Dynamics in secondary metabolite gene clusters in otherwise highly syntenic and stable genomes in the fungal genus *Botrytis*

Claudio A. Valero-Jiménez[1]

Maikel B.F. Steentjes[1]

Jason C. Slot[3]

Xiaoqian Shi-Kunne[1]

Olga E. Scholten[2]

Jan A.L. van Kan[1]*

[1]Laboratory of Phytopathology, Wageningen University, 6708 PB Wageningen, the Netherlands
[2]Wageningen University & Research, Plant Breeding, 6708 PB Wageningen, the Netherlands
[3]Department of Plant Pathology, The Ohio State University, 2021 Coffey Rd., Columbus, OH, USA

* Author for correspondence: Jan van Kan, Laboratory of Phytopathology, Wageningen University, 6708 PB Wageningen, the Netherlands. Phone +31 317483126. Email: jan.vankan@wur.nl.

Abstract

Fungi of the genus *Botrytis* infect >1400 plant species and cause losses in many crops. Besides the broad host range pathogen *B. cinerea*, most other species are restricted to a single host. Long read technology was used to sequence genomes of eight *Botrytis* species, mostly pathogenic on *Allium* species, and the related onion white rot fungus, *Sclerotium cepivorum*. Most assemblies contained <100 contigs, with the *B. aclada* genome assembled in 16 gapless chromosomes. The core genome and pangenome of 16 *Botrytis* species were defined and the secretome, effector and secondary metabolite repertoires analysed. Among those genes, none are shared among all *Allium* pathogens and absent from non-*Allium* pathogens. The genome of each of the *Allium* pathogens contains 8-39 predicted effector genes that are unique for that single species, none stood out as potential determinant for host specificity. Chromosome configurations of common ancestors of the genus *Botrytis* and family Sclerotiniaceae were reconstructed. The genomes of *B. cinerea* and *B. aclada* were highly syntenic with only 19 rearrangements between them. Genomes of *Allium* pathogens were compared with 10 other *Botrytis* species (non-pathogenic on *Allium*) and with 25 Leotiomycetes for their repertoire of secondary metabolite gene clusters. The pattern was complex, with several clusters displaying patchy distribution. Two clusters involved in the synthesis of phytotoxic metabolites are at distinct genomic locations in different *Botrytis* species. We provide evidence that the clusters for botcinic acid production in *B. cinerea* and *B. sinoallii* were acquired by horizontal transfer from taxa within the same genus.

Keywords: ancestral genome, horizontal transfer, necrotroph, secondary metabolite.

22 **Significance statement**

23 We sequenced the genomes of nine plant pathogenic Sclerotiniaceae fungi, most of them
24 infecting onion or related *Allium* species, to identify host range determinants by analysing what
25 these species share and what distinguishes them from their non-*Allium* sister species. Despite
26 being unable to identify host range determining genes, several exciting observations were made.
27 Sclerotiniaceae have stable genomes with similar chromosome architecture. We reconstructed an
28 ancestral genome for all Sclerotiniaceae that contained 16 core chromosomes, as do all extant
29 species for which chromosome numbers are known. Nevertheless, two gene clusters for
30 secondary metabolite biosynthesis were located in entirely different genomic environments in
31 these species. Evidence is presented that one of these gene clusters has undergone horizontal
32 transfer within the genus *Botrytis*.

33 **Introduction**

34 Fungi have great societal impact because of their utility for nutritional, industrial and medical
35 purposes, as well as their pathogenic behaviour on humans and plants. In recent years, the
36 sequencing of fungal genomes has progressed at tremendous pace thanks to their small genome
37 size and decreases in sequencing costs (Spatafora et al. 2017). Many species of industrial fungi
38 from the genera *Aspergillus*, *Penicillium* and *Trichoderma* have been sequenced (e.g. de Vries et
39 al. 2017), while for human pathogens such as *Cryptococcus neoformans*, *Candida* spp., or
40 *Aspergillus fumigatus*, numerous isolates were sequenced to obtain insight in population diversity
41 (e.g. Ashton et al. 2019; Lind et al. 2017). Similarly, many dozens of plant pathogenic fungi species
42 have been sequenced in order to gain insight into their evolution and the traits that enable the
43 infection of plants (Moeller and Stukenbrock 2017). Studies on plant pathogenic fungi have
44 provided evidence for evolutionary adaptations that confer dynamics and plasticity on the genome,
45 such as the presence of repeat-rich, gene-poor genomic regions or the possession of entire
46 "dispensable" or "lineage-specific" chromosomes that contain effector genes which confer the
47 capacity to specifically infect certain host plant species or plant genotypes (Bertazzoni et al. 2018;
48 Dong et al. 2015; Lo Presti and Kahmann 2017; Sipos et al. 2017).

49 The fungal genus *Botrytis* comprises ~35 recognized species that all are pathogenic on
50 plants (Garfinkel et al. 2017; Hyde et al 2014) with the exception of *B. deweyae*, which colonizes
51 *Hemerocallis* (daylily) as an endophyte (Grant-Downton et al. 2014). *Botrytis* spp. are notorious
52 pathogens with a necrotrophic infection behaviour, i.e. they kill host cells and invade the dead cells
53 to acquire nutrients. Two species that have been extensively studied are *B. cinerea* and *B.*
54 *pseudocinerea*, morphologically indistinguishable taxa that cause grey mould on >1400 host plant
55 species (Elad et al. 2016). Other *Botrytis* species are considered to be restricted to a single host or
56 a small number of taxonomically related hosts (Elad et al., 206; Staats et al., 2005). In these cases,

each host plant usually is infected by its own specialized *Botrytis* species. There are two exceptions in the pattern of specialized host-pathogen relationships within the genus: as many as eight *Botrytis* species can infect onion (*Allium cepa*) or other *Allium* species (Staats et al. 2005), and a recent study reported as many as 15 previously unknown, phylogenetically distinct *Botrytis* taxa sampled from peony in Alaska (Garfinkel et al. 2019). Phylogenetic analysis separated the genus *Botrytis* into two distinct clades, and *Botrytis* species that infect *Allium* are widely dispersed throughout the largest clade (Garfinkel et al. 2019; Hyde et al. 2014; Staats et al. 2005). Their closest relatives are often pathogenic on hosts that are phylogenetically distant from *Allium.* For example, the closest relatives of *B. squamosa* (onion leaf blight) are the lily pathogen *B. elliptica* and *Hemerocallis* endophyte *B. deweyae.* Furthermore, the closest relative of *B. aclada* (onion neck rot) is the peony pathogen *B. paeoniae.* By contrast, *B. globosa* and *B. sphaerosperma* are sister taxa and both able to infect *Allium* hosts. The fact that *Allium* pathogens are dispersed over the phylogeny of the genus *Botrytis* suggests that the capacity to infect *Allium* has either been acquired multiple times or lost multiple times, independently, during evolution in the genus.

Pathogens with a necrotrophic lifestyle such as *Botrytis* spp. actively manipulate the cell death balance in their host plant, and in the necrotrophic phase exploit the host cell death machinery by secreting cell death-inducing metabolites and effector proteins (Veloso and van Kan 2018). In the necrotrophic wheat pathogen *Parastagonospora nodorum*, several cell death-inducing effector proteins were identified that contribute to pathogenicity only on wheat genotypes carrying a cognate receptor for these effectors, following an inverse gene-for-gene interaction (Faris et al., 2010; Liu et al. 2009; Liu et al. 2012; Shi et al. 2012; Shi et al. 2015). Each effector-receptor pair contributes in a quantitative manner to disease severity. At least one of the *P. nodorum* effector genes has been horizontally transferred between distinct fungi pathogenic on wheat and barley (Friesen et al. 2006; McDonald et al. 2019).

81 The genome of the generalist *B. cinerea* has been extensively studied in the past decade. A
82 gapless genome assembly was generated comprising 18 contigs, representing (near-)full-length
83 chromosomes. Two contigs are minichromosomes (209 and 247 kbp, respectively) with few genes
84 and neither seems relevant for plant infection (van Kan et al. 2017), indicating that the core genome
85 of *B. cinerea* consists of 16 chromosomes. Light microscopic studies by Shirane et al. (1989)
86 showed that five *Botrytis species* (*B. aclada, B. byssoidea, B. cinerea, B.squamosa,* and *B. tulipae*)
87 all contain 16 mitotic chromosomes. The *B. cinerea* reference assembly was supported by a genetic
88 and optical map (van Kan et al. 2017) and a manually curated community annotation (Ensembl
89 Fungi; Pedro et al. 2019). In a follow-up study, we analysed the genomes of nine *Botrytis* species,
90 mainly pathogens on flower bulb crops, using short read sequence technology (Valero-Jiménez et
91 al. 2019). In the present study, we sequenced the genomes of eight additional host-specific *Botrytis*
92 species and one *Sclerotium* species, most of which are pathogenic on *Allium*, in order to compare
93 their predicted proteome content and possibly identify host range determinants. The comparison
94 focussed on genes that are present in (and possibly shared among) *Allium* pathogens and absent
95 from the non-*Allium* pathogens. The genome assemblies were of sufficiently high quality to analyse
96 chromosome architecture and synteny, and to infer the genome organization of ancestors of the
97 genus *Botrytis* and the family Sclerotiniaceae. Furthermore, analysis of secondary metabolite
98 biosynthetic gene clusters in Sclerotiniaceae and 25 other fungi within the Leotiomycetes showed
99 a patchy distribution of these clusters and provided evidence for two horizontal transfer events of
100 a secondary metabolite biosynthetic gene cluster within the genus *Botrytis*.

101 **Materials and Methods**

102 ***Strains and culture conditions***

103 The fungal isolates that were sequenced are listed in Table S1. For long term storage, all *Botrytis*
104 species were kept as conidial suspensions in 15% glycerol at -80°C, while *S. cepivorum* was stored
105 as sclerotia at room temperature. The fungi were grown on malt extract plates (MEA) at 20°C
106 before DNA extraction.

107 ***DNA and RNA isolation***

108 High molecular weight DNA was isolated from freeze-dried and grinded mycelium upon treatment
109 with cell lysis solution (Qiagen), proteinase K and protein precipitation solution (Qiagen). DNA
110 was precipitated using isopropanol, redissolved in TE buffer and treated with RNase A. The
111 obtained DNA was cleaned using a Salt:chloroform wash (Pacific Biosciences shared protocol).
112 RNA used for producing RNA-seq libraries were pools of RNA isolated from different sources:
113 (1) 5-day old mycelia grown on MEA supplemented with blended onion leaves; (2) conidia; (3)
114 sclerotia; (4) infected onion bulbs; and (5) infected onion leaves. For isolation of RNA, freeze-
115 dried, grinded samples were incubated in Trizol (Ambion, Life Technologies) and treated with
116 chloroform. After adding ethanol to the aqueous phase, the mixture was used as input for an
117 RNeasy Plant Mini Kit (Qiagen) to isolate RNA.

118 ***Sequencing and assembly***

119 All genomes were sequenced with one Pacbio SMRT cell using the Sequel instrument at Keygene
120 N.V. (Wageningen, the Netherlands). De novo assembly was done with HGAP (Chin et al. 2013)
121 and CANU (Koren et al. 2017) using default settings. The resulting assemblies were combined
122 with quickmerge (Chakraborty et al. 2016), then two steps of corrections were done with Arrow,
123 and erroneously merged contigs (based on inspection of mapped reads coverage) were manually
124 corrected. Completeness of the genome assembly was assessed by the Benchmarking Universal

125 Single-Copy Orthologs (BUSCO) (Simao et al. 2015). The transcriptome of each genome was
126 sequenced using strand-specific paired-end libraries with a read length of 2x 150 bp using an
127 Ilumina HiSeq-X sequencer at the Beijing Genome Institute (BGI, Hongkong, China).

128 ***Genome annotation***

129 Genome annotation was performed using the FUNGAP pipeline (Min et al. 2017), which included
130 the annotation by MAKER (Cantarel et al. 2008), AUGUSTUS (Stanke et al. 2006) and BRAKER
131 (Hoff et al. 2015). The gene prediction tools were supported with RNA-Seq libraries. Gene models
132 of the manually curated genome of *B. cinerea* (van Kan et al. 2017), and all the fungal proteins
133 available in the Swissprot database were provided as evidence for gene prediction. Furthermore,
134 the predicted proteins were manually inspected and curated. The genome curation was done in
135 Webapollo (Dunn et al. 2019), and each gene was inspected to confirm that prediction was
136 supported by the evidence tracks (RNA-Seq, *B. cinerea* as reference and the Swissprot proteins);
137 for instance, some gene models were deleted if they were overlapping a repetitive region, while
138 other gene models where changed to have a correct Methionine start, or correct splice junctions.
139 The manual curation was done to all the predicted proteins of *B. aclada*, *B. squamosa* and *S.*
140 *cepivorum*, and to the secretome of all other genomes. The predicted proteins were functionally
141 annotated using the funannotate pipeline (Love et al. 2019).

142 ***Phylogenetic and phylogenomic analysis***

143 The phylogenetic relationships of the *Botrytis* genus and other related species of Sclerotiniaceae
144 were determined between all species sequenced in this study and including the previously
145 sequenced species *B. cinerea* B05.10 (van Kan et al. 2017), and other *Botrytis* species (Valero-
146 Jiménez et al. 2019). The other species that were included were *Sclerotinia sclerotiorum* and
147 *Sclerotinia borealis*, and *Marssonina brunnea* as the outgroup of the tree. The tree was constructed
148 using 4746 single-copy orthologue genes, identified with Orthofinder (Emms and Kelly 2015). The

149 protein sequence for each gene was aligned and concatenated into a single matrix using MAFFT
150 (Kato and Standley 2013), and a maximum likelihood phylogenetic tree was inferred with RAxML
151 v.8.2.10 (Stamatakis 2014) using a generalized time reversible (GTR) plus GAMMA amino acid
152 substitution model with 100 rapid bootstraps. A pan-genome analysis was done to calculate the
153 number of core genes and was estimated using OrthoMCL (Li 2003) implemented in
154 GET_HOMOLOGUES-EST (Contreras-Moreira and Vinuesa 2013) with e-value $1e^{-5}$ and 75%
155 coverage. For the pangenome analysis, only the orthogroups present in at least two species were
156 included.

157 ***Secretome and effector prediction***

158 Genes encoding putatively secreted proteins were identified for each genome using several
159 prediction tools. Signal-P v4.1 (Petersen et al. 2011) was initially used to screen for a signal
160 peptide, followed by TMHMM v.2.0 (Krogh et al. 2001) to identify putative transmembrane
161 domains. Proteins that did not have a signal peptide, or that had a transmembrane domain (a single
162 transmembrane domain in the first 60 amino acids was allowed) were discarded. TargetP was used
163 to predict protein localization (Emanuelsson et al. 2007). Effectors were predicted using the
164 EffectorP tool v1.0 and v2.0 (Sperschneider et al. 2016).

165 ***Ancestral genome reconstruction***

166 The ancestral genome of Botrytis was constructed using the CHROnicle package that comprises
167 SynChro, ReChro and Anchro (Vakirlis et al. 2016). In order to identify conserved synteny blocks,
168 pairwise comparisons between the genomes was done with SynChro. Subsequently, reconstruction
169 of the ancestral chromosome gene order was done with Anchro.

170 ***Secondary metabolite gene cluster analysis***

171 Putative gene clusters that are predicted to be involved in biosynthesis of secondary metabolites
172 were identified using antiSMASH using default settings (antibiotics and Secondary Metabolite

173 Analysis SHell) version 4.0.1 (Weber et al. 2015). The dataset used for this analysis included 45
174 genomes from the order Leotiomycetes that were publicly available and published (S3 Table). BiG-
175 SCAPE version 20181005 (Navarro-Munoz et al. 2019) was used to analyse all the secondary
176 metabolites clusters predicted by antiSMASH. In the BiG-SCAPE analysis a cutoff of 0.65 as well
177 as the MIBiG parameter that included the MIBiG repository version 1.4 of annotated SMC was
178 used (Medema et al. 2015). The output of BiG-SCAPE was visualized using Cytoscape version
179 3.7.1 (Shannon et al. 2003).

180 ***Reconstruction of BGC evolution***

181 Presence/absence and additional fragmented homologs of BOT and BOA genes for each species
182 was confirmed by tblastn against the genome assemblies (Supplementary Data S5 and S6).
183 Pseudogenes were manually identified by inspection of tblastn reports for in-frame stop codons,
184 and interrupted reading frames and truncations that could not be explained by novel intron sites
185 (Supplementary Data S5 and S6).

186 Phylogenetic analyses were performed on all BGC genes, both with and without pseudogenes and
187 outgroup taxa (Supplementary Data S7 and S8). Outgroup taxa were obtained by searching a
188 database of 529 genome annotations (Gluck-Thaler and Slot 2018) using blastp. Protein sequence
189 datasets for each gene were aligned using mafft v. 7.221 (Katoh and Standley 2013), and
190 ambiguously aligned characters were removed using TrimAl v. 1.4 (Capella-Gutierrez et al. 2009).
191 Maximum likelihood analysis was performed in RAxML v. 8.2.9 (Stamatakis 2014) with
192 automated model selection and topological robustness was assessed by 100 bootstrap replicates. In
193 order to evaluate alternative hypotheses versus inferred HGT events we applied minimal
194 topological constraints to exclude putative transferred genes from the donor clade. Constrained
195 trees (Supplementary Data S9) were built with automated model selection and their likelihoods

196 were compared using the Approximately Unbiased test with 10,000 multiscale bootstrap replicates
197 (Shimodaira 2002) as implemented in IQ-TREE v. 1.6.12 (Nguyen et al. 2014).
198 In order to determine synteny in the BOT and BOA loci (Supplementary Data S10) each locus
199 including up to 10 genes on either side of the BOA/BOT genes of interest (if present) were
200 combined and assigned to a homology group using usearch cluster_agg method with a minimum
201 linkage identity of 0.6 in usearch v. 8.0.1517 (Edgar 2010). The loci were then manually aligned
202 according to their homology group and manual blasts were performed to confirm true orthology
203 where ambiguous.
204 Ancestral state reconstructions (Supplementary Data S11) were performed using a
205 substitution matrix weighted against gain of functional genes and pseudogenes, except where HGT
206 was already determined by gene trees and synteny analysis for BOA clusters in Mesquite v 3.6
207 (Maddison and Maddison 2019).

208 **Results**

209 **Sequencing and assembly**

210 Eight *Botrytis* species and *Sclerotium cepivorum* (Table S1) were sequenced using long read single
211 molecule technology at 34-120 X coverage. The genome assembly sizes ranged from 42.98 Mb to
212 61.28 Mb (Table 1). The genomes of six species are similar in size to the previously described
213 genome of *B. cinerea* (43.5 Mb; van Kan et al. 2017), while genomes of *B. squamosa, B. sinoallii*
214 and *S. cepivorum* exceed a size of 54 Mb. The *B. aclada* genome could be assembled into 16 distinct
215 chromosomes, with 8 chromosomes containing telomeric repeats at both ends, and 6 containing a
216 telomeric repeat on one end.

217

218 **Table 1.** Assembly and gene prediction information of *Botrytis* spp. genomes from this study.

Species	Contigs	Assembly Size	Largest Contig	N50	BUSCO complete/partial	Predicted genes	Secretome size	% of secreted proteins
B. byssoidea[a]	59	42.98 Mb	2599 Kb	1263 Kb	98.0 (99.3)	12212	898	7.35
B. globosa[a]	27	45.68 Mb	4093 Kb	2511 Kb	98.0 (99.0)	12073	864	7.16
B. elliptica[a]	137	47.66 Mb	2119 Kb	652 Kb	99.2 (99.9)	12442	932	7.49
B. squamosa[a]	29	54.60 Mb	4659 Kb	2938 Kb	98.7 (99.1)	11963	897	7.5
B. deweyae[a]	76	44.36 Mb	2431 Kb	1076 Kb	98.0 (99.0)	12480	942	7.55
B. sinoallii[a]	47	61.28 Mb	6466 Kb	2252 Kb	98.3 (99.5)	12281	885	7.21
B. porri[a]	31	46.78 Mb	4253 Kb	2706 Kb	98.2 (98.9)	12088	888	7.35
B. aclada[a]	16	48.31 Mb	4155 Kb	3028 Kb	99.1 (99.3)	11870	867	7.30
S. cepivorum[a]	48	55.66 Mb	4533 Kb	1651 Kb	98.2 (99.5)	11107	790	7.11

219

220 [a] Taxa in the table are ordered as they appear in the phylogenetic tree in Figure 1.

221 The most fragmented assembly of the nine species is that of *B. elliptica*, despite its genome size of
222 <48 Mb, with 137 contigs and a contig N50 of 652 Kb. BUSCO analysis indicated that all genomes
223 had a high level of completeness (98.0-99.2%). Prediction of gene models was performed using
224 the FunGAP pipeline and supported by RNAseq data (from *in vitro* samples and infected plant
225 material) and by alignment to the manually curated genome of *B. cinerea* B05.10 (van Kan et al.
226 2017). After prediction by this pipeline, proteomes of *B. aclada, B. squamosa* and *S. cepivorum*
227 were entirely manually curated, while for the other six species, only the (predicted) secreted
228 proteins were manually curated. The curated proteomes of the nine species contain between 11,107
229 and 12,480 genes (Table 1).
230

231 **Phylogenetics and phylogenomics**

232 A phylogenetic tree was constructed based on a concatenated amino acid alignment of 4,746
233 conserved core genes totalling 409,576 positions, using *Marssonina brunnea* (order Helotiales,
234 family Dermataceae) as the outgroup (Fig. 1). The relationship among the *Botrytis* species is fully
235 concordant with previous studies (Hyde et al. 2014; Staats et al. 2005), which divided the genus in
236 two clades based on three protein-coding genes (G3PDH, HSP60 and RPB2). All *Botrytis* species
237 newly sequenced in this study group in Clade 2, which contains taxa that mostly infect monocot
238 host plants (only *B. paeoniae* infects dicots). A pan-genome analysis for 16 *Botrytis* species (eight
239 species sequenced in this study, seven species previously sequenced with short read technology
240 (Valero-Jiménez et al. 2019) and the previously sequenced *B. cinerea* B05.10 (van Kan et al. 2017),
241 indicated that the core genome of *Botrytis* spp. consists of 7,524 orthogroups (>60% of genes
242 within any individual species; Fig. S1a), while the pan-genome consists of 13,856 orthogroups
243 (Fig. S1b).

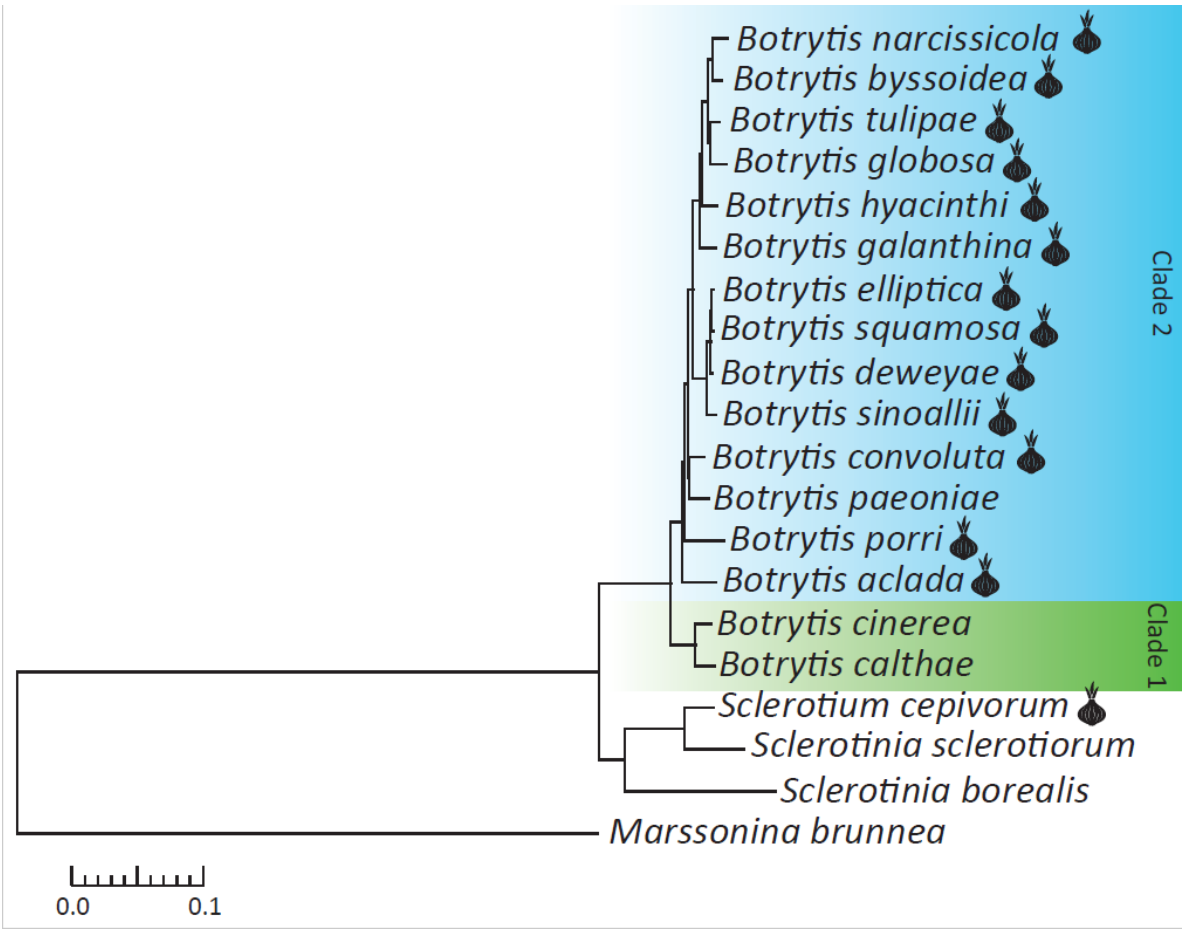

Fig. 1. Phylogenetic tree based on single-copy orthologous genes of different *Botrytis* species and three Sclerotiniaceae, with *Marssonina brunnea* as the outgroup to root the tree. All branches have a high bootstrap support (ML > 90). Two clades previously reported in the genus *Botrytis* are highlighted. The bulb symbols next to the species names indicate species that infect monocotyledonous bulbous plants, species without symbol infect dicot hosts.

Analysis of secreted proteins

Secreted proteins are important tools of plant pathogenic fungi to either manipulate the physiology and immune responses of their host plants (effector proteins) or to decompose the plant tissue that they colonize in order to acquire carbohydrate nutrients (plant cell wall degrading enzymes, PCWDEs). Orthologous groups of all secreted proteins from 16 *Botrytis* species sequenced in this work, as well as previously published (van Kan et al. 2017; Valero-Jiménez et al. 2019) and *S. cepivorum* were determined using Orthofinder. From a total of 14,838 proteins, 14,326 were

256 assigned to 1,116 orthologous groups (Supplementary Data S1). From these, 376 orthologous
257 groups are shared among all 17 species (Fig. S2). Besides orthologous groups shared by all species,
258 171 groups (columns 2-18 in Fig. S2) are common to all species but one, while 454 orthologous
259 groups are unique to a single species (columns 19-37 in Fig. S2). The secretome of *S. cepivorum*
260 lacks 55 secreted proteins that are present in all *Botrytis* species, and contains 83 singletons that
261 are unique to *S. cepivorum*, as to be expected for a species from a distinct genus in the same family.
262 In view of the relevance of secreted effector proteins in fungus-plant interactions, an
263 effector prediction was performed on the set of secreted proteins discussed above. For each of the
264 16 *Botrytis* species and *S. cepivorum,* a total of 121-152 candidate effector genes was identified
265 which were assigned to 244 orthologous groups (Supplementary Data S2). Among these groups,
266 25 are represented in all 17 species and another 25 are shared among all but one species. On the
267 other hand, each of the 17 species contains between 8 and 39 predicted effector genes that remained
268 unassigned to orthologous groups, since they are unique for that single species. There were no
269 predicted effectors which are shared among *Allium* pathogens but absent from non-*Allium*
270 pathogens. Furthermore, pairwise comparisons between related *Botrytis* species with distinct hosts
271 did not identify any effector genes that stood out as potential determinants for host specificity.
272 We also analysed the secreted proteins that are related to the degradation of plant cell wall
273 carbohydrates (Table S2). The genomes of 16 *Botrytis* spp. and *S. cepivorum* contain between 109
274 and 132 plant cell wall degrading enzymes (PCWDEs). *S. cepivorum* has fewer PCWDE-encoding
275 genes than the *Botrytis* species. The PCWDEs were further subdivided depending on their
276 substrate: cellulose, hemicellulose or pectin. The numbers of secreted enzymes capable of
277 degrading cellulose, hemicellulose and pectin were mostly similar among *Botrytis* spp., with some
278 deviations: *B. sinoallii* has notably fewer genes encoding pectinases (22 vs. 27-38 for other species;
279 Table S2).

280 **Secondary metabolite gene clusters**

281 Fungi produce a wide array of secondary metabolites (SM), usually synthesized by proteins
282 encoded by genes that are physically clustered in the genome, referred to as SM biosynthetic gene
283 clusters (BGCs) (Keller et al. 2005). SM contribute to the adaptation and survival in different
284 environments and in the competition with other (micro)organisms (e.g. Chatterjee et al. 2016). In
285 a previous study on nine *Botrytis* genomes assembled from short sequence reads, a patchy
286 absence/presence pattern was observed for orthologs to BGCs that were functionally annotated in
287 *B. cinerea* (Valero-Jiménez et al. 2019). Because of the fragmented assemblies resulting from short
288 read sequencing technology, the latter analysis only considered SM key biosynthetic enzymes, but
289 not the entire gene cluster. In the present study, the analysis of SM gene clusters was extended to
290 all 16 *Botrytis* species (short and long read technology-based), four related taxa from the family
291 Sclerotiniaceae and 25 other taxa from the class Leotiomycetes, for which an annotated genome
292 was publicly available (Table S3). The analysis was conducted by predicting BGCs in all 45
293 genomes using AntiSMASH, and grouping them by families using BiG-SCAPE. The 45
294 Leotiomycete genomes each contained between 3 and 67 BGCs (Supplementary Data S3). The
295 1571 BGCs were grouped over 438 BGC families (Supplementary Data S4), which were further
296 categorized based on their phylogenetic distribution. Category 1 contains 342 families of SM BGCs
297 that are distributed among taxa across Leotiomycetes. This category includes a few BGCs that
298 encode enzymes involved in biosynthesis of common metabolites such as melanin and
299 siderophores, however, the exact chemical structures of compounds produced by the vast majority
300 of BGCs in this category remain unknown. Category 2 contains 36 families of BGCs that are
301 present in Sclerotiniaceae (including the genus *Botrytis*) but not represented in the other 25
302 Leotiomycete taxa. This category includes the BGCs encoding enzymes involved in production of
303 botcinic acid, and other yet unknown compounds. Category 3 contains 60 families of BGCs that

304 are unique to the genus *Botrytis*, such as the cluster involved in production of botrydial, however,
305 all other SMs produced by the other 59 BGCs in this category are unknown.

306 BGCs are commonly annotated on the basis of the type of compound that is produced, often
307 a polyketide (PKS), non-ribosomal peptide (NRPS) or terpene (TS). The evolutionary trajectory of
308 BGCs can be complex, and the distribution of specific BGCs can be scattered throughout the fungal
309 kingdom (Slot and Gluck-Thaler 2019). Several cases of horizontal gene transfer of BGCs have
310 been documented in fungi (Campbell et al. 2012; Navarro-Munoz and Collemare 2020; Ropars et
311 al. 2015; Reynolds et al. 2018). We examined the distribution of the predominant classes of BGCs
312 (PKS, NRPS, TS) over the 45 Leotiomycete species analysed (Fig. 2 for PKS; Fig. S3 for NRPS;
313 Fig. S4 for TS).

314 The distribution of BGCs is largely consistent with phylogenetic patterns, with related
315 fungal taxa containing a similar distribution. A set of 20 PKS families (as identified by BiG-
316 SCAPE) are most abundant in *Botrytis* species. Six families from this set are exclusive to *Botrytis*
317 (highlighted red in Fig. 2), while 14 families are also present in other Sclerotiniaceae or in more
318 distantly related Leotiomycete taxa (highlighted in ochre). Conversely, a set of nine PKS clusters
319 that are most abundant in Leotiomycetes outside the family Sclerotiniaceae have sparse and patchy
320 distributions within the genus *Botrytis* (highlighted in blue).

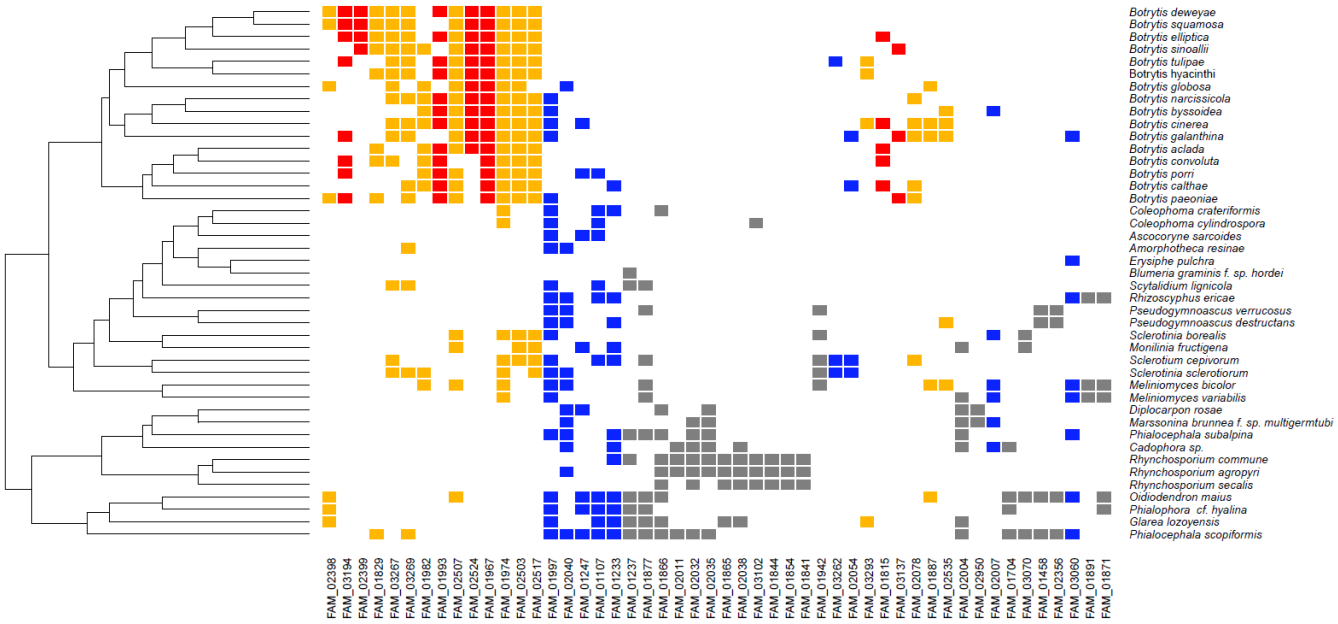

Fig. 2. Distribution of PolyKetide Synthase clusters in 45 Leotiomycetes. The 50 clusters that are most abundant among the 45 Leotiomycetes taxa are displayed. Clusters that are exclusively represented in *Botrytis* are marked red; clusters predominantly in *Botrytis* but also in some other taxa are marked ochre; clusters predominantly in other taxa but also in some *Botrytis* species are marked blue; clusters lacking in all *Botrytis* spp. are marked grey.

Botrytis species possess at least 5 (*B. convoluta*) and at most 11 (*B. cinerea*) NRPS clusters (Fig. S3). Five families of NRPS clusters are unique to the genus *Botrytis* (Fig. S3, highlighted in red), while eight other families are largely confined to the family Sclerotiniaceae, although two of them (FAM_02547 and FAM_02047) are also shared with the distant taxa *Phialophora hyalina* or *Phialocephala scopiformis* (Fig. S3, highlighted in ochre). Notably, *B. cinerea* contains two NRPS clusters that are not shared with any other *Botrytis* species, but have orthologs in several distant Leotiomycetes (Fig. S3, highlighted in blue). The families of terpene cyclase (TS) clusters are relatively simple in pattern, with each *Botrytis* species containing 3-6 TS cluster families (Fig. S4). Eight of the families are exclusively detected in *Botrytis* species (Fig. S4, highlighted in red) while four are also present in other Sclerotiniaceae, and two of the TS cluster families are even detected in distant Leotiomycetes (Fig. S4, highlighted in ochre). The family FAM_03197 is conserved in all Sclerotiniaceae, as well as in 6 other Leotiomycetes while FAM_02531 is present in nine

339 Sclerotiniaceae and six distant Leotiomycetes. Except for the family FAM_02168, involved in the
340 synthesis of the phytoxic metabolite botrydial, the chemical nature of the products of these clusters
341 is unknown.

342 **Ancestral genome reconstruction of the genus *Botrytis* and the family Sclerotinaceae**

343 The high quality of the long read assemblies and the previously published *B. cinerea* genome, as
344 well as the extensive manual curation effort of gene models, enabled us to perform a synteny
345 analysis and a reconstruction of the ancestral chromosome configuration of the genus *Botrytis*, in
346 order to understand the extent and nature of chromosomal rearrangements over the course of
347 evolution of the extant species. *B. elliptica* was excluded from the ancestor reconstruction for two
348 reasons: firstly, the assembly was the most fragmented of all (137 contigs) and secondly, the
349 phylogenetic relation of *B. elliptica* to its sister taxa *B. squamosa* and *B. deweyae* could not be
350 resolved (Fig. 1), which hampered the analysis. The inferred ancestral genome of the entire genus
351 *Botrytis* (AB0) consists of 17 syntenic blocks (Fig. 3). 13 of the 16 *B. cinerea* core chromosomes
352 are entirely syntenic to the AB0 ancestor, and 17 balanced rearrangements (mostly inversions) are
353 inferred between the ancestor AB0 and the extant *B. cinerea* (Table 2; Table S4).

354 **Table 2.** Numbers of balanced genomic rearrangements between inferred ancestral genomes (AB0-A6) and extant
355 *Botrytis* species, as shown in Fig. 3. Further details of the types of rearrangements are provided in Table S4.

	AB0-BCIN[a]	AB0-A1	A1-BACL[b]	A1-A2	A2-A3	A3-A4	A3-A6	A4-A5	A4-BSIN[c]	A5-BSQU[d]	A5-BDEW[e]	A6-BBYS[f]
Inversions	15	0	3	0	2	3	6	1	6	2	1	3
Translocations	1	0	1	0	0	0	1	0	2	2	0	4
Transpositions	1	0	1	0	0	0	0	0	1	0	0	0
Fusions	0	0	0	1	2	1	0	1	2	1	2	0
Fissions	0	0	5	1	1	6	2	1	19	3	40	27
Sum	**17**	**0**	**10**	**2**	**5**	**10**	**9**	**3**	**30**	**8**	**43**	**34**

356 [a]BCIN: *B. cinerea*; [b]BACL: *B. aclada*; [c]BSIN: *B. sinoallii*; [d]BSQU: *B. squamosa*; [e]BDEW: *B. deweyae*; [f]BBYS: *B. byssoidea*

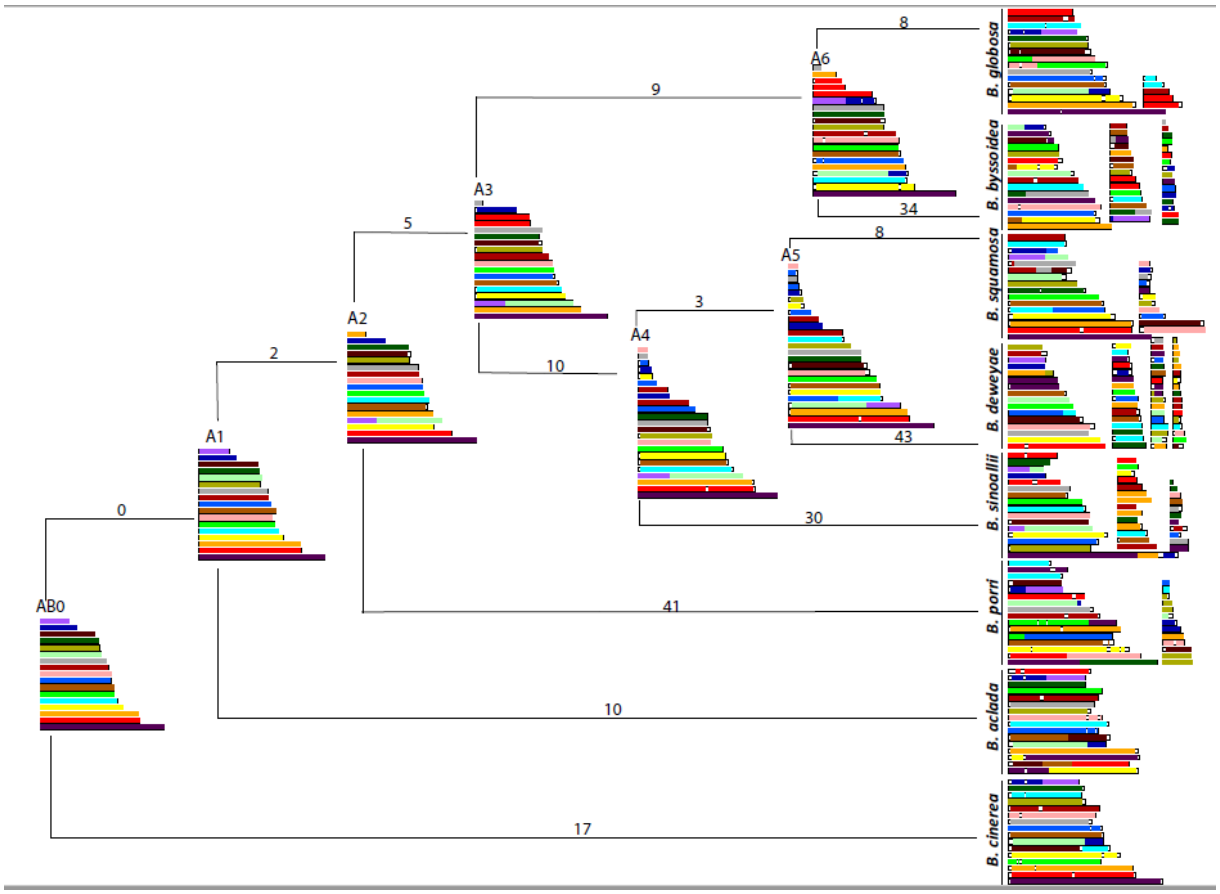

Fig. 3. The most parsimonious evolutionary trajectory from the ancestral (A0) configuration towards extant *Botrytis* species. Coloured boxes represent syntenic blocks. A1-A6 represent intermediate ancestors. Numbers above the branches represent the total number of balanced rearrangements (interchromosomal translocations and fusions/fissions; intrachromosomal inversions) accumulated between two genomes.

The A1 genome is the inferred ancestor of members of clade 2 in the genus *Botrytis*, while *B. cinerea* is the single representative of clade 1 in the analysis (Fig. 1). The inferred A1 genome is identical to AB0 (Fig. 3). The extant *B. aclada* genome contains 10 rearrangements as compared to A1. The A2 intermediate ancestor was inferred to be derived from A1 upon fusion of A1 contigs 13 and 17, and fission of A1 contig 3 (resulting in A2 contigs 5 and 17). Downstream of the A3 intermediate ancestor, the interpretation becomes complex as numbers of contigs increase due to the more fragmented assemblies of some species, e.g. *B. deweyae, B. byssoidea* and *B. sinoallii.* Nonetheless, the number of contigs of intermediate ancestors remains 25 or lower and the number of rearrangements between nodes in the tree ranges from 3 to 43 (Table 2).

372 Reconstruction of ancestral genomes was extended to the family Sclerotiniaceae using the
373 genomes of *Sclerotium cepivorum* (this study) and *Sclerotinia sclerotiorum* (Derbyshire et al.
374 2017) (Fig. S5). Due to the more fragmented assembly of the *S. cepivorum* genome, the inferred
375 common ancestor AS1 comprised 21 syntenic blocks, five of which were quite small and probably
376 represent only parts of chromosomes. However, the common ancestor ABS0 of the family
377 Sclerotiniaceae contains 16 syntenic blocks, and the configuration of ABS0 differs from the
378 ancestral *Botrytis* genome AB0 by just a single rearrangement (Fig. S5).
379

380 **Synteny between *B. aclada* and *B. cinerea***

381 In order to explore genome rearrangements between individual species in more detail, we further
382 examined the synteny between the genomes of *B. aclada* and *B. cinerea* (the most complete and
383 best annotated) by pairwise alignments. *B. cinerea* minichromosome 18 (BCIN18) was excluded
384 from this analysis because it contains only 13 genes, none of which is orthologous to genes in *B.*
385 *aclada.* The second minichromosome of *B. cinerea*, BCIN17, did show some homology to the tip
386 of BACL10 and was therefore included in the analysis. Graphical representation of the alignment
387 (Fig. 4) reveals that four chromosomes represent fully syntenic blocks, though some of these blocks
388 contain segmental inversions of ancestral regions on the same chromosome (not visible in the
389 colour display). In the remaining 12 chromosomes, the alternation of coloured boxes reflects the
390 occurrence of six interchromosomal rearrangements, as well as 13 small translocations or
391 transpositions, of which seven occurred at or close to the telomeres (Fig. 4).

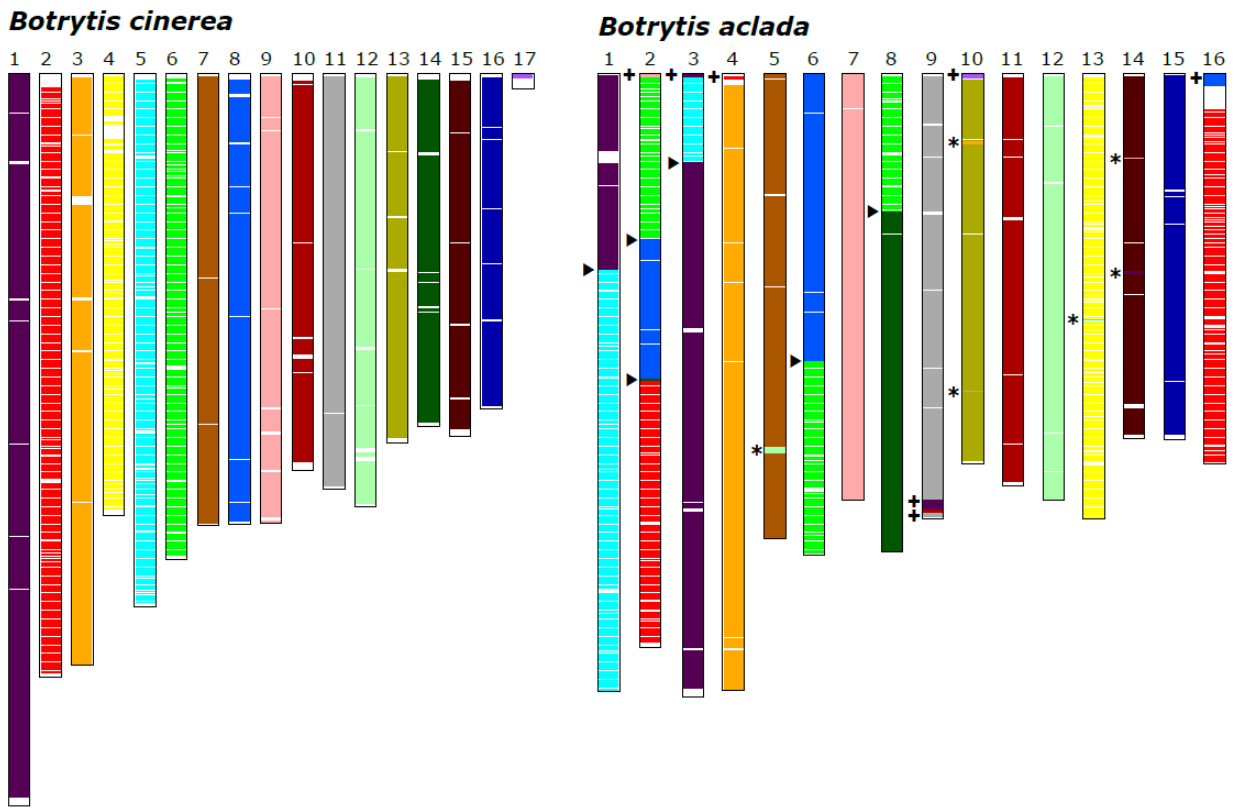

Fig. 4. Synteny analysis between *B. aclada* and *B. cinerea.* The 17 chromosomes of *B. cinerea* are colour-coded uniformly, the corresponding syntenic regions in *B. aclada* have identical colours. White regions reflect repetitive regions or lack of homology. Arrowheads indicate large reciprocal interchromosomal rearrangements. Asterisks indicate small interchromosomal transpositions. Plus symbols indicate interchromosomal telomeric translocations. Intrachromosomal inversions are not indicated.

Strikingly, we noted that SM BGCs were present in some of these translocated segments. Specifically, BACL05 is almost perfectly syntenic to BCIN07, with the exception of an insertion of a cluster of seven genes (Fig. 4, green box marked by an asterisk) representing the BGC for the sesquiterpene metabolite botrydial (Pinedo et al., 2008; Porquier et al. 2016; Siewers et al. 2005), which in *B. cinerea* is located in BCIN12. Conversely, the only difference between BACL12 and BCIN12 is the insertion (in BCIN12) of a segment that exactly contains the BGC for botrydial. Furthermore, BACL9 is entirely syntenic to BCIN11, however, it contains an insertion of the BGC for the phytotoxic metabolite botcinic acid (Dalmais et al. 2011; Porquier et al. 2019) close to the 3'-telomeric region, which in *B. cinerea* is located at the start of BCIN01 (van Kan et al. 2017).

408 **Genomic locations of botrydial and botcinic acid biosynthetic gene clusters**

409 The synteny analyses described above provided indications that SM BGCs occur in regions that
410 possibly underwent translocation at some moment in the evolution of *Botrytis* species. The
411 distribution of botrydial (BOT) and botcinic acid (BOA) BGCs over the Sclerotiniaceae and the
412 genus *Botrytis* appeared to be patchy. Specifically, the BOT cluster is present in 8 *Botrytis* species
413 and absent in other Sclerotiniaceae. We compared the BOT clusters and their flanking sequences
414 in 7 species: *B. aclada, B. cinerea, B. elliptica, B. deweyae, B. porri, B. sinoallii, B. squamosa.* The
415 *B. peaoniae* genome, though containing a BOT cluster, was sequenced by Illumina technology
416 (Valero-Jiménez et al. 2019) and its assembly was too fragmented for synteny analysis. The order
417 of the genes BcBOT1-7 within the cluster was identical in all species, however, the most upstream
418 gene (BcBOT4), was in inverted orientation in *B. aclada* and *B. porri* as compared to the other five
419 species (Fig. 5). The BOT clusters were in all cases flanked by gypsy/copia repeats, with lengths
420 up to 160 kb, either on one side (*B. cinerea, B. deweyae, B.elliptica, B. squamosa*), or on both sides
421 (*B. porri, B. aclada, B. sinoallii*) and some species even contained internal transposon repeats
422 within the BOT cluster (*B. aclada, B. cinerea,* and *B. sinoallii*; Fig. 5). Based on the RNAseq reads
423 used for structural annotation, it was observed that all species that do contain intact BOT clusters
424 express all of the seven genes. As these expression data were based on pooled RNAs, representing
425 multiple fungal tissue types and infection stages, it was not possible to compare the expression
426 levels between species or to determine under which conditions the genes were expressed.

427 The BOA cluster was detected, in whole or in part, in all but one *Botrytis* species (*B. paeoniae*),
428 and in *Sclerotinia sclerotiorum* as well as *Sclerotium cepivorum*. In many cases, the BOA cluster
429 in *Botrytis* species is located close to the end of a contig. It was previously reported that in *B.*
430 *cinerea*, the BOA cluster is at the very start of BCIN01, only 5 kb away from the telomere (van
431 Kan et al. 2017). Alike for the BOT clusters mentioned above, all species that contain intact BOA

clusters express all of the 13 genes, however, the use of pooled RNAs prevented us from comparing expression levels between species or determine under which conditions the genes were expressed.

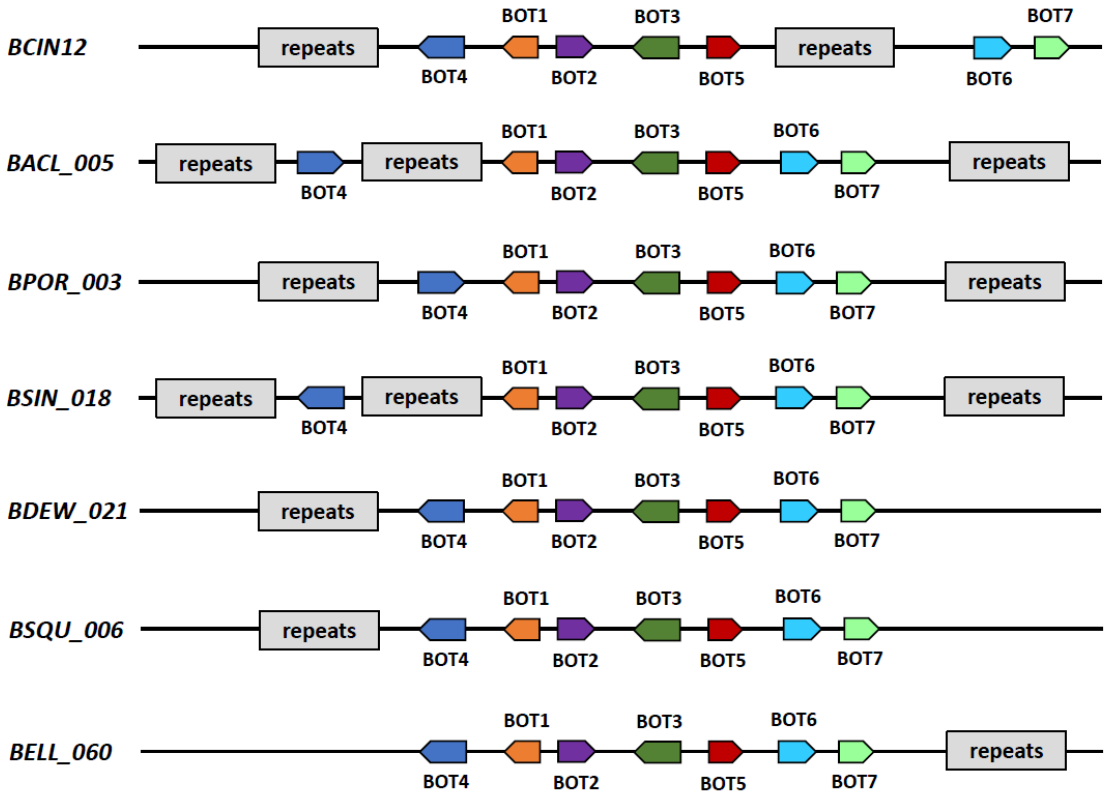

Fig. 5. Organization of BOT clusters in seven *Botrytis* species. BCIN: *B. cinerea*; BACL: *B. aclada*; BPOR: *B. porri*; BSIN: *B. sinoallii*; BSQU: *B. squamosa*; BDEW: *B. deweyae*; BELL: *B. elliptica.* The number of the contig is given behind the species name tag. The seven BOT gene orthologs (not drawn to scale) are colour-coded uniformly, the arrow indicates direction of transcription. Repeats are indicated with a grey box. Repeats are not drawn to scale and range in length from 1-160 kbp.

In view of the high synteny between *Botrytis* species, we examined whether the BOT and BOA clusters in the different species are in syntenic locations as compared to *B. cinerea.* Surprisingly, analysis of flanking genes revealed that BOT clusters are in four distinct genomic regions in the seven *Botrytis* species analysed. None of the species other than *B. cinerea* contained the BOT cluster in a region syntenic to BCIN12 (Fig. 6). The genes directly flanking the BOT cluster in *B. cinerea* (Bcin12g06360 and Bcin12g06440) in all but one of the six species have orthologs that are directly adjacent to one another in these genomes, with intergenic regions ranging from 2-5 kbp.

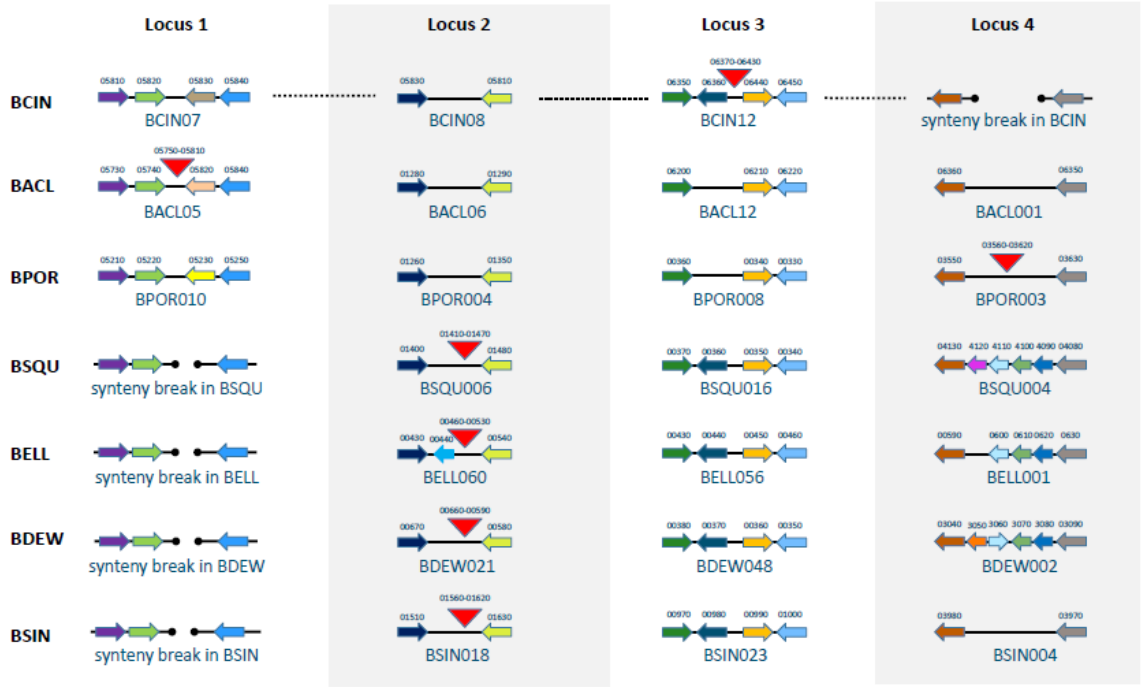

Fig. 6. Distinct genomic locations of BOT clusters in seven *Botrytis* species. Four different loci are provided in the columns. Species name tags are in the left hand margin: *B. cinerea*; BACL: *B. aclada*; BPOR: *B. porri*; BSIN: *B. sinoallii*; BSQU: *B. squamosa*; BDEW: *B. deweyae*; BELL: *B. elliptica.* Contig numbers in the seven species are provided underneath the locus. In each column, orthologous genes are indicated by identical colours. Gene numbers in the contig are provided above the gene, the arrow indicates direction of transcription. The red triangular blocks represent the location of a BOT cluster. Synteny breaks are shown by interrupted lines with dots marking the break.

No indication was found for the occurrence of truncated remnants of BOT genes at this position in the six genomes. Also in all but one of the other species lacking a BOT cluster, orthologs to Bcin12g06360 and Bcin12g06440 are directly adjacent to one another in these genomes. Through similar analyses and reasoning, the BOT cluster in *B. aclada* is present in a unique position that is syntenic to BCIN07, while the BOT cluster in *B. porri* is present in a unique position that is syntenic to positions in five other species (all except in *B. cinerea*, where a synteny break has occurred); lastly, the BOT clusters in *B. squamosa, B. deweyae, B.elliptica* and *B. sinoallii* are all located in a syntenic genomic region, which is equivalent to a location between Bcin08g05830 and Bcin08g05810 (Fig. 6).

465 In *S. sclerotiorum* the BOA cluster is dispersed over two chromosomal locations on
466 SSCL05 (genes BOA1 and BOA2) and SSCL15 (genes BOA3-13). A recent study by Graham-
467 Taylor et al. (2020) reported that SSCL can express the 13 BOA genes in a co-regulated manner
468 despite their spatial separation. For the largest cluster on SSCL15, its flanking genes on both sides
469 are orthologous to syntenic regions in eight *Botrytis* species (BACL006, BBYS014, BCIN06,
470 BDEW005, BELL059, BGLO010, BSIN006, BSQU018) that do not contain any trace of BOA
471 gene remnants. For the smaller cluster on SSCL05, its flanking genes on both sides are orthologous
472 to genes located on BCIN05 (Bcin05g05060 and Bcin05g07100), however the region is not
473 syntenic, since the genes are far separated in *B. cinerea.*

474 **Inheritance and structural evolution of BOT and BOA clusters**

475 BOT and BOA gene loci were carefully examined for evidence of pseudogenization to infer which
476 of the clusters are fully functional (Supplementary Data S5, S6). BOT clusters in *B. sinoallii* and
477 *B. paeoniae* contain one and two pseudogenes, respectively, while six species (*B. aclada, B.*
478 *cinerea, B. elliptica, B. deweyae, B. porri,* and *B. squamosa*) have clusters with seven apparently
479 functional genes (Supplementary Data S5). 17 of the 19 Sclerotiniaceae analysed contained (parts
480 of) BOA clusters, however only seven species (*B. aclada, B. byssoidea, B. cinerea, B. globosa, B.*
481 *porri, B. sinoallii* and *S. sclerotiorum*) appeared to contain a fully functional BGC (Supplementary
482 Data S6). The majority of species contain two or more pseudogenes of catalytic enzymes. The most
483 extreme cases of gene loss were in *B. squamosa, B. deweyae, B.elliptica* and *B. tulipae*, which lost
484 all but one of the BOA cluster genes. By contrast, *B. calthae, B. convoluta,* and *B. narcissicola*
485 contained 2-3 pseudogenes, either in genes encoding accessory enzymes or in the BOA13 gene,
486 which is the transcriptional regulator for the cluster (Porquier et al. 2019). Of the two species
487 outside the genus *Botrytis*, *S. sclerotiorum* contains a functional BOA cluster (Graham-Taylor et

488 al., 2020), whereas *S. cepivorum* lacks four genes, including polyketide synthase gene BOA9, and
489 in addition contains two pseudogenes.

490 Ancestral state reconstructions of genes and pseudogenes (considering horizontal gene
491 transfer [HGT] events, see below) on the *Botrytis* species tree (Fig. 7) suggest that the BOT cluster
492 was gained in the common ancestor of *Botrytis* and has been lost five times; three times leaving no
493 gene remnants (in *B. calthae, B. convoluta* and in the subclade containing *B. galanthina*), and twice
494 leaving a mix of functional genes and pseudogenes (in *B. paeoniae* and *B. sinoallii*). The BOT gene
495 trees (Supplementary Data S7) are in agreement with the species tree and the clusters are thus
496 inferred to be derived from strictly vertical inheritance. Reconstructions of the BOA clusters (Fig.
497 7) revealed a more dynamic process involving twelve losses of cluster function after being gained
498 in the common ancestor of *Botrytis* and *Sclerotinia*, and two recent gains by HGT in *B. cinerea*
499 and *B. sinoallii*. HGT of the two clusters is supported by maximum likelihood gene trees
500 (Supplementary Data S8), which suggest that both clusters were acquired from a relative of *B. porri*
501 or *B. aclada*.

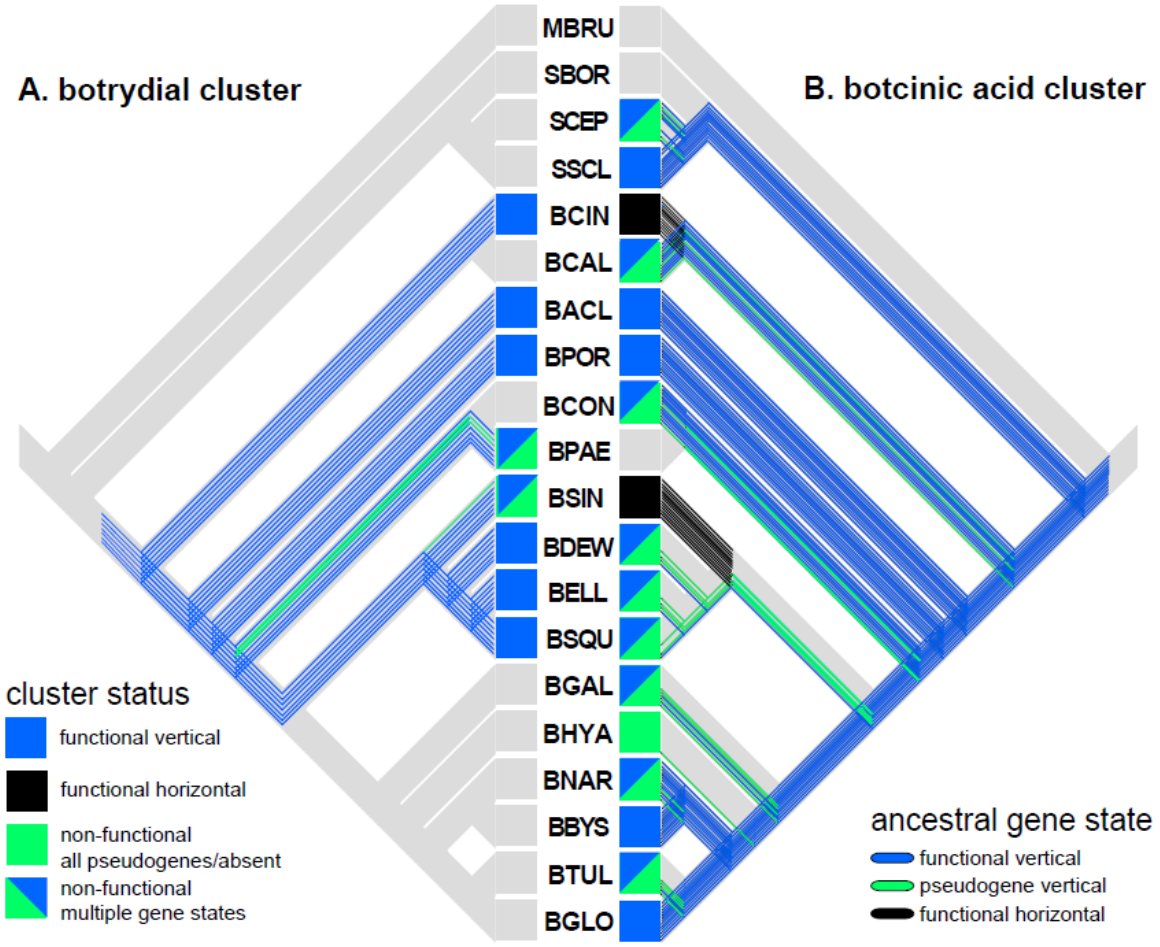

Fig. 7. Ancestral state reconstructions of genes and pseudogenes of the BOT cluster (panel A) and the BOA cluster (panel B) on the phylogenetic tree of 20 Sclerotiniaceae species. The status of the cluster is indicated with coloured boxes: blue = functional cluster, vertically transmitted; black = functional cluster, horizontally transmitted; green = all cluster genes non-functional or absent; green/blue = some genes non-functional or absent; grey = total absence of cluster. The ancestral gene states of 7 BOT genes and 13 BOA genes are indicated with coloured lines in similar way.

Most gene trees became significantly worse than the maximum likelihood trees, according to Approximately Unbiased tests (Shimodaira 2002), when potential HGT homologs were excluded from the putative donor clade (Supplementary Data S9). Strong support for a HGT origin of the functional BOA cluster in *B. sinoallii* comes from two additional observations. First, the inferred HGT cluster is adjoined by a putative amino acid transporter (Bsin003g06700) and alcohol acetyltransferase (Bsin003g06560), which are either adjacent or a few genes removed from the

515 BOA cluster in *B. aclada*; only the homolog of Bsin003g06700 is adjacent to the BOA cluster in
516 *B. porri* (Supplementary Data S10). Secondly, *B. sinoallii* contains an additional, heavily
517 pseudogenized BOA cluster on contig BSIN027, which more closely tracks the species phylogeny
518 (S8 Data) and retains flanking genes that are consistent with the species phylogeny (Supplementary
519 Data S10). The remnants of the ancestral *B. sinoallii* BOA cluster comprise only three pseudogenes
520 that are embedded in a 330 kb genomic region saturated with transposons.

521 The HGT of the BOA cluster to *B. cinerea* is supported by the phylogenetic proximity to
522 *B. aclada* and *B. porri* (Fig.7; Supplementary Data S8), however, it cannot be corroborated by
523 synteny information, as the *B. cinerea* BOA cluster is located at the start of chromosome 1, and the
524 25 kbp region immediately downstream of the cluster is not syntenic with any *Botrytis* species.

525 **Discussion**

526 Following the efforts to sequence *Botrytis cinerea* isolate B05.10 and nine other *Botrytis* species
527 mainly infecting flower bulb crops (Valero-Jiménez et al. 2019), the present study, focussing on 8
528 species from clade 2 of the genus, brings the number of *Botrytis* genome sequences to 16. This
529 represents about half of the currently recognized species in the genus, though a recent study
530 (Garfinkel et al. 2019) identified at least 15 phylogenetically distinct, new taxa sampled from
531 *Paeonia* in Alaska, which remain to be described and named. There is thus far one single fungal
532 genus, i.e. *Verticillium*, for which the genomes of all recognized species have been sequenced (Shi-
533 Kunne et al. 2018). It will take more effort to complete the sequencing of the entire genus *Botrytis.*

534 The present study aimed to identify genes potentially involved in determining host
535 specificity, by comparing genomes of *Botrytis* species pathogenic on *Allium* with each other and
536 with the genomes of their closest relatives pathogenic on other host plants. Specifically, we
537 compared the genomes of the onion (*Allium cepa*) pathogens *B. squamosa* and *B. sinoallii,* with
538 those of their sister taxa *B. elliptica* and *B. deweyae,* which infect lily and *Hemerocallis*,
539 respectively, and we compared the genomes of *B. aclada* (infecting onion) and *B. porri* (infecting
540 *Allium porri,* leek) with that of *B. paeoniae* (infecting the dicot peony). In order to make a
541 meaningful comparison, the effort was made of manually curating all (>11,000) gene models in the
542 genomes of three species (*B. squamosa, B. aclada* and *S. cepivorum*), and manually curating the
543 gene models of all proteins with a (predicted) signal peptide in the other six species. Comparison
544 of the effector repertoires did not reveal candidate effectors that were shared among all *Allium*
545 pathogens but absent in non-*Allium* pathogens. Each of the species analysed contained 8-39
546 predicted effector genes that were unique to the species, however most had no homologs in other
547 fungi and these genes often had little RNA-seq support (even in RNA samples from infected onion
548 tissue), questioning the importance of these predicted genes for pathogenicity on onion. The

549 repertoire of cell wall degrading enzymes was also similar between all 16 *Botrytis* species studied,
550 despite the fact that only three species infect dicot hosts while the vast majority infect monocot
551 hosts. Dicots and monocots are considered to have different compositions of cell wall
552 polysaccharides (Jarvis et al. 1988). Thirteen *Botrytis* species in this study infect monocot hosts
553 from the families *Alliaceae, Amaryllidaceae, Iridiaceae* and *Liliaceae.* Plants from these families
554 contain high levels of pectin in their cell walls as compared to the *Poaceae* (Jarvis et al. 1988),
555 which are more intensively studied as they comprise major staple crops of global relevance: rice,
556 wheat, maize. In view of the high pectin content in the monocot hosts of *Botrytis* species in this
557 study, the large repertoire of pectin degrading enzymes in their genomes appears logical.
558 Altogether, we did not identify (sets of) genes that are shared among the *Allium* pathogens and
559 distinguish them from related species with different hosts. The lack of shared genomic features
560 may reflect the pathology of the *Allium* pathogens, some of which infect the leaves (*B. squamosa*),
561 while others infect the bulb (*B. aclada*) or the roots and scale bases (*S. cepivorum*).

562 Despite the failure to identify host specificity determinants, many interesting features were
563 unravelled by the extensive genome analyses that were performed. The genome of *B. aclada* was
564 assembled into 16 gapless chromosomes, eight of which were full-length (telomere-to-telomere)
565 and six contained telomeric repeats on one end. The *B. aclada* assembly was based on sufficiently
566 high coverage to avoid the requirement for short read-based correction, nor did it require an optical
567 map or genetic map for assembly verification, as was done for *B. cinerea* (van Kan et al. 2017).
568 Cytogenetic studies on four *Botrytis* species (*B. byssoidea, B. cinerea, B. squamosa, B. tulipae*)
569 revealed that they each contain 16 mitotic chromosomes, whereas the same study reported 16 or
570 32 mitotic chromosomes in different isolates of *B. allii* (Shirane et al. 1989). Subsequent studies
571 (Nielsen et al. 2001; Yohalem et al. 2003) revealed that the species earlier named *B. allii* in fact
572 comprised isolates of *B. aclada* (having 16 chromosomes) as well as isolates representing a hybrid

573 of *B. byssoidea* and *B. aclada* (having 32 chromosomes), which is presently still named as *B. allii*
574 (Staats et al. 2005). Strikingly, *Sclerotinia sclerotiorum* also contains 16 chromosomes (Amselem
575 et al. 2011; Derbyshire et al. 2017). These observations suggest a bias for the possession of 16
576 chromosomes in the genus *Botrytis* and possibly even in related genera. Conservation of
577 chromosome numbers is not commonly observed in fungal genera, especially Ascomycota. As an
578 example, the core chromosome numbers in the genus *Fusarium* vary from four (*F. graminearum*)
579 to 12 (*F. fujikuroi*) (Waalwijk et al. 2018). Could this conservation of chromosome numbers in
580 distant species of the same genus be related to functional constraints for sexual reproduction during
581 the evolution of *Botrytis* species? As sexual reproduction requires chromosome pairing during
582 meiosis, any fusion or fission event that affects core chromosome numbers would have serious
583 repercussion on sexual compatibility and the fertility of offspring. We further explored the
584 conservation of chromosome numbers and architecture by examining synteny and reconstructing
585 ancestral genomes of the genus *Botrytis* and the family Sclerotinaceae.

586 The ancestral genome reconstruction inferred as few as 17 syntenic blocks for the common
587 ancestor (AB0) of all *Botrytis* species. The inferred ancestral genome of the Sclerotiniaceae (ABS0)
588 consisted of 16 syntenic blocks, and it differed from the AB0 genome by a single rearrangement.
589 13 of the 16 core chromosomes of *B. cinerea* were represented in these blocks, and only three
590 interchromosomal rearrangements were proposed between the ancestor AB0 and the extant *B.*
591 *cinerea* genome. Moreover, the common ancestor of the entire genus (AB0) was identical to the
592 common ancestor of extant *Botrytis* species in clade 2. Only six interchromosomal rearrangements
593 were proposed between the genome of ancestor A1 and the extant *B. aclada* genome. The genomes
594 of *B. cinerea* and *B. aclada* were thus remarkably syntenic, considering the phylogenetic distance
595 between the two species. Representatives of the two clades within the genus *Botrytis* (Staats et al.,
596 2005) were recently included in molecular clock-based estimates of divergence times for

597 Ascomycota, and these species were estimated to have diverged 5.9 Million years ago (Shen et al.,
598 2020). The maintenance of 16 chromosomes and the stability of their overall configuration would
599 facilitate chromosome pairing during meiosis. This observation thus suggests the occurrence of a
600 strong selection pressure on sexual reproduction within the genus *Botrytis* over time. The
601 suggestion is further supported by the fact that *S. sclerotiorum* also possesses 16 chromosomes
602 (Derbyshire et al. 2017) and that the ancestral genome of the Sclerotiniaceae differs from the
603 ancestral *Botrytis* genome only by a single rearrangement, despite the divergence between the
604 genera *Sclerotinia* and *Botrytis* being estimated to have occurred around 21.5 Million years ago
605 (Shen et al., 2020). The extent of synteny among *Botrytis* species from distinct clades could only
606 have been retained if sexual reproduction in this genus has been prominent over the course of
607 evolution. Of the 22 *Botrytis* species used in the initial phylogeny of the genus (Staats et al. 2005),
608 14 were reported to have a sexual stage while eight were not, including *B. aclada*. Population
609 studies may shed more light on the modes of reproduction of *Botrytis* species. Thus far only *B.*
610 *cinerea*, *B. pseudocinerea*, *B. tulipae* and *B. elliptica* have been subject of population analyses
611 (Fournier et al. 2005; Giraud et al. 1999; Mercier et al. 2019; Soltis et al. 2019; Staats et al. 2007;
612 Walker et al. 2015) while other species have received less attention.

613 While synteny analyses indicated a strong overall conservation of chromosome architecture
614 between *Botrytis* species, it was striking to detect a substantial number of small translocations
615 between *B. cinerea* and *B. aclada*, both in telomeric and internal chromosomal regions. Telomeric
616 translocations are relatively "safe" rearrangements, as they have limited impact on genome
617 architecture and chromatin organization, minimizing the risk of causing major genome stress.
618 However, such rearrangements have the potential risk of (partial or complete) loss of the telomeric
619 region during the translocation. The BOA clusters that were detected in multiple *Botrytis* species
620 were, with two exceptions, located at the end of contigs, presumably because they were flanked by

621 repetitive sequences. In *S. sclerotiorum*, however, the BOA cluster is located internally in
622 chromosome SSCLE15, and it is not flanked by repetitive sequences. Although it seems logical to
623 propose a role of repetitive sequences in the translocation of chromosomal segments (whether
624 telomeric or internal), further studies need to establish such a role. Sequencing multiple isolates of
625 some of the species by long read technology might reveal the frequency of translocation events
626 within a species.

627 It was remarkable to note that the BOT clusters appears to be located in 4 distinct genomic
628 locations in the 7 *Botrytis* species in which it was analysed, and each of the loci was flanked by
629 transposons, and in three cases even interrupted by transposons. It is tempting to speculate that
630 these transposons have played a role in the mobility of the BOT cluster within the genome. The
631 phylogeny of the BOT gene clusters was in full agreement with the species phylogeny, arguing
632 against a horizontal transfer event. Thus the data suggest that there have been independent
633 translocations of the BOT gene cluster to distinct chromosomes, culminating in the four distinct
634 genomic locations presently observed in extant fungal isolates. Only within *B. squamosa, B.*
635 *deweyae* and *B. elliptica,* was the BOT cluster in the equivalent genomic location, as could be
636 expected from their phylogenetic proximity within a subclade of clade 2. This suggests a unique
637 transposition event in the lineage towards the common ancestor of species in this subclade (A5 in
638 Fig. 4). It is not currently possible to estimate the timing of these translocations, nor could the
639 position of the BOT cluster in the ancestral genome be inferred in the Anchro analysis.

640 Polymorphism in genomic locations of SM BGCs was recently described within a
641 collection of *Aspergillus fumigatus* isolates, suggesting that mobility of BGCs may occur even
642 within a single species. In this study, there was even one case of two isolates carrying idiomorph
643 BGCs, i.e. two distinct clusters residing in the same genomic locations (Lind et al. 2017). It will be
644 interesting to analyse multiple isolates of the different *Botrytis* species and explore whether

645 mobility of BGCs occurs within a single species as well. Long read sequence technology will be
646 essential for such purpose, to obtain flanking sequence information that permits to infer the correct
647 genomic locations of the various BGCs.

648 **The evolution and dynamics of BOT and BOA clusters**

649 The BGCs involved in the production of phytotoxic secondary metabolites BOT and BOA were
650 specifically interesting because they trigger (programmed) cell death in dicots (Rossi et al. 2011)
651 and in monocots (our unpublished results) and contribute to the virulence of *B. cinerea* (Dalmais
652 et al. 2011). The unusual observation of the distinct genomic locations of BOT and BOA clusters
653 encouraged us to explore two distinct evolutionary scenarios: that either clusters were vertically
654 transmitted but were able to excise from their location and reinsert at distinct locations; or that
655 clusters were lost and then regained through HGT. We carefully evaluated the functionality,
656 synteny and phylogeny of BOT and BOA genes and avoided assuming that vertical gene
657 duplication is the source of multiple paralogs within a lineage. Indeed half the BOA clusters
658 inferred to be functional in *Botrytis* appear to have been acquired by HGT from other *Botrytis*
659 species, and the functional BOA cluster in *B. sinoallii* is inferred to be a xenolog (horizontally
660 acquired paralog) of the pseudogenized cluster in the same species. The fact that the inferred donor
661 of the BOA cluster in *B. sinoallii* (a taxon closely related to *B. aclada* and *B. porri*), which also is
662 a pathogen of *Allium*, is consistent with host-specific functions selecting for cluster HGT. BGC
663 birth and death processes appear to involve the horizontal replacement of commonly lost clusters;
664 however the trajectories of BOT and BOA contrast in their evolutionary dynamics. While BOT is
665 less frequently lost/non-functionalized and has not been gained by HGT in this dataset, BOA is
666 frequently lost or non-functionalized and also replaced by HGT. It is possible that BOT is more
667 readily retained by natural selection due to its role in microbial competition (Vignatti et al. 2020).

668 This genome comparison has not revealed any host range determinants that enable so many
669 *Botrytis* species (and *S. cepivorum*) to infect *Allium* hosts, likely because fungus-plant interactions
670 may depend on a multitude of factors. Especially the fact that some of these species infect leaf
671 tissue, while others infect the bulb or the root, and some species induce blight symptoms while
672 others cause maceration and rot, adds another layer of complexity when comparing species
673 pathogenic on the same host. The high synteny and conservation of chromosome architecture
674 between such distant species across the genus *Botrytis* is remarkable and contrasts with the
675 dynamics of genome evolution in many other plant pathogens.

Acknowledgements

The authors acknowledge dr. Jeff Rollins (University of Florida, Gainesville, USA), dr. Jürgen Köhl (Wageningen Plant Research, The Netherlands), and dr. Jing Zhang and prof. Guoqing Li (Huazhong Agricultural University, Wuhan, China) for providing fungal isolates, as well as Alexander Wittenberg and Harrie Schneiders of Keygene NV (Wageningen, The Netherlands) for advice and support in preparing high quality DNA samples for the sequencing. Furthermore, the authors are grateful to Laura Vilanova Torren for her contribution in annotation of the genomes.

Data availability statement

The project has been deposited in GenBank under the Bioproject number PRJNA494516.

The Biosamples related to this project have accession numbers SAMN10219759-SAMN10219767. The raw PacBio genomic read data are deposited under accession numbers SRR8062108- SRR8062116.

Assembled genomes are deposited with accession numbers

RCSV00000000	SAMN10219759	BOTACL
RCSW00000000	SAMN10219760	BOTBYS
RCSX00000000	SAMN10219761	BOTDEW
RCSY00000000	SAMN10219762	BOTELL
RCSZ00000000	SAMN10219763	BOTGLO
RCTA00000000	SAMN10219764	BOTPOR
RCTB00000000	SAMN10219765	BOTSIN
RCTC00000000	SAMN10219766	BOTSQU
RCTD00000000	SAMN10219767	SCLCEP

The 12 RNAseq data used for gene prediction are deposited in Genbank under Bioproject number PRJNA494516, with sequence accession numbers SRR8053381- SRR8053392.

Funding

This work was supported by the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs (project 15003).

696 **References**

697 Amselem J, Cuomo CA, van Kan JAL, Viaud M, Benito EP, et al. 2011. Genomic analysis of the necrotrophic
698 fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea.* PLOS Genet 7: e1002230.
699 Ashton PM, Thanh LT, Trieu PN, Van Anh D, Trinh NM, et al. 2019. Three phylogenetic groups have driven
700 the recent population expansion of *Cryptococcus neoformans.* Nature Comm. 10: 2035.
701 Bertazzoni S, Williams AH, Jones DA, Syme RA, Tan KC, Hane JK. 2018. Accessories make the outfit:
702 accessory chromosomes and other dispensable DNA regions in plant-pathogenic fungi. Mol Plant-
703 Microbe Interact 31: 779-788
704 Campbell MA, Rokas A, Slot JC. 2012. Horizontal transfer and death of a fungal secondary metabolic gene
705 cluster. Genome Biol Evol 4: 289-293
706 Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, et al. 2008. MAKER: An easy-to-use annotation pipeline
707 designed for emerging model organism genomes. Genome Res 18: 188-196.
708 Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming
709 in large-scale phylogenetic analyses. Bioinformatics 25: 1972-1973.
710 Chakraborty M, Baldwin-Brown JG, Long AD, Emerson JJ. 2016. Contiguous and accurate de novo assembly
711 of metazoan genomes with modest long read coverage. Nucl Acids Res 44: e147.
712 Chatterjee S, Kuang Y, Splivallo R, Chatterjee P, Karlovsky P. 2016. Interactions among filamentous fungi
713 *Aspergillus niger, Fusarium verticillioides* and *Clonostachys rosea*: fungal biomass, diversity of secreted
714 metabolites and fumonisin production. BMC Microbiol 16: 83.
715 Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, et al. 2013. Nonhybrid finished microbial
716 genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563-569.
717 Contreras-Moreira B, Vinuesa P. 2013. GET_HOMOLOGUES, a versatile software package for scalable and
718 robust microbial pangenome analysis. Appl Environ Microbiol 79: 7696–7701.
719 Dalmais B, Schumacher J, Moraga J, Le Pecheur P, Tudzynski B, et al. 2011. The *Botrytis cinerea* phytotoxin
720 botcinic acid requires two polyketide synthases for production and has a redundant role in virulence with
721 botrydial. Mol Plant Pathol 12: 564-579.
722 de Vries RP, Riley R, Wiebenga A, Aguilar-Osorio G, Amillis S, et al. 2017. Comparative genomics reveals
723 high biological diversity and specific adaptations in the industrially and medically important fungal genus
724 *Aspergillus.* Genome Biol 18:28.
725 Derbyshire M, Denton-Giles M, Hegedus D, Seifbarghy S, Rollins JA, et al. 2017. The finished genome of
726 the plant-infecting fungus *Sclerotinia sclerotiorum*: re-visiting the 'two-speed-genome' hypothesis in the
727 context of broad host-range plant pathogenesis. Genome Biol Evol 9: 593-618.
728 Dong S, Raffaele S, Kamoun S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants.
729 Curr Opin Genet Dev. 35:57–65.

730 Dunn NA, Unni DR, Diesh C, Munoz-Torres M, Harris NL, Yao E, Rasche H, Holmes IH, Elsik CG, Lewis
731 SE. 2019. Apollo: Democratizing genome annotation. PLoS Comput. Biol. 15, e1006790
732 Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26: 2460-
733 2461.
734 Elad Y, Pertot I, Cotes Prado AM, Stewart A. 2016. Plant hosts of *Botrytis* spp. In: Fillinger S, Elad Y,
735 editors. *Botrytis:* the fungus, the pathogen and its management in agricultural systems. Springer
736 International Publishing pp. 413–86.
737 Emanuelsson O, Brunak S, von Heijne G, Nielsen H. 2007. Locating proteins in the cell using TargetP,
738 SignalP and related tools. Nat Protoc 2: 953-971.
739 Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons
740 dramatically improves orthogroup inference accuracy. Genome Biol 16: 157.
741 Faris JD, Zhang ZC, Lu HJ, Lu SW, Reddy L, et al. 2010. A unique wheat disease resistance-like gene
742 governs effector-triggered susceptibility to necrotrophic pathogens. Proc Natl Acad Sci USA 107: 13544-
743 13549.
744 Fournier E, Giraud T, Albertini C, Brygoo Y. 2005. Partition of the *Botrytis cinerea* complex in France using
745 multiple gene genealogies. Mycologia 97: 1251-1267.
746 Friesen TL, Stukenbrock EH, Liu ZH, Meinhardt S, Ling H, et al. 2006. Emergence of a new disease as a
747 result of interspecific virulence gene transfer. Nature Genet 38: 953-956.
748 Garfinkel AR, Lorenzini M, Zapparoli G, Chastagner GA. 2017. *Botrytis euroamericana*, a new species from
749 peony and grape in North America and Europe. Mycologia. 109: 495–507.
750 Garfinkel AR, Coats KP, Sherry DL, Chastagner GA. 2019. Genetic analysis reveals unprecedented diversity
751 of a globally-important plant pathogenic genus. Sci Rep 9: 6671.
752 Giraud T, Fortini D, Levis C, Lamarque C, Leroux P, et al. 1999. Two sibling species of the *Botrytis cinerea*
753 complex, transposa and vacuma, are found in sympatry on numerous host plants. Phytopathol 89: 967-
754 973.
755 Gluck-Thaler E, Slot JC. 2018. Specialized plant biochemistry drives gene clustering in fungi. ISME
756 Journal 12: 1694.
757 Graham-Taylor C, Kamphuis LG, Derbyshire MC (2020) A detailed *in silico* analysis of secondary
758 metabolite biosynthesis clusters in the genome of the broad host range plant pathogenic fungus *Sclerotinia*
759 *sclerotiorum.* BMC Genomics 21: 7.
760 Grant-Downton RT, Terhem RB, Kapralov M, Mehdi S, Rodriguez-Enriquez MJ, et al. 2014. A novel
761 *Botrytis* species is associated with a newly emergent foliar disease in cultivated *Hemerocallis*, PLOS One
762 9: e0089272.

763 Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M. 2015. BRAKER1: Unsupervised RNA-Seq-based
764 genome annotation with Genemark-ET and Augustus. Bioinformatics 32: 767-769.
765 Hyde KD, Nilsson RH, Alias SA, Ariyawansa HA, Blair JE et al. 2014. One stop shop: backbones trees for
766 important phytopathogenic genera. Fungal Diversity 67: 21-125.
767 Jarvis MC, Forsyth W, Duncan HJ. 1988. A survey of the pectic content of nonlignified monocot cell walls.
768 Plant Physiol 88: 309-314.
769 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in
770 performance and usability. Mol Biol Evol 30: 772-780.
771 Keller NP, Turner G, Bennett JW 2005.. Fungal secondary metabolism - from biochemistry to genomics. Nat
772 Rev Microbiol 3: 937–947.
773 Koren S, Walenz BP, Berlin K, Miller JR, Phillippy AM. 2017. Canu: Scalable and accurate long-read
774 assembly via adaptive k-mer weighting and repeat separation. Genome Res 27: 1-15.
775 Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with
776 a hidden markov model: application to complete genomes. J Mol Biol 305: 567-580.
777 Li L 2003. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Res 13: 2178-
778 2189.
779 Lind A, Wisecaver J, Lameiras C, Wiemann P, Palmer J, et al. 2017. Drivers of genetic diversity in secondary
780 metabolic gene clusters within a fungal species. PLOS Biology. 15.e2003583.
781 Liu ZH, Faris JD, Oliver RP, Tan KC, Solomon PS, et al. 2009. SnTox3 acts in effector triggered
782 susceptibility to induce disease on wheat carrying the Snn3 gene. PLOS Path 5: e1000581.
783 Liu ZH, Zhang ZC, Faris JD, Oliver RP, Syme R, et al. 2012. The Cysteine rich necrotrophic effector SnTox1
784 produced by *Stagonospora nodorum* triggers susceptibility of wheat lines harboring Snn1. PLOS Path 8:
785 e1002467.
786 Lo Presti L, Kahmann R. 2017. How filamentous plant pathogen effectors are translocated to host cells. Curr
787 Opinion Plant Biol 38: 19-24.
788 Love J, Palmer J, Stajich J, Esser T, Kastman E, Winter D. 2019, March 24.. nextgenusfs/funannotate:
789 funannotate v1.5.3 Version 1.5.3.
790 Maddison WP, Maddison DR. 2019. Mesquite: a modular system for evolutionary analysis. Version 3.61
791 http://www.mesquiteproject.org.
792 McDonald MC, Taranto AP, Hill E, Schwessinger B, Liu ZH, et al. 2019. Transposon-mediated horizontal
793 transfer of the host-specific virulence protein ToxA between three fungal wheat pathogens. mBIO 10:
794 e01515-19.
795 Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, et al. 2015. Minimum Information about a
796 Biosynthetic Gene cluster. Nat Chem Biol 2015, 11, 625-631.

797 Mercier A, Carpentier F, Duplaix C, Auger A, Pradier JM, et al. 2019.. The polyphagous plant pathogenic
798 fungus *Botrytis cinerea* encompasses host-specialized and generalist populations. Env Microbiol 21:
799 4808-4821.
800 Min B, Grigoriev IV, Choi IG 2017. FunGAP: Fungal Genome Annotation Pipeline using evidence-based
801 gene model evaluation. Bioinformatics 33: 2936–2937.
802 Moeller M, Stukenbrock EH. 2017. Evolution and genome architecture in fungal plant pathogens. Nature
803 Rev Microbiol 15: 756-771.
804 Navarro-Munoz JC, Collemare JM. 2020. Evolutionary histories of Type III Polyketide Synthases in fungi.
805 Front Microbiol 10: 3018.
806 Navarro-Muñoz JC, Selem-Mojica N, Mullowney, MW, Kautsar, S, Tryon JH, et al 2019. A computational
807 framework for systematic exploration of biosynthetic diversity from large-scale genomic data. Nat Chem
808 Biol 16: 60-68.
809 Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2014. IQ-TREE: a fast and effective stochastic
810 algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32: 268-274.
811 Nielsen K, Yohalem DS. 2001. Origin of a polyploid *Botrytis* pathogen through interspecific hybridization
812 between *Botrytis aclada* and *B. byssoidea*. Mycologia 93: 1064-1071.
813 Pedro H, Yates AD, Kersey PJ, De Silva NH. 2019. Collaborative annotation redefines gene sets for crucial
814 phytopathogens. Front Microbiol 10: 2477.
815 Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from
816 transmembrane regions. Nat Methods. 8: 785-786.
817 Pinedo C, Wang CM, Pradier JM, Dalmais B, Choquer M, et al. 2008. Sesquiterpene synthase from the
818 botrydial biosynthetic gene cluster of the phytopathogen *Botrytis cinerea*. ACS Chem Biol. 3:791-801.
819 Porquier A, Moraga J, Morgant G, Dalmais B, Simon A, et al. 2019. Botcinic acid biosynthesis in *Botrytis*
820 *cinerea* relies on a subtelomeric gene cluster surrounded by relics of transposons and is regulated by the
821 Zn(2)Cys(6) transcription factor BcBoa13. Curr Genet 65: 965-980.
822 Porquier A, Morgant G, Moraga J, Dalmais B, Luyten I, et al. 2016. The botrydial biosynthetic gene cluster
823 of *Botrytis cinerea* displays a bipartite genomic structure and is positively regulated by the putative
824 Zn(II)(2)Cys(6) transcription factor BcBot6. Fung Genet Biol 96: 33-46.
825 Reynolds HT, Vijayakumar V, Gluck-Thaler E, Korotkin HB, Matheny PB, Slot JC. 2018. Horizontal gene
826 cluster transfer increased hallucinogenic mushroom diversity. Evol Lett 2: 88-101.
827 Ropars J, de la Vega RCR, Lopez-Villavicencio M, Gouzy J, Sallet E, et al. 2015. Adaptive horizontal gene
828 transfers between multiple cheese-associated fungi. Curr Biol 25: 2562-2569.

829 Rossi FR, Gárriz A, Marina M, Romero FM, Gonzalez ME, et al. 2011. The sesquiterpene botrydial produced
830 by *Botrytis cinerea* induces the hypersensitive response on plant tissues and its action is modulated by
831 salicylic acid and jasmonic acid signaling. Mol Plant-Microbe Interact 24: 888-96.
832 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. 2003. Cytoscape: A software environment for
833 integrated models of biomolecular interaction networks. Genome Res 13: 2498–2504.
834 Shen XX, Steenwyk JL, LaBella AL, Opulente DA, Zhou X, Kominek J, Li Y, Groenewald M, Todd Hittinger
835 C, Rokas A 2020. Genome-scale phylogeny and contrasting modes of genome evolution in the fungal
836 phylum Ascomycota. BioRxiv https://doi.org/10.1101/2020.05.11.088658
837 Shi GJ, Friesen TL, Saini J, Xu SS, Rasmussen JB, Faris JD 2015. The wheat Snn7 gene confers susceptibility
838 on recognition of the *Parastagonospora nodorum* necrotrophic effector SnTox7. Plant Genome 8: DOI
839 10.3835/plantgenome2015.02.0007.
840 Shi GJ, Zhang ZC, Friesen TL, Bansal U, Cloutier S, et al. 2016. Marker development, saturation mapping,
841 and high-resolution mapping of the Septoria nodorum blotch susceptibility gene Snn3-B1 in wheat. Mol
842 Genet Genom 291: 107-119.
843 Shi-Kunne X, Faino L, van den Berg GCM, Thomma BPHJ, Seidl MF. 2018. Evolution within the fungal
844 genus *Verticillium* is characterized by chromosomal rearrangement and gene loss. Env Microbiol 20:
845 1362-1373.
846 Shimodaira H (2002. An approximately unbiased test of phylogenetic tree selection. Systematic Biol 51:
847 492-508.
848 Shirane N, Masuko M, Hayashi Y. 1989. Light microscopic observation of nuclei and mitotic chromosomes
849 of *Botrytis* species. Phytopathology 79: 728-730.
850 Siewers V, Viaud M, Jimenez-Teja D, Collado IG, Gronover CS, et al. 2005. Functional analysis of the
851 cytochrome P450 monooxygenase gene bcbot1 of *Botrytis cinerea* indicates that botrydial is a strain-
852 specific virulence factor Mol Plant-Microbe Interact 18: 602-612.
853 Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome
854 assembly and annotation completeness with single-copy orthologs. Bioinformatics 31: 3210-3212.
855 Sipos G, Prasanna AN, Walter MC, O'Connor E, Balint B, et al. 2017. Genome expansion and lineage-
856 specific genetic innovations in the forest pathogenic fungi Armillaria. Nature Ecol Evol 1: 1931-1941.
857 Slot J, Gluck-Thaler E. 2019. Metabolic gene clusters, fungal diversity, and the generation of accessory
858 functions. Curr Opinion Genet Dev 58: 17-24.
859 Soltis NE, Atwell S, Shi GJ, Fordyce R, Gwinner R, et al. 2019. Interactions of tomato and *Botrytis cinerea*
860 genetic diversity: parsing the contributions of host differentiation, domestication, and pathogen variation.
861 Plant Cell 31: 502-519.

862 Spatafora JW, Aime MC, Grigoriev IV, Martin F, Stajich JE, Blackwell M. 2017. The Fungal Tree of Life:
863 from molecular systematics to genome-scale phylogenies. Microbiol Spectrum 5.
864 Sperschneider J, Gardiner DM, Dodds PN, Tini F, Covarelli L, Singh KB, et al. 2016. EffectorP: Predicting
865 fungal effector proteins from secretomes using machine learning. New Phytol 210: 743-761.
866 Staats M, van Baarlen P, van Kan JAL. 2005. Molecular phylogeny of the plant pathogenic genus *Botrytis*
867 and the evolution of host specificity. Mol Biol Evol 22: 333–346.
868 Staats M, van Baarlen P, van Kan JAL. 2007. AFLP analysis of genetic diversity in populations of *Botrytis*
869 *elliptica* and *Botrytis tulipae* from the Netherlands. Eur J Plant Pathol 117: 219-235.
870 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
871 phylogenies. Bioinformatics 30: 1312-1313.
872 Stanke M, Schöffmann O, Morgenstern B, Waack S. 2006. Gene prediction in eukaryotes with a generalized
873 hidden Markov model that uses hints from external sources. BMC Bioinformatics 7: 62.
874 Vakirlis N, Sarilar V, Drillon G, Fleiss A, Agier N, Meyniel JP, et al. 2016. Reconstruction of ancestral
875 chromosome architecture and gene repertoire reveals principles of genome evolution in a model yeast
876 genus. Genome Res 26: 918–932.
877 Valero Jiménez C, Veloso J, Staats M, van Kan JAL. 2019. Comparative genomics of *Botrytis* spp. BMC
878 Genomics 20: 203.
879 van Kan JAL, Stassen JHM, Mosbach A, van Der Lee TAJ, Faino L, et al. 2017. A gapless genome sequence
880 of the fungus *Botrytis cinerea*. Mol Plant Pathol 18: 75–89.
881 Veloso J, van Kan JAL. 2018. Many shades of grey in *Botrytis*-host plant interactions. Trends Plant Sci 23:
882 613-622.
883 Vignatti P, Gonzalez ME, Jofré EC, Bolívar-Anillo HJ, Moraga J, et al. 2020.. Botrydial confers *Botrytis*
884 *cinerea* the ability to antagonize soil and phyllospheric bacteria. Fung Biol 124: 54-64.
885 Waalwijk C, Taga M, Zheng SL, Proctor RH, Vaughan MM, O'Donnell K. 2018. Karyotype evolution in
886 *Fusarium.* IMA Fungus 9:13-26.
887 Walker AS, Gladieux P, Decognet V, Fermaud M, Confais J, et al. 2015. Population structure and temporal
888 maintenance of the multihost fungal pathogen *Botrytis cinerea*: causes and implications for disease
889 management. Env Microbiol 17: 1261-1274.
890 Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, et al. 2015. antiSMASH 3.0—a comprehensive
891 resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43: W237-243.
892 Yohalem DS, Nielsen K, Nicolaisen M. 2003. Taxonomic and nomenclatural clarification of the onion neck
893 rotting *Botrytis* species. Mycotaxon 85: 175-182.

894 **Supplementary Tables**

895

896 **Supplementary Table S1**

897 Information about the strains used in this study.

898

899 **Supplementary Table S2**

900 Comparison of plant cell wall degrading enzymes (PCWDEs) among *Botrytis* spp. Enzymes were
901 categorized according to their substrate based on information from the CAZY database
902 (www.cazy.org)

903

904 **Supplementary Table S3**

905 The 45 Leotiomycete genomes used for orthogroup analysis of secondary metabolite biosynthetic
906 gene clusters (SM BGCs).

907

908 **Supplementary Table S4**

909 Details of balanced genomic rearrangements between inferred ancestral genomes (AB0-A6) and
910 extant *Botrytis* species, as provided in simplified form in Table 2. Inferences were made using the
911 tool Anchro.

912 **Supplementary Fig. legends**

913 **Supplementary Fig. S1.** Core genome and pan-genome analysis of 16 *Botrytis* species. A)
914 Estimation of *Botrytis* spp. core genome, in which the number of shared genes is plotted as a
915 function of the number of species sequentially added. B) Estimation of *Botrytis* spp. pan-genome
916 size, in which the number of all genes is plotted as a function of the number of species sequentially
917 added.

918

919 **Supplementary Fig. S2.** Presence/absence of orthologous groups of secreted proteins among 16
920 *Botrytis* species and *S. cepivorum* (species listed in descending alphabetical order). Dark circles
921 indicate the presence of different subgroups in each species, while white circles indicate absence.
922 The number of orthologous groups for each subgroup is shown above the bars.

923

924 **Supplementary Fig. S3.** Distribution of orthogroups of Non-Ribosomal Peptide Synthase clusters
925 in 45 Leotiomycetes. The 50 clusters that are most abundant among the 45 Leotiomycetes taxa are
926 displayed. Clusters that are exclusively represented in *Botrytis* are marked red; clusters
927 predominantly in *Botrytis* but also in some other taxa are marked as ochre; clusters predominantly
928 in other taxa but also in some *Botrytis* species are marked as blue; clusters lacking in all *Botrytis*
929 species are marked as grey.

930

931 **Supplementary Fig. S4.** Distribution of orthogroups of Terpene Synthase clusters in 45
932 Leotiomycetes. The 50 clusters that are most abundant among the 45 Leotiomycetes taxa are
933 displayed. Clusters that are exclusively represented in *Botrytis* are marked red; clusters
934 predominantly in *Botrytis* but also in some other taxa are marked as ochre; clusters predominantly
935 in other taxa but also in some *Botrytis* species are marked as blue; clusters lacking in all *Botrytis*
936 species are marked as grey.

937

938 **Supplementary Fig. S5.** The most parsimonious evolutionary trajectory from the ancestral (ABS0)
939 configuration towards extant Sclerotiniaceae. Coloured boxes represent syntenic blocks. A1-A6
940 represent intermediate ancestors. Numbers above the branches represent the total number of
941 balanced rearrangements (interchromosomal translocations and fusions/fissions;
942 intrachromosomal inversions) accumulated between two genomes.

943 **Supplementary Data**

944 **Supplementary Data S1**

945 Analysis of orthologous protein groups in 16 *Botrytis* species and *Sclerotium cepivorum* using
946 Orthofinder. From 14,838 proteins, 14,326 proteins were assigned to 1,116 orthologous groups.

947

948 **Supplementary Data S2**

949 Analysis of orthologous effector protein groups in 16 *Botrytis* species and *Sclerotium cepivorum*
950 using Orthofinder. Each of the 17 species contained a total of 121-152 candidate effector genes,
951 which were assigned to 244 orthologous groups.

952

953 **Supplementary Data S3**

954 BGCs were identified by AntiSMASH in 16 *Botrytis* species, 4 other Sclerotiniaceae and 25 other
955 Leotiomycete taxa. Categories listed according to AntiSMASH.

956

957 **Supplementary Data S4**

958 Grouping of secondary metabolite biosynthetic gene clusters by BiG-SCAPE. The 1571 BGCs
959 were grouped over 438 families.

960

961 **Supplementary Data S5**

962 Manual assessment of functionality of genes in the botrydial biosynthetic gene cluster. Genes were
963 analysed for eight *Botrytis* species and some other fungi containing BOT gene homologs. Genes
964 highlighted in yellow are annotated as pseudogenes, for reasons provided in the comment box.

965

966 **Supplementary Data S6**

967 Manual assessment of functionality of genes in the botcinic acid biosynthetic gene cluster. Genes
968 were analysed for 16 *Botrytis* species and some other fungi containing BOA homologs. Genes
969 highlighted in yellow are annotated as pseudogenes, for reasons provided in the comment box.

970

971 **Supplementary Data S7**

972 Results of Maximum Likelihood phylogenetic analyses of the botrydial biosynthetic gene cluster,
973 with and without pseudogenes included. Pseudogenes are excluded for topological inference and
974 constraint analyses because pseudogenes do not conform to models of protein evolution, which
975 introduces unpredictable error into tree inference. Gene tree topologies are presented in Newick

format, which can be visualized in tree-viewing software such as FigTree (https://github.com/rambaut/figtree/) by copying and pasting into a tree window.

Supplementary Data S8

Results of Maximum Likelihood phylogenetic analyses of the botcinic acid biosynthetic gene cluster, with and without pseudogenes included. Pseudogenes are excluded for topological inference and constraint analyses because pseudogenes do not conform to models of protein evolution, which introduces unpredictable error into tree inference. Gene tree topologies are presented in Newick format, which can be visualized in tree-viewing software as FigTree (https://github.com/rambaut/figtree/) by copying and pasting into a tree window.

Supplementary Data S9

BOA topological constraint analyses. Results of gene tree topology tests conducted in IQtree are presented. Topologies with significantly smaller log likelihoods in the Approximately Unbiased test are indicated by red ($p< 0.01$) or yellow ($p<0.05$), indicating the confidence in the rejection of the null hypothesis of vertical transmission. Constraints and optimal topologies are presented in Newick format, which can be visualized in tree-viewing software such as FigTree (https://github.com/rambaut/figtree/) by copying and pasting into a tree window.

Supplementary Data S10

Shared synteny in the BOA locus. Genes in and flanking the locus with BOA genes were assigned to homologous gene groups, with arbitrary numbers, based on cluster_agg method in Usearch. Genes are shaded as BOA genes (grey), no shared synteny (yellow), or shared synteny (alternate colors). Presumed functional genes are indicated with black text and pseudogenes are indicated with red text.

Supplementary Data S11

Evolutionary hypotheses for BOT and BOA clusters. Hypotheses of origins and diversification of BOT and BOA gene clusters are described as a synthesis of gene phylogenies in relation to the species phylogeny, and evidence of common locus due to shared synteny. The ancestral state reconstruction nexus file is appended.