Draft Genome Sequence Resource of the Citrus Stem-End Rot Fungal Pathogen *Lasiodiplodia theobromae* CITRA15

Qiaolin Zheng,[1] Egem Ozbudak,[1] Guohong Liu,[1,2] Prashant S. Hosmani,[3] Surya Saha,[3] Mirella Flores-Gonzalez,[3] Lukas A. Mueller,[3] Katia Rodrigues-Stuart,[4] Megan M. Dewdney,[4] Youjian Lin,[1] Jiuxu Zhang,[5] Yisel Carrillo Tarazona,[1] Bo Liu,[2] Ricardo Oliva,[6] Mark A. Ritenour,[5] and Liliana M. Cano[1,†]

[1] University of Florida, Institute of Food and Agricultural Sciences IFAS, Department of Plant Pathology, Indian River Research and Education Center, Fort Pierce, FL 34945, U.S.A.
[2] Agricultural Bio-resources Research Institute, Fujian Academy of Agricultural Sciences, Fujian 350003, China
[3] Boyce Thompson Institute, Cornell University, Ithaca NY 14853, U.S.A.
[4] University of Florida, Institute of Food and Agricultural Sciences IFAS, Department of Plant Pathology, Citrus Research and Education Center, Lake Alfred, FL 33850, U.S.A.
[5] University of Florida, Institute of Food and Agricultural Sciences IFAS, Department of Horticultural Sciences, Indian River Research and Education Center, Fort Pierce, FL 34945, U.S.A.
[6] International Rice Research Center IRRI, Plant Pathology, Los Baños, Laguna 4031, Philippines

Abstract

Lasiodiplodia theobromae is a fungal pathogen associated with perennial tropical fruit plants worldwide. In citrus, *L. theobromae* causes stem-end rot (Diplodia stem-end rot), a damaging postharvest disease that is aggravated when trees are also infected with the citrus greening bacteria 'Candidatus Liberibacter asiaticus'. Due to the latent infection of *L. theobromae* during the preharvest stage, it becomes difficult to control the disease by chemical or physical treatment. In the current study, we sequenced and assembled strain CITRA15, the first genome of *L. theobromae* obtained from diseased *Citrus paradise* 'Flame' grapefruit in Florida, and thereby provided a genomic resource for future research on diagnostics, and postharvest and preharvest disease management of citrus and other fruit crops.

Genome Announcement

The genus *Lasiodiplodia* (family Botryosphaeriaceae) includes species of important plant pathogens known to cause leaf spots, gummosis, stem cankers, branch dieback, decline, and fruit rot on a wide range of plants in tropical and subtropical regions (Alves et al. 2008; Bautista-Cruz et al. 2019; Begoude et al. 2010). *Lasiodiplodia* spp. are associated with various diseases in many fruit crops such as cacao (Ali et al. 2020), mango (Shahbaz et al. 2009), papaya (Netto et al. 2014), and citrus (Brown 1986), where they cause vascular streak dieback, decline, stem-end rot and canker, respectively. *Lasiodiplodia theobromae* (Pat.) Griffon & Maubl. (syn. *Diplodia natalensis*) has a broad host range and geographic distribution with an extensive list of synonyms, which include *Botryodiplodia theobromae* Pat. and *Botryosphaeria rhodina* (Berk. & M. A. Curtis) Arx, among others (United States Department of Agriculture database: https://nt.ars-grin.gov/fungaldatabases/fungushost/fungushost.cfm). *D. natalensis*, a fungal name often associated with stem-end rot of citrus, is also considered a synonym, despite the phylogenetic and morphological distinctions between the two genera (Phillips et al. 2013). Nonetheless, the name Diplodia stem-end rot (SER) is retained to differentiate this disease from Phomopsis stem-end rot (*Diaporthe citri*), which also affects *Citrus* spp. Because *L. theobromae* comprises a species complex, a combination of morphology and DNA sequence studies must

Table 1. Summary statistics of the *Lasiodiplodia theobromae* strain CITRA15

Variables	Statistics
Assembly length (Mb)	43.3
Number of contigs	475
N_{50} (bp)	178,039
N_{90} (bp)	56,836
L_{50}	77
Mean contig length (bp)	91,226
Minimum contig length (bp)	634
Maximum contig length (bp)	572,189
BUSCO completeness (%)[a]	95
GC content (%)	54.74
Number of predicted genes	10,496
Number of predicted secreted proteins	1,203

[a] Using the euk odb09 database.

be conducted to discriminate among cryptic species (Alves et al. 2008). Citrus SER caused by *L. theobromae* is one of the most common postharvest diseases in Florida, and mainly causes postharvest decays of citrus fruit. *L. theobromae* has been reported as both an endophyte (Galsurker et al. 2018) and an opportunistic fungal pathogen that invades citrus fruit from previous latent infections, becomes established in necrotic tissues of the fruit button (calyx and disk), and remains inactive until postharvest (Pathak 1997). The Diplodia SER fungus grows into the fruit, usually through cut stems but also through the natural openings formed between the fruit and button after harvest (Brown and Wilson 1967). This fungus is rarely observed in the hanging fruit but it occasionally occurs in over-mature, injured fruit on trees. Ethylene degreening treatment significantly increases the incidence of SER in citrus fruit because ethylene stimulates rapid abscission (Brown 1986; Zhao et al. 2014). *L. theobromae* seems to be most common in the presence of a high amount of dead wood on trees and high rainfall and temperatures. If fungicides are not used continually when the fruit is in storage, untreated citrus can quickly rot. After fungal infection, citrus fruit displays symptoms where the button and fruit attach. A ring softens around the button, and a brown discoloration occurs at the affected areas. Rapid decay develops at the stem end, followed by stylar end rot (Brown 1986).

Recently, Diplodia SER has not only caused the greatest postharvest losses of citrus in Florida, but also exacerbates huanglongbing (HLB)-associated preharvest fruit drop by colonizing under the calyx abscission zone of citrus fruit. *L. theobromae* was consistently present in juice from HLB-symptomatic fruit but not in juice from healthy fruit (Zhao et al. 2014). Furthermore, HLB may increase the amount of weakened or dead wood on the trees, which could promote growth and sporulation of *L. theobromae* and subsequent postharvest SER (Ritenour et al. 2016).

The genomic information of several Botryosphaeriaceae species has been reported, such as three strains isolated from grapevine and including *L. theobromae* LA-SOL3 (Félix et al. 2019), *L. theobromae* AM2As isolated from cacao (Ali et al. 2020), and the related species *D. seriata* F98.1 (Robert-Siegwald et al. 2017), *D. seriata* DS831 (Morales-Cruz et al. 2015), and *D. corticola* CBS112549 isolated from cork oak (Fernandes et al. 2014). There is no genomic resource available for *L. theobromae* strains from citrus. We sequenced and assembled the genome for *L. theobromae* isolate CITRA15 in this study. This important genomic resource will be valuable for researchers working on SER disease and developing effective methodologies to manage *L. theobromae* in citrus and, potentially, in other hosts.

L. theobromae isolate CITRA15 was isolated in 2015 from fruit lesions of grapefruit grown in Citra, Marion County, FL. Mycelia were grown in potato dextrose agar at 25°C, harvested after 10 days, ground to a fine powder in liquid nitrogen, and stored at −80°C prior to DNA and RNA extractions. High molecular weight genomic DNA was extracted from frozen mycelial powder by using the Omniprep for Fungi kit (G-Biosciences, St. Louis, MO, U.S.A.) according to the manufacturer's instructions. Genomic DNA was submitted for PacBio sequencing to the Mount Sinai Institute and two SMRT Pac Bio cells (Pacific Biosciences, Menlo Park, CA, U.S.A.) were used to generate sequencing data. Sequencing reads were assembled with the SMRT-Analysis Tool Hierarchical Genome Assembly Process HGAP3 (Chin et al. 2013) by setting the genome
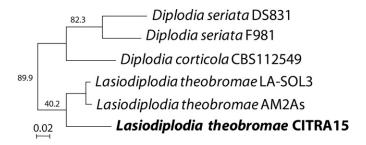
Fig. 1. Phylogenetic relationship of six Botryosphaeriaceae plant-pathogenic fungi. The species tree was inferred as part of the OrthoFinder pipeline (Emms and Kelly 2019) using the species tree inference from all genes (STAG) algorithm (Emms and Kelly 2018) and rooted by species tree root inference from gene duplication events (STRIDE) (Emms and Kelly 2017). STRIDE probability values are shown at internal nodes. Scale bar indicates the number of substitutions per site.

size at 40 Mb and other parameters as default. Total RNA was extracted using the RNeasy minikit (Qiagen Inc.) according to the manufacturer's instructions. RNA was submitted to the University of Florida Interdisciplinary Center for Biotechnology Research ICBR for Nextseq500 RNA-seq Illumina sequencing, and 150-bp paired-end reads were obtained to use as transcript evidence.

The de novo genome draft assembly of *L. theobromae* strain CITRA15 is shown in Table 1. The estimated genome size was 43.3 Mb (43,332,739 bp) with GC content of 54.74%, which was very close to *L. theobromae* AM2As (43.7 Mb) isolated from cacao (Ali et al. 2020) and *L. theobromae* LA-SOL3 (43.9 Mb) isolated from grapevine (Félix et al. 2019). QUAST version 5.0.2 (Gurevich et al. 2013) analysis showed that the *L. theobromae* CITRA15 assembled genome consisted of 475 contigs, with an average size of 91,226 bp, a minimum size of 634 bp, and a maximum size of 572,189 bp. The N_{50} and N_{90} values were 178,039 and 56,836 bp, respectively, and the L_{50} number was 77. The completeness of the assembly was assessed using BUSCO version 3.0.2 (Simão et al. 2015), which estimated the genome sequence to be 95% complete. Gene predictions were obtained using the MAKER2 pipeline (Holt and Yandell 2011), which resulted in a total of 10,496 coding sequences. The proteome was functionally annotated using a blastp search against the nonredundant database (Camacho et al. 2009). Predicted protein analysis with SignalP version 2.0 (Nielsen and Krogh 1998) revealed 1,203 candidates containing a secretion signal peptide. Comparative analysis using OrthoFinder version 2.4 (Emms and Kelly 2019) with five publicly available genomes of Botryosphaeriaceae—that is, *L. theobromae* LA-SOL3 (Félix et al. 2019), *L. theobromae* AM2As (Ali et al. 2020), *D. seriata* F98.1 (Robert-Siegwald et al. 2017), *D. seriata* DS831(Morales-Cruz et al. 2015), and *D. corticola* CBS112549 (Fernandes et al. 2014)—revealed that our sequenced *L. theobromae* isolate CITRA15 shares 96.7% (10,145 genes in orthogroups) of the proteome. We used OrthoFinder to generate a species tree inferred from all orthogroups with the species tree from all genes (STAG) algorithm (Emms and Kelly 2018) and rooted by species tree root inferred from gene duplications (STRIDE) (Emms and Kelly 2017). The species tree showed that *L. theobromae* isolate CITRA15 clustered with reported *L. theobromae* isolates LA-SOL3 and AM2As (Fig. 1). The *L. theobromae* isolate CITRA15 sequence will provide an important resource for researchers working on citrus SER disease management and molecular diagnostics.

Acknowledgments

Literature Cited

Ali, S. S., Asman, A., Shao, J., Balidion, J. F., Strem, M. D., Puig, A. S., Meinhardt, L. W., and Bailey, B. A. 2020. Genome and transcriptome analysis of the latent pathogen *Lasiodiplodia theobromae*, an emerging threat to the cacao industry. Genome 63:37-52.

Alves, A., Crous, P. W., Correia, A., and Phillips, A. J. L. 2008. Morphological and molecular data reveal cryptic speciation in *Lasiodiplodia theobromae*. Fungal Divers. 28:13.

Bautista-Cruz, M. A., Almaguer-Vargas, G., Leyva-Mir, S. G., Colinas-Leon, M. T., Correia, K. C., Camacho-Tapia, M., Robles-Yerena, L., Michereff, S. J., and Tovar-Pedraza, J. M. 2019. Phylogeny, distribution, and pathogenicity of *Lasiodiplodia* species associated with cankers and dieback symptoms of Persian lime in Mexico. Plant Dis. 103:1156-1165.

Begoude, B. A. D., Slippers, B., Wingfield, M., and Roux, J. 2010. Botryosphaeriaceae associated with *Terminalia catappa* in Cameroo, South Africa and Madagascar. Mycol. Prog. 9:101-123.

Brown, G. E. 1986. Diplodia stem-end rot, a decay of citrus increased by ethylene degreening treatment and its control. Proc. Fla. State Hortic. Soc. 99:4.

Brown, G. E., and Wilson, W. 1967. Stem-end rot fungi in oranges: Entry and possible use of pre-harvest fungicides for control. Proc. Fla. State Hortic. Soc. 80:5.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. 2009. BLAST+: Architecture and applications. BMC Bioinf. 10:421.

Chin, C. S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E. E., Turner, S. W., and Korlach, J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat. Methods 10:563-569.

Emms, D. M., and Kelly, S. 2017. STRIDE: Species tree root inference from gene duplication events. Mol. Biol. Evol. 34:3267-3278.

Emms, D. M., and Kelly, S. 2018. STAG: Species Tree Inference from All Genes. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, U.S.A.

Emms, D. M., and Kelly, S. 2019. OrthoFinder: Phylogenetic orthology inference for comparative genomics. Genome Biol. 20:238.

Félix, C., Meneses, R., Gonçalves, M. F. M., Tilleman, L., Duarte, A. S., Jorrin-Novo, J. V., Van de Peer, Y., Deforce, D., Van Nieuwerburgh, F., Esteves, A. C., and Alves, A. 2019. A multi-omics analysis of the grapevine pathogen *Lasiodiplodia theobromae* reveals that temperature affects the expression of virulence- and pathogenicity-related genes. Sci. Rep. 9:13144.

Fernandes, I., Alves, A., Correia, A., Devreese, B., and Esteves, A. C. 2014. Secretome analysis identifies potential virulence factors of *Diplodia corticola,* a fungal pathogen involved in cork oak (*Quercus suber*) decline. Fungal Biol. 118:516-523.

Galsurker, O., Diskin, S., Maurer, D., Feygenberg, O., and Alkan, N. 2018. Fruit stem-end rot. Horticulturae 4:50.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072-1075.

Holt, C., and Yandell, M. 2011. MAKER2: An annotation pipeline and genome-database management tool for second generation genome projects. BMC Bioinf. 12:491.

Morales-Cruz, A., Amrine, K. C., Blanco-Ulate, B., Lawrence, D. P., Travadon, R., Rolshausen, P. E., Baumgartner, K., and Cantu, D. 2015. Distinctive expansion of gene families associated with plant cell wall degradation, secondary metabolism, and nutrient uptake in the genomes of grapevine trunk pathogens. BMC Genomics 16:469.

Netto, M. S. B., Assunção, I. P., Lima, G. S. A., Marques, M. W., Lima, W. G., Monteiro, J. H. A., Balbino, V. Q., Michereff, S. J., Phillips, A. J. L., and Câmara, M. P. S. 2014. Species of *Lasiodiplodia* associated with papaya stem-end rot in Brazil. Fungal Divers. 67:127-141.

Nielsen, H., and Krogh, A. 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6:122-130.

Pathak, V. N. 1997. Post-harvest fruit pathology-present status and future possibilities. Indian Phytopathol. 50:25.

Phillips, A. J., Alves, A., Abdollahzadeh, J., Slippers, B., Wingfield, M. J., Groenewald, J. Z., and Crous, P. W. 2013. The Botryosphaeriaceae: Genera and species known from culture. Stud. Mycol. 76:51-167.

Ritenour, M. A., Zhang, J., and Dewdney, M. M. 2016. Control of stem-end rot of fresh citrus. Citrus Ind. 2016:12-15.

Robert-Siegwald, G., Vallet, J., Abou-Mansour, E., Xu, J., Rey, P., Bertsch, C., Rego, C., Larignon, P., Fontaine, F., and Lebrun, M. H. 2017. Draft genome sequence of *Diplodia seriata* F98.1, a fungal species involved in grapevine trunk diseases. Genome Announce. 5:e00061-17.

Shahbaz, M., Iqbal, Z., Saleem, A., and Anjum, M. A. 2009. Association of *Lasiodiplodia theobromae* with different decline disorders in mango (*Mangifera indica* L.). Pak. J. Bot. 41:10.

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 31:3210-3212.

Zhao, W., Bai, J., McCollum, G., and Baldwin, E. 2014. High incidence of preharvest colonization of huanglongbing-symptomatic citrus sinensis fruit by *Lasiodiplodia theobromae* (*Diplodia natalensis*) and exacerbation of postharvest fruit decay by that fungus. Appl. Environ. Microbiol. 81:364-372.