

Article

Identification and comparison of *Colletotrichum* secreted effector candidates reveal two independent lineages pathogenic to soybean

Thaís R. Bouffleur^{1,2,*}, Nelson S. Massola Júnior¹, Ísis Tikami¹, Serenella A. Sukno², Michael R. Thon² and Riccardo Baroncelli^{2,3,*}

¹ University of São Paulo (USP), Luiz de Queiroz College of Agriculture (ESALQ), Piracicaba, 13418-900, São Paulo, Brazil

² University of Salamanca, Institute for Agrobiotechnology Research (CIALE), Department of Microbiology and Genetics, Villamayor, 37185, Salamanca, Spain

³ University of Bologna, Department of Agricultural and Food Sciences (DISTAL), viale Fanin 44, 40126 Bologna, Italy

* Correspondence: thaïsboffleur@gmail.com and riccardobaroncelli@gmail.com

Abstract: *Colletotrichum* is one of the most important plant pathogenic genera of fungi due to its scientific and economic impact. *Colletotrichum* spp. can infect a wide range of hosts, causing losses in crops of major importance worldwide, such as soybean. In the past, soybean anthracnose was mainly caused by *C. truncatum*, but during the last decade, other species have been identified at an increasing rate, becoming one of the most important limiting factors to soybean production in several regions. To gain a better understanding of the evolutionary origin of soybean anthracnose, we compared the repertoire of effector candidates of four *Colletotrichum* species pathogenic to soybean and eight pathogens of other hosts. Our results show that the four species infecting soybean belong to two lineages and do not share any of the lineage specific effector candidates identified. These results strongly suggest that two *Colletotrichum* lineages have acquired the capability to infect soybean independently. This study also provides, for each lineage, a set of candidate effectors encoding genes that may have important roles in pathogenicity towards soybean offering a new resource useful for further research on soybean anthracnose management.

Keywords: anthracnose; genome sequencing; pathogenicity factors, *Colletotrichum truncatum*, *Colletotrichum orchidearum*, *Glomerella*, *Glycine max*.

1. Introduction

With 257 accepted species classified into species complexes (s.c.) or singletons [1], *Colletotrichum* is considered among the ten most destructive genera of phytopathogenic fungi [2], responsible for losses in many important cultivated crops [3–7]. Several species of *Colletotrichum* have been reported as pathogenic to soybean, being the *C. orchidearum* and *C. truncatum* s.c. found as the complexes with most of the available data associated with symptomatic soybean plants worldwide [8]. While *C. truncatum* is associated with soybean since 1917 [9], *C. musicola*, *C. plurivorum* and *C. sojae*, members of the *C. orchidearum* s.c. were only described in the past 5 years [10–12], however this species complex is misidentified at least since 2003 [8].

A dispute for survival and adaptation marks the evolutionary battle between plants and pathogens throughout history [13–15]. This arms race can be partially described by the “zig-zag” model [13], where the first layer of defense of plants recognizes molecular patterns associated with pathogens (PAMPs) or damage-associated molecular patterns (DAMPs) and active a pattern triggered immune response (PTI) [13,16,17]. On the other

hand, pathogens can overcome this layer of defense releasing effectors, that are secreted proteins that cause alterations in structure or processes of the host cell, suppressing the defense responses or enhancing access to nutrients, promoting the colonization of the host by the pathogen [18]. The recognition of effectors or effector targets by resistance (*R*) genes of the host will then trigger the second layer of defense, called effector-triggered immunity (ETI), being a stronger response than PTI that can lead to a hypersensitive reaction (HR) [14,18]. With the advance of molecular studies, it was shown that the division among PTI and ETI is blurred [19–21], and now an integrated plant immune system has been proposed, where a crosstalk between plant immune receptors is essential to both, PTI and ETI achieve its maximum immune response [22].

Over the past few years, the genomes of at least 43 species of *Colletotrichum* have been sequenced [23–47] (<http://www.colletotrichum.org/genomics/>), including *C. truncatum*, *C. musicola*, *C. plurivorum* and *C. sojae* [45]. The availability of genome sequences from multiple species of *Colletotrichum* enables unprecedented insights into genome composition [48]. An understanding of the pathogenicity mechanisms of *Colletotrichum* and the way that they adapt to their hosts can be a powerful tool in developing sustainable control strategies [49–51]. It is known that the evolution through adaptation of pathogens to different hosts can involve sets of effectors, that can specialize to infect a specific host [52–56], therefore the evolutionary trajectory of host-pathogen interactions can help to clarify the mechanisms underlying the threat of pathogens to crops [57].

The identification of effector candidates is the first step into the functional characterization of these molecules. Until now, several studies on effectors of different species of *Colletotrichum* such as *C. higginsianum* [58,59] *C. orbiculare* [60,61] *C. lentis* [62–64], *C. graminicola* [65–67] *C. simmondsii*, *C. fiorinae*, *C. nymphaeae*, *C. salicis* [30], *C. lindemuthianum* [68], *C. falcatum* [69], *C. fruticola*, *C. siamense*, *C. aenigma*, *C. tropicale*, *C. viniferum* [44] have been published. On the other hand, comparative genomic studies of *Colletotrichum* spp. that infect soybean have not been performed and the number of candidate effectors of *C. truncatum*, *C. plurivorum*, *C. musicola* and *C. sojae*, and how many are unique to these species is unknown. A compilation of candidate effectors of those species may help to identify determinants of host specificity in the *Colletotrichum*-soybean interaction as well as better understanding the mechanisms underlying soybean infection.

To gain a better understanding of the evolutionary origin of soybean anthracnose, we analyzed the repertoire of Lineage Specific Effector Candidates (LSECs) defined as secreted proteins that have no homology to any other protein or that have homology to proteins from other members of the same genus, species or species complex [30]. We analyzed the proteomes of four species of *Colletotrichum* pathogenic to soybean and compared these with eight closely related species of *Colletotrichum* non-pathogenic to this host providing a useful platform for future works regarding soybean anthracnose.

2. Results

2.1 Among the selected *Colletotrichum* species, only *C. truncatum* and members of the *C. orchidearum* s.c. are pathogenic to soybean

The pathogenicity of 10 *Colletotrichum* species selected for comparative genomic analyzes (Table 1) were tested on soybean.

Table 1: *Colletotrichum* strains used in the pathogenicity test and comparative genomics analysis

Strain	Species	Species complex	Host	Origin	*
MAFF 240422	<i>C. orbiculare</i>	<i>C. orbiculare</i>	<i>Cucumis sativus</i>	Japan	[38]
LFN0074	<i>C. musicola</i>	<i>C. orchidearum</i>	<i>Glycine max</i>	Brazil	[45]
LFN00145	<i>C. plurivorum</i>	<i>C. orchidearum</i>	<i>Glycine max</i>	Brazil	[45]

LFN0009	<i>C. sojae</i>	<i>C. orchidearum</i>	<i>Glycine max</i>	Brazil	[45]
1059	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	[45]
Cg-14	<i>C. gloeosporioides</i> s.s.	<i>C. gloeosporioides</i>	<i>Persea americana</i>	Israel	[25]
IMI 349063	<i>C. higginsianum</i>	<i>C. destructivum</i>	<i>Brassica rapa</i>	Trinidad & Tobago	[70]
CBS 168.49	<i>C. tofieldiae</i>	<i>C. spaethianum</i>	<i>Lupinus polyphyllus</i>	Germany	[71]
M1.001	<i>C. graminicola</i>	<i>C. graminicola</i>	<i>Zea mays</i>	USA	[23]
IMI 309357	<i>C. orchidophilum</i>	none	<i>Phalaenopsis</i> sp.	United Kingdom	[36]
IMI 504882	<i>C. fioriniae</i>	<i>C. acutatum</i>	<i>Fragaria x ananassa</i>	New Zealand	[29]
IMI 504889	<i>C. nymphaeae</i>	<i>C. acutatum</i>	<i>Fragaria x ananassa</i>	Denmark	[30]

* Reference of the genome sequences

Assays confirmed that only *C. truncatum* and the three species belonging to the *C. orchidearum* s.c., *C. musicola*, *C. plurivorum* and *C. sojae* cause anthracnose symptoms in soybean, of which *C. truncatum* is the most virulent to the tested soybean cultivar (Monsoy IPRO7739) than the three species belonging to the *C. orchidearum* s.c. (Figure 1). *Colletotrichum gloeosporioides*, *C. higginsianum*, *C. tofieldiae*, *C. orchidophilum*, *C. fioriniae* and *C. nymphaeae* were not pathogenic to soybean (Figure 1).

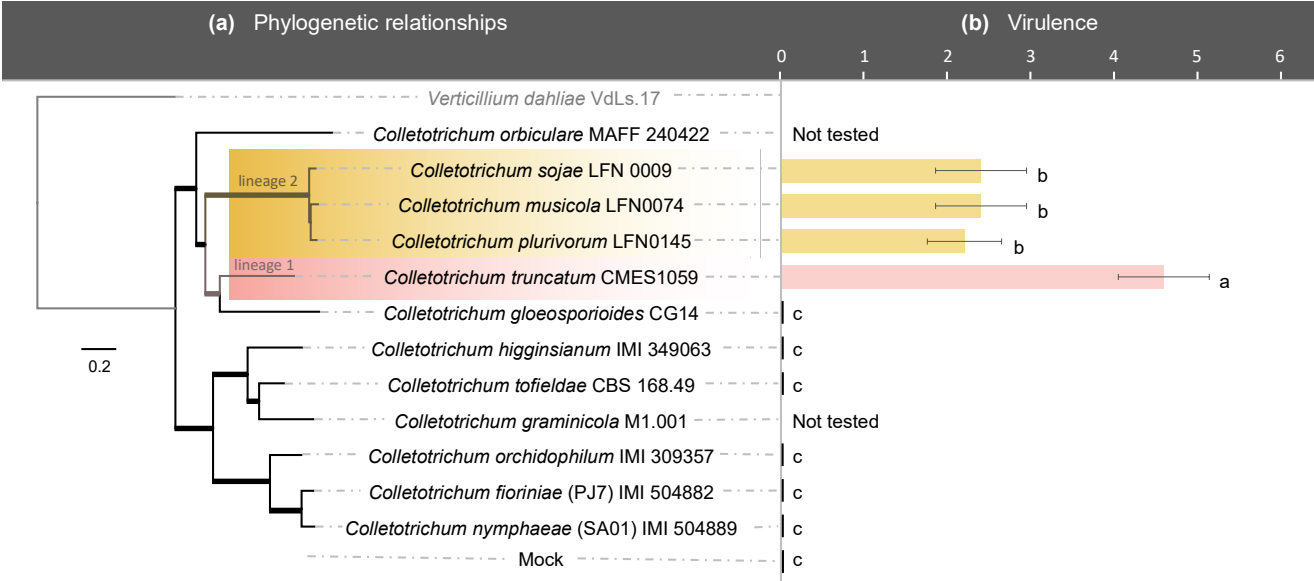


Figure 1. Evolutionary relationships of *Colletotrichum* species. **(a)** Bayesian inference phylogenetic of the strains used in this study. The tree was reconstructed from concatenated nucleotide alignments of the *act* (actin), *chs-1* (chitin synthase), and *gapdh* (glyceraldehyde 3-phosphate dehydrogenase) genes. For each locus the alignment was performed with MAFFT v7.450 (Kato and Standley 2013), exported to MEGA7 (Kumar et al. 2016) and the best-fit substitution model was calculated. Thicker branches represent nodes with bayesian posterior probability equal to 1.00. The scale bar represents the number of expected substitutions per site. **(b)** Level of virulence of *Colletotrichum* species to soybean. Tuckey's test was applied on transformed data $((X+1)^{0.5})$. Equal letters do not differ in the average of virulence among *Colletotrichum* strains in the Tuckey test with p -value = 0.05%. Species belonging to the *C. orchidearum* species complex (s.c) are represented by yellow bars, while *C. truncatum* is represented by the pink bar.

2.2 The majority of candidate effectors of Colletotrichum species are conserved

To better understand the evolution of the two main *Colletotrichum* s. c. that infect soybean worldwide (*C. truncatum* s. c. and *C. orchidearum* s. c.) [8] we conducted *in silico* analyzes (Figure S1) to check if the representative species of those complexes (*C. truncatum*, *C. musicola*, *C. plurivorum* and *C. sojae*) share a unique set of effector candidates.

The proteomes of the 12 *Colletotrichum* species (Table 1) were assigned to 32,018 orthogroups, of which 7,428 are shared among all the proteomes analyzed (Figure 3B). Comparative analysis identified 66 orthogroups comprising 338 genes of *Colletotrichum* spp. common only to the four species infecting soybean, of which only one orthogroup is fully secreted; and 764 orthogroups (2454 genes) shared only between the species belonging to the *C. orchidearum* s.c., of which eight orthogroups are secreted. While 1,214 (1,695 genes); 1,103 (1,126 genes); 760 (771 genes) and 943 (952 genes) orthogroups were specific to *C. truncatum*, *C. musicola*, *C. plurivorum* and *C. sojae*, respectively (Figure 3C).

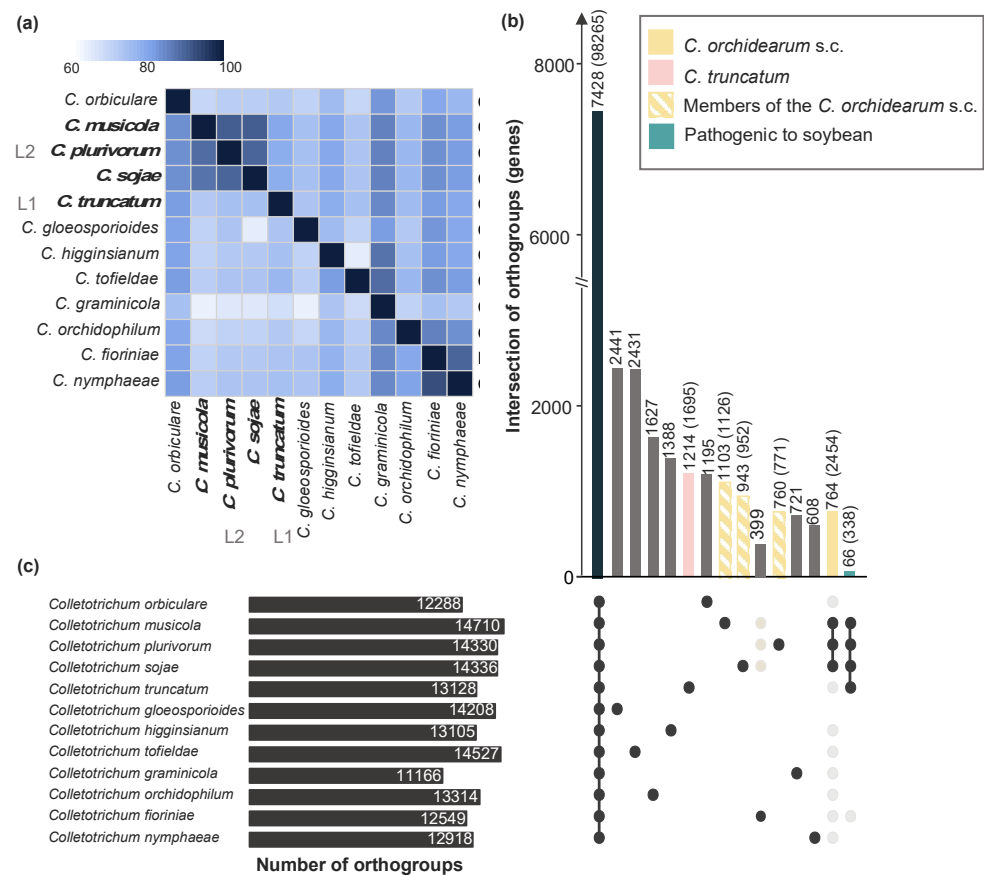


Figure 2. Comparative genomic analysis of *Colletotrichum* species pathogenic and not pathogenic to soybean. Species highlighted in yellow represent the *C. orchidearum* species complex (s.c.), and striped yellow bars correspond to each species belonging to the *C. orchidearum* s.c. (*C. musicola*, *C. plurivorum* and *C. sojae*); while *C. truncatum* is represented in pink. (a) Heatmap showing the percentage of overlapping proteins shared in pairwise comparisons (values correspond to percentage of proteins encoded by the species reported in the y axis that show similarity with those reported by the species reported in the x axis). (b) UpsetR plot of the protein clustering analysis of 12 *Colletotrichum* species. Bars on the upper side represent the number of orthogroups shared by the

species highlighted by the black dots reported on the bottom side. The number of genes corresponding to the orthogroups is in parentheses. (c) Species compared in this study, the bars on the right side of the species name represent the total number of orthogroups in each proteome. L1 (lineage 1); L2 (lineage 2).

The proteomes of the four soybean infecting species of *Colletotrichum* were scanned for the presence of signal peptides, transmembrane (TM)-domains, and glycosylphosphatidylinositol (GPI)-anchors. For further analyses, the secretome of each *Colletotrichum* species was defined based on those proteins with a predicted signal peptide, and absence of TM domains [50] and GPI-anchors. The secretomes of the four species vary between 9-10%, being 1,638; 1,485; 1,495; and 1,447 proteins for *C. truncatum*, *C. musicola*, *C. plurivorum* and *C. sojæ* respectively (Table 2).

Table 2. Secretome size of the four species of *Colletotrichum* that infect soybean, compared in this study.

Species	Proteome	Signal peptide	Absence of TM/GPI anchor	% of secreted proteins
<i>C. truncatum</i>	15,901	2,116	1,638	10
<i>C. plurivorum</i>	15,153	1,989	1,495	10
<i>C. sojæ</i>	16,124	1,931	1,447	9
<i>C. musicola</i>	16,826	1,871	1,485	9

Our results revealed that most of the effector candidates of the four *Colletotrichum* species pathogenic to soybean are present in other microorganisms, corresponding to 80% of *C. truncatum*, 84% of *C. musicola*, 83%, of *C. plurivorum* and 85% of *C. sojæ*. While around 15% of the effector candidates of each species are shared only among the *Colletotrichum* genus (Figure 3). LSECs, with no similarity inside or outside the genus *Colletotrichum* were identified, among those, 11 *C. orchidearum*-LSECs in *C. plurivorum*, 13 *C. orchidearum*-LSECs in *C. musicola*; and 16 *C. orchidearum*-LSECs in *C. sojæ*. We also identified 40 *C. truncatum*-LSECs, 15 *C. musicola*-LSECs, eight *C. plurivorum*-LSECs and nine *C. sojæ*-LSECs. Host-LSECs shared only between the four *Colletotrichum* species that infect soybean were not identified (Figure 3, Table S1). All the sets of s.c. and species-LSECs aforementioned were assigned to their corresponding orthogroups based on the similarity analysis of the proteins (Table S1).

The absence of similarity to proteins with a known function is a common characteristic to effector proteins [72]. All the LSECs were scanned with RunIprScan to identify conserved domains and submitted to a BLAST against the non-redundant database Pathogen Host interactions-base (PHI-base) to check the similarity with known genes of other microorganism species. All LSECs of the four *Colletotrichum* spp. pathogenic to soybean do not have any known domain or similarity in PHI-base (Table S1).

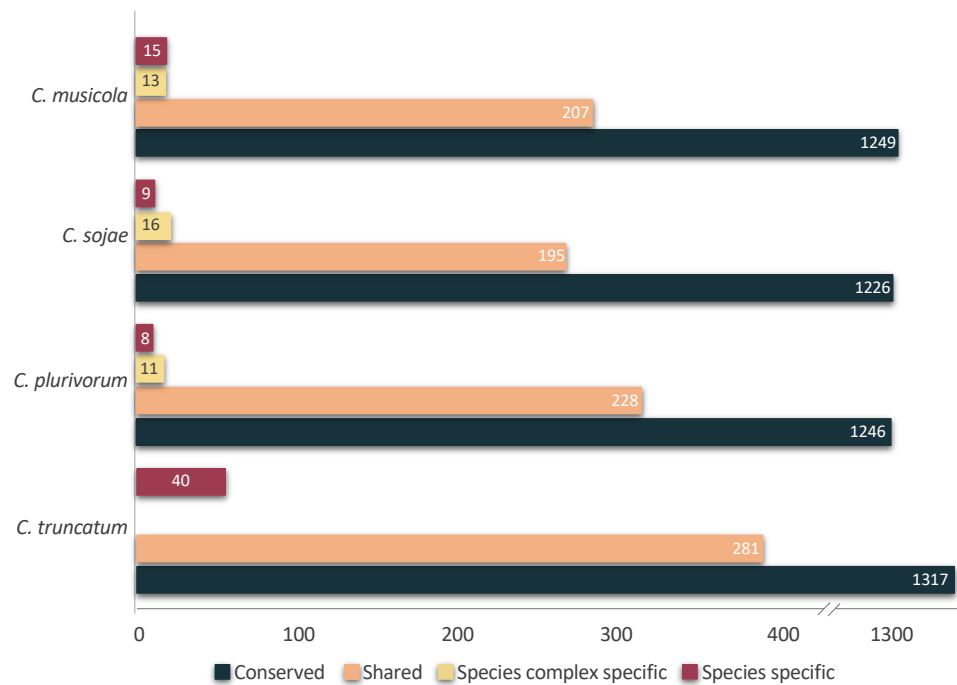


Figure 3. Effector candidates of *Colletotrichum* species pathogenic to soybean. Effector candidates of *C. musicola*, *C. plurivorum*, *C. sojae* and *C. truncatum* with similarity outside the genus *Colletotrichum* are representent in dark green, while effector candidates with similarity with other species of the genus are represented in orange. *C. orchidearum*-Lineage Specific Effector Candidates (LSECs) are represented in yellow and species-LSECs are represented in dark red. Total numbers of candidate effectors are represented in the bars.

We scanned the *C. orchidearum* s.c. LSECs for characteristics commonly observed in effector proteins, such as a high percentage of cysteines (cysteine-rich), with >2% of cysteines in their amino acid sequences [73], repeat-containing proteins [74] and the predicted translocation to different subcellular compartments of the plant cell, such as the chloroplast or mitochondria when they have a transit peptide, to the plant cell nucleus, when they possess nuclear localization signals (NLS) [75] or are delivered to the plant apoplast [76]. All *C. orchidearum*-LSECs, have at least one of the above-mentioned characteristics, from those, six; five; and five LSECs were predicted as effectors by EffectorP 2.0 tool for *C. musicola*, *C. plurivorum*, and *C. sojae*. Among the species-LSECs, 11 *C. musicola*, seven of *C. plurivorum* and eight of *C. sojae* have at least one of these characteristics, of which five, two and three were predicted to be effectors by EffectorP 2.0 tool. Among the *C. truncatum*-LSECs, 34 were predicted to have at least one of those characteristics, being 16 of them predicted by EffectorP 2.0 (Table 3, Table S1).

Table 3. Predicted *C. orchidearum* s.c. and species-LSECs of the four species of *Colletotrichum* pathogenic to soybean, containing characteristics commonly associated with effector proteins in fungi, and the total number of protein sequences predicted as effectors by EffectorP 2.0.

<i>C. orchidearum</i> s.c. SECs							
Species	LSECs	RCP	SL (NLS)	SL (other)	Apoplast	CR	EffectorP
<i>C. musicola</i>	13	5	0	0	9	7	6
<i>C. plurivorum</i>	11	5	1	0	7	6	5
<i>C. sojæ</i>	16	3	2	2	11	8	5
Species-LSECs							
<i>C. truncatum</i>	40	7	5	2	16	21	16
<i>C. musicola</i>	15	2	0	2	4	8	5
<i>C. plurivorum</i>	8	0	0	4	1	4	2
<i>C. sojæ</i>	9	2	0	1	2	6	3

LSECs: Lineage Specific Effector Candidates; RCP: repeat-containing proteins; SL: subcellular localization; NLS: nuclear localization signal; CR: cysteine-rich proteins (>2%); NA: not applicable.

2.3 *C. truncatum* LSECs are expressed and have evolutionary evidence

To confirm the expression of *C. truncatum*-LSECs *in vitro* and in soybean during the infection by *C. truncatum*, samples were collected for RNA sequencing at 12; 48 and 120 hpi, and 21 cDNA libraries were sequenced. A total of 1,202,535,286 raw reads were generated by Illumina HiSeq4000 sequencing. Overall, from 0,02 to 7,56% of the paired-end reads were mapped to the *C. truncatum* genome. 18 *C. truncatum* LSECs have evidence of expression *in planta* and/or *in vitro*. From those, nine are evolutionarily conserved in 18 *C. truncatum* genomes pathogenic to soybean. Another eight *C. truncatum* LSECs are conserved but are not expressed (Figure 4).

LSECs	Evidence of expression	Colletotrichum truncatum strain (accession number)																	
		GO2-03 (SRX7095352)	GO2-06 (SRX7095344)	GO2-12 (SRX7095346)	GO4-07 (SRX7095356)	GO4-08 (SRX7095338)	GO4-17 (SRX7095351)	GO5-11 (SRX7095354)	GO5-14 (SRX7095340)	GO5-25 (SRX7095342)	MT1-01 (SRX7095339)	MT2-05 (SRX7095348)	MT3-01 (SRX7095353)	MT3-21 (SRX7095343)	MT4-05 (SRX7095347)	MT4-13 (SRX7095341)	MT5-12 (SRX7095345)	MT5-26 (SRX7095349)	MT5-32 (SRX7095350)
CTRU02_00830																			
CTRU02_00831																			
CTRU02_00942																			
CTRU02_02134																			
CTRU02_02414																			
CTRU02_02635																			
CTRU02_04831																			
CTRU02_04889																			
CTRU02_05045																			
CTRU02_05072																			
CTRU02_05528																			
CTRU02_05538																			
CTRU02_05841																			
CTRU02_06815																			
CTRU02_07028																			
CTRU02_07468																			
CTRU02_08110																			
CTRU02_08233																			
CTRU02_08935																			
CTRU02_09022																			
CTRU02_09174																			
CTRU02_09190																			
CTRU02_09640																			
CTRU02_10342																			
CTRU02_10581																			
CTRU02_11186																			
CTRU02_11196																			
CTRU02_11952																			
CTRU02_11978																			
CTRU02_12587																			
CTRU02_13132																			
CTRU02_13374																			
CTRU02_13651																			
CTRU02_13804																			
CTRU02_13901																			
CTRU02_14087																			
CTRU02_14662																			
CTRU02_14861																			
CTRU02_15496																			
CTRU02_15750																			

Figure 4. The distribution of 58 LSECs in 18 strains of *Colletotrichum truncatum* and evidence of expression of *C. truncatum* (CMES1059). 18 genomes of *C. truncatum* pathogenic to soybean were scanned for presence/absence of the 58 *C. truncatum*-LSECs using BLAST. Red squares indicate the presence of LSECs by the blasting with coverage >90% and identity >60%. Green squares indicates evidence of expression of LSECs in soybean and/or *in vitro*.

3. Discussion

The availability of four representative *Colletotrichum* genomes of the *C. truncatum* s.c. and *C. orchidearum* s.c. [45] reported as the most distributed s.c. associated with soybean worldwide [8], along with the genomes of several *Colletotrichum* species associated with other hosts [23–46], allowed us to investigate the evolutionary origin of soybean anthracnose, by looking at the repertoire of effector candidates of each species and comparing them with the proteomes of eight additional *Colletotrichum* species non-pathogenic to soybean.

Effectors proteins produced by plant pathogens are secreted proteins, many of which translocated to the apoplast or cytoplasm of the host, where they alter the host defense responses to allow colonization by the pathogen [18,77]. Prediction of effector proteins from proteomes of *Colletotrichum* species has revealed different sets of effector candidates [30,44,62,68]. The evolution of effector proteins rely on the arms-race between plants and pathogens, with the aim of escape detection and evolve the capability of cause disease in different hosts [48], therefore the pathogenicity to specific hosts and/or cultivars can be a result of the evolution of effector proteins from a common ancestor [78,79], as shown for the hemibiotrophic pathogen *Phytophthora infestans* [14], *Venturia* spp. [80] and *Ceratocystis* spp. [81]. Our results revealed effector candidates for the four species pathogenic to soybean. Most of the *C. orchidearum* s.c. and species-LSECs are predicted to be secreted to the plant apoplast, while only a few genes are predicted to be localized to the plant cell nucleus or other subcellular compartments (Table 3, Table S1). These results suggest that the initial contact with the host is determinant for the capability of *Colletotrichum* species to infect soybean.

Initial pathogenicity tests revealed that among the tested *Colletotrichum* isolates, only the four *Colletotrichum* species previously associated with soybean [9–12] were pathogenic to the evaluated soybean cultivar. The three species that belong to the *C. orchidearum* s.c. showed a similar level of virulence, and lower than the level of virulence of *C. truncatum*. In another study, the virulence of one isolate of *C. plurivorum* was compared with five isolates of *C. truncatum*, and overall, the isolate was less virulent than at least one isolate of *C. truncatum* in soybean pods, stems and cotyledons, moreover, the authors reported that pod twisting symptoms were only caused by *C. plurivorum*, when the same stage of soybean development was compared after inoculation with *C. truncatum* [82].

While *C. truncatum* has been associated with soybean since 1917 [9], *C. musicola*, *C. plurivorum* and *C. sojae* were detected in soybean fields only recently [10–12]. Studies have revealed that the *C. orchidearum* s.c. has been misidentified at least since 2003, being *C. truncatum* and *C. orchidearum* s.c. the most associated with soybean until now [8]. Our results show that the four species that infect soybean belong to two lineages and do not share any of the identified LSECs. Moreover, the estimated divergence time of the *C. truncatum* s.c. occurred around 22.9 million years ago (mya), while the *C. orchidearum* s.c. diverged 4.8 mya [83], both of them before the domestication of soybean, that occurred 3000 years ago in China [84]. This evolutionary evidence, along with experimental data and the absence of host-LSECs shared only among the four species of *Colletotrichum* that infect soybean, strongly suggests that the two main *Colletotrichum* lineages associated with soybean have acquired the capability to infect soybean independently. Currently, *C. truncatum* is the most important species associated with soybean anthracnose worldwide [85,86], therefore, we checked if the *C. truncatum*-LSECs are conserved in 18 additional *C. truncatum* genomes pathogenic to soybean. Our results revealed that 17 *C. truncatum* genes have evolutionary evidence of being conserved among the species. This suggests that those effectors might play a role in the virulence of *C. truncatum* to soybean, as microorganisms do not keep useless genes due the high fitness costs of maintaining effector alleles [57,87]. Additionally to *in silico* prediction based genome sequences, an initial list of effector candidates can be narrowed down based on their expression [72]. 18 *C. truncatum*-LSECs have evidence of expression in soybean and/or *in vitro*. The low coverage of RNAseq data was a limiting factor for the analysis of gene expression, therefore LSECs without evidence of expression should not be excluded from the initial dataset and be further investigated.

The identification of sets of LSECs of the *C. orchidearum* s.c. and *C. truncatum* open the field to perform evaluations of the functional role of these genes in soybean infection. Besides cultural and chemical control strategies that have already been described for soybean anthracnose, recent outbreaks of the disease have been reported by researchers and producers [82,88,89], suggesting that the control strategies used are not always effective. This may be a consequence of different *Colletotrichum* species present in soybean fields,

that allied to the suggestion of separate evolution of these species, may imply directly in disease management strategies, as the correct identification of the causal agent is crucial to an efficient control strategy [90,91].

4. Materials and Methods

4.1 Fungal strains used

To gain insights into the repertoire of effector candidates of *Colletotrichum* species pathogenic to soybean, we selected 12 *Colletotrichum* proteomes and correspondent strains (Table 1). Four of them are pathogenic to soybean, including *C. musicola*, *C. plurivorum* and *C. sojae*, members of the *C. orchidearum* s.c. and *C. truncatum*. Eight additional proteomes, isolated from multiple hosts, were included in the analysis: *C. orbiculare*, *C. gloeosporioides* Ssensu Lato, *C. higginsianum*, *C. tofieldiae*, *C. graminicola*, *C. orchidophilum*, *C. fiorinae* and *C. nymphaeae* (Table 1).

4.2 Pathogenicity assays

Pathogenicity assays were performed to confirm the capability of the selected strains to cause soybean anthracnose. Except for *C. graminicola* and *C. orbiculare*, all the strains were retrieved from culture collections to perform the tests (Table 1). Seeds of the soybean cultivar IPRO7739, from Monsoy company, were superficially disinfected with NaClO (1%) for 1 min, then rinsed three times in sterile distilled water (SDW). Disinfected seeds were placed in Petri dishes containing 100 g of sterile sand, soaked with 10 mL of SDW. Each Petri dish contained 5 seeds and were incubated at 25°C for 32 h until germination.

Colletotrichum strains were grown on potato dextrose agar (PDA) culture medium and incubated at 25°C for 15 days. Conidia suspensions were prepared by washing and filtering the cultures and were adjusted to a final concentration of 1×10^6 conidia/mL. Each pre-germinated seed was inoculated with 5 μ L of conidial suspension of each *Colletotrichum* strain as described previously [92]. Water was used as a negative control. Inoculated seedlings were initially incubated in the dark at 25°C for 4 h and then transferred to 100 mL pots filled with sterilized vermiculite and randomly distributed in a greenhouse for 7 days when the severity of anthracnose was evaluated using an adapted diagrammatic scale that ranges from 0 to 5 [93]. Severity data were analyzed with the post-hoc Tuckey method at 0.05 significance level, using the ExpDes R package (v.1.2.0).

4.3 Identification of specific effector protein candidates (SECs) of soybean pathogenic

Colletotrichum species

The proteomes of four species of *Colletotrichum* pathogenic to soybean, and eight additional non-pathogenic *Colletotrichum* species were included in the analysis (Table 1). A phylogeny of the genus *Colletotrichum* was constructed based on publicly available DNA sequences of three nuclear loci belonging to the 12 selected species: *actin* (*ACT*), *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) and *chitin synthase* (*CHS*). The analyses were run from random trees for 5,000,000 generations and sampled every 1000 generations. The predicted proteomes of the 12 *Colletotrichum* spp. were clustered based on similarity with OrthoFinder (v. 2.3.5) [94] and the clusters of proteins were analyzed with the R package UpsetR (v. 1.4.0) [95] to identify unique and shared orthogroups between the species and species complexes.

The prediction of effector candidates of soybean pathogenic species of *Colletotrichum* was made using the proteomes predicted by Rogério et al., (2020). The initial secretome was predicted with SignalP (v.5.0) [96], then sequences containing transmembrane (TM) domains and glycoposphatidylinositol (GPI)-anchors were identified using THMM (v.2.0) [97] and PredGPI [98] respectively, and those proteins that are predicted to have a

signal peptide cleavage site, no transmembrane domains and no GPI-anchors were considered as the initial set of effector candidates for each species of *Colletotrichum*.

The set of effector candidates of each *Colletotrichum* species was submitted individually to a sequence of BLAST searches using an *e*-value threshold of 1E-5 and classified into shared (proteins with homology to proteins from other members of the genus *Colletotrichum*), species-complex specific (those that had homology only within other species from the same s. c.), host-specific (shared only between the four species that infect soybean) and species-specific (those that had no homology to any other protein either within or outside of the same genus) LSECs [30]. The final set of predicted LSECs was scanned with RunIprScan to identify conserved domains and submitted to a BLAST against the non-redundant database of NCBI and Pathogen Host interactions-base (PHI-base) to check the similarity with known genes of other microorganism species; being those proteins with similarity outside the genus *Colletotrichum* considered conserved among microorganisms.

Species-specific and species complex LSECs were characterized. For the prediction of subcellular localization within the plant cell, mature protein sequences were submitted to LOCALIZER [75], and to the prediction of apoplastic LSECs, the proteins were submitted to ApoplastP [76]. The percentage of cysteines was identified in Geneious (v. 2020.10.4) and repeat-containing proteins were predicted using T-REKs [99].

4.5 Genome assembly and gene evolution

To check if *C. truncatum*-LSECs are conserved among the species, Illumina reads of 18 *C. truncatum* strains available in NCBI (Table 4) were trimmed with Trim Galore (v.0.4.5). Forward and reverse reads were merged using Flash (v.1.2.7) [100]. Assemblies of combined and uncombined reads were performed with SPAdes v.3.13.1 using the *C. truncatum* CMES1059 strain genome as a reference.

Table 4. *Colletotrichum truncatum* strains used in the evolutionary analysis

Strain	Species	Species complex	Host	Origin	Accession N°
MT1-01	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095339
MT2-05	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095348
MT3-01	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095353
MT3-21	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095343
MT4-05	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095347
MT4-13	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095341
MT5-12	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095345
MT5-26	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095349
MT5-32	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095350
GO2-03	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095352
GO2-06	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095344
GO2-12	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095346
GO4-07	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095355
GO4-08	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095338
GO4-17	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095351
GO5-11	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095354
GO5-14	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095340
GO5-25	<i>C. truncatum</i>	<i>C. truncatum</i>	<i>Glycine max</i>	Brazil	SRX7095342

4.6 Evidence of expression of *C. truncatum* by RNAseq

To confirm evidence of gene expression of *C. truncatum* in planta, five pre-germinated seeds of soybean cultivars IPRO7739 and IPRO8372 were inoculated with *C. truncatum* (CMES1059) strain as described in 4.2. Hypocotyls fragments of 0.5 cm of five randomly selected plants were collected and pooled together at 12; 48 and 120 hpi. To confirm evidence of expression of *C. truncatum* in vitro, 100mL of potato dextrose liquid culture was inoculated with *C. truncatum* conidia in 250mL Erlenmeyer flasks at 25°C, shaken at 150 rpm. After 120 hpi micelia was collected by filtration and washed with ADE. Harvested plant tissue and fungal micelia was flash-frozen in liquid N₂ and stored at -80°C until RNA extraction. Three biological replicates of the experiment were performed. The collected material was ground using mortar and pestle and total RNA was purified using PureLink RNA Mini Kit (Invitrogen, USA) following the manufacturer's instructions. Total RNA was treated with RNase-free DNase (Life Technologies) to remove DNA contamination. The quantity of total RNA was estimated using Qubit 2.0 fluorometer (Life Technologies) and RNA integrity was checked using Agilent TapeStation 4200 (Agilent Technologies).

Total extracted RNA was sent to Genewiz (South Plainfield, USA) for Illumina sequencing. In total, 21 libraries derived from all the treatments were prepared using NEBNext Ultra RNA Library Prep Kit for Illumina (NEB) using manufacturer's instructions. Sequencing libraries were validated on the Agilent TapeStation (Agilent technologies) and quantified in Qubit 2.0 fluorometer (Invitrogen) and by quantitative PCR (Kapa Biosystems). Libraries were sequenced using Illumina HiSeq4000 (2x150 bp).

The quality of reads was accessed using FastQC (v.0.11.7) and clean reads were obtained by removing reads containing adapters with CutAdaptors (v.1.9.1). Paired-end clean reads were mapped against the *C. truncatum* CMES1059 reference genome [45] using HISAT (v. 2.1.0). Alignments from each library were processed with StringTIE (v.1.3.5) to quantify expression values of transcripts.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1. Figure S1: In silico analysis workflow for the prediction of effector candidates in *C. truncatum*, *C. musicola*, *C. plurivorum* and *C. sojae*. Table S1: Lineage Specific Effector Candidates (LSECs) of *C. truncatum*, *C. musicola*, *C. plurivorum* and *C. sojae*.

Author Contributions: Conceptualization, R.B. and T.R.B.; methodology, T.R.B. and R.B.; formal analysis, T.R.B.; investigation, T.R.B., R.B., S.A.S., M.R.T., I.T., N.S.M., M.R.T.; resources, N.S.M.; data curation, T.B., R.B.; writing—original draft preparation, T.B.; writing—review and editing, R.B., N.S.M., I.T., S.A.S., M.R.T.; supervision, R.B., N.S.M., M.R.T.; funding acquisition, N.S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO – FAPESP, grant number 2017/09178-8, and NATIONAL COUNCIL FOR THE IMPROVEMENT OF HIGHER EDUCATION – CAPES grant number 88887.368016/2019-00 and the Ministerio de Ciencia and Innovación of Spain (grant RTI2018-093611-B-I00).

Institutional Review Board Statement: “Not applicable.”

Informed Consent Statement: “Not applicable.”

Data Availability Statement: All the data analyzed in this work are publicly available at NCBI (<https://www.ncbi.nlm.nih.gov/>) under the accession numbers: *C. orbiculare* (AMCV00000000.1), *C. musicola* (WIGM00000000), *C. plurivorum* (WIGO00000000), *C. sojae* (WIGN00000000), *C. truncatum* (VUJX00000000.1), *C. gloeosporioides* (AMYD00000000.1), *C. higginsianum* (LTAN00000000.1), *C. tofieldiae* (LFHQ00000000.1), *C. graminicola* (ACOD00000000.1), *C. orchidophilum* (MJBS00000000.1), *C. fiorineae* (JARH00000000.1), *C. nymphaeae* (JEMN00000000.1).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Talhinas, P.; Baroncelli, R. *Colletotrichum* Species and Complexes: Geographic Distribution, Host Range and Conservation Status. *Fungal Divers.* **2021**, doi:10.1007/s13225-021-00491-9.
2. Dean, R.; Van Kan, J.A.L.; Pretorius, Z.A.; Hammond-Kosack, K.E.; Di Pietro, A.; Spanu, P.D.; Rudd, J.J.; Dickman, M.; Kahmann, R.; Ellis, J.; et al. The Top 10 Fungal Pathogens in Molecular Plant Pathology: Top 10 Fungal Pathogens. *Mol. Plant Pathol.* **2012**, *13*, 414–430, doi:10.1111/j.1364-3703.2011.00783.x.
3. Vargas, W.A.; Martín, J.M.S.; Rech, G.E.; Rivera, L.P.; Benito, E.P.; Díaz-Minguez, J.M.; Thon, M.R.; Sukno, S.A. Plant Defense Mechanisms Are Activated during Biotrophic and Necrotrophic Development of *Colletotrichum graminicola* in Maize. *Plant Physiol.* **2012**, *158*, 1342–1358, doi:10.1104/pp.111.190397.
4. Amil-Ruiz, F.; Garrido-Gala, J.; Gadea, J.; Blanco-Portales, R.; Muñoz-Mérida, A.; Trelles, O.; de los Santos, B.; Arroyo, F.T.; Aguado-Puig, A.; Romero, F.; et al. Partial Activation of SA- and JA-Defensive Pathways in Strawberry upon *Colletotrichum acutatum* Interaction. *Front. Plant Sci.* **2016**, *7*, doi:10.3389/fpls.2016.01036.
5. Dias, M.D.; Pinheiro, V.F.; Café-Filho, A.C. Impact of Anthracnose on the Yield of Soybean Subjected to Chemical Control in the North Region of Brazil. *Summa Phytopathol.* **2016**, *42*, 18–23, doi:10.1590/0100-5405/2114.
6. Banniza, S.; Warale, R.; Menat, J.; Cohen-Skali, A.; Armstrong-Cho, C.; Bhaduria, V. The Long Path to Understanding the Host–Pathogen Interactions of *Colletotrichum lentis* on Lentil. *Can. J. Plant Pathol.* **2018**, *40*, 199–209, doi:10.1080/07060661.2018.1451391.
7. de Silva, D.D.; Groenewald, J.Z.; Crous, P.W.; Ades, P.K.; Nasruddin, A.; Mongkolporn, O.; Taylor, P.W.J. Identification, Prevalence and Pathogenicity of *Colletotrichum* Species Causing Anthracnose of *Capsicum annuum* in Asia. *IMA Fungus* **2019**, *10*, 8, doi:10.1186/s43008-019-0001-y.
8. Bouffleur, T.R.; Ciampi-Guillardi, M.; Tikami, I.; Rogério, F.; Thon, M.R.; Sukno, S.A.; Massola Júnior, N.S.; Baroncelli, R. Soybean Anthracnose Caused by *Colletotrichum* Species: Current Status and Future Prospects. *Mol. Plant Pathol.* **2021**, doi:10.1111/mpp.13036.
9. Nakata, K.; Takimoto, K. *A List of Crop Diseases in Korea*; Agricultural Experiment Station; Korea, 1934;
10. Bouffleur, T.R.; Castro, R.R.L.; Rogério, F.; Ciampi-Guillardi, M.; Baroncelli, R.; Massola Júnior, N.S. First Report of *Colletotrichum musicola* Causing Soybean Anthracnose in Brazil. *Plant Dis.* **2020**, *104*, 1858, doi:10.1094/PDIS-12-19-2627-PDN.
11. Barbieri, M.C.G.; Ciampi-Guillardi, M.; Moraes, S.R.G.; Bonaldo, S.M.; Rogério, F.; Linhares, R.R.; Massola, N.S. First Report of *Colletotrichum cliviae* Causing Anthracnose on Soybean in Brazil. *Plant Dis.* **2017**, *101*, 1677–1677, doi:10.1094/PDIS-07-16-0963-PDN.
12. Damm, U.; Sato, T.; Alizadeh, A.; Groenewald, J.Z.; Crous, P.W. The *Colletotrichum dracaenophilum*, *C. magnum* and *C. orchidearum* Species Complexes. *Studies Mycol.* **2019**, *92*, 1–46, doi:10.1016/j.simyco.2018.04.001.
13. Jones, J.D.G.; Dangl, J.L. The Plant Immune System. *Nature* **2006**, *444*, 323–329, doi:10.1038/nature05286.
14. Dong, S.; Stam, R.; Cano, L.M.; Song, J.; Sklenar, J.; Yoshida, K.; Bozkurt, T.O.; Oliva, R.; Liu, Z.; Tian, M.; et al. Effector Specialization in a Lineage of the Irish Potato Famine Pathogen. *Science* **2014**, *343*, 552–555, doi:10.1126/science.1246300.
15. Dong, S.; Raffaele, S.; Kamoun, S. The Two-Speed Genomes of Filamentous Pathogens: Waltz with Plants. *Curr. Opin. Genet. Dev.* **2015**, *35*, 57–65, doi:10.1016/j.gde.2015.09.001.

16. Boller, T.; Felix, G. A Renaissance of Elicitors: Perception of Microbe-Associated Molecular Patterns and Danger Signals by Pattern-Recognition Receptors. *Annu. Rev. Plant Biol.* **2009**, *60*, 379–406, doi:10.1146/annurev.arplant.57.032905.105346.
17. Couto, D.; Zipfel, C. Regulation of Pattern Recognition Receptor Signalling in Plants. *Nat. Rev. Immunol.* **2016**, *16*, 537–552, doi:10.1038/nri.2016.77.
18. Win, J.; Chaparro-Garcia, A.; Belhaj, K.; Saunders, D.G.O.; Yoshida, K.; Dong, S.; Schornack, S.; Zipfel, C.; Robatzek, S.; Hogenhout, S.A.; et al. Effector Biology of Plant-Associated Organisms: Concepts and Perspectives. *Cold Spring Harbor Symp. Quant. Biol.* **2012**, *77*, 235–247, doi:10.1101/sqb.2012.77.015933.
19. Kanyuka, K.; Rudd, J.J. Cell Surface Immune Receptors: The Guardians of the Plant's Extracellular Spaces. *Curr. Opin. Plant Biol.* **2019**, *50*, 1–8, doi:10.1016/j.pbi.2019.02.005.
20. Yuan, M.; Jiang, Z.; Bi, G.; Nomura, K.; Liu, M.; Wang, Y.; Cai, B.; Zhou, J.-M.; He, S.Y.; Xin, X.-F. Pattern-Recognition Receptors Are Required for NLR-Mediated Plant Immunity. *Nature* **2021**, *592*, 105–109, doi:10.1038/s41586-021-03316-6.
21. Ngou, B.P.M.; Ahn, H.-K.; Ding, P.; Jones, J.D.G. Mutual Potentiation of Plant Immunity by Cell-Surface and Intracellular Receptors. *Nature* **2021**, *592*, 110–115, doi:10.1038/s41586-021-03315-7.
22. Bjornson, M.; Zipfel, C. Plant Immunity: Crosstalk between Plant Immune Receptors. *Curr. Biol.* **2021**, *31*, R796–R798, doi:10.1016/j.cub.2021.04.080.
23. O'Connell, R.J.; Thon, M.R.; Hacquard, S.; Amyotte, S.G.; Kleemann, J.; Torres, M.F.; Damm, U.; Buiate, E.A.; Epstein, L.; Alkan, N.; et al. Lifestyle Transitions in Plant Pathogenic *Colletotrichum* Fungi Deciphered by Genome and Transcriptome Analyses. *Nat. Genet.* **2012**, *44*, 1060–1065, doi:10.1038/ng.2372.
24. Gan, P.; Ikeda, K.; Irieda, H.; Narusaka, M.; O'Connell, R.J.; Narusaka, Y.; Takano, Y.; Kubo, Y.; Shirasu, K. Comparative Genomic and Transcriptomic Analyses Reveal the Hemibiotrophic Stage Shift of *Colletotrichum* Fungi. *New Phytol.* **2013**, *197*, 1236–1249, doi:10.1111/nph.12085.
25. Alkan, N.; Friedlander, G.; Ment, D.; Prusky, D.; Fluhr, R. Simultaneous Transcriptome Analysis of *Colletotrichum gloeosporioides* and Tomato Fruit Pathosystem Reveals Novel Fungal Pathogenicity and Fruit Defense Strategies. *New Phytol.* **2015**, *205*, 801–815, doi:10.1111/nph.13087.
26. Huo, J.; Wang, Y.; Hao, Y.; Yao, Y.; Wang, Y.; Zhang, K.; Tan, X.; Li, Z.; Wang, W. Genome Sequence Resource for *Colletotrichum scovillei*, the Cause of Anthracnose Disease of Chili. *MPMI* **2021**, *34*, 122–126, doi:10.1094/MPMI-03-20-0055-A.
27. Eaton, M.J.; Edwards, S.; Inocencio, H.A.; Machado, F.J.; Nuckles, E.M.; Farman, M.; Gauthier, N.A.; Vaillancourt, L.J. Diversity and Cross-Infection Potential of *Colletotrichum* Causing Fruit Rots in Mixed-Fruit Orchards in Kentucky. *Plant Dis.* **2021**, *105*, 1115–1128, doi:10.1094/PDIS-06-20-1273-RE.
28. Liu, X.; Li, B.; Yang, Y.; Cai, J.; Shi, T.; Zheng, X.; Huang, G. Pathogenic Adaptations Revealed by Comparative Genome Analyses of Two *Colletotrichum* spp., the Causal Agent of Anthracnose in Rubber Tree. *Front Microbiol.* **2020**, *11*, 1484, doi:10.3389/fmicb.2020.01484.
29. Baroncelli, R.; Sreenivasaprasad, S.; Sukno, S.A.; Thon, M.R.; Holub, E. Draft Genome Sequence of *Colletotrichum acutatum* Sensu Lato (*Colletotrichum fioriniae*). *Genome Announc.* **2014**, *2*, e00112-14, 2/2/e00112-14, doi:10.1128/genomeA.00112-14.
30. Baroncelli, R.; Amby, D.B.; Zapparata, A.; Sarrocco, S.; Vannacci, G.; Le Floch, G.; Harrison, R.J.; Holub, E.; Sukno, S.A.; Sreenivasaprasad, S.; et al. Gene Family Expansions and Contractions Are Associated with Host Range in Plant Pathogens of the Genus *Colletotrichum*. *BMC Genomics* **2016**, *17*, 555, doi:10.1186/s12864-016-2917-6.
31. Gan, P.; Narusaka, M.; Kumakura, N.; Tsushima, A.; Takano, Y.; Narusaka, Y.; Shirasu, K. Genus-Wide Comparative Genome Analyses of *Colletotrichum* Species Reveal Specific Gene Family Losses and Gains during Adaptation to Specific Infection Lifestyles. *Genome Biol. Evol.* **2016**, *8*, 1467–1481, doi:10.1093/gbe/evw089.
32. Hacquard, S.; Kracher, B.; Hiruma, K.; Münch, P.C.; Garrido-Oter, R.; Thon, M.R.; Weimann, A.; Damm, U.; Dallery, J.-F.;

- Hainaut, M.; et al. Survival Trade-Offs in Plant Roots during Colonization by Closely Related Beneficial and Pathogenic Fungi. *Nat. Commun.* **2016**, *7*, 11362, doi:10.1038/ncomms11362.
33. Viswanathan, R.; Prasanth, C.N.; Malathi, P.; Sundar, A.R. Draft Genome Sequence of *Colletotrichum falcatum* - a Prelude on Screening of Red Rot Pathogen in Sugarcane. *J. Genomics* **2016**, *4*, 1–3, doi:10.7150/jgen.13585.
34. Han, J.-H.; Chon, J.-K.; Ahn, J.-H.; Choi, I.-Y.; Lee, Y.-H.; Kim, K.S. Whole Genome Sequence and Genome Annotation of *Colletotrichum acutatum*, Causal Agent of Anthracnose in Pepper Plants in South Korea. *Genom. Data* **2016**, *8*, 45–46, doi:10.1016/j.gdata.2016.03.007.
35. Dallery, J.-F.; Lapalu, N.; Zampounis, A.; Pigné, S.; Luyten, I.; Amselem, J.; Wittenberg, A.H.J.; Zhou, S.; de Queiroz, M.V.; Robin, G.P.; et al. Gapless Genome Assembly of *Colletotrichum higginsianum* Reveals Chromosome Structure and Association of Transposable Elements with Secondary Metabolite Gene Clusters. *BMC Genomics* **2017**, *18*, 667, doi:10.1186/s12864-017-4083-x.
36. Baroncelli, R.; Sukno, S.A.; Sarrocco, S.; Cafà, G.; Le Floch, G.; Thon, M.R. Whole-Genome Sequence of the Orchid Anthracnose Pathogen *Colletotrichum orchidophilum*. *MPMI* **2018**, *31*, 979–981, doi:10.1094/MPMI-03-18-0055-A.
37. Gan, P.; Narusaka, M.; Tsushima, A.; Narusaka, Y.; Takano, Y.; Shirasu, K. Draft Genome Assembly of *Colletotrichum chlorophyti*, a Pathogen of Ehbaceous Plants. *Genome Announc.* **2017**, *5*, e01733-16, doi:10.1128/genomeA.01733-16.
38. Gan, P.; Tsushima, A.; Narusaka, M.; Narusaka, Y.; Takano, Y.; Kubo, Y.; Shirasu, K. Genome Sequence Resources for Four Phytopathogenic Fungi from the *Colletotrichum orbiculare* Species Complex. *MPMI* **2019**, *32*, 1088–1090, doi:10.1094/MPMI-12-18-0352-A.
39. Nakamura, M.; Fujikawa, T.; Nakamori, D.; Iwai, H. Draft Genome Sequence of *Colletotrichum sansevieriae* Sa-1–2, the Anthracnose Pathogen of *Sansevieria trifasciata*. *Data Brief* **2018**, *18*, 691–695, doi:10.1016/j.dib.2018.03.083.
40. Bhadauria, V.; MacLachlan, R.; Pozniak, C.; Cohen-Skalie, A.; Li, L.; Halliday, J.; Banniza, S. Genetic Map-Guided Genome Assembly Reveals a Virulence-Governing Minichromosome in the Lentil Anthracnose Pathogen *Colletotrichum lentis*. *New Phytol.* **2019**, *221*, 431–445, doi:10.1111/nph.15369.
41. Silva Junior, W.J. da; Falcão, R.M.; Sousa-Paula, L.C. de; Sbaraini, N.; Vieira, W.A. dos S.; Lima, W.G.; Paiva Junior, S. de S.L.; Staats, C.C.; Schrank, A.; Benko-Iseppon, A.M.; et al. Draft Genome Assembly of *Colletotrichum musae*, the Pathogen of Banana Fruit. *Data Brief* **2018**, *17*, 256–260, doi:10.1016/j.dib.2018.01.002.
42. Lelwala, R.V.; Korhonen, P.K.; Young, N.D.; Scott, J.B.; Ades, P.K.; Gasser, R.B.; Taylor, P.W.J. Comparative Genome Analysis Indicates High Evolutionary Potential of Pathogenicity Genes in *Colletotrichum tanacetii*. *PLoS One* **2019**, *14*, e0212248, doi:10.1371/journal.pone.0212248.
43. Gan, P.; Tsushima, A.; Hiroyama, R.; Narusaka, M.; Takano, Y.; Narusaka, Y.; Kawaradani, M.; Damm, U.; Shirasu, K. *Colletotrichum ahisoi* Sp. Nov., an Anthracnose Pathogen of *Perilla frutescens* in Japan: Molecular Phylogenetic, Morphological and Genomic Evidence. *Sci. Rep.* **2019**, *9*, 13349, doi:10.1038/s41598-019-50076-5.
44. Gan, P.; Hiroyama, R.; Tsushima, A.; Masuda, S.; Shibata, A.; Ueno, A.; Kumakura, N.; Narusaka, M.; Hoat, T.X.; Narusaka, Y.; et al. Telomeres and a Repeat-rich Chromosome Encode Effector Gene Clusters in Plant Pathogenic *Colletotrichum* Fungi. *Environ. Microbiol.* **2021**, 1462–2920.15490, doi:10.1111/1462-2920.15490.
45. Rogério, F.; Bouffleur, T.R.; Ciampi-Guillardi, M.; Sukno, S.A.; Thon, M.R.; Massola Júnior, N.S.; Baroncelli, R. Genome Sequence Resources of *Colletotrichum truncatum*, *C. plurivorum*, *C. musicola*, and *C. sojae*: Four Species Pathogenic to Soybean (*Glycine max*). *Phytopathology* **2020**, PHYTO-03-20-010, doi:10.1094/PHYTO-03-20-0102-A.
46. Meng, Y.; Ren, Y.; Wang, W.; Gleason, M.L.; Zhang, R.; Sun, G. A Genome Sequence Resource for the Geographically Widespread Anthracnose Pathogen *Colletotrichum asianum*. *Plant Dis.* **2020**, *104*, 2044–2047, doi:10.1094/PDIS-01-20-0034-A.
47. Baroncelli, R.; Pensec, F.; Da Lio, D.; Bouffleur, T.R.; Vicente, I.; Sarrocco, S.; Picot, A.; Baraldi, E.; Sukno, S.; Thon, M.R.; et al. Complete Genome Sequence of the Plant Pathogenic Fungus *Colletotrichum lupini*. *MPMI* **2021**, doi:10.1094/MPMI-07-21-0173-A.

48. Lo Presti, L.; Lanver, D.; Schweizer, G.; Tanaka, S.; Liang, L.; Tollot, M.; Zuccaro, A.; Reissmann, S.; Kahmann, R. Fungal Effectors and Plant Susceptibility. *Annu. Rev. Plant Biol.* **2015**, *66*, 513–545, doi:10.1146/annurev-arplant-043014-114623.
49. Dussert, Y.; Mazet, I.D.; Couture, C.; Gouzy, J.; Piron, M.C.; Kuchly, C.; Bouchez, O.; Rispe, C.; Mestre, P.; Delmotte, F.; et al. A High-Quality Grapevine Downy Mildew Genome Assembly Reveals Rapidly Evolving and Lineage-Specific Putative Host Adaptation Genes. *Genome Biol. Evol.* **2019**, *11*, 954–969, doi:10.1093/gbe/evz048.
50. Depotter, J.R.L.; Doehlemann, G. Target the Core: Durable Plant Resistance against Filamentous Plant Pathogens through Effector Recognition. *Pest. Manag. Sci.* **2020**, *76*, 426–431, doi:10.1002/ps.5677.
51. Mousavi-Derazmahalleh, M.; Chang, S.; Thomas, G.; Derbyshire, M.; Bayer, P.E.; Edwards, D.; Nelson, M.N.; Erskine, W.; Lopez-Ruiz, F.J.; Clements, J.; et al. Prediction of Pathogenicity Genes Involved in Adaptation to a Lupin Host in the Fungal Pathogens *Botrytis Cinerea* and *Sclerotinia sclerotiorum* via Comparative Genomics. *BMC Genomics* **2019**, *20*, 385, doi:10.1186/s12864-019-5774-2.
52. Ma, L.-J.; van der Does, H.C.; Borkovich, K.A.; Coleman, J.J.; Daboussi, M.-J.; Di Pietro, A.; Dufresne, M.; Freitag, M.; Grabherr, M.; Henrissat, B.; et al. Comparative Genomics Reveals Mobile Pathogenicity Chromosomes in *Fusarium*. *Nature* **2010**, *464*, 367–373, doi:10.1038/nature08850.
53. Poppe, S.; Dorsheimer, L.; Happel, P.; Stukenbrock, E.H. Rapidly Evolving Genes Are Key Players in Host Specialization and Virulence of the Fungal Wheat Pathogen *Zymoseptoria tritici* (*Mycosphaerella graminicola*). *PLoS Pathog.* **2015**, *11*, e1005055, doi:10.1371/journal.ppat.1005055.
54. Petre, B.; Lorrain, C.; Stukenbrock, E.H.; Duplessis, S. Host-Specialized Transcriptome of Plant-Associated Organisms. *Curr. Opin. Plant Biol.* **2020**, *56*, 81–88, doi:10.1016/j.pbi.2020.04.007.
55. Palma-Guerrero, J.; Ma, X.; Torriani, S.F.F.; Zala, M.; Francisco, C.S.; Hartmann, F.E.; Croll, D.; McDonald, B.A. Comparative Transcriptome Analyses in *Zymoseptoria tritici* Reveal Significant Differences in Gene Expression among Strains during Plant Infection. *MPMI* **2017**, *30*, 231–244, doi:10.1094/MPMI-07-16-0146-R.
56. Liao, J.; Huang, H.; Meusnier, I.; Adreit, H.; Ducasse, A.; Bonnot, F.; Pan, L.; He, X.; Kroj, T.; Fournier, E.; et al. Pathogen Effectors and Plant Immunity Determine Specialization of the Blast Fungus to Rice Subspecies. *eLife* **2016**, *5*, e19377, doi:10.7554/eLife.19377.
57. Sánchez-Vallet, A.; Fouché, S.; Fudal, I.; Hartmann, F.E.; Soyer, J.L.; Tellier, A.; Croll, D. The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens. *Annu. Rev. Phytopathol.* **2018**, *56*, 21–40, doi:10.1146/annurev-phyto-080516-035303.
58. Kleemann, J.; Rincon-Rivera, L.J.; Takahara, H.; Neumann, U.; Ver Loren van Themaat, E.; van der Does, H.C.; Hacquard, S.; Stüber, K.; Will, I.; Schmalenbach, W.; et al. Correction: Sequential Delivery of Host-Induced Virulence Effectors by Appressoria and Intracellular Hyphae of the Phytopathogen *Colletotrichum higginsianum*. *PLoS Pathog.* **2012**, *8*, doi:10.1371/annotation/0f398a0c-dfda-4277-b172-4ff9cb31aec3.
59. Robin, G.P.; Kleemann, J.; Neumann, U.; Cabre, L.; Dallery, J.-F.; Lapalu, N.; O’Connell, R.J. Subcellular Localization Screening of *Colletotrichum higginsianum* Effector Candidates Identifies Fungal Proteins Targeted to Plant Peroxisomes, Golgi Bodies, and Microtubules. *Front. Plant Sci.* **2018**, *9*, doi:10.3389/fpls.2018.00562.
60. Irieda, H.; Maeda, H.; Akiyama, K.; Hagiwara, A.; Saitoh, H.; Uemura, A.; Terauchi, R.; Takano, Y. *Colletotrichum orbiculare* Secretes Virulence Effectors to a Biotrophic Interface at the Primary Hyphal Neck via Exocytosis Coupled with SEC22-Mediated Traffic. *Plant Cell* **2014**, *26*, 2265–2281, doi:10.1105/tpc.113.120600.
61. Irieda, H.; Takano, Y. Identification and Characterization of Virulence-Related Effectors in the Cucumber Anthracnose Fungus *Colletotrichum orbiculare*. *Physiol. Mol. Plant Pathol.* **2016**, *95*, 87–92, doi:10.1016/j.pmpp.2016.01.006.
62. Bhadauria, V.; MacLachlan, R.; Pozniak, C.; Banniza, S. Candidate Effectors Contribute to Race Differentiation and Virulence of the Lentil Anthracnose Pathogen *Colletotrichum lentis*. *BMC Genomics* **2015**, *16*, 628, doi:10.1186/s12864-015-1836-

2.

63. Bhadauria, V.; Banniza, S.; Vandenberg, A.; Selvaraj, G.; Wei, Y. Overexpression of a Novel Biotrophy-Specific *Colletotrichum truncatum* Effector, CtNUDIX, in Hemibiotrophic Fungal Phytopathogens Causes Incompatibility with Their Host Plants. *Eukaryot. Cell* **2013**, *12*, 2–11, doi:10.1128/EC.00192-12.
64. Bhadauria, V.; Vijayan, P.; Wei, Y.; Banniza, S. Transcriptome Analysis Reveals a Complex Interplay between Resistance and Effector Genes during the Compatible Lentil-*Colletotrichum lentis* Interaction. *Sci. Rep.* **2017**, *7*, 42338, doi:10.1038/srep42338.
65. Sanz-Martín, J.M.; Pacheco-Arjona, J.R.; Bello-Rico, V.; Vargas, W.A.; Monod, M.; Díaz-Mínguez, J.M.; Thon, M.R.; Sukno, S.A. A Highly Conserved Metalloprotease Effector Enhances Virulence in the Maize Anthracnose Fungus *Colletotrichum graminicola*. *Mol. Plant Pathol.* **2016**, *17*, 1048–1062, doi:https://doi.org/10.1111/mpp.12347.
66. Vargas, W.A.; Sanz-Martín, J.M.; Rech, G.E.; Armijos-Jaramillo, V.D.; Rivera, L.P.; Echeverría, M.M.; Díaz-Mínguez, J.M.; Thon, M.R.; Sukno, S.A. A Fungal Effector with Host Nuclear Localization and DNA-Binding Properties Is Required for Maize Anthracnose Development. *MPMI* **2015**, *29*, 83–95, doi:10.1094/MPMI-09-15-0209-R.
67. Eisermann, I.; Weihmann, F.; Krijger, J.-J.; Kröling, C.; Hause, G.; Menzel, M.; Pienkny, S.; Kiesow, A.; Deising, H.B.; Wirsal, S.G.R. Two Genes in a Pathogenicity Gene Cluster Encoding Secreted Proteins Are Required for Appressorial Penetration and Infection of the Maize Anthracnose Fungus *Colletotrichum graminicola*. *Environ. Microbiol.* **2019**, *21*, 4773–4791, doi:10.1111/1462-2920.14819.
68. de Queiroz, C.B.; Correia, H.L.N.; Santana, M.F.; Batista, D.S.; Vidigal, P.M.P.; Brommonschenkel, S.H.; de Queiroz, M.V. The Repertoire of Effector Candidates in *Colletotrichum lindemuthianum* Reveals Important Information about *Colletotrichum* Genus Lifestyle. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 2295–2309, doi:10.1007/s00253-019-09639-9.
69. Prasanth, C.N.; Viswanathan, R.; Malathi, P.; Sundar, A.R. Comparative Transcriptome Analysis of Candidate Secretory Effector Proteins from *Colletotrichum falcatum* Infecting Sugarcane. *Agri. Gene* **2019**, *13*, 100089, doi:10.1016/j.aggene.2019.100089.
70. Zampounis, A.; Pigné, S.; Dallery, J.-F.; Wittenberg, A.H.J.; Zhou, S.; Schwartz, D.C.; Thon, M.R.; O'Connell, R.J. Genome Sequence and Annotation of *Colletotrichum higginsianum*, a Causal Agent of Crucifer Anthracnose Disease. *Genome Announc.* **2016**, *4*, e00821-16, /ga/4/4/e00821-16.atom, doi:10.1128/genomeA.00821-16.
71. Hacquard, S.; Kracher, B.; Hiruma, K.; Münch, P.C.; Garrido-Oter, R.; Thon, M.R.; Weimann, A.; Damm, U.; Dallery, J.-F.; Hainaut, M.; et al. Survival Trade-Offs in Plant Roots during Colonization by Closely Related Beneficial and Pathogenic Fungi. *Nat. Commun.* **2016**, *7*, 11362, doi:10.1038/ncomms11362.
72. Depotter, J.R.L.; Doehlemann, G. Target the Core: Durable Plant Resistance against Filamentous Plant Pathogens through Effector Recognition. *Pest. Manag. Sci.* **2020**, *76*, 426–431, doi:10.1002/ps.5677.
73. Lu, S.; Edwards, M.C. Genome-Wide Analysis of Small Secreted Cysteine-Rich Proteins Identifies Candidate Effector Proteins Potentially Involved in *Fusarium graminearum*-wheat Interactions. *Phytopathology* **2016**, *106*, 166–176, doi:10.1094/PHYTO-09-15-0215-R.
74. Mesarich, C.H.; Bowen, J.K.; Hamiaux, C.; Templeton, M.D. Repeat-Containing Protein Effectors of Plant-Associated Organisms. *Front. Plant Sci.* **2015**, *6*, doi:10.3389/fpls.2015.00872.
75. Sperschneider, J.; Catanzariti, A.-M.; DeBoer, K.; Petre, B.; Gardiner, D.M.; Singh, K.B.; Dodds, P.N.; Taylor, J.M. LOCALIZER: Subcellular Localization Prediction of Both Plant and Effector Proteins in the Plant Cell. *Sci. Rep.* **2017**, *7*, 44598, doi:10.1038/srep44598.
76. Sperschneider, J.; Dodds, P.N.; Singh, K.B.; Taylor, J.M. ApoplastP: Prediction of Effectors and Plant Proteins in the Apoplast Using Machine Learning. *New Phytol.* **2018**, *217*, 1764–1778, doi:10.1111/nph.14946.
77. Hogenhout, S.A.; Van der Hoorn, R.A.L.; Terauchi, R.; Kamoun, S. Emerging Concepts in Effector Biology of Plant-Associated Organisms. *MPMI* **2009**, *22*, 115–122, doi:10.1094/MPMI-22-2-0115.

78. Stergiopoulos, I.; van den Burg, H.A.; Okmen, B.; Beenen, H.G.; van Liere, S.; Kema, G.H.J.; de Wit, P.J.G.M. Tomato Cf Resistance Proteins Mediate Recognition of Cognate Homologous Effectors from Fungi Pathogenic on Dicots and Monocots. *PNAS* **2010**, *107*, 7610–7615, doi:10.1073/pnas.1002910107.
79. Stergiopoulos, I.; Kourmpetis, Y.A.I.; Slot, J.C.; Bakker, F.T.; De Wit, P.J.G.M.; Rokas, A. In Silico Characterization and Molecular Evolutionary Analysis of a Novel Superfamily of Fungal Effector Proteins. *Mol. Biol. Evol.* **2012**, *29*, 3371–3384, doi:10.1093/molbev/mss143.
80. Prokchorchik, M.; Won, K.; Lee, Y.; Segonzac, C.; Sohn, K.H. Whole Genome Enabled Phylogenetic and Secretome Analyses of Two *Venturia nashicola* Isolates. *Plant Pathol. J.* **2020**, *36*, 98–105, doi:10.5423/PPJ.NT.10.2019.0258.
81. Fourie, A.; de Jonge, R.; van der Nest, M.A.; Duong, T.A.; Wingfield, M.J.; Wingfield, B.D.; Barnes, I. Genome Comparisons Suggest an Association between *Ceratocystis* Host Adaptations and Effector Clusters in Unique Transposable Element Families. *Fungal Genet. Biol.* **2020**, *143*, 103433, doi:10.1016/j.fgb.2020.103433.
82. Dias, M.D.; Fonseca, M.E.N.; Dias-Neto, J.J.; Santos, M.D.M.; Pandolfo, G.M.; Boiteux, L.S.; Café-Filho, A.C. Biology, Pathogenicity, and Haplotype Analyses of *Colletotrichum cliviae*: A Novel Soybean Anthracnose Agent in Warm Tropical Areas. *Trop. Plant Pathol.* **2018**, *43*, 439–451, doi:10.1007/s40858-018-0249-6.
83. Bhunjun, C.S.; Phukhamsakda, C.; Jayawardena, R.S.; Jeewon, R.; Promputtha, I.; Hyde, K.D. Investigating Species Boundaries in *Colletotrichum*. *Fungal Divers.* **2021**, *107*, 107–127, doi:10.1007/s13225-021-00471-z.
84. Hymowitz, T. On the Domestication of the Soybean. *Econ Bot* **1970**, *24*, 408–421, doi:10.1007/BF02860745.
85. Rogério, F.; Ciampi-Guillardi, M.; Barbieri, M.C.G.; Bragança, C. a. D.; Seixas, C.D.S.; Almeida, A.M.R.; Massola, N.S. Phylogeny and Variability of *Colletotrichum truncatum* Associated with Soybean Anthracnose in Brazil. *J. Appl. Microbiol.* **2017**, *122*, 402–415, doi:10.1111/jam.13346.
86. Sharma, S.K.; Gupta, G.K.; Ramteke, R. *Colletotrichum truncatum* [(Schw.) Andrus & W.D. Moore], the Causal Agent of Anthracnose of Soybean [*Glycine max* (L.) Merrill.] – a Review. *Soybean Res.* **2011**, *9*, 31–52.
87. Fouché, S.; Plissonneau, C.; Croll, D. The Birth and Death of Effectors in Rapidly Evolving Filamentous Pathogen Genomes. *Curr. Opin. Microbiol.* **2018**, *46*, 34–42, doi:10.1016/j.mib.2018.01.020.
88. Rogério, F.; Ciampi-Guillardi, M.; Barbieri, M.C.G.; Bragança, C.A.D.; Seixas, C.D.S.; Almeida, A.M.R.; Massola, N.S. Phylogeny and Variability of *Colletotrichum truncatum* Associated with Soybean Anthracnose in Brazil. *J. Appl. Microbiol.* **2017**, *122*, 402–415, doi:10.1111/jam.13346.
89. Rogério, F.; Gladieux, P.; Massola, N.S.; Ciampi-Guillardi, M. Multiple Introductions without Admixture of *Colletotrichum truncatum* Associated with Soybean Anthracnose in Brazil. *Phytopathology* **2019**, *109*, 681–689, doi:10.1094/PHYTO-08-18-0321-R.
90. Cai, L.; Hyde, K.D.; Taylor, P.W.J.; Weir, B.S.; Waller, J.M.; Abang, M.M.; Zhang, J.Z.; Yang, Y.H.; Phoulivong, S.; Liu, Z.Y.; et al. A Polyphasic Approach for Studying *Colletotrichum*. *Fungal Diver.* **2009**, *39*, 183–204.
91. Jayawardena, R. Notes on Currently Accepted Species of *Colletotrichum*. *Mycosphere* **2016**, *7*, 1192–1260, doi:10.5943/mycosphere/si/2c/9.
92. Dubrulle, G.; Picot, A.; Madec, S.; Corre, E.; Pawtowski, A.; Baroncelli, R.; Zivy, M.; Balliau, T.; Le Floch, G.; Pensec, F. Deciphering the Infectious Process of *Colletotrichum lupini* in Lupin through Transcriptomic and Proteomic Analysis. *Microorganisms* **2020**, *8*, 1621, doi:10.3390/microorganisms8101621.
93. Yang, H.-C.; Haudenschild, J.S.; Hartman, G.L. Multiplex Real-Time PCR Detection and Differentiation of *Colletotrichum* Species Infecting Soybean. *Plant Dis.* **2015**, *99*, 1559–1568, doi:10.1094/PDIS-11-14-1189-RE.
94. Emms, D.M.; Kelly, S. OrthoFinder: Solving Fundamental Biases in Whole Genome Comparisons Dramatically Improves Orthogroup Inference Accuracy. *Genome Biol.* **2015**, *16*, 157, doi:10.1186/s13059-015-0721-2.
95. Conway, J.R.; Lex, A.; Gehlenborg, N. UpSetR: An R Package for the Visualization of Intersecting Sets and Their Properties.

Bioinformatics **2017**, *33*, 2938–2940, doi:10.1093/bioinformatics/btx364.

96. Almagro Armenteros, J.J.; Tsirigos, K.D.; Sønderby, C.K.; Petersen, T.N.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 5.0 Improves Signal Peptide Predictions Using Deep Neural Networks. *Nat. Biotechnol.* **2019**, *37*, 420–423, doi:10.1038/s41587-019-0036-z.

97. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L.L. Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes¹Edited by F. Cohen. *J. Mol. Biol.* **2001**, *305*, 567–580, doi:10.1006/jmbi.2000.4315.

98. Pierleoni, A.; Martelli, P.; Casadio, R. PredGPI: A GPI-Anchor Predictor. *BMC Bioinf.* **2008**, *9*, 392, doi:10.1186/1471-2105-9-392.

99. Jorda, J.; Kajava, A.V. T-REKS: Identification of Tandem REpeats in Sequences with a K-MeanS Based Algorithm. *Bioinformatics* **2009**, *25*, 2632–2638, doi:10.1093/bioinformatics/btp482.

100. Magoc, T.; Salzberg, S.L. FLASH: Fast Length Adjustment of Short Reads to Improve Genome Assemblies. *Bioinformatics* **2011**, *27*, 2957–2963, doi:10.1093/bioinformatics/btr507.