

 Open access • Posted Content • DOI:10.1101/249565

Nuclear and mitochondrial genomes of the hybrid fungal plant pathogen *Verticillium longisporum* display a mosaic structure — Source link

Depotter, Beveren Fv, Thomas A. Wood, Bart P. H. J. Thomma ...+1 more authors

Institutions: Wageningen University and Research Centre

Published on: 17 Jan 2018 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Verticillium longisporum, Genome, Verticillium, Genomics and Nuclear gene

Related papers:

- [Homogenization of sub-genome secretome gene expression patterns in the allodiploid fungus *Verticillium longisporum*](#)
- [Analysis of the hybrid genomes of two field isolates of the soil-borne fungal species *Verticillium longisporum*](#)
- [Verticillium longisporum and plant immunity responses in Arabidopsis](#)
- [The Ascomycete *Verticillium longisporum* Is a Hybrid and a Plant Pathogen with an Expanded Host Range](#)
- [Dynamic virulence-related regions of the plant pathogenic fungus *Verticillium dahliae* display enhanced sequence conservation](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/nuclear-and-mitochondrial-genomes-of-the-hybrid-fungal-plant-546st370gg>

1

Research Article

2 **Nuclear and mitochondrial genomes of the hybrid fungal plant pathogen**

3 ***Verticillium longisporum* display a mosaic structure**

4

5 Authors: Jasper R.L. Depotter^{1,2†}, Fabian van Beveren^{1†}, Grady C.M. van den Berg¹, Thomas

6 A. Wood^{2‡}, Bart P.H.J. Thomma^{1*‡}, Michael F. Seidl^{1*‡}

7

8 ¹Laboratory of Phytopathology, Wageningen University, Droevendaalsesteeg 1, 6708 PB

9 Wageningen, The Netherlands

10 ²Department of Crops and Agronomy, National Institute of Agricultural Botany, Huntingdon

11 Road, CB3 0LE Cambridge, United Kingdom

12

13 † These authors contributed equally to this work

14 ‡ These authors contributed equally to this work

15 *For correspondence:

16 Bart P.H.J. Thomma, Laboratory of Phytopathology, Wageningen

17 University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands. Tel. 0031-317-

18 484536, e-mail: bart.thomma@wur.nl

19 Michael F. Seidl, Laboratory of Phytopathology, Wageningen

20 University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands. Tel. 0031-317-

21 484536, e-mail: michael.seidl@wur.nl

22 **ABSTRACT**

23 Allopolyploidization, genome duplication through interspecific hybridization, is an important
24 evolutionary mechanism that can enable organisms to adapt to environmental changes or
25 stresses. This increased adaptive potential of allopolyploids can be particularly relevant for
26 plant pathogens in their quest for host immune response evasion. Allodiploidization likely
27 caused the shift in host range of the fungal pathogen plant *Verticillium longisporum*, as *V.*
28 *longisporum* mainly infects Brassicaceae plants in contrast to haploid *Verticillium* spp. In this
29 study, we investigated the allodiploid genome structure of *V. longisporum* and its evolution in
30 the hybridization aftermath. The nuclear genome of *V. longisporum* displays a mosaic
31 structure, as numerous contigs consists of sections of both parental origins. *V. longisporum*
32 encountered extensive genome rearrangements, whereas the contribution of gene conversion
33 is negligible. Thus, the mosaic genome structure mainly resulted from genomic
34 rearrangements between parental chromosome sets. Furthermore, a mosaic structure was also
35 found in the mitochondrial genome, demonstrating its bi-parental inheritance. In conclusion,
36 the nuclear and mitochondrial genomes of *V. longisporum* parents interacted dynamically in
37 the hybridization aftermath. Conceivably, novel combinations of DNA sequence of different
38 parental origin facilitated genome stability after hybridization and consecutive niche
39 adaptation of *V. longisporum*.

40

41 **Key words:** whole-genome duplication, allopolyploidy, genomic rearrangement, *Verticillium*
42 stem striping, *Verticillium* wilt, *Brassica*

43 INTRODUCTION

44 Whole-genome duplication (WGD) is an important evolutionary mechanism that facilitates
45 environmental adaptation (Van de Peer et al. 2017; Mallet 2005). The duplication of genomic
46 content increases genomic plasticity, leading to an augmented adaptive potential of organisms
47 that underwent WGD. Consequently, polyploids have been associated with increased
48 invasiveness (te Beest et al. 2012) and resistance to environmental stresses (Lohaus & Van de
49 Peer 2016). For instance, numerous plant species that survived the Cretaceous-Palaeogene
50 mass extinction, 66 million years ago, underwent a WGD which is thought to have
51 contributed to their increased survival rates (Vanneste et al. 2014a; Vanneste et al. 2014b).
52 Both genome copies involved in WGD may have the same species origin, i.e.
53 autopolyploidization, or originate from different species as a result of interspecific
54 hybridization, i.e. allopolyploidization. In general, allopolyploids are believed to have a
55 higher adaptive potential than autopolyploids due to the increased genetic divergence between
56 the chromosome sets.

57 The impact of allopolyploidization has mainly been investigated in plants, as
58 approximately a tenth of all plant species consists of allopolyploids (Barker et al. 2015). In
59 contrast, allopolyploidization in fungi is far less intensively investigated (Campbell et al.
60 2016). Nonetheless, allopolyploidization impacted the evolution of numerous fungal species,
61 including the economically important baker's yeast *Saccharomyces cerevisiae* (Marcet-
62 Houben & Gabaldón 2015). The increased adaptive potential enabled allopolyploid fungi to
63 develop desirable traits that can be exploited in industrial bioprocessing (Peris et al. 2017).
64 For instance, at least two recent hybridization events between *S. cerevisiae* and its close
65 relative *Saccharomyces eubayanus* gave rise to *Saccharomyces pastorianus*, a species with
66 high cold tolerance and good maltose/maltotriose utilization capabilities, which is exploited in

67 the production of lager beer that requires barley to be malted at low temperatures (Gibson &
68 Liti 2015).

69 Allopolyploid genomes experience a so-called “genome shock” upon hybridization,
70 inciting major genomic reorganizations that can manifest by genome rearrangements,
71 extensive gene loss, transposon activation, and alterations in gene expression (Doyle et al.
72 2008). These early stage alterations are primordial for hybrid survival, as divergent evolution
73 is principally associated with incompatibilities between the parental genomes (Matute et al.
74 2010). Allopolyploids benefit from a thorough re-organization where negative interactions
75 between the parental genomes are purged. Frequently, heterozygosity is lost for many regions
76 in the allopolyploid genome (Mixão & Gabaldón 2018). This can be a result of the direct loss
77 of a duplicated gene copy through deletion or gene conversion, a process where one of the
78 copies substitutes its homeologous counterpart (McGrath et al. 2014). Gene conversion and
79 the homogenization of complete chromosomes played a pivotal role in the evolution of the
80 osmotolerant yeast species *Pichia sorbitophila* (Louis et al. 2012). In total, two of its seven
81 chromosome pairs consist of partly heterozygous, partly homozygous sections, whereas two
82 chromosome pairs are completely homozygous. Gene conversion may eventually result in
83 chromosomes consisting of sections of both parental origins as “mosaic genomes”
84 (Stukenbrock et al. 2012). However, mosaic genomes can also arise through recombination
85 between chromosomes of the different parents, such as in the hybrid yeast
86 *Zygosaccharomyces parvii* (Ortiz-Merino et al. 2017).

87 Plant pathogens are often thought to evolve while being engaged in arms races with
88 their hosts; pathogens evolve to evade host immunity while plant hosts attempt to intercept
89 pathogen ingress (Cook et al. 2015). Due to the increased adaptation potential,
90 allopolyploidization has been proposed as a potent driver in pathogen evolution (Depotter et
91 al. 2016b). Allopolyploids often have different pathogenic traits than their parental lineages,

92 such as higher virulence (Husson et al. 2015; Brasier & Kirk 2001) and shifted host ranges
93 (Inderbitzin et al. 2011b; Zeise & Tiedemann 2002). Within the fungal genus *Verticillium*,
94 allodiploidization resulted in the emergence of a novel pathogen on brassicaceous plants;
95 *Verticillium longisporum* (Inderbitzin et al. 2011b; Depotter et al. 2017b). Similar to haploid
96 *Verticillium* spp., *V. longisporum* is thought to have a predominant asexual reproduction as a
97 sexual cycle has never been described and populations are not outcrossing (Depotter et al.
98 2017b; Short et al. 2014). *V. longisporum* is sub-divided into three lineages, each representing
99 a separate hybridization event (Inderbitzin et al. 2011b). The economically most important
100 lineage is A1/D1 that originates from hybridization between *Verticillium* species A1 and D1
101 that have hitherto not been found in their haploid states. *V. longisporum* lineage A1/D1 is the
102 main causal agent of *Verticillium* stem striping on oilseed rape (Novakazi et al. 2015) and its
103 economic importance as emerging pathogen is increasing worldwide (Depotter et al. 2017a).
104 A recent study revealed that lineage A1/D1 can be further divided into two genetically distinct
105 populations, which have been named ‘A1/D1 West’ and ‘A1/D1 East’ after their geographic
106 occurrence in Europe (Depotter et al. 2017b). Nevertheless, both populations were shown to
107 originate the same hybridization event (Depotter et al. 2017b).

108 *V. longisporum* is assumed to have largely conserved its allodiploid state as the sizes
109 of its sub-genomes resemble those of haploid *Verticillium* spp. (Depotter et al. 2017b; Shi-
110 Kunne et al. forthcoming; Fogelqvist et al. 2018). Nevertheless, not all genes are present in
111 heterozygous copies, as its nuclear ribosomal internal transcribed spacer region is derived
112 only from one of the parents (Inderbitzin et al. 2011b). Here, we investigated the evolution of
113 the allodiploid genome of *V. longisporum* and determined to what extent heterozygosity is
114 lost.

115 MATERIAL AND METHODS

116 Genome analysis

117 Genome assemblies of the two *V. longisporum* strains (VLB2 and VL20) and *V. dahliae* strain
118 JR2 were previously published (Faino et al. 2015; Depotter et al. 2017b). Telomeric regions
119 were determined based on the telomeric repeat pattern: TAACCC/GGGTTA (minimum three
120 repetitions). Furthermore, additional repeats were identified and characterized using
121 RepeatModeler (v1.0.8). De-novo-identified repeats were combined with the repeat library
122 from RepBase (release 20170127) (Bao et al. 2015). The exact coordinates of the repeats were
123 extracted with the software RepeatMasker (v4.0.6) (Smit et al. 2015). Genome-wide sequence
124 identities between *Verticillium* strains were calculated with dnadiff (Kurtz et al. 2004).
125 Homologous genes were retrieved by nucleotide BLAST (v2.2.31+). Here, only hits with a
126 minimal coverage of 80% with each other were selected.

127

128 Gene annotation

129 *Verticillium* isolates JR2, VLB2 and VL20 were grown for 3 days in potato dextrose broth.
130 Total RNA was extracted based on TRIzol RNA extraction (Simms et al. 1993). cDNA
131 synthesis, library preparation (TreSep RNA-Seq short-inser library), and Illumina sequencing
132 (single-end 50bp) was performed at the Beijing Genome Institute (BGI, Hong Kong, China).
133 In total, ~2Gb of filtered reads were mapped on the *Verticillium* genomes using TopHat
134 (Trapnell et al. 2009) with Bowtie2 (Langmead & Salzberg 2012). Gene annotation was
135 performed with the BRAKER1 1.9 pipeline (Hoff et al. 2016) using GeneMark-ET
136 (Lomsadze et al. 2014) and AUGUSTUS (Stanke et al. 2008). Predicted genes with internal
137 stop codons were removed from the analysis.

138

139 Parental origin determination

140 Sub-genomes were divided based on the differences in sequence identities between species
141 A1 and D1 with *V. dahliae*. *V. longisporum* genomes of VLB2 and VL20 were aligned to the
142 complete genome of *V. dahliae* JR2 using NUCmer from the MUMmer package v3.23 (Kurtz
143 et al. 2004; Faino et al. 2015). Here, only 1-to-1 alignments longer than 10kb and a minimum
144 of 80% identity were retained. Subsequent alignments were clustered together. The average
145 nucleotide identity was determined for every cluster and used for sub-genome segregation.

146 The parental origin determination based on sequence identities of the exonic regions of
147 genes, which was performed by BLAST (v2.6.0+). Here, hits with a minimum subject and
148 query coverage of 80% were used. Furthermore, similar to Louis et al. (2012), differences in
149 GC-content between homolog genes present in two copies were calculated accordingly:

$$150 \quad dGC_{\text{gene}} = 2 * GC_{\text{gene}} / (GC_{\text{gene}} + GChomolog)$$

$$151 \quad GC_{\text{gene}} = GC\% \text{ of gene}$$

$$152 \quad GChomolog = GC\% \text{ of homolog}$$

$$153 \quad dGC_{\text{gene}} = GC \text{ ratio from the mean GC\% value}$$

154

155 **Gene conversion and genomic rearrangements**

156 Double copy genes were retrieved by nucleotide BLAST (v2.6.0+) and the sequence identity
157 determined. Here, hits with a minimum subject and query coverage of 80% were used.
158 Nucleotide BLAST (v2.6.0+) was also used to find the corresponding homologous genes in *V.*
159 *longisporum* strains VLB2 and VL20, which were present in two copies in both strains.

160 The VLB2 genome assembly was aligned to VL20 to find breaks in synteny using
161 NUCmer from the MUMmer package v3.23 (Kurtz et al. 2004). Subsequent alignments were
162 clustered if they aligned to the same contig with the same orientation and order as the
163 reference genome. In order to confirm the breaks in synteny, filtered *V. longisporum* reads of
164 VLB2 were aligned to the *V. longisporum* VL20 genome with the Burrows-Wheeler Aligner

165 (BWA) and further processed with the samtools package (v1.3.1) (Li et al. 2009). Breaks in
166 syntenic clusters were then visualized and determined using the R package Sushi (Phanstiel et
167 al. 2014) and the Integrative Genomics Viewer (Robinson et al. 2011). The association
168 between breaks with repeats was tested through permutation. First, the fraction of breaks
169 flanked by repeats was determined. Here, breaks were assigned to reside in a “repeat-rich”
170 region if a 1 kb window around the break consisted for more than 10% of repeats. Then, the
171 *V. longisporum* VL20 genomes was divided into windows of 1 kb using BEDTools (v2.26.0)
172 to calculate the significance of the break/repeat association (Quinlan & Hall 2010). In total,
173 10,000 permutations were executed with the same amount of windows to determine the
174 random distribution of repeat-rich regions.

175

176 **Phylogenetic tree construction**

177 The phylogenetic tree of the nuclear DNA was based on the nucleotide sequences of the
178 ascomycete set BUSCO orthologs present in clade Flavnonexudans *Verticillium* spp. (Simão
179 et al. 2015). In total, 1,194 genes were included and concatenated for tree construction. The
180 mitochondrial phylogenetic tree was based on the nucleotide sequence of complete or partial
181 mitochondrial genomes. Nuclear and mitochondrial genomes of *Verticillium* spp. other than
182 *V. longisporum* were previously sequenced and assembled (Shi-Kunne et al. forthcoming;
183 Faino et al. 2015; Jelen et al. 2016). Whole-genome alignments for tree construction were
184 performed by mafft (v7.271) (default settings) (Katoh & Standley 2013; Katoh et al. 2002),
185 and subsequently the likelihood phylogenetic tree was reconstructed using RAxML with the
186 GTRGAMMA substitution model (v8.2.0) (Stamatakis 2014). The robustness of the inferred
187 phylogeny was assessed by 100 rapid bootstrap approximations.

188 **RESULTS**

189 ***V. longisporum* displays a mosaic genome structure**

190 The genomes of two *V. longisporum* strains were analysed to investigate the impact of
191 hybridization on the genome structure. Previously, *V. longisporum* strains VLB2 and VL20,
192 belonging to ‘A1/D1 West’ and ‘A1/D1 East’, were sequenced with the PacBio RSII platform
193 and annotated into genomes of 72.9 and 72.3 Mb in size, respectively (Depotter et al. 2017b).
194 These genome sizes exceed double the amount of the telomere-to-telomere sequenced *V.*
195 *dahliae* strains JR2 (36.2 Mb) and VdLs17 (36.0 Mb) (Faino et al. 2015). Despite the
196 previous observation that the two *V. longisporum* strains belong to distinct populations within
197 the A1/D1 lineage, the genomes show an extremely high degree of sequence identity
198 (99.91%). We used RepeatModeler (V1.0.8) in combination with RepeatMasker to
199 determined that 14.28 and 13.90% of the *V. longisporum* strain VLB2 and VL20 genomes is
200 composed of repeats, respectively (Table 1) (Smit et al. 2015). Intriguingly, this is more than
201 double the repeat content as in *V. dahliae* strain JR2, for which 6.49% of the genome was
202 annotated as repeat using the same methodology. The *V. longisporum* genomes were also
203 screened for telomere-specific repeats (TAACCC/GGGTTA) to estimate the number of
204 chromosomes. In total, 29 and 30 telomeric regions were found in the VLB2 and VL20
205 genomes, respectively, that were consistently situated at the end of contigs, suggesting that *V.*
206 *longisporum* contains at least 15 chromosomes (Table 1). Six out of 45 and four out of 44
207 contigs in strains VLB2 and VL20, respectively, were flanked on both ends by telomeric
208 repeats and therefore likely represent complete chromosomes (Table 1). For comparison, *V.*
209 *dahliae* strains contain 8 chromosomes (Faino et al. 2015).

210 In allopolyploid organisms, parental origin determination is elementary to investigate
211 genome evolution in the hybridization aftermath. As species D1 is phylogenetically closer
212 related, and consequently has a higher sequence identity, to *V. dahliae* than species A1, *V.*

213 *longisporum* genomic regions were previously provisionally assigned to either species D1 or
214 A1 (Depotter et al. 2017b). Here, we determined the parental origin of *V. longisporum*
215 genomic regions more precisely. The difference in phylogenetic distance of species A1 and
216 D1 to *V. dahliae* caused that *V. longisporum* genome alignments to *V. dahliae* displayed a
217 bimodal distribution with one peak at 93.1% and another peak at 98.4% sequence identity that
218 represent the two parents with a dip in between at 96.0% (Figure S1). In order to separate the
219 two sub-genomes, regions with an average sequence identity to *V. dahliae* <96% were
220 assigned to species A1, whereas regions with an identity of $\geq 96\%$ were assigned to species
221 D1 (Figure 1). In this manner, 36.2 Mb of *V. longisporum* strain VLB2 was assigned to
222 species A1 and 35.7 Mb to species D1. For *V. longisporum* strain VL20, 36.3 Mb was
223 assigned to species A1 and 35.2 Mb to species D1. Only 1.0 and 0.8 Mb of strains VLB2 and
224 VL20 could not be aligned to *V. dahliae* and thus remained unassigned, respectively.

225 To trace the chromosome sets of the original parents of the hybrid, the parental origin
226 of individual contigs was determined. In total, 8 of the 10 largest contigs of *V. longisporum*
227 strain VLB2 as well as strain VL20 consist of regions originating from both species A1 and
228 species D1 (Figure 1). Thus, parental chromosome sets cannot be separated from one another
229 as *V. longisporum* apparently evolved a mosaic genome structure in the hybridization
230 aftermath.

231

232 **Genomic rearrangements are responsible for the mosaic genome**

233 Typically, a mosaic structure of a hybrid genome can originate from gene conversion or from
234 chromosomal rearrangements between DNA strands of different parental origin (Mixão &
235 Gabaldón 2018). To analyse the extent of gene conversion, genes were predicted for the *V.*
236 *longisporum* strains VLB2 and VL20. To aid gene annotation with the BRAKER1 1.9
237 pipeline (Hoff et al. 2016), ~2 Gb of filtered RNA-seq reads were generated from fungal

238 cultures in potato dextrose broth. In total, 19,123 and 18,784 genes were predicted for *V.*
239 *longisporum* strains VLB2 and VL20 respectively, which is ~90% higher than the amount of
240 genes that were predicted for *V. dahliae* strain JR2 in this manner (9,909 genes) (Table 1). As
241 expected, the divergence of species A1 and D1 was also reflected at the level of gene
242 sequences based on sequence identity and GC-content (Figure 1, S1). In total, 9,531 and
243 9,402 genes were assigned to the species A1 sub-genome of the strains VLB2 and VL20,
244 respectively, whereas the number of genes in the species D1 sub-genomes was 9,468 and
245 9,243 for these strains, respectively (Figure 2). Thus, the amount of genes is similar in the two
246 sub-genomes for both *V. longisporum* strains. Over 80% of the *V. longisporum* genes are
247 present in two copies whereas, similar to *V. dahliae*, almost all genes (97-98%) are present in
248 one copy within each of the *V. longisporum* sub-genomes. Moreover, of the 7,620 genes that
249 are present in two copies in VLB2 and VL20, only 5 genes were found to be highly identical
250 (<1%, nucleotide sequence identity) in VLB2, whereas the corresponding gene pair in VL20
251 was more diverse (>1%, nucleotide sequence identity) (Figure 3). In *V. longisporum* strain
252 VL20, no highly identical copies were found that are more divergent in VLB2. Collectively,
253 these findings indicate that most *V. longisporum* genes have a copy of a different parental
254 origin and that gene conversion played a minor role in during evolution of the mosaic
255 genome.

256 Considering that gene conversion played a minor role during genome evolution, the
257 mosaic genome structure of *V. longisporum* is likely to originate from rearrangements
258 between the chromosomes of different parental origin. To identify the location of genomic
259 rearrangements, the genome of *V. longisporum* strain VLB2 was aligned to that of strain
260 VL20 (Figure 4). Extensive rearrangements were observed between the two *V. longisporum*
261 strains, as 87 putative syntenic breaks were found. In order to confirm these synteny breaks,
262 individual long-reads of VLB2 were aligned to the VL20 genome assembly to assess if breaks

263 were supported by read mapping (Figure S2). In total, 60 synteny breaks could be confirmed
264 by read mapping. As genomic rearrangements are often associated with repeat-rich genome
265 regions, the synteny break points were tested for their occurrence in repeat-rich regions. In
266 total, 34 of the 60 (57%) confirmed synteny break points were flanked by repeats, which is
267 significantly more than what would be expected from random sampling (mean = 18.5%, $\sigma =$
268 0.05%) (Figure S3). In conclusion, it appears that genomic rearrangements, rather than gene
269 conversion, are the main driver behind the mosaic structure of the *V. longisporum* genome.

270

271 ***V. longisporum* loses heterozygosity through deletions**

272 In each of the *V. longisporum* isolates, 17% of the genes occur only in a single copy.
273 Although gene conversion played a minor role in the hybridization aftermath, loss of
274 heterozygosity may occur through gene loss or, alternatively, single-copy genes may originate
275 from parent-specific contributions. However, as 12% of the singly copy genes in strain VLB2
276 are present in two copies in strain VL20, and 16% of the single copy genes in VL20 are
277 present in two copies in VLB2, gene deletion seems to be an on-going process in *V.*
278 *longisporum* evolution since both strains are derived from the same hybridization event
279 (Depotter et al. 2017b). Thus, in the general absence of gene conversion in the *V. longisporum*
280 genome, loss of heterozygosity is mainly caused by deletions.

281

282 **Also the mitochondrial genome has a bi-parental origin**

283 To determine how *Verticillium* species A1 and D1 relate to other species in the clade
284 Flavnonexudans, a phylogenetic tree was constructed based on 1,194 Ascomycota
285 Benchmarking Universal Single-Copy Orthologs (BUSCOs) that are present in all members
286 of *Verticillium* clade Flavnonexudans (Figure 5). Species A1 diverged before the last common
287 ancestor of *V. alfalfa*, *V. dahliae* and *V. nonalfalfae*. In contrast, species D1 only recently

288 diverged from *V. dahliae* after the last common ancestor of *V. alfalfae*, *V. dahliae* and *V.*
289 *nonalfalfae*. In addition to genomic DNA, the *V. longisporum* clade Flavnonexudans
290 phylogeny was also determined based on mitochondrial DNA (mtDNA) (Figure 5). The
291 mitochondrial genomes of the haploid clade Flavnonexudans spp. were previously sequenced
292 and found to be between 25-27 kb in size (Shi-Kunne et al. forthcoming; Jelen et al. 2016).
293 The *V. longisporum* mtDNA was assembled along with the nuclear genome of VLB2 and
294 VL20, which resulted in a mitochondrial genome of 26.1 kb. Unanticipatedly, the
295 phylogenetic position of *V. longisporum* based on mtDNA did not correspond to the
296 previously determined phylogenetic positions of either of the parents based on the nuclear
297 DNA (Figure 5), as it appeared that *V. longisporum* diverged more recently from the *V.*
298 *alfalfae/V. nonalfalfae* cluster than *V. dahliae*. As hybrid genome structures may impede
299 truthful phylogenetic resolution (Linder & Rieseberg 2004), we hypothesized that a hybrid
300 origin of the *V. longisporum* mitochondrial genome caused the discrepancy between the
301 phylogenetic trees based on nuclear and mitochondrial DNA. In order to elucidate the hybrid
302 nature, differences in mitochondrial sequence identities of *V. longisporum* or *V. nonalfalfae*
303 compared with *V. dahliae* were used to determine the parental origin of *V. longisporum*
304 mitochondrial genomic regions (Figure 6). As *V. nonalfalfae* diverged more recently from *V.*
305 *dahliae* than species A1, yet before the divergence of *V. dahliae* and species D1, mtDNA that
306 originates from species A1 and D1 should have lower and higher sequence identity,
307 respectively, with *V. dahliae* than with *V. nonalfalfae*. Indeed, the *V. longisporum*
308 mitochondrial genome consists of sections with higher and lower identity to *V. dahliae* than to
309 *V. nonalfalfae* confirming the hybrid nature of the *V. longisporum* mitochondrial genome
310 (Figure 6). The hybrid nature of the *V. longisporum* mtDNA was furthermore confirmed with
311 a phylogenetic analysis based on a 3.5 kb region that displays 0.7% higher average sequence
312 identity to *V. dahliae* than to *V. nonalfalfae*. This mtDNA region placed *V. longisporum* in the

313 same phylogenetic position of species D1 with a divergence after the last common ancestor of
314 *V. dahliae*, *V. alfalfae* and *V. nonalfalfae* (Figure 6). In contrast, a *V. longisporum* mtDNA
315 region of the same length that has on average 0.4% lower sequence identity to *V. dahliae* than
316 *V. nonalfalfae* placed *V. longisporum* in the same phylogenetic position as species A1 (Figure
317 6). In conclusion, in addition to the nuclear genome, also the mtDNA of *V. longisporum*
318 displays a mosaic structure after recombination of the DNA of the two individual parents.

319 **DISCUSSION**

320 Divergent evolution often fixates genomic incompatibilities between populations, leading to
321 reproductive isolation and eventually even speciation (Seehausen et al. 2014). *Verticillium*
322 species A1 and D1 are distinct with a genome-wide nucleotide divergence of 7.6%, yet both
323 species overcame putative incompatibilities and hybridized into a stable allodiploid (Depotter
324 et al. 2016a). Conceivably, extensive genome alterations occurred after hybridization,
325 facilitating the *V. longisporum* genome to reach a stable equilibrium. The dynamic nature of
326 the *V. longisporum* genome during the hybridization aftermath is displayed by its mosaic
327 structure, not only in the nuclear genome, but even in the mitochondrial genome (Figure 1).
328 Mosaicism in *V. longisporum* is not driven by homogenization that played a negligible role in
329 the hybridization aftermath (Figure 3). Rather, *V. longisporum* mosaic genome structure is
330 caused by extensive genomic rearrangements after hybridization (Figure 4, S2). Genomic
331 rearrangements are major drivers of evolution and facilitate adaptation to novel or changing
332 environments (Seidl & Thomma 2014). This mode of evolution is not specific to the hybrid
333 nature of *V. longisporum* as *V. dahliae* similarly encountered extensive chromosomal
334 reshuffling (de Jonge et al. 2013; Faino et al. 2016). In *V. dahliae*, genomic rearrangements
335 are associated with the occurrence of lineage-specific regions that are derived from segmental
336 duplications, and that are crucial for fungal aggressiveness as they are enriched for transcribed
337 transposable elements and *in planta*-expressed genes (de Jonge et al. 2013; Faino et al. 2016).
338 Genomic rearrangements often result from double-strand DNA breaks that are erroneously
339 repaired based on templates that display high sequence similarity (Seidl & Thomma 2014). As
340 expected, the majority of the synteny breaks between the genomes of *V. longisporum* strains
341 VLB2 and VL20 reside in repeat-rich genome regions (Figure S3) as, due to their abundance,
342 repetitive sequences are more likely to act as a substrate for unfaithful repair (Seidl &
343 Thomma 2014). Nevertheless, 43% of the synteny breaks identified in *V. longisporum* are not

344 associated with repeat-rich regions. However, the presence of two genomes provides
345 orthologous sequences with sufficient identity to mediate unfaithful repair. Double-strand
346 DNA breaks can be generated by transposable elements (TEs). Transposon activity is
347 typically constrained by epigenetic mechanisms such as DNA methylations. However, these
348 constrained may be alleviated upon genome challenges, such as allopolyploidization (Slotkin
349 & Martienssen 2007). The “genome-shock” that *V. longisporum* encountered upon
350 hybridization may have induced TE activity. The higher repeat content of the *V. longisporum*
351 genome compared to the haploid *V. dahliae* may suggest that TE proliferation occurred after
352 hybridization, leading to a modest genome expansion, resulting in a genome size of *V.*
353 *longisporum* that is more than double when compared with *V. dahliae*, whereas *V.*
354 *longisporum* contains only 90% more genes (Table 1). Similarly, it has been suggested that
355 hybridization led to TE proliferation in allopolyploid root-knot nematodes, *Meloidogyne* spp.,
356 as TEs cover a ± 1.7 times higher proportion of their genomes when compared with the
357 closely related, yet non-hybrid, *Meloidogyne hapla* (Blanc-Mathieu et al. 2017).

358 Whole-genome duplication events are usually followed by extensive gene loss, often
359 leading to reversion to the original ploidy state (Maere et al. 2005). However, the so-called
360 ‘haploidization’ of *V. longisporum* has only proceeded to a limited extent, as 80% of the
361 genes are present in two copies (Figure 2), whereas the haploid *V. dahliae* genome contains
362 only 1% of its genes in two copies. Perhaps *V. longisporum* hybridized only recently, with
363 gene-loss being an on-going process that by now has only progressed marginally, but that will
364 lead to further losses over time. Alternatively, the retention of genes in two copies is of an
365 evolutionary advantage, as the two copies make an additive contribution or their redundancy
366 provides a source for functional divergence (Gu et al. 2002; Blanc & Wolfe 2014).
367 Furthermore, the majority of the two parental genomes may also be retained to maintain

368 genomic balance, as stoichiometric difference of in interacting genes may have detrimental
369 outcomes (Birchler & Veitia 2012).

370 Hybridization can lead to incongruences between phylogenetic trees based on genome
371 sections of different parental origin (Linder & Rieseberg 2004). High-resolution tree
372 construction showed that species A1 diverged before the last common ancestor of *V. alfalfae*,
373 *V. dahliae* and *V. nonalfalfae*, whereas species D1 only recently diverged from *V. dahliae* as
374 previously reported (Inderbitzin et al. 2011a). Intriguingly, the phylogenetic positions of the
375 parental genomes did not correspond to that of the mitochondrial genome of *V. longisporum*
376 (Figure 6). Mitochondrial DNA is subjected to laws different from the Mendelian principles
377 of segregation and independent assortment as it is maternally inherited in most sexual
378 eukaryotes, including numerous fungal species (Basse 2010). However, for particular
379 organisms, bi-parental mitochondrial inheritance is common. For instance, in *S. cerevisiae*
380 heteroplasmy can be maintained for ~20 generations, allowing parental mitochondrial
381 genomes to recombine (Fritsch et al. 2014). The mitochondrial genome of *V. longisporum* has
382 been inherited bi-parentally, as the mitochondrial genome is a mosaic consisting of alternating
383 regions derived from the A1 or D1 parent (Figure 6). Similar to *V. longisporum*, two incipient
384 species of the budding yeast *Saccharomyces paradoxus* contributed to the mitochondrial
385 genome of their natural hybrid offspring (Leducq et al. 2017). Bi-parental inheritance is often
386 associated with hybridization, although it is uncertain if hybridization facilitates paternal
387 leakage or if bi-parental inheritance is a common phenomenon that is easier to detect in
388 hybrids (Barr et al. 2005).

389

390 **Conclusion**

391 The *V. longisporum* genome consists of two near to complete genomes of its haploid parents.
392 Rearrangements between these parental chromosome sets occurred, resulting in a mosaic

393 genome structure. *V. longisporum* genomes display high plasticity, as 60 synteny breaks were
394 confirmed between strains with high nucleotide identity. Conceivably, the absence of meiotic
395 constraints and the presence of orthologous DNA clusters provide ample opportunities for the
396 *V. longisporum* genome to recombine. Inter-parental genomic recombination and functional
397 diversification of homeologs give *V. longisporum* an additional potential to adapt to
398 environmental alterations, which haploid *Verticillium* spp. do not have. This may have
399 enabled *V. longisporum* to alter its host range and cause disease on Brassicaceous plants.

400 **ACKNOWLEDGEMENTS**

401 The authors would like to thank the Marie Curie Actions programme of the European
402 Commission that financially supported the research of J.R.L.D. Work in the laboratories of
403 B.P.H.J.T. and M.F.S is supported by the Research Council Earth and Life Sciences (ALW)
404 of the Netherlands Organization of Scientific Research (NWO).

405

406 **REFERENCES**

- 407 Bao W, Kojima KK, Kohany O. 2015. Repbase Update, a database of repetitive elements in
408 eukaryotic genomes. *Mob. DNA*. 6:11.
- 409 Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA. 2015. On the relative abundance of
410 autopolyploids and allopolyploids. *New Phytol*. 210:391–398.
- 411 Barr CM, Neiman M, Taylor DR. 2005. Inheritance and recombination of mitochondrial
412 genomes in plants, fungi and animals. *New Phytol*. 168:39–50.
- 413 Basse CW. 2010. Mitochondrial inheritance in fungi. *Curr. Opin. Microbiol*. 13:712–719.
- 414 te Beest M et al. 2012. The more the better? The role of polyploidy in facilitating plant
415 invasions. *Ann. Bot*. 109:19–45.
- 416 Birchler JA, Veitia RA. 2012. Gene balance hypothesis: connecting issues of dosage
417 sensitivity across biological disciplines. *Proc. Natl. Acad. Sci*. 109:14746–14753.
- 418 Blanc-Mathieu R et al. 2017. Hybridization and polyploidy enable genomic plasticity without
419 sex in the most devastating plant-parasitic nematodes. *PLoS Genet*. 13:e1006777.
- 420 Blanc G, Wolfe KH. 2004. Functional divergence of duplicated genes formed by polyploidy
421 during Arabidopsis evolution. *Plant Cell*. 16:1679–1691.
- 422 Brasier CM, Kirk SA. 2001. Comparative aggressiveness of standard and variant hybrid alder
423 phytophthoras, *Phytophthora cambivora* and other *Phytophthora* species on bark of *Alnus*,
424 *Quercus* and other woody hosts. *Plant Pathol*. 50:218–229.
- 425 Campbell MA, Ganley ARD, Gabaldón T, Cox MP. 2016. The case of the missing ancient
426 fungal polyploids. *Am. Nat*. 188:602–614.
- 427 Cook DE, Mesarich CH, Thomma BPHJ. 2015. Understanding plant immunity as a
428 surveillance system to detect invasion. *Annu. Rev. Phytopathol*. 53:541–563.
- 429 Depotter JRL, Deketelaere S, et al. 2016a. *Verticillium longisporum*, the invisible threat to
430 oilseed rape and other brassicaceous plant hosts. *Mol. Plant Pathol*. 17:1004–1016.

- 431 Depotter JRL, Rodriguez-Moreno L, Thomma BPHJ, Wood TA. 2017a. The emerging British
432 *Verticillium longisporum* population consists of aggressive *Brassica* pathogens.
433 *Phytopathology*. 107:1399-1405.
- 434 Depotter JRL, Seidl MF, van den Berg GCM, Thomma BPHJ, Wood TA. 2017b. A distinct
435 and genetically diverse lineage of the hybrid fungal pathogen *Verticillium longisporum*
436 population causes stem striping in British oilseed rape. *Environ. Microbiol.* 19:3997–4009.
- 437 Depotter JRL, Seidl MF, Wood TA, Thomma BPHJ. 2016b. Interspecific hybridization
438 impacts host range and pathogenicity of filamentous microbes. *Curr. Opin. Microbiol.* 32:7–
439 13.
- 440 Doyle JJ et al. 2008. Evolutionary genetics of genome merger and doubling in plants. *Annu.*
441 *Rev. Genet.* 42:443–461.
- 442 Faino L et al. 2015. Single-molecule real-time sequencing combined with optical mapping
443 yields completely finished fungal genome. *MBio.* 6:e00936-15.
- 444 Faino L et al. 2016. Transposons passively and actively contribute to evolution of the two-
445 speed genome of a fungal pathogen. *Genome Res.* 26:1091–1100.
- 446 Fogelqvist J et al. 2018. Analysis of the hybrid genomes of two field isolates of the soil-borne
447 fungal species *Verticillium longisporum*. *BMC Genomics.* 19:14.
- 448 Fritsch ES, Chabbert CD, Klaus B, Steinmetz LM. 2014. A genome-wide map of
449 mitochondrial DNA recombination in yeast. *Genetics.* 198:755–771.
- 450 Gibson B, Liti G. 2015. *Saccharomyces pastorianus*: genomic insights inspiring innovation
451 for industry. *Yeast.* 32:17–27.
- 452 Gu Z, Nicolae D, Lu H-S, Li W-H. 2002. Rapid divergence in expression between duplicate
453 genes inferred from microarray data. *Trends Genet.* 18:609–613.
- 454 Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M. 2016. BRAKER1: unsupervised
455 RNA-seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics.*

456 32:767–769.

457 Husson C et al. 2015. Evidence for homoploid speciation in *Phytophthora alni* supports
458 taxonomic reclassification in this species complex. *Fungal Genet. Biol.* 77:12–21.

459 Inderbitzin P, Bostock RM, et al. 2011a. Phylogenetics and taxonomy of the fungal vascular
460 wilt pathogen *Verticillium*, with the descriptions of five new species. *PLoS One.* 6:e28341.

461 Inderbitzin P, Davis RM, Bostock RM, Subbarao KV. 2011b. The ascomycete *Verticillium*
462 *longisporum* is a hybrid and a plant pathogen with an expanded host range. *PLoS One.*
463 6:e18260.

464 Jelen V, de Jonge R, Van de Peer Y, Javornik B, Jakše J. 2016. Complete mitochondrial
465 genome of the *Verticillium*-wilt causing plant pathogen *Verticillium nonalfalfae*. *PLoS One.*
466 11:e0148525.

467 de Jonge R et al. 2013. Extensive chromosomal reshuffling drives evolution of virulence in an
468 asexual pathogen. *Genome Res.* 23:1271–1282.

469 Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple
470 sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30:3059–3066.

471 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7:
472 improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.

473 Kurtz S et al. 2004. Versatile and open software for comparing large genomes. *Genome Biol.*
474 5:R12.

475 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods.*
476 9:357–359.

477 Leducq J-B et al. 2017. Mitochondrial recombination and introgression during speciation by
478 hybridization. *Mol. Biol. Evol.* 34:1947–1959.

479 Li H et al. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.*
480 25:2078–2079.

- 481 Linder CR, Rieseberg LH. 2004. Reconstructing patterns of reticulate evolution in plants.
482 Am. J. Bot. 91:1700–1708.
- 483 Lohaus R, Van de Peer Y. 2016. Of dups and dinos: evolution at the K/Pg boundary. Curr.
484 Opin. Plant Biol. 30:62–69.
- 485 Lomsadze A, Burns PD, Borodovsky M. 2014. Integration of mapped RNA-Seq reads into
486 automatic training of eukaryotic gene finding algorithm. Nucleic Acids Res. 42:e119.
- 487 Louis VL et al. 2012. *Pichia sorbitophila*, an interspecies yeast hybrid, reveals early steps of
488 genome resolution after polyploidization. G3. 2:299–311.
- 489 Maere S et al. 2005. Modeling gene and genome duplications in eukaryotes. Proc. Natl. Acad.
490 Sci. 102:5454–5459.
- 491 Mallet J. 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol. 20:229–237.
- 492 Marcet-Houben M, Gabaldón T. 2015. Beyond the whole-genome duplication: phylogenetic
493 evidence for an ancient interspecies hybridization in the baker’s yeast lineage. PLOS Biol.
494 13:e1002220.
- 495 Matute DR, Butler IA, Turissini DA, Coyne JA. 2010. A test of the snowball theory for the
496 rate of evolution of hybrid incompatibilities. Science. 329:1518–1521.
- 497 McGrath CL, Gout J-F, Johri P, Doak TG, Lynch M. 2014. Differential retention and
498 divergent resolution of duplicate genes following whole-genome duplication. Genome Res.
499 24:1665–1675.
- 500 Mixão V, Gabaldón T. 2018. Hybridization and emergence of virulence in opportunistic
501 human yeast pathogens. Yeast. 35:5-20.
- 502 Novakazi F et al. 2015. The three lineages of the diploid hybrid *Verticillium longisporum*
503 differ in virulence and pathogenicity. Phytopathology. 105:662–673.
- 504 Ortiz-Merino RA et al. 2017. Evolutionary restoration of fertility in an interspecies hybrid
505 yeast, by whole-genome duplication after a failed mating-type switch. PLOS Biol.

506 15:e2002128.

507 Van de Peer Y, Mizrachi E, Marchal K. 2017. The evolutionary significance of polyploidy.

508 Nat. Rev. Genet. 18:411-424.

509 Peris D et al. 2017. Hybridization and adaptive evolution of diverse *Saccharomyces* species

510 for cellulosic biofuel production. Biotechnol. Biofuels. 10:78.

511 Phanstiel DH, Boyle AP, Araya CL, Snyder MP. 2014. Sushi.R: flexible, quantitative and

512 integrative genomic visualizations for publication-quality multi-panel figures. Bioinformatics.

513 30:2808–2810.

514 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic

515 features. Bioinformatics. 26:841–842.

516 Robinson JT et al. 2011. Integrative Genomics Viewer. Nat. Biotechnol. 29:24–26.

517 Seehausen O et al. 2014. Genomics and the origin of species. Nat. Rev. Genet. 15:176–192.

518 Seidl MF, Thomma BPHJ. 2014. Sex or no sex: Evolutionary adaptation occurs regardless.

519 BioEssays. 36:335–345.

520 Shi-Kunne X, Faino L, van den Berg GCM, Thomma BPHJ, Seidl MF. Evolution within the

521 fungal genus *Verticillium* is characterized by chromosomal rearrangements and gene losses.

522 Environ. Microbiol. Accepted

523 Short DPG, Gurung S, Hu X, Inderbitzin P, Subbarao KV. 2014. Maintenance of sex-related

524 genes and the co-occurrence of both mating types in *Verticillium dahliae*. PLoS One.

525 9:e112145.

526 Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO:

527 assessing genome assembly and annotation completeness with single-copy orthologs.

528 Bioinformatics. 31:3210–3212.

529 Simms D, Cizdziel P, Chomczynski P. 1993. TRIzol: a new reagent for optimal single-step

530 isolation of RNA. Focus. 15:99–102.

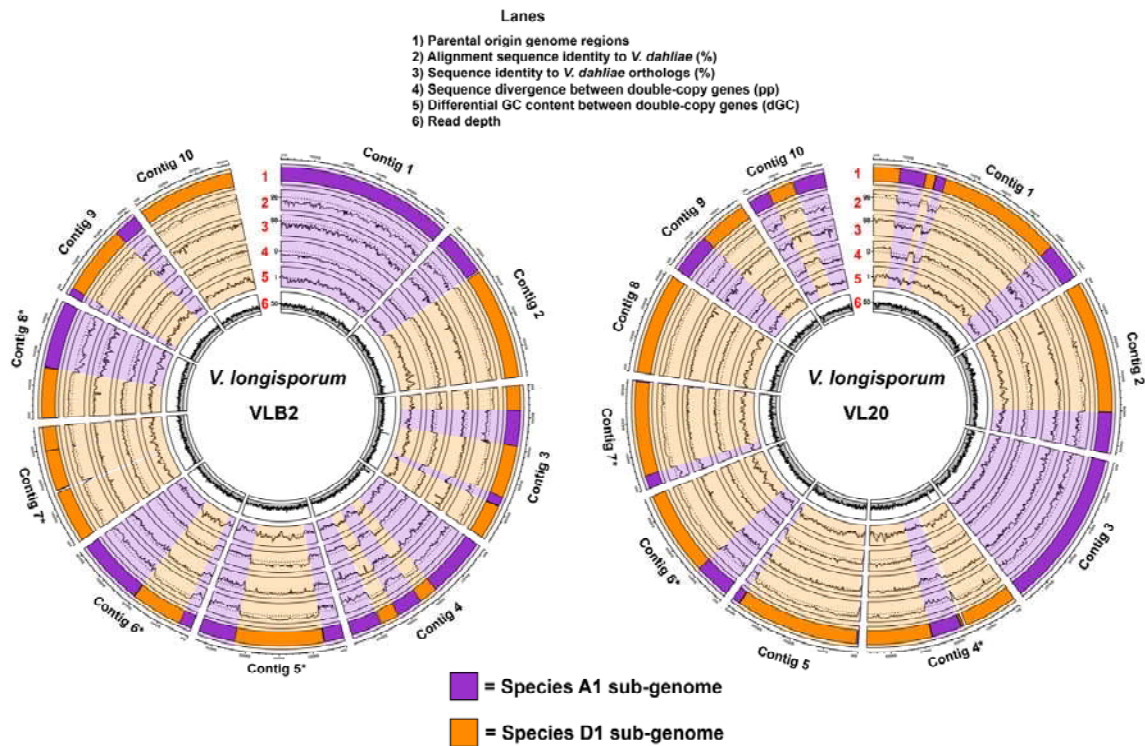
- 531 Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the
532 genome. *Nat. Rev. Genet.* 8:272–285.
- 533 Smit AFA, Hubley R, Green P. 2015. RepeatMasker Open-4.0. Available from:
534 <http://www.repeatmasker.org>
- 535 Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of
536 large phylogenies. *Bioinformatics.* 30:1312–1313.
- 537 Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntenically mapped
538 cDNA alignments to improve de novo gene finding. *Bioinformatics.* 24:637–644.
- 539 Stukenbrock EH, Christiansen FB, Hansen TT, Dutheil JY, Schierup MH. 2012. Fusion of
540 two divergent fungal individuals led to the recent emergence of a unique widespread pathogen
541 species. *Proc. Natl. Acad. Sci.* 109:10954–9.
- 542 Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-
543 Seq. *Bioinformatics.* 25:1105–1111.
- 544 Vanneste K, Baele G, Maere S, Van de Peer Y. 2014a. Analysis of 41 plant genomes supports
545 a wave of successful genome duplications in association with the Cretaceous-Paleogene
546 boundary. *Genome Res.* 24:1334–1347.
- 547 Vanneste K, Maere S, Van de Peer Y. 2014b. Tangled up in two: a burst of genome
548 duplications at the end of the Cretaceous and the consequences for plant evolution. *Philos.*
549 *Trans. R. Soc. B Biol. Sci.* 369:20130353.
- 550 Zeise K, Tiedemann Av. 2002. Host specialization among vegetative compatibility groups of
551 *Verticillium dahliae* in relation to *Verticillium longisporum*. *J. Phytopathol.* 150:112–119.

552 **TABLES**

553 **Table 1: Comparison *V. longisporum* and *V. dahliae* genomes.**

	<i>V. longisporum</i>	<i>V. longisporum</i>	<i>V. dahliae</i>
	VLB2	VL20	JR2
Genome size	72.9 Mb	72.3 Mb	36.2 Mb
Number of contigs	45	44	8
Complete chromosomes	6	4	8
Number of genes	19,123	18,784	9,909
Telomere regions	29	30	16
Repeat content	14.28	13.90	6.49
BUSCO completeness	99.0%	98.3%	98.7%

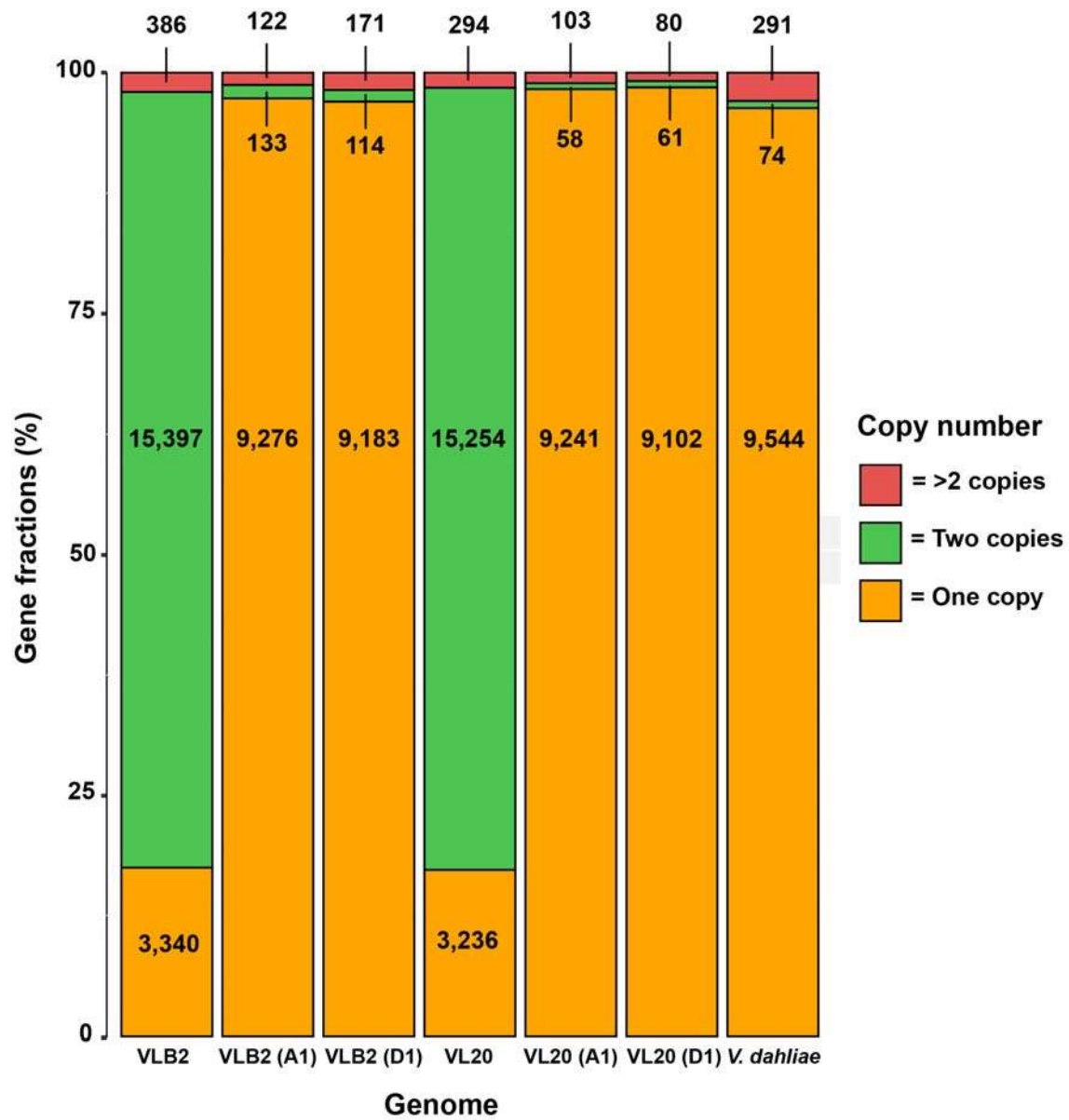
554 **FIGURE LEGENDS**



555

556 **Fig. 1: Determination of parental origin of *Verticillium longisporum* genome sections.**

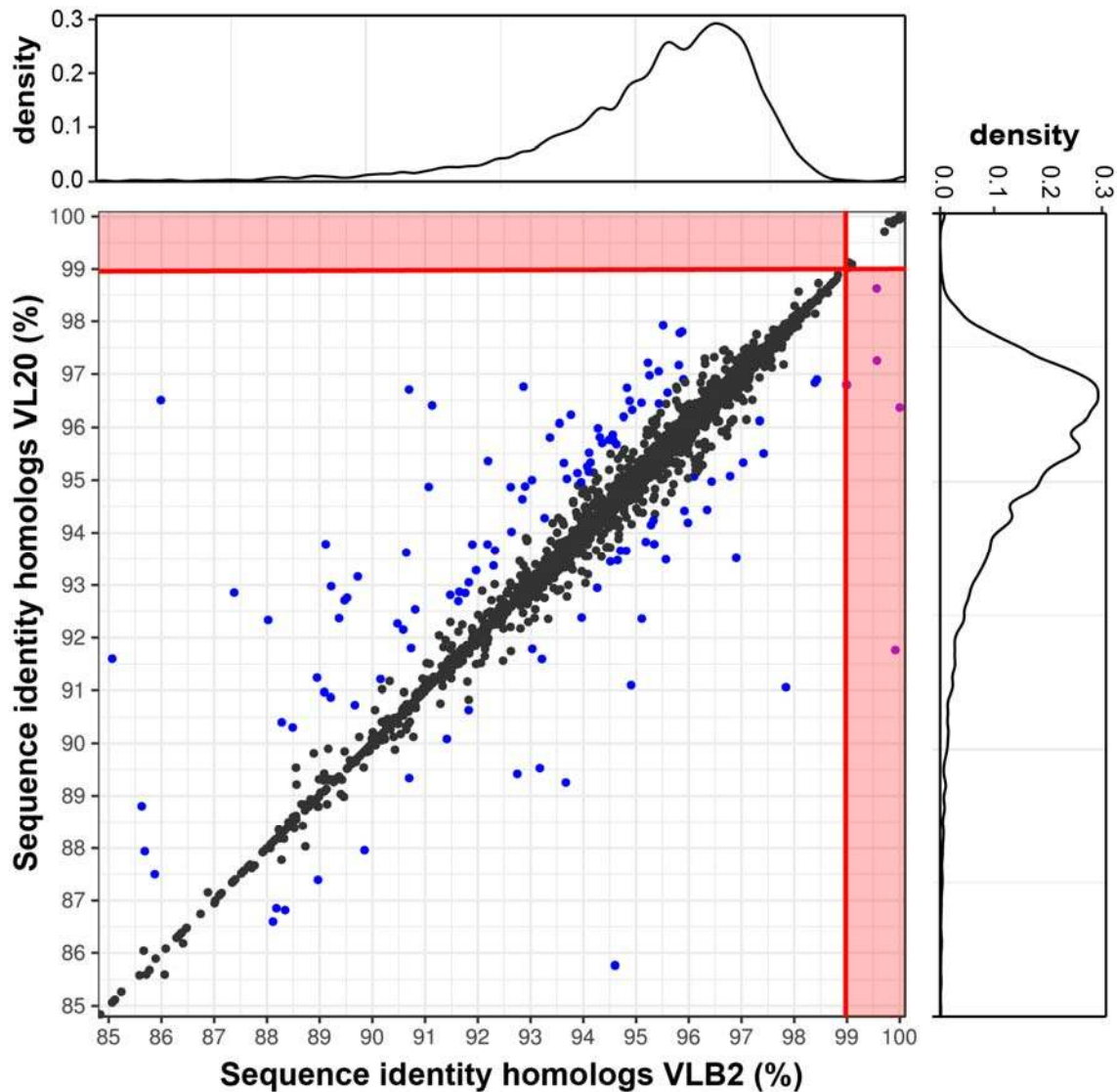
557 The ten largest contigs of the genome assemblies of *V. longisporum* strains VLB2 and VL20
558 are depicted. Lane 1: regions with a sequence identity of at least 96% were assigned to
559 Species D1 (orange), whereas ones with lower sequence identity to Species A1 (purple). Lane
560 2: sequence identity of *V. longisporum* alignments to *V. dahliae*. Lane 3: Sequence identity
561 between exonic regions of *V. longisporum* and *V. dahliae* orthologs. Lane 4: Difference in
562 sequence identity in percent point (pp) between exonic regions of *V. longisporum* double-
563 copy genes. Only gene pairs with an ortholog in *V. dahliae* are depicted. Alleles with a higher
564 identity to *V. dahliae* are depicted as a positive pp difference, whereas the corresponding
565 homolog as a negative pp difference. Lane 5: the relative difference in GC content (dGC)
566 between genes in double copy. Lane 6: Read depth with non-overlapping windows of 10 kb.
567 Data points of lanes 3-5 represent the average value of a window of eleven genes, which
568 proceed with a step of one gene.



569

570 **Fig. 2: Gene copy number distribution within *Verticillium* (sub-)genomes.** “(A1)” and
571 “(D1)” represent species A1 and D1 sub-genomes, respectively, of the *V. longisporum* strains
572 VLB2 and VL20. For *V. dahliae*, the strain JR2 was used.

573



574

575 **Fig. 3: The contribution of gene conversion to *V. longisporum* genome evolution.**

576 Sequence identities between double-copy genes, present in *V. longisporum* VLB2 and VL20,

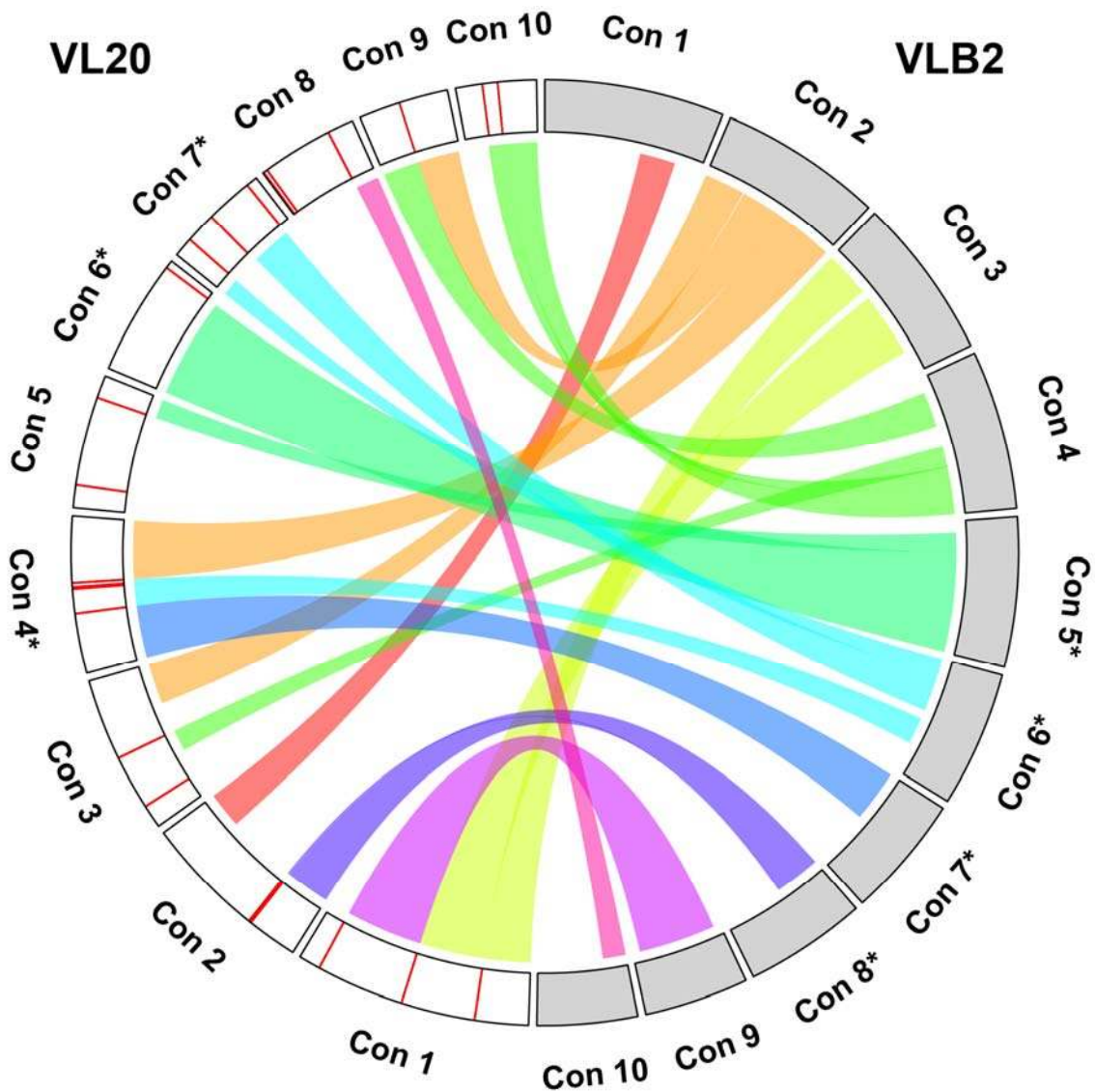
577 are depicted. Gene pairs that encountered gene conversion (purple dots in the red zones) have

578 sequence divergence of more than one percent in one *V. longisporum* strain and less than one

579 percent in the other strain. In other cases, pairs that differ less than one percent are depicted as

580 a black dot, whereas a difference higher than one percent is depicted as a blue dot.

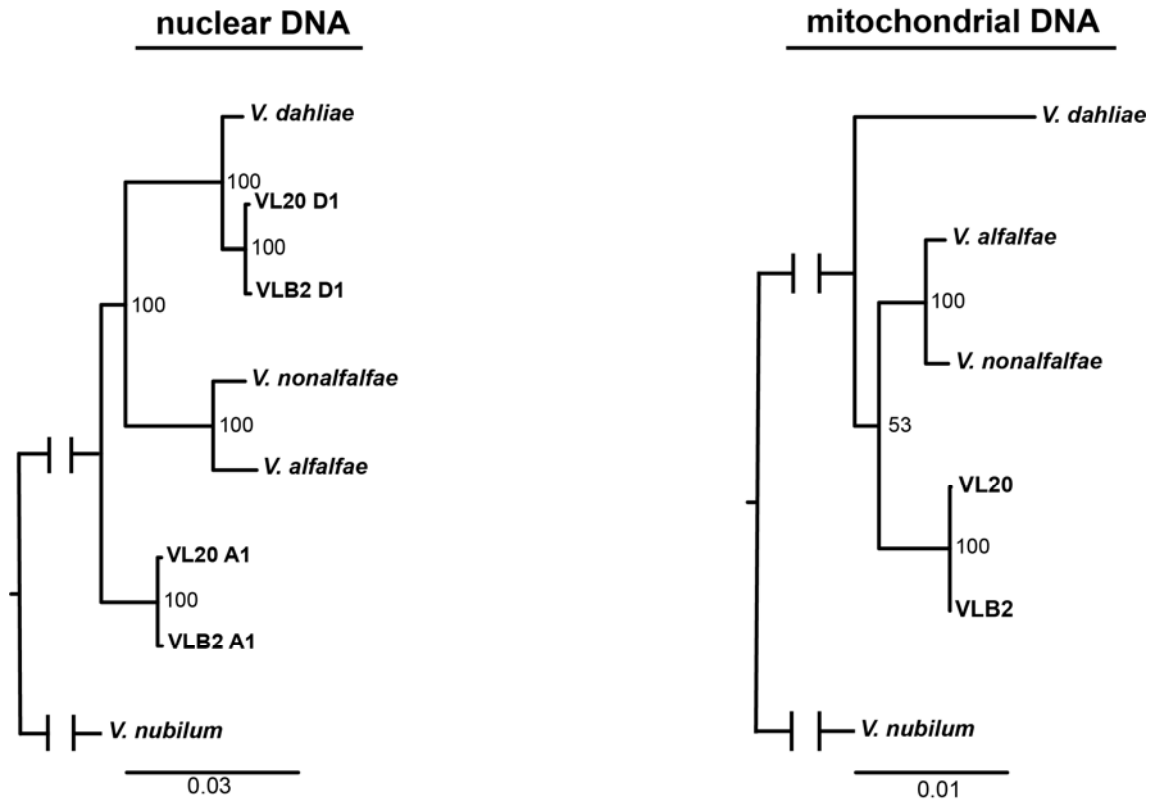
581



582

583 **Fig. 4: The contribution of genomic rearrangements to *V. longisporum* genome**
584 **evolution.** The ten largest contigs of the *V. longisporum* strains VLB2 (displayed in grey) and
585 VL20 (displayed in white) are depicted with complete chromosomes indicated by asterisks.
586 Ribbons indicate syntenic genome regions between the two strains. Red bars on the contigs
587 indicate syntenicity breaks that are confirmed by discontinuity in alignment of VLB2 reads to
588 VL20 genome.

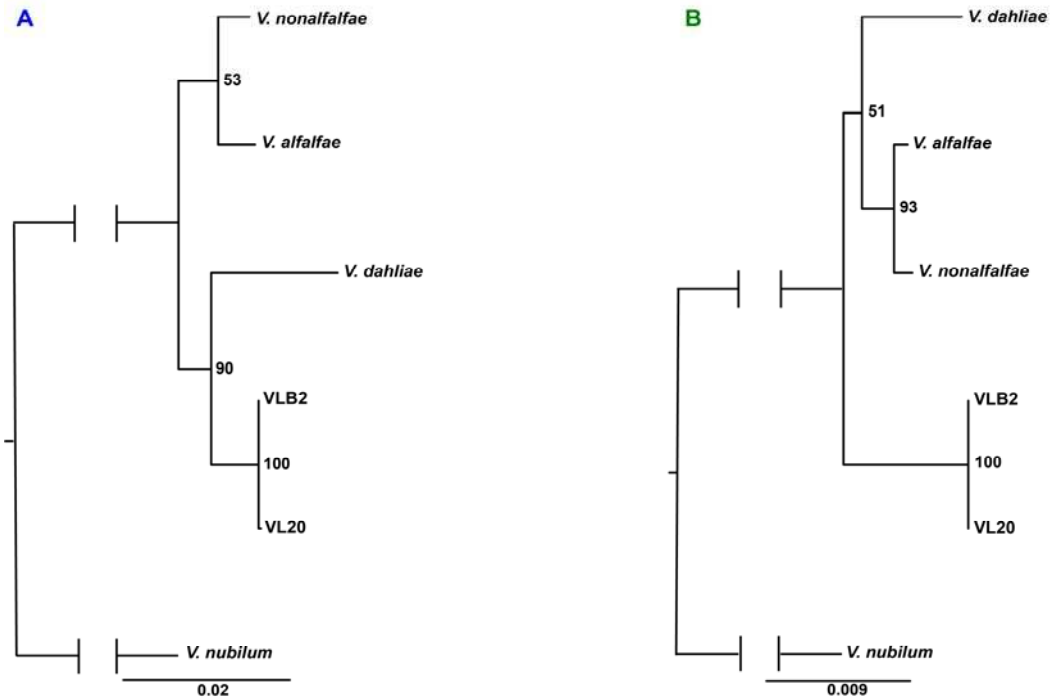
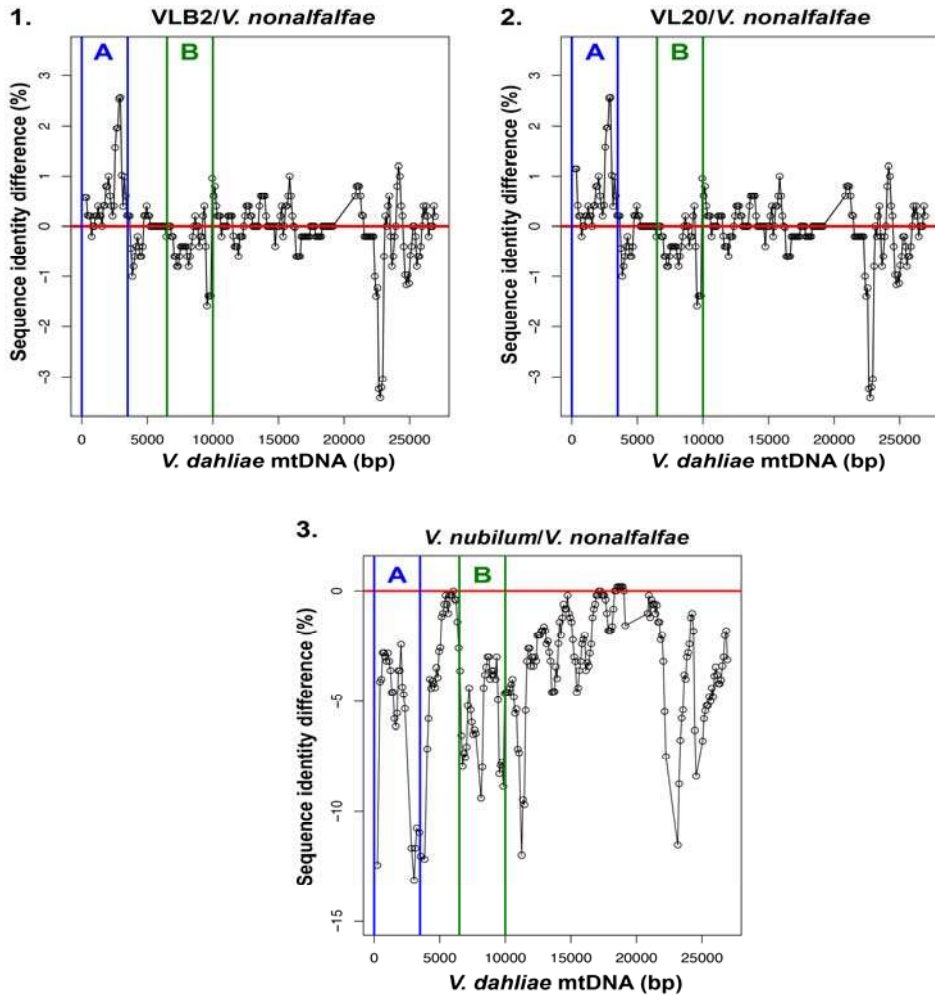
589



590

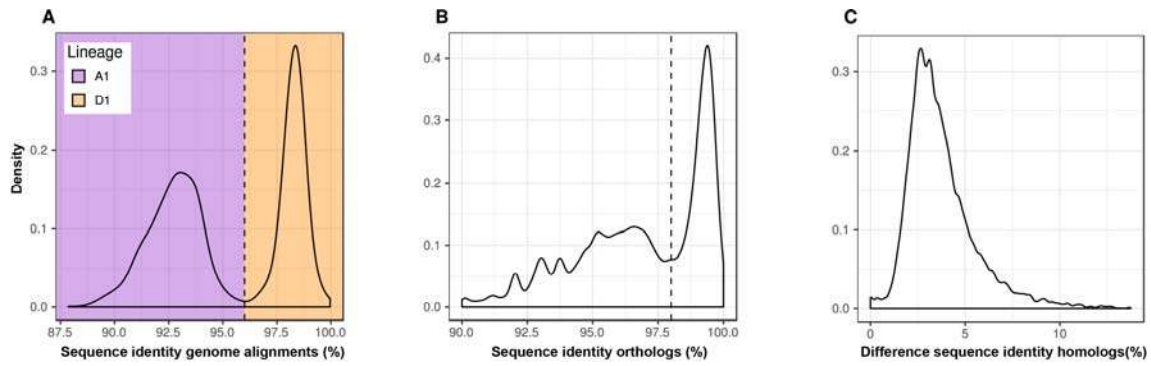
591 **Fig. 5: Phylogenetic positioning of *Verticillium longisporum* nuclear sub-genomes and**
592 **mitochondrial genome to other *Verticillium* species of clade Flavnonexudans.** The nuclear
593 phylogenetic tree was constructed with 1,194 orthologous genes, whereas the mitochondrial
594 phylogenetic tree was based on complete mitochondrial genomes. Maximum-likelihood
595 phylogeny analysis of *Verticillium* spp. was rooted on *Verticillium nubilum* and the robustness
596 of the tree was assessed using 100 bootstrap replicates.

597



599 **Fig. 6: The bi-parental origin to mitochondrial genome of *Verticillium longisporum*.** The
600 *V. dahliae* mitochondrial genome was divided in 500 bp sliding windows with 100 bp steps.
601 The differences in sequence identity of these windows with other *Verticillium* genomes were
602 determined for: 1. *V. longisporum* strain VLB2 and *V. nonalfalfae*, 2. *V. longisporum* strain
603 VL20 and *V. nonalfalfae*, and 3. *V. nubilum* and *V. nonalfalfae*. The two phylogenetic trees
604 are constructed with maximum-likelihood based on region A and B of the mitochondrial *V.*
605 *dahliae* genome, both 3.5 kb in size. The phylogenetic trees were rooted on *Verticillium*
606 *nubilum* and the robustness of the tree was assessed using 100 bootstrap replicates.

607 **SUPPLEMENTARY MATERIAL**



608

609 **Fig. S1: Lines of evidence for the parental origin of *V. longisporum* genomic regions.** (A)

610 Distribution of sequence identity of *V. longisporum* alignments to *V. dahliae*. (B) The

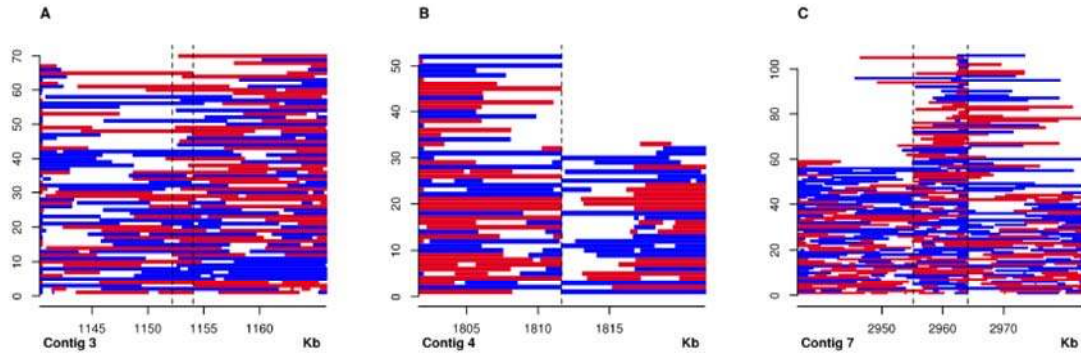
611 distribution of the sequence identity between *V. longisporum* exonic regions of genes and

612 their *V. dahliae* orthologs. (C) Distribution of sequence identity between exonic regions of *V.*

613 *longisporum* homologs that are present in two copies. Strains VLB2 and JR2 were used for *V.*

614 *longisporum* and *V. dahliae*, respectively.

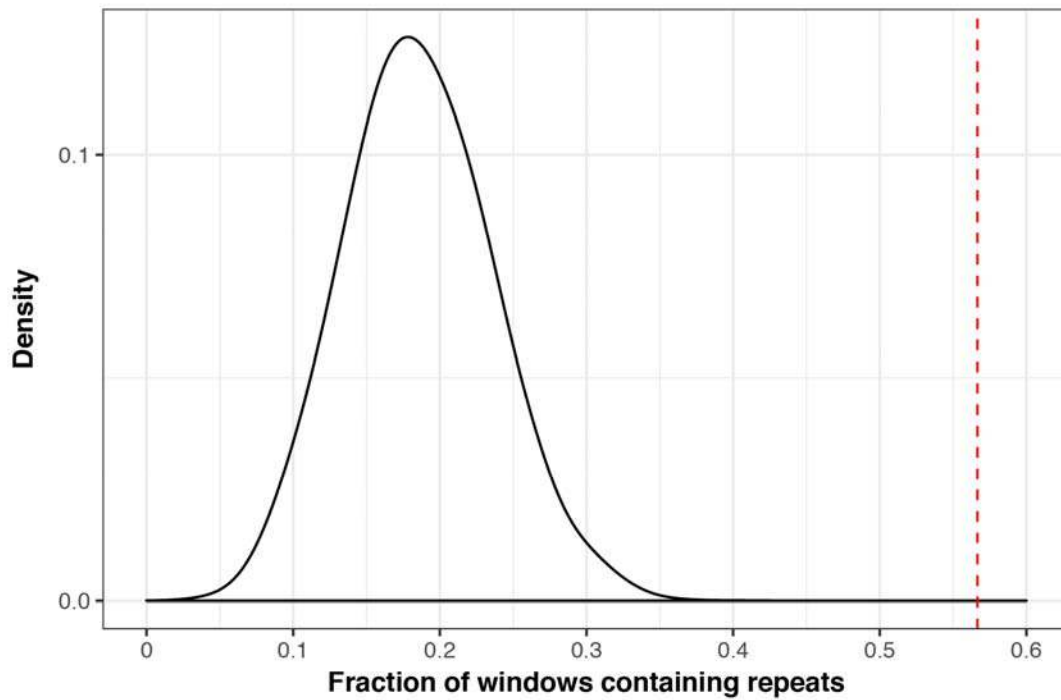
615



616

617 **Fig. S2: Synteny break confirmation by mapping *V. longisporum* VLB2 reads to the**
618 **VL20 genome.** Red and blue bars represent forward and reverse aligned VLB2 reads to the
619 VL20 genome. The dashed lines show the suggested position in synteny break through
620 genome alignment. (A) In particular cases, read alignments did not confirm the break in
621 synteny. (B) In other cases, breaks in synteny were confirmed as reads abruptly stopped and
622 started on these genome positions. (C) Breaks were also considered truthful if regions showed
623 overlap in repeat-rich regions where read overlap between adjacent genome regions is
624 lacking.

625



626

627 **Fig. S3: The association of synteny breaks with repetitive elements.** The black curve
628 represents the fraction of 60 randomly chosen 1 kb windows in the *V. longisporum* VL20 that
629 are repeat-rich, which has been permuted 10,000 times. The red line indicates the fraction of
630 true breaks that lay in a 1 kb window enriched for repeats (57%).