Gene cluster conservation provides insight into cercosporin biosynthesis and extends production to the genus Colletotrichum

Ronnie de Jonge a,b,c,d,1,2, Malaika K. Ebert e,f,g,h,i, Callie R. Huit-Roelj h,i, Paramita Pal h, Jeffrey C. Suttle e, Rebecca E. Spanner e,f, Jonathan D. Neubauer h, Wayne M. Jurick I, Karina A. Stott e,h, Gary A. Secor I, Bart P. H. J. Thomma g, Yves Van de Peer b,c,d,j, Craig A. Townsend h,i, and Melvin D. Bolton e,f,2

a Plant-Microbe Interactions, Department of Biology, Science4Life, Utrecht University, 3584 CH, Utrecht, The Netherlands; b Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052, Ghent, Belgium; c VIB Center for Plant Systems Biology, B-9052 Ghent, Belgium; d Bioinformatics Institute Ghent, Ghent University, B-9052 Ghent, Belgium; e Red River Valley Agricultural Research Center, US Department of Agriculture, Fargo, ND 58102; f Department of Plant Pathology, North Dakota State University, Fargo, ND 58108; g Laboratory of Phytopathology, Wageningen University, 6708 PB Wageningen, The Netherlands; h Department of Chemistry, The Johns Hopkins University, Baltimore, MD 21218; i Food Quality Laboratory, US Department of Agriculture, Beltsville, MD 20705; and Department of Genetics, Genomics Research Institute, University of Pretoria, 0028 Pretoria, South Africa

Edited by Paul Schulze-Lefert, Max Planck Institute for Plant Breeding Research, Cologne, Germany, and approved May 4, 2018 (received for review July 18, 2017)

Species in the genus Cercospora cause economically devastating diseases in sugar beet, maize, rice, soy bean, and other major food crops. Here, we sequenced the genome of the sugar beet pathogen Cercospora beticola and it found it encodes 63 putative secondary metabolism gene clusters, including the cercosporin toxin biosynthesis (CTB) cluster. We show that the CTB gene cluster has experienced multiple duplications and horizontal transfers across a spectrum of plant pathogenic fungi, including the wide-host range Colletotrichum genus as well as the rice pathogen Magnaporthe oryzae. Although cercosporin biosynthesis has been thought to rely on an eight-gene CTB cluster, our phylogenomic analysis revealed gene collinearity adjacent to the established cluster in all CTB cluster-harboring species. We demonstrate that the CTB cluster is larger than previously recognized and includes cercosporin facilitator protein, previously shown to be involved with cercosporin autoresistance, and four additional genes required for cercosporin biosynthesis, including the final pathway enzymes that install the unusual cercosporin methylenedioxy bridge. Lastly, we demonstrate production of cercosporin by Colletotrichum floriniae, the first known cercosporin producer within this agriculturally important genus. Thus, our results provide insight into the intricate evolution and biology of a toxin critical to agriculture and broaden the production of cercosporin to another fungal genus containing many plant pathogens of important crops worldwide.

Among the most speciose genera in all of Fungi are the Cercosporinae (1). First described in 1863 (2), the genus has sustained a long history, largely due to notoriety as the causal agent of leaf spot diseases in a wide range of plants, including agriculturally important crops such as sugar beet, soybean, maize, and rice that together account for hundreds of millions of dollars in lost revenue annually to growers worldwide (3–8). Although Cercospora spp. share several characteristics associated with pathogenicity, such as penetration through natural openings and extracellular growth during the biotrophic stage of infection, most rely on the production of the secondary metabolite (SM) cercosporin (1) to facilitate infection (9, 10). Studies spanning nearly 60 y have made cercosporin a model perylenequinone (11), a class of SMs characterized by a core pentacyclic conjugated chromophore that gives rise to its phototoxicity. When exposed to ambient light, cercosporin is a potent producer of reactive oxygen species in the presence of oxygen (12), with a quantum efficiency of >80% (13). This small molecule is lipo-phlic and can readily penetrate plant leaves, leading to indiscriminate cellular damage within minutes of exposure (14).

Indec, cercosporin is nearly universally toxic to a wide array of organisms, including bacteria, mammals, plants, and most fungal species, with the key exception of cercosporin-producing fungi, which exhibit cercosporin autoresistance. To date, cercosporin has been reported to be produced only by Cercospora spp., with the single exception of the brassica pathogen Pseudocercosporella capsellae (15). However, Pseudocercosporella and Cercospora are phylogenetically closely related, residing in a large clade within Mycosphaerellaceae (16).

In contrast to the large body of information on cercosporin biology spanning several decades (17, 18), the cercosporin toxin natural product | perylenequinone | secondary metabolism | cercosporin | Cercospora

Significance

Species in the fungal genus Cercospora cause diseases in many important crops worldwide. Their success as pathogens is largely due to the secretion of cercosporin during infection. We report that the cercosporin toxin biosynthesis (CTB) gene cluster is ancient and was horizontally transferred to diverse fungal plant pathogens. Because our analyses revealed genes adjacent to the established CTB cluster with similar evolutionary trajectories, we evaluated their role in Cercospora beticola to show that four are necessary for cercosporin biosynthesis. Lastly, we confirmed that the apple pathogen Colletotrichum floriniae produces cercosporin, the first case outside the family Mycosphaerellaceae. Other Colletotrichum plant pathogens also harbor the CTB cluster, which points to a wider role that this toxin may play in virulence.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The sequences reported in this paper have been deposited in the National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA270309 (BioProject ID: PRJNA270309). Custom code and auxiliary data are available through Zenodo (https://doi.org/10.5281/zenodo.1156551) and Figshare (http://doi.org/10.6084/m9.figshare.4056522), respectively.


To whom correspondence may be addressed. Email: r.dejonge@uu.nl, ctownsend@jhu.edu, or melvin.bolton@ars.usda.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1712798115/-/DCSupplemental.

Published online May 29, 2018.

www.pnas.org/cgi/doi/10.1073/pnas.1712798115

PNAS | vol. 115 | no. 24 | E5459-E5466

Agricultural Sciences
biosynthesis (CTB) gene cluster was only recently resolved in *Cercospora nicotianae* (19). The keystone enzyme for cercosporin biosynthesis, CTB1, bears all the hallmarks of an iterative, nonreducing polyketide synthase (PKS) (20). Using CTB1 as a point of reference, the complete *C. nicotianae* CTB gene cluster was determined to consist of eight contiguous genes, of which six are believed to be responsible for cercosporin assembly (CTB1, 2, 3, 5, 6, and 7) (19, 21). The zinc finger transcription factor CTB8 coregulates expression of the cluster (19), while the major facilitator superfamily transporter CTB4 exports the final metabolite (22). Downstream of the CTB cluster are two ORFs encoding truncated transcription factors, while loci designated as ORF9 and ORF10 upstream of the CTB cluster are not regulated by light and are not believed to encode proteins with metabolic functions (19). Consequently, the clustering of eight genes with demonstrated coregulation by light that are flanked by ORFs with no apparent role in cercosporin biosynthesis has suggested that cercosporin production relies on the eight-gene CTB cluster (19). In this study, we used an evolutionary comparative genomics approach to show that the CTB gene cluster underwent multiple duplication events and was transferred horizontally across large taxonomic distances. Since these horizontal transfer events included genes adjacent to the canonical eight-gene CTB cluster, we used reverse genetics to show that the CTB cluster included additional genes in *Cercospora beticola*, including one gene that was previously shown to be involved with cercosporin autoresistance (23) and four previously unrecognized genes involved with biosynthesis. The CTB cluster was found in several *Colletotrichum* spp., and we confirmed that the apple pathogen *Colletotrichum fioriniae* can also produce cercosporin. As all earlier understanding of cercosporin biosynthesis has been unwittingly limited by a truncated set of genes in *Cercospora* spp., the full dimension of the gene cluster provides deeper insight into the evolution, biosynthesis, and dissemination of a fungal toxin critical to worldwide agriculture.

**Results**

**SM Cluster Expansion in *C. beticola***. *C. beticola* strain 09-40 was sequenced to 100-fold coverage and scaffolded with optical and genome maps, resulting in 96.5% of the 37.06-Mbp assembly being placed in 12 supercontigs, of which 10 are assumed to be chromosomes. Despite their ubiquitous presence in nature and cropping systems, genome sequences of *Cercospora* spp. are not well represented in public databases. Therefore, to aid comparative analysis within the *Cercospora* genus, we also sequenced the genome of *Cercospora berteroae* and reassembled the genome of *Cercospora canescens* (24) (*SI Appendix, Table S1*). To identify gene clusters responsible for biosynthesis of aromatic polyketides in *C. beticola*, we mined the genome to identify all SM clusters (25) and compared these with predicted clusters in related Dothideomycetes. The *C. beticola* genome possesses a total of 63 predicted SM clusters of several classes, representing an expanded SM repertoire with almost twice the number as compared with closely related Dothideomycetes fungi, which average 34 SM clusters (*SI Appendix, Table S2* and Dataset S1). Notably, *C. beticola* encodes 23 candidate nonribosomal peptide synthetase clusters, which is considerably higher than most Dothideomycetes fungi, which have an average of 13 (26). To identify the *C. beticola* PKS cluster responsible for cercosporin biosynthesis, we compared the sequence of the *C. nicotianae* CTB cluster (19) with predicted PKS clusters of *C. beticola*. To fill in sequencing gaps between genes in the *C. nicotianae* CTB cluster, we sequenced the genome of *C. nicotianae*, which showed that *C. beticola* PKS CBET3_00833 (CbCTB1) and flanking genes (CBET3_00830 – CBET3_00837) were ~96% identical to *C. nicotianae* CTB1 to CTB8, and all genes were collinear, strongly suggesting that this region houses the CTB cluster in *C. beticola* (*SI Appendix, Fig. S1*).

**Repeated Duplication and Lateral Transfer of the Cercosporin Biosynthetic Cluster**. To study the evolutionary relationships of *C. beticola* PKSs, we conducted large-scale phylogenomic analyses that included various previously characterized PKSs from selected species (27). Because resolving orthologous relationships among PKSs can predict the type of SM that will be synthesized, we first built a phylogenetic tree of the conserved core β-ketocarboxyl synthase domains of each PKS that resulted in separating PKS enzymes into four major groups (*SI Appendix, Fig. S2A*). Among the eight *C. beticola* nonreducing PKSs, phylogenetic analysis revealed significant similarity among CBCTB1, CBET3_10910-RA, and CBET3_11350-RA, which cluster at the base of the cercosporin clade (*SI Appendix, Fig. S2B*). Interestingly, genes flanking CBET3_10910-RA, but not CBET3_11350-RA, were also strikingly similar to CBCTB cluster genes (Fig. 1). Consequently, we hypothesized that the CBET3_10910 SM cluster is the result of a CTB cluster duplication. Since duplicated SM gene clusters appeared to be relatively rare in fungi (28), we investigated the origin and specificity of the CTB cluster and the putative duplication by searching for CBCTB1 homologs against a selected set of 48 published Ascomycetes proteomes (*SI Appendix, Table S3*) representing a diverse group of fungal orders. We identified CBCTB1 orthologs in the *Cercospora* spp. *C. berteroae* and *C. canescens* and confirmed its presence in *Cladosporium fulvum* (27) and *Parastagonospora nodorum* (29). Surprisingly, seven additional orthologs were identified in Sordariomycetes species (*Colletotrichum orbiculare*, *Colletotrichum gloeosporioides*, *C. fioriniae*, *Colletotrichum graminicola*, *Colletotrichum higginsianum*, and *Magnaporthe oryzae*), as well as one within Leotiomycetes (*Sclerotinia sclerotiorum*) (*SI Appendix, Fig. S3*), representing diverse taxa harboring CTB1. Analysis of sequence identity showed that intraspecies (CBCTB1 to CBET3_10910-RA) sequence identity (45%) was lower than the interspecies [e.g., CbCTB1 and *C. fulvum* CTB1 (Clafu1_196875)] sequence identity (55%) (*SI Appendix, Table S4*), suggesting that the CTB1 duplication event occurred before Dothideomycetes speciation.

To develop a “phylogenetic roadmap” that may explain CTB1 evolution, we used the process of “reconciliation” that takes into account both species and gene histories (30). Although not conclusive, reconciliation considers the costs of evolutionary events (i.e., gene duplications, transfers, and/or losses) to explain the most parsimonious evolutionary route to the present scenario (31). Reconciliation of the species tree (*SI Appendix, Fig. S4*) with the CTB1 protein tree revealed that the predicted evolutionary history of CTB1 can be characterized by four duplications, three
transfers, and widespread loss to most species analyzed (SI Appendix, Fig. S5A) and further corroborates our hypothesis that the CTB1 duplication event (D1) occurred early in Dothideomycetes speciation. Reconciliation also revealed an ancient CTB1 ortholog in S. sclerotiorum (SI Appendix, Fig. S5A), suggesting that CTB1 arose before speciation of Dothideomycetes. Duplications D2 to D4 arose after lateral transfer (T1) of CTB1 into the last common ancestor of Glomerellales. CTB1 was then transferred (T2) from a common ancestor in Glomerellales to M. oryzae (SI Appendix, Fig. S5A).

We extended the search for CTB cluster protein orthologs by scanning the 48 proteomes for homologs of CbCTB2 (CBET3_00830) to CbCTB8 (CBET3_00837), followed by phylogenetic tree construction and subtree selection (SI Appendix, Fig. S3B–N). This resulted in the identification of orthologs in the same set of species previously listed to contain CTB1, with the only exceptions in cases where CTB gene homologs were lost in a species. Although the loss of CTB6 and CTB7 orthologs limits reconciliation analysis of these gene families, reconciliation of the subtrees for CTB2, CTB3, CTB4, CTB5, and CTB8 (SI Appendix, Fig. S5B–H) supported a similar scenario as proposed for CTB1, involving at least two duplications (D1 and D2) and two horizontal transfer events (T1 and T2) that explain the present-day CTB scenario (Fig. 2). However, an alternative explanation involving a single transfer to an ancestral Glomerellales species followed by widespread loss in most species in this lineage, except for M. oryzae and the analyzed Colletotrichum spp. (Fig. 2 and SI Appendix, Table S5), cannot be ruled out by our analyses at this stage.

**Extension of the Predicted Cercosporin Biosynthetic Cluster Based on Microsynteny.** To further examine the CTB clusters across all recipient species, we generated pairwise alignments relative to the C. beticola CTB cluster and flanks. To our surprise, we observed a striking level of similarity outside the known eight CTB genes on the 3′ end of the cluster (Fig. 3) in all CTB-containing genomes. To investigate whether the amount of microsynteny observed for the CTB cluster and these flanking genes can be reasonably expected when comparing Dothideomycetes and Sordariomycetes genomes, we assessed the genome-wide microsynteny between the genomes of C. beticola and C. gloeosporioides and between C. beticola and M. oryzae. This analysis identified the CTB cluster together with its flanking genes as having the highest level of microsynteny among all regions in the genome between C. beticola and C. gloeosporioides and showed that the observed CTB microsynteny between C. beticola and M. oryzae was also higher than the genome-wide average (Fig. 4). Likewise, sequence identity of most CTB proteins between C. beticola and Colletotrichum spp., and to a lesser degree with M. oryzae, is higher compared with the genome-wide average (SI Appendix, Figs. S6 and S7). In contrast, sequence conservation of CTB8, a Zn2Cys6 transcription factor previously implicated for transcriptional regulation of the CTB cluster (19), appears much lower than that of other CTB and non-CTB proteins and, therefore, is suggestive of positive diversifying selection. Considering the level of microsynteny and protein conservation, we hypothesized that these flanking genes are part of the C. beticola CTB cluster. To test this proposal, we first determined the relative expression of all eight established C. beticola CTB genes as well as a number of flanking genes (CBET3_00828 to CBET3_00848) under light (cercosporin-inducing) vs. dark (cercosporin-repressing) conditions, which showed that all candidate CTB genes on the 3′ flank were induced in the light, except CBET3_00846 and CBET3_00848 (SI Appendix, Table S7). Functional annotation of these genes revealed one nonconserved phenylalanine amine lyase (CBET3_00840), the cercosporin facilitator protein (CFP) (23) (CBET3_00841), a candida α-ketoglutarate–dependent dioxygenase (CBET3_00842), a dehydratase (CBET3_00843), a β-ig-h3 fasciclin (CBET3_00844), a laccase (CBET3_00845), zinc finger domain-containing protein (CBET3_00846), and protein phosphatase 2A (CBET3_00847; SI Appendix, Table S7), several of which have functions associated with multidomain enzymes or polyketide biosynthesis in fungi or bacteria (19, 22, 32–36). Phylogenetic analyses of these flanking genes and reconciliation of their respective protein phylogenies (SI Appendix, Figs. S3 and S5) with the species tree (SI Appendix, Fig. S4) suggest that all genes except CBET3_00840, CBET3_00846, and CBET3_00847 have undergone highly similar evolutionary trajectories as the established CTB cluster genes (Fig. 2 and SI Appendix, Fig. S5), suggesting that the CTB cluster was transferred as a whole at least once, followed by species-specific evolutionary trajectories involving frequent gene loss as well as gene gain (Fig. 2). We further evaluated the hypothesis of horizontal cluster transfer using a comparative topology test that examines whether the determined tree topologies that support horizontal cluster transfer are significantly better than constrained topologies that would not support transfer. Tree topologies were compared using the Approximately Unbiased test (37), implemented in CONSEL (38) as previously described by Wisecaver and Rokas (39).

**Fig. 2.** Phylogenetic roadmap of CTB cluster evolution. Phylogenetic road-map detailing the proposed evolutionary trajectory of the CTB cluster involving horizontal gene transfer events from Capnodiales to Glomerellales (T1) and another from Glomerellales to Magnaportheales (T2), as well as multiple duplications (D1 to D4) and frequent gene loss (x). Cladogram of the phylogenetic relationship of Cercospora spp. and 45 other sequenced fungi. The unscaled tree was constructed using CVTree. Duplication nodes are marked with blue stars, losses are indicated by x’s, and transfers are highlighted by green arrows. Species without the CTB cluster are depicted in gray; those encompassing it are in black. An alternative and slightly less parsimonious scenario involving a single transfer from Capnodiales into the last common ancestor of Magnaportheales and Glomerellales is shown by the dashed arrow.
Fig. 3. Synteny and rearrangements of the conserved \textit{C. beticola} cercosporin biosynthetic cluster. The cercosporin biosynthetic cluster in \textit{C. beticola} (top line) and flanking genes are conserved in \textit{C. fulvum}, \textit{C. higginsianum}, \textit{C. graminicola}, \textit{M. oryzae}, and \textit{P. nodorum}. For all species, the displayed identifiers are transcript IDs, and the corresponding sequences can be retrieved from JGI MycoCosm or ORCAE. \textit{CTB} orthologs are colored relative to the \textit{C. beticola} \textit{CTB} cluster genes; the color key and annotated functions are highlighted below the \textit{CTB} cluster graphic. \textit{Cercospora}-specific \textit{CTB} genes \textit{CTB6} and \textit{CTB7} are underlined.

Constrained topologies in which we force either a monophyletic origin of all Dothideomycetes proteins or a monophyletic origin of all Sordariomycetes proteins were significantly worse than trees without such constraint (\textit{SI Appendix}, Table S6). Thus, the comparative topology tests support the previously determined topologies, which suggest horizontal cluster transfer.

\textbf{CTB Genes Essential for Cercosporin Biosynthesis.} To confirm individual gene contributions for cercosporin production, we generated single-gene deletion mutants of all candidate genes from \textit{CBET3_00840} to \textit{CBET3_00845} and tested their ability to produce cercosporin. Initial assays of selected mutants showed that cercosporin production in \textit{D\textit{CBET3_00844}} and \textit{D\textit{CBET3_00845}} mutants was abolished, while \textit{D\textit{CBET3_00842}} mutants accumulated only a red, cercosporin-like metabolite that migrated differently in potato dextrose agar (PDA) culture plates and TLC (\textit{SI Appendix}, Fig. S8). To provide more definitive analyses of cercosporin production, HPLC profiles were obtained from all candidate \textit{CTB} gene mutants and compared with WT cercosporin (Fig. 5A). Unlike other analyzed mutants, \textit{D\textit{CTB3_00840}} and \textit{D\textit{CBET3_00846}} produced compounds with HPLC profiles like cercosporin (Fig. 5A), suggesting that these genes are not involved with cercosporin biosynthesis. Together, these results corroborate our hypothesis that the \textit{CTB} cluster extends to at least \textit{CBET3_00845} at the 3' side and includes four additional \textit{CTB} biosynthetic genes, as well as \textit{ChCIP}. Consequently, we propose naming genes \textit{CBET3_00842}, \textit{CBET3_00843}, \textit{CBET3_00844}, and \textit{CBET3_00845} as \textit{CTB9} to \textit{CTB12}, respectively (\textit{SI Appendix}, Table S7).

\textbf{Precercosporin Isolation and Characterization.} To characterize the red metabolite that accumulated in the \textit{D\textit{CTB3_00840}} and \textit{D\textit{CBET3_00845}} \textit{CTB}10 mutants (Fig. 5A and \textit{SI Appendix}, Fig. S8), an ethyl acetate extract of the collected mycelia was analyzed by reverse-phase HPLC. At 280 nm, a single peak was observed in both mutant extracts with identical retention times and UV-Vis spectra (Fig. 5). This peak was compared with a reference sample of cercosporin produced by wild-type \textit{C. beticola}. The retention time of this peak was shorter than that of cercosporin, suggesting a more polar metabolite. Comparison of the UV-Vis spectra (Fig. 5 B–D) of the unknown compound and cercosporin revealed nearly identical chromophores, suggesting close structural relation. The exact mass of the metabolite from the mutant was determined (\textit{Δ\textit{S42/CTB9}}: m/z 537.1762, \textit{Δ\textit{S43/CTB10}}: m/z 537.1757, [M+H\textsuperscript{+}]), consistent with the elemental composition C\textsubscript{29}H\textsubscript{29}O\textsubscript{10}. This mass is 2 Da greater than that of cercosporin (+2 hydrogens), which led to a proposed structure for precercosporin (2) (Fig. 6). Alternative hydroquinones of cercosporin could be excluded simply on the basis of the UV-Vis spectral information and chemical instability. The presence of a free phenol in precercosporin in place of the unusual seven-membered methylenedioxy of cercosporin is consonant with the red shift of the long wavelength (\textlambda\textsubscript{max}) and the shorter HPLC retention time.

To firmly support the tentative structure of precercosporin, the crude extract of \textit{Δ\textit{S42/CTB9}} was further purified by reverse-phase HPLC. To obtain sufficient material for \textit{H}\textsuperscript{1}-NMR analysis, extractions were performed quickly and in low light and reduced temperature to slow apparent polymerization of precercosporin. The relative instability of precercosporin compared with cercosporin suggests a possible role for the methylenedioxy bridge in overall stability. Immediately evident from the \textit{H}-NMR spectrum (\textit{SI Appendix}, Fig. S9A), apart from its overall similarity to that of cercosporin itself, was not only the absence of the methylenedioxy singlet at 65.74 diagnostic of cercosporin but also the appearance of a new methoxyl signal at 64.28 and a phenol at 69.25. Consistent with the new asymmetry in precercosporin, two strongly hydrogen-bonded peri-hydroxy groups could be seen far downfield at around 15 ppm, and two aryl hydrogens were observed at 66.92 and 86.87. That these latter resonances are observed only in pairs (as are the two side-chain methyl doublets at around 0.6 ppm) and the other signals are doubled implies that precercosporin is formed as a single atropisomer having a helical configuration likely identical to that of cercosporin, although it is conceivable that \textit{CTB9} or \textit{CTB10} sets the final stereochemistry.

\textit{13C}-NMR spectroscopy data were obtained by growing a larger number of PDA plates of \textit{Δ\textit{S42/CTB9}} supplemented with 2 mM [1-\textit{13C}]-sodium acetate and 2 mM [2-\textit{13C}]-sodium acetate to evenly enrich all polyketide-derived carbons (~3% per site). Working quickly to isolate and purify precercosporin in low light and low temperature as described earlier, both 1D and heteronuclear single quantum coherence spectra of precercosporin were acquired (\textit{SI Appendix}, Fig. S9 B and C). As seen in the \textit{H}-NMR spectrum, breaking the symmetry of cercosporin was evident in the observation of all 29 carbons in the \textit{13C}-NMR spectrum.
spectra, which notably revealed three methoxyl groups and diagnostic doubling of all resonances, save two overlapping pairs of signals. This behavior is fully in accord with the assigned structure of precercosporin.

Identification of Cercosporin from *C. fioriniae*. Because our initial phylogenomic analyses suggested that several *Colletotrichum* spp. harbored *CTB* clusters (Figs. 2 and 3), we questioned whether the *CTB* cluster can be found in additional *Colletotrichum* spp. *CTB* protein orthology analysis revealed that 8 of the 13 *Colletotrichum* spp. hosted at Ensembl Fungi ([https://fungi.ensembl.org/index.html](https://fungi.ensembl.org/index.html)) encode a similar set of *CTB* proteins as observed in *C. higginsianum* ([SI Appendix](https://www.pnas.org/content/115/24/E5463), Table S8). These eight species are plant pathogens of crops such as apple, safflower, melon, and cucumber; a variety of *Brassica* and cereal crops; as well as various tree species (40–46) ([SI Appendix](https://www.pnas.org/content/115/24/E5463), Figs. S10 and S11 and Table S8). Remarkably, many species have lost several *CTB* genes, such as the endophyte *Colletotrichum tofieldiae*, which has lost the cluster entirely ([SI Appendix](https://www.pnas.org/content/115/24/E5463), Figs. S10 and S11 and Table S8).

Earlier reports suggested the production of a red pigment by some *Colletotrichum* spp. such as the apple pathogen *C. fioriniae* (47, 48); therefore, we questioned whether the red pigment was cercosporin. As a first step, two *C. fioriniae* strains (HC89 and HC91) from our collection that were previously isolated from apple were assayed for cercosporin production using the KOH assay (49). No cercosporinlike pigment was observed in the medium under the same conditions that stimulate cercosporin production in *C. beticola*. Since epigenetic modifiers have been used to induce production of SMs in fungal species (50, 51), we considered whether this strategy could be used to induce cercosporin production in *C. fioriniae*. Medium augmented with the histone deacetylase inhibitor trichostatin A (TSA) (50) induced...
production of a red cercosporin-like compound. To characterize this red metabolite, mycelia from both C. fioriniae strains were extracted with ethyl acetate. Reverse-phase HPLC analysis of extracts of both strains revealed a peak with a retention time and UV-Vis spectrum consistent with cercosporin in both extracts (Fig. 7 A and B). The presence of cercosporin was confirmed by ultra-performance LC (UPLC) electrospray ionization (ESI)-MS (Fig. 7C). Moreover, complementary resequencing of both isolates on the Illumina platform followed by automated genome assembly and gene prediction confirmed the presence of a CTB cluster in both genomes (SI Appendix, Fig. S11 and Table S8).

To assess whether C. fioriniae produces cercosporin during apple infection, apples were inoculated with the pathogen, and tissue samples were collected every other day from 1 to 14 d postinoculation. Regardless of time point or extraction methodology, we were unable to detect cercosporin from infected apple. However, by infiltrating apples with defined amounts of cercosporin, we determined that the isolation efficiency is ∼5% (SI Appendix, Figs. S12 and S13). Consequently, any cercosporin produced by C. fioriniae during infection is likely rapidly bound or modified by apple tissue such that its recovery by extraction is poor and below our level of detection by HPLC or ESI-MS. To provide additional support for the involvement of cercosporin in C. fioriniae infection, we infiltrated cercosporin into apple fruit and visually compared the phenotype of the infiltrated area to C. fioriniae-infected apple. At 5 d postinfiltration/inoculation, cercosporin-infiltrated apple exhibited similar symptomology as C. fioriniae-infected apple (SI Appendix, Fig. S14). We also quantified expression of CTB1 using qRT-PCR since it is highly expressed during the necrotrophic phase of infection (60 hpi) (43). Again, we were unable to detect cercosporin from infected Arabidopsis leaves regardless of the extraction methodology or time point after inoculation. The relatively low amount of fungal biomass associated with the early time points when the CTB cluster is induced during Arabidopsis infection, we hypothesize that the amount of cercosporin produced in planta is below our current limits of detection.

**Discussion**

Several hypotheses exist for the maintenance of SM biosynthetic genes as clusters. In one, unlinked SM pathway genes are at a greater risk for dissociation during meiotic recombination (52) or chromosomal rearrangements (53). Additionally, clustering may facilitate strict coordination of gene expression, which may be particularly important during the biosynthesis of SMs that have potentially toxic or chemically unstable intermediates to ensure their efficient conversion to final end products (54). Horizontal transfer and maintenance of the CTB cluster specifically among plant pathogens suggests that it was critical for disease development in diverse pathosystems, including rice blast caused by M. oryzae and various anthracnose diseases caused by Colletotrichum spp. on many different crops. The CTB clusters in C. higginsianum and C. graminicola were reported as one of the few SM clusters between these species that are microsyntenic (43). Moreover, O’Connell et al. (43) detected specific up-regulation of the CTB cluster in C. higginsianum during colonization of Arabidopsis. Indeed, 9 of 14 C. higginsianum CTB genes were among the top 100 most highly expressed genes in planta. Recent analysis of natural selection processes in C. graminicola identified orthologs of CTB genes CTB1 and CFP among the ∼80 genes undergoing significant positive selection (55), further suggesting a role in pathogenicity. Interestingly, the CTB clusters of Colletotrichum spp. and M. oryzae contain additional genes: two short-chain dehydrogenases, an additional desaturase, a ferric-chelate reductase, and an NmrA-like family protein, which has been reported (56) to act as a negative transcriptional regulator.

The identification of cercosporin production in two isolates of C. fioriniae may have significant implications for the apple packing, storage, and processing industries. Bitter rot, caused by Colletotrichum spp., is one of the top pre- and postharvest pathogens of apple (57). This disease is a major problem for the apple industry because it limits fresh fruit in the field and during storage and has a quiescent stage, allowing decay to occur on seemingly high-quality apples, only to come out of storage rotten (57, 58). Hence, contamination of processed apple products with cercosporin could be a significant health hazard. For example, other fungal-produced toxins (e.g., patulin, citrinin, and penicillenic acid) can contaminate processed apple products (59). Patulin, produced by Penicillium spp., is the most troubling because it is carcinogenic; consequently, the United States and Europe have strict patulin limits in fruit juices and processed pome-fruit products (59, 60). It may be necessary to will focus on the role of cercosporin production during the Colletotrichum-apple fruit interactions in addition to assaying processed fruit products made from apples with bitter rot symptoms to determine levels of the toxin in fruit. Although only C. fioriniae and C. higginsianum strains were analyzed for the ability to produce cercosporin, the identification of highly similar CTB clusters in other Colletotrichum spp. (SI Appendix, Figs. S10 and S11 and Table S8) suggests that cercosporin production may be widespread in this genus. Future studies directed toward analysis of CTB cluster expression among various C. fioriniae isolates and apple
cercosporin will be necessary to confirm whether cercosporin is necessary for virulence of this pathogen. The microsynteny outside the established CTB cluster prompted us to test whether the flanking genes in *C. beticola* are also required for cercosporin biosynthesis. Notably, we observed that these flanking genes, similar to the established CTB genes, were up-regulated under cercosporin-inducing conditions. Furthermore, targeted gene replacement of CTB9, CTB10, CTB11, and CTB12 completely abolished cercosporin biosynthesis, while replacement of CTB9 and CTB10 resulted in the accumulation of a red metabolite, defined here as precercosporin. We thus conclude that the CTB cluster is significantly larger than previously described (19).

The isolation and characterization of an intermediate in the cercosporin biosynthetic pathway (precercosporin) strongly suggests that formation of the unique seven-membered methylenedioxy bridge in the final product is the result of a two-step process requiring three genes. First, one of two precursor aryl methoxyl groups of *ent* (+)-calphostin D (Fig. 6) is removed, followed by oxidative ring closure by CTB9, an apparent α-ketoglutarate-dependent dioxygenase, in collaboration with CTB10. The precise role of CTB10, a putative dehydratase, in ring closure is unclear, but it could serve to facilitate closure of the unfavorable seven-membered methylenedioxy ring. In contrast, a single cytochrome P450 is known to convert two aryl ortho-methoxyl groups into the relatively more common five-membered methylenedioxy group in alkaloid biosynthesis (61). We attribute the single demethylation to an oxidative process possibly carried out by the flavin-dependent enzymes CTB5 or CTB7. CTB6 correlates with the SDR NAD(P)H-binding superfamily of oxidoreductases and could install the side-chain hydroxyl groups stereospecifically. Owing to the extreme instability of most pathway intermediates and the role feedback inhibition in response to these metabolites could play, our experience dictates that analysis of pathway knockouts alone will not lead to the full determination of cercosporin biosynthesis. Biochemical evaluation of the individual enzymes, as has been done with CTB3 (21) with synthetic substrates and product standards, will be necessary to accomplish this task.

A tentative cercosporin biosynthesis scheme was recently proposed (21) without knowledge of the expanded CTB cluster. However, in light of the identification of precercosporin and the presence of the fasciclin/laccase pair (CTB11/12) in the previously proposed biosynthetic pathway (21) will have to be revised. While these investigations will be reported in due course, we suspect the fasciclin/laccase pair (CTB11/12) may act early in the pathway to dimerize the product of CTB3 (21) to the first perylenequinone intermediate, which would have precedent in the α-ketoglutarate-dependent dioxygenase pair (CTB11/12) may act early in the pathway. As noted above.

Despite sustained research on cercosporin for several decades, there are significant knowledge gaps in cercosporin biosynthesis. Our data have shed light on cercosporin biology and will have significant impact on cercosporin research in particular and on perylenequinone research in general. The finding that at least one species in the important plant pathogenic genus *Colletotrichum* can produce cercosporin has significant implications for disease management. Moreover, since *C. fioriniae* may secrete cercosporin into apple food products that may be directly consumed by humans, the toxic effects of cercosporin on human health may need to be considered.

### Materials and Methods

For further information, see SI Appendix, SI Materials and Methods and figshare under DOI: 10.6084/m9.figshare.4056522. Custom code is permanently archived at Zenodo under DOI: 10.5281/zenodo.1156551. Fungal genomic DNA was isolated from mycelia scraped from the surface of agar Petri plates. Library preparations and sequencing on the Illumina platform was performed by BGI Americas Corp. For *C. beticola*, three genomic libraries with an insert size (500 bp, 5 Kbp, and 10 Kbp) were sequenced. For *C. berteroa*, *C. nicotianae*, and *C. fioriniae* strains HB9 and HC91, single, short insert libraries (500 bp) were sequenced. For *C. beticola* specifically, optical maps were prepared using the Augus (OpGen) and BioNano Genomics platforms and subsequently used to scaffold contigs into large supercontigs. A combination of ab initio gene prediction, homologous protein alignment, and transcript alignment followed by extensive manual curation was used to prepare draft gene models for *C. beticola*. A curated Augustus parameters were used for automated protein-coding gene modeling in the case of *C. berteroa*, *C. canescens*, and *C. nicotianae*. Gene-assembly names and annotations, if applicable, are deposited in the NCBI GenBank database and listed under BioProject PRJNA270309. Accession numbers for *C. beticola*, *C. berteroa*, *C. nicotianae*, and *C. fioriniae* strains HB9 and HC91 are KLM000000000, PEN000000000, POS500000000, and PNM300000000, respectively.

Mycological plucks of wild-type and mutant *C. beticola* were placed on top of eight “thin” PDA (Difco) plates (3.0 mL of PDA per 50-mm Petri plate). Cultures were incubated at 22 °C for 1 wk under continuous light. PDA and mycelia were ground under liquid nitrogen and lyophilized to dryness twice. The resulting powder was resuspended in water acidified with HCl (pH < 1), allowed to sit 10 min, and filtered. The filtrate was extracted thrice with ethyl acetate. These extracts were pooled, washed with brine, and evaporated to dryness. The extracted metabolites were resuspended in 500 µL of methanol and analyzed by HPLC on an Agilent 1200 fitted with a Kinetex XB-C18 column (4.6 × 75 mm, 2.6 µm; Phenomenex). Injections of 1 µL were run at 1.25 mL/min with a linear gradient of 5% A/95% B to 95% A/5% B over 10.8 min, where solvent A was acetonitrile + 0.1% formic acid and solvent B was 0.1% formic acid. Chromatograms were monitored at 436, 280, and 210 nm, and UV-Vis spectra were recorded over a range of 210 to 800 nm. Mass spectra data were obtained from a Waters Acquity Xevo-G2 UPLC-ESI-MS in positive-ion mode.

ACKNOWLEDGMENTS. We thank W. Underwood and T. L. Friesen (US Department of Agriculture (USDA) Agricultural Research Service (ARS)) for review of the manuscript, A. G. Newman for helpful discussions and a reference sample of cercosporin, and N. Metz (USDA-ARS) for excellent technical assistance. Prof. M. Muller and L. Fürtges (University of Freiburg) are thanked for discussion of fungal laccases and unpublished results. This project was supported by a Long-Term Fellowship of the European Molecular Biology Organization (ALTF 359-2013) and a postdoctoral fellowship of the Research Foundation Flanders (FWO 12B8116N) (to R.d.J.); the USDA-ARS Current Research Information System Project 3060-22000-049 (to M.D.B.); National Institutes of Health Grants R01 ES01670 (to C.A.T.) and T32 GM080189 (to C.R.H.-R.); and Netherlands Organization for Scientific Research Grant 833.13.007 (to M.K.E.). Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.


