeats, thus suggesting that genome instability caused
427 by recombination of inverted repeats can drive mitochondrial diversity in Peronosporaceae.
428 Because of this instability, we can foresee that rearrangements are likely to reoccur in this
429 region. Thus, the inverted repeats itself would be a poor marker to robustly separate the defined

430 haplogroups. To distinguish the mitochondrial haplogroups the inverted repeats could be
431 augmented by the identified haplotype-specific, single-nucleotide variants.

432 The mitochondrial genes *cox2* and *nad1* have been previously used to separate *P. effusa*
433 isolates into two district haplogroups that were found to be associated with geographical origin
434 of the isolates tested (Choi *et al.*, 2011). Here, we used 24 *P. effusa* isolates and robustly
435 recovered these two distinct haplogroups based on *cox2* and *nad1*, whole-mitochondrial
436 genome alignment, and structural variation. Even though most *P. effusa* isolates that were
437 isolated outside the USA can be found in Haplogroup I, we were unable to establish a clear
438 correlation between the mitochondrial haplogroup and the geographical origin of the isolates.
439 Isolates with similar virulence spectrum are often simultaneously found at different geographical
440 locations (Irish *et al.*, 2003; Satou *et al.*, 2006), which is likely due to the globalized spinach
441 market where new spinach varieties are distributed and sold worldwide. Possibly, the sexually
442 produced oospores of *P. effusa* that are abundant in spinach fields drive the repeated
443 emergence of new races (Dhillon *et al.*, 2020), followed by a rapid spread over a large area.
444 Therefore, it is difficult to determine where a novel resistance-breaking isolate first occurred,
445 and in our case, to know the exact geographical origin of the denominated *P. effusa* races.

446 *P. effusa* can reproduce both sexually and asexually (Kandel *et al.*, 2020). It has been
447 suggested that sexual reproduction is the main mechanism behind increased adaptability, while
448 asexual reproduction allows the rapid propagation of successful genotypes thereby causing
449 major disease epidemics (Drenth *et al.*, 2019). As a result, sampling in the field during an
450 epidemic will be greatly biased towards isolates that have been asexually reproducing,
451 obscuring the role of sexual reproduction. Consequently, most *P. effusa* field isolates at a certain
452 time and geographic location are the result of asexual reproduction (Lyon *et al.*, 2016). Asexual
453 reproduction can introduce significant phenotypic changes in resistant breaking *P. effusa*
454 isolates (e.g., *Pe4* and *Pe6, Pe12 and Pe14*). Various mechanisms for asexual evolution have
455 been documented in other oomycete and fungal plant pathogens. In *Phytophthora infestans*
456 asexually evolving populations increase their genotypic variation by increase in ploidy and in
457 gene copy number (Cooke *et al.*, 2012; Knaus *et al.*, 2020). Additionally, phenotypic differences
458 can result from changes in gene expression (Cooke *et al.*, 2012; Depotter *et al.*, 2021). The

459 forest pathogen *Phytophthora ramorum* has only been observed to reproduce asexually, with
460 extensive mitotic recombination and gene copy number variation driving the genotypic diversity
461 (Dale *et al.*, 2019). In fungi it has been suggested that high levels of genotypic diversity can be
462 generated by asexual reproduction through large structural variation (Seidl and Thomma, 2014;
463 McDonald and Stukenbrock, 2016). For example, different races of the asexual fungal plant
464 pathogen *Verticillium dahliae* evolve by extensive chromosomal rearrangements including large-
465 scale gene losses (De Jonge *et al.*, 2012; Faino *et al.*, 2016; Chavarro-Carrero *et al.*, 2021). In
466 *P. effusa* loss of heterozygosity is suggested to contribute to the observed variation in asexually
467 reproducing isolates (Lyon *et al.*, 2016). From the available 13 *P. effusa* races that were
468 discovered after 1990, nine have evolved through sexual recombination (*Pe5, Pe7* - *Pe11*, *Pe13*,
469 *Pe15*, and *Pe16*). Thus, we hypothesise that in the period when *P. effusa* is under strong
470 evolutionary pressure from the introduction of resistant spinach varieties, sexual reproduction
471 has driven the fast emergence of new *P. effusa* races. Evolving sexually is often connected to
472 a more stable genome structure. For example, in *Phytophthora infestans* sexually evolving
473 populations are predominantly diploid and have less gene copy number variation than the
474 asexual populations (Cooke *et al.*, 2012). These findings show that there are multiple
475 mechanisms for evolutionary adaptation, both with asexual and sexual reproduction (Seidl and
476 Thomma, 2014).

477 Up to today, the genetic underpinnings of the race structure and evolution in *P. effusa* remain
478 unknown. Based on the findings in other plant pathogenic oomycetes and fungi, it can be
479 anticipated that the emergence of novel races is linked to genotypic changes in specific effector
480 genes. Effector proteins are secreted by pathogens to modulate host physiology, often by
481 deregulating immune responses, and establish infection. Resistant hosts produce immune
482 receptors that can detect these effectors and trigger strong immune responses thereby stopping
483 colonization. Effector genes often localize in repeat-rich genomic regions that are notoriously
484 challenging to assemble with short-read sequencing that has been commonly used until recently
485 (Gibriel *et al.*, 2016). High quality, complete genome assemblies of multiple *P. effusa* isolates
486 are thus invaluable to fully uncover the molecular underpinning of race emergence in *P. effusa*
487 (Thomma *et al.*, 2016). The newly generated high-quality genome assembly of *P. effusa* is much

488 longer and more repetitive compared with previous genome assemblies (Fletcher *et al.*, 2021),
489 providing a valuable resource to start comparative research between races. To link genotypic
490 variation to differences in resistance-breaking on spinach cultivars, emphasis should be given
491 to the systematic mining of effector genes and their variation between genetically and
492 phenotypically distinct *P. effusa* isolates. This research will greatly benefit from the here
493 developed phylogenetic framework and insights in the proposed sexual recombination between
494 different *P. effusa* races, which will facilitate our understanding of *P. effusa*-spinach interaction
495 and assist in sustainable production of spinach through knowledge-drive resistance breeding.

496 ## Experimental Procedures

497 ### *Peronospora effusa* infection on soil-grown spinach and spore isolation

498 We sowed spinach plants in potting soil (Primasta, NL) and kept them under long-day conditions
499 (16-hour light, 21 °C). Two to three weeks after germination, the spinach plants were inoculated
500 with *P. effusa* by spraying them with *P. effusa* spores suspended in water using a spray gun.
501 Following inoculation, the lids of the plastic trays were sprayed with water and covered to keep
502 the plants humid and dark. After 24 hours, we placed the plants under 9-hour light and 16 °C.
503 The lids of the boxes were again sprayed with water 7-10 days after inoculation, creating a
504 humid environment that promotes the sporulation of *P. effusa.*

505 To harvest *P. effusa* spores for Oxford nanopore sequencing, we collected leaves with
506 sporulating *P. effusa* from spinach plants and placed them in a glass bottle with tap water. The
507 spores were brought into suspension by shaking the bottle vigorously. Soil and other large
508 contaminants were removed by filtering the spore suspension over a 50-µm nylon mesh filter
509 (Merck Millipore, USA). To remove small biological contaminants, the remaining filtrate was first
510 filtered over a 11-µm nylon mesh filter (Merck Millipore, USA) using the Merck™ All-Glass Filter
511 Holder (47 mm) and a vacuum pump. Afterwards, the remaining spores were resuspended in
512 the All-Glass Filter Holder using 100 mL of autoclaved tap water and filtered.

513 ### High-molecular weight DNA extraction protocol

514 To isolate high-molecular weight (HMW) DNA, the collected *P. effusa* spores were freeze-dried
515 and lysed using glass beads in a tissue lyser (2 x 30 seconds on 30 Hz). The lysed spores were
516 incubated in extraction buffer (3% CTAB, 2% PVP, 1.25 M NaCl, 200 mM Tris-HCl, pH 8.5, 25
517 mM EDTA, pH 8.0) for 30 minutes at 65 °C, gently inverting the tube several times every 5
518 minutes. Then RNase was added, and the sample was incubated for 30 minutes at 37 °C, after
519 which 0.5% 2-Mercaptoethanol was added, followed by another 15-minute incubation at 37 °C.
520 HMW DNA was isolated from the lysate by subsequent phenol/chloroform/IAA extraction,
521 chloroform/IAA washing, RNase treatment, another phenol/chloroform/IAA extraction,
522 chloroform washing, and lastly, isopropanol precipitation.

523 DNA purity was determined with Nanodrop (Thermo Fisher Scientific, USA) to measure
524 A260/280 and A260/230 ratios. DNA concentrations were estimated using a fluorimeter and
525 fluorescent DNA-binding dye (Qubit^{TM} dsDNA BR Assay Kit; Thermo Fisher Scientific, USA),
526 according to the manufacturer's protocol. DNA integrity was confirmed by 0.4% agarose gel
527 electrophoresis performed at 4 °C at 20 V overnight.

528 **Genome sequencing using Oxford Nanopore**

529 We obtained long-read sequencing data for four *P. effusa* isolates (Pe1, Pe4, Pe11, Pe16) with
530 Oxford Nanopore sequencing technology (Oxford Nanopore, UK). The ligation-based
531 sequencing kit from Oxford Nanopore was used for library preparation (ONT - SQK-LSK109;
532 Oxford Nanopore, UK) following the manufacturer's protocol, apart from DNA-nick repair. This
533 was performed as described by the protocol of Illumina sequence library preparation (kit: NEB -
534 M6630; protocol: E6040; Illumina, USA). We used a Nanopore MinION flowcell (R10) for real-
535 time sequencing, and base-calling of the raw long-read sequencing data was performed using
536 Guppy (version 4.4.2; default settings). The raw long-read sequencing data were checked for
537 contamination using Kraken2 (version 2.0.9; default settings).

538 **Genome sequencing using Illumina**

539 We obtained paired-end short-read sequencing data of 24 *P. effusa* isolates and the
540 *Peronospora farinosa* f. sp. *betae* isolate *ES-15* on the Illumina sequencing platform (Illumina,
541 USA). Sequencing libraries were constructed with the Illumina TruSeq DNA PCR-Free kit.
542 Fragment-size distribution was determined before and after the library preparation using the
543 Agilent Bioanalyzer 2100 with HS-DNA chip (Agilent Technologies). The final library size was
544 approximately 550 bp). Libraries were paired-end sequenced (two times 150-bp reads) on the
545 Illumina NextSeq platform (Utrecht Sequencing Facility) in high output mode.

546 **Mitochondrial genome assembly and annotation**

547 We selected on average 155 Mb of long-read sequences from the four sequenced *P. effusa*
548 isolates based on their sequence alignment to 26 publicly available oomycete mitochondrial
549 genomes. To this end, we aligned the long-read sequences to these 26 mitochondrial genomes
550 using mimimap2 (version 2.17) (Li, 2018) with recommended settings for Nanopore reads. The

551 extracted reads were subsequently assembled with Canu (version 2.0) (Koren *et al.*, 2017) with
552 recommended settings for the assembly of Nanopore reads. We mapped the high-quality
553 Illumina paired-end short-reads of the four *P. effusa* isolates to their respective draft assembly
554 with BWA-mem (version 0.7.17; default settings) (Li and Durbin, 2009), and sequencing errors
555 in the draft mitochondrial genome assembly were corrected with Pilon (version 1.23; default
556 settings) (Walker *et al.*, 2014). Genes were annotated according to the universal genetic code
557 and based on sequence similarity with known mitochondrial genes using GeSeq (Tillich *et al.*,
558 2017). Inverted repeats were detected by self-alignment of mitochondrial genome sequences
559 and visualised using Circoletto (version 07.09.16) (Darzentas, 2010).

560 Clinker (version 0.0.12) (Gilchrist and Chooi, 2021) was used to visualise the mitochondrial
561 genome structure, by aligning the mitochondrial protein sequences of *Pe1* together with 17
562 Peronosporaceae species and using those alignments as ankers to visualise the order of genes
563 in the genomes.

564 **Variant calling on mitochondrial and nuclear genomes**

565 We used Illumina paired-end genomic short-read data for 26 *P. effusa* isolates to discover
566 single-nucleotide variants. We first filtered short reads by quality using Trimmomatic (version
567 0.39-1; MINLEN:36 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15) (Bolger *et al.*, 2014) and
568 removed Illumina sequencing artefacts with Fastp (version 0.20.0) (Chen *et al.*, 2018). The
569 filtered short-read data were aligned to the newly generated mitochondrial genome sequences
570 or to the publicly available nuclear genome assemblies of *Pe1* with BWA-mem (version 0.7.17;
571 default settings) (Li and Durbin, 2009; Klein *et al.*, 2020). We identified single-nucleotide variants
572 (SNPs) using the GATK joint variant calling pipeline (version 4.1.9.0) following the best practices
573 for germline short-variant discovery. We filtered individual variants based on the hard-filtering
574 best practices of GATK: Quality by depth < 4.0, Fisher Strand > 60.0, Strand Odds Ratio > 3.0,
575 RMS Mapping Quality < 20.0, Mapping Quality Rank Sum Test < -3.0, Read Position Rank Sum
576 Test < -1.0, Read Position Rank Sum Test > 3.5.

577 **Mitochondrial genome phylogeny**

578 The SNPs identified in each of the 26 *P. effusa* isolates were incorporated in the *Pe1*
579 mitochondrial genome sequence to reconstruct an isolate-specific mitochondrial sequence with
580 FastaAlternateReferenceMaker (GATK version 4.1.9.0; default settings). The resulting isolate-
581 specific sequences were aligned using MAFFT (version 7.471; default settings) (Nakamura *et*
582 *al.*, 2018) and a maximum-likelihood genome phylogeny was constructed using IQ-TREE
583 (version 1.6.12; default settings) (Nguyen *et al.*, 2015). Branch support was evaluated using
584 1,000 bootstrap replicates. Similarly, we reconstructed a maximum-likelihood phylogeny based
585 on the sequences of the mitochondrial genes *cox2* and *nad1* that were extracted from the
586 mitochondrial genomes of 18 Peronosporaceae.

587 **Nuclear genome nucleotide diversity**

588 Based on the data for all 26 *P. effusa* isolates, we generated a joint VCF file with both variant
589 and invariant sites, using the variant calling pipeline described above with the added option "-
590 all-sites" in the tool GenotypeGVCFs (GATK version 4.1.9.0). We calculated the nucleotide
591 diversity between the *P. effusa* isolates with pixy (version 1.0.4; --chunk_size 1000000)
592 (Korunes and Samuk, 2021). We performed hierarchical clustering of the genome-wide
593 nucleotide diversity and visualized the data as a hierarchically clustered heatmap with seaborn
594 (version 0.11.1) (Waskom, 2021).

595 The single nucleotide variants of the 26 *P. effusa* isolates were transformed into a distance
596 matrix with PGDSpider (version 2.1.1.5) (Lischer and Excoffier, 2012), which was used for the
597 following analysis: first, we performed a principal component analysis and visualised the first
598 two principal components on a scatterplot in RStudio (libraries: vcfR version 1.12.0, adegenet
599 version 2.1.3) (Jombart and Ahmed, 2011; Knaus and Grünwald, 2017). Second, we constructed
600 a decomposition network using the Neighbor-Net algorithm with SplitsTree (version 4.17.0)
601 (Huson and Bryant, 2006). We calculated the branch confidence of the network using 1,000
602 bootstrap replicates. PHI recombination test was calculated using phipack (version 1.1; default
603 settings) (Bruen *et al.*, 2006). The single nucleotide nuclear variants of the 26 *P. effusa* isolates
604 were additionally filtered for Minor Allele Frequency < 0.1 (169,016 biallelic SNPs), and
605 transformed into the STRUCTURE format with PGDSpider (version 2.1.1.5). We performed

606 STRUCTURE analysis using fastStructure (version 1.0; default settings) (Raj *et al.*, 2014) with
607 various k values. The results were visualised on a stacked bar-plot in RStudio (libraries: vcfR
608 version 1.12.0, poppr version 2.9.0) (Kamvar *et al.*, 2015; Knaus and Grünwald, 2017).

609 **Acknowledgments**

610 This project was financially supported by the TopSector TKI Tuinbouw & Uitgangsmaterialen,
611 the Netherlands, and four private breading companies: Enza Zaden Research & Development
612 BV, Pop Vriend Research B.V., Rijk Zwaan Breeding, and Syngenta.

613 **Data availability**

614 For this study we sequenced 16 denominated races of *P. effusa*. Cultures of these can be
615 requested for research from Naktuinbouw, the Netherlands (Correll *et al.*, 2015) (Supplementary
616 Table 4). The raw sequence data and genome assemblies generated in this study are available
617 at NCBI under BioProject PRJNA772192.

618 **Conflict of Interest**

619 Authors declare that they have no conflicting interests.

620 **References**

621 Achaz, G., Coissac, E., Netter, P., and Rocha, E.P.C. (2003) Associations between inverted
622 repeats and the structural evolution of bacterial genomes. *Genetics* **164**: 1279–1289.

623 Barrett, L.G., Thrall, P.H., Dodds, P.N., Van Der Merwe, M., Linde, C.C., Lawrence, G.J., and
624 Burdon, J.J. (2009) Diversity and evolution of effector loci in natural populations of the plant
625 pathogen *Melampsora lini*. *Mol Biol Evol* **26**: 2499–2513.

626 Barsoum, M., Kusch, S., Frantzeskakis, L., Schaffrath, U., and Panstruga, R. (2020) Ultraviolet
627 mutagenesis coupled with next-generation sequencing as a method for functional
628 interrogation of powdery mildew genomes. *Mol Plant-Microbe Interact* **33**: 1008–1021.

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

631 Bourret, T.B., Choudhury, R.A., Mehl, H.K., Blomquist, C.L., McRoberts, N., and Rizzo, D.M.
632 (2018) Multiple origins of downy mildews and mito-nuclear discordance within the
633 paraphyletic genus *Phytophthora*. *PLoS One* **13**: 1–23.

634 Bruen, T.C., Philippe, H., and Bryant, D. (2006) A simple and robust statistical test for detecting
635 the presence of recombination. *Genetics* **172**: 2665–2681.

636 Cai, G. and Scofield, S.R. (2020) Mitochondrial genome sequence of *Phytophthora sansomeana*
637 and comparative analysis of *Phytophthora* mitochondrial genomes. *PLoS One* **15**:
638 e0231296.

639 Chavarro-Carrero, E.A., Vermeulen, J.P., E. Torres, D., Usami, T., Schouten, H.J., Bai, Y., et
640 al. (2021) Comparative genomics reveals the in planta-secreted *Verticillium dahliae* Av2
641 effector protein recognized in tomato plants that carry the V2 resistance locus. *Environ*
642 *Microbiol* **23**: 1941–1958.

643 Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018) Fastp: An ultra-fast all-in-one FASTQ
644 preprocessor. In *Bioinformatics*. Oxford Academic, pp. i884–i890.

645 Choi, Y.J., Thines, M., Han, J.G., and Shin, H.D. (2011) Mitochondrial phylogeny reveals
646 intraspecific variation in *Peronospora effusa*, the spinach downy mildew pathogen. *J*
647 *Microbiol* **49**: 1039–1043.

648 Cook, D.E., Mesarich, C.H., and Thomma, B.P.H.J. (2015) Understanding Plant Immunity as a
649 Surveillance System to Detect Invasion. *Annu Rev Phytopathol* **53**: 541–563.

650 Cooke, D.E.L., Cano, L.M., Raffaele, S., Bain, R.A., Cooke, L.R., Etherington, G.J., et al. (2012)
651 Genome analyses of an aggressive and invasive lineage of the Irish potato famine
652 pathogen. *PLoS Pathog* **8**: e1002940.

653 Correll, J. and Smilde, D. (2021) Denomination of Pe: 18 and 19, two new races of downy mildew
654 in spinach. *Phantum* 1–2.

655 Correll, J., Toit, L. du, Koike, S., and Ettekoven, K. van (2015) Guidelines for Spinach Downy
656 Mildew: *Peronospora farinosa f. sp. spinaciae (Pfs)*.

657 Correll, J.C., Bluhm, B.H., Feng, C., Lamour, K., du Toit, L.J., and Koike, S.T. (2011) Spinach:
658 Better management of downy mildew and white rust through genomics. *Eur J Plant Pathol*
659 **129**: 193–205.

660 Dale, A.L., Feau, N., Everhart, S.E., Dhillon, B., Wong, B., Sheppard, J., et al. (2019) Mitotic
661 recombination and rapid genome evolution in the invasive forest pathogen *Phytophthora*
662 *ramorum*. *MBio* **10**: e02452-18.

663 Darzentas, N. (2010) Circoletto: Visualizing sequence similarity with Circos. *Bioinformatics* **26**:
664 2620–2621.

665 Depotter, J.R.L., Zuo, W., Hansen, M., Zhang, B., Xu, M., and Doehlemann, G. (2021) Effectors
666 with different gears: Divergence of ustilago maydis effector genes is associated with their
667 temporal expression pattern during plant infection. *J Fungi* **7**: 1–14.

668 Dhillon, B., Feng, C., Villarroel-Zeballos, M.I., Castroagudin, V.L., Bhattarai, G., Klosterman,
669 S.J., and Correll, J.C. (2020) Sporangiospore viability and oospore production in the
670 spinach downy mildew pathogen, *Peronospora effusa*. *Plant Dis* **104**: 2634–2641.

671 Drenth, A., McTaggart, A.R., and Wingfield, B.D. (2019) Fungal clones win the battle, but
672 recombination wins the war. *IMA Fungus* **10**: 1–6.

673 Faino, L., Seidl, M.F., Shi-Kunne, X., Pauper, M., Van Den Berg, G.C.M., Wittenberg, A.H.J.,
674 and Thomma, B.P.H.J. (2016) Transposons passively and actively contribute to evolution
675 of the two-speed genome of a fungal pathogen. *Genome Res* **26**: 1091–1100.

676 Feng, C., Saito, K., Liu, B., Manley, A., Kammeijer, K., Mauzey, S.J., et al. (2018) New races
677 and novel strains of the spinach downy mildew pathogen *Peronospora effusa*. *Plant Dis*
678 **102**: 613–618.

679 Fletcher, K., Klosterman, S.J., Derevnina, L., Martin, F., Bertier, L.D., Koike, S., et al. (2018)
680 Comparative genomics of downy mildews reveals potential adaptations to biotrophy. *BMC*
681 *Genomics* **19**: 8–10.

682 Fletcher, K., Shin, O.-H., Clark, K.J., Feng, C., Putman, A.I., Correll, J.C., et al. (2021) Ancestral
683 chromosomes for the Peronosporaceae inferred from a telomere-to-telomere genome
684 assembly of *Peronospora effusa*. *bioRxiv* 1–36.

685 Galtier, N., Nabholz, B., GlÉmin, S., and Hurst, G.D.D. (2009) Mitochondrial DNA as a marker
686 of molecular diversity: A reappraisal. *Mol Ecol* **18**: 4541–4550.

687 Gibriel, H.A.Y., Thomma, B.P.H.J., and Seidl, M.F. (2016) The age of effectors: Genome-based
688 discovery and applications. *Phytopathology* **106**: 1206–1212.

689 Gilchrist, C.L.M. and Chooi, Y.-H. (2021) clinker & clustermap.js: automatic generation of gene
690 cluster comparison figures. *Bioinformatics* **37**: 2473–2475.

691 Haas, B.J., Kamoun, S., Zody, M.C., Jiang, R.H.Y., Handsaker, R.E., Cano, L.M., et al. (2009)
692 Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora*
693 *infestans*. *Nature* **461**: 393–398.

694 Huson, D.H. and Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies.
695 *Mol Biol Evol* **23**: 254–267.

696 Irish, B.M., Correll, J.C., Koike, S.T., Schafer, J., and Morelock, T.E. (2003) Identification and
697 cultivar reaction to three new races of the spinach downy mildew pathogen from the United
698 States and Europe. *Plant Dis* **87**: 567–572.

699 Jombart, T. and Ahmed, I. (2011) adegenet 1.3-1: new tools for the analysis of genome-wide
700 SNP data. *Bioinformatics* **27**: 3070–3071.

701 De Jonge, R., Van Esse, H.P., Maruthachalam, K., Bolton, M.D., Santhanam, P., Saber, M.K.,
702 et al. (2012) Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens
703 uncovered by genome and RNA sequencing. *Proc Natl Acad Sci U S A* **109**: 5110–5115.

704 Kamvar, Z.N., Brooks, J.C., and Grünwald, N.J. (2015) Novel R tools for analysis of genome-
705 wide population genetic data with emphasis on clonality. *Front Genet* **0**: 208.

706 Kandel, S.L., Hulse-Kemp, A.M., Stoffel, K., Koike, S.T., Shi, A., Mou, B., et al. (2020)
707 Transcriptional analyses of differential cultivars during resistant and susceptible
708 interactions with *Peronospora effusa*, the causal agent of spinach downy mildew. *Sci Rep*
709 **10**: 1–13.

710 Klein, J., Neilen, M., van Verk, M., Dutilh, B.E., and van den Ackerveken, G. (2020) Genome
711 reconstruction of the non-culturable spinach downy mildew *Peronospora effusa* by
712 metagenome filtering.

713 Knaus, B.J. and Grünwald, N.J. (2017) vcfr: a package to manipulate and visualize variant call
714 format data in R. *Mol Ecol Resour* **17**: 44–53.

715 Knaus, B.J., Tabima, J.F., Shakya, S.K., Judelson, H.S., and Grünwald, N.J. (2020) Genome-
716 wide increased copy number is associated with emergence of dominant clones of the Irish
717 potato famine pathogen *Phytophthora infestans*. *MBio* **11**: 1–13.

718 Koike, S., Smith, R., and Schulbach, K. (1992) Resistant cultivars, fungicides combat downy
719 mildew of spinach. *Calif Agric* **46**: 29–30.

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

723 Korunes, K.L. and Samuk, K. (2021) pixy: Unbiased estimation of nucleotide diversity and
724 divergence in the presence of missing data. *Mol Ecol Resour* **21**: 1359–1368.

725 Li, H. (2018) Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**: 3094–
726 3100.

727 Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler
728 transform. *Bioinformatics* **25**: 1754–1760.

729 Linck, E. and Battey, C.J. (2019) Minor allele frequency thresholds strongly affect population
730 structure inference with genomic data sets. *Mol Ecol Resour* **19**: 639–647.

731 Lischer, H.E.L. and Excoffier, L. (2012) PGDSpider: An automated data conversion tool for

732 connecting population genetics and genomics programs. *Bioinformatics* **28**: 298–299.

733 Lyon, R., Correll, J., Feng, C., Bluhm, B., Shrestha, S., Shi, A., and Lamour, K. (2016) Population
734 structure of *Peronospora effusa* in the southwestern United States. *PLoS One* **11**:
735 e0148385.

736 McDonald, B.A. and Stukenbrock, E.H. (2016) Rapid emergence of pathogens in agro-
737 ecosystems: Global threats to agricultural sustainability and food security. *Philos Trans R*
738 *Soc B Biol Sci* **371**: 20160026.

739 McMullan, M., Gardiner, A., Bailey, K., Kemen, E., Ward, B.J., Cevik, V., et al. (2015) Evidence
740 for suppression of immunity as a driver for genomic introgressions and host range
741 expansion in races of *Albugo candida*, a generalist parasite. *Elife* **2015**: e04550.

742 Miller, M.E., Nazareno, E.S., Rottschaefer, S.M., Riddle, J., Santos Pereira, D. Dos, Li, F., et al.
743 (2020) Increased virulence of *Puccinia coronata f. sp. avenae* populations through allele
744 frequency changes at multiple putative Avr loci. *PLoS Genet* **16**: 1–30.

745 Mohd-Assaad, N., McDonald, B.A., and Croll, D. (2019) The emergence of the multi-species
746 NIP1 effector in Rhynchosporium was accompanied by high rates of gene duplications and
747 losses. *Environ Microbiol* **21**: 2677–2695.

748 Möller, M. and Stukenbrock, E.H. (2017) Evolution and genome architecture in fungal plant
749 pathogens. *Nat Rev Microbiol* **15**: 756–771.

750 Nakamura, T., Yamada, K.D., Tomii, K., and Katoh, K. (2018) Parallelization of MAFFT for large-
751 scale multiple sequence alignments. *Bioinformatics* **34**: 2490–2492.

752 Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. (2015) IQ-TREE: A fast and
753 effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*
754 **32**: 268–274.

755 Raj, A., Stephens, M., and Pritchard, J.K. (2014) FastSTRUCTURE: Variational inference of
756 population structure in large SNP data sets. *Genetics* **197**: 573–589.

757 Ribera, A., Bai, Y., Wolters, A.M.A., van Treuren, R., and Kik, C. (2020) A review on the genetic
758 resources, domestication and breeding history of spinach (*Spinacia oleracea L.*). *Euphytica*
759 **216**: 1–21.

760 Satou, M., Nishi, K., Kubota, M., Fukami, M., Tsuji, H., and Ettekoven, K. Van (2006)
761 Appearance of race Pfs:5 of spinach downy mildew fungus, *Peronospora farinosa f. sp.*
762 *spinaciae*, in Japan. *J Gen Plant Pathol* **72**: 193–194.

763 Seidl, M.F. and Thomma, B.P.H.J. (2014) Sex or no sex: Evolutionary adaptation occurs
764 regardless. *BioEssays* **36**: 335–345.

765 Thomma, B.P.H.J., Seidl, M.F., Shi-Kunne, X., Cook, D.E., Bolton, M.D., van Kan, J.A.L., and
766 Faino, L. (2016) Mind the gap; seven reasons to close fragmented genome assemblies.
767 *Fungal Genet Biol* **90**: 24–30.

768 Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R., and Greiner, S.
769 (2017) GeSeq - Versatile and accurate annotation of organelle genomes. *Nucleic Acids*
770 *Res* **45**: W6–W11.

771 Voineagu, I., Narayanan, V., Lobachev, K.S., and Mirkin, S.M. (2008) Replication stalling at
772 unstable inverted repeats: Interplay between DNA hairpins and fork stabilizing proteins.
773 *Proc Natl Acad Sci U S A* **105**: 9936–9941.

774 Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014) Pilon:
775 An integrated tool for comprehensive microbial variant detection and genome assembly
776 improvement. *PLoS One* **9**: e112963.

777 Wang, W., Schalamun, M., Morales-Suarez, A., Kainer, D., Schwessinger, B., and Lanfear, R.
778 (2018) Assembly of chloroplast genomes with long- and short-read data: A comparison of
779 approaches using Eucalyptus pauciflora as a test case. *BMC Genomics* **19**: 1–15.

780 Waskom, M.L. (2021) seaborn: statistical data visualization. *J Open Source Softw* **6**: 3021.

781 Yuan, X., Feng, C., Zhang, Z., and Zhang, C. (2017) Complete mitochondrial genome of
782 *Phytophthora nicotianae* and identification of molecular markers for the oomycetes. *Front*
783 *Microbiol* **8**: 1–14.

784