Functional Annotation of circRNAs of Tea Leaves During Infection by the Tea Leaf Spot Pathogen *Didymella segeticola*

Shilong Jiang,^{1,2} Yuanyou Yang,¹ Di Guo,^{1,3} Yu Wang,^{1,2} Honglin Huang,¹ Yuqin Yang,^{1,3} Delu Wang,⁴ Xuanli Jiang,^{1,†} and Zhuo Chen^{1,†}

¹ Key Laboratory of Green Pesticide and Agricultural Bioengineering, Ministry of Education, Guizhou University, Guiyang, Guizhou 550025, China

² Agricultural College, Guizhou University, Guiyang, Guizhou 550025, China

³ College of Tea Science, Guizhou University, Guiyang, Guizhou 550025, China

⁴ College of Forestry, Guizhou University, Guiyang, Guizhou 550025, China

Funding

The project was supported by the National Key Research Development Program of China (2017YFD0200308) and its post-subsidy project (2018-5262), National Natural Science Foundation of China (No. 31860515 and 21977023), Program of Introducing Talents to Chinese Universities (111 Program, D20023), and China Agriculture Research System of MOF and MARA (D09).

Keywords

bioinformatics, circular RNA, fungal pathogens, host-pathogen interactions, tea leaf spot

Circular RNA Announcement

Didymella segeticola (Q. Chen) Q. Chen, Crous & L. Cai (syn. *Phoma segeticola*) is an important phytopathogen that causes leaf spot on tea, *Camellia sinensis* (L.) O. Kuntze, which results in huge tea production losses. The functional annotation and analysis of circular RNAs (circRNAs) of tea leaves during *D. segeticola* infection could reveal the host–pathogen interaction mechanisms. In this study, a model of host tea (*C. sinensis* ‘Fuding-dabaicha’) leaves infected by *D. segeticola* was constructed, and then circRNAs of tea leaves were sequenced on the Illumina HiSeq 4000 platform. In total, 231 and 118 differentially expressed circRNAs from tea leaves were significantly up- and downregulated, respectively, by the *D. segeticola* infection. A Kyoto Encyclopedia of Genes and Genomes analysis indicated that the differentially expressed circRNA-hosting genes were significantly enriched in the plant–pathogen interaction pathway, including 15 *S* genes (gene with significantly differential expression was annotated in special database). As the differentially expressed circRNA-hosting gene in the pathway, the *U-box domain-containing protein 35* gene was enriched in response to stress according to a Gene Ontology (GO) analysis. The expression profiling of circRNAs and the annotation of circRNA-hosting genes of tea leaves during *D. segeticola* infection will provide a resource for future research on host–pathogen interactions.

As a member of the *Didymellaceae* family, *Didymella segeticola* is located in the clade of *Didymella*, and it causes leaf spot on Tibetan thistle (Chen et al. 2015), tea (*C. sinensis*; Ren et al. 2019; Zhao et al. 2018), and tobacco (Guo et al. 2020). Tea leaf spot caused by *D. segeticola* results in huge tea leaf production losses and decreases in tea leaf quality. It often occurs during cold spells in late spring and in high-altitude tea plantations (Ren et al. 2019; Zhao et al. 2018). Because of the complicated pathogenesis and epidemic rule of tea leaf spot, there are currently no effective control measures. Studying RNA–RNA interactions increases our understanding of pathogenesis-related mechanisms, which enables the development of novel control strategies (Cui et al. 2020). In a previous study, messenger RNAs (mRNAs) and microRNAs (miRNAs) of *D. segeticola* and tea leaves were sequenced, and some putative target genes of miRNAs originating from *D. segeticola* and tea host were predicted (Yang et al. 2021). As a novel type of RNA species, circular RNAs (circRNAs) possess

[†]Corresponding authors: X. Jiang; jxl32371@163.com, and Z. Chen; gychenzhuo@aliyun.com

e-Xtra: Supplementary tables are published online.

The author(s) declare no conflict of interest.

Accepted for publication 15 September 2021.

Copyright © 2022 The Author(s). This is an open access article distributed under the CC BY-NC-ND 4.0 International license.

some biological functions involved in host–pathogen interactions, disease resistance, and environmental responses by sponging miRNAs (Panda 2018; Yu et al. 2019). The different responses of rice host to *Magnaporthe oryzae* infection involve the diversity of circRNAs originating from the host, and circR5g05160 can enhance rice immunity against *M. oryzae* (Fan et al. 2020). Nevertheless, there are no reports of circRNAs of tea hosts during infection by *D. segeticola*.

In this study, the third or fourth leaves on the twigs of tea (*C. sinensis* ‘Fuding-dabaicha’) plants were inoculated by puncturing them four times, forming holes approximately 8 mm in diameter, using sterilized syringe needles. A single conidium of *D. segeticola* strain CGMCC3.20152 was incubated on potato dextrose agar (PDA) at 28°C in the dark for 10 days to obtain a vigorously growing colony. A mycelial plug (diameter = 8 mm) was cut from the edge of a colony growing on the PDA medium. The neighboring four sites were inoculated with the mycelial plugs of *D. segeticola* strain CGMCC3.20152. For the uninfected treatment, a sterile PDA plug was inoculated into the neighboring four holes on the tea leaves. At 48 h postinoculation, a lesion formed on each tea leaf, and the area neighboring the lesion was harvested using sterile hole puncher. In total, 15 healthy twigs were pooled to form one treatment group for both the infected and uninfected treatments. The experiment was conducted independently with three biological replicates (total of six samples from two treatment groups).

Total RNA was isolated and purified from tea leaves of both treatment groups using TRIzol reagent (Invitrogen, Carlsbad, CA, U.S.A.) following the manufacturer’s procedure. The RNA amount and purity of each sample was quantified using a NanoDrop ND-1000 (NanoDrop, Wilmington, DE, U.S.A.). The RNA integrity (RIN) was assessed by an Agilent 2100 with RIN number greater than 7.0. Approximately 5 µg of total RNA was used to deplete ribosomal RNA according to the protocol of the Ribo-Zero rRNA Removal Kit (Illumina, San Diego, CA, U.S.A.). After removing ribosomal RNAs, the left RNAs were fragmented into small pieces using divalent cations under high temperature. Then, the cleaved RNA fragments were reverse-transcribed to create the cDNA, which were next used to synthesize U-labeled second-stranded DNAs with *E. coli* DNA polymerase I, RNase H, and dUTP. An A-base was then added to the blunt ends of each strand, preparing them for ligation to the indexed adapters. Each adapter contained a T-base overhang for ligating the adapter to the A-tailed fragmented DNA. Single- or dual-index adapters were ligated to the fragments, and size selection was performed with AMPureXP beads. After the U-labeled second-stranded DNAs were conducted using the heat-labile UDG enzyme, the ligated products were amplified with PCR under the following conditions: initial denaturation at 95°C for 3 min; 8 cycles of denaturation at 98°C for 15 s, annealing at 60°C for 15 s, and extension at 72°C for 30 s; and then final extension at 72°C for 5 min. The average insert size for the final cDNA library was 300 bp (±50 bp). Finally, the paired-end sequencing was performed on an Illumina HiSeq 4000 (LC Bio, China), and 2 × 150-bp paired-end reads (PE150) were produced following the vendor’s recommended protocol.

Cutadapt (<https://github.com/marcelm/cutadapt>) was used to remove the reads that contained adaptor contamination, low-quality bases, and undetermined bases (Martin 2011). Then, sequence quality was verified using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Bowtie2 (version 2.0.0-beta4; <https://github.com/BenLangmead/bowtie2>) and TopHat2 software (<http://ccb.jhu.edu/software/tophat>) were used to map reads to tea genome (<http://tpdb.shengxin.ren/index.html>; Kim et al. 2013; Langmead and Salzberg 2012). The remaining reads (unmapped reads) were mapped to the genome of *D. segeticola* (BioProject PRJNA516041) using TopHat-Fusion software (<http://tophat-fusion.sourceforge.net/>; Kim and Salzberg 2011). CIRCexplorer 2 software (<https://github.com/YangLab/CIRCexplorer2>) was used to de novo assemble the mapped reads of all the samples to circRNA. According to circRNA structural characteristics and splicing sequence characteristics, circRNA was identified using the following criteria: (i) mismatch ≤ 2; (ii) back-spliced (BS) junction reads ≥ 1; and (iii) two splice sites are less than 100 kb in the genome. CircRNAs were divided into the types of circRNA and intronic circRNA (ciRNA) using CIRCexplorer 2 software (Zhang et al. 2014, 2016). Because the reads containing BS sites do not exist in the genome, the numbers of circRNAs were determined mainly using BS sites. BS reads were identified in unmapped reads by TopHat-Fusion and CIRCexplorer. All samples generated unique circRNAs. The differentially expressed circRNAs were selected using the following criteria: log₂ (fold change) > 1 or log₂ (fold change) < −1 and statistical significance (*P* < 0.05) using the R package-empirical analysis of DGE in R (edgeR, version 3.13; Robinson et al. 2010). The differentially expressed (DE) circRNA-hosting genes were

Table 1. Circular RNA (circRNA) sequences from tea (*Camellia sinensis* [L.] O. Kuntze) leaves during infection by the fungal pathogen *Didymella segeticola*

Sequence	Sample source	
	Uninfected treatment ^w	Infected treatment ^x
Raw reads	141,647,262	135,085,399
Valid reads	139,330,285	132,648,511
Valid bases	20.90 G	19.89 G
All CircRNAs	Upregulation	231 ^y (infected treatment vs. uninfected treatment)
	Downregulation	118 ^y (infected treatment vs. uninfected treatment)
Type 1: circRNAs	Upregulation	228 (infected treatment vs. uninfected treatment)
	Downregulation	117 (infected treatment vs. uninfected treatment)
Type 2: ciRNAs ^z	Upregulation	3 (infected treatment vs. uninfected treatment)
	Downregulation	1 (infected treatment vs. uninfected treatment)

^w Tea leaves were punctured four times using sterilized syringe needles, producing clustered holes approximately 8 mm in diameter, and then sterile PDA was used to inoculate on the adaxial side of each leaf for the uninfected treatment.

^x Procedure was the same as in the uninfected treatment, except that a mycelial plug was used to inoculate the tea leaves for the infected treatment.

^y The values indicate the numbers of differentially expressed circRNAs with up- or downregulation between infected versus uninfected treatment.

^z ciRNA = intronic circRNA.

annotated in the databases of Gene Ontology (<http://geneontology.org>) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (<https://www.kegg.jp/kegg>) using BLAST (Conesa et al. 2005).

In this study, infection by *D. segeticola* significantly up- and downregulated 231 and 118 circRNAs in tea leaves, respectively (Table 1). In total, 345 circRNAs and 4 ciRNAs of tea leaves were differentially expressed after infection by *D. segeticola*. The KEGG database annotation indicated that the DE circRNA-hosting genes were enriched in the plant–pathogen interaction pathway and composed of *U-box domain-containing protein 35* (TEA028029.1), *probable inactive receptor-like protein kinase At3g56050* (TEA018206.1), *probable LRR receptor-like serine/threonine-protein kinase At5g10290* (TEA029912.1), *WRKY transcription factor WRKY24* (TEA028505.1), *calmodulin-domain protein kinase 5 isoform 1* (TEA023488.1), *mitochondrial processing peptidase* (TEA024744.1), *probable calcium-binding protein CML21* (TEA023345.1), *suppressor of the G2 allele of skp1 (SGT1)* (TEA017998.1), and so on. Interestingly, U-box domain-containing protein 35, an alternative name for plant U-box protein 35, contains 2 domains of E3 ubiquitin ligase and serine/threonine-protein kinase. U-box domain-containing protein 35 is involved in the protein ubiquitination pathway. In addition, the circRNA-hosting gene *WRKY transcription factor WRKY24* was bound with circRNA1292, ciRNA20, and ciRNA41. The GO database annotation indicated that the DE circRNA-hosting genes were enriched in the “response to stress” of the GO term and composed of *U-box domain-containing protein 35*, *RNA polymerase II C-terminal domain phosphatase-like 1* (TEA018696.1), *seven transmembrane MLO family proteins, putative isoform 1* (TEA012477.1), and so on.

This dataset provides an important resource for studying the mechanisms underlying *D. segeticola* and tea host interactions. The sequences of circRNAs from tea during *D. segeticola* infection have been deposited in GenBank under SRA accession PRJNA689791 and PRJNA755815. In addition, the bioinformatics analysis of circRNAs is presented in the supplementary tables. The sequences and functional analysis of circRNAs originating from tea leaves during *D. segeticola* infection will provide a valuable resource for future research on trait-specific genes of the pathogen and disease resistance in the tea host.

Literature Cited

- Chen, Q., Zhang, K., Zhang, G., and Cai, L. 2015. A polyphasic approach to characterize two novel species of *Phoma* (*Didymellaceae*) from China. *Phytopathology* 197: 267–281.
- Conesa, A., Götts, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676.
- Cui, J., Jiang, N., Hou, X.-X., Wu, S.-H., Zhang, Q., Meng, J., and Luan, Y.-S. 2020. Genome-wide identification of lncRNAs and analysis of ceRNA networks during tomato resistance to *Phytophthora infestans*. *Phytopathology* 110:456–464.
- Fan, J., Quan, W., Li, G.-B., Hu, X.-H., Wang, Q., Wang, H., Li, X.-P., Luo, X., Feng, Q., Hu, Z.-J., Feng, H., Pu, M., Zhao, J.-Q., Huang, Y.-Y., Li, Y., Zhang, Y., and Wang, W.-M. 2020. circRNAs are involved in the rice-*Magnaporthe oryzae* interaction. *Plant Physiol.* 182:272–286.

- Guo, Z.-N., Xie, H.-L., Wang, H.-C., Huang, Y., Chen, Q.-L., Xiang, L.-G., Yu, Z.-H., and Yang, X.-H. 2020. Leaf spot caused by *Didymella segeticola* on tobacco in China. *Plant Dis.* 104:1559.
- Kim, D., and Salzberg, S. L. 2011. TopHat-Fusion: An algorithm for discovery of novel fusion transcripts. *Genome Biol.* 12:R72. doi: 10.1186/gb-2011-12-8-r72.
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. 2013. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14:R36. doi: 10.1186/gb-2013-14-4-r36.
- Langmead, B., and Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9:357-359.
- Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17:10-12.
- Panda, A. C. 2018. Circular RNAs act as miRNA sponges. *Adv. Exp. Med. Biol.* 1087: 67-79.
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. 2010. EdgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139-140.
- Ren, Y.-F., Li, D.-X., Zhao, X.-Z., Wang, Y., Bao, X.-T., Wang, X., Wu, X., Wang, D.-L., Song, B.-A., and Chen, Z. 2019. Whole genome sequences of the tea leaf spot pathogen *Didymella segeticola*. *Phytopathology* 109:1676-1678.
- Yang, R., Jiang, S.-L., Li, D.-X., Yin, Q.-X., Wu, X., Wang, Y., Wang, D.-L., and Chen Z. 2021. Integrated mRNA and small RNA sequencing for analyzing leaf spot pathogen *Didymella segeticola* and its host, tea (*Camellia sinensis*), during infection. *Mol. Plant-Microbe Interact.* 34:127-130.
- Yu, Y., Zhang, Y., Chen, X., and Chen, Y. 2019. Plant noncoding RNAs: Hidden players in development and stress responses. *Annu. Rev. Cell Dev. Biol.* 35:407-431.
- Zhang, X.-O., Dong, R., Zhang, Y., Zhang, J.-L., Luo, Z., Zhang, J., Chen, L.-L., and Yang, L. 2016. Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.* 26:1277-1287.
- Zhang, X.-O., Wang, H.-B., Zhang, Y., Lu, X., Chen, L.-L., and Yang, L. 2014. Complementary sequence-mediated exon circularization. *Cell* 159:134-147.
- Zhao, X.-Z., Wang, Y., Li, D.-X., Ren, Y.-F., and Chen, Z. 2018. Morphological characterization and phylogenetic analysis of the pathogen *Phoma segeticola* var. *camelliae* causing a new tea disease. *Acta Phytopathol. Sin.* 48:556-559.