Whole-Genome Sequence of Diaporthe citri Isolate NFHF-8-4, the Causal Agent of Citrus Melanose

Xiang-Yu Liu, 1,2 Chingchai Chaisiri, 1,2 Yang Lin, 2 Wei-Xiao Yin, 2 and Chao-Xi Luo 1,2,†

- ¹ Key Lab of Horticultural Plant Biology, Ministry of Education, Huazhong Agricultural University, Wuhan 430070, China
- ² College of Plant Science & Technology and Hubei Key Laboratory of Plant Pathology, Huazhong Agricultural University, Wuhan 430070, China

Abstract

Diaporthe species are the causal agents of melanose, stem-end rot, and gummosis diseases of citrus. D. citri is the predominant species on different citrus varieties. These diseases exceedingly reduce quality and marketability of fresh fruits. Melanose on fruits especially causes massive economic losses. The infection mechanisms of D. citri are still unclear and the genome sequence of D. citri has not been released. In order to systemically explore the interaction between citrus and D. citri, we sequenced the whole-genome of D. citri NFHF-8-4, which was isolated from a sample with melanose in Jiangxi Province. The NFHF-8-4 genome sequence will provide valuable information for studying the development process, infection process, and resistance to fungicides mechanisms in D. citri.

Genome Announcement

Diaporthe spp. (anamorph: Phomopsis spp.) are the causal agents of citrus melanose. They affect the young leaves and fruits and cause symptoms of small spots, scab-like lesions, tear-drop or star melanose (Nelson 2008). Diaporthe spp. belong to the kingdom Fungi, division Ascomycota, class Sordariomycetes, order Diaporthales, family Diaporthaceae, genus Diaporthe (Gomes et al. 2013). Conidia are the primary source of inoculum in fields and are dispersed to fruits, leaves, and twigs by rain splash. Fruits and leaves are susceptible to conidia until they mature (Agostini et al. 2003). When Diaporthe spp. attack citrus plants, the defense response is triggered and melanose symptoms are produced to suppress the expansion of pathogens (Arimoto and Homma 1988). In Diaporthe spp., D. citri was the dominant species causing melanose disease on different citrus varieties and tissues (Huang et al. 2013). To date, research on citrus melanose has focused on fungicide control in fields (Gopal and Lakshmi 2014). There are few studies on the infection process and mechanism of pathogenicity in D. citri.

Several Diaporthe species have been described in China (Gao et al. 2017; Long et al. 2019; Zhou and Hou 2020). Until now, four genome sequences of Diaporthe spp. were reported, i.e., D. aspalathi, causing southern stem canker disease of soybean (Li et al. 2016), D. ampelina, an endophytic fungus of Commiphora wightii (Savitha et al. 2016),

The author(s) declare no conflict of interest.

Accepted for publication 24 February 2021.

Copyright © 2021 The Author(s). This is an open access article distributed under the CC BY-NC-ND 4.0 International license.

e**-X**tra*

Funding

This work was supported by the National Key Research and Development Program of China (number 2017YFD020200103) and the Fundamental Research Funds for the Central Universities (number 2662020ZKPY018).

Keywords

citrus disease, Diaporthe citri, genome, melanose

[†]Corresponding author: C.-X. Luo; cxluo@mail.hzau.edu.cn

^{*}The e-Xtra logo stands for "electronic extra" and indicates there are supplementary materials published online.

Table 1. Summarized genome information of the Diaporthe citri isolate NFHF-8-4

Assembly statistics	Scaffolds
Total sequence number	34
Total sequenced length (bp)	63,685,968
Maximum sequence length (bp)	12,370,252
Minimum sequence length (bp)	11,195
N ₅₀ length (bp)	5,472,022
N ₉₀ length (bp)	1,665,382
GC (%)	46.72
Total genes length (bp)	26,007,773
Total number of genes	15,921
Average gene length (bp)	1,633.55
Genes percentage of genome (%)	40.81
Total exons length (bp)	22,061,877
Total number of exons	46,011
Average length of exons (bp)	479.49
Average exons per gene	2.89
Exons percentage of genome (%)	34.62
Total coding sequences (CDSs) length (bp)	22,047,369
Average CDS length (bp)	1,384.71
CDSs percentage of genome (%)	34.6
Average intron length (bp)	131.14

D. helianthin, causing stem canker of sunflower (Baroncelli et al. 2016), and *D. longicolla*, causing Phomopsis seed decay of soybean (Li et al. 2017). The details of *Diaporthe* genomes are shown in Supplementary Table S1.

In August 2016, *D. citri* isolate NFHF-8-4 was isolated from Nanfeng mandarin fruit (*Citrus reticulata* cv. Nanfengmiju) from Jiangxi Province, China (Supplementary Table S2) as described previously (Chaisiri et al. 2020). The isolate was grown on potato dextrose agar (PDA) at 25°C for 5 days in the dark. Agar plugs containing actively growing mycelium were taken from PDA and were transferred to 250-ml flasks containing 30 ml of potato dextrose broth (200 ml of potato juice and 20 g of dextrose per liter of medium). Flasks were shaken at 120 rpm for 4 days at 25°C. DNA was extracted following the method described previously (Hu et al. 2011). Genomic sequencing was performed at Personalbio (Shanghai Personal Biotechnology Co., Ltd., Shanghai, China).

A whole-genome shotgun strategy of genome assembly was adopted. The genome of NFHF-8-4 was sequenced with the Illumina Miseg platform and the Pacific Biosciences (Pac-Bio) sequencing platform, respectively. After correction by AdapterRemoval (v 1.5.4) (Schubert et al. 2016) and Quake (v 0.3) (Kelley et al. 2010), 11,138,480 high-quality reads and 3,028,709,328 clean bases were gained in an Illumina Miseq library. Altogether, the PacBio Sequel platform produced 10,271,744,961bp of high-quality bases and 970,379 sequences. Before assembling, the genome size was estimated as 57 Mb with the Illumina reads in a k-mer analysis. Next, the genome of the D. citri NFHF-8-4 isolate was assembled from scratch with data of PacBio sequencing platform by Falcon. Then, the assembled contigs were polished with data from Illumina Miseq sequencing in pilon v 1.18 (Walker et al. 2014). Finally, the gaps in the genome were filled with GapCloser. The total assembly length was 63.68 Mb (approximately 271× coverage of the genome). The genome consisted of 34 contigs ($N_{50} = 5,472,022$ bp, $N_{90} = 1,665,382$ bp), GC content was 46.72%, and maximum scaffold size was 12,370,252 bp. Statistics of gene assembly results were shown in Table 1. The repeat sequences were identified and masked by RepeatModeler (v1.0.4) and Repeatmasker (v4.0.5) (Tempel 2012). The interspersed repeats accounted for 0.27% of the genome assembly including 0.18% retroelements, 0.08% DNA transposons, and 0.01% unclassified repeats. Also, 0.01% of satellites, 1.4% of simple repeats, and 0.19% of lowcomplexity elements of the genome were identified. The assembled genome was assessed by BUSCO, v 2.0 (Seppey et al. 2019), which estimated the conserved protein-coding sequences to be 97.30% of the entire protein-coding sequence. Genes were predicted by Augustus (v 3.03), glimmerHMM (v 3.0.1), and SNAP (v 2006-07-28). Homologous gene prediction was obtained from protein sequences of related species through exonerate (v 2.2.0). The results above were integrated by EvidenceModeler (v r2012-06-25) (Haas et al. 2008). In total, 15,921 genes were predicted, the average gene length was 1,633.55 bp. Approximately 40.81% of the whole genome was covered in genes. In 26,007,773 bp sequences within genes, 22,047,369 bp were coding sequences. Statistics of gene assembly results are shown in Table 1.

The protein-coding genes were annotated using non-redundant (NR), Gene Ontology (GO), EuKaryotic Orthologous Groups (KOG), and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. Among 15,921 genes, 14,837 genes (87.54%) had matched to genes in the National Center for Biotechnology Information (NCBI) nr database, and 7,880 genes (30.42%) had matched to genes in molecular function, cellular component, and biological process GO terms. Most genes were involved in biological process. Functional categorization and distribution are shown in Supplementary Figure S1. Further analysis of potential genes by KOG (Supplementary Fig. S2) showed 922 (11.7%) genes were about secondary metabolites biosynthesis, transport, and catabolism, while 888 genes were about carbohydrate transport and metabolism. Only a few genes were involved in cell motility and extracellular structures. The genes were further divided to 44 different categories by KEGG pathway and more details are shown in Supplementary Fig. S3. Twelve pathways belong to 'metabolism', which was the most highly enriched group. Organismal systems were also enriched, such as 'aging', 'circulatory system', 'development', 'digestive system', 'endocrine system', 'environmental adaptation', 'excretory system', 'immune system', 'nervous system', and 'sensory system'. In addition, cellular processes, environmental information processing, genetic information processing and human diseases were also observed.

Citrus melanose is one of the most common diseases affecting fruit quality in the world, but research into infection process and pathogenicity of *D. citri* is still limited. Sequencing and analysis of genome *D. citri* NFHF-8-4 may facilitate research on mechanisms of pathogenicity and fungicide resistance.

Data Availability

The details of the NFHF-8-4 isolate genome are available in the NCBI database under accession number JACTAD000000000, BioProject number PRJNA660204.

Literature Cited

- Agostini, J. P., Bushong, P. M., Bhatia, A., and Timmer, L. W. 2003. Influence of environmental factors on severity of citrus scab and melanose. Plant Dis. 87:1102-1106.
- Arimoto, Y., and Homma, Y. 1988. Studies on citrus melanose and citrus stemend rot by *Diaporthe citri* (Faw.) Wolf. Part 9. Effect of light and temperature on the self-defense reaction of citrus plants. Ann. Phytopathol. Soc. Jpn. 54:282-289.
- Baroncelli, R., Scala, F., Vergara, M., Thon, M. R., and Ruocco, M. 2016. Draft whole-genome sequence of the *Diaporthe helianthi* 7/96 strain, causal agent of sunflower stem canker. Genom. Data 10:151-152.
- Chaisiri, C., Liu, X. Y., Lin, Y., Li, J. B., Xiong, B., and Luo, C. X. 2020. Phylogenetic analysis and development of molecular tool for detection of *Diaporthe citri* causing melanose disease of citrus. Plants 9:329.
- Gao, Y., Liu, F., Duan, W., Crous, P. W., and Cai, L. 2017. *Diaporthe* is paraphyletic. IMA Fungus 8:153-187.
- Gomes, R. R., Glienke, C., Videira, S. I. R., Lombard, L., Groenewald, J. Z., and Crous, P. W. 2013. *Diaporthe*: A genus of endophytic, saprobic and plant pathogenic fungi. Persoonia 31:1-41.
- Gopal, K., Mukunda Lakshmi, L., Sarada, G., Nagalakshmi, T., Gouri Sankar, T., Gopi, V., and Ramana, K. T. V. 2014. Citrus melanose (*Diaporthe citri* Wolf): A review. Int. J. Curr. Microbiol. App. Sci. 3:113-124.
- Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., White, O., Buell, C. R., and Wortman, J. R. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. Genome Biol. 9:R7.
- Hu, M. J., Cox, K. D., Schnabel, G., and Luo, C. X. 2011. *Monilinia* species causing brown rot of peach in China. PLoS One 6:e24990.
- Huang, F., Hou, X., Dewdney, M. M., Fu, Y. S., Chen, G. Q., Hyde, K. D., and Li, H. Y. 2013. *Diaporthe* species occurring on citrus in China. Fungal Divers. 61:237-250.

- Kelley, D. R., Schatz, M. C., and Salzberg, S. L. 2010. Quake: Quality-aware detection and correction of sequencing errors. Genome Biol. 11:R116.
- Li, S., Darwish, O., Alkharouf, N. W., Musungu, B., and Matthews, B. F. 2017. Analysis of the genome sequence of *Phomopsis longicolla*: A fungal pathogen causing Phomopsis seed decay in soybean. BMC Genomics 18:688.
- Li, S., Song, Q., Martins, A. M., and Cregan, P. 2016. Draft genome sequence of Diaporthe aspalathi isolate MS-SSC91, a fungus causing stem canker in soybean. Genom. Data 7:262-263.
- Long, H., Zhang, Q., Hao, Y. Y., Shao, X. Q., Wei, X. X., Hyde, K. D., Wang, Y., and Zhao, D. G. 2019. *Diaporthe* species in south-western China. Myco-Keys 57:113-127.
- Nelson, S. 2008. Citrus melanose. Plant disease, PD-59, Cooperative Extension Service, College of Tropical Agriculture and Human Resources, University of Hawaii, Manoa, HI, U.S.A.
- Savitha, J., Bhargavi, S. D., and Praveen, V. K. 2016. Complete genome sequence of the endophytic fungus *Diaporthe (Phomopsis) ampelina*. Genome Announc. 4:e00477-16.
- Schubert, M., Lindgreen, S., and Orlando, L. 2016. AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. BMC Res. Notes 9:88.
- Seppey, M., Manni, M., and Zdobnov, E. M. 2019. BUSCO: Assessing genomeaAssembly and annotation completeness. Pages 227-245 in: Gene Prediction: Methods and Protocols. M. Kollmar. ed. Humana. New York.
- Tempel, S. 2012. Using and Understanding RepeatMasker. Pages 29-51 in: Mobile Genetic Elements. Y. Bigot, ed. Humana Press, Totowa, NJ, U.S.A.
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., and Earl, A. M. 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963.
- Zhou, H., and Hou, C. L. 2020. Three new species of *Diaporthe* from China based on morphological characters and DNA sequence data analyses. Phytotaxa 434:201-202.