Draft Genome Sequence of *Alternaria longipes* Causing Tobacco Brown Spot

Zhonghong Feng,¹ Yaxin Li,¹ Xueyan Ma,¹ Yangbo Duan,¹ Rong Zhang,^{1,†} Tom Hsiang,² Yongchun Niu,³ and Guangyu Sun^{1,†}

¹ State Key Laboratory of Crop Stress Biology in Arid Areas and College of Plant Protection, Northwest A&F University, Yangling, Shaanxi Province 712100, China

² School of Environmental Sciences, University of Guelph, Guelph, Ontario N1G 2W1, Canada

³ Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Funding

This research was supported by China Agriculture Research System of MOF and MARA (CARS-27).

Keywords

Alternata section, genomics, genome comparison, pathogen diversity

Abstract

Alternaria is a cosmopolitan fungal genus associated with diverse hosts. Tobacco brown spot caused by *Alternaria longipes* is one of the most destructive diseases of tobacco. *A. longipes* can also infect many other plants, some animals and even humans. Here, we report a genome assembly of *A. longipes* CBS 540.94 using Oxford Nanopore Technologies. A total of 15 contigs were assembled, and the genome size was 37.5 Mb with contig N50 of 4.33 Mb. This genome resource will provide information for further research on comparative genomics of the genus *Alternaria* and be a valuable resource in investigations of the molecular interactions of pathogen and hosts.

Alternaria longipes has been reported as a plant-pathogenic fungus as well as having saprophytic and endophytic activity. *A. longipes* has a broad range of hosts and induces many severe diseases, such as leaf blight of carrots (*Daucus carota* L.), oil palm (*Elaeis guineensis* Jacq.) and alfalfa (*Medicago sativa* L.) (Maiti et al. 2007; Suwannarach et al. 2015; Vintal et al. 2002) and leaf spot of China root (*Smilax china* L.) and sunflower (*Helianthus annuus* L.) (Addrah et al. 2019; Long et al. 2009). It was isolated from healthy dragon tree (*Paulownia fortunei* (Seem.) Hemsl.) (Luo et al. 2010), and Indian mangrove (*Avicennia officinalis* L.) as an endophyte (Ranganathan and Mahalingam 2019). *A. longipes* has also been associated with cutaneous phaeohyphomycosis in humans (Gené et al. 1995). Tobacco brown spot is one of the dominant leaf diseases of tobacco (*Nicotiana* spp.) in the major tobacco producing areas across the world. Losses from this disease can be up to 90% (LaMondia 2001).

The type strain CBS 540.94 of *A. longipes* (Ellis & Everh.) E.W. Mason was originally isolated from diseased tobacco (*Nicotiana tabacum* L.) leaves in 1892 (Tisdale and Wadkins 1931) and deposited at CBS-KNAW, Utrecht, Netherlands. Its pathogenicity was reconfirmed as causing brown spot of tobacco in this research. The genomic DNA was isolated by a modified CTAB procedure (Chen et al. 1993) and sent for sequencing on an Oxford Nanopore Technologies sequencer GridION X5 (Nextomics Biosciences, Wuhan, China). The raw reads were checked by LongQC before assembly (Fukasawa et al. 2020). NextDenovo software (<https://github.com/Nextomics/NextDenovo>) (Tang et al. 2020) with default parameters was used to assemble the reads and to correct the original sequencing errors of long reads. NextPolish software (<https://github.com/Nextomics/NextPolish>) (Hu et al. 2020) was used to correct the base errors (SNP/InDel) from assembly based on the Illumina data. BUSCO (benchmarking universal single-copy ortholog) v1.2 with the universal fungal library

[†]Corresponding authors: G. Y. Sun; sgy@nwsuaf.edu.cn, and R. Zhang; rongzh@nwsuaf.edu.cn

The author(s) declare no conflict of interest.

Accepted for publication 29 September 2021.

Table 1. Genome features of *Alternaria longipes* CBS 540.94 in comparison with *A. longipes* cx1 and BMP0313

	<i>A. longipes</i> CBS 540.94	<i>A. longipes</i> cx1 ^a	<i>A. longipes</i> BMP0313 ^b
Sequencing method	Nanopore	Illumina	Illumina
Assembly size (Mb)	37.50	35.71	36.30
Contigs	15	2,836	3,412
Contig N50 (Mb)	4.33	0.32	0.13
GC content (%)	50.97	51.03	51.00 ^c
Predicted genes	15,842	12,690	13,219
Repetitive sequences (%)	4.48	3.07	1.37 ^c

^a Hou et al. (2016).^b Dang et al. (2015). The genome resources of *A. longipes* BMP0313 can be downloaded from the JGI Genome Portal.^c These data were calculated from the published assembly.

fungi_odb9 was used to assess the completeness of the genome assembly (Simão et al. 2015). For gene predictions, Augustus v3.1 (Stanke and Morgenstern 2005) was used to train the gene models from 24 species selected within the section *Alternaria* (Dang et al. 2015), and *A. taquetica* was obtained as the optimal gene prediction species model. Then GeneMark-ES (Ter-Hovhannisyan et al. 2008) was self-trained for structure prediction. Finally, MAKER2 v2.31.8 (Holt and Yandell 2011) integrated the results trained by Augustus and GeneMark-ES based on protein sequences of the 24 species of section *Alternaria* (Dang et al. 2015). Repetitive sequences were predicted using RepeatMasker v4.0.5 (<http://www.repeatmasker.org>) with a combined database (Dfam_Consensus and RepBase, 20170127).

Around 24.10 Gb clean reads (642× coverage) were obtained. Finally, the genome of *A. longipes* was assembled into 15 contigs (ctg1–15) with a total genome size of 37.50 Mb (Table 1). The strains *A. longipes* cx1 and *A. longipes* BMP0313 were previously sequenced on Illumina HiSeq platform (Dang et al. 2015; Hou et al. 2016), and so we performed comparisons of our assembly with the previous assemblies. The number of contigs in our study was much lower than previously reported assemblies, and the N50 value higher than *A. longipes* cx1 (2,836 contigs, N50 = 0.32 Mb) (Hou et al. 2016) and *A. longipes* BMP0313 (3,412 contigs, N50 = 0.13 Mb) (Dang et al. 2015). The N50 value (4.33 Mb) also was much larger than assemblies of other *Alternaria* species reported (2.10 to 3.10 Mb; Rajarammohan et al. 2019). The telomeric repetitive sequence (TTAGGG/CCCTAA) was discovered in six contigs on both ends (ctg1–6), and six contigs on single ends (ctg7–12). For one contig, the length was less than 2 Mb (0.70 Mb) with two end telomeric repetitive sequences; this was speculated to be a conditionally dispensable chromosome (CDC).

The *A. longipes* genome assembly was approximately 97% complete (1,409 complete BUSCOs, 27 fragmented BUSCOs, and two missing BUSCOs out of 1,438 total BUSCO sequences searched). In this genome assembly, 4.48% repetitive sequences were identified, which was lower than that of *A. brassicae* J3 (9.33%), *A. brassicicola* abra43 (9.30%), and *A. solani* altNL03003 (5.71%), but higher than *A. alternata* ATCC34957 (2.71%), *A. alternata* PN1 (2.43%), or *A. alternata* PN2 (2.64%) (Rajarammohan et al. 2019). The predicted number of protein-coding genes was 15,842. This number of genes was substantially more than was previously reported for *A. longipes* cx1 (12,690) (Hou et al. 2016) or *A. longipes* BMP0313 (13,219) (Dang et al. 2015). This result may be because the genome assembly based on third generation sequencing was improved over previous assemblies, which could have allowed for prediction of a greater number of genes. There also exists the possibility that our isolate sequenced (CBS 540.94) has more genes than those previously sequenced (cx1 or BMP0313). This draft genome information will be valuable for analyzing evolutionary relationships and conducting comparative genomics among *Alternaria* spp., and for revealing the molecular mechanisms of pathogenicity in *Alternaria* species and host-pathogen interactions.

This genome assembly has been deposited at GenBank under the accession number JAHLEY000000000 (BioProject: PRJNA736258; BioSample: SAMN19641085).

Literature Cited

- Addrah, M. E., Zhang, Y., Zhang, J., Liu, L., Zhou, H., Chen, W., and Zhao, J. 2019. Fungicide treatments to control seed-borne fungi of sunflower seeds. *Pathogens* 9:29.
- Chen, X., Line, R. F., and Leung, H. 1993. Relationship between virulence variation and DNA polymorphism in *Puccinia striiformis*. *Phytopathology* 83:1489-1497.
- Dang, H., Pryor, B., Peever, T., and Lawrence, C. B. 2015. The *Alternaria* genomes database: a comprehensive resource for a fungal genus comprised of saprophytes, plant pathogens, and allergenic species. *BMC Genomics* 16:239.
- Fukasawa, Y., Ermini, L., Wang, H., Carty, K., and Cheung, M. S. 2020. LongQC: A quality control tool for third generation sequencing long read data. *G3 (Bethesda)* 10:1193-1196.
- Gené, J., Azón-Masoliver, A., Guarro, J., Ballester, F., Pujol, I., Llovera, M., and Ferrer, C. 1995. Cutaneous phaeohyphomycosis caused by *Alternaria longipes* in an immunosuppressed patient. *J. Clin. Microbiol.* 33:2774-2776.
- Holt, C., and Yandell, M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12:491.
- Hou, Y., Ma, X., Wan, W., Long, N., Zhang, J., Tan, Y., Duan, S., Zeng, Y., and Dong, Y. 2016. Comparative genomics of pathogens causing brown spot disease of tobacco: *Alternaria longipes* and *Alternaria alternata*. *PLoS One* 11:e0155258.
- Hu, J., Fan, J., Sun, Z., and Liu, S. 2020. NextPolish: a fast and efficient genome polishing tool for long-read assembly. *Bioinformatics* 36:2253-2255.
- LaMondia, J. A. 2001. Outbreak of brown spot of tobacco caused by *Alternaria alternata* in Connecticut and Massachusetts. *Plant Dis.* 85:230.
- Long, Y., Hsiang, T., and Huang, J. 2009. First report of leaf spot of *Smilax china* caused by *Alternaria longipes* in China. *Plant Pathol.* 58:800.
- Luo, J., Nazierbieke, W., Li, K., Yao, F., and Borathybay, E. 2010. Isolation and antimicrobial activity of endophytic fungi from *Paulownia fortune*. *Microbiol. China* 37:1481-1485.
- Maiti, C. K., Sen, S., and Acharya, A. 2007. First report of leaf blight of *Medicago sativa* caused by *Alternaria longipes* in India. *J. Plant Pathol.* 89:S69.
- Rajarammohan, S., Paritosh, K., Pental, D., and Kaur, J. 2019. Comparative genomics of *Alternaria* species provides insights into the pathogenic lifestyle of *Alternaria brassicae* - a pathogen of the Brassicaceae family. *BMC Genomics* 20:1036.
- Ranganathan, N., and Mahalingam, G. 2019. Secondary metabolite as therapeutic agent from endophytic fungi *Alternaria longipes* strain VITN14G of mangrove plant *Avicennia officinalis*. *J. Cell. Biochem.* 120:4021-4031.
- 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.
- Stanke, M., and Morgenstern, B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 33(Suppl. 2):W465-W467.
- Suwanarach, N., Kumla, J., and Lumyong, S. 2015. First report of *Alternaria* leaf blight disease on oil palm caused by *Alternaria longipes* in Thailand. *Phytoparasitica* 43:57-59.
- Tang, B., Wang, Z., Liu, Q., Zhang, H., Jiang, S., Li, X., Wang, Z., Sun, Y., Sha, Z., Jiang, H., Wu, X., Ren, Y., Li, H., Xuan, F., Ge, B., Jiang, W., She, S., Sun, H., Qiu, Q., Wang, W., Wang, Q., Qiu, G., Zhang, D., and Li, Y. 2020. High-quality genome assembly of *Eriocheir japonica sinensis* reveals its unique genome evolution. *Front. Genet.* 10:1340.
- Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y. O., and Borodovsky, M. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res.* 18:1979-1990.
- Tisdale, W. B., and Wadkins, R. F. 1931. Brown spot of tobacco caused by *Alternaria longipes* (E. & E.), n. comb. *Phytopathology* 21:641-660.
- Vital, H., Shtienberg, D., Shlevin, E., and Ben-Noon, E. 2002. First report of leaf blight caused by *Alternaria longipes* on carrots in Israel. *Plant Dis.* 86:186.