AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of "*Candidatus* Nardonella dryophthoridicola" Strain NARMHE1, Endosymbiont of *Metamasius hemipterus* (Coleoptera, Curculionidae, Dryophthorinae)

Luciano Palmieri,[a] Ronaldo Pavarini,[b] Prashant P. Sharma[a]

[a]Department of Integrative Biology, University of Wisconsin-Madison, Madison, Wisconsin, USA
[b]Faculdade de Ciências Agrárias do Vale do Ribeira, Campus de Registro, São Paulo State University-UNESP, Registro, São Paulo, Brazil

ABSTRACT Here, we report the draft genome and annotation of "*Candidatus* Nardonella dryophthoridicola" strain NARMHE1, obtained via Oxford Nanopore sequencing of the ovaries of its host, the weevil *Metamasius hemipterus*, from a population from southeast Brazil.

Metamasius (family, Curculionidae; subfamily, Dryophthorinae) harbors obligate endosymbiont *Nardonella*, a group of *Gammaproteobacteria* that have cospeciated with their hosts (1–3). *Metamasius hemipterus* is a pest on sugarcane and other crops and has invasive potential, threatening agriculture (4–7). However, genomic sampling of *Nardonella* remains limited, particularly for host species that exhibit invasive potential. Here, we generate the genome of the *Nardonella* associated with *M. hemipterus*, found on cultivated *Bactris gasipaes* (Arecaceae).

M. hemipterus adults were collected (Pariquera-açu, São Paulo, Brazil; −24.608873, −47.896800) using scent bait traps (8). Larvae were extracted from stems and fixed in ethanol. To detect *Nardonella* presence, 15 females were dissected, and their midguts and ovaries were separated. We also dissected gut tissues of 15 larvae. Samples were immersed in 2% bleach for 60 s and dissected in 1× phosphate-buffered saline (PBS). DNA extractions were performed using Qiagen DNeasy blood and tissue kit following the manufacturer's protocol with modifications (overnight proteinase K incubation, two double-distilled water [ddH_2O; 56°C] elutions, and 10 min final incubation). The DNA concentration was verified with Qubit double-stranded DNA (dsDNA) high-sensitivity (HS) assay kit (Life Technologies). Samples were prepared according to Celero PCR workflow with enzymatic fragmentation (Tecan Genomics). The quality and quantity of the finished libraries were assessed with Agilent TapeStation (Agilent) and Qubit dsDNA HS assay kit. Sequencing was performed using Illumina NovaSeq6000 (2×150 bp).

Raw reads (host and symbiont) were queried in BLASTn (9, 10) to identify sequences of *Nardonella* against the NCBI genome database (E value cutoff, 10^{-6}; see supplemental information for program options). Sequences of interest were extracted using Seqtk1.3 (https://github.com/lh3/seqtk). Ovaries presented 5× more *Nardonella* sequences (Fig. 1A) and were selected for sequencing using Oxford Nanopore Technologies. No shearing/fragmentation was performed on input DNA (LSK110 kits were used for library preparation and run on two GridION-MinION flow cells; high-accuracy base calling min_qscore was 9). Long reads were processed using the same filtering procedures described for Illumina (Table 1). Long reads were deduplicated with BBMap version 38.94 (https://sourceforge.net/projects/bbmap/); sequences shorter than 130 bp were removed with Filtlong version 0.2.0.1 (https://github.com/rrwick/Filtlong). The 30,930-bp-long reads were assembled twice independently using Canu version 2.2 (11) and Flye version 2.8.3 (12). Final contigs for each assembly were corrected using unfiltered

Editor Irene L. G. Newton, Indiana University, Bloomington

Address correspondence to Luciano Palmieri, lucianopalmierirocha@gmail.com.

The authors declare no conflict of interest.

Received 1 August 2022
Accepted 12 October 2022
Published 31 October 2022

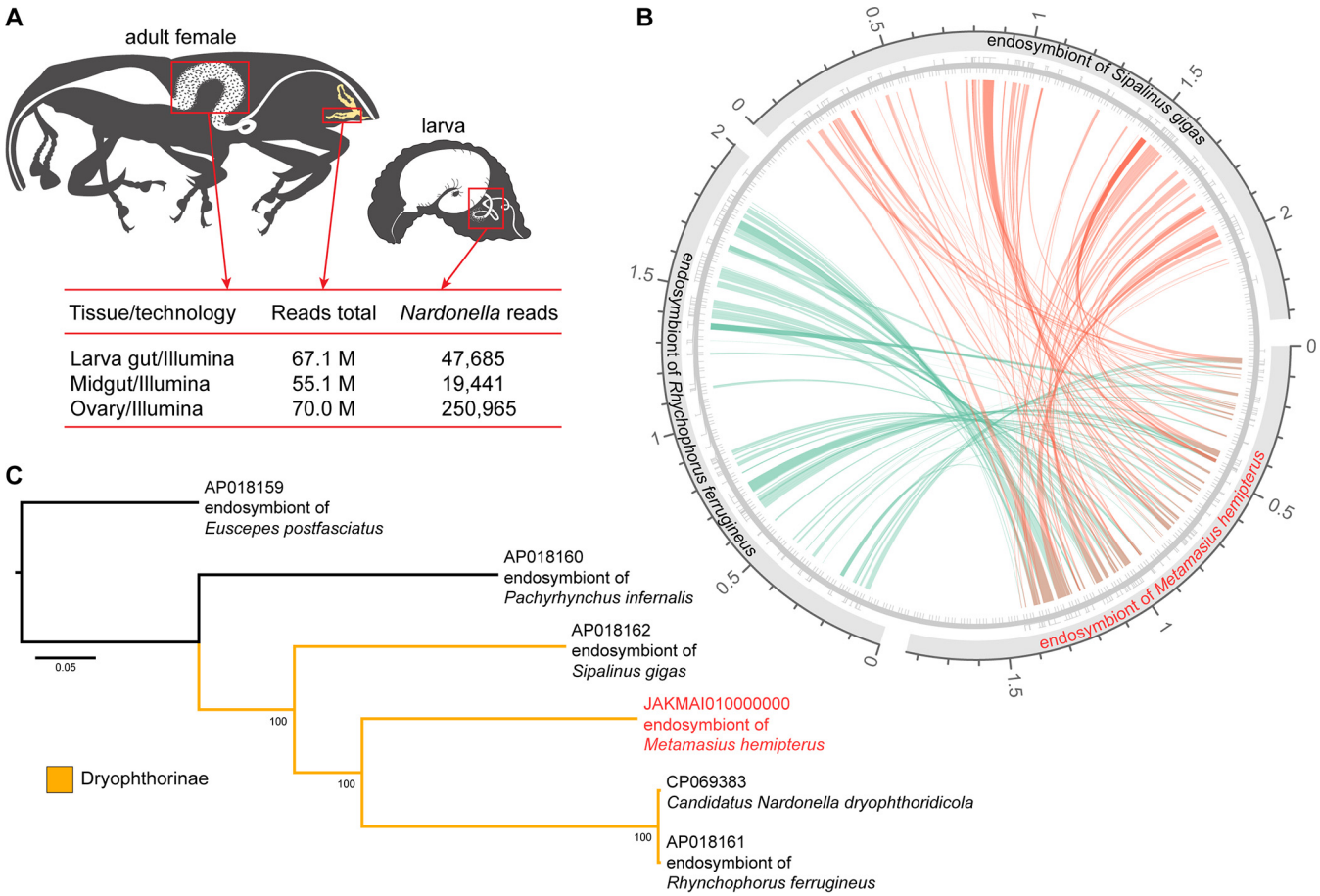

FIG 1 "*Candidatus* Nardonella dryophthoridicola" strain NARMHE1 draft genome assembly. (A) Determination of the most suitable tissue for recovering symbiont sequences. (B) Synteny circus plot showing rearrangement of some relevant genes on resulting draft genome compared with other Dryophthorinae endosymbiont *Nardonella* genomes. (C) Maximum-likelihood tree constructed on RAxML with all available *Nardonella* genomes. GenBank accession numbers of the sequences used are AP018159, endosymbiont of *Euscepes postfasciatus*; AP018160, endosymbiont of *Pachyrhynchus infernalis*; AP018161, endosymbiont of *Rhynchophorus ferrugineus*; AP018162, endosymbiont of *Sipalinus gigas*; CP069383, "*Candidatus* Nardonella dryophthoridicola"; and JAKMAI010000000, endosymbiont of *Metamasius hemipterus* (this study).

long reads with Medaka version 1.6.0 (https://github.com/nanoporetech/medaka). Additional polishing using short reads was made with Polypolish version 0.4.3 (13), with a further long-read polishing made with poLCA version 4.0.6 (14). To improve contiguity, we used quickmerge version 0.3 (15). The final assembly was composed of seven contigs (Table 1). Scaffolding was performed using homology between contigs and reference genomes. Potential misassembles were corrected with RagTag version 2.1.0 (https://github.com/malonge/RagTag). After correction, we used reference protein sequences of the *Rhynchophorus ferrugineus* endosymbiont, GenBank accession no. AP018161, to orient contigs. Stretches of 100 "Ns" were placed between adjacent sequences to indicate gap regions. The final genome was annotated using PGAP version 6.0 (16) (Table 1). Although some genes are incomplete, there was high similarity of genetic composition to other *Nardonella* genomes (Fig. 1B). To ascertain identification, we aligned our contigs to other *Nardonella* genomes with Mauve version 2.4.0 (17). A maximum-likelihood (ML) tree was inferred with RAxML version 8.2.11 (18) (GTR+gamma; 1,000 bootstraps). NARMHE1 formed a strongly supported clade with other Dryophthorinae endosymbionts (Fig. 1C). The positioning of *Nardonella* strains on the phylogeny emulates their dryophthorid hosts (19), suggesting a coevolutionary process (20).

Data availability. All code and software parameters used to produce the results, as well as ONT quality control reports and phylogenetic matrix, are described in the supplemental material publicly available on GitHub repository Nardonella-NARMHE1-genome (https://github.com/LucPalmieri/Nardonella-NARMHE1-genome). The genome version described in this paper is the first version, and it is under GenBank accession no. JAKMAI010000000. Raw

TABLE 1 Combined assembly summary

Characteristic	Value	GenBank accession no.
Long-read characteristics		
No. of reads after deduplication	30,930	
Mean read length (bp)	883	
Longest read length (bp)	14,328	
Shortest read length (bp)	135	
N_{50} (bp)	3,170	
Contig characteristics (no. of bp)		
Contig 1	12,016	JAKMAI010000001
Contig 2	28,356	JAKMAI010000002
Contig 3	16,370	JAKMAI010000003
Contig 4	6,418	JAKMAI010000004
Contig 5	12,133	JAKMAI010000005
Contig 6	13,069	JAKMAI010000006
Contig 7	91,945	JAKMAI010000007
Genome characteristics		
Size (bp)	180,307	
GC content (%)	19.80	
No. of genes	199	
No. of CDSs[a]	144	
No. of RNAs	36	
No. of pseudogenes	19	
Complete rRNAs	5S, 16S, 23S	

[a] CDS, coding DNA sequence.

Illumina and Nanopore sequences are available on NCBI Sequence Read Archive under the accession numbers SRR21424116 and SRR20324089, respectively.

ACKNOWLEDGMENTS

We thank Leonardo Souza Carvalho (CHN-UFPI) for the logistical support with handling of samples. We also thank the University of Wisconsin Biotechnology Center DNA Sequencing Facility for providing next-generation sequencing consultation and services.

Components of this work were supported by National Science Foundation grant number IOS-2016141 and Binational Science Foundation grant number 2019216.

REFERENCES

1. Nardon P, Lefèvre C, Delobel B, Charles H, Heddi A. 2002. Occurrence of endosymbiosis in Dryophthoridae weevils: cytological insights into bacterial symbiotic structures. Symbiosis 30:227–241.
2. Lefèvre C, Charles H, Vallier A, Delobel B, Farrell B, Heddi A. 2004. Endosymbiont phylogenesis in the Dryophthoridae weevils: evidence for bacterial replacement. Mol Biol Evol 21:965–973. https://doi.org/10.1093/molbev/msh063.
3. Hosokawa T, Koga R, Tanaka K, Moriyama M, Anbutsu H, Fukatsu T. 2015. Nardonella endosymbionts of Japanese pest and non-pest weevils (Coleoptera: Curculionidae). Appl Entomol Zool 50:223–229. https://doi.org/10.1007/s13355-015-0326-y.
4. Weissling TJ, Giblin-Davis RM. 2017. Silky cane weevil, Metamasius hemipterus sericeus (Olivier) (Insecta: Coleoptera: Curculionidae), p 1–5. IFAS Extension, University of Florida, Gainesville, FL.
5. Molet T. 2013. CPHST pest datasheet for Metamasius hemipterus. USDA-APHIS-PPQ-CPHST. USDA. http://download.ceris.purdue.edu/file/2231.
6. Frank JH, Fish D. 2008. Potential biodiversity loss in Florida Bromeliad Phytotelmata due to Metamasius callizona (Coleoptera: Dryophthoridae), an invasive species. Florida Entomol 91:1–8. https://doi.org/10.1653/0015-4040(2008)091[0001:PBLIFB]2.0.CO;2.
7. Thorn MJ, Seltzer JL, Hendon AG, Keasler BL, Hill JG. 2019. The first detection of Metamasius hemipterus (L.) (Coleoptera: Curculionidae) from coastal Mississippi, USA. Trans Am Entomol Soc 145:407–412. https://doi.org/10.3157/061.145.0304.
8. Soliman EP, Garcia VA, Pavarini R, Lima RC, Nomura ES, Pavarini GMP. 2010. Evaluation of the Attractiveness of Different Baits to Rhynchophorus palmarum (Coleoptera: Curculionidae) at Peach Palm Cultivation (Bactris gasipaes). Nucleus 7:197–202. (In Portuguese.) https://doi.org/10.3738/1982.2278-350.
9. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421.
10. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
11. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.
12. 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.
13. Wick RR, Holt KE. 2022. Polypolish: short-read polishing of long-read bacterial genome assemblies. PLoS Comput Biol 18:e1009802. https://doi.org/10.1371/journal.pcbi.1009802.
14. Zimin AV, Salzberg SL. 2020. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. PLoS Comput Biol 16:e1007981. https://doi.org/10.1371/journal.pcbi.1007981.
15. Chakraborty M, Baldwin-Brown JG, Long AD, Emerson JJ. 2016. Contiguous and accurate de novo assembly of metazoan genomes with modest

long read coverage. Nucleic Acids Res 44:gkw654. https://doi.org/10.1093/nar/gkw654.
16. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.
17. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403. http://www.genome.org/cgi/doi/10.1101/gr.2289704.
18. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
19. Chamorro ML, de Medeiros BAS, Farrell BD. 2021. First phylogenetic analysis of Dryophthorinae (Coleoptera, Curculionidae) based on structural alignment of ribosomal DNA reveals Cenozoic diversification. Ecol Evol 11:1984–1998. https://doi.org/10.1002/ece3.7131.
20. Zhang G, Browne P, Zhen G, Johnston A, Cadillo-Quiroz H, Franz N. 2017. Endosymbiont diversity and evolution across the weevil tree of life. bioRxiv. https://doi.org/10.1101/171181.