



Pathogenic Adaptations Revealed by Comparative Genome Analyses of Two *Colletotrichum* spp., the Causal Agent of Anthracnose in Rubber Tree

Xianbao Liu, Boxun Li, Yang Yang, Jimiao Cai, Tao Shi, Xiaolan Zheng and Guixiu Huang*

Environment and Plant Protection Institute, Chinese Academy of Tropical Agricultural Sciences (CATAS), Key Laboratory of Integrated Pest Management on Tropical Crops, Ministry of Agriculture, Key Laboratory for Monitoring and Control of Tropical Agricultural Pests, Haikou, China

OPEN ACCESS

Edited by:

Koichi Tamano,
National Institute of Advanced
Industrial Science and Technology
(AIST), Japan

Reviewed by:

Abbot Okotie Oghenekaro,
University of Manitoba, Canada
Leandro Ferreira Moreno,
Fungal Biodiversity Centre (KNAW),
Netherlands

*Correspondence:

Guixiu Huang
hgxiu@vip.163.com

Specialty section:

This article was submitted to
Systems Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 07 February 2020

Accepted: 08 June 2020

Published: 21 July 2020

Citation:

Liu X, Li B, Yang Y, Cai J, Shi T,
Zheng X and Huang G (2020)
Pathogenic Adaptations Revealed by
Comparative Genome Analyses
of Two *Colletotrichum* spp.,
the Causal Agent of Anthracnose
in Rubber Tree.
Front. Microbiol. 11:1484.
doi: 10.3389/fmicb.2020.01484

Colletotrichum siamense and *Colletotrichum australisense* cause *Colletotrichum* leaf disease that differ in their symptoms in rubber tree (*Hevea brasiliensis*), and pathogenicity of these two fungal species is also not identical on different cultivars of rubber tree. This divergence is often attributed to pathogen virulence factors, namely carbohydrate-active enzymes (CAZymes), secondary metabolites (SM), and small-secreted protein (SSP) effectors. The draft genome assembly and functional annotation of potential pathogenicity genes of both species obtained here provide an important and timely genomic resource for better understanding the biology and lifestyle of *Colletotrichum* spp. This should pave the way for designing more efficient disease control strategies in plantations of rubber tree. In this study, the genes associated with these categories were manually annotated in the genomes of *C. australisense* GX1655 and *C. siamense* HBCG01. Comparative genomic analyses were performed to address the evolutionary relationships among these gene families in the two species. First, the size of genome assembly, number of predicted genes, and some of the functional categories differed significantly between the two congeners. Second, from the comparative genomic analyses, we identified some specific genes, certain higher abundance of gene families associated with CAZymes, CYP450, and SM in the genome of *C. siamense*, and Nep1-like proteins (NLP) in the genome of *C. australisense*.

Keywords: cell wall degrading enzymes, *Colletotrichum*, genome, *Hevea brasiliensis*, necrosis- and ethylene-inducing peptides, secondary metabolites associated genes, symptoms

INTRODUCTION

Hevea brasiliensis (Willd. ex A. Juss.) Müll. Arg., commonly known as the Para rubber tree, is native to the tropical rain forests of Central and South America, and the primary source of natural rubber. *Colletotrichum* leaf disease (CLD) caused by *Colletotrichum* species has become a serious biotic factor limiting rubber tree cultivation in plantations. The disease has been recorded and confirmed in most rubber growing countries such as China, Malaysia, Sri Lanka (Wastie, 1967; Liu et al., 1987; Jayasinghe et al., 1997). The fungus causing this disease had been originally

identified as *Colletotrichum heveae* Petch, but then presumed to be *Colletotrichum gloeosporioides* (Penzig & Sacc.). Yet *C. gloeosporioides* as well as *Colletotrichum acutatum* are often associated with anthracnose disease occurrences in rubber trees (Brown and Soepena, 1994; Jayasinghe et al., 1997). Furthermore, *C. dematium*, *C. crassipes*, *C. karstii*, and *C. boninense* specimens have also been isolated from infected rubber leaves, but were deemed of less economic importance than the aforementioned two *Colletotrichum* species because they happen sporadically (Wastie and Janardhanan, 1970; Cai et al., 2016; Jiang et al., 2016).

Taxonomically, much of the recent research and treatments have primarily focused on studying the different *Colletotrichum* species complexes. According to the latest work and updates, the genus consists of fourteen species complexes (Jayawardena et al., 2016; Marin-Felix et al., 2017; Damm et al., 2018). Within just the last decade, many cryptic and new species have been reported from rubber leaves with CLD symptoms; for example, five *Colletotrichum* species belonging to the *C. acutatum* species complex were reported from Sri Lanka (Damm et al., 2012; Hunupolagama et al., 2017). In China, however, *Colletotrichum siamense*, *Colletotrichum fructicola*, and *Colletotrichum ledongense* of the *C. gloeosporioides* species complex were newly associated with CLD of rubber tree (Liu et al., 2018). Two new species (*Colletotrichum australisinese* and *C. bannaense*) of the *C. acutatum* species complex were also discovered, although *C. siamense* (the *C. gloeosporioides* species complex) and *C. australisinese* (the *C. acutatum* species complex) are currently recognized as the predominant species in that country. However, the infectious symptoms on rubber tree arising from *C. siamense* and *C. australisinese* are significantly different (Liu et al., 2018). The thorough characterization of species causing anthracnose and knowledge of their respective molecular mechanisms of infection is critical for devising strategies to effectively control the spread of the disease in rubber tree plantations.

The *Colletotrichum* species have evolved a diverse range of lifestyles to colonize and obtain nutrients from their host species of monocot and/or dicot plants, ranging functionally from biotrophs, necrotrophs, hemibiotrophs to endophytes (De Silva et al., 2017). In recent years, progress has advanced toward a better understanding of their infection process at the molecular level and its correlation with the life history of each fungus (Villa-Rivera et al., 2017). The entire genomic sequence of some *Colletotrichum* species belonging to different species complexes, are now publicly available, providing an impetus to research of *Colletotrichum*–host tree interactions. Genus-wide comparative genome analyses of *Colletotrichum* species focusing on important gene classes, such as secretory proteases and carbohydrate-active enzymes (CAZymes), candidate effectors, and secondary metabolite (SM) biosynthesis genes, have revealed the core genes conserved in the genus *Colletotrichum*. Additionally, information on the amplification and contraction of certain gene families that could be attributed to their specific life histories and host range have been provided (O'Connell et al., 2012; Gan et al., 2013; Baroncelli et al., 2016; Gan et al., 2016; Buiate et al., 2017; Rao and Nandineni, 2017; Liang et al., 2018).

In this study, we report on the genome sequences of *C. siamense* in the *C. gloeosporioides* species complex (CGsc) and of *C. australisinese* in the *C. acutatum* species complex (CAsc), which, respectively, cause different symptoms in diseased rubber trees. Comparative genomic analyses were carried out between the two species. We laid special emphasis on analyzing the gene families encoding CAZymes, SM biosynthesis, and necrosis and ethylene-inducing peptide 1 (Nep1)-like proteins (NLPs), given their already known roles in plant–pathogen interactions, and characterized the content of these genes and their variation. Comparative genomics revealed reductions of gene families encoding CAZymes and SM biosynthesis within *C. australisinese* belonging to the CAsc. We also found evidence for an expansion of NLPs. Based on these patterns, we hypothesize that for *Colletotrichum* species with a broad host range, particularly those within CAsc, Lineage Specific Effector protein Candidates (LSECs) have diminished roles in interactions with plants, perhaps instead relying mainly on CAZymes, proteases, and NLPs for successful host colonization.

MATERIALS AND METHODS

Fungal Cultures and Infection Conditions

Colletotrichum australisinese GX1655 and *C. siamense* HBCG01 isolates were obtained from distinctive leaf lesions associated with the different disease symptoms (the brown conical or raised spots vs. the anthracnose spots) they, respectively, cause in rubber trees. Their species identity was confirmed by a polyphasic approach (Liu et al., 2018). The GX1655 isolate was cultured on potato dextrose agar (PDA) and preserved as conidial suspension amended with 15% glycerol at -80°C . It was afterward deposited in the China General Microbiological Culture Collection center (CGMCC) under the accession number CGMCC3.18886.

Genome Sequencing and Assembly

Genomic DNA of freshly collected mycelia was extracted with the QIAGEN® Genomic DNA kit. The BluePippin system was used for the library construction. Genome sequencing was performed on a GridION X5 platform at the Nextomics Genomic Sequencing Center (Wuhan, China). Raw reads were trimmed with GraphMap and Minimap software tools¹ to remove low quality reads and any reads having adaptor contamination (mean qscore >7). Clean reads were then *de novo* assembled using Canu v1.3 (Koren et al., 2017), for which the correction of each assembly was then assessed using Bwa 0.7.12-r1039 (Walker et al., 2014) and “Benchmarking Universal Single-Copy Orthologs” (BUSCO v3.0) (Simão et al., 2015; Seppey et al., 2019). The generated genomes were deposited at NCBI as BioProjects under accession PRJNA597657 and PRJNA597926.

Gene Prediction

The transposable elements present in assembled genome were identified using a combination of *de novo* and homology-based approaches. Specifically, RepeatMasker v4.0.5

¹<https://github.com/lbcb-sci/graphmap2>

(Tarailo-Graovac and Chen, 2009) and RepeatProteinMask (v4.0.5, a package within RepeatMasker) were used to identify known sequences within the DNA repeat database (RepBase v21.01) (Bao et al., 2015). For the *de novo* predictions, RepeatModeler V1.0.8 and LTRfinder² were jointly employed to build a *de novo* repeat library, after which RepeatMasker identified the repeats by using both the *de novo* repeat database and RepBase. Tandem repeats were annotated with RepeatMasker and the Tandem Repeats Finder (TRF; v4.07) (Benson, 1999).

Initial gene models were derived as statistically combined consensus models from both *de novo* and homology-based predictions. For *de novo* predictions, we used a Combined approach of Augustus (v3.2.1) (Stanke et al., 2008), GeneD (v1.4) (Parra et al., 2000), GenScan (Burge and Karlin, 1998), SNAP (Korf, 2004) and Glimmer (Majoros et al., 2004). For the homology-based predictions, the protein sequences from *Colletotrichum nymphaeae* and *Colletotrichum simmondsii* in the CAsc (Baroncelli et al., 2016), *C. fructicola* in the CGsc (Gan et al., 2013), and *Colletotrichum higginsianum* (Zampounis et al., 2016) were mapped onto the target genome, by using TblastN with an *E*-value cutoff of 1e-5. The ensuing aligned sequences as well as corresponding query proteins were then filtered and transferred into GeneWise (Birney, 2004) and GeMoMa (Keilwagen et al., 2018) to search for their accurately spliced alignments. Gene models predicted by the *de novo* and homology approaches were integrated using EvidenceModeler (EVM) (Haas et al., 2008), to generate a consensus gene set.

Functional Annotation of Genes

The predicted genes were annotated using five well known protein databases: Cluster of Orthologous Group of proteins (COG), gene ontology (GO), Swiss-Prot (using an *E*-value threshold = 1e-05), Non-Redundant Protein Database (NR), and Kyoto Encyclopedia of Genes and Genomes (KEGG). Carbohydrate-utilizing enzymes were identified and classified based on a BLASTP search against Carbohydrate-Active Enzyme (CAZy) database³ with a cut-off *E*-value of 1e-5 (Liang et al., 2018). The cytochrome P450 genes were identified by retaining only the best hits found, i.e., P450s with >40% shared identity were assigned to the same family, through a BLASTP search (using an *E*-value threshold = 1e-5) against the cytochrome P450 database⁴ (Fischer et al., 2007; Liang et al., 2018). Potential SM clusters were identified using antiSMASH fungal version 5.0.0rc1 (Medema et al., 2011). The adenylation domain (A domain) of non-ribosomal peptide synthetases (NRPS) proteins and polyketide synthases (PKS)-NRPS hybrids, the keto-synthase domains (KS) (also referred to as N- and C-terminal docking domains) of PKS proteins and PKS-NRPS hybrids were separately used for the phylogenetic analyses (Buiate et al., 2017). Necrosis and ethylene-inducing peptide 1-like proteins (NLPs) genes were identified and classed based on a search for the conserved motif GHRHDWE in the predicted proteins and a TblastN search against the current assembly of genomes, by using the six

previously published NLP sequences belonging to three different types (i.e., type-1, type-2, and type-3) (Kleemann et al., 2012; Oome and Van den Ackerveken, 2014). Only full-length protein sequences were used in that analysis. Those proteins transported out of the cell and into the extracellular space were identified using a battery of tools: namely, Signal-P (v4.1) (Petersen et al., 2011), TargetP⁵ TMHMM (v2.0) (Krogh et al., 2001), GPI-SOM (Fankhauser and MaËser, 2005), and WoLF PSORT (v0.2) (Horton et al., 2007), which were run sequentially.

Construction of Phylogenetic Trees

For phylogenetic analysis, the genomes and/or proteomes of 10 *Colletotrichum* species (Gan et al., 2013; Baroncelli et al., 2014a,b, 2016; Hacquard et al., 2016; Zampounis et al., 2016) were downloaded from NCBI. Predicted proteins encoded by a total of 12 *Colletotrichum* species genomes were filtered, and clustered into orthologous groups by Orthofinder version 2.3.8 (Emms and Kelly, 2019). Single copy ortholog groups were then extracted for phylogenomic tree construction. Independent ortholog groups were aligned with MAFFT version 7⁶ and then concatenated (Baroncelli et al., 2016). A substitution model and its parameter values were selected using ProtTest version 3.4 (Abascal et al., 2005; Baroncelli et al., 2016). A maximum-likelihood (ML) phylogenetic tree with 1000 bootstraps was constructed with RAXML version 8.1.1 (Stamatakis, 2006; Liang et al., 2018). *Verticillium alfalfae* (Klosterman et al., 2011) was used as outgroup. In the phylogenetic analysis of SM clusters and NLP between *C. australisinese* and *C. siamense*, amino acid sequences were aligned using MUSCLE of MEGA 6, for which the best-fitting amino-acid substitution model and parameter settings were chosen based on the “model test” of MEGA 6. To construct each phylogenetic tree, the maximum likelihood algorithm implemented in MEGA 6 was used (Kumar et al., 2016), with *n* = 1000 bootstrap replications.

RESULTS

Infection Conditions

In rubber tree plantations, *C. australisinese* mainly infects tender and immature leaves (phenophase: copper brown and light green leaves), as revealed initially by many brown conical or raised spots. As the disease progresses, the infected leaves become visibly wrinkled and twisted, eventually falling off the petiole (Figures 1A–C). On a mature leaf, the symptoms appear only as conical or raised spots, latterly perforated (Figure 1D). By contrast, the anthracnose spots caused by *C. siamense* infection appear as concentric rings that occur generally along the leaf margins, and occasionally in the middle of a mature leaf (Saha et al., 2002). These lesions are large, and may coalesce to form discernable larger-sized spots. The central portion of this spot is light brown and papery. Another noteworthy leaf spot symptom associated with *C. siamense* infection is the almost circular papery lesions having

²http://tlife.fudan.edu.cn/tlife/ltr_finder/

³www.cazy.org

⁴<http://cyped.biocatnet.de/>

⁵<http://www.cbs.dtu.dk/services/TargetP/>

⁶<http://mafft.cbrc.jp/alignment/server>

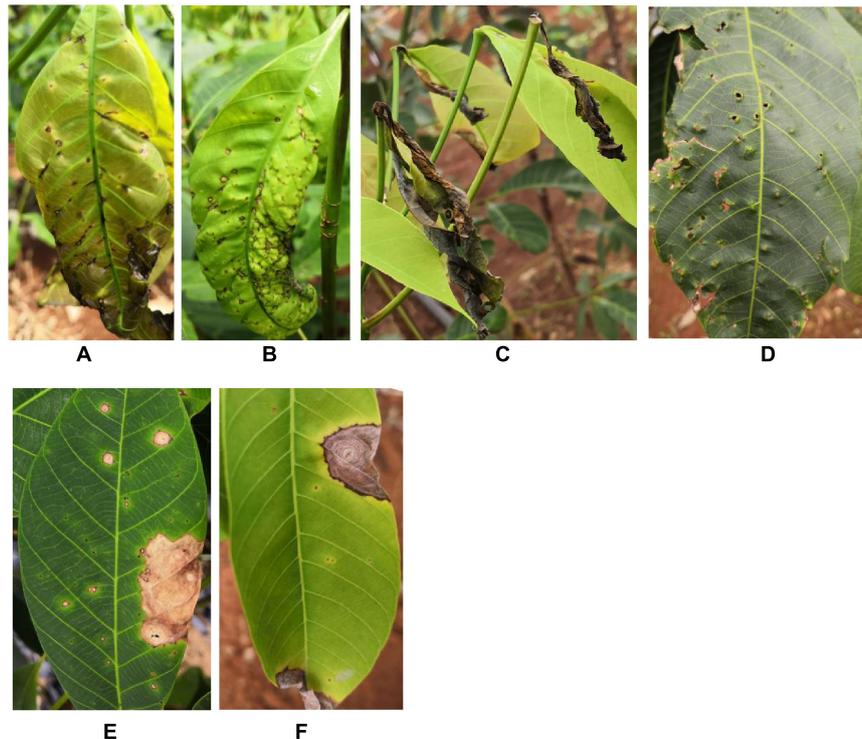


FIGURE 1 | Symptoms caused by *Colletotrichum siamense* and *Colletotrichum australisense* on rubber trees in the field. **(A,B,D)** Raised or wrinkled spots caused by *C. australisense* on the copper brown, light green, and mature leaves. **(C)** Die-back symptoms caused by *C. australisense* on rubber tree leaves. **(E,F)** Anthracnose symptoms caused by *C. siamense* on mature rubber leaves.

a dark brown center surrounded by a yellow halo (Saha et al., 2002; **Figures 1E,F**).

Genome Sequencing and Assembly and Gene Prediction

The genome assembly size of *C. australisense* GX1655 is distinctly different from *C. siamense*. BUSCO v3.0 was used to provide an estimate of assembly completeness. According to this analysis, the assembly of the *C. australisense* genome covered 95.5% of the total gene space while that of *C. siamense* covered 99.3% of it (**Table 1**).

For the *C. australisense* GX1655 genome, a total of 15,189 protein-coding genes were predicted, of which 26.9% (4095), 46.8% (7114), 58.4% (8870), 0.84% (127), and 98.2% (14,915) of the putative proteins could be, respectively, annotated based on COG, GO, Swiss-Prot, KEGGs, and NR, respectively. In the BUSCO analysis, 95.5% of the fungal core genes had “complete” hits and 95.2% had “complete and single-copy” hits, and ca. 5.65% of the assembly consisted of repeat sequences. From the *C. siamense* HBCG01 genome a total of 15,493 protein-coding genes were predicted—slightly more than obtained for the *C. australisense* GX1655—of which 29.1% (4,519), 50.4% (7,812), 62.4% (9,672), 2.76% (427), and 99.0% (15,438) could be annotated based on COG, GO, Swiss-Prot, KEGGs and NR, respectively. According to BUSCO, 98.3 and 96.9% of this species fungal core genes had hits of “complete” and

“complete and single-copy,” respectively, with repeat sequences amounting to ca. 1.85% of the assembly. For both species, the thoroughness of their genes’ annotation was demonstrated by the high percent values (**Table 1**). In the categorization of all genes based on GO, *C. siamense* HBCG01 in general has more genes in cellular component, molecular function, biological process. *C. australisense* GX1655 has only higher abundance of genes involved in signal transducer activity and molecular transducer activity (molecular function), membrane-enclosed lumen (cellular component), biological regulation

TABLE 1 | Assembly and gene prediction information of *Colletotrichum* spp. genomes.

	<i>C. siamense</i>	<i>C. australisense</i>
Assembly size (Mb)	58.4	55.3
Max contig (Mb)	7.62	10.04
N50 (Mb)	4.58	5.69
BUSCO complete	99.3%	95.5%
BUSCO partial	99.0%	95.5%
Repeat elements	1.85%	5.65%
Number of predicted genes	15493	15189
Average gene length (bp)	1612.97	1586.30
Average exons number per gene	2.90	3.17
Average introns number per gene	1.90	2.17

and cellular component organization or biogenesis (biological process) (**Supplementary Figure 1**).

Phylogenetic Analyses

In orthogroups analysis, we found 13,018 orthogroups in the *C. australisinese*, and 12,930 orthogroups in the *C. siamense*. A total of 4,782 single-copy orthologs were found in all species used in this study. 42 orthogroups had no homologs in any other species used in this study, and hence represented the species-specific orthogroups that are unique to *C. australisinese*. However, 5 orthogroups are species-specific to *C. siamense*. To understand the evolutionary relationships among species within the genus *Colletotrichum*, a phylogenomic tree was constructed based on the concatenated alignment of 4,782 single-copy orthologs obtained from the genomes of 12 *Colletotrichum* species. The inferred phylogeny showed that *C. australisinese* was most closely related to *C. nymphaeae*. *C. siamense* clustered with *C. fructicola*. The CAsc and CGsc appear to be evolutionary very distant (**Figure 2**).

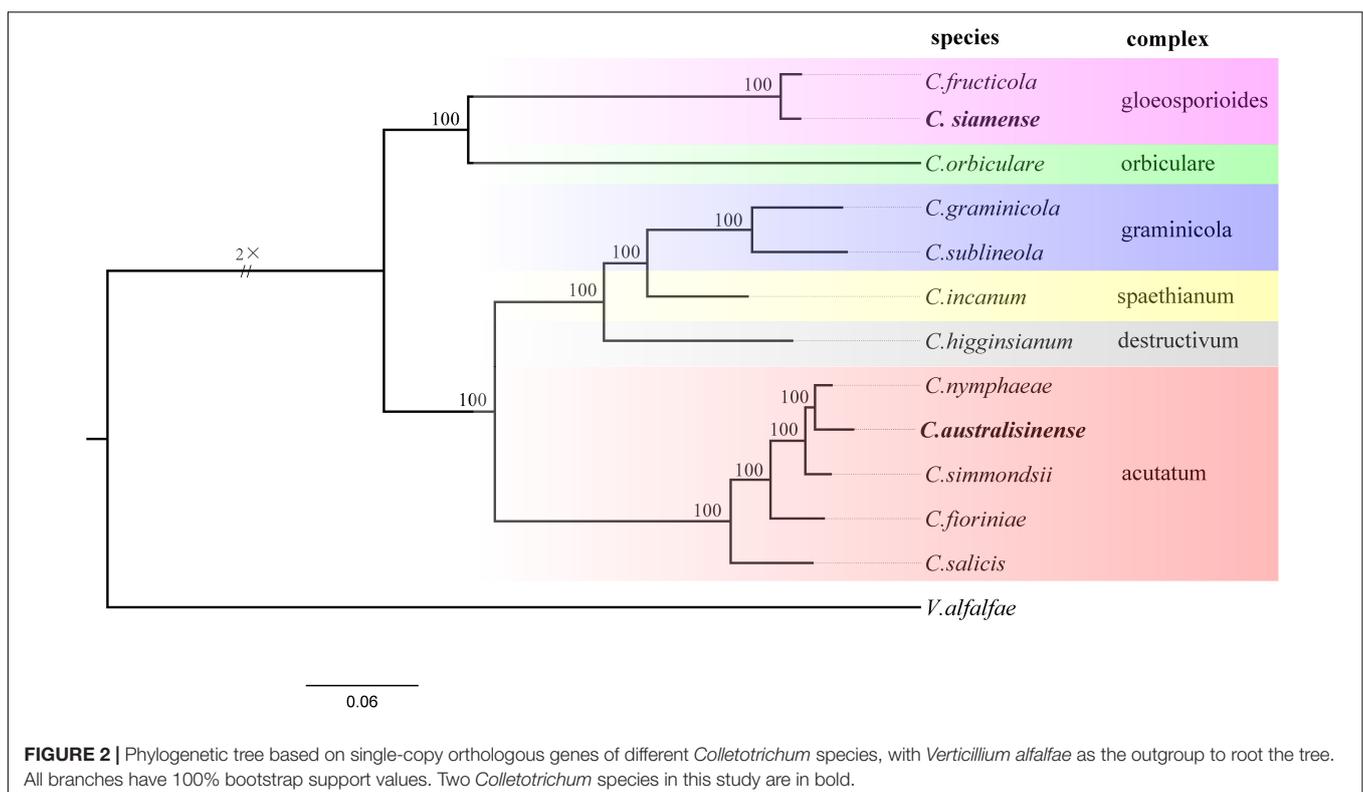
Secreted CAZymes

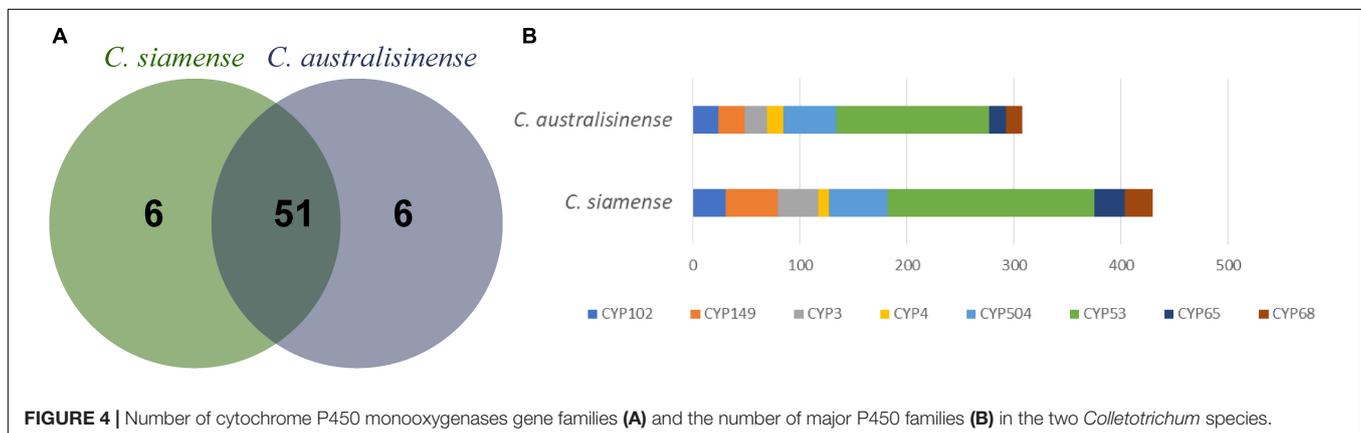
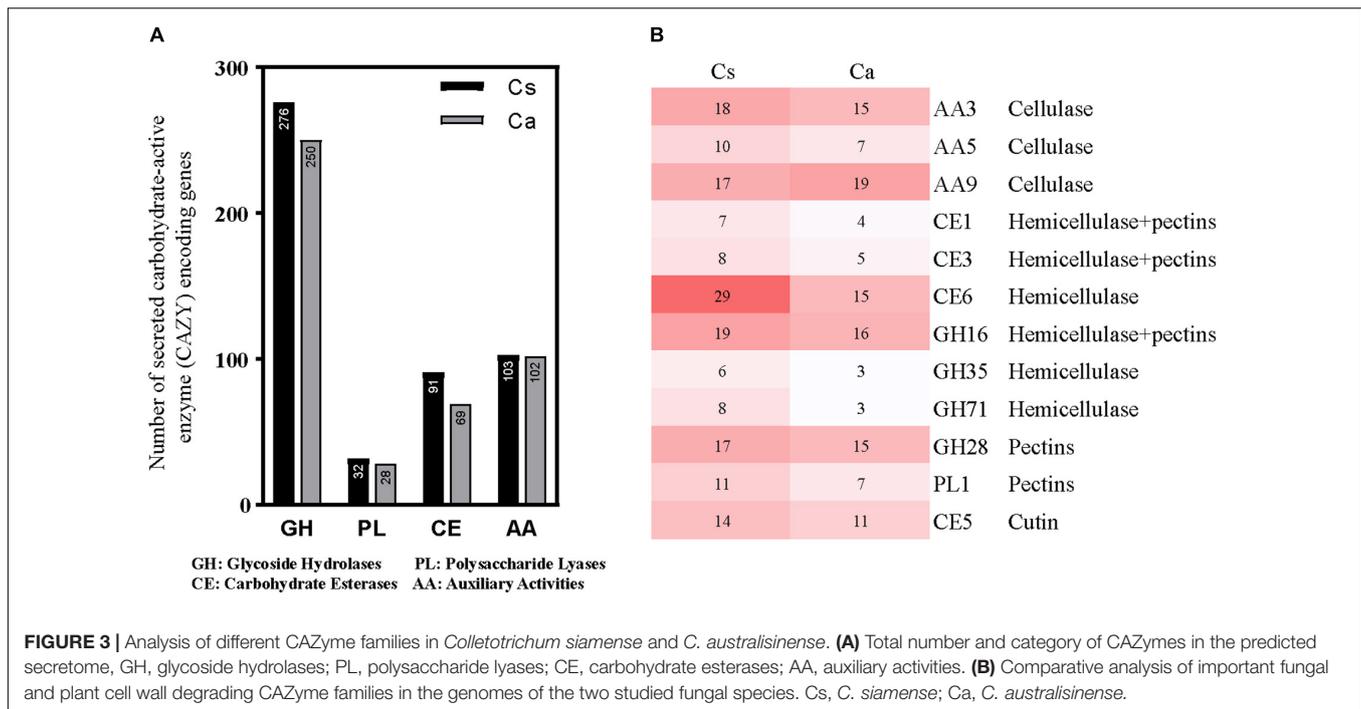
The plant cell wall primarily consists of three polysaccharide types (cellulose, hemicelluloses and pectins). Carbohydrate-metabolizing enzymes are vital for the degradation and utilization of these polysaccharides during host colonization by pathogenic fungi. In all, 451 genes of *C. australisinese* GX1655 associated with 88 CAZyme families were predicted to have secreted proteins; however, *C. siamense* HBCG01 harbored a larger repertoire of CAZymes in which 507 genes associated with

92 CAZyme families were predicted to have secreted proteins. The key classes of CAZymes, such as auxiliary activities (AA), carbohydrate esterases (CE), glycoside hydrolases (GH), and polysaccharide lyases (PL), were compared between species (**Figure 3A**). Evidently, *C. siamense* HBCG01 had higher copy numbers of these families, especially of CE and GH. Among the most amplified gene families were those involved in pectin degradation, such as PL1, GH28, and GH78, as well as those participating in the degradation of hemicelluloses and cellulose, such as CE1, CE3, CE6, GH16, GH35, GH71, and GH92 (**Figure 3B**). Hence, *C. siamense* evolution underwent a strong amplification of plant cell wall-degrading enzymes (PCWDEs). Auxiliary activities (AA) is another pertinent class of genes not belonging to CAZymes but instead linked to both biomass degradation and microbe-plant interaction (e.g., involved in breaking down lignin; Levasseur et al., 2013; Baroncelli et al., 2016). We found the total number of AA genes to be roughly similar in both species.

Cytochrome P450s

Cytochromes P450 (CYPs) are proteins of the superfamily containing heme as its cofactor, otherwise known as hemoproteins (Gonzalez and Gelboin, 1992). A total of 674 and 817 genes were predicted, respectively in *C. australisinese* and *C. siamense*, for which both numbers were associated with 57 families of CYPs (**Figure 4A**), which pointed to a significant higher abundance in P450 gene families of *C. siamense*. The family that expanded most was CYP53. Research on fungal P450s has revealed that the P450 family CYP53 can serve as a novel



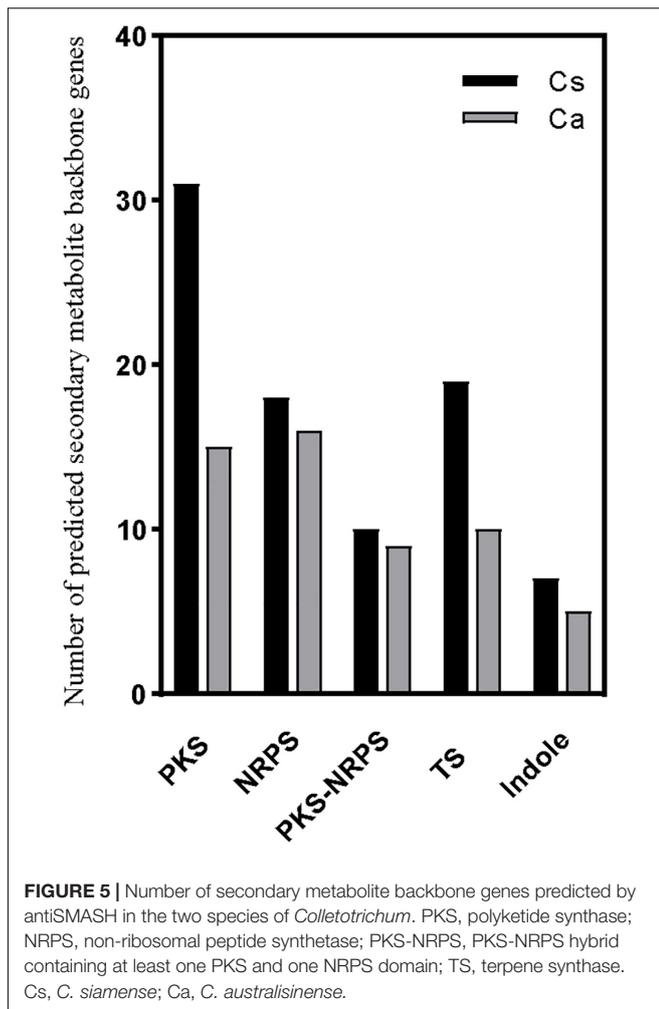


alternative anti-fungal drug target (Jawallapersand et al., 2014), and the CYP53 gene was earlier shown to be essential for fungal species survival (Fraser et al., 2002). The CYP149, CYP68, and CYP65 families corresponding to cytochrome P450s related to SM biosynthesis (Jimbo et al., 2000; Liang et al., 2018) also have obvious higher abundance (Figure 4B).

SM Synthesis Capacity

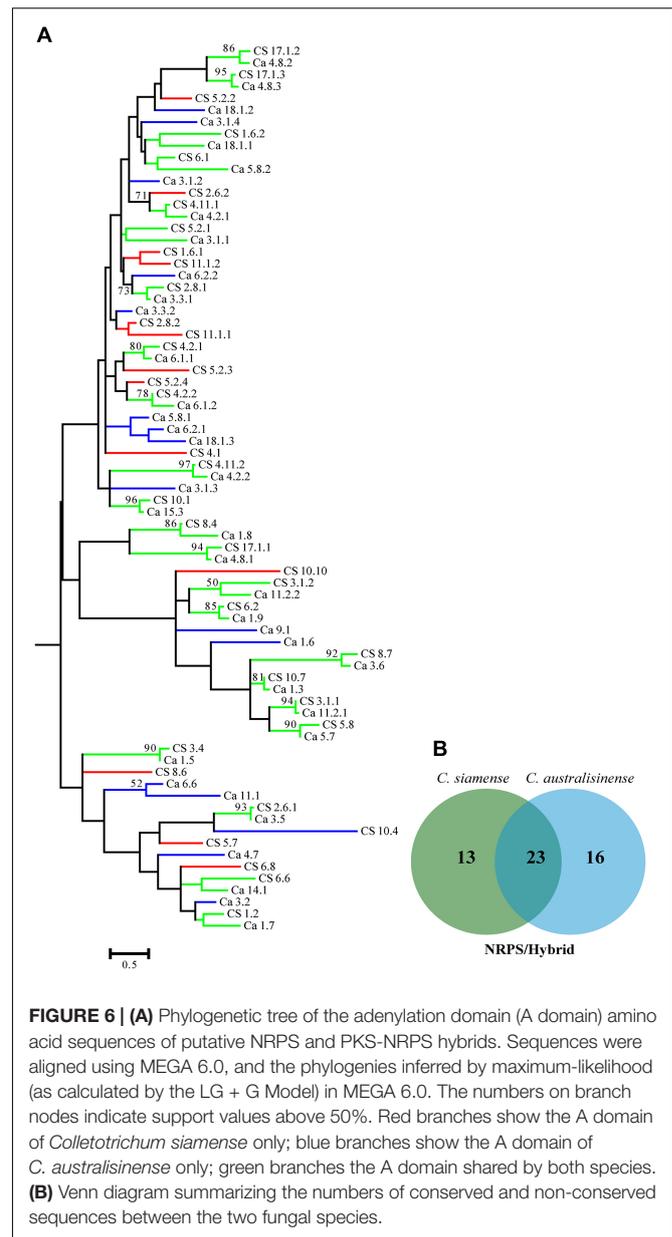
The fungal genus *Colletotrichum* produces an enormous array of SM that could figure prominently in its pathogenesis of plants. The biosynthesis of SM requires precursor genes, most of which are located inside fungal gene clusters. It is imperative to identify the genes involved in the codification of these enzymes, because these genes are often involved in different mechanisms of pathogenicity and/or virulence (Boettger and Hertweck, 2013; Moraga et al., 2018). Here, by

using antiSMASH, 85 SM gene clusters in the *C. siamense* genome were predicted: 31 type 1 polyketide synthases (t1PKS) and 18 NRPS genes, 7 t1PKS-NRPS hybrid genes, 19 terpene (TS) and 7 indole genes, and one each of t1PKS- type 3 polyketide synthases (t3PKS), NRPS-TS, and t1PKS-indole hybrid genes. Yet there were fewer SM gene clusters in the *C. australisinese* genome, whose 55 clusters comprised 15 t1PKS and 16 NRPS genes, 7 t1PKS-NRPS hybrid genes, 10 TS and 5 indole genes, and one each of NRPS-betalactone and NRPS-TS hybrid genes. The t1PKSs commonly responsible for the biosynthesis of macrolides were the most abundant clusters detected in the *C. siamense* genome (Figure 5). Some of these genes showed homology to known clusters in other fungi: eight gene clusters of *C. siamense* showed 100% similarity with alternapyrone (*Alternaria solani*), dimethylcoprogen (*Alternaria solani*), chrysogine (*Fusarium graminearum*), clavric acid



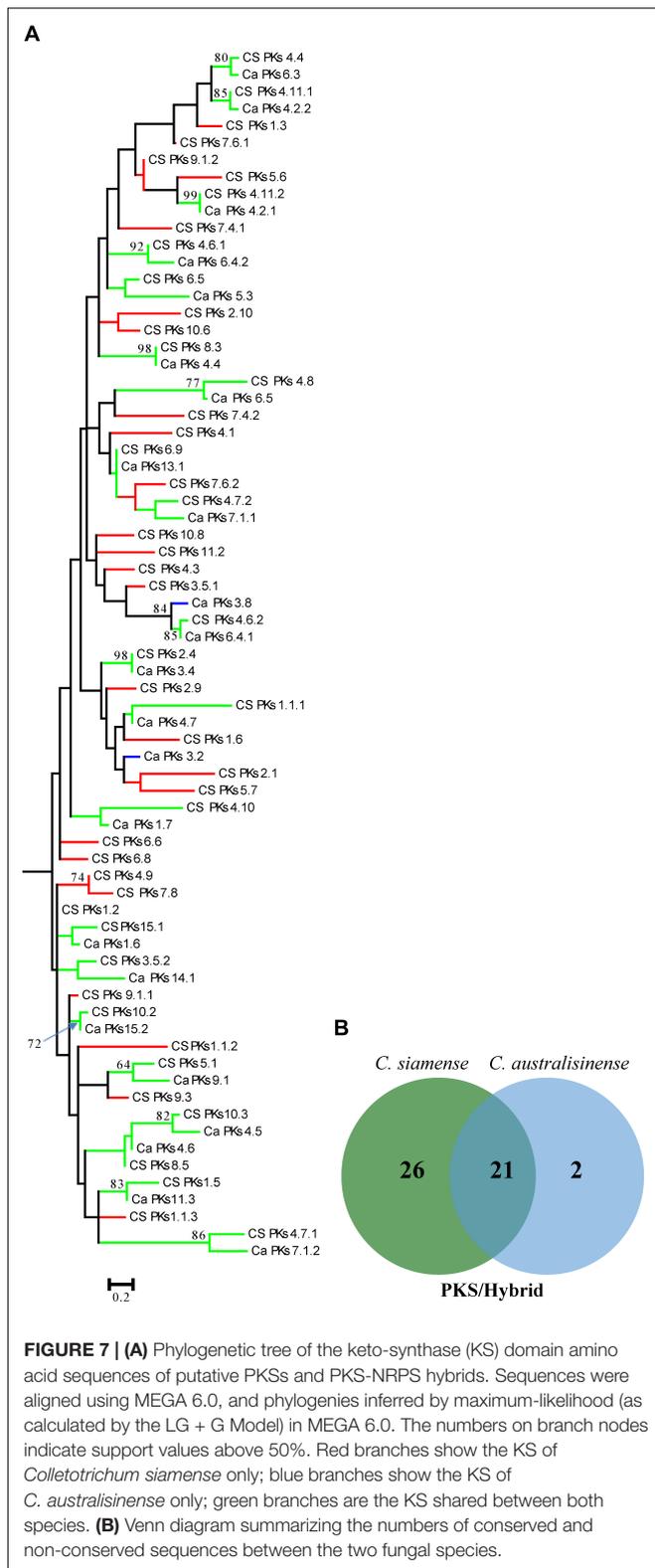
(*Hypholoma sublateritium*), fusarin (*Fusarium verticillioides*), phomopsins (*Phomopsis leptostromiformis*), gibberellin, and 1,3,6,8-tetrahydroxynaphthalene (*Nodulisporium* sp. ATCC74245) biosynthetic gene clusters. Of these clusters, in *C. australisinsense* genome only alternapyrone and 1,3,6,8-tetrahydroxynaphthalene biosynthetic gene clusters were found, which showed 100% similarity. Furthermore, two species-specific gene clusters of *C. australisinsense* showed 100% similarity with the pyranonigrin E (*Aspergillus niger* ATCC 1015) and deoxysambucinol/sambucinol/roridin E (*Fusarium sambucinum*) biosynthetic gene clusters.

We also investigated the relationships among the putative SM-associated proteins using the adenylation domain (A domain) of NRPS proteins and the keto-synthase (KS) of PKS in the two *Colletotrichum* spp. This phylogenetic analysis revealed a rich diversity, with relatively few NRPS-associated protein ortholog families conserved across both species (Figures 6A,B). Nonetheless, 11 PKS gene clusters—but none that included the PKS-NRPS hybrid—were apparently shared between *C. siamense* and *C. australisinsense*. Interestingly, *C. siamense* HBCG01 had a larger repertoire of keto-synthase (KS) in these PKS clusters (Figures 7A,B).



Necrosis and Ethylene-Inducing Peptide 1 (Nep1)-Like Proteins (NLPs)

Nep1-like proteins (NLP) are perhaps best known for their cytotoxic activity in dicot plants (Oome and Van den Ackerveken, 2014). These proteins function as effectors that boost pathogen virulence during the host colonization process by disintegrating the plasma membrane of plant cells. NLP are taxonomically widespread among microbes having very different lifestyles (Oome and Van den Ackerveken, 2014). Here, NLPs were identified through a search for the conserved motif GHRHDWE in the predicted proteins. Eleven proteins were predicted in *C. australisinsense*, and likewise eight in *C. siamense*, which belonged to the NEP1-like protein (NLP) family (Oome and Van den Ackerveken, 2014). Two NLPs of *C. siamense*,

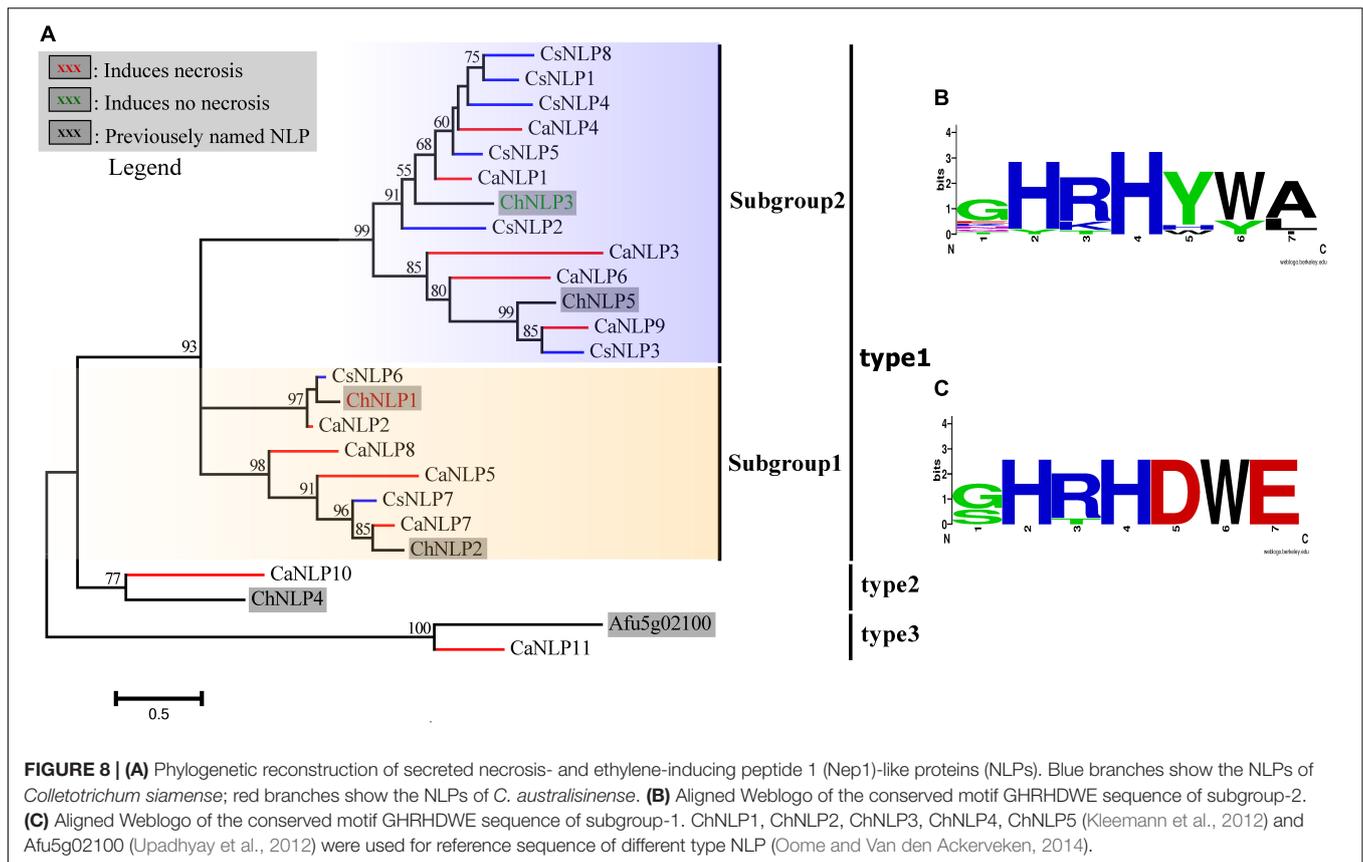


lacking their N-terminal sequence, were not included in this study, perhaps because of incomplete genome sequencing or due to being pseudogenes. Clear differences in both the

type and number of NLPs between the two *Colletotrichum* species were obtained. Among the predicted NLPs, nine of *C. australisinsense* were identified as type-1, plus one each of type-2 and type-3. In stark contrast, all NLPs for *C. siamense* belonged to type-1. According to the substitutions in the conserved GHRHDWE heptapeptide motif, diversification of type-1 NLP is discernable, in that it could be divided into two subgroups (i.e., subgroup-1 and subgroup-2). The subgroup-1 NLP contained a near fully conserved motif, whose members are confirmed participants in the induction of necrotrophy. However, the aspartic acid and glutamic acid residues (underlined) of the conserved heptapeptide motif (GHRHDWE) are not conserved in subgroup-2 NLP, whose members lack necrosis-inducing activity (Figure 8; Kleemann et al., 2012; Oome and Van den Ackerveken, 2014).

DISCUSSION

In our study, the whole-genome sequencing, assembly, and functional annotation of *C. australisinsense* and *C. siamense* characterized by distinct lifestyles on rubber tree were carried out. Pathogenicity-related genes were classified and compared, and marked duplication/reduction events were also identified. *Colletotrichum* is a large genus of *Ascomycete* fungi, containing many species with different life-histories capable of causing anthracnose or blight to a wide range of host plants (Jayawardena et al., 2016). The majority of *Colletotrichum* spp. are known for their hemibiotrophic lifestyle, which includes two successive stages of infection, biotrophy followed by necrotrophy. The particular life history of species in this genus depends on multiple factors, including the host identity, type of host tissue infected, and local environmental conditions (Perfect et al., 1999; De Silva et al., 2017). Although *C. australisinsense* and *C. siamense* form the dominant populations causing *Colletotrichum* leaf disease in rubber trees, they have a different life history. *C. australisinsense* has only been found on the rubber tree. By contrast, many *C. siamense* has a wide host range, and can colonize many host plant species, such as those in the genera *Carica*, *Coffea*, *Mangifera*, *Annona*, *Musa*, and *Camellia* (Udayanga et al., 2013; Liu et al., 2015). Genus-wide comparative genome analyses indicate that *Colletotrichum* species have tailored profiles for some enzymes according to their specific lifestyle, including CAZymes, SM synthetases, and peptidases (Gan et al., 2016; Buiate et al., 2017; De Silva et al., 2017; Liang et al., 2018). In the pathogenic process, these genes are expressed dynamically to fulfill stage-specific pathogenic functions (O'Connell et al., 2012; Gan et al., 2013; Rao and Nandineni, 2017). Effectors and SM enzymes are induced before penetration and during biotrophy, whereas most degradative enzyme and transporters are upregulated later, at the necrotrophic stage of growth (O'Connell et al., 2012). Hemibiotrophy in *C. orbiculare* is characterized by distinct stage-specific gene expression profiles of expanded classes of potential pathogenicity genes (Gan et al., 2013). The comparative genomic research has ever yielded key insights into the *Colletotrichum* fungi's evolution (Liang et al., 2018). The gene family expansions and contractions



are thought to be critical for shaping the host specificity of members of this genus (Gan et al., 2016; Baroncelli et al., 2016).

When compared with *C. australisense*, our results suggests that the major polysaccharide degrading enzyme (CAZyme) families underwent a species-specific increase of pectin degrading CAZymes (GH28, GH78, PL1), cutinases (CE5), and hemicelluloses and cellulose degrading CAZymes (CE1, CE3, CE6, GH16, GH35, GH71, GH92) in *C. siamense*, the species that causes anthracnose symptoms in a wide variety of hosts. Thus, *C. siamense* has a stronger ability to degrade plant cell wall components, such as xyloglucan, xylan, pectin, and cellulose. This offensive trait may help explain why, of the two *Colletotrichum* spp. studied, it is *C. siamense* which mainly infects the mature leaves of rubber tree, causing large anthracnose lesions. In contrast, the *C. australisense* infection of leaves only causes little raised spots that slowly expand.

P450s are diverse, heme-thiolate proteins, which enable primary and SM synthesis, fungal pathogenicity, and detoxification of plant-derived antimicrobial compounds. In this respect, there seems to be little difference between *C. siamense* and *C. australisense*, both of which contain nearly the same P450 categories. However, the gene number of particular P450 families has strongly expanded in *C. siamense*, especially that of CYP65, while CYP68 was found closely associated with the SM synthesis, and CYP102, CYP149, and CYP3 with toxic substance metabolism. These results correspond well with the number of SM gene clusters predicted in two species, which

may explain how these two pathogens can detoxify phytoalexins (Soby et al., 1996).

The SMs produced by fungal phytopathogens are typically associated with their respective pathogenicity and host range (Rao and Nandineni, 2017). When applied to host leaves, these phytotoxic metabolites will induce symptoms similar to those of anthracnose caused by *Colletotrichum* species, and they have been shown to play a significant role in pathogenesis and infection mechanisms (García-Pajón and Collado, 2003). The comparative analysis of the SM gene clusters, as predicted through antiSMASH, revealed considerably more of such clusters in *C. siamense*, especially PKS (31 vs. 15) and TS (19 vs. 10) are twofold higher than in *C. australisense*. Although most of the KS genes in the predicted *C. australisense* PKS clusters were shared with *C. siamense*, the latter did harbor many specific KS genes that could contribute to its genetic potential to generate diverse SMs. It has been appreciated for some time now that PKSs are important in facilitating host penetration by *Colletotrichum* species (Takano et al., 1995, 1997).

Nep1-like proteins (NLPs) induce necrosis and ethylene production when they infiltrate the extracellular space in leaves of dicot plant species (but not monocots) (Oome and Van den Ackerveken, 2014). The NLPs are taxonomically widespread among microbes having very different lifestyles. Based on their phylogenetic analysis of NLP sequences, Oome and Van den Ackerveken (2014) defined four groups belonging to three phylogenetic types. A great number of newly identified NLP

sequences, from a broad range of diverse organisms, have been studied to date, providing much valuable insight into the evolution and functions of these remarkable proteins (de Oliveira et al., 2012; Kleemann et al., 2012; Zhou et al., 2012; Oome and Van den Ackerveken, 2014; Chen et al., 2018). Importantly, each NLP type differs in its ability to induce necrosis. It has been suggested that an NLP acts in a manner extracellular to the plant cell to confer its cytotoxic activity, which may be achieved via different mechanisms (Kleemann et al., 2012; Oome and Van den Ackerveken, 2014). In this study, we found that the number and categories of NLPs in *C. australisnense* exceeded those of *C. siamense*: In the previous work, it was confirmed that ChNLP1 and ChNLP2 of *C. higginsianum* were upregulated during the transition from biotrophy to necrotrophy, the former having been shown to induce necrosis in plants. Most of the NLPs in *C. siamense* belong to subgroup II, including ChNLP3 and ChNLP5, which are unable to induce necrosis in infected host plants. Since the genes encoding non-cytotoxic NLP are expressed during the biotrophic and early stages of infection by these (hemi)biotrophic pathogens, this suggests they play a role in the penetration or establishment of infection (Kanneganti et al., 2006; Kleemann et al., 2012). Previous studies have revealed lineage-specific expansions of NLP families within the *C. acutatum* species complex. It may well be that a lineage specific evolution is unique to members of the *C. acutatum* species complex (Baroncelli et al., 2016; Liang et al., 2018).

The remarkable repertoire of NLPs found in the *C. acutatum* species complex might be linked to their potent ability to infect multiple host species. Some of these proteins may function as necrosis-inducing proteins, whereas other NLPs may contribute to overcoming early plant defense responses (Baroncelli et al., 2016). We suggest that *C. australisnense*—which belongs to the *C. acutatum* species complex—has a stronger ability to induce necrosis than does *C. siamense* of the *C. gloeosporioides* species complex. For species within the *acutatum* complex, infection strategies are usually based on intracellular hemibiotrophy in addition to subcuticular, intramural necrotrophy (De Silva et al., 2017). Nonetheless, infection by *C. gloeosporioides* was reportedly hemibiotrophic, in that both intracellular hemibiotrophic and intramural

necrotrophic infections can occur (Kim et al., 2004; O'Connell et al., 2012; Moraes et al., 2013; De Silva et al., 2017). For the intracellular hemibiotrophic strategy, transitioning from one state to another can vary depending on the host plant identity and ontogenic stage and the particular infecting *Colletotrichum* species, of which some engage in rather unique infection behaviors. Species in the *acutatum* complex commonly begin their life cycle with a short biotrophic phase before switching to a necrotrophic stage (De Silva et al., 2017). These developmental aspects may help to explain why *C. australisnense* mainly infects tender and immature leaves of rubber tree, leading to rapidly wrinkled lesions.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the GenBank, accession nos. PRJNA597657 and PRJNA597926.

AUTHOR CONTRIBUTIONS

XL conducted the experiments and analyzed the results. BL, YY, JC, TS, and XZ collected *Colletotrichum* isolates and performed the experiments. GH revised and approved the final version of the manuscript. All authors reviewed the manuscript.

FUNDING

This work was financially supported by the National Natural Science Fund Project (No. 31571966) and the project of support plan for modern agricultural talents of the Ministry of Agriculture (No. 0316001).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.01484/full#supplementary-material>

REFERENCES

- Abascal, F., Zardoya, R., and Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*. 21, 2104–2105. doi: 10.1093/bioinformatics/bti263
- Bao, W., Kojima, K. K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA* 6:11. doi: 10.1186/s13100-015-0041-9
- Baroncelli, R., Amby, D. B., Zapparata, A., Sabrina, S., Giovanni, V., and Gaétan Le, F. (2016). Gene family expansions and contractions are associated with host range in plant pathogens of the genus *Colletotrichum*. *BMC Genomics* 17:555. doi: 10.1186/s12864-016-2917-6
- Baroncelli, R., Sanz-Martín, J. M., Rech, G. E., Sukno, S., and Thon, M. R. (2014a). Draft genome sequence of *Colletotrichum sublineola*, a destructive pathogen of cultivated sorghum. *Genome Announc.* 2, 10–11.
- Baroncelli, R., Sreenivasaprasad, S., Sukno, S., Thon, M. R., and Holub, E. (2014b). Draft Genome Sequence of *Colletotrichum acutatum* Sensus Lato (*Colletotrichum fioriniae*). *Genome Announc.* 2:e00112-14. doi: 10.1128/genomeA.00112-14
- Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573–580. doi: 10.1093/nar/27.2.573
- Birney, E. (2004). GeneWise and Genomewise. *Genome Res.* 14, 988–995. doi: 10.1101/gr.1865504
- Boettger, D., and Hertweck, C. (2013). Molecular diversity sculpted by fungal PKS-NRPS hybrids. *Chembiochem* 14, 28–42. doi: 10.1002/cbic.201200624
- Brown, A. E., and Soepena, H. (1994). Pathogenicity of *Colletotrichum acutatum* and *C. gloeosporioides* on leaves of *Hevea* spp. *Mycol. Res.* 98, 264–266. doi: 10.1016/s0953-7562(09)80453-x
- Buiate, E. A. S., Xavier, K. V., Moore, N., Torres, M. F., Farman, M. L., and Schardl, C. L. (2017). A comparative genomic analysis of putative pathogenicity genes in the host-specific sibling species *Colletotrichum graminicola* and *Colletotrichum sublineola*. *BMC Genomics* 18:67. doi: 10.1186/s12864-016-3457-9
- Burge, C. B., and Karlin, S. (1998). Finding the genes in genomic DNA. *Curr. Opin. Struct. Biol.* 8, 346–354. doi: 10.1016/s0959-440x(98)80069-9

- Cai, Z. Y., Liu, Y. X., Shi, Y. P., Mu, H. J., and Li, G. H. (2016). First report of leaf anthracnose caused by *Colletotrichum Karstii* of rubber tree in China. *Plant Dis.* 100:2528. doi: 10.1094/PDIS-04-16-0577-PDN
- Chen, X. R., Huang, S. X., Zhang, Y., Zhang, Y., Sheng, G. L., Li, Y. P., et al. (2018). Identification and functional analysis of the NLP-encoding genes from the phytopathogenic oomycete *Phytophthora capsici*. *Mol. Genet. Genomics* 293, 541–555.
- Damm, U., Cannon, P. F., Woudenberg, J. H. C., and Crous, P. W. (2012). The *Colletotrichum acutatum* species complex. *Stud. Mycol.* 73, 37–113. doi: 10.3114/sim0010
- Damm, U., Sato, T., Alizadeh, A., Groenewald, J. Z., and Crous, P. W. (2018). The *Colletotrichum dracaenophilum*, *C. magnum* and *C. orchidearum* species complexes. *Stud. Mycol.* 92, 1–46. doi: 10.1016/j.simyco.2018.04.001
- de Oliveira, G. A. P., Pereira, E. G., Dias, C. V., Souza, T. L. E., Ferretti, G. D. S., and Cordeiro, Y. (2012). Monilophthora perniciosa necrosis- and ethylene-inducing protein 2 (MpNep2) as a metastable dimer in solution: structural and functional implications. *PLoS One* 7:e45620. doi: 10.1371/journal.pone.0045620
- De Silva, D. D., Crous, P. W., Ades, P. K., Hyde, K. D., and Taylor, P. W. J. (2017). Life styles of *Colletotrichum* species and implications for plant biosecurity. *Fungal Biol. Rev.* 31, 155–168. doi: 10.1016/j.fbr.2017.05.001
- Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20, 238.
- Fankhauser, N., and MaËser, P. (2005). Identification of GPI anchor attachment signals by a Kohonen self-organizing map. *Bioinformatics* 21, 1846–1852. doi: 10.1093/bioinformatics/bti299
- Fischer, M., Knoll, M., Sirim, D., Wagner, F., Funke, S., and Pleiss, J. (2007). The cytochrome P450 engineering database: a navigation and prediction tool for the cytochrome P450 protein family. *Bioinformatics* 23, 2015–2017. doi: 10.1093/bioinformatics/btm268
- Fraser, J. A., Davis, M. A., and Hynes, M. J. (2002). The gene *gmdA*, encoding an amidase and *bzuA*, encoding a cytochrome P450, are required for benzamide utilization in *Aspergillus nidulans*. *Fungal Genet. Biol.* 35, 135–146. doi: 10.1006/fgbi.2001.1307
- Gan, P., Ikeda, K., Irieda, H., Narusaka, M., O'Connell, R. J., and Narusaka, Y. (2013). Comparative genomic and transcriptomic analyses reveal the hemibiotrophic stage shift of *Colletotrichum* fungi. *New Phytologist*. 197, 1236–1249. doi: 10.1111/nph.12085
- Gan, P., Narusaka, M., Kumakura, N., Tsushima, A., Takano, Y., and Narusaka, Y. (2016). Genus-Wide comparative genome analyses of *Colletotrichum*, species reveal specific gene family losses and gains during adaptation to specific infection lifestyles. *Genome Biol. Evol.* 8, 1467–1481. doi: 10.1093/gbe/evw089
- García-Pajón, C. M., and Collado, I. G. (2003). Secondary metabolites isolated from *Colletotrichum* species. *Nat. Prod. Rep.* 20, 426–431. doi: 10.1039/b302183c
- Gonzalez, F. J., and Gelboin, H. V. (1992). Human cytochromes P450: evolution and cDNA-directed expression. *Environ. Health Perspect.* 98, 81–85. doi: 10.1289/ehp.929881
- Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., et al. (2008). Automated eukaryotic gene structure annotation using evidencemodeler and the program to assemble spliced alignments. *Genome Biol.* 9:R7. doi: 10.1186/gb-2008-9-1-r7
- Hacquard, S., Kracher, B., Hiruma, K., Münch, P. C., Garrido-Oter, R., and Thon, M. R. (2016). Survival trade-offs in plant roots during colonization by closely related beneficial and pathogenic fungi. *Nat. Commun.* 7:11362. doi: 10.1038/ncomms11362
- Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., et al. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 35, 585–587.
- Hunupolagama, D. M., Chandrasekharan, N. V., Wijesundera, W. S. S., Kathriarachchi, H. S., Fernando, T. H. P. S., and Wijesundera, R. L. C. (2017). Unveiling members of *Colletotrichum acutatum* species complex causing *Colletotrichum* leaf disease of *Hevea brasiliensis* in sri lanka. *Curr. Microbiol.* 74, 747–756. doi: 10.1007/s00284-017-1238-6
- Jawallapersand, P., Mashele, S. S., Kovačić, L., Stojan, J., and Komel, R. (2014). Cytochrome P450 monooxygenase CYP53 family in fungi: comparative structural and evolutionary analysis and its role as a common alternative antifungal drug target. *PLoS One* 9:e107209. doi: 10.1371/journal.pone.0107209
- Jayasinghe, C. K., Fernando, T. H. P. S., and Priyanka, U. M. S. (1997). *Colletotrichum acutatum* is the main cause of *Colletotrichum* leaf disease of rubber in Sri Lanka. *Mycopathologia* 137, 53–56.
- Jayawardena, R. S., Hyde, K. D., Damm, U., Cai, L., Cai, L., Liu, M., et al. (2016). Notes on currently accepted species of *Colletotrichum*. *Mycosphere* 7, 1192–1260. doi: 10.5943/mycosphere/si/2c/9
- Jiang, G. Z., Shi, Z. W., Liu, Y. X., and Hu, C. Y. (2016). A New *Colletotrichum* Isolated from *Hevea brasiliensis*. *Trop. Agric. Sci. Technol.* 39, 5–8.
- Jimbo, M., Muramoto, K., Kamiya, H., and Yamaguchi, M. (2000). Cloning of the *Microcystis aeruginosa* M228 Lectin (MAL) Gene. *Biochem. Biophys. Res. Commun.* 273, 499–504. doi: 10.1006/bbrc.2000.2961
- Kanneganti, T. D., Huitema, E., Cakir, C., and Kamoun, S. (2006). Synergistic interactions of the plant cell death pathways induced by *Phytophthora infestans* Nep1-like protein PiNPP1.1 and INF1 elicitor. *Mol Plant Microbe Interact.* 19, 854–863. doi: 10.1094/mpmi-19-0854
- Keilwagen, J., Hartung, F., Paulini, M., Twardziok, S. O., and Grau, J. (2018). Combining RNA-seq data and homology-based gene prediction for plants, animals and fungi. *BMC Bioinformatics* 19:189. doi: 10.1186/s12859-018-2203-5
- Kim, K.-H., Yoon, J.-B., Park, H.-G., Park, E. W., and Kim, Y. H. (2004). Structural modifications and programmed cell death of chili pepper fruit related to resistance responses to *Colletotrichum gloeosporioides* infection. *Phytopathology* 94, 1295–1304. doi: 10.1094/PHTO.2004.94.12.1295
- Kleemann, J., Rincon-Rivera, L. J., Takahara, H., Neumann, U., van Themaat, E. V. L., and van der Does, H. C. (2012). Sequential delivery of host-induced virulence effectors by appressoria and intracellular hyphae of the phytopathogen *Colletotrichum higginsianum*. *PLoS Pathog.* 8:e1002643. doi: 10.1371/journal.ppat.1002643
- Klosterman, S. J., Subbarao, K. V., Kang, S., Veronese, P., Gold, S. E., and Thomma, B. P. H. J. (2011). Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. *PLoS Pathog.* 7:e1002137. doi: 10.1371/journal.ppat.1002137
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116
- Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics* 5:59. doi: 10.1186/1471-2105-5-59
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 33:870.
- Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., and Henrissat, B. (2013). Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels.* 6:41. doi: 10.1186/1754-6834-6-41
- Liang, X., Wang, B., Dong, Q., Li, L., Rollins, J. A., and Zhang, R. (2018). Pathogenic adaptations of *Colletotrichum* fungi revealed by genome wide gene family evolutionary analyses. *PLoS One* 13:e0196303. doi: 10.1371/journal.pone.0196303
- Liu, X., Li, B., Cai, J., Zheng, X., Feng, Y., and Huang, G. (2018). *Colletotrichum* species causing anthracnose of rubber trees in China. *Sci. Rep.* 8:10435. doi: 10.1038/s41598-018-28166-7
- Liu, F., Weir, B. S., Damm, U., Crous, P. W., Yang, Y., and Liu, B. (2015). Unravelling *Colletotrichum* species associated with *Camellia*, with the combined use of ApMat and GS loci to resolve species in the *C. gloeosporioides* complex. *Pers. Mol. Phyl. Evol. Fungi* 35, 63–86. doi: 10.3767/003158515x687597
- Liu, X. J., Yang, Y. T., and Leng, H. Q. (1987). Identification of species and forms of *colletotrichum gloeosporioides* in rubber growing regions in south China. *Chin. J. Trop. Crops* 8, 93–101.
- Majoros, W. H., Pertea, M., and Salzberg, S. L. (2004). TigrScan and GlimmerHMM: two open-source ab initio eukaryotic gene-finders. *Bioinformatics* 20, 2878–2879. doi: 10.1093/bioinformatics/bth315
- Marin-Felix, Y., Groenewald, J. Z., Cai, L., Chen, Q., Marincowitz, S., and Barnes, I. (2017). Genera of phytopathogenic fungi: GOPHY 1. *Stud. Mycol.* 86, 99–216. doi: 10.1016/j.simyco.2017.04.002
- Medema, M. H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., and Fischbach, M. A. (2011). antiSMASH: rapid identification, annotation and

- analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* 39, W339–W346. doi: 10.1093/nar/gkr466
- Moraes, S. R. G., Tanaka, F. A. O., and Massola, N. S. (2013). Histopathology of *Colletotrichum gloeosporioides* on guava fruit (*Psidium guajava* L.). *Rev. Bras. Frutic.* 35, 657–664. doi: 10.1590/S0100-29452013000200039
- Moraga, J., Gomes, W., Pinedo, C., Cantoral, J. M., Hanson, J. R., and Carbu, M. (2018). The current status on secondary metabolites produced by plant pathogenic *Colletotrichum* species. *Phytochem. Rev.* 18, 215–239. doi: 10.1007/s11101-018-9590-0
- O'Connell, R. J., Thon, M. R., Hacquard, S., Amyotte, S. G., Kleemann, J., and Torres, M. F. (2012). Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nat. Genet.* 44, 1060–1065. doi: 10.1101/gr.10.4.511
- Oome, S., and Van den Ackerveken, G. (2014). Comparative and functional analysis of the widely occurring family of nep1-like proteins[J]. *Mol. Plant Microbe Interact.* 27, 1081–1094. doi: 10.1094/mpmi-04-14-0118-r
- Parra, G., Blanco, E., and Guigo, A. (2000). GeneID in *Drosophila*. *Genome Res.* 10, 511–515. doi: 10.1101/gr.10.4.511
- Perfect, S. E., Hughes, H. B., O'Connell, R. J., and Green, J. R. (1999). *Colletotrichum*: a model genus for studies on pathology and fungal–plant interactions. *Fungal Genet. Biol.* 27, 186–198. doi: 10.1006/fgbi.1999.1143
- Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701
- Rao, S., and Nandineni, M. R. (2017). Genome sequencing and comparative genomics reveal a repertoire of putative pathogenicity genes in chilli anthracnose fungus *Colletotrichum truncatum*. *PLoS One* 12:e0183567. doi: 10.1371/journal.pone.0183567
- Saha, T., Kumar, A., Ravindran, M., Jacob, C. K., Roy, B., and Nazeer, M. A. (2002). Identification of *Colletotrichum acutatum* from rubber using random amplified polymorphic DNAs and ribosomal DNA polymorphisms. *Mycol. Res.* 106, 215–221. doi: 10.1017/s0953756201005342
- Seppy, M., Manni, M., and Zdobnov, E. M. (2019). BUSCO: assessing genome assembly and annotation completeness. *Methods Mol. Biol.* 1962, 227–245. doi: 10.1007/978-1-4939-9173-0_14
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210. doi: 10.1093/bioinformatics/btv351
- Soby, S., Caldera, S., Bates, R., and VanEtten, H. (1996). Detoxification of the phytoalexins maackiain and medicarpin by fungal pathogens of alfalfa. *Phytochemistry* 41, 759–765. doi: 10.1016/0031-9422(95)00684-2
- Stamatakis, A. (2006). RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446
- Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* 24, 637–644. doi: 10.1093/bioinformatics/btn013
- Takano, Y., Kubo, Y., Kuroda, I., and Furusawa, I. (1997). Temporal transcriptional pattern of three melanin biosynthesis genes, PKS1, SCD1 and THR1, in appressorium-differentiating and nondifferentiating conidia of *Colletotrichum lagenarium*. *Appl. Env. Microbiol.* 63, 351–354. doi: 10.1128/aem.63.1.351-354.1997
- Takano, Y., Kubo, Y., and Shimizu, K. (1995). Structural analysis of PKS1, a polyketide synthase gene involved in melanin biosynthesis of *Colletotrichum lagenarium*. *Mol. Gen. Genet.* 249, 162–167. doi: 10.1007/bf00290362
- Tarailo-Graovac, M., and Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics* 25, 4.10.1–4.10.14. doi: 10.1002/0471250953.bi0410s25
- Udayanga, D., Manamgoda, D., Liu, X., Chukeatirote, E., and Hyde, K. (2013). What are the common anthracnose pathogens of tropical fruits?. *Fungal Diversity* 61, 165–179. doi: 10.1007/s13225-013-0257-2
- Upadhyay, S. K., Gautam, P., Pandit, H., Singh, Y., Basir, S. F., and Madan, T. (2012). Identification of fibrinogen-binding proteins of *Aspergillus fumigatus* using proteomic approach. *Mycopathologia* 173, 73–82. doi: 10.1007/s11046-011-9465-z
- Villa-Rivera, M. G., Conejo-Saucedo, U., Lara-Marquez, A., Cano-Camacho, H., Lopez-Romero, E., and Zavala-Paramo, M. G. (2017). The role of virulence factors in the pathogenicity of *Colletotrichum* sp. *Curr. Prot. Peptide Sci.* 18, 1–14. doi: 10.2174/1389203717666160813160727
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., and Sakhthikumar, S. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. doi: 10.1371/journal.pone.0112963
- Wastie, R. L. (1967). Gloeosporium leaf disease of rubber in West Malaysia. *Planter* 43, 553–565.
- Wastie, R. L., and Janardhanan, P. S. (1970). Pathogenicity of *Colletotrichum gloeosporioides*, *C. dematium* and *C. crassipes* to leaves of *Hevea brasiliensis*. *Trans. Br. Mycol. Soc.* 54, 150–152.
- Zampounis, A., Pigné, S., Dallery, J. F., Wittenberg, A. H. J., Zhou, S., and Schwartz, D. C. (2016). Genome sequence and annotation of *Colletotrichum higginsianum*, a causal agent of crucifer anthracnose disease. *Genome Announc.* 4:e00821-16. doi: 10.1128/genomeA.00821-16
- Zhou, B. J., Jia, P. S., Gao, F., and Guo, H. S. (2012). Molecular characterization and functional analysis of a necrosis- and ethylene-inducing, protein-encoding gene family from *Verticillium dahliae*. *Mol Plant Microbe Interact.* 25, 964–975.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Liu, Li, Yang, Cai, Shi, Zheng and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.