JOURNAL OF NEMATOLOGY

Research Paper | DOI: 10.2478/jofnem-2022-0047

e2022-1 | Vol. 54

Draft Genome of the Insect-Parasitic Nematode *Bradynema listronoti*

Dave T. Ste-Croix, Annie-Ève Gagnon and Benjamin Mimee*

Agriculture and Agri-Food Canada, Saint-Jean-sur-Richelieu Research and Development Centre, Saint-Jean-sur-Richelieu, QC J3B 3E6, Canada

*E-mail: benjamin.mimee@agr.gc.ca

This paper was edited by Erik C. Andersen.

Received for publication June 2, 2022.

Abstract

Bradynema listronoti is an insect-parasitic nematode known to infect the carrot weevil, *Listronotus oregonensis*. We present the first sequence for this species and for any Allantonematidae, produced with a combination of short and long reads. The draft genome of *B. listronoti* is 80.6 Mb in size, assembled in 152 scaffolds.

Keywords

Bradynema listronoti, entomoparasitic nematode, genomics, MinION, NovaSeq, whole-genome amplification

The genus *Bradynema* (Nematoda: Allantonematidae), first characterized by Strassen in 1892, currently comprises nine species of insect-parasitic nematodes with a diverse set of hosts. In 2007, a novel species, *Bradynema listronoti*, was isolated from the hemocoel of adult carrot weevils, *Listronotus oregonensis* (Coleoptera: Curculionidae), an important pest of horticultural crops (Zeng *et al*., 2007). It was shown that *B. listronoti* castrates its host by inhibiting the maturation of the reproductive system in females (Gagnon *et al*., 2019), and its potential as a biological control agent is currently investigated. In comparison to entomopathogenic nematodes, this group has been studied very little, despite its importance in the regulation of insect populations. In this study, we combined short- and long-read sequencing data to assemble a high-quality genome for *B. listronoti*.

B. listronoti specimens (strain AAFC-BrLi-01) used in this study were isolated from a colony of infected carrot weevils at the Saint-Jean-sur-Richelieu Research and Development Centre of Agriculture and Agri-Food Canada. This colony was established following the initial isolation of *B. listronoti* by Zeng *et al*. (2007). Genomic DNA was obtained from an individual juvenile (J3) nematode and a pool of J3. For the single nematode, DNA was extracted and amplified using the REPLI-g Advanced DNA Single Cell Kit (Qiagen, Hilden, Germany). To resolve the hyperbranching generated by this whole-genome amplification (WGA) protocol, 5 µg of WGA gDNA was treated with T7 endonuclease, and short fragments were removed using the Circulomics SRE XS Kit (Circulomics, Baltimore, MD). From this size-selected material, two Nanopore MinION libraries were generated, from 1 µg of genomic material each, using the Ligation Sequencing Kit chemistry (SQK-LSK109, Oxford Nanopore Technologies, Oxford, United Kingdom). Short fragments were eliminated using the L Fragment Buffer provided in the kit prior to sequencing on MinION R9.4.1 flow cells. Using the same WGA gDNA, a short-read shotgun sequencing library was prepared using the NEBNext Ultra II DNA Library Preparation Kit and paired-end-sequenced on the Illumina NovaSeq6000 platform (2 × 150 bp) at the McGill University and Génome Québec Innovation Centre. In a second experiment, high-molecular weight (HMW) gDNA was prepared from a pool of ~15,000 J3 nematodes for downstream assembly scaffolding purposes. Genomic DNA was extracted using the Monarch HMW DNA Extraction Kit for Tissue (New England Biolabs, Ipswich, MA) following the manufacturer's recommendations. The sequencing library was prepared from 1 µg of genomic material using the Ligation Sequencing Kit (Q20+) chemistry

Open Access. Published by Sciendo. CC BY

© 2022 His Majesty the King in Right of Canada. This is an Open Access article licensed under the Creative Commons CC BY 4.0 license, https://creativecommons.org/licenses/by/4.0/

(SQK-LSK112) and sequenced on a MinION R10.4 flow cell. Messenger RNA was extracted and amplified from multiple *B. listronoti* stages (egg, J2, J3, and J4) using the method described in Ste-Croix *et al*. (2021) with three biological replicates of each. Sequencing libraries were prepared using the NEBNext Ultra II DNA Library Preparation Kit and paired-end-sequenced on the NovaSeq6000 platform (2 × 150 bp). Genome size was estimated using GenomeScope (Vurture *et al*., 2017) with k-mer frequency distributions derived from single-nematode whole-genome shotgun sequencing data.

The different sequencing technologies generated 33 Gb of long-read sequences alongside 5.3 Gb of Q20+ long-read sequences derived from HMW gDNA, and 20.7 Gb of paired-end short reads. Additionally, mRNA sequencing generated 706 M reads for an estimated 210 Gb of data. For long- and short-read sequencing data, sequences were trimmed and filtered based on quality using Guppy (v 6.0.6) and Fastp (v 0.20.1 – Chen *et al*., 2018), respectively. The genome was assembled using sequencing data generated from a single nematode to reduce the level of heterozygosity and facilitate assembly. First, long-read sequences were assembled using Flye (v 2.9-b1768 – Kolmogorov *et al*., 2019) in default setting. Second, short-reads were assembled using HASLR (v 0.8a1 – Haghshenas *et al*., 2020) assisted by long-read sequences. These assemblies were then merged using Quickmerge (v 0.3) prior to scaffolding using LINKS (v 1.8.7 – Warren *et al*., 2015) with Q20+ long reads derived from the pool of J3. This assembly was further scaffolded using information provided by RNA-seq with Rascaf (v 1.0.2 – Song *et al*., 2016). Finally, this scaffolded assembly was polished and corrected using five cycles of Pilon (v 1.23 – Walker *et al*., 2014) supplemented with single-nematode whole-genome sequencing data. The final corrected draft was analyzed for contaminating scaffolds with BlobToolKit (v 3.1.6 – Challis *et al*., 2020), and the quality and completeness of the genome were evaluated using BUSCO with the built-in Augustus gene predictor option set (v 5.3.2 – eukaryota_odb10 – Manni *et al*., 2021). The genome was then masked using RepeatMasker (v 4.1.2-p1 – Smit *et al*., 2015), annotated using BRAKER (Stanke *et al*., 2006, 2008; Li *et al*., 2009; Barnett *et al*., 2011; Lomsadze *et al*., 2014; Hoff *et al*., 2016, 2019; Bruna *et al*., 2021), retrained using the RNA-seq data. The draft genome size of *B. listronoti* was 80,637,718 bp assembled into 152 scaffolds, with an average G + C content of 33.08%, an L50 of 7 contigs, and an N50 of 4,373,361 bp. Completeness analysis using eukaryotic BUSCO genes yielded a score of 75.3% (C: 75.3% [S: 61.6%, D: 13.7%], F: 9.0%, M: 15.7%, n: 255). This is slightly lower than that obtained with the genome of *Deladenus siricidicola* (C: 87.1% [S: 86.3%, D: 0.8%], F: 5.5%, M: 7.4%, n: 255), a phylogenetically related species of insect-parasitic nematode (assembly accession GCA_009724625.1). Also, 13.7% of these BUSCO genes were found to be duplicated, indicating that a low level of haplotypic duplication may remain. This could also explain the small difference between the genome size (80.6 Mb) and its estimation by GenomeScope (69.8 Mb). BRAKER predicted 16,831 protein-coding genes. Functional annotation was accomplished by querying these predictions against the NCBI nr database using BLASTP (v. 2.6.0 – Altschul *et al*., 1990). This is the first genome release for *B. listronoti* and for any Allantonematidae species.

Data availability and accession numbers

All the raw sequences and genome assembly files are available under the NCBI BioProject PRJNA842945. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JANLNO000000000.

Acknowledgements

The authors thank Julie Frenette and Danielle Thibodeau for technical assistance with insect and nematode rearings, and Joël Lafond-Lapalme and Pierre-Yves Véronneau for bioinformatics assistance. This work was funded by Agriculture and Agri-Food Canada (APMS, J-002846).

Literature Cited

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. Journal of Molecular Biology 215:403–410.

Barnett, D. W., Garrison, E. K., Quinlan, A. R., Strömberg, M. P., and Marth, G. T. 2011. BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics 27:1691–1692

Bruna, T., Hoff, K. J., Lomsadze, A., Stanke, M., and Borodovsky, M. 2021. BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS Supported by a Protein Database. NAR Genomics and Bioinformatics 3:lqaa108.

Challis, R., Richards, E., Rajan, J., Cochrane, G., and Blaxter, M. 2020. BlobToolKit – Interactive quality assessment of genome assemblies. G3 Genes|Genomes|Genetics 10:1361–1374.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. 2018. fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890.

Gagnon, A.-È., Boivin, G., Bélair, G., and Mimee, B. 2019. Prevalence of a nematode castrator of the carrot weevil and impact on fecundity and survival. Parasitology 146:702–707.

Haghshenas, E., Asghari, H., Stoye, J., Chauve, C., and Hach, F. 2020. HASLR: Fast hybrid assembly of long reads. Cold Spring Harbor Laboratory. doi/10.1101/2020.01.27.921817.

Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M., and Stanke, M. 2016. BRAKER1: Unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. Bioinformatics 32:767–769.

Hoff, K. J., Lomsadze, A., Borodovsky, M., and Stanke, M. 2019. Whole-genome annotation with BRAKER. Methods in Molecular Biology 1962:65–95.

Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P. A. 2019. Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology 37:540–546.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.; and 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.

Lomsadze, A., Burns, P. D., and Borodovsky, M. 2014. Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. Nucleic Acids Research 42:e119.

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., and Zdobnov, E. M. 2021. BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Molecular Biology and Evolution 38:4647–4654.

Smit, A., Hubley, R., and Green, P. 2015. RepeatMasker Open-4.0. 2013–2015. Institute for Systems Biology. Available at: http://repeatmasker.org.

Song, L., Shankar, D. S., and Florea, L. 2016. Rascaf: Improving genome assembly with RNA sequencing data. The Plant Genome 9(3):1–12.

Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. 2008. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics 24:637–644.

Stanke, M., Schöffmann, O., Morgenstern, B., and Waack, S. 2006. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinformatics 7:62.

Ste-Croix, D. T., St-Marseille, A. F. G., Lord, E., Bélanger, R. R., Brodeur, J., and Mimee, B. 2021. Genomic profiling of virulence in the soybean cyst nematode using single-nematode sequencing. Phytopathology 111:137–148.

Vurture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., and Schatz, M. C. 2017. GenomeScope: Fast reference-free genome profiling from short reads. Bioinformatics 33:2202–2204.

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.

Warren, R. L., Yang, C., Vandervalk, B. P., Behsaz, B., Lagman, A., Jones, S. J M., and Birol, I. 2015. LINKS: Scalable, alignment-free scaffolding of draft genomes with long reads. GigaScience 4:35. doi/10.1186/s13742-015-0076-3

Zeng, Y., Giblin-Davis, R., Bélair, G., Boivin, G., and Thomas, K. 2007. *Bradynema listronoti* n. sp. (Nematoda: Allantonematidae), a parasite of the carrot weevil *Listronotus oregonensis* (Coleoptera: Curculionidae) in Quebec, Canada. Nematology 9:608–622.