Complete Mitochondrial Genome Sequence of *Colletotrichum siamense* Isolated in South Korea

Sung-Eun Cho,[a] Ji Yeon Oh,[b] Dong-Hyeon Lee[b]

aForest Biodiversity Division, Korea National Arboretum, Pocheon, South Korea
bDivision of Forest Insect Pests and Diseases, National Institute of Forest Science, Seoul, South Korea

Sung-Eun Cho and Ji Yeon Oh contributed equally to this study. Author order was determined by the degree to which the workload was apportioned.

ABSTRACT The complete mitochondrial genome of *Colletotrichum siamense* is characterized. The circular genome has a size of 52,710 bp, with a GC content of 34.45%, and contains 15 protein-coding genes, 23 tRNA genes, and 2 rRNA genes.

Among the top 10 fungal pathogens of scientific and economic importance, the genus *Colletotrichum* Corda 1831 is considered the eighth most important phytopathogenic fungus (1); it causes typical symptoms of anthracnose on fruit, vegetable, and ornamental hosts worldwide, leading to severe losses in the yield and quality of hosts (2).

In 2020, severely infected fruit of pecan, *Carya illinoinensis* (Wangenh.) K. Koch, showing distinct anthracnose symptoms was observed in pecan orchards in Uiseong (36°21′31.5″N, 128°27′15.9″E) and Miryang (35°22′54.9″N, 128°48′06.5″E) in South Korea, and the causal agent of anthracnose on pecan was identified as *Colletotrichum siamense* Prihastuti, L. Cai & K. D. Hyde 2009 (3).

The Korean isolate of *C. siamense*, which was deposited in the Korean Agricultural Culture Collection (KACC), National Institute of Agricultural Sciences, South Korea (http://genebank.rda.go.kr) (accession no. KACC 49782), was obtained from the culture collection (Cultures of Dong-Hyeon) of the National Institute of Forest Science, South Korea (https://nifos.forest.go.kr) (accession no. CDH2020-20) (3). Genomic DNA was extracted from the mycelium using Maxwell 16 DNA purification kits (Promega, USA). An Illumina paired-end (PE) library was constructed and sequenced using the Illumina HiSeq X platform with 151-bp PE reads. Raw sequencing data (2.2 Gb) were trimmed using the quality_trim program in the CLC Assembly Cell package v. 4.2.1 (Qiagen, Denmark) with default parameters (Phred scores of >20) and then were used for *de novo* assembly of the mitochondrial genome as in a previous study (4). In brief, trimmed high-quality read sequences were *de novo* assembled using the clc_novo_assemble program in CLC Assembly Cell with default parameters, and then mitochondrial contigs were selected and ordered by similarity searches with mitochondrial sequences from NCBI Organelle Genome Resources (https://www.ncbi.nlm.nih.gov/genome/organelle). The selected contigs were merged and generated a draft mitochondrial genome sequence. Subsequently, gaps were filled and errors were corrected by PE read mapping and manual curation as in a previous study (5). The final complete mitochondrial genome sequence was annotated using the GeSeq (6) and Artemis (7) programs with mitochondrial reference genomes (GenBank accession no. KX885100, KX885103, KX885101, KX034082, and KX885105). In addition, the precise gene regions were determined by manual curation based on BLAST searches against the mitochondrial reference genomes.

The mitochondrial genome of the Korean isolate of *C. siamense* is a circular molecule of 52,710 bp, with a GC content of 34.45%. Gene prediction and annotation identified 15 protein-coding genes, 23 tRNA genes, and 2 rRNA genes in the genome. Comparative analysis with the mitochondrial genomes of Chinese isolates of *C. siamense* (GenBank accession no.
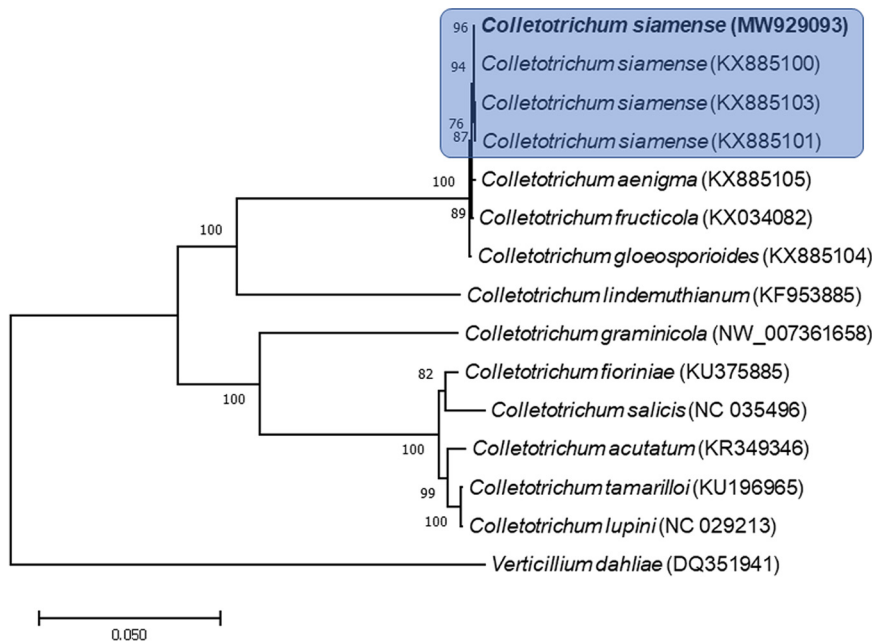
FIG 1 ML phylogenetic tree of the mitochondrial genomes of the *Colletotrichum siamense* Korean isolate and related species. Eleven protein-coding sequences, including *atp8*, *atp9*, *cox2*, *nad5*, *cob*, *atp6*, *nad1*, *cox1*, *rps3*, *nad6*, and *cox3*, that were conserved in the mitochondrial genomes of 15 species underwent multiple alignment using MAFFT (http://mafft.cbrc.jp/alignment/server/index.html) and were used to generate the phylogenetic tree with MEGA v. 7.0 (9). The bootstrap support values (>50%) from 1,000 replicates are indicated on the nodes. The Korean isolate, *Colletotrichum siamense* KACC 49782, obtained and used in the phylogenetic analysis is represented in bold. The GenBank accession no. of mitochondrial genome sequences used for this tree are indicated in parentheses.

KX885098 to KX885103) (8), based on sequence alignment using BLAST and MAFFT (http://mafft.cbrc.jp/alignment/server/index.html), revealed that the genome of the Korean isolate showed 99.2 to 99.9% similarity to the Chinese isolates and had 35 polymorphic sites, including 15 single-nucleotide polymorphisms (SNPs) and 20 indels. These findings indicated that the Korean isolate of *C. siamense* is distinct from the Chinese isolates.

Phylogenetic analysis using the Korean isolate of *C. siamense* with other taxa was performed using a maximum likelihood (ML) method with conserved protein-coding sequences and revealed that the Korean isolate is located in the same clade as other *Colletotrichum* strains and close associates (Fig. 1).

Data availability. The genome sequence data that support the findings of this study are available in NCBI GenBank under the accession no. MW929093. The associated BioProject, BioSample, and SRA no. are PRJNA722052, SAMN18744303, and SRR14241033, respectively.

REFERENCES

1. Dean R, Van Kan JAL, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, Foster GD. 2012. The top 10 fungal pathogens in molecular plant pathology. Mol Plant Pathol 13:414–430. https://doi.org/10.1111/j.1364-3703.2011.00783.x.
2. Hyde KD, Cai L, Cannon PF, Crouch JA, Crous PW, Damm U, Goodwin PH,

Chen H, Johnston PR, Jones EBG, Liu ZY, McKenzie EHC, Moriwaki J, Noireung P, Pennycook SR, Pfenning LH, Prihastuti H, Sato T, Shivas RG, Tan YP, Taylor PWJ, Weir BS, Yang YL, Zhang JZ. 2009. *Colletotrichum*: names in current use. Fungal Divers 39:147–182.
3. Oh JY, Heo JI, Lee DH. 2021. First report of anthracnose on pecan (*Carya*

illinoiensis) caused by *Colletotrichum siamense* in Korea. Plant Dis 105:3296. https://doi.org/10.1094/PDIS-12-20-2709-PDN.

4. Lee HO, Choi JW, Baek JH, Oh JH, Lee SC, Kim CK. 2018. Assembly of the mitochondrial genome in the *Campanulaceae* family using Illumina low-coverage sequencing. Genes (Basel) 9:383. https://doi.org/10.3390/genes9080383.

5. Kim K, Lee SC, Lee J, Yu Y, Yang K, Choi BS, Koh HJ, Waminal NE, Choi HI, Kim NH, Jang W, Park HS, Lee J, Lee HO, Joh HJ, Lee HJ, Park JY, Perumal S, Jayakodi M, Lee YS, Kim B, Copetti D, Kim S, Kim S, Lim KB, Kim YD, Lee J, Cho KS, Park BS, Wing RA, Yang TJ. 2015. Complete chloroplast and ribosomal sequences for 30 accessions elucidate evolution of *Oryza* AA genome species. Sci Rep 5:15655. https://doi.org/10.1038/srep15655.

6. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq: versatile and accurate annotation of organelle genomes. Nucleic Acids Res 45:W6–W11. https://doi.org/10.1093/nar/gkx391.

7. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. 2012. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics 28:464–469. https://doi.org/10.1093/bioinformatics/btr703.

8. Liang X, Tian X, Liu W, Wei T, Wang W, Dong Q, Wang B, Meng Y, Zhang R, Gleason ML, Sun G. 2017. Comparative analysis of the mitochondrial genomes of *Colletotrichum gloeosporioides* sensu lato: insights into the evolution of a fungal species complex interacting with diverse plants. BMC Genomics 18:171. https://doi.org/10.1186/s12864-016-3480-x.

9. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054.