Botrytis cinerea strains infecting grapevine and tomato display contrasted repertoires of accessory chromosomes, transposons and small RNAs

Adeline Simon[1], Alex Mercier[1], Pierre Gladieux [2], Benoit Poinssot [3], Anne-Sophie Walker[1] and Muriel Viaud[1]*

[1] Université Paris-Saclay, INRAE, UR BIOGER, 78850 Thiverval-Grignon, France

[2] PHIM Plant Health Institute, Univ Montpellier, INRAE, CIRAD, Institut Agro, 34398 Montpellier, France

[3]Agroécologie, CNRS, INRAE, Institut Agro Dijon, Univ. Bourgogne, Université Bourgogne Franche-Comté, 21000 Dijon, France

[*] Corresponding author: muriel.viaud@inrae.fr

Abstract

The fungus *Botrytis cinerea* is a polyphagous pathogen that encompasses multiple host-specialized lineages. While several secreted proteins, secondary metabolites and retrotransposons-derived small RNAs have been characterized as virulence factors, their role in host specialization remain unknown. The aim of this study was to identify the genomic correlates of host-specialization in populations of *B. cinerea* associated with grapevine and tomato. Using PacBio sequencing, we produced complete assemblies of the genomes of strains Sl3 and Vv3 that represent the French populations T and G1 of *B. cinerea*, specialized on tomato and grapevine, respectively. Both assemblies revealed 16 core chromosomes that were highly syntenic with chromosomes of the reference strain B05.10. The main sources of variation in gene content were the subtelomeric regions and the accessory chromosomes, especially the chromosome BCIN19 of Vv3 that was absent in Sl3 and B05.10. The repertoires and density of transposable elements were clearly different between the genomes of Sl3 and Vv3 with a larger number of subfamilies (26) and a greater genome coverage in Vv3 (7.7%) than in Sl3 (14 subfamilies, 4.5% coverage). An Helitron-like element was found in almost all subtelomeric regions of the Vv3 genome, in particular in the flanking regions of a highly duplicated gene encoding a Telomere-Linked Helicase, while both features were absent from the Sl3 and B05.10 genomes. Different retrotransposons in the Sl3 and the Vv3 strains resulted in the synthesis of distinct sets of small RNAs. Finally, extending the study to additional strains indicated that the accessory chromosome BCIN19 and the small RNAs producing retrotransposons Copia_4 and Gypsy_7 are common features of the G1 population that are scarcely if ever found in strains isolated from other populations. This research reveals that accessory chromosomes, repertoires of transposons and their derived small RNAs differ between populations of *B. cinerea* specialized on different hosts. The genomic data characterized in our study pave the way for further studies aiming at investigating the molecular mechanisms underpinning host specialization in a polyphagous pathogen.

Keywords: Fungi, PacBio genome sequence, secondary metabolism, helicase, Helitron, host specialization*, Vitis vinifera, Solanum lycopersicum*

28 **Introduction**

29

30 While specialist phytopathogenic fungi are highly specific to a single host plant, some other fungi stand
31 out for having a broad host range and are qualified as generalist. Nevertheless, these so-called
32 generalist pathogens could actually correspond to multiple co-existing populations that show a certain
33 level of host specialization (Gladieux et al., 2018; Stukenbrock & McDonald, 2008). Such scenario
34 provides an excellent opportunity to investigate the molecular determinants of host specialization. In
35 this regard, the grey mould fungus *Botrytis cinerea* has recently become a model of choice as both
36 powerful functional genomic tools and population data are available. This ascomycete species is a
37 necrotrophic pathogen that infects more than 1400 host plant species belonging to 580 genera causing
38 significant damages in grapevine and in most cultivated fruits (Elad et al., 2015). Multiple studies
39 showed that *B. cinerea* produces an arsenal of Plant Cell Wall Degrading Enzymes (PCWDEs) and
40 enzymes conferring the ability to cope with oxidative stress (Choquer et al., 2007). In addition, some
41 small molecules were identified as effectors able to kill or manipulate host cells: small secreted
42 proteins, secondary metabolites (*i.e*. botrydial and botcinic acid) and finally the recently characterized
43 small interfering RNAs (siRNAs) (Weiberg et al., 2013; Mbengue et al., 2016; Veloso & van Kan, 2018).
44 These fungal siRNAs highjack the silencing machinery of the host plant to silence genes involved in the
45 defence process. All these fungal weapons have been mainly investigated in the model strain B05.10
46 whose genome is fully sequenced and assembled thanks to the PacBio technology (van Kan et al.,
47 2017). Nevertheless, phenotypic variability in *B. cinerea* strains as well as the identification of strain-
48 specific virulence factors suggested that the infection strategy may differ from one strain to another
49 (Choquer et al., 2007). In addition, several studies conducted in different parts of the world revealed
50 that populations of *B. cinerea* are structured, and the host plant was recognized as the factor with the
51 highest explanatory power for this structure, ahead of geography (reviewed in Walker, 2015). Notably
52 French populations of *B. cinerea* isolated from tomato and grapevine are differentiated from each

53 other (Walker et al., 2015), and have higher aggressiveness on their host-of-origin than other strains,
54 indicating host specialization (Mercier et al., 2019). Recently, Illumina sequencing of 32 representative
55 isolates confirmed the subdivision of these *B. cinerea* French populations into two genetic clusters on
56 grapevine (G1 and G2 populations) and another, single cluster on tomato that diverged from the G2
57 population (T population; Sup. Fig. S1; Mercier et al., 2021). These genomic data also allowed to
58 investigate the molecular differences underlying host specialization. By characterizing single-
59 nucleotide polymorphisms in Illumina short-read data, genes with footprints of positive selection
60 and/or divergent selection in the genomes of populations specialized to different hosts were
61 identified. These candidate genes that represent possible determinants of host specialization were
62 enriched in genes encoding PCWDEs and transporters and in genes involved in the oxidative stress
63 response. The same study also highlighted a limited number of candidate genes that are population-
64 specific including a couple of PCWDE-coding genes and one secondary metabolism gene cluster
65 (Mercier et al., 2021). Though this previous study provided significant information about the evolution
66 of *B. cinerea* genes in populations and identified candidate genes possibly involved in host
67 specialization, the lack of a complete assembly of the sequenced genomes did not allow to investigate
68 other mechanisms potentially involved in specialization, such as (i) variation in the presence of
69 accessory chromosomes, (ii) chromosomal rearrangements, and (iii) differences in repertoires of
70 transposable elements (TEs). This last point is of particular interest since retrotransposons were shown
71 to be the source of synthesis of siRNA acting as effectors in *B. cinerea* (Weiberg et al., 2013), and could
72 therefore be considered as pathogenicity factor (Porquier et al., 2021). High variability in the repertoire
73 of TEs across fungal populations and the fact that retrotransposons are the main source of siRNAs raise
74 the possibility that these TEs could be involved in host specialization.

75 In this study, our objective was to investigate whether strains from French populations of *B. cinerea*
76 specialized on tomato *versus* grapevine differ in genomic content in terms of core and accessory
77 chromosomes, TEs and small RNA repertoires. We characterized genome sequences produced with
78 the PacBio technology for two representative strains: the Sl3 strain that belongs to the T population

79 specialized on tomato and the Vv3 strain that belongs to the G1 population specialized on grapevine.
80 The full assembly of Sl3 and Vv3 genomes provided a complete view of the core and accessory
81 chromosomes as well as the full repertoires of genes and TEs. Small RNAs produced by the repertoires
82 of retrotransposons were also compared. Finally, extending the study to additional strains allowed to
83 identify which of the genomic features are associated to the populations specialized on tomato or
84 grapevine.

85

86 **Results**

87

88 **The genome architecture of the *Botrytis cinerea* strains Sl3 and Vv3 differ from the B05.10 strain in**
89 **their accessory chromosomes**

90 Genome assemblies based on PacBio Sequel sequencing data identified 16 core chromosomes (CCs)
91 and two accessory chromosomes (ACs) in both *Botrytis cinerea* strains Sl3 and Vv3 collected on tomato
92 and grapevine. Sequencing reads were assembled *de novo* using a combination of HGAP4 (SMRT-LINK
93 v5.0.1) (https://github.com/PacificBiosciences/) and CANU v1.6 (Koren et al., 2017). Both Sl3 and Vv3
94 assemblies consisted of 18 main contigs. The genome assemblies were estimated to be 43.2 Mb (Sl3)
95 and 44.9 Mb (Vv3), which is close to the 42.6Mb of the B05.10 genome (van Kan et al., 2017). Contigs
96 of the new genome assemblies were ordered and oriented as the previously described chromosomes
97 of B05.10 (BCIN01 to BCIN18) (Fig. 1), and then compared to the genome of B05.10 with QUAST
98 (Gurevich et al., 2013) (Sup. Table S1). The length (1.9-4.2 Mb) and GC content (41-43%) of the 16
99 largest contigs of Sl3 and Vv3 were similar to those of the 16 CCs defined in B05.10. In addition, the
100 high percentage of B05.10 genome coverage (94-99%) confirmed that the 16 largest newly sequenced
101 contigs corresponded to CCs BCIN01 to BCIN16 of *B. cinerea*. The length of CCs was homogenous across
102 the three strains, with less than 400 kb variation per CC (BCIN03: 3.6 Mb in Vv3, 3.3 Mb in Sl3 and 3.2
103 Mb in B05.10). The length (0.2-0.5 Mb) and GC content (<40%) of the two smallest contigs of Sl3 and
104 Vv3 was typical of ACs, as those found in B05.10. The percentage of B05.10 genome coverage indicated
105 that both strains had the AC BCIN17 (coverage >96%) while only Sl3 had an AC related to BCIN18 (52%
106 coverage). The second putative AC of Vv3 did not share any similarity with the genome of B05.10 and
107 corresponded to a new AC that we numbered BCIN19.

108 A criterion to evaluate the quality of a genome assembly is the presence of telomeres at the terminal
109 regions of the putative full chromosomes. For Sl3, telomeric repeats (*i.e*. TTAGGG repetitions; Levis et
110 al., 1997a) were found at both ends of eight chromosomes and at one end of eight other chromosomes

111 (Sup. Table S1). Only the assemblies of CCs BCIN1 and BCIN12 lacked both telomeric repeats. Thus,
112 telomeric repeats were missing at twelve chromosomal ends in the Sl3 assembly. In comparison,
113 telomeric repeats were missing at nine chromosomal ends in the B05.10 assembly (van Kan et al.,
114 2017), though not necessarily at the same chromosomal ends as in Sl3. Telomeric repeats in Vv3
115 contigs were retrieved only at the 5' end of CC BCIN2. To investigate the relative scarcity of telomeric
116 repeats in the Vv3 assembly, we searched the raw reads for telomeric repeats. We found 317 and 3712
117 occurrences of the telomeric repeat $(TTAGGG)_3$ and its reverse complement in Vv3 and Sl3 reads,
118 respectively. The relatively low abundance of telomeric repeats in sequencing reads of Vv3 indicated
119 that telomeric sequences were relatively rare in the library. This could be due to degradation of these
120 sequences during the preparation of libraries, as previously described (Tolios et al., 2015).

121 As the assemblies of the Vv3 and Sl3 genomes revealed different pairs of ACs, we further investigated
122 whether there was a correlation between the content in ACs and the host of origin. We used the 35
123 available genomic sequences of *B. cinerea* strains from three distinct populations specialized on
124 tomato (T population) or grape (G1 and G2 populations) as well as those of additional strains isolated
125 from bramble (Rf1, Rf2), hydrangea (Hm1), tomato (T4) and grapevine (BcDW1) (Mercier et al., 2021 ;
126 Amselem et al., 2011; Blanco-Ulate et al., 2013). Percentage of chromosome coverage showed
127 different distributions of the three ACs BCIN17, BCIN18 and BCIN19 (AC17 to 19; Table 1). The AC
128 BCIN17 first identified in B05.10 was detected in 30 other strains, including strains belonging to
129 populations T, G1 and G2 as well as in Rf1, Hm1 and BcDW1. The AC BCIN18 also identified in B05.10
130 was detected in seven other strains, including strains belonging to the T and G2 populations, Hm1 and
131 BcDW1. The AC BCIN19 identified in the Vv3 strain, was specific to the G1 group (detected in 11 strains
132 out of 12). Among populations T, G1 and G2, only G1 showed a relatively homogenous content in ACs
133 with the systematic presence of BCIN17, very high frequency of BCIN19 and absence of BCIN18.

134

135

136 **The core chromosomes of B05.10, Sl3 and Vv3 strains show a high level of synteny**

137 A vast majority of the genes identified in the reference strain B05.10 were retrieved in the Sl3 and Vv3
138 genomes. B05.10 gene annotation (van Kan et al., 2017) was transferred to Sl3 and Vv3 with the LIFTOFF
139 annotation mapping tool (Shumate & Salzberg, 2021). 11 661 and 11 609 genes were transferred to
140 Sl3 and Vv3 respectively, *i.e.* 99.6% and 99.2% of the 11 707 genes previously predicted in B05.10.
141 Altogether, 11 590 genes were retrieved in all three genomes, the vast majority in CCs (11 572 genes)
142 and very few in ACs (BCIN17: 18 genes). To assess synteny, we used SYNCHRO (Drillon et al., 2014), a
143 tool that reconstructs synteny blocks between pairwise comparisons of multiple genomes. As shown
144 in Sup. Fig. S2, the three sets of CCs were largely syntenic. When considering major synteny breaks, *i.e.*
145 those involving at least four adjacent genes that were not organized as in the B05.10 reference
146 genome, only one and two possible inversion events could be identified in Sl3 and Vv3, respectively
147 (Sup. Fig. S3). Two rearrangements between the terminal regions of chromosomes were also observed,
148 both accompanied by gene duplication. The first one involved a secondary metabolism gene cluster
149 with the polyketide synthase BcPKS7 as a key enzyme (Bcin10g00010-40). This cluster and the two 3'
150 flanking genes were located in two different regions in B05.10 and Vv3, *i.e.* the 5' end of BCIN10 in
151 B05.10 *versus* the 3' end of BCIN03 in Vv3. (Sup. Fig. S4). The rearrangement also resulted in the
152 duplication of a gene encoding a putative mitochondrial NADPH cytochrome b5 reductase in Vv3. The
153 second major subtelomeric rearrangement was observed when comparing the genomes of B05.10 and
154 Sl3. The nine first genes of BCIN08 in B05.10 (Bcin08g00010-90) were localized at the start of
155 chromosomes 2 and 15 in Sl3, while BCIN08 in Sl3 started with Bcin08g00080 (Sup. Fig. S5). Hence, the
156 genome of Sl3 contained two copies of seven of these genes and three copies of two of them. The
157 duplicated genes were predicted to encode a P450 monooxygenase, a laccase (*Bclcc11*), a glycoside
158 hydrolase possibly acting on hemicellulose or pectin chains (GH43 family; Lombard et al., 2014), a
159 glycosyl transferase, a heterokaryon incompatibility protein, a pyridine transferase as well as two
160 secreted proteins with unknown functions.

161 **Subtelomeric regions of core chromosomes and accessory chromosomes provide variation in gene**
162 **content**

163 The CCs of *B. cinerea* showed Presence Absence Variation (PAV) of genes in particular for secondary
164 metabolism gene clusters. To identify genes present in the Sl3 or Vv3 genomes but not in the B05.10
165 genome, we used the FGENESH *ab initio* gene-finder (Solovyev et al., 2006) and excluded the gene
166 models mapping in the B05.10 genome (Sup. Table S2). As shown in Fig. 1, *de novo* genes annotation
167 identified new genes in the CCs especially in their subtelomeric regions. When considering groups of
168 contiguous genes that were not shared between the three genomes, three secondary metabolism
169 gene clusters were identified based on the presence of genes encoding key enzymes. The first one was
170 only detected in Sl3 and included the gene encoding the previously identified SesquiTerpene Cyclase
171 BcSTC6 (BCIN_Sl3_03_394) (Amselem et al., 2011). The second gene cluster was identified in a
172 subtelomeric region of Vv3 and B05.10 (Bcin02g00013 to Bcin02g00016) and included a gene encoding
173 a non-ribosomal peptide synthetase-like similar to MelA , an enzyme involved in the biosynthesis of an
174 α-keto acid dimer in *Aspergillus terreus* (Geib et al., 2016). Finally, the third gene cluster was in a
175 subtelomeric region shared by Sl3 and Vv3 (BCIN_Sl3_10_11 to 26; BCIN_Vv3_10_2 to 14) and included
176 a gene predicted to encode a DiTerpene Synthase (DTC). When looking into previous Illumina
177 sequencing data (Mercier et al., 2021), these three secondary metabolism genes clusters had different
178 distributions: while the *Bcstc6* and the newly identified *Bcdtc* clusters were detected in strains of the
179 three different populations, the *BcmelA* cluster was present in all strains of the G1 and G2 groups but
180 absent from the T group (Table 1).

181 Fourteen subtelomeric genes coding for a protein predicted to be a helicase were identified in the
182 genome of Vv3 strain but not in the ones of B05.10 and Sl3 strains. This protein showed 51% similarity
183 with the Telomere-Linked Helicases (TLHs) identified in *Magnaporthe oryzae* (Gao et al., 2002;
184 Rehmeyer et al., 2009) and in *Ustilago maydis* (Sánchez-Alonso & Guzmán, 1998). It also shared the
185 same predicted domains *i.e.* two C2H2 zinc finger motifs, a helicase ATP-binding domain, a helicase C

186 domain and a specific TLH domain (Rehmeyer et al., 2009) ; Sup. Fig. S6). As in *M. oryzae* and *U. maydis*,
187 the 14 Vv3 helicase-encoding genes were localized in subtelomeric regions (11 regions). They were
188 located between 9 bp and 63 Kbp from chromosome ends, and six of them were actually the first genes
189 detected at the end of the chromosome. Incomplete copies of these helicase-encoding genes were
190 also detected in eight other subtelomeric regions (Sup. Fig. S7). Because of their homology with TLHs
191 of other fungi and of their genomic localization, these newly identified helicases were called BcTLHs.
192 PCR screening of *B. cinerea* populations further indicated that one or several copies of *Bctlh* are present
193 in a majority of strains specialized on grapevine, but also in few strains specialized on tomato or
194 isolated from other hosts (Table 1).

195 Accessory chromosomes were an additional source of PAV of genes among the B05.10, Sl3 and Vv3
196 strains (Fig. 1; Sup. Table S1). Firstly, the AC BCIN18 that was present in both Sl3 and B05.10 strains
197 was found to be only partially conserved as it carried seven genes that were shared between the two
198 strains, and nine contiguous genes that were present only in B05.10. Secondly, the newly identified AC
199 BCIN19 of the strain Vv3 carried 78 genes that were not present in the Sl3 and B05.10 strains. These
200 78 genes mostly encoded proteins with unknown functions, which is reminiscent of the gene content
201 of ACs BCIN17 and BCIN18 (van Kan et al., 2017). Nevertheless, a high proportion of the predicted
202 proteins displayed IPR domains that could be related to nucleic acids binding and modifications (eight
203 ribonucleases, two helicases, two zinc finger C2H2 proteins), transport through vesicles (five dynamins)
204 and peptidase activity (five peptidases) (Sup. Table S2).

205

206 **B05.10, Sl3 and Vv3 strains show different repertoires and densities of transposable elements**

207 Different subfamilies of Transposable Elements (TEs) were identified in the B05.10, Sl3 and Vv3 strains.
208 The complete annotation of TEs in the Sl3 and Vv3 genomes was carried out using a *de novo* approach
209 with the REPET package (Flutre et al., 2011; Amselem et al., 2015). Consensus sequences representing
210 all possible TEs among the Sl3 and Vv3 genomes were identified. After manual curation, 19 and 33

211 consensus TE sequences were retained for Sl3 and Vv3, respectively (Sup. File S1). They were classified
212 based on their structure and sequence similarities to characterized eukaryotic transposons (Hoede et
213 al., 2014; Wicker et al., 2007). Fifteen consensus were previously identified in the B05.10 genome using
214 the same pipeline and parameters as in the present study (Porquier et al., 2016). Comparison of the
215 B05.10, Sl3 and Vv3 consensus TE sequences revealed a total of 33 different subfamilies: 13 class I LTR
216 (retrotransposons), seven class II TIR (DNA transposons), one class II Helitron and 12 subfamilies with
217 either low genome coverage or vague classification (Fig. 2; Sup. Table S3). The subfamily *Boty*/Gypsy_1
218 element (Diolez et al., 1995) was built from several *Boty* consensus as they were identical in their
219 central part and only differed at terminal regions (Sup. Fig. S8), as previously observed in B05.10
220 (Porquier et al., 2016, 2021). In total, eight subfamilies were shared between the three genomes, *i.e.*
221 two subfamilies of Copia retrotransposons, four subfamilies of Gypsy retrotransposons including the
222 *Boty*_Gypsy_1 element and two subfamilies of the Tc1-Mariner DNA transposons, including the *Flipper*
223 element (Levis et al., 1997b). Fifteen, four, and three subfamilies were exclusive to the Vv3, Sl3 or
224 B05.10 genomes, respectively, and three subfamilies were shared between two genomes. Libraries of
225 subfamily sequences were used to annotate TE copies in the whole genomes. For each subfamily, both
226 full-length and truncated copies were retrieved in the genomes.

227 Genome coverage by subfamilies of TEs differed among the strains. Percentages of coverage for each
228 subfamily in each genome were computed (Fig. 2; Sup. Table S3). Total coverage by TEs was higher for
229 the Vv3 genome (7.7%), in which the largest number of subfamilies was identified, than in Sl3 and
230 B05.10 (4.5% and 3.7%, respectively) (Porquier et al., 2016). Transposable Elements were detected at
231 all chromosome ends but the 3' end of BCIN08 in Sl3 (Fig. 1), and also inside chromosomes, with a
232 significantly higher coverage in ACs (27.9-35.4%) than in CCs (1.9-11.3%). The eight shared subfamilies
233 were responsible for about 3% of the coverage in each genome whereas other subfamilies covered
234 approximately 1% of B05.10 or Sl3 genomes and 3.6% of Vv3 genome. The specific subfamilies
235 responsible for the higher coverage in Vv3 were mainly Copia and Gypsy retrotransposons and a TE
236 that was similar to a Helitron and covered 1.37% of the Vv3 genome.

237 The subtelomeric regions of the chromosomes of the Vv3 strain were invaded by a Helitron-like TE.
238 Helitron is a unique class of DNA transposons in eukaryotes and the representative elements of this
239 class encode a protein with two enzymatic domains corresponding to the helicase and the nuclease
240 activities that are required for transposition (Kojima, 2019). As only the helicase domain was clearly
241 identified in the TE found in Vv3, it was called a Helitron-like element (Sup. Fig. S9). Helitrons are also
242 known for their capacity to capture genes, which increases their size (Castanera et al., 2014). In
243 addition to a helicase-coding gene, the 15.5 kb sequence of the Helitron-like element found in Vv3
244 revealed two captured genes as well as relics of *Boty* (*i.e.* Long Terminal Repeats). For one of the
245 captured genes, the predicted protein was a secreted pectate-lyase (IPR002022), similar to enzymes
246 involved in the maceration and soft rotting of plant tissue. Notably, 33 complete or uncomplete copies
247 of the Helitron-like TE carried this putative virulence gene in the genome of Vv3. Unlike most TEs that
248 were randomly distributed along chromosomes, the Helitron-like TEs were mainly localized at the
249 chromosome ends (Sup. Fig. S10). Looking closer at their precise localization, we noticed that they
250 were found in 3' and/or 5' of all the 14 complete copies of the helicase encoding gene *Bctlh* (Sup. Fig.
251 S7). Finally, to investigate how this Helitron-like TE was distributed among the populations of *B.*
252 *cinerea*, we designed PCR specific primers. The results indicated that all G1 strains specialized on
253 grapevine carried one or several copies of the Helitron-like TE while the element was scarcely detected
254 in the T population and detected in two strains out of four in the G2 population (Table 1).

255 Fungi have developed defense mechanisms against invasion by TEs, including a Repeat-Induced Point
256 mutation (RIP) machinery that inactivates repeated sequences by causing Cytosine to Thymine
257 mutations and therefore decreases GC content (Amselem et al., 2015). The occurrence of RIP in *B.*
258 *cinerea* was previously revealed by the analysis of TEs in the strain B05.10 (Porquier et al., 2016, 2019,
259 2021). We therefore investigated the possible inactivation by RIP of TE copies from Sl3 and Vv3 by
260 computing the GC content of complete copies (Fig. 3; Sup. Fig. S11). The majority of the subfamilies
261 displayed copies with a large range of GC contents. For example, the Flipper DNA transposon showed
262 both copies with a GC content similar to those previously observed in a mobile copy (39%; Levis et al.,

263 1997b) and copies with a lower GC content (<20%) suggesting inactivation by RIP. For each subfamily
264 of class I and II TEs identified in B05.10, copies with different GC content were observed, also
265 suggesting a history of RIP. The same pattern was observed for Sl3 except for subfamily Copia_2 for
266 which all four copies had a GC content > 40%. In contrast, the Vv3 genome included seven subfamilies
267 of TEs, *i.e.* the Copia_2 and 4, Gypsy_5, 6 and 7, Mariner_4 and Helitron-like subfamilies, that contain
268 only copies with GC content >40% indicating the absence of RIP. Apart from Copia_2, these subfamilies
269 are absent in the Sl3 and B05.10 genomes.

270

271 **Retrotransposons of the Sl3 and Vv3 strains produce different sets of small RNAs**

272 As retrotransposons and especially unripped copies are known to be the origin of the synthesis of small
273 RNAs in *B. cinerea* (Weiberg et al., 2013; Porquier et al., 2021), we investigated whether the different
274 repertoires of TEs in the Sl3 and Vv3 strains could lead to the production of different sets of small
275 RNAs. The two strains were grown both on grape and tomato juice solid media to partially mimic the
276 conditions that the fungus encounters on the host plants (Simon et al., 2013). The corresponding small
277 RNA libraries were Illumina-sequenced. After quality filtering, only the reads between 20 and 24
278 nucleotides were kept. A Principal Component Analysis (PCA) of the repertoires of small RNAs in the
279 four samples indicated that the two strains harbour different repertoires, and that the use of grape
280 *versus* tomato juice culture medium had little impact on those repertoires (Sup. Fig. S12, A).

281 In order to identify which TEs were involved in the production of small RNAs, the reads from the Sl3
282 and Vv3 strains were mapped against the consensus sequences of the 33 identified TE subfamilies. The
283 results were similar with the two culture media and indicated that the TE-derived small RNAs were
284 produced by retrotransposons corresponding to seven subfamilies (Table 2, grey lines):

285 - Three out of six retrotransposons shared between Sl3 and Vv3 (Boty/Gypsy_1, Gypsy_3 and Gypsy_4)
286 produced small RNAs in both strains.

287 - The unique retrotransposon found in Sl3 but not in Vv3 (Copia_6) produced small RNAs.

288 - Three out of six Vv3-specific retrotransposons found in Vv3 but not in Sl3 (Copia_4, Gypsy_6 and
289 Gypsy_7) produced small RNAs.

290 In addition to these qualitative differences, quantitative differences were also observed with Gypsy_3
291 producing the higher amount of small RNAs in the Sl3 strain (291.074 and 283.036 Reads Per Millions
292 (RPMs) in the grape and tomato juice media, respectively (Sup. Table S4) and Copia_4 producing the
293 higher amount of small RNAs in Vv3 strain (115.863 and 168.212 RPMs in the grape and tomato juice
294 media, respectively).

295 Sequences of the seven small RNA-producing TEs all showed the expected structure for Copia and
296 Gypsy classes of transposons (Wicker et al., 2007), except Gypsy_4 that lacked the GAG domain and
297 Copia_4 that had a small additional one (Sup. Fig S13). Small RNAs mapped almost all along the
298 sequences suggesting that the whole elements could be converted into small RNAs (except from the
299 GAG domain of Gypsy_3). All seven small RNA-producing TEs were in numerous complete copies (six
300 to 80; Sup. Fig. S11) all over the chromosomes of Vv3 or Sl3 strains. The three Vv3 specific small RNA-
301 producing TE, *i.e.* Copia_4, Gypsy_6 and 7, only showed copies with relatively high GC content (44.4%
302 +/-0.0, 47.4% +/-0.1 and 43.6% +/- 0.0 respectively). For the other small RNA-producing TEs, *i.e*
303 Gypsy_3, Gypsy_4, Boty/Gypsy_1 and Copia_6, copies with varied levels of GC were identified.
304 Nevertheless, mapping of the small RNA reads on the genomes showed that only copies with relatively
305 high GC content could generate small RNAs (Data not shown).

306 Among *B. cinerea* small RNAs, the three siRNAs previously identified as effectors (*i.e.* siR5, siR3.1 and
307 siR3.2) were 21 bp in size and displayed the nucleotide U in first position (Weiberg et al., 2013). We
308 investigated which proportion of the small RNAs produced by the Vv3 and Sl3 strains shared the same
309 features (Sup. Fig. S14). For all seven TEs, the generated small RNAs were mainly composed of 22
310 nucleotide sequences (54-69%), followed by 21 nucleotide sequences (18-36%). Notably, a vast
311 majority of them (97%) showed an uracil (U) in first position. Finally, the three siRNA previously

312 identified as effectors in the strain B05.10 (*i.e.* siR5, siR3.1 and siR3.2) were identified among the four
313 libraries. Mapping data indicated that siR5 was produced by Boty/Gypsy_1 while siR3.1 and siR3.2
314 were produced by Gypsy_3, as previously demonstrated in the B05.10 strain (Weiberg et al., 2013;
315 Porquier et al., 2021).

316

317 **A set of retrotransposons-derived small RNAs is associated with the G1 population specialized on**
318 **grapevine**

319 As the repertoires of TE-derived small RNAs produced by the strains Vv3 and Sl3 were partially
320 different, we further investigated whether there was a correlation between the previously identified
321 populations of *B. cinerea* specialized on grapevine or on tomato (Mercier et al., 2021) and the
322 production of some of these small RNAs. The study was extended to a total of six strains from the T
323 population, four strains from the G1 population and three strains from the G2 population, as well as
324 to reference B05.10. All strains were grown on grape juice medium and small RNAs were investigated
325 as described above. Principal component analysis of small RNA repertoires (Sup. Fig. S12, B)
326 differentiated samples from the G1 population from samples from the G2 and T populations. To
327 compare the repertoires of TE-derived small RNAs from the 14 strains, we mapped the reads on the
328 TEs identified in B05.10, Sl3 and /or Vv3 strains (Table 2; Sup. Table S4). Mapping data indicated that
329 the same seven retrotransposons that produced small RNAs in B05.10, Sl3 or Vv3 strains could be
330 responsible for the production of small RNAs in the other tested strains. An exception was observed
331 for the Vv9 strain as its small RNAs did not map to any of the known TEs. It should be noted here that
332 for Vv9 and the other ten strains where only unassembled genome sequences are available (Illumina
333 data), additional unknown retrotransposons may be involved in the production of small RNAs. From
334 the presented analysis, the following distribution of retrotransposons derived-small RNAs in the
335 different populations could be observed:

336 - Small RNAs produced by the shared Boty/Gypsy_1 and Gypsy_4 TEs were identified in most of the
337 strains from the three populations and in the strain B05.10.

338 - Small RNAs produced by the shared Gypsy_3 TE were identified in most of the strains of the T and G1
339 populations and in the strain B05.10, but not in three strains of the G2 population.

340 - Small RNAs produced by the Sl3-identified Copia_6 TE were only identified in Sl3 and another strain
341 of the T population (Sl6).

342 - Small RNAs produced by the Vv3-identified Copia_4, Gypsy_6 and Gypsy_7 TEs were exclusively
343 identified in G1 strains. This could explain why the G1 strains separated from the other strains in the
344 PCA analysis (Sup. Fig. S12, B). In the genome of Vv3, Copia_4, Gypsy_6 and Gypsy_7 were detected in
345 10, 12 and 23 complete copies, respectively (Sup. Fig. S11). Copia_4 was localized on several CCs, while
346 Gypsy_6 and Gypsy-7 were found both on CCs and ACs (Sup. Fig. S15). The G1-specific AC BCIN19
347 revealed seven complete copies of Gypsy_7.

348 We further focused on the potential G1-specific small RNAs by investigating the presence of the three
349 corresponding retrotransposons in strains from the T, G1 and G2 populations and strains from other
350 hosts. Copia_4, Gypsy_6 and Gypsy_7 were searched by PCR using one pair of specific PCR primers per
351 TE which were expected to amplify at least the complete and unripped copies of TEs (Table 1). The
352 results indicated that the only small RNA generating-TE that was clearly detected in all strains of the
353 G1 population but not in any strain of G2 and T populations was Copia_4. To a lesser extent, Gypsy_7
354 was also found to be enriched in the G1 population.

355

356 **Discussion**

357

358 While the genomes of several strains of the polyphagous pathogen *B. cinerea* have already been
359 sequenced allowing studies of genetic variation (Atwell et al., 2015; Mercier et al., 2021), only one
360 gapless genome, whose of the model strain B05.10, was available so far (van Kan et al., 2017). In this
361 study, we used the PacBio technology to sequence the full genomes of the Sl3 and Vv3 strains that
362 represent two populations of *B. cinerea* specialized to tomato (T population) and grapevine (G1
363 population; Mercier et al., 2021). We also used additional strains of these two populations and strains
364 from a second population specialized to grapevine (G2) to extend our comparative analyses to a larger
365 set of pathogens.

366

367 **The full genomic assemblies of *B. cinerea* strains from different populations reveal highly syntenic**
368 **core chromosomes**

369 The long-reads generated with the PacBio technology allowed us to generate genome assemblies of
370 43.2 Mb and 44.9 Mb, for the strains Sl3 and Vv3, respectively. These assembly lengths are close to
371 that of the reference strain B05.10 (42.6 Mb). Additionally, almost all genes (>99%) previously
372 annotated in the strain B05.10 were identified in the generated assemblies. As shown by sequence
373 comparison and gene annotation, the three strains share 16 CCs with a high level of synteny with only
374 few events of inversions or translocations. The same number of CCs was also observed in the other
375 genomes of *Botrytis* and Sclerotiniaceae species (Derbyshire et al., 2017; Valero-Jiménez et al., 2020).
376 The high level of synteny between the CCs of *B. cinerea* strains is consistent with their interfertility, as
377 synteny is needed for appropriate chromosome pairing during meiosis. Direct evidence for interfertility
378 comes from a fertile progeny that was obtained by crossing Sl3 and Vv3 (M. Viaud, *unpublished data*).

379 Indirect evidence for interfertility can be found in population genetic analyses, which previously
380 indicated gene flow between the T and G1 groups to which Sl3 and Vv3 belong (Mercier et al., 2021).

381

382 **The accessory chromosomes BCIN19 is specific to the *B. cinerea* G1 population specialized on**
383 **grapevine**

384 The full genomic assemblies generated in this study revealed that Sl3 and Vv3 strains have different
385 pairs of ACs characterized by their small sizes (0.2-0.5 Mbp), their low GC content (25-38%) and a high
386 density of TEs (28-35%). While the Sl3 strain carries two ACs similar to those of B05.10 (BCIN17 and
387 BCIN18), the Vv3 strain carries BCIN17 and a newly characterized AC named BCIN19. Extending our
388 study to 35 *B. cinerea* strains confirmed the dispensability of BCIN17, BCIN18 and BCIN19 and did not
389 support a strict association between ACs and the host of origin. Nevertheless, our data suggest a high
390 frequency of the BCIN17/BCIN19 pair of ACs in the G1 group with BCIN19 being specific to this group.
391 Exploring further the distribution and conservation of the ACs in larger samples would allow to define
392 more precisely to which extent some parts of the accessory genome are associated with a specific
393 population. Though the function of fungal ACs remains largely elusive, several studies on
394 phytopathogenic species reported that ACs could carry genes encoding effectors or genes involved in
395 the synthesis of toxic metabolites (Bertazzoni et al., 2018; Meena et al., 2017; Yang et al., 2020). Our
396 study indicates that ACs are an important source of variation in gene content in *B. cinerea*, but at this
397 stage, none of the 78 genes carried by the ACs has a predicted function indicating a direct role in
398 pathogenicity. Accessory chromosomes could also be considered as reservoirs of TEs. Interestingly,
399 BCIN19 carries several copies of the Gypsy_7 retrotransposon that is specific of the G1 population and
400 generates small RNAs.

401

402

403 **The genome of the *B. cinerea* strain Vv3 has a larger repertoire of transposable elements than the**
404 **Sl3 and B05.10 strains**

405 Transposons are important features of fungal genomes playing key roles in genome structure, genome
406 plasticity (Fouché et al., 2021), and production of small RNAs (Weiberg et al., 2013). TEs identified in
407 the three fully sequenced and assembled genomes of *B. cinerea* strains B05.10, Sl3 and Vv3 cover 3.7,
408 4.5 and 7.7% of their genomic sequences, respectively. Variation in TE content explains most of the
409 variation in total assembly length, with the genomic sequences varying between 41.0 and 41.4 Mb
410 when TEs are excluded. As already observed in the B05.10 strain (van Kan et al., 2017), TEs identified
411 in the Sl3 and Vv3 strains are frequent in the subtelomeric regions of the CCs, as well as in ACs where
412 they can cover up to one third of the sequence. While previous studies already indicated variation in
413 the presence of TEs (e.g *Boty* and *Flipper*) in *B. cinerea* strains (Amselem et al., 2011; Porquier et al.,
414 2021), our study is the first that compares TEs between fully assembled genomes. Our data clearly
415 show that the repertoires of TEs are very different between the three sequenced strains and 21 new
416 subfamilies of TEs were recovered, mostly from the Vv3 strain. As in many fungi, TEs of *B. cinerea* could
417 be inactivated by RIP (Amselem et al., 2015; Porquier et al., 2016, 2019, 2021). In our analysis, the GC
418 content of TE copies suggests that RIP is less active in the Vv3 strain, consistent with its larger TE
419 content. This larger content is due to retrotransposons, TIR transposons, as well as to a Helitron-like
420 TE. Several subfamilies of the TEs identified in this strain, including retrotransposons and the Helitron-
421 like TE, do not show any traces of RIP and could possibly be able to transpose in new *loci.* Both inactive
422 and active TEs can additionally generate genomic rearrangements by homologous recombination as
423 discussed below.

424

425 **Rearrangements at chromosome ends are a source of gene gains and losses in *B. cinerea***

426 Our synteny and PAV analyses indicate that the few rearrangements occurring in the CCs correspond
427 mainly to exchanges of chromosomes ends, and that most gene gains or losses also happen in these

428 TE-rich regions. In many fungi, subtelomeric regions are enriched in secondary metabolism gene
429 clusters (Kjærbølling et al., 2020). Therefore, rearrangements in these regions could contribute to the
430 intraspecific diversity of the biosynthetic clusters (Olarte et al., 2019). In the genome of *B. cinerea*, six
431 out of the approximately 40 identified secondary metabolism gene clusters are in subtelomeric regions
432 (Amselem et al., 2011; van Kan et al., 2017). The present study provides evidence that the evolution of
433 some of these subtelomeric clusters is subjected to chromosomal ends exchange (PKS7 cluster) as well
434 as to gene gains and losses (MelA-like and DTC clusters). In other studies, partial or total loss of the
435 subtelomeric gene cluster responsible for the biosynthesis of the botcinic acid phytotoxin was
436 observed in some rare strains of *B. cinerea* and in other species of *Botrytis* (Plesken et al., 2021; Valero-
437 Jiménez et al., 2020).

438 In addition to gene gains and losses, our study revealed that duplication of *B. cinerea* genes occurs in
439 subtelomeric regions. For example, the four colocalized genes (Bcin08g00060 to 90) possibly related
440 to detoxification of reactive oxygen species or plant cell wall degradation and previously identified as
441 duplicated in strains specialized to tomato (T group; Mercier et al., 2021) were localized on two
442 different chromosomes ends of the Sl3 strain.

443 Subtelomeric regions of all three B05.10, Sl3 and Vv3 strains were found to be enriched in TEs, and TEs
444 were detected in the exchanged chromosomal ends mentioned above. Similar observations made in
445 other fungal genomes led to the hypothesis that TEs may be responsible for repeat-driven
446 recombination and gene rearrangements (Fouché et al., 2021; Lind et al., 2017; Olarte et al., 2019). If
447 these exchanges of chromosomes ends are later followed by a sexual cycle in which chromosome
448 pairing occurs between CCs that have different extremities, it could result in gene gain or loss. Indeed,
449 in this scenario, the progeny would include individuals with two copies of one initial subtelomeric
450 region or none of them. This two-steps mechanism could have an important role in the evolution of
451 the repertoire of secondary metabolites produced by *B. cinerea.*

452

453 **A helicase encoding gene and a Helitron-like element shape the subtelomeric regions of the genome**
454 **of the *B. cinerea* strain Vv3**

455 Analysis of the genome of the *B. cinerea* strain Vv3 revealed that subtelomeric regions show a
456 particular structure with the co-occurrence of a Telomere-Linked Helicase (TLH) encoding gene and a
457 Helitron-like TE that were not described so far in this species. Genes encoding TLHs were previously
458 described in the subtelomeric regions of the fungal pathogens *M. oryzae* and in *U. maydis* (Gao et al.,
459 2002; Rehmeyer et al., 2009; Sánchez-Alonso & Guzmán, 1998). The Vv3 genome contains 14 copies
460 of the *Bctlh* gene, all in subtelomeric regions and, strikingly, all flanked in 3' and/or 5' by a copy of a
461 Helitron-like TE. Helitrons were described in other fungi *i.e. Pleurotus ostreatus* and *Fusarium*
462 *oxysporum* (Castanera et al., 2014; Chellapan et al., 2016) but, to our knowledge, this is the first study
463 that shows a strong enrichment of this family of TEs in subtelomeric regions. One hypothesis to explain
464 the localization of the Helitron-like TE of *B. cinerea* could be that it preferentially inserts in the flanking
465 regions of the *Bctlh* gene copies. Nevertheless, as no nuclease domain was detected in the sequence
466 of Helitron-like TE, this may question its ability to transpose (Kojima, 2019). Alternatively, the
467 occurrence of the *Bcthl* gene and Helitron-like TE in the subtelomeric regions could arise from
468 recombination between chromosome ends followed by sexual crosses leading to duplications as
469 discussed above. Indeed, when these mechanisms are repeated in successive generations, they could
470 lead to the homogenization of the subtelomeric regions. Previous studies revealed a drastic reduction
471 of the number of Restriction Fragment Length Polymorphisms (RFLPs) in subtelomeric regions of *B.*
472 *cinerea* strains isolated from grapevine suggestion a possible homogenization of these regions (Levis
473 et al., 1997a). Based on our analysis, one may wonder whether this homogenization is due to the
474 amplification of the *Bcthl* genes and Helitron-like TEs by exchanges and duplications of chromosomes
475 ends.

476 Another original feature of the 15.5 kb sequence of the Vv3 Helitron-like TE is that, as in some other
477 Helitrons elements (Castanera et al., 2014), it carries captured genes. It would be interesting to

478 evaluate the expression of these captured genes especially the one encoding a putative secreted
479 pectate-lyase and to test whether the numerous copies (33) present in the Vv3 genome confer a higher
480 ability to degrade host tissues. Our analysis to *B. cinerea* populations indicated that the Helitron-like
481 TE is present in all the strains of the G1 group, in about half of the strains of the G2 group and in a
482 minority of strains of the T group which further questions its potential role in the interaction with the
483 host plant.

484

485 **Both Gypsy and Copia retrotransposons generates small RNAs in *B. cinerea***

486 *Botrytis cinerea* produces siRNA that are acting as effectors in the host plant where they can highjack
487 the silencing machinery to impair the expression of genes involved in the defence process (Weiberg et
488 al., 2013). As in other fungal models, these small RNAs are mainly synthesized from retrotransposons.
489 The work just published by Porquier et al., (2021) actually demonstrated that the Gypsy_3 TE could be
490 considered as a pathogenicity factor by itself as the introduction of Gypsy_3 in a strain lacking this TE
491 resulted in the production of siRNA and an enhanced aggressiveness on tomato and on *Arabidopsis*
492 *thaliana*. In our study, we demonstrated that the Sl3 and Vv3 strains that have very different
493 repertoires of TEs are consequently able to produce different sets of TE-derived small RNAs. Seven
494 retrotransposons, five Gypsy subfamilies but also two Copia subfamilies, were identified as responsible
495 for their production. A common feature of all the copies of TEs that generate small RNAs in the Vv3
496 and Sl3 strains is that they show a relatively high GC content. By contrast, the AT-rich copies of the
497 same subfamilies do not produce small RNAs. These data corroborate those of Porquier et al. (2021),
498 who observed that the production of small RNAs in the B05.10 strain correlates positively with the
499 retrotransposon GC content and negatively with a RIP index. This suggests that only retrotransposons
500 that are not inactivated by RIP are able to be expressed and therefore to generate small RNAs.
501 Furthermore, the small RNAs generated by the seven retrotransposons present in Sl3 and/or Vv3
502 strains all show the same features, *i.e.* a size of 21-22 nucleotides and a preference for a 5' terminal U.

503 These features are believed to be required to associate with the plant AGO1 and activate the gene
504 silencing process (Weiberg et al., 2013).

505 Among the seven small RNA generating-retrotransposons that were identified, three Gypsy subfamilies
506 *i.e*. Boty/Gypsy_1, Gypsy_3 and Gypsy_4 are present in both Vv3 and Sl3 strains as well as in B05.10
507 (Porquier et al., 2021) and are responsible for the synthesis of a common set of small RNAs that
508 includes the characterized effectors siR5, siR3.1 and siR3.2 (Weiberg et al., 2013). In addition to the
509 shared repertoire of small RNAs mentioned above, the Vv3 and Sl3 strains produce additional small
510 RNAs through the Copia_4, Gypsy_6, Gypsy_7 (Vv3) and Copia_6 (Sl3) retrotransposons. These newly
511 characterized TEs and the small RNAs that they generate therefore confirm that the difference in the
512 repertoires of retrotransposons present in *B. cinerea* strains Sl3 and Vv3 has strong impact on the
513 production of potential small RNA effectors.

514

515 **The *B. cinerea* G1 population specialized on grapevine produces a specific set of small RNAs**

516 Investigating the repertoires of small RNAs and the presence of some of the small RNA-producing TEs
517 in a larger set of strains *B. cinerea* highlighted some significative differences between the three
518 considered populations. While some sets of TE-derived small RNAs are present in strains belonging to
519 different populations, some seem to be specific to one genetic group. Small RNAs produced by
520 Boty/Gypsy_1, Gypsy_3 and Gypsy_4 are not only produced by the B05.10, Sl3 and Vv3 strains but
521 appear to be commonly produced in the T and G1 populations which could suggest that this set plays
522 a conserved role that is important in the interaction with different hosts. This is indeed the case for
523 the characterized Gypsy_3-derived siRNAs that were shown to silence genes both in *Arabidopsis*
524 *thaliana* and in tomato (Weiberg et al., 2013 ; Porquier et al., 2021). Regarding the three analyzed
525 strains of the G2 population, only Boty and Gypsy_4-derived small RNAs could be identified but
526 uncharacterized TEs may generate other small RNAs. A full assembly of the genome of one the G2
527 strains would be required to answer this question.

528 Finally, our data revealed that two retrotransposons, Copia_4, and to a lesser extent Gypsy_7, occur
529 at high frequency in the G1 population where they jointly generate an important proportion of the
530 arsenal of TE-derived small RNAs. The genome of the Vv3 strain harbors ten complete copies of
531 Copia_4 on CCs, and 23 complete copies of Gypsy_7 shared between the CCs and the AC BCIN19 that
532 is specific to the G1 population. By contrast, these two TEs were very rarely detected by PCR in other
533 strains of the T and G2 populations (this study) and they were not either retrieved in the genomic
534 sequences of *B. cinerea* strains B05.10 (van Kan et al., 2017), T4 (Amselem et al., 2011) and BcDW1
535 (Blanco-Ulate et al., 2013). Using a larger number of strains from the G1 and G2 populations as well as
536 strains from other hosts would be valuable to investigate further the distribution of these sources of
537 small RNAs. If they are confirmed to be mainly present in the G1 population specialized on grapevine,
538 it will be tempting to hypothesize that they have been maintained because they have a significant role
539 in this interaction, either directly with the host *i.e.* through the silencing of grapevine genes, either
540 indirectly *e.g.* by acting on other microorganisms present in the same ecological niche. Nevertheless,
541 the fact that the G2 strains do not have the Copia_4 and Gypsy_7 TEs suggests that their potential role
542 is not needed in all *B. cinerea* populations found on grapevine. Functional studies will be required to
543 answer these questions.

544

545 **Conclusion**

546 Our study was conducted in order to investigate the genomic determinants of host specialization in *B.*
547 *cinerea.* Previous work revealed widespread signatures of positive selection in the T population
548 specialized to tomato, with genes under positive selection encoding cellulases, pectinases and
549 enzymes involved in the oxidative stress response suggesting that these activities may contribute to
550 the specialization on tomato (Mercier et al., 2021). The present study substantially extends these
551 findings by revealing that populations of *B. cinerea* specialized on different hosts harbor different sets
552 of accessory chromosomes, repertoires of transposons and their derived small RNAs. Hence, we

553 identified genomic features that are specific to the main population specialized to grapevine in France
554 (the G1 population). The retrotransposons Copia_4 and Gypsy_7 allow the production of a G1-specific
555 set of small RNAs with structures similar to the known effector siRNA (Weiberg et al., 2013). In
556 addition, strains of the G1 population specifically carry the newly characterized BCIN19 accessory
557 chromosome that includes several copies of the Gypsy-7 elements and newly identified genes whose
558 functions remained to be characterized. Our characterization and analysis of new genomic data in
559 populations of *Botrytis* specialized to different hosts pave the way for new molecular investigations of
560 the mechanisms underlying host specialization in this polyphagous pathogen.

561

562

563 **Methods**

564 **Genome sequencing and assembly**

565 The *Botrytis cinerea* Sl3 and Vv3 strains were isolated from the Champagne region (France) respectively
566 from tomato (cultivar Moneymaker) and grapevine berries (Pinot noir), as previously described
567 (Mercier et al., 2021). For DNA sequencing, these strains were cultivated three days in liquid malt
568 medium. High-molecular-weight DNA was extracted using a sarkosyl procedure in which the ethanol-
569 precipitated DNA was fished with a glass hook to avoid centrifugation and DNA breaks. Sl3 and Vv3
570 genomic DNA were sequenced using a PacBio Sequel sequencer (KeyGene, Wageningen, NL). Total
571 lengths of respectively 8.3Gb and 4.7Gb reads were obtained with mean lengths of 8.4Kb and 11.3Kb.
572 The theorical coverages of the genome were thus of 195X and 120X. Reads were assembled with
573 HGAP4 (SMRT-LINK v5.0.1) (https://github.com/PacificBiosciences/) and CANU v1.6 (Koren et al., 2017).
574 HGAP4 was run with the set of parameters recommended for fungal genomes (https://pb-
575 falcon.readthedocs.io/en/latest/parameters.html). CANU was run with default parameters and an
576 expected coverage of 42Mb. HGAP assemblies were polished with pilon (Walker et al., 2014) using Sl3
577 or Vv3 Illumina reads (Mercier et al., 2021). CANU assemblies were polished with both arrow
578 (https://github.com/PacificBiosciences/) and PILON. Mitochondrial, ribosomal and small redundant
579 contigs were removed, and final versions of assemblies were manually compiled from polished runs of
580 HGAP and CANU assemblies. Both genome assemblies were evaluated and compared with B05.10
581 referent genome (van Kan et al., 2017) using QUAST (Gurevich et al., 2013).

582

583 **Transposable elements prediction**

584 Transposable elements (TEs) were annotated using the REPET package
585 (https://urgi.versailles.inra.fr/Tools/REPET ; (Amselem et al., 2015) as previously described in Porquier
586 et al., 2016 for the genome of the *B. cinerea* reference strain B05.10. Briefly, the TEDENOVO pipeline

587 (Flutre et al., 2011) was used to detect repeated elements in the genome and to provide a consensus
588 sequence for each family. Consensus sequences were then classified using the PASTEC tool (Hoede et
589 al., 2014), based on the Wicker hierarchical TE classification system (Wicker et al., 2007). After manual
590 correction, the resulting library of consensus sequences was used to annotate TE copies in the whole
591 genome using the TEANNOT pipeline (Quesneville et al., 2005). Consensus sequences identified from
592 different genomes were compared and considered as the same subfamily based on a bidirectional best
593 hit approach (blastn with evalue < 1e-10 and coverage > 70%). Presence of the subfamilies in the
594 genomes of additional strains (Table 1) was tested by PCR using the MyTaq polymerase (Bioline) and
595 the primers listed in Sup. Table S5.

596

597 **Gene prediction, synteny**

598 The structural annotation of B05.10 genes (van Kan et al., 2017)
599 (http://fungi.ensembl.org/Botrytis_cinerea/) was transferred to Vv3 and Sl3 genomes using the LIFTOFF
600 annotation mapping tool (Shumate & Salzberg, 2021). Genes were also predicted *de novo* using the
601 FGENESH *ab initio* gene-finder (Solovyev et al., 2006) (http://www.softberry.com/berry.phtml) with the
602 *Botrytis* matrix. Presence/absence of genes in other genomes was determined with blastn analyses.
603 The synteny between B05.10, Sl3 and Vv3 genes was analysed with SYNCHRO (Drillon et al., 2014), which
604 detects ortholog proteins with Reciprocal Best Hit (RBH). SYNCHRO was run with a best score threshold
605 (min_sim_RBH) of 80, a length ratio threshold (max_diflen) of 1.3 and a delta value of 3 (medium
606 stringency). Duplications of gene clusters were further explored with CLINKER (Gilchrist & Chooi, 2021).
607 Presence of the gene *Bctlh* in the genomes of additional strains (Table 1) was tested by PCR using the
608 MyTaq polymerase and the primers indicated in Sup. Table S5.

609

610

611 **Small RNA sequencing and analysis**

612 *Botrytis cinerea* strains were cultivated 48 hours on solid media made from grape juice or tomato juice
613 supplemented with agar and covered with a sheet of cellophane as previously described (Simon et al.,
614 2013). RNAs were extracted with the TRIzol reagent (InvitroGen) and submitted to a DNAse treatment
615 (DNA-free kit, Ambion). Total RNAs were purified with miRNeasy kit (Qiagen) which allows the
616 selection of the sRNA fraction less than 100 bases. Small RNA (sRNA) libraries were prepared and
617 sequenced at Integragen (https://www.integragen.com/). Libraries were generated following
618 (Vigneault et al., 2012), with adjustments to improve ligation, from at least 1 µg of extracted total RNA
619 with a RIN greater than 7. Libraries were sequenced on Illumina HiSeq4000. Quality of raw sequence
620 data was checked with FASTQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/).
621 CUTADAPT (Martin, 2011) was used to remove adapters and processed reads shorter than 16 bases were
622 discarded. Reads were then pre-processed using FASTX-TOOLKIT
623 (http://hannonlab.cshl.edu/fastx_toolkit/) as in (Weiberg et al., 2013), *i.e.* low-quality sequences were
624 masked with FASTQ_QUALITY_FILTER -q 30 -p 70, low-complexity sequences were masked with
625 FASTX_ARTIFACTS_FILTER, and finally identical sequences were counted with FASTX_COLLAPSER. Only
626 sequences 20-24 nucleotides length and with more than five reads per million in at least one library
627 were kept for further analyses. Quality controls were applied, including Principal Component Analyses
628 (PCA), to visualize the distance between samples. Reads were mapped against B05.10 (van Kan et al.,
629 2017), Sl3 and Vv3 genomes and against TE consensus with the GLINT software (https://forge-
630 dga.jouy.inra.fr/projects/glint/wiki); only perfect matches (100% identity and 100% coverage) were
631 considered.

632

633 **Acknowledgements**

634 We are grateful to the following bioinformatics platforms and partners for providing computational
635 support and/or storage resources : Bioinfo Genotoul,

(https://doi.org/10.15454/1.5572369328961167E12), CATI BARIC (https://www.cesgo.org/catibaric/), INRAE-URGI, INRAE-LIPME Bioinfo (Jérôme Gouzy, Sébastien Carrère, Erika Sallet) and INRAE-BIOGER Bioinfo (https://bioinfo.bioger.inrae.fr/ ; Nicolas Lapalu). We thank Alexander H. J. Wittenberg (KeyGene, Wageningen, NL) and Jean-Paul Saraiva (Integragen, Evry, France) for the PacBio and Illumina sequencing, respectively.

Funding

This work was supported by a grant overseen by the French National Research Agency (ANR) as part of the "Investments d'Avenir" Programme (LabEx BASC; ANR-11-LABX-0034) and by the INRAE department SPE. The BIOGER unit also benefits from the support of "Saclay Plant Science-SPS" (ANR-17-EUR-0007). AM was supported by a grant from the Doctoral School « Sciences du Végétal », Université Paris-Saclay.

Availability of data and materials

The genomic raw data were deposited into the NCBI SRA under the accessions PRJNA752967 for Sl3 and PRJNA752962 for Vv3. Genome assemblies were deposited at NCBI under the accessions CP080979-CP080996 for Sl3 and CP080961-CP080978 for Vv3. Small RNAs raw data and the matrix of sequence counts were deposited into the NCBI GEO under accession GSE181592. Furthermore, the genomic and annotation (gene/TE) fasta and gff files, as well as Sl3 and Vv3 genome browsers are available at the INRAE BIOGER Bioinformatics platform (https://bioinfo.bioger.inrae.fr/).

658 **References**

659

660 Amselem, J., Cuomo, C. A., van Kan, J. A. L., Viaud, M., Benito, E. P., Couloux, A., Coutinho, P. M., de
661 Vries, R. P., Dyer, P. S., Fillinger, S., Fournier, E., Gout, L., Hahn, M., Kohn, L., Lapalu, N.,
662 Plummer, K. M., Pradier, J. M., Quévillon, E., Sharon, A., … Dickman, M. (2011). Genomic
663 analysis of the necrotrophic fungal pathogens sclerotinia sclerotiorum and botrytis cinerea.
664 *PLoS Genetics*, *7*(8). https://doi.org/10.1371/journal.pgen.1002230

665 Amselem, J., Lebrun, M. H., & Quesneville, H. (2015). Whole genome comparative analysis of
666 transposable elements provides new insight into mechanisms of their inactivation in fungal
667 genomes. *BMC Genomics*, *16*(1). https://doi.org/10.1186/s12864-015-1347-1

668 Atwell, S., Corwin, J. A., Soltis, N. E., Subedy, A., Denby, K. J., & Kliebenstein, D. J. (2015). Whole
669 genome resequencing of Botrytis cinerea isolates identifies high levels of standing diversity.
670 *Frontiers in Microbiology*, *6*(SEP). https://doi.org/10.3389/FMICB.2015.00996

671 Bertazzoni, S., Williams, A. H., Jones, D. A., Syme, R. A., Tan, K.-C., & Hane, J. K. (2018). Accessories
672 Make the Outfit: Accessory Chromosomes and Other Dispensable DNA Regions in Plant-
673 Pathogenic Fungi. *Molecular Plant-Microbe Interactions®*, *31*(8), 779–788.
674 https://doi.org/10.1094/mpmi-06-17-0135-fi

675 Blanco-Ulate, B., Allen, G., Powell, A. L. T., & Cantu, D. (2013). Draft genome sequence of Botrytis
676 cinerea BcDW1, inoculum for noble rot of grape berries. *Genome Announcements*, *1*(3).
677 https://doi.org/10.1128/genomeA.00252-13

678 Castanera, R., Pérez, G., López, L., Sancho, R., Santoyo, F., Alfaro, M., Gabaldón, T., Pisabarro, A. G.,
679 Oguiza, J. A., & Ramírez, L. (2014). Highly expressed captured genes and cross-kingdom domains
680 present in Helitrons create novel diversity in pleurotus ostreatus and other fungi. *BMC*

681 *Genomics*, *15*(1). https://doi.org/10.1186/1471-2164-15-1071

682 Chellapan, B. V., Van Dam, P., Rep, M., Cornelissen, B. J. C., & Fokkens, L. (2016). Non-canonical
683 Helitrons in Fusarium oxysporum. *Mobile DNA*, *7*(1). https://doi.org/10.1186/S13100-016-0083-
684 7

685 Choquer, M., Fournier, E., Kunz, C., Levis, C., Pradier, J. M., Simon, A., & Viaud, M. (2007). Botrytis
686 cinerea virulence factors: New insights into a necrotrophic and polyphageous pathogen. In
687 *FEMS Microbiology Letters* (Vol. 277, Issue 1, pp. 1–10). FEMS Microbiol Lett.
688 https://doi.org/10.1111/j.1574-6968.2007.00930.x

689 Derbyshire, M., Denton-Giles, M., Hegedus, D., Seifbarghi, S., Rollins, J., Kan, J. Van, Seidl, M. F.,
690 Faino, L., Mbengue, M., Navaud, O., Raffaele, S., Hammond-Kosack, K., Heard, S., & Oliver, R.
691 (2017). The complete genome sequence of the phytopathogenic fungus Sclerotinia sclerotiorum
692 reveals insights into the genome architecture of broad host range pathogens. *Genome Biology*
693 *and Evolution*, *9*(3), 593–618. https://doi.org/10.1093/gbe/evx030

694 Diolez, A., Marches, F., Fortini, D., & Brygoo, Y. (1995). Boty, a long-terminal-repeat retroelement in
695 the phytopathogenic fungus Botrytis cinerea. *Applied and Environmental Microbiology*, *61*(1),
696 103–108. https://doi.org/10.1128/AEM.61.1.103-108.1995

697 Drillon, G., Carbone, A., & Fischer, G. (2014). SynChro: A fast and easy tool to reconstruct and
698 visualize synteny blocks along eukaryotic chromosomes. *PLoS ONE*, *9*(3).
699 https://doi.org/10.1371/journal.pone.0092621

700 Elad, Y., Vivier, M., & Fillinger, S. (2015). Botrytis, the good, the bad and the ugly. In *Botrytis - The*
701 *Fungus, the Pathogen and its Management in Agricultural Systems* (pp. 1–15). Springer
702 International Publishing. https://doi.org/10.1007/978-3-319-23371-0_1

703 Flutre, T., Duprat, E., Feuillet, C., & Quesneville, H. (2011). Considering transposable element

704 diversification in de novo annotation approaches. *PLoS ONE*, *6*(1).
705 https://doi.org/10.1371/journal.pone.0016526

706 Fouché, S., Oggenfuss, U., Chanclud, E., & Croll, D. (2021). A devil's bargain with transposable
707 elements in plant pathogens. *Trends in Genetics : TIG*, *38*(3), 222–230.
708 https://doi.org/10.1016/J.TIG.2021.08.005

709 Gao, W., Khang, C. H., Park, S. Y., Lee, Y. H., & Kang, S. (2002). Evolution and organization of a highly
710 dynamic, subtelomeric helicase gene family in the rice blast fungus Magnaporthe grisea.
711 *Genetics*, *162*(1), 103–112. https://doi.org/10.1093/genetics/162.1.103

712 Geib, E., Gressler, M., Viediernikova, I., Hillmann, F., Jacobsen, I. D., Nietzsche, S., Hertweck, C., &
713 Brock, M. (2016). A Non-canonical Melanin Biosynthesis Pathway Protects Aspergillus terreus
714 Conidia from Environmental Stress. *Cell Chemical Biology*, *23*(5), 587–597.
715 https://doi.org/10.1016/j.chembiol.2016.03.014

716 Gilchrist, C. L. M., & Chooi, Y.-H. (2021). clinker & clustermap.js: automatic generation of gene cluster
717 comparison figures. *Bioinformatics*, *37*(16), 2473–2475.
718 https://doi.org/10.1093/bioinformatics/btab007

719 Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J. L. N., Nhani, A., Chen, L., Terauchi, R.,
720 Lebrun, M. H., Tharreau, D., Mitchell, T., Pedley, K. F., Valent, B., Talbot, N. J., Farman, M., &
721 Fournier, E. (2018). Gene Flow between Divergent Cereal- and Grass-Specific Lineages of the
722 Rice Blast Fungus Magnaporthe oryzae. *MBio*, *9*(1). https://doi.org/10.1128/MBIO.01219-17

723 Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome
724 assemblies. *Bioinformatics*, *29*(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

725 Hoede, C., Arnoux, S., Moisset, M., Chaumier, T., Inizan, O., Jamilloux, V., & Quesneville, H. (2014).
726 PASTEC: An automatic transposable element classification tool. *PLoS ONE*, *9*(5).

https://doi.org/10.1371/journal.pone.0091929

Kjærbølling, I., Vesth, T., Frisvad, J. C., Nybo, J. L., Theobald, S., Kildgaard, S., Petersen, T. I., Kuo, A., Sato, A., Lyhne, E. K., Kogle, M. E., Wiebenga, A., Kun, R. S., Lubbers, R. J. M., Mäkelä, M. R., Barry, K., Chovatia, M., Clum, A., Daum, C., … Andersen, M. R. (2020). A comparative genomics study of 23 Aspergillus species from section Flavi. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-019-14051-y

Kojima, K. K. (2019). Structural and sequence diversity of eukaryotic transposable elements. *Genes and Genetic Systems*, *94*(6), 233–252. https://doi.org/10.1266/ggs.18-00024

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/gr.215087.116

Levis, C., Giraud, T., Dutertre, M., Fortini, D., & Brygoo, Y. (1997a). Telomeric DNA of Botrytis cinerea: a useful tool for strain identification. *FEMS Microbiology Letters*, *157*(2), 267–272. https://doi.org/10.1111/J.1574-6968.1997.TB12783.X

Levis, C., Fortini, D., & Brygoo, Y. (1997b). Flipper, a mobile Fot1-like transposable element in Botrytis cinerea. *Molecular and General Genetics*, *254*(6), 674–680. https://doi.org/10.1007/s004380050465

Lind, A. L., Wisecaver, J. H., Lameiras, C., Wiemann, P., Palmer, J. M., Keller, N. P., Rodrigues, F., Goldman, G. H., & Rokas, A. (2017). Drivers of genetic diversity in secondary metabolic gene clusters within a fungal species. *PLoS Biology*, *15*(11). https://doi.org/10.1371/JOURNAL.PBIO.2003583

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, *42*(D1).

750 https://doi.org/10.1093/nar/gkt1178

751 Martin, M. (2011). TECHNICAL NOTES. In *EMBnet.journal* (Vol. 17, Issue 1). http://www-
752 huber.embl.de/users/an-

753 Mbengue, M., Navaud, O., Peyraud, R., Barascud, M., Badet, T., Vincent, R., Barbacci, A., & Raffaele,
754 S. (2016). Emerging trends in molecular interactions between plants and the broad host range
755 fungal pathogens botrytis cinerea and sclerotinia sclerotiorum. In *Frontiers in Plant Science* (Vol.
756 7, Issue MAR2016). Frontiers Media S.A. https://doi.org/10.3389/fpls.2016.00422

757 Meena, M., Gupta, S. K., Swapnil, P., Zehra, A., Dubey, M. K., & Upadhyay, R. S. (2017). Alternaria
758 toxins: Potential virulence factors and genes related to pathogenesis. In *Frontiers in*
759 *Microbiology* (Vol. 8, Issue AUG). Frontiers Media S.A.
760 https://doi.org/10.3389/fmicb.2017.01451

761 Mercier, A., Carpentier, F., Duplaix, C., Auger, A., Pradier, J. M., Viaud, M., Gladieux, P., & Walker, A.
762 S. (2019). The polyphagous plant pathogenic fungus Botrytis cinerea encompasses host-
763 specialized and generalist populations. *Environmental Microbiology*, *21*(12), 4808–4821.
764 https://doi.org/10.1111/1462-2920.14829

765 Mercier, A., Simon, A., Lapalu, N., Giraud, T., Bardin, M., Walker, A.S., Viaud, M., & Gladieux, P.
766 (2021). Population genomics reveals molecular determinants of specialization to tomato in the
767 polyphagous fungal pathogen Botrytis cinerea in France. *Phytopathology®*.
768 https://doi.org/10.1094/phyto-07-20-0302-fi

769 Olarte, R. A., Menke, J., Zhang, Y., Sullivan, S., Slot, J. C., Huang, Y., Badalamenti, J. P., Quandt, A. C.,
770 Spatafora, J. W., & Bushley, K. E. (2019). Chromosome rearrangements shape the diversification
771 of secondary metabolism in the cyclosporin producing fungus Tolypocladium inflatum. *BMC*
772 *Genomics*, *20*(1). https://doi.org/10.1186/s12864-018-5399-x

773 Plesken, C., Pattar, P., Reiss, B., Noor, Z. N., Zhang, L., Klug, K., Huettel, B., & Hahn, M. (2021). Genetic
774 Diversity of Botrytis cinerea Revealed by Multilocus Sequencing, and Identification of B. cinerea
775 Populations Showing Genetic Isolation and Distinct Host Adaptation. *Frontiers in Plant Science*,
776 *12*. https://doi.org/10.3389/fpls.2021.663027

777 Porquier, A., Moraga, J., Morgant, G., Dalmais, B., Simon, A., Sghyer, H., Collado, I. G., & Viaud, M.
778 (2019). Botcinic acid biosynthesis in Botrytis cinerea relies on a subtelomeric gene cluster
779 surrounded by relics of transposons and is regulated by the Zn2Cys6 transcription factor
780 BcBoa13. *Current Genetics*, *65*(4), 965–980. https://doi.org/10.1007/s00294-019-00952-4

781 Porquier, A., Morgant, G., Moraga, J., Dalmais, B., Luyten, I., Simon, A., Pradier, J. M., Amselem, J.,
782 Collado, I. G., & Viaud, M. (2016). The botrydial biosynthetic gene cluster of Botrytis cinerea
783 displays a bipartite genomic structure and is positively regulated by the putative Zn(II)2Cys6
784 transcription factor BcBot6. *Fungal Genetics and Biology*, *96*, 33–46.
785 https://doi.org/10.1016/j.fgb.2016.10.003

786 Porquier, A., Tisserant, C., Salinas, F., Glassl, C., Wange, L., Enard, W., Hauser, A., Hahn, M., &
787 Weiberg, A. (2021). Retrotransposons as pathogenicity factors of the plant pathogenic fungus
788 Botrytis cinerea. *Genome Biology*, *22*(1). https://doi.org/10.1186/s13059-021-02446-4

789 Quesneville, H., Bergman, C. M., Andrieu, O., Autard, D., Nouaud, D., Ashburner, M., & Anxolabehere,
790 D. (2005). Combined evidence annotation of transposable elements in genome sequences. *PLoS*
791 *Computational Biology*, *1*(2), 0166–0175. https://doi.org/10.1371/journal.pcbi.0010022

792 Rehmeyer, C. J., Li, W., Kusaba, M., & Farman, M. L. (2009). The telomere-linked helicase (TLH) gene
793 family in Magnaporthe oryzae: Revised gene structure reveals a novel TLH-specific protein
794 motif. In *Current Genetics* (Vol. 55, Issue 3, pp. 253–262). Curr Genet.
795 https://doi.org/10.1007/s00294-009-0240-3

796 Sánchez-Alonso, P., & Guzmán, P. (1998). Organization of chromosome ends in Ustilago maydis.

RecQ-like helicase motifs at telomeric regions. *Genetics*, *148*(3), 1043–1054. https://doi.org/10.1093/genetics/148.3.1043

Shumate, A., & Salzberg, S. L. (2021). Liftoff: accurate mapping of gene annotations. *Bioinformatics*, *37*(12), 1639–1643. https://doi.org/10.1093/bioinformatics/btaa1016

Simon, A., Dalmais, B., Morgant, G., & Viaud, M. (2013). Screening of a Botrytis cinerea one-hybrid library reveals a Cys2His2 transcription factor involved in the regulation of secondary metabolism gene clusters. *Fungal Genetics and Biology*, *52*, 9–19. https://doi.org/10.1016/j.fgb.2013.01.006

Solovyev, V., Kosarev, P., Seledsov, I., & Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biology*, *7 Suppl 1*(Suppl 1). https://doi.org/10.1186/gb-2006-7-s1-s10

Stukenbrock, E. H., & McDonald, B. A. (2008). The Origins of Plant Pathogens in Agro-Ecosystems. *Http://Dx.Doi.Org/10.1146/Annurev.Phyto.010708.154114*, *46*, 75–100. https://doi.org/10.1146/ANNUREV.PHYTO.010708.154114

Tolios, A., Teupser, D., & Holdt, L. M. (2015). Preanalytical conditions and DNA isolation methods affect telomere length quantification in whole blood. *PLoS ONE*, *10*(12). https://doi.org/10.1371/journal.pone.0143889

Valero-Jiménez, C. A., Steentjes, M. B. F., Slot, J. C., Shi-Kunne, X., Scholten, O. E., & van Kan, J. A. L. (2020). Dynamics in secondary metabolite gene clusters in otherwise highly syntenic and stable genomes in the fungal genus botrytis. *Genome Biology and Evolution*, *12*(12), 2491–2507. https://doi.org/10.1093/GBE/EVAA218

van Kan, J. A. L., Stassen, J. H. M., Mosbach, A., Van Der Lee, T. A. J., Faino, L., Farmer, A. D., Papasotiriou, D. G., Zhou, S., Seidl, M. F., Cottam, E., Edel, D., Hahn, M., Schwartz, D. C.,

Dietrich, R. A., Widdison, S., & Scalliet, G. (2017). A gapless genome sequence of the fungus Botrytis cinerea. *Molecular Plant Pathology*, *18*(1), 75–89. https://doi.org/10.1111/mpp.12384

Veloso, J., & van Kan, J. A. L. (2018). Many Shades of Grey in Botrytis–Host Plant Interactions. In *Trends in Plant Science* (Vol. 23, Issue 7, pp. 613–622). Elsevier Ltd. https://doi.org/10.1016/j.tplants.2018.03.016

Vigneault, F., Ter-Ovanesyan, D., Alon, S., Eminaga, S., Christodoulou, D. C., Seidman, J. G., Eisenberg, E., & Church, G. M. (2012). High-throughput multiplex sequencing of miRNA. *Current Protocols in Human Genetics*, *Chapter 11*(SUPPL.73). https://doi.org/10.1002/0471142905.hg1112s73

Walker, A. S. (2015). Diversity within and between species of Botrytis. In *Botrytis - The Fungus, the Pathogen and its Management in Agricultural Systems* (pp. 91–125). Springer International Publishing. https://doi.org/10.1007/978-3-319-23371-0_6

Walker, A. S., Gladieux, P., Decognet, V., Fermaud, M., Confais, J., Roudet, J., Bardin, M., Bout, A., C. Nicot, P., Poncet, C., & Fournier, E. (2015). Population structure and temporal maintenance of the multihost fungal pathogen Botrytis cinerea: Causes and implications for disease management. *Environmental Microbiology*, *17*(4), 1261–1274. https://doi.org/10.1111/1462-2920.12563

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., & Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*, *9*(11). https://doi.org/10.1371/journal.pone.0112963

Weiberg, A., Wang, M., Lin, F. M., Zhao, H., Zhang, Z., Kaloshian, I., Huang, H. Da, & Jin, H. (2013). Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science*, *342*(6154), 118–123. https://doi.org/10.1126/science.1239705

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., & Schulman, A. H. (2007). A unified classification system for eukaryotic transposable elements. In *Nature Reviews Genetics* (Vol. 8, Issue 12, pp. 973–982). Nature Publishing Group. https://doi.org/10.1038/nrg2165

Yang, H., Yu, H., & Ma, L. J. (2020). Accessory chromosomes in fusarium oxysporum. In *Phytopathology* (Vol. 110, Issue 9, pp. 1488–1496). American Phytopathological Society. https://doi.org/10.1094/PHYTO-03-20-0069-IA

852 ## Supplemental information:

853 **File S1**: Consensus sequences of the repeated elements identified in the genomes of *B. cinerea* strains
854 Sl3 and Vv3.

855 **Table S1:** Chromosomes of the *B. cinerea* strains B05.10, Sl3 and Vv3.

856 **Table S2:** Genes predicted in the genomes of *B. cinerea* strains B05.10, Sl3 and Vv3.

857 **Table S3**: Repertoires of Transposable Elements (TEs) in the *B. cinerea* strains B05.10, Sl3 and Vv3.

858 **Table S4**: Libraries of small RNAs isolated from *B. cinerea* strains belonging to different populations.

859 **Table S5:** List of PCR primers used to detect the presence of Transposable Elements (TEs) or the gene
860 encoding the Telomere-Linked Telomerase (*BcTLH*).

861

862 **Figure S1 :** Genomic genealogy of *B. cinerea* strains showing the populations T, G1 and G2.

863 **Figure S2:** Synteny between the Core Chromosomes (CCs) of the *B. cinerea* strains B05.10, Sl3 and Vv3.

864 **Figure S3**: Main inversions events observed between the genomes of the *B. cinerea* strains Sl3 and Vv3.

865 **Figure S4:** Localization of the PKS7 secondary metabolism gene cluster in the genomes of the *B. cinerea*
866 strains B05.10 and Vv3.

867 **Figure S5**: Duplication of four contiguous genes in the genome of the *B. cinerea* strain Sl3.

868 **Figure S6:** Structure of the Telomere-Linked Helicase (TLH) identified in the *B. cinerea* strain Vv3.

869 **Figure S7:** Positions of the genes encoding the Telomere-Linked Helicase (BcTLH) and of the Helitron-
870 like Transposable Elements (TE) in the genome of the *B. cinerea* strain Vv3.

871 **Figure S8:** Alignment of the ten consensus sequences corresponding to the Boty_Gypsy_1
872 Transposable Element (TE) identified in the genomes of the *B. cinerea* strains B05.10, Sl3 or Vv3.

873 **Figure S9:** Helitron-like Transposable Element (TE) identified in the genome of the *B. cinerea* strain
874 Vv3.

875 **Figure S10:** Localization of the main superfamilies of Transposable Elements (TEs) in the genomes of
876 the *B. cinerea* strains Sl3 and Vv3.

877 **Figure S11:** GC content of complete copies of the main subfamilies of Transposable Elements (TEs) in
878 the genomes of the *B. cinerea* strains B05.10, Sl3 and Vv3.

879 **Figure S12:** Principal Component Analysis (PCA) of the repertoires of small RNAs in different strains of
880 *B. cinerea*.

881 **Figure S13:** Retrotransposons that generate small RNAs in the *B. cinerea* strains Sl3 and/or Vv3.

882 **Figure S14:** Small RNAs produced by the *B. cinerea* strains Sl3 and Vv3: size distribution and 5'
883 nucleotide.

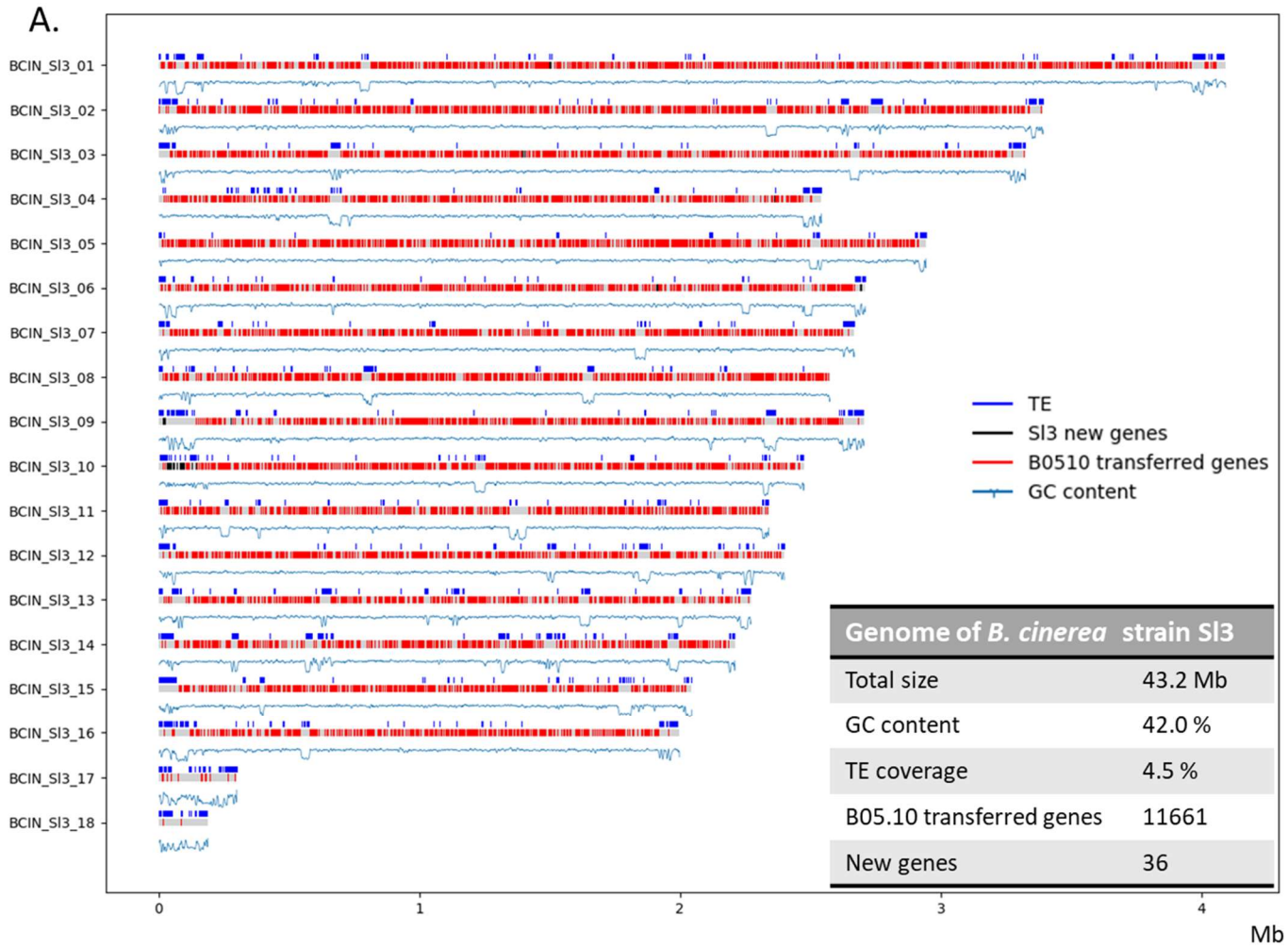
884 **Figure S15:** Positions of complete copies of the TE Copia_4, Gypsy_6 and Gypsy_7 in the *B. cinerea* Vv3
885 genome.

886

Figures & Tables

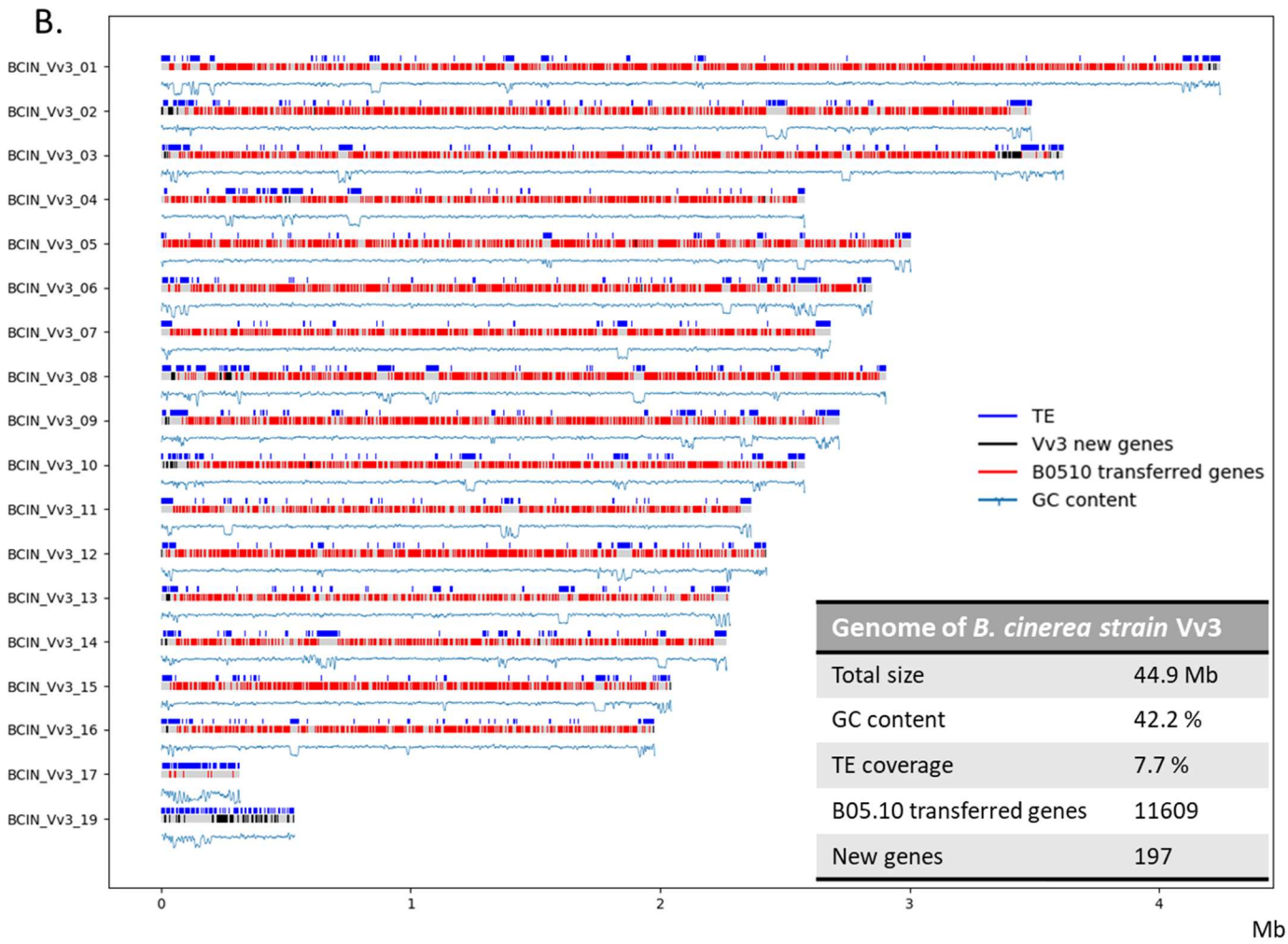
Table 1: **Distribution of Accessory Chromosomes (ACs) and a selection of dispensable genes and Transposable Elements (TEs) in *B. cinerea* strains from different populations.** The genomic features highlighted in the present study, using complete assemblies of Sl3 and Vv3 genomes, were investigated in populations specialized either on grapevine (G1 and G2) either on tomato (T; Mercier et al., 2021). The distribution of ACs and secondary metabolism key genes were investigated using Illumina genomic data (Mercier et al., 2021) while those of the TEs and of the gene encoding the telomere-linked helicase (BcHTL) were investigated by PCR. a. van Kan et al., 2017. b. Mercier et al., 2021. c. Amselem et al., 2011. d. Blanco-Ulate et al., 2013. Nd, not determined. ✔, presence. -, absence. (✔), partial presence of an AC. Nd, Not determined.

Population	Strain	Host	Location	Ref	Accessory chromosomes (Illumina sequencing; Ref b)			Presence of gene (Illumina sequencing; Ref b)			Presence of TE or gene (PCR detection)				
					AC 17	AC 18	AC 19	*Bcstc6*	*BcmelA*	*Bcdtc*	*Bctlh*	Helitron-like	Copia_4	Gypsy_6	Gypsy_7
Nd	B05.10	unknown	unknown	a	✔	✔	-	-	✔	-	-	-	-	-	-
T	Sl1	*Solanum lycopersicum*	Bourgogne	b	✔	-	-	-	-	✔	-	-	-	-	-
	Sl2		Champagne	b	✔	-	-	-	-	✔	-	-	-	-	-
	Sl3		Champagne	b	✔	(✔)	-	✔	-	✔	-	-	-	-	-
	Sl4		Provence	b	✔	-	-	-	-	✔	-	-	-	-	-
	Sl5		Provence	b	✔	✔	-	-	-	✔	✔	-	-	-	-
	Sl6		Rhône-Alpes	b	✔	-	-	-	-	-	-	-	-	-	-
	Sl7		Rhône-Alpes	b	✔	-	-	-	-	✔	-	-	-	-	-
	Sl8		Rhône-Alpes	b	✔	-	-	-	-	-	-	-	-	-	-
	Sl9		Occitanie	b	-	✔	-	-	-	✔	-	-	-	-	-
	Sl10		Occitanie	b	-	-	-	-	-	✔	-	✔	-	-	-
	Sl11		Occitanie	b	✔	-	-	-	-	-	✔	-	-	-	-
	Sl12		Occitanie	b	✔	-	-	-	-	-	✔	-	-	-	-
	Sl13		Occitanie	b	✔	-	-	✔	-	✔	✔	-	-	-	-
G1	Vv1	*Vitis vinifera*	Champagne	b	✔	-	✔	✔	✔	✔	✔	✔	✔	-	✔
	Vv2		Champagne	b	✔	-	✔	✔	✔	✔	✔	✔	✔	-	-
	Vv3		Champagne	b	✔	-	✔	-	✔	✔	✔	✔	✔	✔	✔
	Vv4		Champagne	b	✔	-	✔	-	✔	✔	✔	✔	✔	-	✔
	Vv5		Champagne	b	✔	-	✔	✔	✔	✔	✔	✔	✔	-	✔
	Vv6		Champagne	b	✔	-	✔	-	✔	✔	-	✔	✔	-	✔
	Vv7		Provence	b	✔	-	✔	-	✔	✔	-	✔	✔	-	✔
	Vv10		Provence	b	✔	-	✔	✔	✔	✔	✔	✔	✔	✔	✔
	Vv12		Provence	b	✔	-	✔	✔	✔	✔	✔	✔	✔	-	✔
	Vv13		Provence	b	✔	-	-	-	✔	✔	✔	✔	✔	✔	✔
	Vv14		Provence	b	✔	-	✔	-	✔	✔	✔	✔	✔	-	✔
	Vv16		Champagne	b	✔	-	✔	✔	✔	✔	✔	✔	✔	-	✔
G2	Vv8	*Vitis vinifera*	Provence	b	✔	-	-	-	✔	✔	✔	✔	-	-	✔
	Vv9		Provence	b	✔	✔	-	-	✔	✔	✔	-	-	-	-
	Vv11		Provence	b	✔	✔	-	-	✔	-	✔	-	-	-	-
	Vv15		Provence	b	-	-	-	✔	✔	✔	-	✔	-	-	-
Nd	Rf1	*Rubus fruticosus*	Bourgogne	b	✔	-	-	✔	✔	-	-	-	-	✔	-
Nd	Rf2	*Rubus fruticosus*	Champagne	b	(✔)	-	-	-	-	-	✔	✔	✔	✔	✔
Nd	Hm1	*Hydrangea macrophylla*	Anjou	b	✔	✔	-	-	✔	-	-	-	-	✔	-
Nd	T4	*Solanum lycopersicum*	Provence	c	-	-	-	✔	-	-	-	-	-	-	-
Nd	BcDW1	*Vitis vinifera*	California	d	✔	✔	-	✔	✔	✔	Nd	Nd	Nd	Nd	Nd



Genome of *B. cinerea* strain Sl3	
Total size	43.2 Mb
GC content	42.0 %
TE coverage	4.5 %
B05.10 transferred genes	11661
New genes	36

899



Genome of *B. cinerea strain* Vv3	
Total size	44.9 Mb
GC content	42.2 %
TE coverage	7.7 %
B05.10 transferred genes	11609
New genes	197

Figure 1: Genome organization in *B. cinerea* strains Sl3 (A) and Vv3 (B). The two karyoplots show the 18 chromosomes identified in each strain, genes previously predicted in the B05.10 reference strain (van Kan et al., 2017; red), newly detected genes (black), Transposable Elements (TEs; dark blue) and the GC content (blue) along the Core Chromosomes (CCs BCIN_01 to 16) and the Accessory Chromosomes (ACs BCIN_17 to 19).

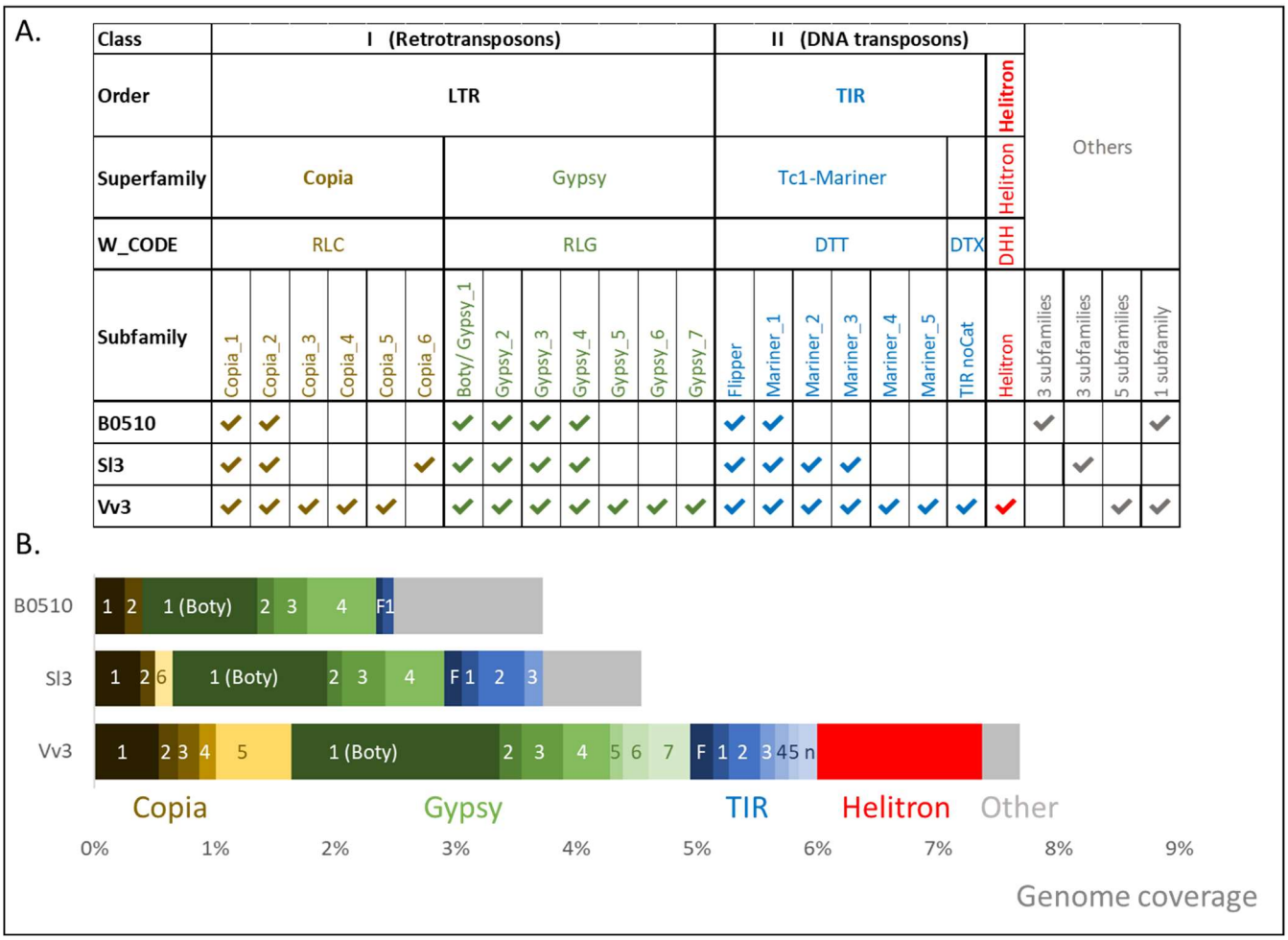


Figure 2: Transposable Elements (TEs) in the genomes of the *B. cinerea* strains B05.10, Sl3 and Vv3. **A**. Subfamilies of TEs identified in the Sl3 and Vv3 strains were classified according to Wicker et al., (2007), as previously done for the strain B05.10 (Porquier et al., 2016). In the present study, all *Boty* consensus were merged into a single subfamily. **B.** Total genome coverage was detailed for the different subfamilies. For more details, see Sup. Table 3.

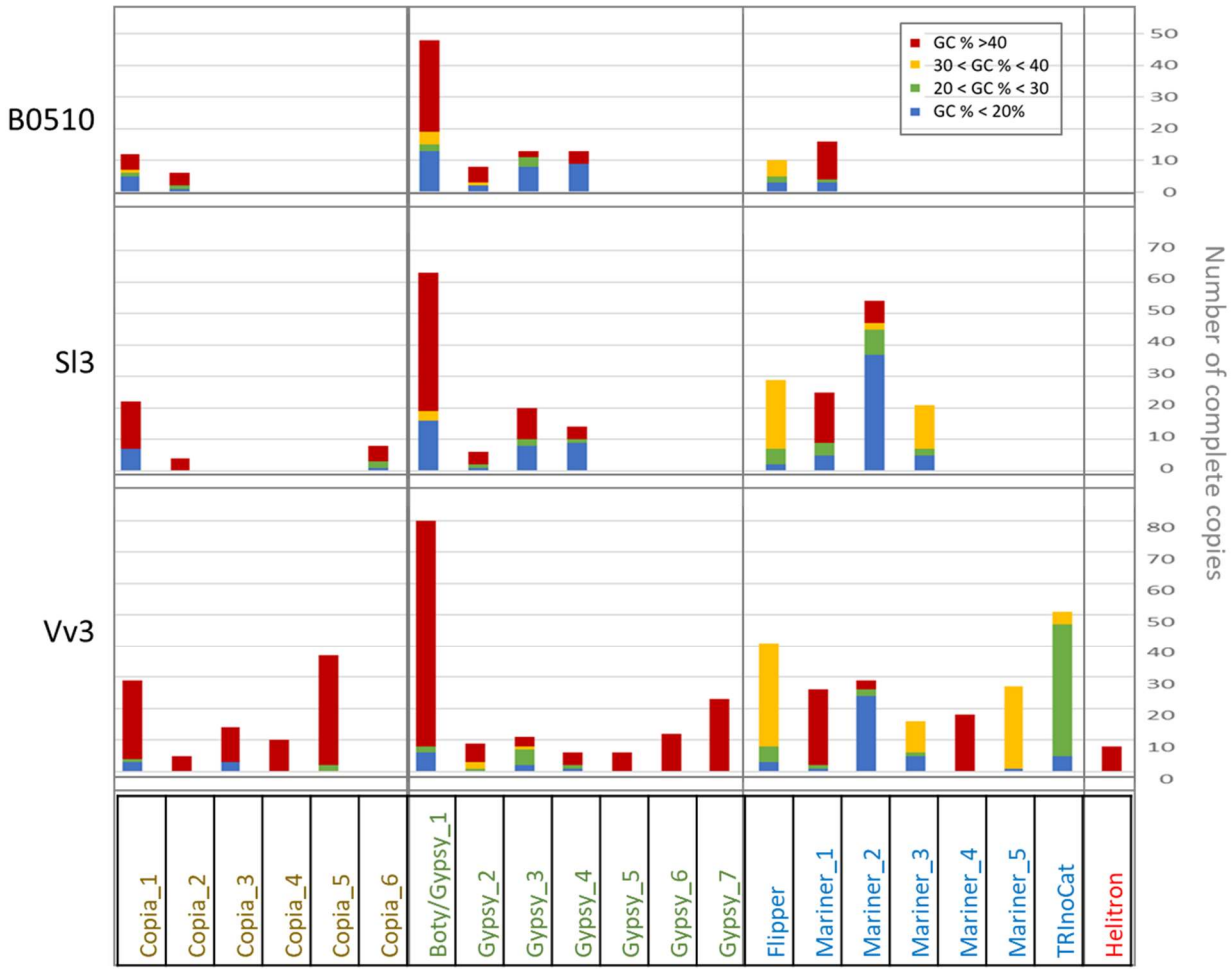


Figure 3: GC content of complete copies of the main Transposable Elements (TEs) in the genomes of *B. cinerea* strains B05.10, Sl3 and Vv3. The complete copies of each subfamily of TEs were classified according to their GC content. See Sup. Fig. S11 for more details.

Table 2: Small RNAs produced by the Transposable Elements (TEs) identified in the *B. cinerea* strains B05.10, SI3 and Vv3. Small RNA libraries were made from strains cultivated on grape juice medium except two that were done from strains cultivated on tomato juice medium (indicated by -T*). Reads from the 16 libraries were mapped on the TEs to identify the small RNA-producing retrotransposons. For small RNAs, ✔ indicates that at least 300 Reads Per Million (RPMs) from the library are mapping on the corresponding TE. Genomes in which the TEs were identified are reminded at the top of the table (✔, presence. -, absence). For more details, see Sup. Table S4.

Subfamily name			Copia_1	Copia_2	Copia_3	Copia_4	Copia_5	Copia_6	Boty_ Gypsy_1	Gypsy_2	Gypsy_3	Gypsy_4	Gypsy_5	Gypsy_6	Gypsy_7
Genomes in which the TE was identified		B0510	✔	✔	-	-	-	-	✔	✔	✔	✔	-	-	-
		SI3	✔	✔	-	-	-	✔	✔	✔	✔	✔	-	-	-
		Vv3	✔	✔	✔	✔	✔	-	✔	✔	✔	✔	✔	✔	✔
Libraries of small RNAs		B0510							✔		✔	✔			
	T population	SI1							✔		✔				
		SI2							✔		✔	✔			
		SI3						✔	✔		✔	✔			
		SI3-T*						✔	✔		✔	✔			
		SI6	✔					✔	✔		✔	✔			
		SI11							✔			✔			
		SI13							✔		✔	✔			
	G1 population	Vv1				✔			✔		✔	✔		✔	✔
		Vv3				✔			✔		✔	✔		✔	✔
		Vv3-T*				✔			✔		✔	✔		✔	✔
		Vv5				✔			✔		✔	✔		✔	✔
		Vv14				✔			✔		✔	✔		✔	✔
	G2 population	Vv8							✔			✔			
		Vv9													
		Vv15							✔			✔			