

Telomere-to-Telomere Genome Assembly of *Bursaphelenchus okinawaensis* Strain SH1

Simo Sun,^a Ryoji Shinya,^b Mehmet Dayi,^{a,c} Akemi Yoshida,^d Paul W. Sternberg,^e  Taisei Kikuchi^a

^aDepartment of Infectious Diseases, Faculty of Medicine, University of Miyazaki, Miyazaki, Japan

^bSchool of Agriculture, Meiji University, Kawasaki, Japan

^cForestry Vocational School, Duzce University, Duzce, Turkey

^dLaboratory of Genomics, Frontier Science Research Center, University of Miyazaki, Miyazaki, Japan

^eDivision of Biology and Biological Engineering, Caltech, Pasadena, California, USA

ABSTRACT *Bursaphelenchus okinawaensis* is a self-fertilizing, hermaphroditic, fungus-feeding nematode used as a laboratory model for the genus *Bursaphelenchus*, which includes the important pathogen *Bursaphelenchus xylophilus*. Here, we report the nearly complete genome sequence of *B. okinawaensis*. The 70-Mbp assembly contained six scaffolds (>11 Mbp each) with telomere repeats on their ends, indicating complete chromosomes.

B *ursaphelenchus okinawaensis* is a fungus-feeding nematode associated with long-horn beetles (*Monochamus maruokai*) and the beetles' host trees (1). Because these nematodes self-fertilize, *B. okinawaensis* has recently emerged as a laboratory model for the genus *Bursaphelenchus* (2), which includes the important plant pathogen *Bursaphelenchus xylophilus* (3). Here, we generated a nearly complete genome sequence of *B. okinawaensis*.

Botrytis cinerea grown on autoclaved barley grains was fed to *Bursaphelenchus okinawaensis* strain SH1 (maintained at Meiji University) for 13 days. Mixed-stage worms were collected using a modified Baermann funnel technique (4). Briefly, worm culture was suspended in distilled water (dH₂O) complemented with streptomycin, amphotericin B, and penicillin (antibiotic/antimycotic [anti/anti]; Gibco), and live worms were passed through a sieve lined with Kimwipes (Crecia) followed by discontinuous sucrose gradient centrifugation to remove culture debris (5). Cleaned worms were incubated in a worm lysis solution (buffer G2 with 0.8 mg/ml proteinase K [Qiagen], 50 mM dithiothreitol [Wako], and 0.5 mg/ml RNase A [Invitrogen]; 55°C; 4 h) following two freeze-thaw treatments. High-molecular-weight genomic DNA was extracted via phenol-chloroform extraction and ethanol precipitation. A Nanopore library was prepared from 1 μg of genomic DNA using a ligation sequencing kit (SQK-LSK109; Oxford Nanopore Technologies) according to the manufacturer's protocol. A single 24-h sequencing run was performed with an R9.4.1 MinION flow cell; thereafter, 2.8 Gbp of sequence data (173,396 reads; N_{50} 33.6 kbp) was obtained. The Nanopore reads were base called to generate FASTQ files using the Guppy v4.0.15 basecaller (Oxford Nanopore Technologies) with the supplied dna_r9.4.1_450bps_hac configuration and were quality checked using NanoPlot v1.31.0 (6). An Illumina sequencing library was separately prepared from 1 μg of extracted DNA using a TruSeq DNA sample preparation kit (Illumina) according to the manufacturer's protocol. The library was sequenced on an Illumina HiSeq 2000 instrument according to the manufacturer's protocol, producing 101-bp paired-end reads (4.2 Gbp). Raw Illumina sequence data were subjected to the Real-Time Analysis (RTA) v1.12.4.2 analysis pipeline (Illumina). A Hi-C library was prepared from ~10,000 fresh worms using an Arima-HiC kit (Arima Genomics) followed by a Colibri ES

Citation Sun S, Shinya R, Dayi M, Yoshida A, Sternberg PW, Kikuchi T. 2020. Telomere-to-telomere genome assembly of *Bursaphelenchus okinawaensis* strain SH1. Microbiol Resour Announc 9:e01000-20. <https://doi.org/10.1128/MRA.01000-20>.

Editor Antonis Rokas, Vanderbilt University

Copyright © 2020 Sun et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Taisei Kikuchi, tasei_kikuchi@med.miyazaki-u.ac.jp.

Received 7 September 2020

Accepted 25 September 2020

Published 22 October 2020

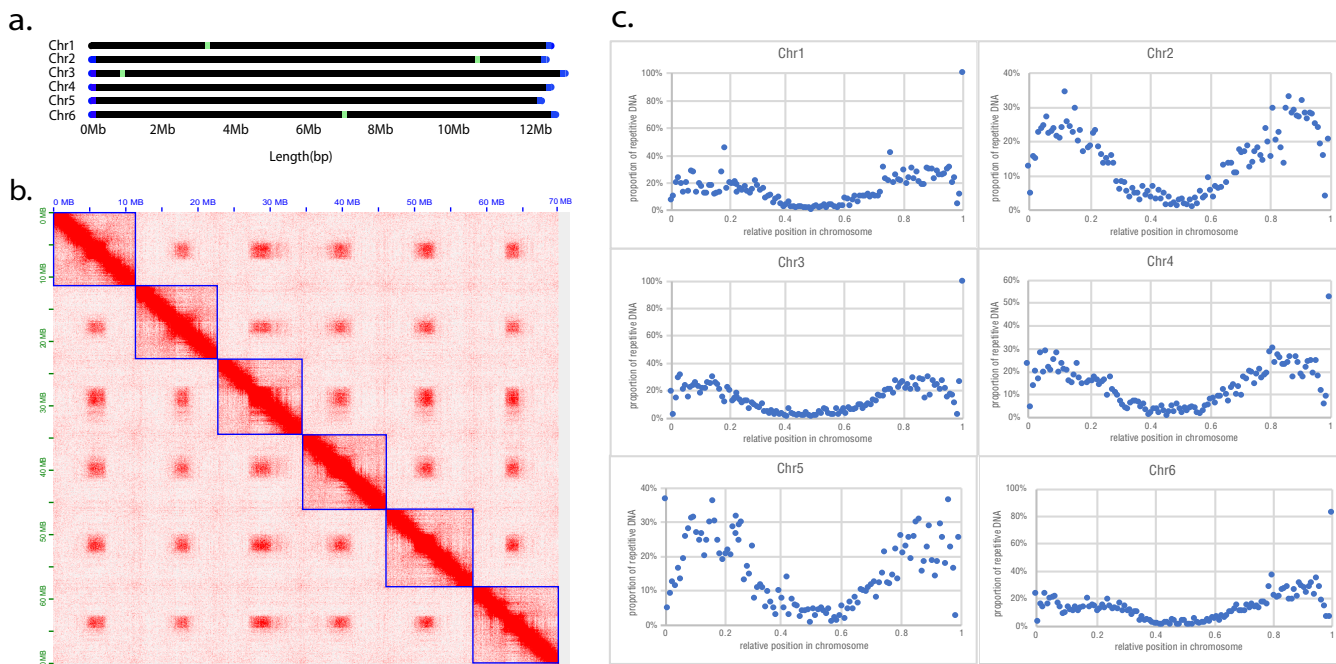


FIG 1 Six chromosomes of the *Bursaphelenchus okinawaensis* genome assembly. (a) Scaffold map of the 70-Mbp *B. okinawaensis* genome. Blue boxes represent telomere repeat signatures, and green boxes indicate remaining gaps in the assembly. (b) Hi-C contact map representing 6 chromosomes (blue boxes) of the *B. okinawaensis* genome. (c) Nonrandom distribution of repetitive DNA sequences in the 6 chromosomes. A species-specific repeat library was generated using RepeatModeler v1.0.8 (<https://github.com/Dfam-consortium/RepeatModeler>) with the default options, and repeats were identified using RepeatMasker v4.0.9 (<http://www.repeatmasker.org/>) with the default options.

DNA library prep kit (Thermo Fisher Scientific) according to the manufacturers' protocols and was sequenced using a MiSeq system with the MiSeq reagent kit v3 (101 cycles \times 2), and the 3.6 million short reads were quality checked using the Hi-C quality control pipeline (<https://phasegenomics.github.io/2019/09/19/hic-alignment-and-qc.html>).

Nanopore long reads were assembled using Flye v2.7.1 (7) with the following parameters: --genome size 70 M and -iteration 4. The Flye assembly was highly contiguous, comprising eight long (1- to 12-Mb) contigs and one short (6-kb) contig. Base correction was then performed with the Illumina paired-end reads with two rounds of Pilon v1.23 (8). To scaffold contigs and confirm assembly fidelity, we performed Hi-C analysis on the Flye assembly using the 3D-DNA pipeline v180114 with the following parameters: -g 50 -r 2 --editor-coarse-resolution 2500000 --editor-coarse-region 10000000 --editor-fine-resolution 1000000 --polisher-input-size 20000000 --splitter-input-size 20000000 (9). Juicebox v1.11.08 (10) was used for visualization of the Hi-C results.

The resulting 70.0-Mbp assembly had a GC content of 36.2%. It comprised six scaffolds (>11 Mbp each) with four gaps and one ~ 6 -kbp contig. BUSCO v3.1 completeness analyses (11) of the assembled genome revealed that 92.1% of core eukaryote genes are present in this assembly. Notably, we identified a telomere repeat signature (TTAGGC) $_n$ at both ends of five scaffolds, indicating complete chromosomal sequences at one end of the sixth scaffold (Fig. 1).

Data availability. The *B. okinawaensis* v2 assembly has been deposited in DDBJ/EMBL/GenBank under BioProject number [PRJEB40023](https://www.ncbi.nlm.nih.gov/bioproject/PRJEB40023). The raw Illumina, Nanopore, and Hi-C read data are available in the Sequence Read Archive with accession numbers [DRR243691](https://www.ncbi.nlm.nih.gov/sra/DRR243691), [DRR243689](https://www.ncbi.nlm.nih.gov/sra/DRR243689), and [DRR243690](https://www.ncbi.nlm.nih.gov/sra/DRR243690), respectively.

ACKNOWLEDGMENTS

We thank Asuka Konosu, Melis Konno, Ryusei Tanaka, Igor Antoschekhin, and the Jacobs Genome Facility for their technical assistance.

This work was funded in part by Japan Society for the Promotion of Science (JSPS) KAKENHI grant numbers 26460510 and 19H03212 and JST CREST grant number JPMJCR18S7.

REFERENCES

1. Kanzaki N, Maehara N, Aikawa T, Togashi K. 2008. First report of parthenogenesis in the genus *Bursaphelenchus* Fuchs, 1937: a description of *Bursaphelenchus okinawaensis* sp. nov. isolated from *Monochamus maruokai* (Coleoptera: Cerambycidae). *Zool Sci* 25:861–873. <https://doi.org/10.2108/zsj.25.861>.
2. Shinya R, Hasegawa K, Chen A, Kanzaki N, Sternberg PW. 2014. Evidence of hermaphroditism and sex ratio distortion in the fungal feeding nematode *Bursaphelenchus okinawaensis*. *G3 (Bethesda)* 4:1907–1917. <https://doi.org/10.1534/g3.114.012385>.
3. Jones JT, Moens M, Mota M, Li H, Kikuchi T. 2008. *Bursaphelenchus xylophilus*: opportunities in comparative genomics and molecular host-parasite interactions. *Mol Plant Pathol* 9:357–368. <https://doi.org/10.1111/j.1364-3703.2007.00461.x>.
4. Thorne G. 1961. *Principles of nematology*. McGraw-Hill, New York, NY.
5. Kikuchi T, Aikawa T, Kosaka H, Pritchard L, Ogura N, Jones JT. 2007. Expressed sequence tag (EST) analysis of the pine wood nematode *Bursaphelenchus xylophilus* and *B. mucronatus*. *Mol Biochem Parasitol* 155:9–17. <https://doi.org/10.1016/j.molbiopara.2007.05.002>.
6. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>.
7. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
8. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
9. Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP, Aiden EL. 2017. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* 356:92–95. <https://doi.org/10.1126/science.aal3327>.
10. Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. 2016. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst* 3:99–101. <https://doi.org/10.1016/j.cels.2015.07.012>.
11. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.