●ISME

ARTICLE

Transcriptional response to host chemical cues underpins the expansion of host range in a fungal plant pathogen lineage

Stefan Kusch [1,4], Justine Larrouy[1,5,7], Heba M. M. Ibrahim [1,2,6,7], Shantala Mounichetty[1], Noémie Gasset[1], Olivier Navaud[1], Malick Mbengue[1], Catherine Zanchetta[3], Céline Lopez-Roques[3], Cécile Donnadieu[3], Laurence Godiard [1] and Sylvain Raffaele [1✉]

The host range of parasites is an important factor in assessing the dynamics of disease epidemics. The evolution of pathogens to accommodate new hosts may lead to host range expansion, a process the molecular bases of which are largely enigmatic. The fungus *Sclerotinia sclerotiorum* has been reported to parasitize more than 400 plant species from diverse eudicot families while its close relative, *S. trifoliorum*, is restricted to plants from the *Fabaceae* family. We analyzed *S. sclerotiorum* global transcriptome reprogramming on hosts from six botanical families and reveal a flexible, host-specific transcriptional program. We generated a chromosome-level genome assembly for *S. trifoliorum* and found near-complete gene space conservation in two representative strains of broad and narrow host range *Sclerotinia* species. However, *S. trifoliorum* showed increased sensitivity to the *Brassicaceae* defense compound camalexin. Comparative analyses revealed a lack of transcriptional response to camalexin in the *S. trifoliorum* strain and suggest that regulatory variation in detoxification and effector genes at the population level may associate with the genetic accommodation of *Brassicaceae* in the *Sclerotinia* host range. Our work proposes transcriptional plasticity and the co-existence of signatures for generalist and polyspecialist adaptive strategies in the genome of a plant pathogen.

INTRODUCTION

The range of hosts that pathogens can infect is determined by genetic and environmental factors. This host range is an important factor in assessing the dynamics of disease epidemics. Specialists parasitize one or few hosts, such as the cereal powdery mildews infecting only cereals [1] while generalists live on a range of diverse hosts. Natural selection by host populations and environmental factors drives frequent host switches and variations in pathogen–host range. Specialization occurs when a pathogen adapts to a specific host and enters a co-evolutionary arms race with it. In some instances, adaptation following host switching and host jumps involves the ability to efficiently colonize new hosts while retaining the ability to colonize the original host lineage, resulting in host range expansions [2, 3].

Adaptation to the gene-for-gene type of plant resistance is a paradigmatic example of co-evolutionary arms race [4]. Plant resistance genes typically recognize specific pathogenic proteins called effectors and mount a resistance reaction upon perception. Adapted pathogens evolved to avoid recognition by modification or loss of the respective effector [5]. This involves rapid adaptation, for example by selective sweeps [6] leaving characteristic patterns of variation in the genome of plant pathogens [7, 8]. In the potato late blight pathogen *Phytophthora infestans*, the resulting genetic variation is notably responsible for a tradeoff in effector activity on targets from different hosts [9] and distinctive

two-speed-genome architecture [10, 11]. In addition, balancing selection allows polymorphisms to persist in the gene pool and increases the standing genetic diversity in populations [12]. Arms race models generally assume an isolated pathogen co-evolving with one host *via* pairwise selection. However, pathogen genomes often evolve in response to selection caused by more than one host under diffuse co-evolution [13, 14]. Genomic signatures of diffuse selection and molecular adaptations associated with interaction with multiple hosts are largely unresolved [14].

Theories of evolutionary transitions suggest that genetic accommodation of pathogens to new hosts could entail general-purpose molecular bases supporting the colonization of any host (true generalist) or multiple modules dedicated to the colonization of specific hosts (polyspecialist) [3, 15]. Core effectors that act on conserved plant targets could serve as general-purpose molecules contributing to the accommodation of new hosts [2]. For example, the plant peroxidase inactivating effector PEP1 is conserved among the *Ustilaginaceae* and functional even in distantly related non-host plant species [16]. Comparative genomic studies have highlighted the role of the expansion of gene families related to secondary metabolism and detoxification in host range expansion in insect [17] and fungal [18] plant pathogens. A recent comparative study reported the transition from specialized one-speed genomes toward adaptive two-speed genomes correlated with increased host range in ergot fungi [19].

---

[1]Laboratoire des Interactions Plantes Microorganismes Environnement (LIPME), INRAE, CNRS, Castanet Tolosan Cedex, France. [2]Department of Genetics, Faculty of Agriculture, Cairo University, Giza, Egypt. [3]INRAE, US 1426, GeT-PlaGe, Genotoul, Castanet-Tolosan, France. [4]Present address: Unit of Plant Molecular Cell Biology, Institute for Biology I, RWTH Aachen University, Aachen, Germany. [5]Present address: Department of Pest Management and Conservation, Lincoln University, Lincoln, Canterbury, New Zealand. [6]Present address: Division of Plant Biotechnics, KU Leuven, Leuven, Belgium. [7]These authors contributed equally: Justine Larrouy, Heba M. M. Ibrahim. ✉email: sylvain.raffaele@inrae.fr

However, *P. infestans* and the head blight pathogen *Fusarium graminearum* exhibit typical two-speed-genome architecture and high degrees of host specialization, indicating that genome transitions cannot be unambiguously associated to directional host range variation [10, 11, 20]. The analysis of seven fungi from the *Metarhizium* genus of entomopathogens showed an expansion in genes encoding G protein-coupled receptors, proteases, transporters, enzymes for detoxification, and secondary metabolite biosynthesis, coinciding with increased host range [21]. In this lineage, horizontal gene transfers contributed to host range expansion [22]. By contrast, genome size was inversely correlated with host range in *Helicoverpa* butterflies [23]. The genome of the generalist aphid *Myzus persicae* has a gene count half that of *Acyrthosyphon pisum*, which is specialized on pea. Instead, *M. persicae* colonizes diverse host plants through rapid transcriptional induction of specific gene clusters [24]. These studies support the polyspecialist model of host range expansion in which pathogen genomes harbor several specialized gene modules, resulting from gene family expansion or differential gene expression.

The white mold fungus *Sclerotinia sclerotiorum* is a typical generalist plant pathogen reported to infect more than 400 plant species [25]. Its genome lacks signatures of selective sweeps [26] and two-speed architecture [27]. Host colonization by *S. sclerotiorum* is supported by division of labor enabling cooperation between cells of invasive hyphae [28]. In addition, the *S. sclerotiorum* genome shows signatures of adaptive translation, a selective process shaping the optimization of codon usage to increase protein synthesis efficiency. Codon optimization is particularly clear in genes expressed during plant colonization and encoding predicted secreted proteins [29]. Division of labor and codon optimization increase *S. sclerotiorum* fitness independently of the host genotype and therefore constitute genomic signatures of a true generalist. In addition, a study of a few effector candidate genes suggested the existence of plant host-specific expression patterns [30], and various small RNAs are differentially expressed on *Arabidopsis thaliana* and common bean (*Phaseolus vulgaris*) [31]. Furthermore, various isoforms originating from host-specific alternative splicing add additional plasticity to the transcriptomic profile of *S. sclerotiorum* [32] and suggest polyspecialist adaptations in this species. Nevertheless, the extent to which *S. sclerotiorum* genome harbors signatures of adaptation to a polyspecialist lifestyle has not been fully elucidated.

Here, we tested whether *S. sclerotiorum* strain 1980 activates distinct gene sets for the infection of diverse host species. Using global gene expression data obtained during the infection of plants covering six botanical families [28, 32, 33] we reveal core and flexible host-specific transcriptional programs. Genes activated specifically during the colonization of plants from the *Brassicaceae* family are largely conserved in the newly sequenced genome of the sister species *Sclerotinia trifoliorum* strain SwB9, unable to infect *Brassicaceae* plants. We show that lack of transcriptional response to camalexin in SwB9 coincides with increased sensitivity of several *S. trifoliorum* strains to this *Brassicaceae* defense compound. Interspecific comparison of promoter sequences suggests that regulatory variation may associate with the genetic accommodation of *Brassicaceae* in the *Sclerotinia* host range. Our work associates adaptive plasticity of a broad host range pathogen with specific responses to different host plants and exemplifies the co-existence of signatures for generalism and polyspecialism in the genome of a plant pathogen.

MATERIALS AND METHODS
Biological material
The fungal isolates *S. sclerotiorum* isolate 1980 [25] and *S. trifoliorum* SwB9 [34] were used in this study. These strains were chosen for their strong, but not extreme, aggressiveness on their respective hosts [34, 35] and for

originating both from hosts in the *Fabaceae* family. Camalexin sensitivity was also tested for *S. sclerotiorum* strains Fr.B5 [34], CU6.1, MB21, MB52, S55 [26], Blo.P144, and Blo.P154, *S. Trifolium* strains Be.B9, Li.A6 and Sw.B8 [34], *Sclerotinia minor* CBS 112.17, *Sclerotinia nivalis* MAFF 21347, and *Sclerotinia kitajimana* MAFF 410428. The fungi were cultivated on potato dextrose agar (PDA) at 23 °C or stored on PDA at 4 °C. *A. thaliana* accession Col-0 and the T-DNA insertion lines *cyp79b2 cyp79b3* [36] and *pad3*-1 [37] were grown in Jiffy pots under controlled conditions at 23 °C, with a 9-h light period at intensity of 120 μmol m^{-2} s^{-1} for up to 5 weeks.

RNA sequencing and gene expression analyses
RNA was collected in triplicates as described in [28] and [33], and from fungi cultivated on potato dextrose agar (Fluka) with DMSO or camalexin (125 μM for *S. sclerotiorum*, 25 μM for *S. trifoliorum*). Quality and concentrations of RNA were assessed with Agilent bioanalyzer nanochips. RNA sequencing (RNA-seq) was performed by Fasteris (Switzerland, Plan-les-Ouates) to produce Illumina reads (125 bp, paired-end) on a HiSeq2500 sequencer. Quality and adapter trimming of reads and reads mapping were performed as in [38] against the *S. sclerotiorum* isolate 1980 reference genome [27]. FPKM (fragments per kilobase of transcript per million mapped reads) tables were generated using the Cufflinks function cuffnorm with -compatible-hits-norm -library-norm-method classic-fpkm [39] (Dataset S1: Table S1). Differential expression analysis was run on 7423 protein-coding genes under a limma-edgeR pipeline [40] with cut-offs of $p \leq 0.01$ and $\log_2$ fold change (LFC) ≤ -1 (downregulated) or ≥ 1 (upregulated) using *S. sclerotiorum* gene expression on PDA as a reference (Dataset S1: Table S2). R v.3.5.1 [41] was used for statistical analysis and generation of plots. Codes for data analysis are deposited at https://github.com/stefankusch/sclerotinia_2020. GO and PFAM annotations (Dataset S1: Table S3, Table S4) and enrichment analyses were performed as described in [33].

S. trifoliorum genome assembly and comparative analysis
High molecular weight DNA isolation was performed as in [38]. Library preparation and sequencing were done at the GeT-PlaGe core facility, INRAE Toulouse, France, as in [38] with the following modifications: For one flow cell, 5 μg of purified DNA was sheared at 8 kb using the megaruptor1 system (Diagenode), followed by a DNA damage repair step on 2 μg of sample. Then an END-repair, dA-tailing of double-stranded DNA fragments, and adapters ligation were performed on the library. The library was loaded onto one R9.4.1 flow cell and was sequenced on a GridION instrument at 0.05 pmol within 48 h. The Canu v1.6 [42] assembly yielded 48 contigs between 3,548,298 and 6,307 bp and a total genomic length of 40,161,953 bp. We did four cycles of polishing with Pilon [43] and then used Blobtools v1.0 [44] to identify contigs with *Sclerotiniaceae* identity, as described before [38] (Fig. S1). The final genome assembly (Table 1) was subjected to repeat masking (RepeatMasker v4.0.7 [45], RepBase-20170127) prior to ab initio gene annotation with BRAKER2 [46] where we included all RNA-seq data of *S. trifoliorum* SwB9 produced in this study, i.e., cultivated in vitro on PDA (1×), PDA with DMSO (3×), and with 25 μM camalexin (3×), and from infection of *P. vulgaris* (1×). All gene models were manually curated via Web Apollo [47]. The genomes of *S. sclerotiorum* Ss1980 [27] and *S. trifoliorum* SwB9 were compared by synteny using MUMmer3 [48]; synteny plotting was performed with genoPlotR [49]. The proteomes of *S. sclerotiorum* Ss1980 [27], *Myrioesclerotinia sulcatula* MySu01 [38], *Botrytis cinerea* B05.10 [50], and *S. trifoliorum* SwB9 were compared via OrthoFinder [51].

Detection of cis-elements
The 1000-bp upstream sequences of all genes of *S. sclerotiorum* Ss1980 [27] and *S. trifoliorum* SwB9 were extracted using bedtools v2.25.0 [52]. The MEME-Suite 5.1.1. at http://meme-suite.org [53, 54] was used for motif discovery and enrichment (MEME, DREME), motif scanning (FIMO), and motif comparison (TOMTOM) against the *Saccharomyces cerevisiae* YEAS-TRACT_20130918.meme database. The *S. cerevisiae* motif-binding proteins were queried against the *S. sclerotiorum* Ss1980 proteome via BLASTP to identify the potential motif-binding orthologues DNA-binding proteins.

RESULTS
Host-specific transcriptome reprogramming in S. sclerotiorum
To investigate transcriptional reprogramming in *S. sclerotiorum* during the colonization of hosts from different botanical families,

Feature	*S. sclerotiorum* 1980	*S. trifoliorum* SwB9
Genome [bp]	38,906,597	39,909,921
Estimated size [kmer]	–	38.5 Mb
Genomic contigs	–	42
Mitochondrial contigs	1	3
Mitochondrial genome [bp]	132,532	226,016
Full chromosomes	16	2
Number chromosomes	16	16[a]
Largest contig [bp]	3,951,982	3,548,298
Average length [bp]	2,431,662	836,707
GC content	41.6%	39.9%
Coverage nanopore	–	86X
N_{50}	2,434,682	2,127,167
N_{90}	1,815,632	623,328
BUSCO[c]	–	97.4%
Repeat elements	6.2%	3.4%
Genes (BRAKER2)	–	11,290
Genes (curated)	10,768 (11,130)[b]	10,626
BUSCO[c] annotation	99.3%	97.8%

[a]Estimated by number of telomeric repeats in the assembly.
[b]The number in brackets includes the transposable elements annotated as genes in [27].
[c]Using the database for ascomycete core genes (ascomycota_odb9).

we performed RNA-seq analysis during infection of *A. thaliana*, *Solanum lycopersicum*, *Helianthus annuus*, *P. vulgaris*, *Ricinus communis* and *Beta vulgaris* (Fig. 1A). To control for variations in the kinetics of pathogen colonization on different hosts, samples were collected at similar infection stages by macro-dissection of disease lesion edge [28]. We performed differential expression analysis on 7524 expressed genes using mycelium grown in vitro as reference. We identified 1712 differentially expressed genes (DEGs, 1120 genes upregulated and 592 genes downregulated) in lesion edge across six hosts. The number of upregulated genes ranged from 110 on *H. annuus* to 777 on *B. vulgaris* (7.1-fold variation) and the number of downregulated genes ranged from 16 on *S. lycopersicum* to 409 on *R. communis* (25.6-fold variation) (Fig. 1A). Hierarchical clustering and principal component analysis showed clear separation of *S. sclerotiorum* gene expression according to the infected host (Fig. 1B, Fig. S2, Dataset S1: Tables S5 and S6). In colony edge, 53 DEGs (4.7%) were upregulated during the colonization of all host plants, 483 DEGs (43.1%) where upregulated on at least two host plants, and 584 DEGs (52.1%) showed specific upregulation on one host (Fig. 1C). The number of genes upregulated on one host only represented 0% on *S. lycopersicum*, 4.5% on *P. vulgaris*, 14.7% on *A. thaliana*, 15.5% on *H. annuus*, 32.3% on *R. communis* and 39.6% on *B. vulgaris*. These results indicate that the colonization of some hosts relies on core virulence genes while other hosts trigger the activation of host-specific fungal virulence programs.

To document the functional diversity of *S. sclerotiorum* genes differentially expressed during host colonization, we analyzed GO and PFAM annotation enrichment with genes upregulated *in planta*. We found 11 GOs and 94 PFAMs significantly enriched with upregulated genes (chi-squared adjusted p val < 0.01) during the colonization of at least one plant (Dataset S1: Table S7 and S8). Thirteen of these PFAMs were enriched in genes upregulated during the colonization of five or six host species, including

galactosidase (PF10435, PF13363, PF13364, PF16499), glycosyl hydrolase (GH) (PF00150, PF01301, PF00295), sugar transporter (PF00083), cytochrome p450 (PF00067) and domain of unknown function DUF4965 (PF16335) domains. Forty PFAMs were enriched in genes upregulated during the colonization of a single host species (Fig. S3). Next, we examined the differential expression pattern of the largest gene families enriched in upregulated genes using hierarchical clustering of LFC values. We identified 59 expressed genes harboring a cytochrome p450 domain into eight hierarchical clusters (a–h, Fig. 1D). The colonization of each host lead to a specific signature of p450 genes upregulated in *S. sclerotiorum*. We identified 147 expressed genes harboring a GH domain that classified into 11 hierarchical clusters (Fig. S4). The colonization of each host upregulated specific sets of GH genes. The 1120 upregulated genes included 246 secretome genes (21.9%), coding for putative secreted proteins. Secretome genes represented 8.5% of the 7423 expressed genes, 18.3% of genes upregulated on one host, 23.2% of genes upregulated on two, three, or four hosts, and 35.5% of genes upregulated on five or six hosts. This enrichment is consistent with a prominent role of fungal secreted proteins in the interaction with host plants, and indicate that secreted proteins contribute to a larger part of the *S. sclerotiorum* core infection program than the host-specific infection programs.

Arabidopsis-specific induced genes are conserved in the genome of the nonpathogenic *S. trifoliorum*

S. trifoliorum is closely related to *S. sclerotiorum* but has a host range restricted to plants in the Asterids and Fabids families [55]. Unlike *S. sclerotiorum*, *S. trifoliorum* colonizes *A. thaliana* very poorly to not at all. To gain insights into the evolution of *A. thaliana*-specific upregulated genes in the *Sclerotinia* lineage, we performed genome sequencing of *S. trifoliorum* isolate SwB9 using Illumina short-read and nanopore long-read data. We assessed completeness of the genome with BUSCO [56] and found 97.4% of highly conserved ascomycete genes present, suggesting near-complete gene space coverage. According to synteny analysis 67.2% of the *S. trifoliorum* genome aligned to 68.8% of the *S. sclerotiorum* genome, with a number of large-scale inversions and an apparent chromosome arm exchange (Fig. 2A, B). Overall, we generated a near-chromosome assembly for *S. trifoliorum* SwB9 of sufficient quality for gene space comparison with *S. sclerotiorum*. We predicted 10,626 unique gene models in the genome of *S. trifoliorum* SwB9. The *S. trifoliorum* proteome comprises 2,020 proteins with putative transmembrane domains and 705 proteins contained predicted secretion signal peptides, 73 of which are possible effector candidates according to an EffectorP 2.0 search [57]. We found 9155 orthogroups (OGs) shared between the two *Sclerotinia* species containing 10,085 (*S. sclerotiorum*) and 9922 (*S. trifoliorum*) genes, respectively (Fig. 2C, Dataset S1: Table S9). 685 *S. sclerotiorum* genes had no ortholog in *S. trifoliorum*, including 249 *S. sclerotiorum* genes from 199 OGs that did not contain *S. trifoliorum* genes. Out of 306 *S. sclerotiorum* genes induced at the edge of colonies on *A. thaliana*, 300 (98%) had orthologs in *S. trifoliorum*. Out of 45 *S. sclerotiorum* genes upregulated specifically on *A. thaliana*, 42 (93.3%) had orthologs in *S. trifoliorum* (Dataset S1: Table S10). Twenty-seven of the 457 (5.9%) DEGs on the common host *P. vulgaris* were absent in *S. trifoliorum*, while all 28 *P. vulgaris*-specific DEGs had *S. trifoliorum* orthologues. In total, 1625 of the 1712 DEGs in *S. sclerotiorum* had orthologues in *S. trifoliorum* SwB9 (95.0%). Therefore, expansion of the host range of *Sclerotiniaceae* fungi to *A. thaliana* largely relied on genes acquired prior to the divergence between *S. sclerotiorum* and *S. trifoliorum*.

S. trifoliorum is sensitive and not responsive to phytoalexins

Plants in the order Brassicales like *A. thaliana* produce tryptophan-derived defensive metabolites such as indolic glucosinolates
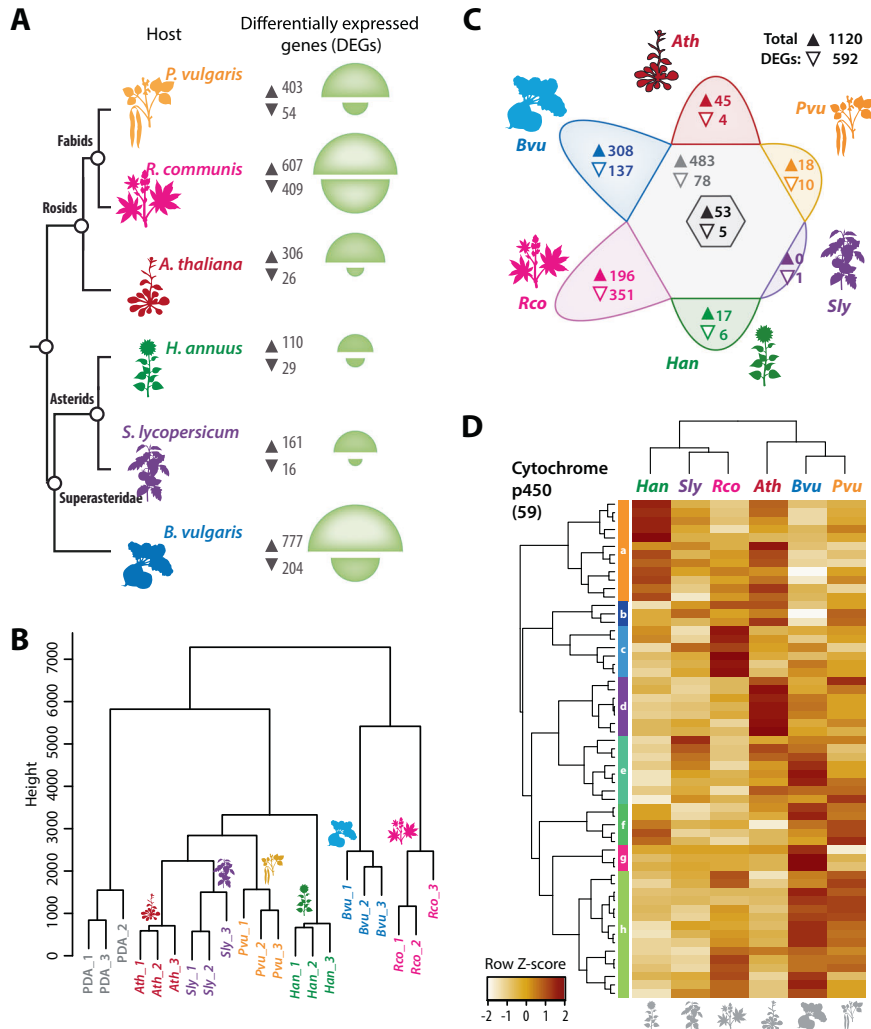
Fig. 1 Differential regulation of *S. sclerotiorum* transcriptome during the colonization of plants from six botanical families. A Number of *S. sclerotiorum* genes expressed differentially (DEGs) on each host compared to in vitro-grown colonies. Bubbles are sized proportionally to the number of DEGs in each treatment, labels showing the corresponding number of genes. The upper half of bubbles correspond to genes upregulated, the lower half to genes downregulated. **B** Hierarchical clustering of RNA-seq samples based on the expression of DEGs. Numbers in branch labels correspond to biological replicates. **C** Distribution of DEGs according to host species. For each sector, the upper value (Δ) corresponds to upregulated genes, the lower value (∇) to downregulated genes. The central dark gray hexagon shows DEGs detected on all six hosts, the light gray hexagon shows DEGs detected on 2 to 5 hosts. **D** Hierarchical clustering of log2 fold change of expression for 59 expressed genes encoding cytochrome p450. Eight hierarchical clusters labeled a-h were delimited. *Ath, Arabidopsis thaliana; Bvu, Beta vulgaris; Han, Helianthus annuus; Pvu, Phaseolus vulgaris; Rco, Ricinus communis; Sly, Solanum lycopersicum*.

(iGLs), the indole alkaloid camalexin, and the indole phytoalexin brassinin. The infection of plants from the Brassicales by some pathogens involves metabolizing these plant defense compounds into non-toxic derivatives [58, 59]. To test whether the toxicity of tryptophan-derived plant defense metabolites could explain the inability of *S. trifoliorum* to colonize Brassicales, we compared the sensitivity of *S. sclerotiorum* and *S. trifoliorum* to five major defense tryptophan-derivatives produced by *A. thaliana* using an in vitro growth assay (Fig. 3A). Both *S. sclerotiorum* and *S. trifoliorum* tolerated similar concentrations of tryptophan, raphanusamic acid, indole-3-carboxylic acid and indole-3ylmethylamine (Figure S5). However, in contrast to *S. sclerotiorum* the growth of *S. trifoliorum* was completely inhibited by 125 μM camalexin and 250 μM brassinin. The four *S. trifoliorum* strains we tested were unable to grow on 125 μM camalexin, while all eight *S. sclerotiorum* strains tested grew without major defects (Fig. S5). The growth of strains of the closely related *S. minor, S. nivalis*, and *S. kitajimana* was also drastically impaired on 125 μM camalexin (Fig. S5). To verify that these compounds contribute to plant resistance to *S. trifoliorum*,

we compared the colonization of wild type and *cyp79b2/b3 A. thaliana* plants, defective in iGLs biosynthesis [60] (Fig. 3B). Three days post inoculation (dpi), *S. sclerotiorum* had fully colonized 56% of Col-0 wild type leaves, while *S. trifoliorum* hardly grew out of the inoculation plug (0% of leaves fully colonized), consistent with the inability of *S. trifoliorum* to colonize *A. thaliana*. The *cyp79b2/b3* mutant was more susceptible to *S. sclerotiorum*, harboring 89% of leaves fully colonized at 3 dpi. Remarkably, *cyp79b2/b3* plants appeared susceptible to *S. trifoliorum* in this assay, as 67% of leaves were fully colonized at 3 dpi. These results indicate that sensitivity to tryptophan-derived defense metabolites contribute to the inability of *S. trifoliorum* to infect *A. thaliana*.

This prompted us to explore the extent to which *S. sclerotiorum* and *S. trifoliorum* differ in their transcriptional response to camalexin. For this, we performed RNA sequencing of *S. sclerotiorum* and *S. trifoliorum* colonies grown on PDA plates with camalexin (Dataset S1: Table S11). We identified 323 genes upregulated in *S. sclerotiorum* on camalexin (Dataset S1: Table S12). Among those, 180 (55.7%) were also induced in at least one of the
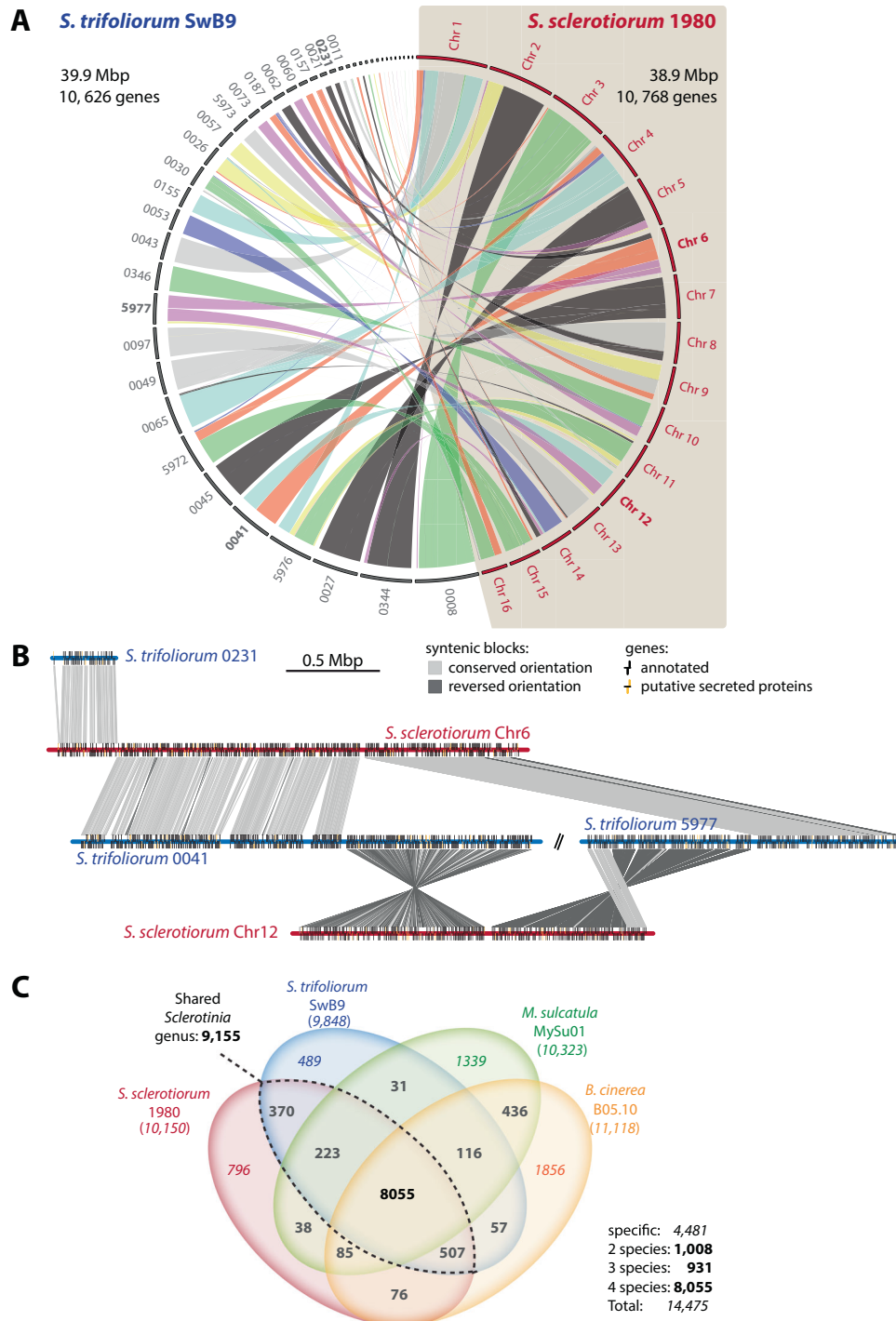
Fig. 2 Synteny and orthology relationships between *S. sclerotiorum* and *S. trifoliorum* genomes. A Overall synteny between *S. sclerotiorum* chromosomes (red) and *S. trifoliorum* contigs (gray). Colored ribbons connect syntenic regions across genomes. Chromosomes and contigs shown in B are labeled with bold fonts. **B** Synteny of *S. sclerotiorum* chromosomes 6 and 12 against the *S. trifoliorum* assembled contigs. **C** Venn diagram summarizing the results from the Orthofinder analysis between the proteomes of *S. sclerotiorum* 1980, *S. trifoliorum* SwB9, *Myriosclerotinia sulcatula* MySu01, and *Botrytis cinerea* B05.10. Values in bold correspond to number of orthogroups, values in italics correspond to number of genes.

plant species (Dataset S1: Table S12) and 301 had orthologs in *S. trifoliorum*. Only 42 genes were upregulated in *S. trifoliorum* on camalexin (Dataset S1: Table S12), among which 40 had orthologs in *S. sclerotiorum*. Hierarchical clustering of expression for the 341 DEGs and their orthologs from the two species identified a cluster (b) of 70 upregulated during the colonization of *A. thaliana* and on camalexin in *S. sclerotiorum* (Fig. 3C, D). Only 3 orthologs of these

genes in *S. trifoliorum* were induced on camalexin (Fig. 3D, Dataset S1: Table S12). We detected 23 genes encoding putative secreted proteins in this dataset of which 22 were conserved in *S. trifoliorum* SwB9. Nineteen of the respective proteins were predicted to be secreted in *S. trifoliorum* (Dataset S1: Table S13). One of these (Sscle05g046060, SCLTRI_001855) encoded an effector candidate according to EffectorP 2.0 analysis. The
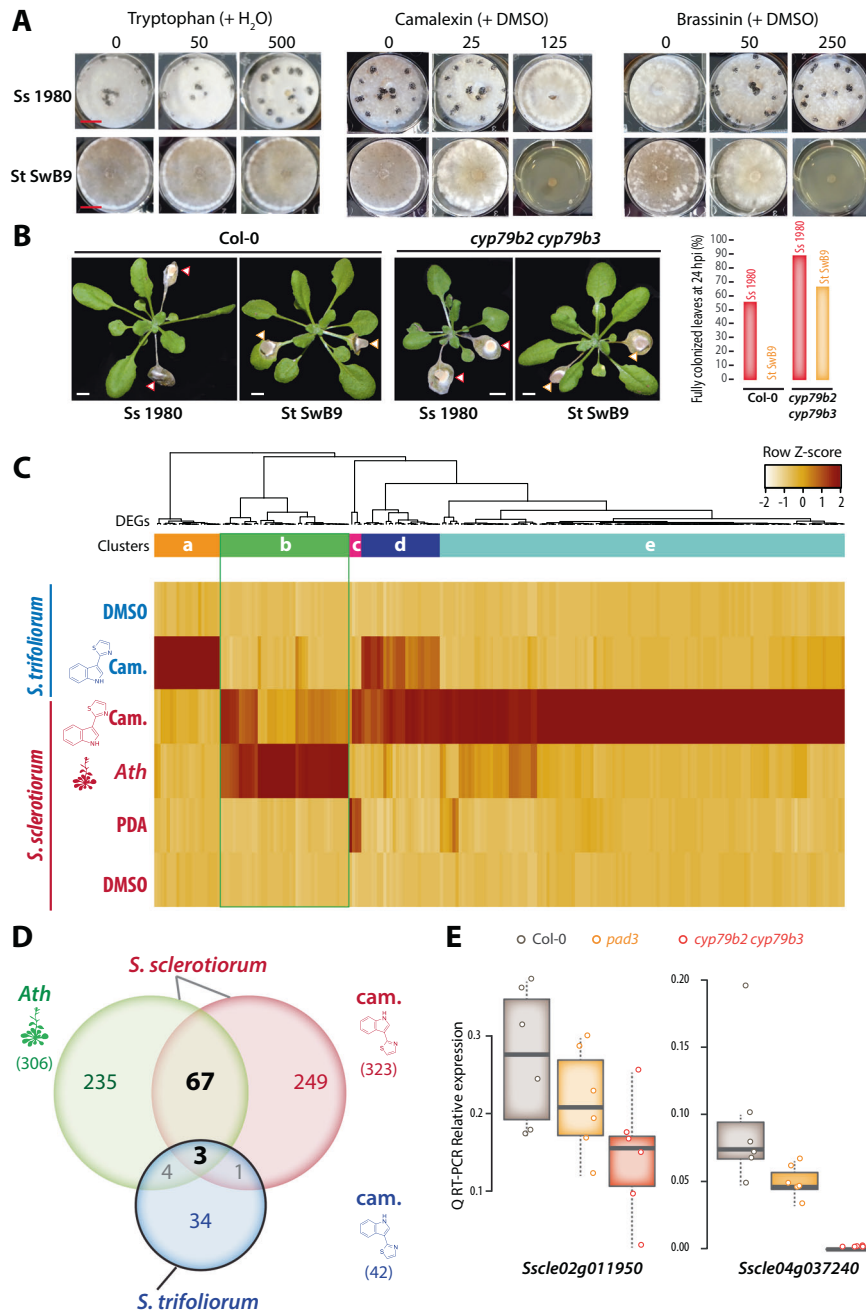
Fig. 3 S. trifoliorum is more sensitive than S. sclerotiorum to phytoalexins in vitro and in planta. A Phytoalexin tolerance plate assay. *S. sclerotiorum* 1980 and *S. trifoliorum* SwB9 were cultivated on potato dextrose agar (PDA) containing phytoalexins at different concentrations. The solvent used for each compound is indicated between brackets. Photos were taken after seven days; the experiment was conducted three times with similar results. Scale bar: 1 cm. **B** The *A. thaliana* wild type Col-0 and the indole glucosinolate and camalexin deficient mutant *cyp79b2 cyp79b3* were infected with *S. sclerotiorum* 1980 and *S. trifoliorum* SwB9. Photos were taken three days after inoculation. Arrowheads indicate agar plugs with *S. sclerotiorum* (red) or *S. trifoliorum* (yellow). Scale bar: 1 cm. The bar chart indicates the proportion of inoculated leaves fully colonized by each fungus for $n = 9$ or 10 leaves. **C** FPKM expression values for differentially expressed genes during the colonization of *A. thaliana* (Ath, *S. sclerotiorum* only) and upon camalexin treatment (Cam., *S. sclerotiorum* and *S. trifoliorum*). Expression of orthologs of DEGs from the other *Sclerotinia* species are shown for comparison purposes. **D** Venn diagram illustrating the number of DEGs in *S. sclerotiorum* during the colonization *A. thaliana*, *S. sclerotiorum* growth on camalexin and orthologs of *S. trifoliorum* DEGs during growth on camalexin. The number between brackets corresponds to complete gene sets. **E** Relative expression at 72 h post inoculation for two *S. sclerotiorum* genes determined by quantitative reverse transcription PCR (Q RT-PCR) on *A. thaliana* wild type plants, *cyp79b2/cyp79b3* and *pad3* mutants. Values shown are for 6 independent biological replicates averaged over two technical replicates. DMSO dimethyl sulfoxide, PDA potato dextrose agar.

putative effector is 87 amino acids in length in both species, 12 of which were variable between the two species (Fig. S6). A BLASTP search against the nr database ($E < 1e-10$) revealed broad conservation of this effector in *Sclerotiniaceae* and in *Fusarium* sp., suggesting that Sscle05g046060 could represent a core

effector with a conserved function in virulence against *Brassicaceae*.

We propose that the 70 *S. sclerotiorum* genes upregulated on camalexin and during the colonization of *A. thaliana* could respond to camalexin *in planta*. To confirm that camalexin and
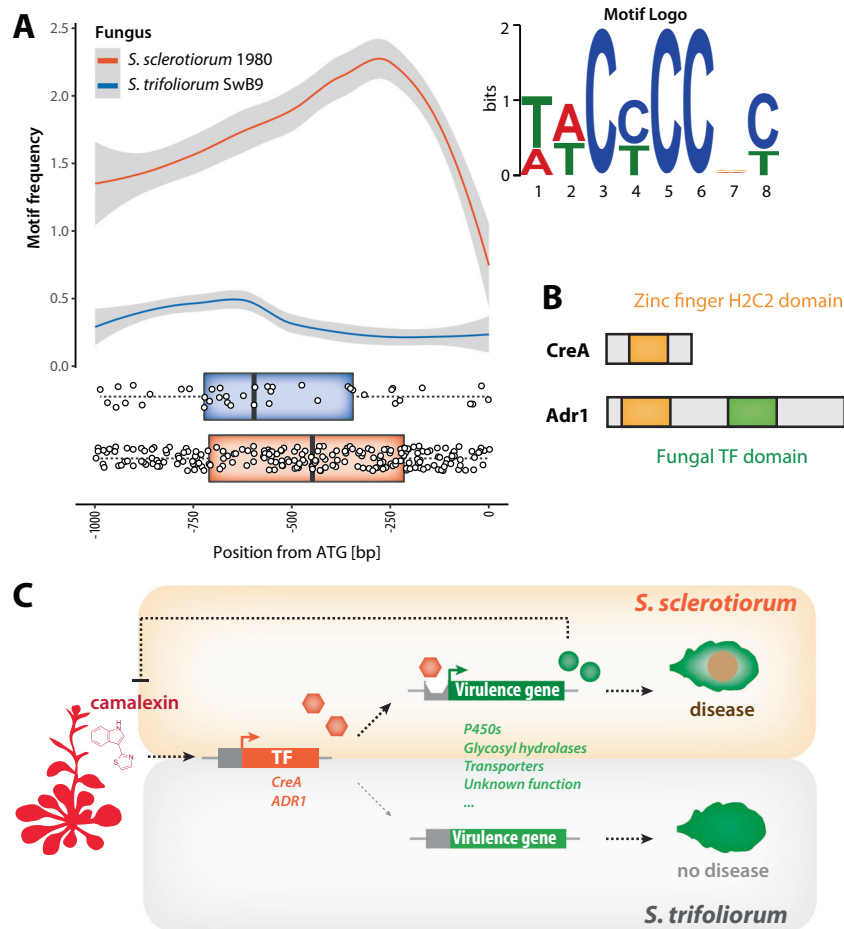
Fig. 4 **_Cis_-regulatory variation in _S. sclerotiorum_ genes induced by camalexin and during the colonization of _A. thaliana_. A** Top: sequence logo of the WWCCCCRC _cis_-regulatory element enriched in _S. sclerotiorum_ genes induced by camalexin and _A. thaliana_ infection by comparison to their _S. trifoliorum_ orthologs. Bottom: distribution of the WWCCCCRC element in the 1 kbp upstream sequence of _S. sclerotiorum_ genes induced by camalexin and _A. thaliana_ infection and their _S. trifoliorum_ orthologs. **B** Domain structure of _S. sclerotiorum_ orthologs of CreA and Adr1 transcription factors known to bind the WWCCCCRC _cis_ element in yeast. Length of the boxes is proportional to number of amino acids. **C** Graphical summary illustrating how variation in _cis_-regulatory regions (gray boxes) leads to differential activation (colored plain arrowheads) of fungal genes (orange and green boxes) and could have contributed to the evolution of virulence on _Arabidopsis_ in _Sclerotinia_. Hexagons and circles are fungal proteins produced by genes of the corresponding colors. The block arrow indicates functional inhibition. TF transcription factor.

iGLs produced by _A. thaliana_ are required to trigger the induction of _S. sclerotiorum_ genes, we compared by quantitative RT-PCR the expression of seven _S. sclerotiorum_ genes during the colonization of wild type, _cyp79b2/b3_, and _pad3_ plants, defective in camalexin biosynthesis (Fig. 3E, Fig. S7). At 72 hpi, the expression of _Sscle07g055350_ and _Sscle15g106410_ was not different during infection of _A. thaliana_ wild type and mutant plants. The expression of _Sscle02g011950_ and _Sscle02g022130_ was strongly reduced during infection of _cyp79b2/b3_ but not _pad3_ as compared to wild type. The expression of _Sscle04g037240_, _Sscle08g067130_, and _Sscle16g108230_ was significantly reduced both during infection of _cyp79b2/b3_ and _pad3_ mutants as compared to wild type (Welch's _t_ test _p_ value < 0.05). These results indicated that host-derived camalexin modulates the expression of _S. sclerotiorum_ genes during infection but that _S. trifoliorum_ does not significantly reprogram its transcriptome in response to this compound.

Cis-regulatory variation associates with the evolution of camalexin responsiveness in the _Sclerotinia_ genus
We propose that transcriptome plasticity in host responsive genes contributed to host range variation in the _Sclerotiniaceae_. In particular, regulatory variation in the 70 _S. sclerotiorum_ genes

upregulated both on _A. thaliana_ and on camalexin in vitro (Fig. 3, Dataset S1: Table S13) may have facilitated the colonization of plants from the _Brassicaceae_. To support this hypothesis, we first analyzed the conservation of these 70 genes in 670 species across the fungal kingdom (Figure S8). 91% of these genes were detected in more than 100 fungal species, suggesting that gene presence/absence polymorphism played a limited role in the evolution of responsiveness to plant-derived camalexin in _S. sclerotiorum_. Second, we analyzed the _cis_-elements in the 1000-bp upstream sequences of these genes. The WWCCCCRC motif was significantly enriched in these 70 _S. sclerotiorum_ genes but not in the upstream sequences of the 69 orthologous _S. trifoliorum_ genes (Fig. 4A). Using published yeast protein-DNA ChIP data we found five proteins of _Saccharomyces cerevisiae_, Mig1, Mig2, Mig3, Adr1, and Rsf2, that are capable of binding WWCCCCRC-like motifs. By homology searches and phylogenetic analyses, we identified potential orthologous proteins in _S. sclerotiorum_ and _S. trifoliorum_ (Fig. S9, File S1). Probable orthologs of yeast _Mig1-3_ (_Multicopy Inhibitor of GAL gene expression_) and _Aspergillus CreA_ (_Carbon catabolite repressor A_) encoding a protein harboring a central zinc-finger H2C2 domain were _Sscle01g002690_ and _Scltri_002966_ (Fig. 4B). The expression of this gene was detectable in both fungi at FPKM > 100 and was slightly induced by camalexin (LFC

0.60) and during *A. thaliana* colonization (LFC 0.38) in *S. sclerotiorum* (Fig. 4C). Both *Sscle01g002690* and *Scltri_002966* contained two and four occurrences, respectively, of the WWCCCCRC motif in their upstream sequences. Probable orthologs of yeast *ADR1* (*Alcohol Dehydrogenase II synthesis Regulator*) and *RSF2* (*Respiration factor 2*) encoding a protein harboring a N-terminal zinc-finger H2C2 domain and a C-terminal fungal specific transcription factor domain PF04082 were *Sscle07g055670* and *Scltri_004270* (Fig. 4B). The expression of *CreA* (FPKM > 100) and *ADR1* (FPKM > 20) was detectable in vitro in *S. trifoliorum* and *S. sclerotiorum* and *in planta* in *S. sclerotiorum*. They are therefore prime candidates for mediating large-scale transcriptome reprogramming in response to camalexin in the *Sclerotinia* lineage. These results suggest that *cis*-regulatory variation in targets of SsCreA and SsADR1 contributed to the evolution of transcriptional responsiveness to camalexin in *S. sclerotiorum* (Fig. 4C).

DISCUSSION
We analyzed the global transcriptome of *S. sclerotiorum* during the colonization of hosts from six botanical families, providing a unique opportunity to test for the existence of generalist and host-specific transcriptomes in this fungal pathogen. While previous investigations revealed adaptations to a true generalist lifestyle supporting the colonization of any host [28, 29], we provide here molecular evidence for polyspecialism, the use of multiple independent modules dedicated to the colonization of specific hosts. These findings indicate that adaptation to new hosts can select for generalist and polyspecialist features within a single genome. We highlight a key role of regulatory variation in conserved fungal genes for the expansion of host range in this pathogen lineage.

We identified a subset of host species triggering specialized transcriptome reprogramming in *S. sclerotiorum*. Genes related to detoxification of host defense compounds were enriched in the specialized transcriptomes, while the core transcriptome over-represented functions associated with carbohydrate catabolism and sugar transport. Host-specific regulation of pathogen genes confers the ability to quantitatively modulate the virulence program to match requirements for specific hosts. Transcriptional plasticity contributes to successful colonization of different host plants in aphids and a range of hemi-biotrophic fungal and oomycete pathogens [24, 61]. Host-specialized transcriptomes are often small and consist of secreted proteins with roles in modulating host plant defense responses and nutrient assimilation [61] as for the head blight pathogen *F. graminearum* [62, 63] and the septoria leaf blotch pathogen *Zymoseptoria tritici* [64]. A host-specialized transcriptome has been suggested previously for *S. sclerotiorum* when infecting *B. napus* and lupin (*Lupinus angustifolius*), although the majority of induced genes were induced on both hosts [65]. Our analysis on six dicot host species revealed that 52% of *S. sclerotiorum* genes upregulated *in planta* were host-specific. Predicting how the relative proportion of host-specific transcripts would differ by including more hosts in the analysis remains challenging since the number of host-specific induced genes varied considerably according to host (from 0% on *S. lycopersicum* to 36.9% on *B. vulgaris*). Clear host-specific gene regulation suggests that transcriptome-based reverse ecology approaches are feasible, allowing for instance to identify new host defense mechanisms based on pathogen transcriptome data.

Gene losses and gene gains by recombination, horizontal gene transfer, or copy number variation are efficient means to modify the host range [66, 67]. For instance, horizontal gene transfer contributed to host range expansion of the *Metarhizium* genus [22]. Also, gene copy number variation is often driven by repetitive and transposable elements and contributes to pathogenicity and can provide high variability of diverging paralogs of effectors, for example in powdery mildew fungi [68]. These mechanisms can enable rapid adaptation to new hosts and contribute to host range expansion in some lineages. Nevertheless, gene presence/absence polymorphism and coding sequence changes can have detrimental effects, in particular in unstable environments. For instance, pathogen gene loss may be advantageous on host carrying a matching R protein, but detrimental otherwise. Because coding but not *cis*-regulatory mutations are sensitive to frameshift, coding sequence variation is more likely to be detrimental and pleiotropic. In addition, transcription factor binding sites are short in comparison with coding sequence and therefore more likely to be neofunctionalized. This is the reasoning behind the *cis*-regulatory hypothesis stating that mutations that alter the regulation of gene expression are more likely to contribute to phenotypic evolution [69]. In agreement with this, we provide evidence that *cis*-regulatory variation contributes to the evolution of camalexin responsiveness in *Sclerotinia*. *S. trifoliorum* is closely related to *S. sclerotiorum* but has only been reported on plants from the *Fabaceae* family and rare cases on hosts from the *Asteraceae* and *Plantaginaceae* [55, 70]. All *Sclerotiniaceae* species and strains we tested were highly sensitive to camalexin, except for *S. sclerotiorum* strains. In spite of a high level of similarity in the genome of *S. sclerotiorum* strain 1980 and *S. trifoliorum* strain SwB9, their transcriptomes upon camalexin treatment appeared clearly distinct. Our transcriptomic analysis on camalexin focused on the respective highest tolerable concentration for each of the two fungi in order to ensure a comparable physiological state amid camalexin pressure. *S. sclerotiorum* generally displays little host preference [71] and further investigations will be required to determine the extent of variation in camalexin sensitivity and transcriptomic response to camalexin at the intraspecific level. Our findings highlight regulatory variation as an adaptive strategy for fungal pathogens jumping to new host plants. We propose that a certain degree of transcriptional plasticity or stochasticity for genes encoding promiscuous enzymes, conserved effectors, and their regulators, enable pathogen survival as endophytes on non-host plants. The persistence of a founder endophyte population may lead to the fixation of alleles with an adaptive expression pattern enabling the emergence of a pathogenic lifestyle on new hosts. Importantly, this implies that the gene pool the common ancestor of both *Sclerotinia* species was pre-equipped to adapt on non-host plants, and that a certain degree of fluctuations and randomness in gene regulation could enable the emergence of new traits. In this scenario, host range expansion does not require the acquisition of new genes from horizontal gene transfer or outcrossing. In the context of the impact of plant pathogens on food security, this observation calls for monitoring not only crop pathogen populations but also their close relatives infecting wild species in the same environment.

Our current analysis suggests that promiscuous enzymes with a flexible transcriptional pattern in *Sclerotinia* mediate the detoxification of host defense metabolites in the appropriate context. In addition, conserved "core" effectors with conserved plant targets could enable host jumps or host range expansions (e.g. [2]), followed by or in parallel with transcriptional plasticity. We discovered that the 70 *S. sclerotiorum* genes induced on camalexin and *A. thaliana* contain one predicted effector (Sscle_05g046060) and 23 secreted proteins (Dataset S1: Table S13). Among them, 22 genes are conserved in *S. trifoliorum* SwB9, 20 of which are conserved in over 100 fungal species, probably being part of a core secretome. Current functional information is not sufficient to determine whether these genes harbor a conserved function, contribute to *Sclerotinia* virulence, or act on plant antifungal chemicals. Polymorphisms in the amino acid sequence of core secreted proteins could lead to functional innovations enabling for instance the detoxification of a broader range of compounds in the context of host range expansion and speciation, providing a complementary adaptive mechanism to the regulatory variation uncovered in this work.

S. Kusch et al.

The current study focused on *S. sclerotiorum* infection of *A. thaliana* and relationship to its typical secondary metabolites, indole glucosinolates and camalexin, owing to the tractability of this system. We also found brassinin to be more toxic on *S. trifoliorum* than on *S. sclerotiorum*. *S. sclerotiorum* uses the brassinin glucosyltransferase SsBGT1 (*Sscle01g003110*) to detoxify this defense compound [59], a gene conserved in *S. trifoliorum* (*SCLTRI_002931*). In our assays, SsBGT1 was induced on all plants except *S. lycopersicum*. *S. sclerotiorum* SsBGT1 and its *S. trifoliorum* ortholog were induced to similar levels on camalexin (LFC 1.31 and 1.27 respectively). This data made brassinin a less straightforward determinant of *Sclerotinia* host range than camalexin. Nevertheless, future studies on the evolution of brassinin detoxification in the *Sclerotiniacea* should provide insights into host adaptation in these fungal pathogens.

Besides *A. thaliana*, *S. sclerotiorum* exhibited clear host-specific transcriptomes on castor bean, sugar beet, and common bean. On castor bean and sugar beet, genes related to the detoxification of host compounds were enriched in the specific transcriptomes. This may reflect the specific response to phytotoxins common to these host plants, such as the red beet antimicrobial phenolic secondary metabolites known as betalains [72] and the ribosome-inactivating protein ricin from castor bean [73]. By contrast, tomato, sunflower, and common bean induced limited host-specific transcriptomes. Flavonoids are common antimicrobial phytotoxins produced by many vascular plants and abundant in sunflower and beans for example [74]. While there are more than 10,000 flavonoid structures, *S. sclerotiorum* may use a conserved pathway to tackle these toxins. Indeed, the quercetin dioxygenase SsQDO (*Sscle07g059700*) catalyzes the cleavage of the flavonol carbon skeleton, thus targeting a range of flavonoids [75].

Our promoter region analyses identified a motif enriched in the *cis*-elements of *S. sclerotiorum* genes but not their orthologs in *S. trifoliorum*. In baker's yeast, the motif is recognized by the zinc-finger transcriptional regulators ADR1 and Mig1-3, which have two likely orthologues in *Sclerotinia* and *Botrytis*. Mig1-3 are orthologous to the *Aspergillus* carbon catabolite repressor CreA, which regulates plant cell wall-degrading enzymes in *A. nidulans* [76] and human mycosis disease in *A. fumigatus* [77]. *A. flavus creA* mutants are defective in aflatoxin biosynthesis and crop colonization [78]. In the apple blue mold pathogen *Penicillium expansum*, CreA acts as positive regulator of the mycotoxins patulin and citrinin, and *P. expansum creA* mutants are near-avirulent on apples [79]. The *S. sclerotiorum* SsCreA (*Sscle01g002690*) could therefore be involved in the tight host-specific regulation of mycotoxin biosynthesis and plant cell wall-degrading enzymes. The targets of these transcription factors in *S. sclerotiorum* could direct the appropriate response to plant defense compounds, including activating the detoxification machinery for glucosinolates. In addition, DNA-binding proteins may have acquired the capability of binding the WWCCCCRC motif in *S. sclerotiorum*. For example, the putative C2H2 fungal transcription factor Sscle07g059580 was induced both by camalexin and on *A. thaliana* and might have evolved a specific role in the response to *Brassicaceae*. The evolution of camalexin responsiveness may therefore have involved *cis*- and *trans*-regulation. Further analysis of the transcription factors SsCreA, SsADR1, and Sscle07g059580 regarding their binding efficacy of the WWCCCCRC motif, the repertoire of their genomic targets, and their specific function in response to *Brassicaceae* is required to answer these questions.

Our findings reveal that host range expansion can be supported by regulatory variation in genes conserved in related non-adapted fungal pathogen species. The genetic requirements for such an evolutionary trajectory need to be determined in order to identify nonpathogenic species with a high potential for jumping to new hosts through regulatory variation. One explanation could be that core effectors, effector innovation, and regulatory features contribute to various degrees to host jumps and expansions. This mechanism enables the emergence of new disease with no or limited gene flow between strains and species, and could underlie the emergence of new epidemics originating from wild plants in agricultural settings.

REFERENCES
1. Troch V, Audenaert K, Wyand RA, Haesaert G, Höfte M, Brown JKM. *Formae speciales* of cereal powdery mildew: close or distant relatives? Mol Plant Pathol. 2014;15:304–14.
2. Thines M. An evolutionary framework for host shifts—jumping ships for survival. New Phytol. 2019;224:605–17.
3. Nylin S, Janz N. Butterfly host plant range: an example of plasticity as a promoter of speciation? Evolut Ecol. 2009;23:137–46.
4. Flor HH. Current status of the gene-for-gene concept. Annu Rev Phytopathol. 1971;9:275–96.
5. Jones JDG, Dangl JL. The plant immune system. Nature. 2006;444:323–9.
6. Sánchez-Vallet A, Fouché S, Fudal I, Hartmann FE, Soyer JL, Tellier A, et al. The genome biology of effector gene evolution in filamentous plant pathogens. Annu Rev Phytopathol. 2018;56:21–40.
7. Möller M, Stukenbrock EH. Evolution and genome architecture in fungal plant pathogens. Nat Rev Microbiol. 2017;15:756–71.
8. Raffaele S, Kamoun S. Genome evolution in filamentous plant pathogens: why bigger can be better. Nat Rev Microbiol. 2012;10:417–30.
9. Dong S, Stam R, Cano LM, Song J, Sklenar J, Yoshida K, et al. Effector specialization in a lineage of the Irish potato famine pathogen. Science. 2014;343:552–5.
10. Dong S, Raffaele S, Kamoun S. The two-speed genomes of filamentous pathogens: Waltz with plants. Curr Opin Genet Dev. 2015;35:57–65.
11. Raffaele S, Farrer RA, Cano LM, Studholme DJ, MacLean D, Thines M, et al. Genome evolution following host jumps in the Irish potato famine pathogen lineage. Science. 2010;330:1540–3.
12. Thrall PH, Laine A-L, Ravensdale M, Nemri A, Dodds PN, Barrett LG, et al. Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. Ecol Lett. 2012;15:425–35.
13. Janzen DH. When is it coevolution? Evolution. 1980;34:611–2.
14. Ebert D, Fields PD. Host–parasite co-evolution and its genomic signature. Nat Rev Genet. 2020;21:754–68.
15. West-Eberhard MJ. Phenotypic plasticity and the origins of diversity. Annu Rev Ecol Syst. 1989;20:249–78.
16. Hemetsberger C, Mueller AN, Matei A, Herrberger C, Hensel G, Kumlehn J, et al. The fungal core effector Pep1 is conserved across smuts of dicots and monocots. New Phytol. 2015;206:1116–26.
17. Simon J-C, D'Alencon E, Guy E, Jacquin-Joly E, Jaquiery J, Nouhaud P, et al. Genomics of adaptation to host-plants in herbivorous insects. Brief Funct Genomics. 2015;14:413–23.
18. Baroncelli R, Amby DB, Zapparata A, Sarrocco S, Vannacci G, Le Floch G, et al. Gene family expansions and contractions are associated with host range in plant pathogens of the genus. Colletotrichum BMC Genomics. 2016;17:555.
19. Wyka SA, Mondo SJ, Liu M, Dettman J, Nalam V, Broders KD. Whole-genome comparisons of ergot fungi reveals the divergence and evolution of species within the genus *Claviceps* are the result of varying mechanisms driving genome evolution and host range expansion. Genome Biol Evol. 2021;13: evaa267.
20. Laurent B, Palaiokostas C, Spataro C, Moinard M, Zehraoui E, Houston RD, et al. High-resolution mapping of the recombination landscape of the phytopathogen *Fusarium graminearum* suggests two-speed genome evolution. Mol Plant Pathol. 2018;19:341–54.
21. Hu X, Xiao G, Zheng P, Shang Y, Su Y, Zhang X, et al. Trajectory and genomic determinants of fungal-pathogen speciation and host adaptation. Proc Natl Acad Sci USA. 2014;111:16796–801.
22. Zhang Q, Chen X, Xu C, Zhao H, Zhang X, Zeng G, et al. Horizontal gene transfer allowed the emergence of broad host range entomopathogens. Proc Natl Acad Sci USA. 2019;116:7982–9.
23. Zhang S, Gu S, Ni X, Li X. Genome size reversely correlates with host plant range in *Helicoverpa* species. Front Physiol. 2019;10:29.

24. Mathers TC, Chen Y, Kaithakottil G, Legeai F, Mugford ST, Baa-Puyoulet P, et al. Rapid transcriptional plasticity of duplicated gene clusters enables a clonally reproducing aphid to colonise diverse plant species. Genome Biol. 2017;18:27.

25. Boland GJ, Hall R. Index of plant hosts of *Sclerotinia sclerotiorum*. Can J Plant Pathol. 1994;16:93–108.

26. Derbyshire MC, Denton-Giles M, Hane JK, Chang S, Mousavi-Derazmahalleh M, Raffaele S, et al. A whole genome scan of SNP data suggests a lack of abundant hard selective sweeps in the genome of the broad host range plant pathogenic fungus *Sclerotinia sclerotiorum*. PLoS ONE. 2019;14:e0214201.

27. Derbyshire MC, Denton-Giles M, Hegedus DD, Seifbarghi S, Rollins JA, Van Kan JAL, et al. The complete genome sequence of the phytopathogenic fungus *Sclerotinia sclerotiorum* reveals insights into the genome architecture of broad host range pathogens. Genome Biol Evol. 2017;9:593–618.

28. Peyraud R, Mbengue M, Barbacci A, Raffaele S. Intercellular cooperation in a fungal plant pathogen facilitates host colonization. Proc Natl Acad Sci USA. 2019;116:3193–201.

29. Badet T, Peyraud R, Mbengue M, Navaud O, Derbyshire MC, Oliver RP, et al. Codon optimization underpins generalist parasitism in fungi. eLife. 2017;6:e22472.

30. Guyon K, Balagué C, Roby D, Raffaele S. Secretome analysis reveals effector candidates associated with broad host range necrotrophy in the fungal plant pathogen *Sclerotinia sclerotiorum*. BMC Genomics. 2014;15:336.

31. Derbyshire M, Mbengue M, Barascud M, Navaud O, Raffaele S. Small RNAs from the plant pathogenic fungus *Sclerotinia sclerotiorum* highlight host candidate genes associated with quantitative disease resistance. Mol Plant Pathol. 2019;20:1279–97.

32. Ibrahim HMM, Kusch S, Didelon M, Raffaele S. Genome-wide alternative splicing profiling in the fungal plant pathogen *Sclerotinia sclerotiorum* during the colonization of diverse host families. Mol Plant Pathol. 2021;33:880–3.

33. Sucher J, Mbengue M, Dresen A, Barascud M, Didelon M, Barbacci A, et al. Phylotranscriptomics of the Pentapetalae reveals frequent regulatory variation in plant local responses to the fungal pathogen *Sclerotinia sclerotiorum*. Plant Cell. 2020;32:1820–44.

34. Vleugels T, Baert J, van Bockstaele E. Morphological and pathogenic characterization of genetically diverse *Sclerotinia* isolates from European red clover crops (*Trifolium pratense* L.). J Phytopathol. 2013;161:254–62.

35. Barbacci A, Navaud O, Mbengue M, Barascud M, Godiard L, Khafif M, et al. Rapid identification of an Arabidopsis NLR gene as a candidate conferring susceptibility to *Sclerotinia sclerotiorum* using time-resolved automated phenotyping. Plant J. 2020;103:903–17.

36. Zhao Y, Hull AK, Gupta NR, Goss KA, Alonso JM, Ecker JR, et al. Trp-dependent auxin biosynthesis in *Arabidopsis*: involvement of cytochrome P450s CYP79B2 and CYP79B3. Genes Dev. 2002;16:3100–12.

37. Zhou N, Tootle TL, Glazebrook J. Arabidopsis *PAD3*, a gene required for camalexin biosynthesis, encodes a putative cytochrome P450 monooxygenase. Plant Cell. 1999;11:2419–28.

38. Kusch S, Ibrahim HMM, Zanchetta C, Lopez-Roques C, Donnadieu C, Raffaele S. A chromosome-scale genome assembly resource for *Myriosclerotinia sulcatula* infecting sedge grass (*Carex* sp.). Mol Plant-Microbe Interact. 2020;33:880–3.

39. Trapnell C, Williams BA, Pertea GM, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28:511–5.

40. Law CW, Alhamdoosh M, Su S, Smyth GK, Ritchie ME. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. F1000Research. 2016;5:1408.

41. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2018.

42. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27:722–36.

43. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE. 2014;9:e112963.

44. Laetsch DR, Blaxter ML. BlobTools: Interrogation of genome assemblies. F1000Research. 2017;6:1287.

45. Smit AFA, Hubley R, Green P. *RepeatMasker Open-4.0*. 2013-2015. http://www.repeatmasker.org.

46. Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M. BRAKER1: unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. Bioinformatics. 2016;32:767–9.

47. Lee E, Helt GA, Reese JT, Munoz-Torres MC, Childers CP, Buels RM, et al. Web Apollo: a web-based genomic annotation editing platform. Genome Biol. 2013;14:R93.

48. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu CM, et al. Versatile and open software for comparing large genomes. Genome Biol. 2004;5:R12.

49. Guy L, Roat Kultima J, Andersson SGE. genoPlotR: comparative gene and genome visualization in R. Bioinformatics. 2010;26:2334–5.

50. Van Kan JAL, Stassen JHM, Mosbach A, Van Der Lee TAJ, Faino L, Farmer AD, et al. A gapless genome sequence of the fungus *Botrytis cinerea*. Mol Plant Pathol. 2017;18:75–89.

51. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20:238.

52. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26:841–2.

53. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37:W202–208.

54. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. Nucleic Acids Res. 2015;43:W39–49.

55. Navaud O, Barbacci A, Taylor A, Clarkson JP, Raffaele S. Shifts in diversification rates and host jump frequencies shaped the diversity of host range among *Sclerotiniaceae* fungal plant pathogens. Mol Ecol. 2018;27:1309–23.

56. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31:3210–2.

57. Sperschneider J, Dodds PN, Gardiner DM, Singh KB, Taylor JM. Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. Mol Plant Pathol. 2018;19:2094–110.

58. Chen J, Ullah C, Reichelt M, Beran F, Yang Z-L, Gershenzon J, et al. The phytopathogenic fungus *Sclerotinia sclerotiorum* detoxifies plant glucosinolate hydrolysis products via an isothiocyanate hydrolase. Nat Commun. 2020;11:3090.

59. Sexton AC, Minic Z, Cozijnsen AJ, Pedras MSC, Howlett BJ. Cloning, purification and characterisation of brassinin glucosyltransferase, a phytoalexin-detoxifying enzyme from the plant pathogen *Sclerotinia sclerotiorum*. Fungal Genet Biol. 2009;46:201–9.

60. Stotz HU, Sawada Y, Shimada Y, Hirai MY, Sasaki E, Krischke M, et al. Role of camalexin, indole glucosinolates, and side chain modification of glucosinolate-derived isothiocyanates in defense of *Arabidopsis* against *Sclerotinia sclerotiorum*. Plant J. 2011;67:81–93.

61. Petre B, Lorrain C, Stukenbrock EH, Duplessis S. Host-specialized transcriptome of plant-associated organisms. Curr Opin Plant Biol. 2020;56:81–88.

62. Lysøe E, Seong K-Y, Kistler HC. The transcriptome of *Fusarium graminearum* during the infection of wheat. Mol Plant-Microbe Interact. 2011;24:995–1000.

63. Harris LJ, Balcerzak M, Johnston A, Schneiderman D, Ouellet T. Host-preferential *Fusarium graminearum* gene expression during infection of wheat, barley, and maize. Fungal Biol. 2016;120:111–23.

64. Kellner R, Bhattacharyya A, Poppe S, Hsu TY, Brem RB, Stukenbrock EH. Expression profiling of the wheat pathogen *Zymoseptoria tritici* reveals genomic patterns of transcription and host-specific regulatory programs. Genome Biol Evol. 2014;6:1353–65.

65. Allan J, Regmi R, Denton-Giles M, Kamphuis LG, Derbyshire MC. The host generalist phytopathogenic fungus *Sclerotinia sclerotiorum* differentially expresses multiple metabolic enzymes on two different plant hosts. Sci Rep. 2019;9:19966.

66. Gluck-Thaler E, Slot JC. Dimensions of horizontal gene transfer in eukaryotic microbial pathogens. PLOS Pathog. 2015;11:e1005156.

67. Menardo F, Praz CR, Wyder S, Ben-David R, Bourras S, Matsumae H, et al. Hybridization of powdery mildew strains gives rise to pathogens on novel agricultural crop species. Nat Genet. 2016;48:201–5.

68. Frantzeskakis L, Kracher B, Kusch S, Yoshikawa-Maekawa M, Bauer S, Pedersen C, et al. Signatures of host specialization and a recent transposable element burst in the dynamic one-speed genome of the fungal barley powdery mildew pathogen. BMC Genomics. 2018;19:381.

69. Stern DL, Orgogozo V. The loci of evolution: How predictable is genetic evolution? Evolution. 2008;62:2155–77.

70. Farr DF, Rossman AY. Fungal databases, U.S. National Fungus Collections, ARS, USDA. 2021. https://nt.ars-grin.gov/fungaldatabases/.

71. Leyronas C, Bardin M, Berthier K, Duffaud M, Troulet C, Torres M, et al. Assessing the phenotypic and genotypic diversity of *Sclerotinia sclerotiorum* in France. Eur J Plant Pathol. 2018;152:933–44.

72. Kumar S, Brooks MS-L. Use of red beet (*Beta vulgaris* L.) for antimicrobial applications - a critical review. Food Bioprocess Technol. 2018;11:17–42.

73. Polito L, Bortolotti M, Battelli MG, Calafato G, Bolognesi A. Ricin: an ancient story for a timeless plant toxin. Toxins. 2019;11:324.

74. Weston LA, Mathesius U. Flavonoids: their structure, biosynthesis and role in the rhizosphere, including allelopathy. J Chem Ecol. 2013;39:283–97.

75. Chen J, Ullah C, Reichelt M, Gershenzon J, Hammerbacher A. *Sclerotinia sclerotiorum* circumvents flavonoid defenses by catabolizing flavonol glycosides and aglycones. Plant Physiol. 2019;180:1975–87.

76. de Vries RP, Visser J. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microbiol Mol Biol Rev. 2001;65:497–522.

77. Beattie SR, Mark KMK, Thammahong A, Ries LNA, Dhingra S, Caffrey-Carr AK, et al. Filamentous fungal carbon catabolite repression supports metabolic plasticity and stress responses essential for disease progression. PLOS Pathog. 2017;13: e1006340.

78. Fasoyin OE, Wang B, Qiu M, Han X, Chung K-R, Wang S. Carbon catabolite repression gene *creA* regulates morphology, aflatoxin biosynthesis and virulence in *Aspergillus flavus*. Fungal Genet Biol. 2018;115:41–51.

79. Tannous J, Kumar D, Sela N, Sionov E, Prusky D, Keller NP. Fungal attack and host defence pathways unveiled in near-avirulent interactions of *Penicillium expansum creA* mutants on apples. Mol Plant Pathol. 2018;19:2635–50.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

SK and SR designed the project. RNA-seq sampling was performed by JL, ON, MM, and SK. SK and SR performed data analysis, statistical analysis, and plotting. Genome nanopore sequencing was done by CZ, CR, and CD, assembly of S. *trifoliorum* SwB9 v1 genome by SK, and gene annotation and manual curation by SK and HMMI. SK, SM, and LG conducted the in vitro camalexin tolerance assay for various *Sclerotinia* strains. Further experiments and pathogen assays were performed by SK, SR, JL, and NG. SK and SR drafted the manuscript, all authors edited and proofread the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41396-021-01058-x.

Correspondence and requests for materials should be addressed to S.R.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.