A secreted protease-like protein in *Zymoseptoria tritici* is responsible for avirulence on *Stb9* resistance gene in wheat

Reda AMEZROU[1], Colette AUDÉON[1], Jérôme COMPAIN[2], Sandrine GÉLISSE[1], Aurélie DUCASSE[1], Cyrille SAINTENAC[3], Nicolas LAPALU[1,2], Simon ORFORD[4], Daniel CROLL[5], Joëlle AMSELEM[2], Sabine FILLINGER[1], Thierry C MARCEL[1*]

[1] Université Paris-Saclay, INRAE, UR BIOGER, 91120 Palaiseau, France.
[2] Université Paris-Saclay, INRAE, UR URGI, 78026 Versailles, France.
[3] UCA, INRAE, GDEC, 63000 Clermont-Ferrand, France.
[4] Crop Genetics, John Innes Centre, Norwich NR4 7UH, UK.
[5] University of Neuchâtel, Laboratory of Evolutionary Genetics, 2000 Neuchâtel, Switzerland.

[*] Corresponding author: thierry.marcel@inrae.fr

Running title: The avirulence factor *AvrStb9* from *Zymoseptoria tritici*

Key words: Avirulence, Gene-for-gene interaction, Protease, GWAS, Wheat, *Zymoseptoria tritici*

Word count (main text): 6,396

Summary

- *Zymoseptoria tritici* is the fungal pathogen responsible for Septoria tritici blotch on wheat. Disease outcome in this pathosystem is partly determined by isolate-specific resistance, where wheat resistance genes recognize specific fungal factors triggering an immune response. Despite the large number of known wheat resistance genes, fungal molecular determinants involved in such cultivar-specific resistance remain largely unknown.
- We identified the avirulence factor *AvrStb9* using association mapping and functional validation approaches. Pathotyping *AvrStb9* transgenic strains on *Stb9* cultivars, near isogenic lines and wheat mapping populations, showed that *AvrStb9* interacts with *Stb9* resistance gene, triggering a strong immune response.
- *AvrStb9* encodes an unusually large avirulence gene with a predicted secretion signal and a protease domain. It belongs to a S41 protease family conserved across different filamentous fungi in the Ascomycota class and may constitute a core effector. *AvrStb9* is also conserved among a global *Z. tritici* population and carries multiple amino acid substitutions caused by strong positive diversifying selection.

- These results demonstrate the contribution of an 'atypical' conserved effector protein to fungal avirulence and the role of sequence diversification in the escape of host recognition, adding to our understanding of host-pathogen interactions and the evolutionary processes underlying pathogen adaptation.

Introduction

Filamentous fungal pathogens are responsible for the most damaging crop diseases (Savary *et al.*, 2019). Their ability to cause disease is facilitated by fungal proteins called effectors, which are delivered into the plant apoplast or cytoplasm during the infection process (Lo Presti *et al.*, 2015; Rocafort et *al.*, 2020). Effectors play critical role in the infection by manipulating plant immune system and other cellular functions. To counteract fungal infections, plants have evolved efficient defense systems relying on the recognition of different fungal factors (Jones & Dangl, 2006). The first type of defense involves plant receptors recognizing generic fungal molecules, called pathogen-associated molecular patterns (PAMPs). This PAMP-triggered immunity (PTI) provides basal defense against non-adapted pathogens. The second type of defense is activated with the recognition of fungal effectors by plant receptors, defined as resistance (*R*) genes. Effector-triggered immunity (ETI) induces a specific and strong immune response thought to limit the infection (Jones & Dangl, 2006). This interaction typically follows the gene-for-gene model, where host *R* gene products recognize pathogen effectors, which are then called avirulence (*Avr*) genes (Flor, 1956). However, ETI exerts a strong evolutionary pressure on pathogen populations leading to the selection of virulent isolates overcoming host immune response (Lo Presti et al., 2015). Most of these virulent isolates displayed mutations in the avirulence gene leading to variant proteins not recognized by the corresponding R protein (Sánchez-Vallet et *al.*, 2018). These observations emphasize the importance of pathogen *Avr* genes surveillance and the adequate spatiotemporal deployment of *R* genes in the field.

A common assumption is that most *Avr* genes encode small secreted proteins (SSP), with no predicted functional domain and no sequence homology in other species (Stergiopoulos & de Wit, 2009; Franseschetti et *al.*, 2017). It is often overlooked that *Avr* genes may also code for proteins not strictly falling into this category. For instance, some *Avr* genes from *Cladosporium fulvum* (*Avr2* and *Avr4*), *Magnaporthe oryzae* (*Avr-Pita*) or *Melampsora lini* (*AvrM14*) encode proteins that share sequence similarities with known enzymes or enzyme inhibitors (reviewed in Franseschetti et *al.*, 2017). However, only few such 'atypical' avirulence factors have been described in literature so far and their identification will broaden our knowledge on the complex molecular cross-talk between plants and pathogens.

Zymoseptoria tritici is an ascomycete fungal pathogen that causes the major wheat foliar disease Septoria tritici blotch (STB). *Z. tritici* is apoplastic with a biphasic infection process (Petit-Houdenot et *al.*, 2021). After penetration into wheat leaves through stomata, the fungus colonizes the apoplast for 8-11 days while the host remains symptomless (Battache et *al.*, 2022). This long asymptomatic phase is

72 followed by a rapid onset of necrotic symptoms, associated with the development of fungal asexual
73 fruiting bodies called pycnidia, which contain the pycnidiospores that further spread the disease
74 (Sánchez-Vallet *et al.*, 2015). STB is currently the economically most important wheat disease in Europe
75 and is mainly controlled using fungicides and genetic resistances (Fones & Gurr, 2015; Petit-Houdenot
76 et *al.*, 2021). However, this disease remains a problem due to the ability of fungal populations to
77 overcome control measures (*i.e.*, fungicide resistance, host resistance breakdown) and adapt to climate
78 changes (Fones & Gurr, 2015). Despite its economic importance, the molecular basis of *Z. tritici*
79 avirulence remains largely unknown. Only two avirulence genes have been characterized so far,
80 *AvrStb6*, recognized by the resistance protein *Stb6* leading to a strong immune response; and *Avr3D1*
81 which is thought to trigger quantitative resistance in cultivars carrying *Stb7* resistance gene (Zhong *et*
82 *al.*, 2017; Kema *et al.*, 2018; Meile *et al.*, 2018). Both genes encode SSPs located in highly plastic
83 genomic regions rich in transposable elements thought to favor their rapid diversification (Zhong *et al.*,
84 2017; Meile *et al.*, 2018). In addition to *Stb6* and *Stb7*, 21 major isolate-specific STB resistance genes
85 have been mapped in wheat, but their corresponding avirulence genes remain unknown (Brown *et al.*,
86 2015; Yang *et al.*, 2018; Langlands-Perry *et al.*, 2022).
87 Here, we aimed to broaden our understanding of avirulence in fungal pathogens, the mechanisms
88 underlying gene-for-gene interactions in the *Z. tritici*–wheat pathosystem and the evolutionary processes
89 related to the escape of host recognition. We first showed that a large protein encoding a predicted
90 secreted protease, *AvrStb9*, is responsible for avirulence in a subset of fungal strains. We demonstrated
91 that recognition *AvrStb9* by hosts carrying *Stb9* resistance gene triggers a strong immune response
92 following the gene-for-gene model. We next showed that *AvrStb9* belongs to a family of proteins with
93 a predicted protease S41 domain conserved across diverse ascomycetes lineages and that mutations in
94 *AvrStb9*, including some in the predicted protease domain, led to variants escaping *Stb9*-mediated
95 recognition.
96

97 ## Materials and Methods

98 ### Fungal material

99 A total of 103 *Z. tritici* strains were used for GWAS analysis. The strains derive from a French *Z. tritici*
100 collection collected in 2009 on cultivars 'Apache' (n = 44), 'Premio' (n = 42), 'Soissons' (n = 9),
101 'Caphorn' (n = 3), 'Alixan' (n = 1), 'Bermude' (n = 1), 'Dinosaur' (n = 1) and 'Garcia' (n = 1); one
102 isolate was collected in the UK on cultivar 'Humbert'. The two isolates used for genetic mapping of the
103 resistance and fungal transformation are IPO323 and IPO09593. IPO323 is a Dutch isolate collected on
104 cultivar 'Arminda' in 1981 and used as the reference strain for *Z. tritici* (Goodwin *et al.*, 2011).
105 IPO09593 is part of the 103 strains and has been isolated from 'Apache' in 2009. IPO323 is virulent on
106 cultivars carrying *Stb9* resistance gene while IPO09593 is avirulent on *Stb9*-cultivars. The three
107 additional isolates used to construct deletion and allele swap transformants, namely IPO09359

108 ('Apache', 2009), IPO09139 ('Apache', 2009) and SYN32 (unknown cultivar, 2006), are avirulent on
109 *Stb9*-cultivars.

110 **Plant material**

111 We performed quantitative trait loci (QTL) mapping in a recombinant doubled haploid (DH) population
112 comprised of 65 lines, developed from the F_1 progeny of the cultivars 'Beaver' and 'Soissons' (Singh *et*
113 *al.*, 2009). A panel of 220 modern French bread wheat (BW) varieties was used for GWAS analysis;
114 details on the BW panel used here can be found in Touzy et *al.* (2019). We also characterized *Z. tritici*
115 wild-type and mutant strains on near isogenic lines (NILs) carrying or not the *Stb9* resistance gene in
116 the genetic background of cultivar 'Courtot', developed at INRAE GDEC (Clermont-Ferrand, France).
117 The 'Courtot' NILs were developed following three backcrosses starting with F_1 'Courtot' x 'Chinese
118 spring' (carrying a susceptible allele of *Stb9* gene). To keep the *Stb9* susceptible allele in the progenies,
119 each plant was genotyped with single sequence repeat markers wmc317 and barc129 and phenotyped
120 with the *Stb9* avirulent isolate IPO89011. A BC_3F_1 plant heterozygous at the *Stb9* locus was self-
121 fertilized. BC_3F_2 plants homozygous either for the susceptible *Stb9* allele ('Courtot' – Stb9) or for the
122 resistant allele ('Courtot' + Stb9) were selfed and constitute the 'Courtot' NILs.

123 **Pathogenicity assays**

124 Pathogenicity assays were performed following the procedure described by Langlands-Perry *et al.*
125 (2022). Briefly, 16 days old plants were marked to delimit a 75 mm segment on the first true leaf. The
126 segment was inoculated with a paintbrush dipped into a 10^6 spores.mL^{-1} inoculum solution with one
127 drop of Tween®20 added per 15 mL. After inoculation, the plants were watered and covered with
128 transparent polyethylene bags for 72 hours to initiate the infection. The plants were kept in a climate
129 chamber with a 16 h photoperiod (300 µmol.m^{-2}.s^{-1}), 80-90% relative humidity and 18°C night/22°C
130 day. Phenotyping of the fungal population was performed by visually assessing strain virulence on
131 the cultivar 'Soissons' at 21 days post inoculation (dpi). Visual assessments were made of percent of
132 leaf area covered by necrosis (PLACN) and by pycnidia (PLACP) within the inoculated segment of each
133 leaf. Each of the 103 strains was tested in three replicates with three leaves per replicate. Phenotyping
134 of *AvrStb9* mutant strains was performed following the same procedure.
135 For the host mapping populations, the Beaver/Soissons DH lines were inoculated with the strain
136 IPO09593, and the 220 wheat varieties were inoculated with the strains IPO09593 and IPO323. Here,
137 each interaction was tested in two replicates with three leaves per replicate and visual assessments of
138 PLACN and PLACP were recorded at 14, 20 and 26 dpi. We calculated the area under disease progress
139 curve for necrosis (AUDPCn) and sporulation (AUDPCs) with the formula:

140 $$\mathrm{AUDPC} = \frac{\sum[(t_{i+1} - t_i) \times (y_i + y_{i+1})]}{2}$$

With, $t_{i+1}-t_i$: number of days between two observations; y_i: PLACN or PLACP at day i; y_{i+1}: PLACN or PLACP at day i+1. AUDPC were calculated relative to their maximum theoretical value, RAUDPCn and RAUDPCs respectively. For all assays, disease severity data were averaged between replicates and used as the input phenotype in GWAS and QTL analyses.

Fungal whole-genome sequencing and variant calling

Fungal spores were obtained after 7 days growth in liquid cultures of yeast extract-peptone-dextrose Broth (Difco™ YPD Broth), and then freezed-dried during 24 h. High molecular weight DNA was extracted using a phenol-chloroform based procedure. Illumina paired-end libraries were prepared for each sample and DNA fragments sequenced with 2x100 bp reads on a HiSeq-2000 sequencing system at Genewiz, Inc. (formerly Beckman Coulter Genomics). Depth of sequence coverage was variable depending on the isolate, ranging from 16 to 303 genome equivalents with an average of 61 genome equivalents. For each strain, reads were trimmed using trimmomatic v0.32 (Bolger *et al.*, 2014) and then mapped on the reference genome IPO323 (Goodwin *et al.*, 2011) using the mem algorithm from BWA v0.7.7 (Li & Durbin, 2009) with default options. Samtools v0.1.19 and Picard tools v1.106 (Li *et al.*, 2009) were used to filter reads (optic duplicates, secondary alignments and reads with a mapping quality < 30) and keep only pairs in which both reads met quality checks. SNP calling was done with Freebayes v0.9 with the option --report-monomorphic (Garrison & Marth, 2012). All positions detected as low-complexity regions or transposable elements (TE) were excluded using RepeatMasker 4.0 with default options and REPET (Flutre *et al.*, 2011), respectively. Quality filters were applied on the gVCF using in-house scripts. Depth of coverage (DP) calculated at each position follows a normal distribution around the depth main peak; a DP filter was set at twice the standard deviation from the mean. All the gVCF were stored into a matrix reduced to keep only the positions where at least one strain has a detected SNP in a non-filtered position. The SNP matrix was parsed in .tped format, converted to .ped/.map and filtered for < 0.5 missing data and > 0.1 minor allele frequency (MAF) using PLINK v.1.9 (Chang *et al.*, 2015).

GWAS analyses

For Genome-Wide Association Analysis (GWAS), we used the linear mixed model implemented in the R package 'GAPIT' (Lipka *et al.*, 2012). To account for population structure and reduce spurious associations, a kinship matrix using the VanRaden method and the first three principal components computed from SNP data were included a covariate in the model. We used a Bonferroni threshold of $p < 2.67e\text{-}7$ (α=0.05) to declare significant SNPs associated with phenotypic variations. Linkage Disequilibrium (LD) measures (r^2) for the significant SNPs were obtained with PLINK v.1.9 using the --r2 command and LD blocks were visualized using the R package 'LDheatmap'. SNPs with an $r^2 \geq 0.5$ were considered in strong LD.

176 **Genetic mapping of *Stb9***

177 A genetic linkage map was available for the Beaver/Soissons DH population and contains 312
178 polymorphic loci: 121 SSR, 49 AFLP, 135 DArT markers and 7 major genes (Singh *et al.*, 2009). The
179 map covers 1,526 cM with an average density of one marker per 12.6 cM. QTL mapping was conducted
180 using RAUDPCs and RAUDPCn phenotypes obtained from the isolate IPO09593 (*AvrStb9*) with R/qtl
181 software v.1.48-1 (Broman *et al.*, 2003). Linkage analysis was performed for each trait with an initial
182 Simple Interval Mapping (SIM), followed by a Composite Interval Mapping (CIM). For SIM, 1000
183 genome-wide permutations were used to calculate the significant logarithm of odds (LOD) threshold.
184 Only QTLs that showed *p*-values less than 0.05 were considered significant. Association mapping in the
185 BW panel was done following the exact same procedure as in the fungal population (see GWAS
186 analyses). The 220 BW varieties have been previously genotyped on an Affymetrix Axiom 410K single
187 nucleotide polymorphism (SNP) array (Danguy des Déserts *et al.*, 2021). Procedure of allele calling,
188 filtering and quality check of genotype data are described in Balfourier *et al.*, 2022. For GWAS analysis,
189 we excluded markers with less than 5% minor allele frequency and more than 10% missing data and
190 kept a set of 186,961 SNPs that passed quality control criteria and could be anchored on the IWGSC
191 annotation v2.1 (Zhu *et al.*, 2021). We used filtered SNP data and phenotypes (RAUDPCs and
192 RAUDPCn) from the isolates IPO323 (*avrStb9*) and IPO09593 (*AvrStb9*) to perform marker-trait
193 associations.

194 **Gene annotation and expression analysis**

195 Predicted genes in the candidate region were retrieved from the *Z. tritici* reference genome annotation
196 (Grandaubert *et al.*, 2015). We used RNA-seq data (Rudd *et al.*, 2015) mapping onto the reference
197 genome IPO323 to manually curate *Zt_1_693* (*AvrStb9)* gene model. Canonical and non-canonical
198 splicing sites were confirmed by *de-novo* transcriptome assembly obtained with DRAP (Cabau *et al.*,
199 2017). The final gene model differed from those from previous annotations due to errors in predicting
200 correctly introns and includes two missing exons. Candidate genes were screened for the presence of a
201 secretion signal using SignalP 3.0 server (Dyrløv Bendtsen *et al.*, 2004). Blastp was used to search for
202 *AvrStb9* orthologous and paralogous proteins against the reference strain IPO323 and closely-related
203 *Zymoseptoria* species. Protein domains were searched using Interproscan and structure homology
204 prediction was performed in the HHpred server; <https://toolkit.tuebingen.mpg.de/tools/hhpred>. For
205 gene expression analysis, RNA-seq data samples from *in vitro* and *in planta* (1, 4, 9, 14 and 21 dpi)
206 conditions were obtained from Rudd *et al.* (2015), and gene expression quantified as fragments per
207 kilobase of exon per million fragments mapped (FPKM) using Cuffdiff v.2.2.1 (Trapnell *et al.*, 2012).
208 *AvrStb9* DNA sequences are deposited in NCBI Genbank (accession numbers OP376566 and
209 OP376567).

210 **Construction of disruption plasmids**

211 For *AvrStb9* disruption, the hygromycin resistance marker (*hph*) was cloned between the 1,000 bp up-
212 and downstream flanking regions of the ORF, respectively. The following DNA fragments were
213 generated and cloned into pCAMBIA0380 plasmids (GenBank: AF234290.1). The up- and downstream
214 flanks were obtained by PCR on gDNA of IPO323 with primer pairs 1&2 and 5&6 (Table S2)
215 respectively using Q5 Hot Start High fidelity DNA polymerase (NEB, Evry, France). The *hph* was
216 amplified from pCAMBIA0380 with the primers 3 and 4 (Table S2). pCAMBIA0380 was digested by
217 *Eco*RI and *Xho*I and used for the cloning of three PCR fragments with Gibson Assembly cloning kit
218 (NEB, Evry, France) according to the supplier's indications. The ligation product was introduced into
219 NEB 5α competent *Escherichia coli* cells (NEB, Evry, France), transformants selected on kanamycin
220 (50 µg.mL^{-1}) containing LB-broth. Colonies were screened by PCR with the primer-pairs used for
221 cloning. Four positive colonies were chosen for plasmid extraction (Macherey & Nagel, Düren,
222 Germany) and restriction analysis. The construct was then introduced by electroporation into the
223 *Agrobacterium tumefaciens* Agl1 strain. Transformants were selected and isolated on YEB (5 g. L^{-1} beef
224 extract, 1 g. L^{-1} yeast extract, 5 g. L^{-1} peptone, 5 g. L^{-1} sucrose, 0.5 g. L^{-1} MgCl2, 1.5% agar) with
225 kanamycin (50 µg. mL^{-1}).

226 **Gene replacement constructs**

227 To swap *AvrStb9* alleles between virulent (IPO323) and avirulent (IPO09593) strains, the following
228 replacement cassettes were constructed. The respective *AvrStb9* alleles were amplified with primers 11
229 and 12 (Table S2) from IPO323 or IPO09593 gDNA. These fragments cover the ORF with 1 kb
230 upstream of the START codon and 360 bp downstream of the STOP codon. The *hph* (amplified with
231 primers 9 and 10 from pCAMBIA0380 plasmid DNA) and the 1 kb upstream flank of *AvrStb9* gene
232 (amplified with primers 7 and 8) were ligated to *AvrStb9* alleles into the *EcoRI-XhoI* digested
233 pCAMBIA0380, with Gibson Assembly cloning kit (NEB, Evry, France). Plasmids were introduced
234 into NEB 5α competent *E. coli* cells (NEB, Evry, France) and transformants selected on kanamycine
235 (50 µg.mL^{-1}) containing LB-broth. Colonies were screened by PCR with the primer-pairs used for
236 cloning. Four positive colonies were chosen for plasmid extraction (Macherey & Nagel, Düren,
237 Germany) and sequence analysis. These constructs named respectively pR323 and pR9593 were then
238 introduced by electroporation into the *Agrobacterium tumefaciens* Agl1 strain. Transformants were
239 selected and isolated on YEB with kanamycin (50 µg.mL^{-1}).

240 **Generation of *Z. tritici* transformants**

241 The transformation procedure was adapted from Zwiers & de Waard (2001) and was performed as
242 previously described (Omrane *et al.*, 2015). For *AvrStb9* deletion, the following *Z. tritici* strains were
243 transformed with the disruption plasmid: IPO323, IPO09593, IPO09359, IPO09139 and SYN32.
244 Transformants were selected and isolated twice on hygromycin (100 µg. mL^{-1}) containing medium. The
245 replacement of the *AvrStb9* ORF by the *hph* gene was checked by PCR on gDNA with the primer pair

246 13 and 14. The gene deletion is expected to give a 1.6 kb fragment instead of the native 2.4 kb amplicon.
247 The total number of isolated and analyzed transformants were: IPO323 (n=144), IPO09593 (n=23),
248 IPO09359 (n=26), IPO09139 (n=46), SYN32 (n=38). One single transformant produced the expected
249 PCR profile, namely SYN32_ΔT12.

250 For *AvrStb9* allele swaps, IPO323 was transformed with plasmids pR323 and pR9593. Transformants
251 selected and isolated twice on hygromycin containing medium (100 µg. mL^{-1}). The replacement of
252 *AvrStb9* alleles in the transformants was checked by PCR using primers 15 and 16, leading to a single
253 1.6 kb amplicon in replacement mutants and to an additional 238 bp amplicon in ectopic transformants.
254 Transformants with the expected single band size were PCR checked with primers 17 and 18 as these
255 primers allow to distinguish the deletion of the amino acid residues 66-71 in the IPO09593 background
256 from the IPO323 one, and subsequently submitted to sequence analysis with primers indicated in Table
257 S2.

258 **Phylogenetic analysis.** We used OrthoMCL algorithm to search for *AvrStb9* orthologous and
259 paralogous protein sequences using a cutoff E-value less than 1e-20 and filtering for low complexity
260 regions (Li *et al.*, 2003); <https://orthomcl.org/orthomcl/app/>. Of the 148 putative *AvrStb9*
261 orthologs/paralogs identified through OrthoMCL, we selected 122 by removing redundant information
262 using the following criteria: (*i*) when multiple identical sequences were present, one was arbitrarily
263 chosen and retained; and (*ii*) when multiple sequences from different strains of the same species were
264 present, an arbitrarily chosen strain (the reference strain, if known) was retained. A multiple sequence
265 alignment was then generated with Clustal Omega v1.2.4 (Sievers *et al.*, 2011) and alignment cleaning
266 was performed using trimAL with the --gappyout option (Capella-Gutiérrez *et al.*, 2009). We
267 constructed a maximum likelihood tree of selected *AvrStb9* orthologs/paralogs using RaxML with the
268 GAMMA JTT model and 100 bootstrap replicates (Stamatakis, 2014). Phylogenetic trees were
269 visualized using iTOL (Letunic & Bork, 2016); <https://itol.embl.de/>.

270 **Population genetics and selection analyses**

271 In addition to the fungal population used for genetic analyses, we extended our sequence diversity and
272 selection analyses to include a global *Z. tritici* population originating from four distinct geographic
273 locations, including the middle East (n=30, Israel), Australia (n=27, Wagga Wagga), Europe (n=33,
274 Switzerland) and North America (n=56, Oregon, USA). Procedures of *de novo* genome assemblies and
275 sequence extraction were previously described (Hartmann *et al.*, 2017). We used DnaSP v6 (Rozas *et*
276 *al.*, 2017) to calculate summary statistics of the population genetic parameters associated with *AvrStb9*.
277 Sliding window analyses of nucleotide diversity (π) and Tajima's D were conducted in a window length
278 of 200 bp and a step size of 20 bp. To test the selective pressure acting on *AvrStb9*, we estimated the
279 rates of non-synonymous (dN) and synonymous (dS) substitutions using CodeML from the PAML
280 package (Nielsen & Yang, 1998). A dN/dS (ω) ratio of 1 indicated neutrality, whereas $\omega < 1$ or > 1

281 suggested purifying or diversifying selection, respectively. However, the presence of intragenic
282 recombination could bias the ω rate testing. To account for these issues, we performed the analysis with
283 a non-recombining phylogenetic tree generated from the RDP4 software using RaxML with 100
284 bootstraps (Martin *et al.*, 2015).

285 **Results**

286 **Mapping of a candidate gene conferring avirulence on the cultivar 'Soissons'.**

287 'Soissons' is a French bread wheat variety, developed by Florimond Desprez and released in 1987,
288 which represented 10 to 40% of the French wheat acreage between 1991 and 2002. 'Soissons' is
289 moderately susceptible to STB (Suffert *et al.*, 2013) but it has been postulated to carry the resistance
290 gene *Stb9* based on screening with differentially virulent isolates (data not shown). Infection assays of
291 the fungal population on cultivar 'Soissons' showed that ~12% of the French isolates are avirulent (Fig.
292 S1). We conducted GWAS to identify the *Avr* gene interacting with the resistance in 'Soissons'.

293 A total of 103 *Z. tritici* strains were re-sequenced, generating 718,810 SNPs passing quality filters. Using
294 full SNP data, we investigated the population structure and kinship and found very little genetic
295 substructure and relatedness in the fungal population (Fig. S2a,b). For instance, the first three PCs
296 accounted for only 3.11% of the total genetic variation. Using a linear mixed model to associate genetic
297 variations with pycnidia (PLACP) and necrosis (PLACN) variations on 'Soissons', we identified a
298 cluster of 26 SNPs on chromosome 1 from 2,092,563 bp to 2,096,143 bp, significantly associated with
299 PLACN at the Bonferroni threshold (Fig. 1a). The same SNP cluster was also associated with variation
300 in PLACP but failed to reach the genome-wide significance (Fig. S3). We also detected a significant
301 association with PLACN variations (p=3.801e-8; position 2,451,010 bp) on a segment of chromosome
302 7. However, the region falls in a broad section of the chromosome without any *in planta* expressed genes
303 (Rudd *et al.*, 2015). The significant SNPs on chromosome 1 mapped to a region containing two predicted
304 genes, *Zt_1_692* and *Zt_1_693*. The lead SNP (p=7.31e-10; position 2,095,552 bp) mapped to the fourth
305 exon of *Zt_1_693*, but was in strong linkage disequilibrium with other variants spanned across the two
306 predicted genes (Fig. 1b). *Zt_1_692* is not expressed *in planta* so we were unable to use RNA-seq data
307 to evaluate the quality of the gene annotation and blastp searches yielded only weak hits. Without any
308 corroborating evidence for a biological role, we concluded that *Zt_1_692* is likely a pseudogene or an
309 annotation error. In contrast, *Zt_1_693* is highly expressed *in planta* (Fig. S4), with a peak expression
310 at 9 days post infection coinciding with the switch from asymptomatic to necrotrophic phase (Fig. 1d).
311 Further, allele analyses of the lead SNP clearly discriminated between avirulent and virulent phenotypes
312 (Fig. 1c). Taken together, we considered *Zt_1_693* as the best candidate gene to explain avirulence on
313 the wheat cultivar 'Soissons'.

314 ***Zt_1_693* is involved in a gene-for-gene relationship with *Stb9* resistance gene.**

315 To determine whether *Zt_1_693* is an avirulence factor recognized by an *R* gene in the cultivar
316 'Soissons', we sought to generate *Zt_1_693* deletion mutants (KO) in the virulent strain IPO323 and
317 four different avirulent strains by targeted gene replacement. We also switched alleles between avirulent
318 (IPO09395) and virulent (IPO323) strains using targeted gene replacement. However, we could not
319 obtain any deletion mutant neither in IPO323 (144 transformants tested), nor in the avirulent strains
320 IPO09593, IPO09359 and IPO09139 (n=95). Only a single deletion mutant was obtained in strain
321 SYN32 out of 32 transformants, producing small colonies and weak symptoms on the susceptible
322 cultivar 'Taichung-29' (Table S3). This suggests that *Zt_1_693* could be essential for *Z. tritici* growth,
323 and pathogenicity. Nevertheless, using allele swapping we were able to generate mutants expressing the
324 virulent and avirulent versions of *Zt_1_693* in IPO323 genetic background, and performed infection
325 assays on cultivars 'Soissons', 'Courtot' which carries the *Stb9* gene and a 'Courtot' near isogenic line
326 lacking *Stb9*.

327 IPO323 gene replacement transformants carrying the avirulent allele (n=6) from IPO09593 (referred to
328 as *Avrstb9*) were avirulent on 'Soissons' and 'Courtot', but not on 'Courtot' lacking *Stb9* (Fig. 2a).
329 Inversely, gene replacement transformants with the virulent allele (referred to as *avrstb9*) from IPO323
330 (n=2) resulted in virulence on all tested cultivars, with PLACP values at 26 dpi ranging from 32 to 100%
331 (Fig. 2a). Interestingly, Sanger sequencing of IPO323$_{Avrstb9}$ transformants revealed that four strains were
332 chimeric with the avirulent allele not being introduced in full length, retaining some polymorphisms of
333 the virulent haplotype between residues 276 and 628. All four chimeric strains were virulent on *Stb9*-
334 containing cultivars, highlighting the importance of these polymorphisms in virulence (Table 1; Table
335 S3; Fig. S6). The fact that all IPO323 mutants carrying the complete *AvrStb9* allele were unable to infect
336 *Stb9*-containing cultivars, while mutants expressing the *avrStb9* allele produced symptoms on all three
337 cultivars, suggest that *Zt_1_693* encodes the avirulence factor *AvrStb9*, which triggers complete
338 resistance in cultivars carrying the cognate resistance gene.

339 To further validate the *AvrStb9-Stb9* gene-for-gene relationship, we performed two independent genetic
340 analyses using a bi-parental mapping population and a panel of wheat varieties. The isolate-specific
341 resistance to IPO09593 (*AvrStb9*) was mapped by a QTL analysis on Beaver/Soissons recombinant
342 double haploid progenies. A single major resistance QTL was detected on the long arm of the wheat
343 chromosome 2B, likely corresponding to the *Stb9* locus (Chartrain *et al.*, 2009). This QTL had a log-
344 likelihood (LOD) of 10.77 and 15.33 for the RAUDPC of necrosis and sporulation, respectively (Fig.
345 2b). As expected, the resistance allele originated from 'Soissons' suggesting that the cultivar indeed
346 carries the *Stb9* resistance. We confirmed this interaction by performing a GWAS analysis on a wheat
347 panel phenotyped using the virulent and avirulent strains. Association mapping results revealed a strong
348 signal associated with RAUDPCs for the avirulent strain IPO09593 (eight SNPs above the Bonferroni-
349 corrected cutoff) in the telomeric region of chromosome 2B, but not with the virulent strain IPO323
350 (Fig. 2c). This confirms that IPO09593 triggers a strong resistance from varieties carrying a resistance

351 QTL in the vicinity of the *Stb9* gene. Additional details on *Stb9*-associated SNPs are provided in the
352 supplemental Table S1. Finally, based on the lead SNP (AX-89505650; position 808,036,267 bp on
353 chromosome 2B), phenotypic and allele analyses showed that ca. 8.2% of the French varieties likely
354 carry *Stb9*.

355 Altogether, our combined functional and genetic analyses demonstrate that *Zt_1_693* is the avirulence
356 gene *AvrStb9* and is involved in a gene-for-gene relationship with *Stb9* resistance in wheat.

357 **_AvrStb9_ encodes a secreted protein with a conserved protease domain shared among diverse**
358 **ascomycetes lineages.**

359 Unlike other known fungal effectors, *AvrStb9* encodes a large protein and shares sequence similarities
360 with different fungal species. The re-annotated *AvrStb9* gene model has five introns and encodes a 727
361 amino acid long protein with eight cysteine residues (Fig. 4a; 4b). This protein has a predicted signal
362 peptide sequence of 18 amino acids at its N-terminus and two serine-protease domains (C-terminal
363 processing peptidase S41 family; positions 70-726; Fig. 4b) according to InterProScan (Panther entry
364 PTHR-37049, e-value: 2.7 e-64). HHpred searches in protein structure databases identified the secreted
365 protease CPAF as the strongest hit (e-value of 1.5e-27), which is involved in pathogenesis of the human
366 pathogenic bacterium *Chlamydia trachomatis* (Huang et *al.,* 2008).

367 To investigate whether *AvrStb9* originated before or after *Z. tritici* speciation, we used blastp searches
368 against *Z. tritci* reference genome IPO323 and closely-related species from the *Zymoseptoria* family.
369 We identified *AvrStb9* and its paralogs from IPO323 and all its closest relatives *Z. pseudotritici, Z.*
370 *ardabiliae* and *Z. brevis* (Fig. S7a). Interestingly, most of these proteins had a predicted signal peptide,
371 and they all displayed a C-terminal processing peptidase S41 domain. Using IPO323 *in planta* gene
372 expression, we found that three out of the five paralogs (*Zt_4_489, Zt_10_226, Zt_11_236*) had an
373 expression profile comparable to *AvrStb9* (Fig. S7b). According to these expression patterns, these genes
374 could potentially play a role in infection.

375 Using OrthoMCL, we identified 122 putative *AvrStb9* orthologs/paralogs with an average identity of
376 33.6% and an average e-value of 8e-73 (Table S4). *AvrStb9* orthologs/paralogs are found in a broad
377 range of ascomycetes lineages belonging to Eurotiomycetes, Leotiomycetes, Sordariomycetes,
378 Orbiliomycetes and Dothideomycetes (Fig. 3; Table S4). Most of fungal species (48 out of 65 species)
379 identified from the orthogroup are pathogenic, indicating that the well conserved *AvrStb9*
380 orthologs/paralogs may be involved in host interaction processes. Interestingly, ~87% of these proteases
381 possess a predicted signal peptide.

382 **Diversification of _AvrStb9_ leads to evasion of host recognition**

383 Non-synonymous substitutions in avirulence gene sequences often mediate the evasion of host
384 recognition. Analysis of sequence polymorphisms between the avirulent and virulent strains revealed a

385 6 amino acid deletion and 12 non-synonymous substitutions (Table 1; Fig. 4b). We crossed this data
386 with the fungal population phenotypes and identified eight mutations potentially involved in host
387 evasion for which more than ~94% of strains carrying the mutation were virulent on *Stb9* cultivars
388 (Table 1). Five of these mutations fall into the virulence-domain cited above, namely A431, F473, S525,
389 D565, L628. All of these mutations except D565 are located in one of the two putative protease domains
390 (Fig. 4b), with F473, S525 and L628 being relatively conserved residues among proteins belonging to
391 the CPAF structural family (in 34-57% of the analyzed proteins respectively; data not shown).

392 In order to investigate *AvrStb9* diversity at a larger scale, we extended our analysis to include four global
393 *Z. tritici* populations. We found 44 different *AvrStb9* alleles in these populations, which translate to 38
394 different protein variants, mostly population-specific (Fig. 4c). Next, we analyzed *AvrStb9* nucleotide
395 diversity including ~500 bp of up- and downstream sequences from the start and stop codons,
396 respectively. We observed both a drop-in nucleotide diversity and Tajima's D in the *Avr* coding sequence
397 compared to the flanking regions (Fig. 4a). A negative Tajima's D is indicative of an excess of low
398 frequency polymorphisms, indicating population size expansion and/or selective sweeps.

399 The codon-based model analysis of *AvrStb9* coding sequence was used to investigate the selection
400 pressure exerted on the avirulence gene. Likelihood ratio tests (LRT) for different models indicated that
401 positive selection was highly probable, as selection models (M2 and M8) showed a significantly better
402 data fit compared to the neutral models M1 and M7 (Table S5). Selection analysis revealed 23 residues
403 (3.16% of the protein sequence) under positive diversifying selection (BEB probability > 0.95) in the
404 M2 model. Interestingly, S525F and L628Y, had the strongest positive selection probability (BEB
405 probability of = 1) among the five mutations potentially responsible of the virulence according to the
406 previous analyses (Table 1). We postulate that these adaptive mutations are under strong selective
407 pressure most probably to counteract recognition by the *Stb9* resistance gene.

408 **Discussion**

409 Gene-for-gene interactions in plant pathosystems are described when a host resistance gene recognizes
410 a pathogen's avirulence product that triggers a strong host immune response. Although 23 isolate-
411 specific *Stb* resistance genes have been mapped in wheat, only two have been cloned so far and their
412 corresponding fungal avirulence genes remain largely unknown. Here, we report the identification and
413 characterization of a novel avirulence gene, *AvrStb9*, which triggers complete resistance in wheat
414 cultivars carrying *Stb9* resistance gene. Additionally, we performed a high-resolution mapping of *Stb9*
415 locus and delivered tightly-linked SNP markers for its use in breeding.

416 *AvrStb9* is highly expressed during the asymptomatic phase of *Z. tritici* infection and drastically
417 downregulated upon the switch to the late necrotrophic phase. In a previous study, *AvrStb9* was among
418 the top 100 highly expressed genes during the transition from biotroph to necrotrophy stage (Palma-
419 Guerrero *et al.*, 2017), suggesting a function in host colonization during the asymptomatic phase or the

420 transition to the necrotrophic phase. Recognition of the *Avr* allele leads to full resistance in cultivars
421 carrying *Stb9* following the gene-for-gene model. A previous study mapped the *Stb9* resistance locus to
422 the long arm of chromosome 2B but the genomic interval remained very large (Chartrain *et al.*, 2009).
423 We used *AvrStb9* strain to trigger wheat resistance and perform a precise mapping of the resistance
424 locus, which revealed that *Stb9* resides in the telomeric region of wheat chromosome 2B in an interval
425 of 79.99-81.10Mb ($r^2 > 0.50$). Only two STB resistance genes have been cloned to date, *Stb6* and *Stb16q*,
426 both encoding receptor-like kinase proteins (Saintenac et *al.*, 2018; 2021), and our findings will help in
427 the cloning of *Stb9* resistance gene.

428 *AvrStb9* is an atypical fungal avirulence gene, since it encodes a relatively large secreted protein (727
429 aa). Up to now, avirulence genes from filamentous fungal pathogens encode small (< 300 aa) secreted
430 proteins rich in cysteine. Only three exceptions to this rule are the *Ace1* gene from *Magnaporthe grisea*
431 (Collemare *et al.*, 2008) encoding a large polyketide synthase (PKS) fused to a nonribosomal peptide
432 synthetase (NRPS) likely involved in secondary metabolism, and two *AvrMla* genes from *Blumeria*
433 *graminis hordei* which encode large proteins (Bourras *et al.*, 2016). Furthermore, in contrast to what has
434 been shown for most avirulence genes, *AvrStb9* has a predicted functional domain corresponding to the
435 family of C-terminal processing S41 proteases and show sequence similarities with proteins from
436 diverse fungal species. In plant pathogenic bacteria, some type III secreted effectors injected in host
437 plant cells are similar to catalytic enzymes involved in a variety of reactions such as E3 ligation or
438 proteolysis (Hann *et al.*, 2010). For instance, *Pseudomonas syringae* secretes the cysteine protease
439 *HopN1* to suppress the host innate response and promote virulence (López-Solanilla *et al.*, 2004).

440 The importance of proteases in fungal pathogenicity has previously been reported (Jashni *et al.*, 2015b;
441 Chandrasekaran *et al.*, 2016). *Fusarium oxysporum* f. sp. *lycopersicum* secretes a serine protease and a
442 metalloprotease that cleave host chitinases to prevent the degradation of fungal cell walls and facilitate
443 host colonization (Jashni *et al.*, 2015a). The maize anthracnose fungus *Colletotrichum graminicola*
444 secretes a highly conserved metalloprotease effector that enhances virulence (Sanz-Martín *et al.*, 2016).
445 Similarly, the rice blast fungus *Magnaporthe oryaze* secretes during infection *Avr-Pita*, an effector
446 protein with a predicted zinc metalloproteases domain that belongs to a family of proteases that include
447 other avirulence genes recognized by the resistance gene *Pi-ta* (Orbach *et al.*, 2000; Khang *et al.*, 2008).
448 Our protein analyses (OrthoMCL, InterProScan, HHpred) indicated that AvrStb9 has sequence and
449 structural features typical of a protease. Protein structure-based homology searches identified the
450 secreted protease CPAF from the human pathogenic bacterium *C. trachomatis* as closest homolog.
451 CPAF degrades host transcription factors and therefore directly helps *Chlamydia* evading the host
452 defenses. We postulate that AvrStb9 could play a similar role in proteolytic degradation of host protein
453 substrates. However, the enzymatic activity of AvrStb9 is yet to be investigated, as our analysis relied
454 primarily on sequence and structure comparisons. Further, the fact that we were hardly able to produce

455 *AvrStb9*-disrupted mutants indicates that the gene may play an important role in *Z. tritici* survival, and
456 that the proteolytic activity likely extends to include substrates essential for fungal growth.

457 The broad conservation of AvrStb9 in *Z. tritici* world population, its sister species and some other fungal
458 species suggested that this protein is a core effector shared among these lineages. Irieda et *al.* (2019)
459 reported the conservation of *NIS1* effector that was initially detected from two *Colletotrichum* spp. in a
460 wide range of Ascomycota and Basidiomycota lineages. *NIS1* targets plant immune kinases to suppress
461 PAMP-triggered immunity (Irieda *et al.*, 2019). We hypothesize that S41 proteases may have been
462 anciently acquired by fungi, and that the expansion of particular serine protease families is associated
463 with the evolution of fungal lifestyle. Indeed, S41 serine proteases are suspected to have a role in
464 programmed cell death and pathogenic life style adaptation in fungi (Muszewska *et al.*, 2017). We also
465 identified multiple paralogs from multiple fungal species, likely resulting from gene duplication events,
466 highlighting either the redundancy of these genes or their evolution to interact with different substrates.
467 The functional redundancy of virulence genes has been a technical challenge to study their biological
468 functions, as disruption mutants often lead to the conclusion that these genes are not essential for
469 pathogenicity (Plummer *et al.*, 2004; Have *et al.*, 2010). These redundancies may also increase the
470 organism's genetic robustness against null mutations or act in compensation of a gene loss (Gu *et al.*,
471 2003). In *Z. tritici* for instance, the different *AvrStb9* paralogs had a secretion signal and show an
472 expression pattern consistent with a role as effectors. Taken altogether, we hypothesize that *AvrStb9* and
473 related S41 proteases constitute a set of core effectors gained early in the evolution of fungal-plant
474 interactions.

475 Recognition of effectors by host resistance exerts an evolutionary pressure that favors sequence
476 diversification, effector deletion or acquisition to counteract the immune response. We analyzed
477 AvrStb9 sequence diversity and found a broad conservation of its protein residues, unlike the other
478 known *Z. tritici* avirulence genes, *AvrStb6* and *Avr3D1* (Brunner & McDonald, 2018; Meile *et al.*, 2018).
479 Nevertheless, we identified few substitutions undergoing strong positive diversifying selection.
480 Additional data from chimeric AvrStb9 transformants, from *Z. tritici* populations and from sequence
481 conservation among the CPAF family indicate S525 and L628 as probably the most important residues
482 involved in virulence, although combinations of mutations may be necessary. Based on the structural
483 model for CPAF function, one may speculate that a functional protease is required to escape host
484 recognition.

485 This is corroborated by the fact that, we did not observe in any of the global strains a loss of *AvrStb9* or
486 other deleterious mutations. Furthermore, *AvrStb9* deletion proved to be nearly impossible (only one
487 poorly growing transformant had the gene deletion out of the 271 tested). We therefore suggest that
488 *AvrStb9* plays an essential role in the biology of *Z. tritici*. It is tempting to speculate that the functional
489 AvrStb9 protein actively degrades (through proteolysis) both fungal protein(s) and host defense

490 protein(s). Additional functional studies, including mutating the S41 protease active sites in virulent and
491 avirulent alleles will shed light on the role of the protease activity of AvrStb9 in its diverse biological
492 activities (fungal growth, wheat infection, recognition by *Stb9*). But also, the function of the wheat
493 resistance factor Stb9 is intriguing. Does it work as classical receptor protein, recognizing the avirulent
494 AvrStb9 protein? Does it inhibit AvrStb9 protease activity or, on the contrary, may the Stb9 resistance
495 protein be the target of AvrStb9 protease?

496 In conclusion, we identified a new major avirulence factor of *Z. tritici* that triggers a strong immune
497 response in hosts carrying the corresponding resistance gene. Unlike what has been described for most
498 avirulence genes, *AvrStb9* encodes a large protein with a predicted protease domain and likely belongs
499 to a family of conserved fungal S41 proteases. This is an important step in deciphering molecular
500 mechanisms underlying wheat – *Z. tritici* interactions, and shed light on the role of 'atypical' effectors
501 (*i.e.*, non-SSPs) in fungal avirulence. It opens new research avenues to broaden our knowledge on host-
502 pathogen interactions, in particular the characterization of enzyme activity of effectors, which will
503 greatly help our understanding of the biology of protease effectors, commonly found in fungal
504 pathogens.

505 **Data Availability**

506 Genome sequencing data for the 103 French *Z. tritici* isolates is available at NCBI BioProject
507 PRJNA777581 (accessions SRR16762555 to SRR16762657). *AvrStb9* gene sequences are available at
508 NCBI Genbank (accession numbers OP376566 and OP376567).

509 **Acknowledgements**

510 We thank David Gouache, Henriette Goyeau, Gert H.J. Kema, Marc-Henri Lebrun, Frédéric Suffert and
511 Anne-Sophie Walker for their contribution in establishing the fungal population; Johann Confais for his
512 help in the virulence assays; Robert King and Jason Rudd for providing IPO323 RNA-seq data; Florence
513 Cambon, the VégéPôle and GENTYANE platforms for the development of the Courtot NILs; Marc-
514 Henri Lebrun for critical reading of the manuscript and valuable feedback; and Sophie Bouchet and
515 Fanny Hartmann for assistance with the initial GWAS analyses. This work was supported by the French
516 National Research Agency (program BIOADAPT, grant no. ANR-12-ADAP-0009-04), the French Fund
517 to support Plant Breeding (FSOV 2008 H) and by BASF France SAS (V.I.S.A. project). INRAE
518 BIOGER benefits from the support of Saclay Plant Sciences-SPS (ANR-17-EUR-0007).

519 **Author contributions**

520 T.C.M and S.F. conceived and designed the study. A.D. and S.G. performed virulence assays, C.A.
521 performed the cloning and genetic transformations. S.O. provided DH seeds and linkage mapping data.
522 C.S. provided the 'Courtot' NILs. J.C., N.L., J.A. and R.A performed bioinformatics analysis. D.C.

523 provided sequence data from the global *Z. tritici* populations. R.A, T.C.M and S.F. analyzed the data
524 and wrote the manuscript. All authors have reviewed the manuscript.

525 **References**

526 **Balfourier F, Bouchet S, Robert S, De Oliveira R, Rimbert H, Kitt J, Choulet F, International**
527 **Wheat Genome Sequencing Consortium, BreedWheat Consortium, Paux E. 2022**. Worldwide
528 phylogeography and history of wheat genetic diversity. *Science Advances* **5**: eaav0536.

529 **Battache M, Lebrun MH, Sakai K, Soudière O, Cambon F, Langin T, Saintenac C. 2022.** Blocked
530 at the Stomatal Gate, a Key Step of Wheat Stb16q-Mediated Resistance to *Zymoseptoria tritici*.
531 *Frontiers in Plant Science* **13:** 921074.

532 **Bolger AM, Lohse M, Usadel B**. **2014**. Trimmomatic: a flexible trimmer for Illumina sequence data.
533 *Bioinformatics* **30**: 2114–2120.

534 **Bourras S, McNally KE, Müller MC, Wicker T, Keller B**. **2016**. Avirulence genes in cereal powdery
535 mildews: the gene-for-gene hypothesis 2.0. *Frontiers in Plant Science* **7**: 241.

536 **Broman KW, Wu H, Sen Ś, Churchill GA**. **2003**. R/qtl: QTL mapping in experimental crosses.
537 *Bioinformatics* **19**: 889–890.

538 **Brown JKM, Chartrain L, Lasserre-Zuber P, Saintenac C**. **2015**. Genetics of resistance to
539 *Zymoseptoria tritici* and applications to wheat breeding. *Fungal Genetics and Biology* **79**: 33–41.

540 **Brunner PC, McDonald BA**. **2018**. Evolutionary analyses of the avirulence effector AvrStb6 in global
541 populations of *Zymoseptoria tritici* identify candidate amino acids involved in recognition. *Molecular*
542 *Plant Pathology* **19**: 1836–1846.

543 **Cabau C, Escudié F, Djari A, Guiguen Y, Bobe J, Klopp C. 2017.** Compacting and correcting Trinity
544 and Oases RNA-Seq *de novo* assemblies. *PeerJ* **5**: e2988.

545 **Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T**. **2009**. trimAl: a tool for automated alignment
546 trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973.

547 **Chandrasekaran M, Thangavelu B, Chun SC, Sathiyabama M**. **2016**. Proteases from
548 phytopathogenic fungi and their importance in phytopathogenicity. *Journal of General Plant Pathology*
549 **82**: 233–239.

550 **Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ**. **2015**. Second-generation
551 PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**: 7.

552 **Chartrain L, Sourdille P, Bernard M, Brown JKM**. **2009**. Identification and location of *Stb9*, a gene
553 for resistance to Septoria tritici blotch in wheat cultivars Courtot and Tonic. *Plant Pathology* **58**: 547–
554 555.

555 **Collemare J, Pianfetti M, Houlle A-E, Morin D, Camborde L, Gagey M-J, Barbisan C, Fudal I,**
556 **Lebrun M-H, Böhnert HU**. **2008**. *Magnaporthe grisea* avirulence gene *ACE1* belongs to an infection-
557 specific gene cluster involved in secondary metabolism. *New Phytologist* **179**: 196–208.

558 **Danguy des Déserts A, Bouchet S, Sourdille P, Servin B**. **2021**. Evolution of recombination
559 landscapes in diverging populations of bread wheat. *Genome Biology and Evolution* **13**: evab152.

560 **Dyrløv Bendtsen J, Nielsen H, von Heijne G, Brunak S**. **2004**. Improved prediction of signal peptides:
561 SignalP 3.0. *Journal of Molecular Biology* **340**: 783–795.

562 **Flor HH**. **1956**. The complementary genic systems in flax and flax rust**Joint contribution from the
563 Field Crops Research Branch, Agricultural Research Service, United States Department of Agriculture
564 and the North Dakota Agricultural Experiment Station. In: Demerec MBT-A in G, ed. Academic Press,
565 29–54.

566 **Flutre T, Duprat E, Feuillet C, Quesneville H**. **2011**. Considering transposable element diversification
567 in *de novo* annotation approaches. *PLOS ONE* **6**: e16526.

568 **Fones H, Gurr S**. **2015**. The impact of Septoria tritici blotch disease on wheat: An EU perspective.
569 *Fungal Genetics and Biology* **79**: 3–7.

570 **Franceschetti M, Maqbool A, Jiménez-Dalmaroni MJ, Pennington HG, Kamoun S, Banfield**
571 **MJ. 2017.** Effectors of filamentous plant pathogens: commonalities amid diversity. *Microbiology and*
572 *Molecular Biology Reviews* **81:** e00066–16.

573 **Garrison E, Marth G**. **2012**. Haplotype-based variant detection from short-read sequencing. *arXiv*:
574 1207.3907.

575 **Goodwin SB, Ben M'Barek S, Dhillon B, Wittenberg AHJ, Crane CF, Hane JK, Foster AJ, van**
576 **der Lee TAJ, Grimwood J, Aerts A, *et al.* 2011**. Finished genome of the fungal wheat pathogen
577 *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth
578 pathogenesis. *PLOS Genetics* **7**: e1002070.

579 **Grandaubert J, Bhattacharyya A, Stukenbrock EH**. **2015**. RNA-seq-based gene annotation and
580 comparative genomics of four fungal grass pathogens in the genus Zymoseptoria identify novel orphan
581 genes and species-specific invasions of transposable elements. *G3 Genes|Genomes|Genetics* **5**: 1323–
582 1333.

583 **Gu Z, Steinmetz LM, Gu X, Scharfe C, Davis RW, Li W-H**. **2003**. Role of duplicate genes in genetic
584 robustness against null mutations. *Nature* **421**: 63–66.

585 **Hann DR, Gimenez-Ibanez S, Rathjen JP**. **2010**. Bacterial virulence effectors and their activities.
586 *Current Opinion in Plant Biology* **13**: 388–393.

587 **Hartmann FE, McDonald BA, Croll D**. **2018**. Genome-wide evidence for divergent selection between
588 populations of a major agricultural pathogen. *Molecular Ecology* **27**: 2725–2741.

589 **Hartmann FE, Sánchez-Vallet A, McDonald BA, Croll D**. **2017**. A fungal wheat pathogen evolved
590 host specialization by extensive chromosomal rearrangements. *The ISME Journal* **11**: 1189–1204.

591 **Have A ten, Espino JJ, Dekkers E, van Sluyter SC, Brito N, Kay J, González C, van Kan JAL**.
592 **2010**. The *Botrytis cinerea* aspartic proteinase family. *Fungal Genetics and Biology* **47**: 53–65.

593 **Huang Z, Feng Y, Chen D, Wu X, Huang S, Wang X, Xiao X, Li W, Huang N, Gu L, *et al.* 2008.**
594 Structural basis for activation and inhibition of the secreted chlamydia protease CPAF. *Cell Host*
595 *Microbe* **4:** 529-42.

596 **Irieda H, Inoue Y, Mori M, Yamada K, Oshikawa Y, Saitoh H, Uemura A, Terauchi R, Kitakura**
597 **S, Kosaka A, *et al.* 2019.** Conserved fungal effector suppresses PAMP-triggered immunity by targeting
598 plant immune kinases. *Proceedings of the National Academy of Sciences* **116**: 496–505.

599 **Jashni MK, Dols IHM, Iida Y, Boeren S, Beenen HG, Mehrabi R, Collemare J, de Wit PJGM**.
600 **2015a**. Synergistic action of a metalloprotease and a serine protease from *Fusarium oxysporum* f. sp.
601 *lycopersici* cleaves chitin-binding tomato chitinases, reduces their antifungal activity, and enhances
602 fungal virulence. *Molecular Plant-Microbe Interactions®* **28**: 996–1008.

603 **Jashni MK, Mehrabi R, Collemare J, Mesarich CH, de Wit PJGM**. **2015b**. The battle in the
604 apoplast: Further insights into the roles of proteases and their inhibitors in plant–pathogen interactions.
605 *Frontiers in Plant Science* **6**: 584.

606 **Jones JDG, Dangl JL**. **2006**. The plant immune system. *Nature* **444**: 323–329.

607 **Kema GHJ, Mirzadi Gohari A, Aouini L, Gibriel HAY, Ware SB, van den Bosch F, Manning-**
608 **Smith R, Alonso-Chavez V, Helps J, Ben M'Barek S, *et al.* 2018**. Stress and sexual reproduction
609 affect the dynamics of the wheat pathogen effector AvrStb6 and strobilurin resistance. *Nature Genetics*
610 **50**: 375–380.

611 **Khang CH, Park S-Y, Lee Y-H, Valent B, Kang S**. **2008**. Genome organization and evolution of the
612 *AVR-Pita* avirulence gene family in the *Magnaporthe grisea* species complex. *Molecular Plant-Microbe*
613 *Interactions®* **21**: 658–670.

614 **Langlands-Perry C, Cuenin M, Bergez C, Ben Krima S, Gélisse S, Sourdille P, Valade R, Marcel**
615 **TC**. **2022**. Resistance of the wheat cultivar 'Renan' to Septoria leaf Blotch explained by a combination
616 of strain specific and strain non-specific QTL mapped on an ultra-dense genetic map. *Genes* **13**: 100.

617 **Letunic I, Bork P**. **2016**. Interactive tree of life (iTOL) v3: an online tool for the display and annotation
618 of phylogenetic and other trees. *Nucleic Acids Research* **44**: W242–W245.

619 **Li H, Durbin R**. **2009**. Fast and accurate short read alignment with Burrows–Wheeler transform.
620 *Bioinformatics* **25**: 1754–1760.

621 **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R,**
622 **Subgroup 1000 Genome Project Data Processing**. **2009**. The Sequence Alignment/Map format and
623 SAMtools. *Bioinformatics* **25**: 2078–2079.

624 **Li L, Stoeckert Jr CJ, Roos DS**. **2003**. OrthoMCL: identification of ortholog groups for eukaryotic
625 genomes. *Genome research* **13**: 2178–2189.

626 **Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z**. **2012**.
627 GAPIT: genome association and prediction integrated tool. *Bioinformatics* **28**: 2397–2399.

628 **Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, Zuccaro A, Reissmann S,**
629 **Kahmann R**. **2015**. Fungal effectors and plant susceptibility. *Annual Review of Plant Biology* **66**: 513–
630 545.

631 **López-Solanilla E, Bronstein PA, Schneider AR, Collmer A**. **2004**. HopPtoN is a *Pseudomonas*
632 *syringae* Hrp (type III secretion system) cysteine protease effector that suppresses pathogen-induced
633 necrosis associated with both compatible and incompatible plant interactions. *Molecular Microbiology*
634 **54**: 353–365.

635 **Martin DP, Murrell B, Golden M, Khoosal A, Muhire B**. **2015**. RDP4: Detection and analysis of
636 recombination patterns in virus genomes. *Virus Evolution* **1**.

637 **Meile L, Croll D, Brunner PC, Plissonneau C, Hartmann FE, McDonald BA, Sánchez-Vallet A**.
638 **2018**. A fungal avirulence factor encoded in a highly plastic genomic region triggers partial resistance
639 to Septoria tritici blotch. *New Phytologist* **219**: 1048–1061.

640 **Monod M, Capoccia S, Léchenne B, Zaugg C, Holdom M, Jousson O**. **2002**. Secreted proteases from
641 pathogenic fungi. *International Journal of Medical Microbiology* **292**: 405–419.

642 **Muszewska A, Stepniewska-Dziubinska MM, Steczkiewicz K, Pawlowska J, Dziedzic A, Ginalski**
643 **K**. **2017**. Fungal lifestyle reflected in serine protease repertoire. *Scientific Reports* **7**: 9147.

644 **Nielsen R, Yang Z**. **1998**. Likelihood models for detecting positively selected amino acid sites and
645 applications to the HIV-1 envelope gene. *Genetics* **148**: 929–936.

646 **Omrane S, Sghyer H, Audéon C, Lanen C, Duplaix C, Walker A-S, Fillinger S. 2015**. Fungicide
647 efflux and the MgMFS1 transporter contribute to the multidrug resistance phenotype in *Zymoseptoria*
648 *tritici* field isolates. *Environmental Microbiology* **17**: 2805–2823.

649 **Orbach MJ, Farrall L, Sweigard JA, Chumley FG, Valent B. 2000**. A telomeric avirulence gene
650 determines efficacy for the rice blast resistance gene *Pi-ta*. *The Plant cell* **12**: 2019–2032.

651 **Palma-Guerrero J, Ma X, Torriani SFF, Zala M, Francisco CS, Hartmann FE, Croll D, McDonald**
652 **BA. 2017**. Comparative transcriptome analyses in *Zymoseptoria tritici* reveal significant differences in
653 gene expression among strains during plant infection. *Molecular Plant-Microbe Interactions®* **30**: 231–
654 244.

655 **Petit-Houdenot Y, Lebrun MH, Scalliet G. 2021.** Understanding plant-pathogen interactions in
656 Septoria tritici blotch infection of cereals. In: Achieving durable disease resistance in cereals. London,
657 UK: *Burleigh Dodds Science Publishing*, 263-302.

658 **Plummer KM, Clark SJ, Ellis LM, Loganathan A, Al-Samarrai TH, Rikkerink EHA, Sullivan**
659 **PA, Templeton MD, Farley PC. 2004**. Analysis of a secreted aspartic peptidase disruption mutant of
660 *Glomerella cingulata*. *European Journal of Plant Pathology* **110**: 265–274.

661 **Rocafort M, Fudal I, Mesarich CH. 2020.** Apoplastic effector proteins of plant-associated fungi and
662 oomycetes. *Current Opinion in Plant Biology* **56:** 9-19.

663 **Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE,**
664 **Sánchez-Gracia A. 2017**. DnaSP 6: DNA Sequence Polymorphism analysis of large data sets.
665 *Molecular Biology and Evolution* **34**: 3299–3302.

666 **Rudd JJ, Kanyuka K, Hassani-Pak K, Derbyshire M, Andongabo A, Devonshire J, Lysenko A,**
667 **Saqi M, Desai NM, Powers SJ, *et al.* 2015**. Transcriptome and metabolite profiling of the infection
668 cycle of *Zymoseptoria tritici* on wheat reveals a biphasic interaction with plant immunity involving
669 differential pathogen chromosomal contributions and a variation on the hemibiotrophic lifestyle
670 definition. *Plant Physiology* **167**: 1158–1185.

671 **Saintenac C, Cambon F, Aouini L, Verstappen E, Ghaffary SMT, Poucet T, Marande W, Berges**
672 **H, Xu S, Jaouannet M, *et al.* 2021**. A wheat cysteine-rich receptor-like kinase confers broad-spectrum
673 resistance against Septoria tritici blotch. *Nature Communications* **12**: 433.

674 **Saintenac C, Lee W-S, Cambon F, Rudd JJ, King RC, Marande W, Powers SJ, Bergès H, Phillips**
675 **AL, Uauy C, *et al.* 2018**. Wheat receptor-kinase-like protein Stb6 controls gene-for-gene resistance to
676 fungal pathogen *Zymoseptoria tritici*. *Nature Genetics* **50**: 368–374.

677 **Sánchez-Vallet A, McDonald MC, Solomon PS, McDonald BA. 2015**. Is *Zymoseptoria tritici* a
678 hemibiotroph? *Fungal Genetics and Biology* **79**: 29–32.

679 **Sánchez-Vallet, A, Fouché S, Fudal I, Hartmann FE, Soyer JL, Tellier A, Croll D. 2018.** The
680 Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens. *Annual Review of*
681 *Phytopathology* **56:** 21-40.

682 **Sanz-Martín JM, Pacheco-Arjona JR, Bello-Rico V, Vargas WA, Monod M, Díaz-Mínguez JM,**
683 **Thon MR, Sukno SA**. **2016**. A highly conserved metalloprotease effector enhances virulence in the
684 maize anthracnose fungus *Colletotrichum graminicola*. *Molecular Plant Pathology* **17**: 1048–1062.

685 **Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A**. **2019**. The global burden
686 of pathogens and pests on major food crops. *Nature Ecology & Evolution* **3**: 430–439.

687 **Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M,**
688 **Söding J, *et al.* 2011**. Fast, scalable generation of high-quality protein multiple sequence alignments
689 using Clustal Omega. *Molecular Systems Biology* **7**: 539.

690 **Singh D, Simmonds J, Park RF, Bariana HS, Snape JW**. **2009**. Inheritance and QTL mapping of leaf
691 rust resistance in the European winter wheat cultivar 'Beaver.' *Euphytica* **169**: 253–261.

692 **Stamatakis A**. **2014**. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
693 phylogenies. *Bioinformatics (Oxford, England)* **30**: 1312–1313.

694 **Stergiopoulos I, de Wit PJGM**. **2009**. Fungal Effector Proteins. *Annual Review of Phytopathology* **47**:
695 233–263.

696 **Suffert F, Sache I, Lannou C**. **2013**. Assessment of quantitative traits of aggressiveness in
697 *Mycosphaerella graminicola* on adult wheat plants. *Plant Pathology* **62**: 1330–1341.

698 **Su M-Y, Som N, Wu C-Y, Su S-C, Kuo Y-T, Ke L-C, Ho M-R, Tzeng S-R, Teng C-H, Mengin-**
699 **Lecreulx D, *et al.* 2017**. Structural basis of adaptor-mediated protein degradation by the tail-specific
700 PDZ-protease Prc. *Nature Communications* **8**: 1516.

701 **Touzy G, Rincent R, Bogard M, Lafarge S, Dubreuil P, Mini A, Deswarte J-C, Beauchène K, Le**
702 **Gouis J, Praud S. 2019.** Using environmental clustering to identify specific drought tolerance QTLs in
703 bread wheat (*T. aestivum* L.). *Theoretical and Applied Genetics* **132**: 2859-2880.

704 **Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL,**
705 **Pachter L**. **2012**. Differential gene and transcript expression analysis of RNA-seq experiments with
706 TopHat and Cufflinks. *Nature Protocols* **7**: 562–578.

707 **Yang N, McDonald MC, Solomon PS, Milgate AW**. **2018**. Genetic mapping of *Stb19*, a new resistance
708 gene to *Zymoseptoria tritici* in wheat. *Theoretical and Applied Genetics* **131**: 2765–2773.

Zhong Z, Marcel TC, Hartmann FE, Ma X, Plissonneau C, Zala M, Ducasse A, Confais J, Compain J, Lapalu N, *et al.* 2017. A small secreted protein in *Zymoseptoria tritici* is responsible for avirulence on wheat cultivars carrying the *Stb6* resistance gene. *New Phytologist* **214**: 619–631.

Zhu T, Wang L, Rimbert H, Rodriguez JC, Deal KR, de Oliveira R, Choulet F, Keeble-Gagnère G, Tibbits J, Rogers J, *et al.* 2021. Optical maps refine the bread wheat *Triticum aestivum* cv. Chinese Spring genome assembly. *The Plant Journal* **107**: 303–314.

Zwiers L-H, de Waard MA. **2001**. Efficient *Agrobacterium tumefaciens*-mediated gene disruption in the phytopathogen *Mycosphaerella graminicola*. *Current Genetics* **39**: 388–393.

717 **Tables**

718 **Table 1.** Polymorphisms in AvrStb9 protein sequences between the wild type and transgenic mutant strains. Difference in amino acid residues between the
719 virulent (avrstb9) and avirulent (Avrstb9) phenotypes are highlighted in orange while residues potentially involved in virulence are highlighted in red.

Strain	**Mutant ID**	**Codon position / Amino acid residue**													**Phenotype (on *Stb9*)**
		58	**66-71**	**73**	**188**	**276**	**349**	**370**	**431**	**473**	**512**	**525**	**565**	**628**	
Wild-type strains															
IPO323 (avrstb9)	-	R	PCNSTI	N	Y	T	A	T	A	F	S	S	D	L	Virulent
IPO09593 (Avrstb9)	-	Q	Δ	D	V	N	V	N	D	Y	T	F	N	Y	Avirulent
IPO323 Avrstb9 transformants															
IPO323$_{Avrstb9}$ hph#5	E	Q	Δ	D	V	N	V	N	D	Y	T	F	N	Y	Avirulent
IPO323$_{Avrstb9}$ hph#6	F	Q	Δ	D	V	N	V	N	D	Y	T	F	N	Y	Avirulent
IPO323$_{Avrstb9}$ hph#9	H	Q	Δ	D	V	N	V	N	D	Y	T	F	N	Y	Avirulent
IPO323$_{Avrstb9}$ hph#12	K	Q	Δ	D	V	N	V	N	D	Y	T	F	N	Y	Avirulent
IPO323$_{Avrstb9}$ hph#14	L	Q	Δ	D	V	N	V	N	D	Y	T	F	N	Y	Avirulent
IPO323$_{Avrstb9}$ hph#18	N	Q	Δ	D	V	N	V	N	D	Y	T	F	N	Y	Avirulent
IPO323 Avrstb9 chimeric transformants															
IPO323$_{Avrstb9}$ hph#3	I	Q	Δ	D	V	T	A	T	A	F	S	S	D	L	Virulent
IPO323$_{Avrstb9}$ hph#11	J	Q	Δ	D	Y	T	A	T	A	F	S	S	D	L	Virulent
IPO323$_{Avrstb9}$ hph#7	G	Q	Δ	D	Y	T	A	T	A	F	S	S	D	L	Virulent
IPO323$_{Avrstb9}$ hph#17	M	Q	Δ	D	V	N	A	T	A	F	S	S	D	L	Virulent
IPO323 avrstb9 transformants															
IPO323$_{avrstb9}$ hph#3	O	R	PCNSTI	N	Y	T	A	T	A	F	S	S	D	L	Virulent

IPO323$_{avrstb9}$ hph#4	P	R	PCNSTI	N	Y	T	A	T	A	F	S	S	D	L	Virulent
Percentage of virulent isolates in the fungal population	87.5	98.86	86.63	98.86	98.86	54.54	88.63	**98.86**	**98.86**	89.77	**98.86**	**98.86**	**95.54**		
Probability of dN/dS>1	0.97	-	0.99	0.99	0.58	0.54	0.59	0.93	0.60	**0.99**	**1.00**	0.65	**1.00**		

720

Figures

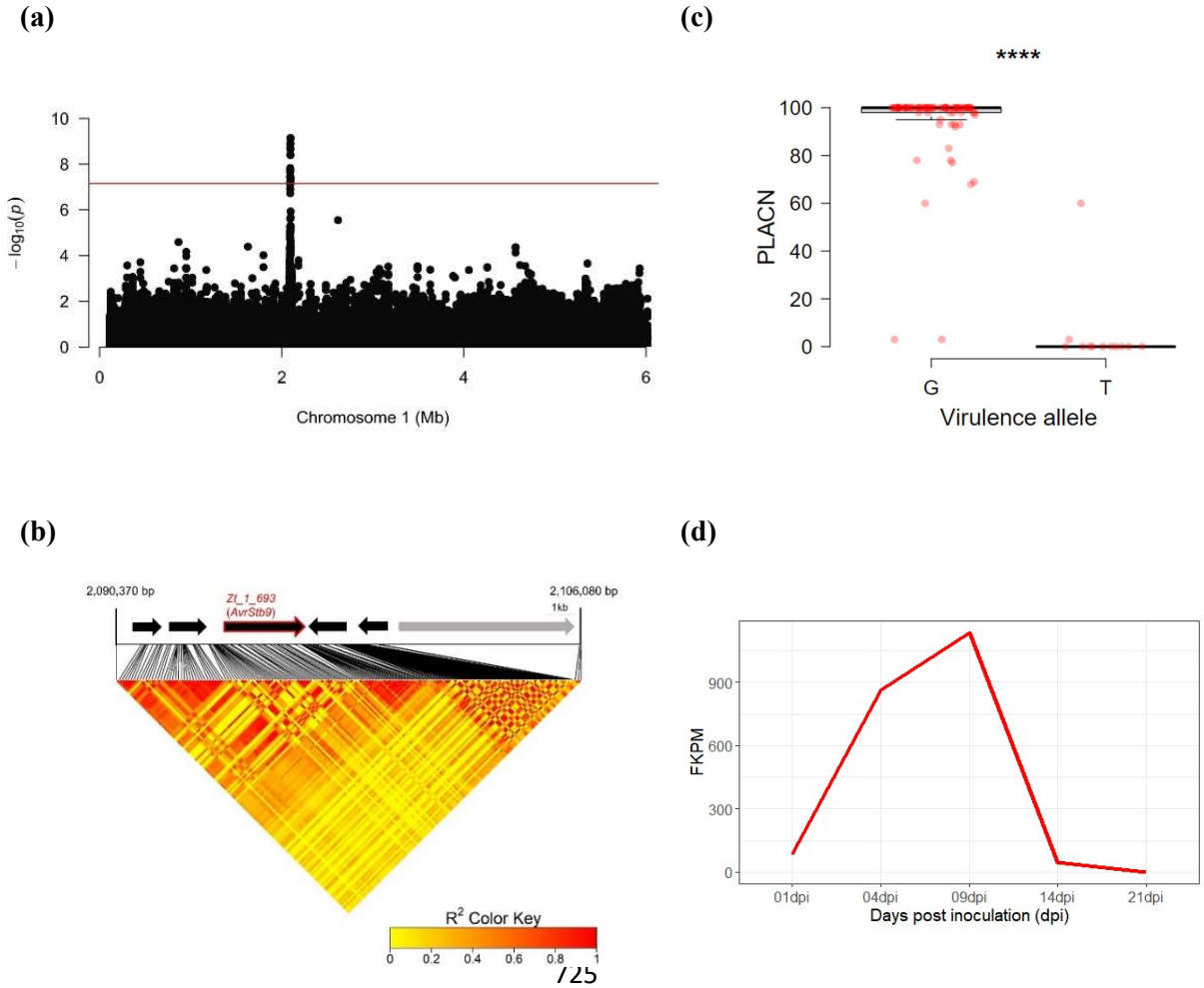


Figure 1. GWAS identifies a candidate gene controlling avirulence on the cultivar 'Soissons'. (a) Manhattan plot representing the association of multiple SNPs with PLACN on cultivar 'Soissons'. The X and Y axes indicate positions on chromosome 1 and −log10 (*p*-value) for associations, while the red line represents the Bonferroni threshold (α=0.05). (b) Top: Gene annotations (black arrows) and transposable elements (grey arrow) on the reference genome IPO323. The candidate gene is highlighted in red. Bottom: Pairwise linkage disequilibrium plot around the lead significant SNP. The association peak is associated with several polymorphisms in the gene *Zt_1_693*. (c) The difference in PLACN of isolates carrying the virulent and avirulent alleles (based on the lead SNP). Significance difference from the mean was computed using a two-sided Mann-Whitney test (****, $p<0.0001$). (d) Gene expression profile of *Zt_1_693* measured as fragments per kilobase of exon per million mapped fragments (FPKM) throughout the time course of *Z. tritici* infection on a susceptible wheat cultivar (Rudd *et al.*, 2015).

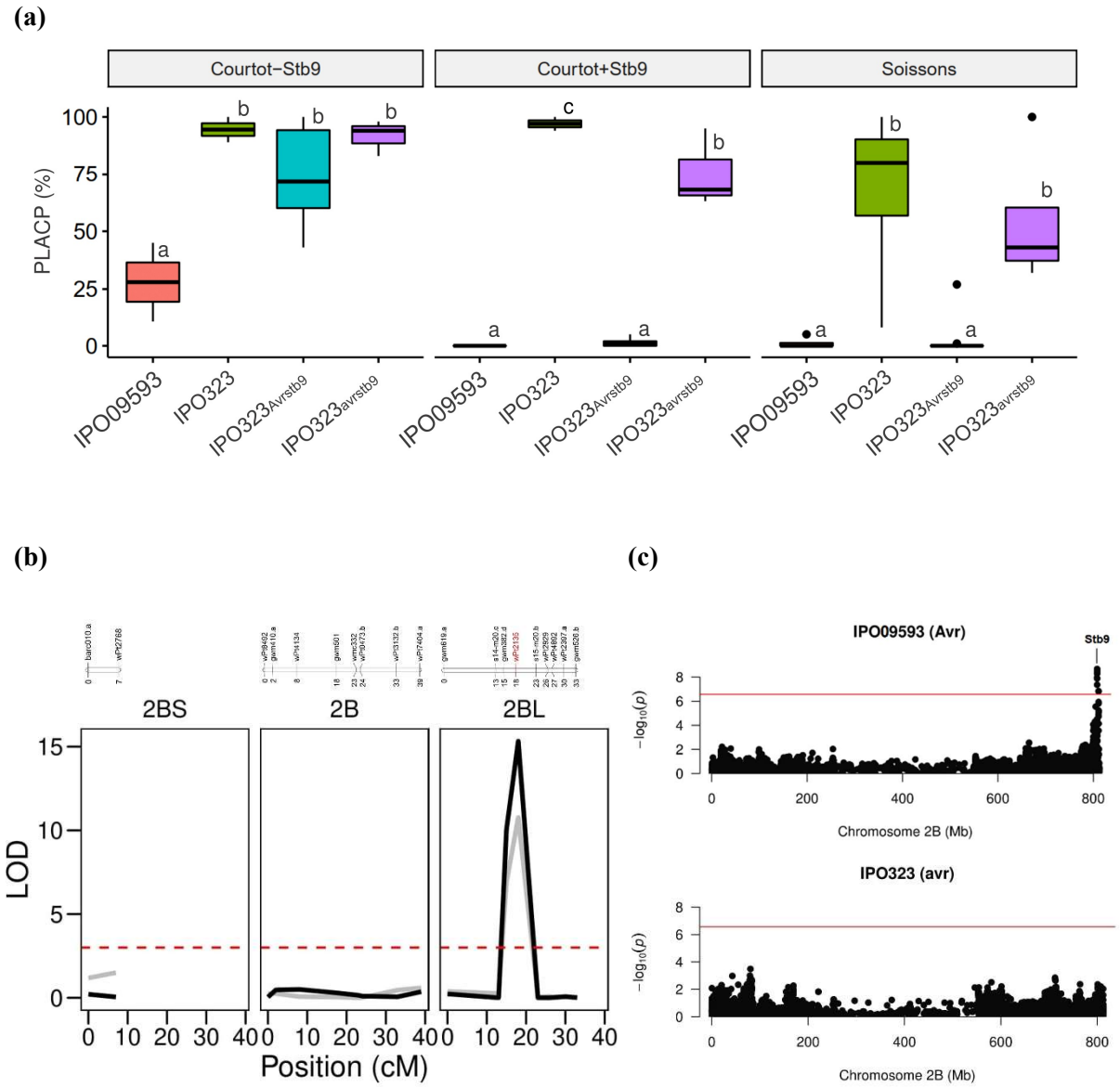
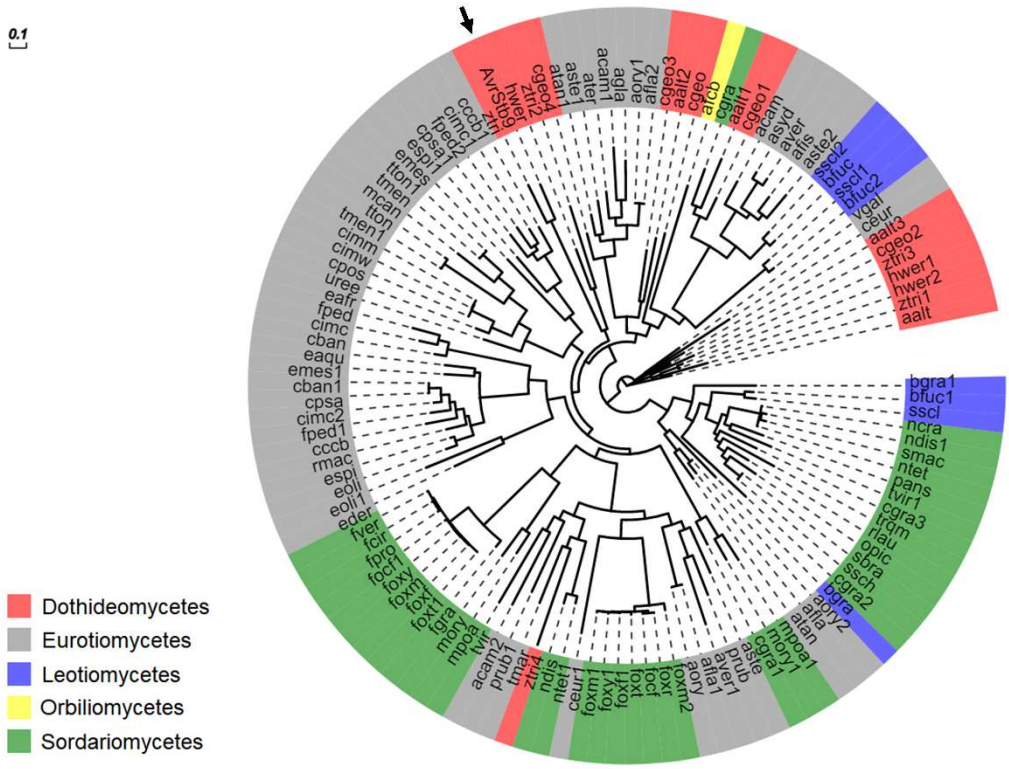


Figure 2. *Zt_1_693* encodes for the avirulence gene *AvrStb9* and triggers *Stb9* resistance. (a) PLACP at 26 dpi measured after inoculation with the avirulent isolate IPO09593, the virulent isolate IPO323 and IPO323 mutants expressing the avirulent allele (IPO323$_{Avrstb9}$) or the virulent allele (IPO323$_{avrstb9}$) on cultivars 'Courtot' (*Stb9*), a 'Courtot' mutant lacking *Stb9* and 'Soissons' (details on the phenotype of each mutant is provided in Fig. S6; Table S6). Statistical significance was determined using a Tukey's HSD test. (b) Genetic mapping of *Stb9* resistance gene using linkage mapping. Top panel: Linkage map and marker position in cM. Bottom panel: Logarithm of odds (LOD) score plot of the QTL analysis (Relative AUDPC for sporulation and necrosis in black and grey, respectively) possibly containing the *Stb9* locus mapped to the long arm of the chromosome 2B. The horizontal dashed line indicates a suggestive linkage threshold (LOD = 3). The position of the marker in the QTL peak is highlighted in red. (c) Manhattan plots representing association mapping results of relative AUDPC sporulation of the avirulent (top) and virulent (bottom) isolates in the bread wheat panel. Multiple SNPs in the telomeric region of chromosome 2B associated with the avirulent isolate are located in the *Stb9* genomic region.



765

766 **Figure 3.** Phylogenetic tree of selected *AvrStb9* orthologs/paralogs belonging to five distinct classes of
767 Ascomycota lineages. Leaf labels are color-coded according to class and represent abbreviated names of the fungal
768 species (Table S4). AvrStb9 is indicated with a black arrow.

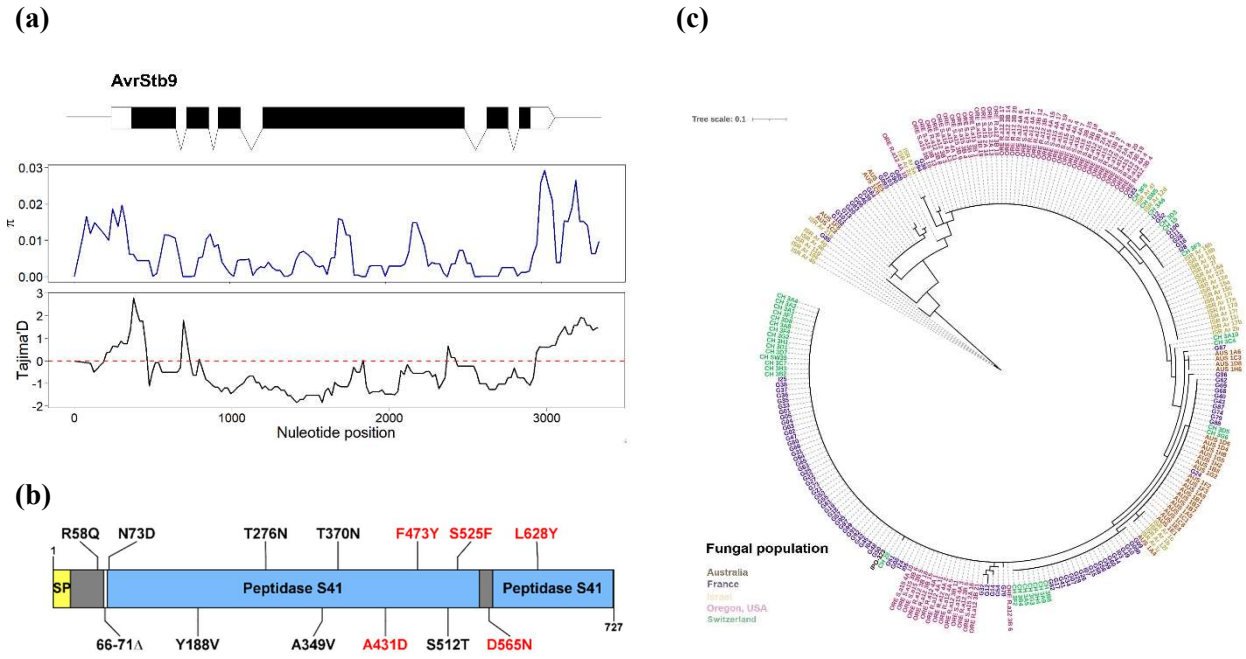


Figure 4. *AvrStb9* diversity and phylogeny in the global *Z. tritici* population. (a) The top panel shows the *AvrStb9* gene structure. Middle and low panels are sliding window analyses for nucleotide diversity (π) and Tajima's D across the *AvrStb9* locus, respectively, with a window size of 200 bp and a step size of 20 bp. The plots show a reduced nucleotide diversity and a drop in Tajima's D values, likely corresponding to a selective sweep spanning the region. (b) Schematic diagram of *AvrStb9* protein sequence with an annotation of amino acid substitutions between virulent and avirulent alleles. Substitutions potentially involved in virulence are highlighted in red. SP=Signal peptide. (c) A phylogenetic tree of *AvrStb9* protein sequence generated from a global *Z. tritici* population and the reference isolate IPO323.