AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Rhynchophorus ferrugineus* Endocytobiont "*Candidatus* Nardonella dryophthoridicola" Strain NardRF

Bessem Chouaia,[a] Matteo Montagna,[b,c] Pompeo Suma,[d] Franco Faoro[b]

[a]Department of Molecular Sciences and Nanosystems, Ca' Foscari University of Venice, Venice, Italy
[b]Department of Agricultural and Environmental Sciences, University of Milan, Milan, Italy
[c]Interuniversity Center for Studies on Bioinspired Agro-Environmental Technology (BAT Center), Università di Napoli Federico II, Portici, Italy
[d]Department of Agriculture, Food and Environment, University of Catania, Catania, Italy

ABSTRACT We report the complete genome sequence and annotation of "*Candidatus* Nardonella dryophthoridicola" strain NardRF, obtained by sequencing its host bacteriome, *Rhynchophorus ferrugineus*, using Oxford Nanopore technology.

The bacterium "*Candidatus* Nardonella dryophthoridicola" is a Gram-negative gammaproteobacterial endocytobiont (Fig. 1). Specifically, it is an intracellular obligate mutualist associated with weevils (1). The bacterium plays a crucial role in cuticle hardening by supplying tyrosine to its host (2). Unlike the second weevil-associated symbiont, "*Candidatus* Sodalis pierantonius," it is maintained within a functional bacteriome for its host's entire life cycle (3–5).

We used long-read sequencing to investigate the genome sequence of "*Ca.* Nardonella dryophthoridicola" strain NardRF, associated with an Italian population of *Rhynchophorus ferrugineus*. The insect hosts were sampled from a single palm tree in the region of Catania in 2017. The pupae were kept at 25°C, 24-h dark, until molting into adults. Ten newly emerged adults were dissected to extract their bacteriomes. The bacteriomes were then pooled for DNA extraction using the DNeasy blood and tissue kit (Qiagen, Italy) following the manufacturer's instructions for animal tissue extraction. The DNA integrity was verified by 0.8% agarose gel electrophoresis at 90 V for 1 h. The DNA purity and concentration were measured with a NanoDrop 100 spectrophotometer (Thermo Fisher Scientific, Italy) and Qubit double-stranded DNA (dsDNA) high-sensitivity assay kit.

Long-read sequencing was performed using the R9.5 flow cell on a MinION Mk1B device. For the library preparation, 2.5 μg of nonsheared and non-size-selected total genomic DNA was used following the 1D ligation sequencing kit (SQK-LSK 108) protocol. Then, 0.5 μg of the final DNA was loaded onto the flow cell. The sequencing was run for 48 h using MinKNOW v18.03.1. Base calling was then run on the fast5 files using Guppy v4.4.1 (6) with the high-accuracy algorithm and a quality cutoff of 7. Reads longer than 500 bp were used for the subsequent analyses. All tools were run with default parameters unless otherwise specified.

The metagenomics fastq reads (host and symbiont) were first assembled using miniasm (7). Contigs identified as "*Ca.* Nardonella dryophthoridicola" were identified using BLASTn (E value cutoff, 10^{-6}) against the NCBI nonredundant (nr) database. These contigs were extracted and used to refine the assembly. The contigs were used to map and extract the "*Ca.* Nardonella dryophthoridicola" long reads using minimap2 v2.17 (8). The 836,116 reads were then reassembled using Flye v2.8.1 (9). The resulting genome was circularized using Circlator v1.5.5 (10) with the options –merge_min_id 85 and –merge_breaklen 1000 as advised for Oxford Nanopore reads. The circular genome was corrected using the publicly available Illumina short reads (SRA accession

Citation Chouaia B, Montagna M, Suma P, Faoro F. 2021. Complete genome sequence of *Rhynchophorus ferrugineus* endocytobiont "*Candidatus* Nardonella dryophthoridicola" strain NardRF. Microbiol Resour Announc 10: e00355-21. https://doi.org/10.1128/MRA.00355-21.

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

Address correspondence to Bessem Chouaia, bessem.chouaia@unive.it.

Received 5 April 2021
Accepted 31 May 2021
Published 1 July 2021

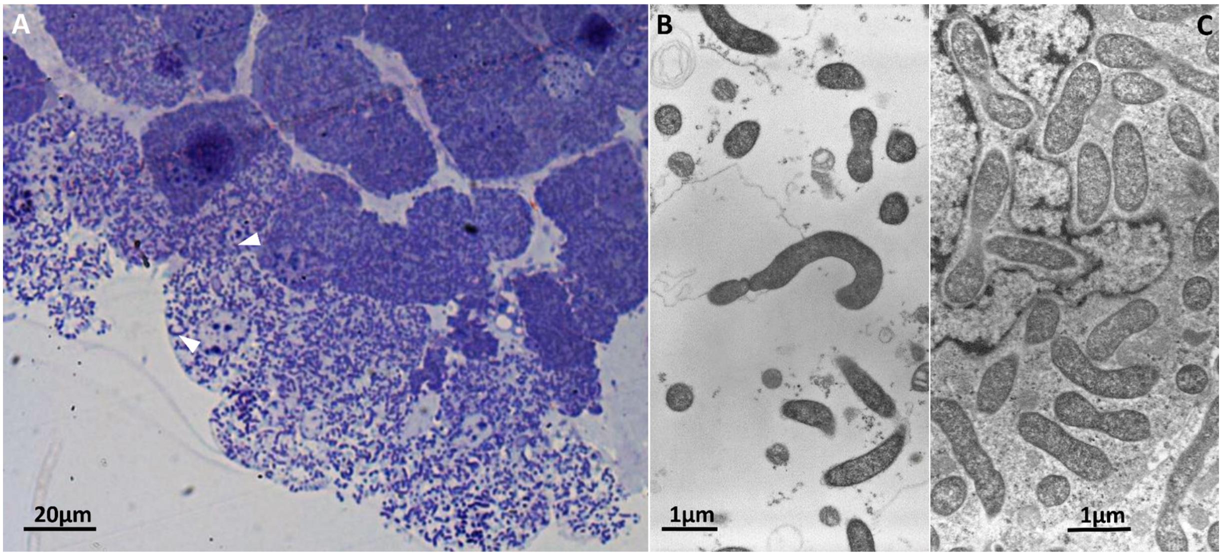

FIG 1 The endocytobiont "*Candidatus* Nardonella dryophthoridicola." (A) Semithick cross-section of the *Rhynchophorus ferrugineus* bacteriome in which it is possible to observe the bacterial cells, stained with toluidine blue, within the host cell (white arrowheads). (B and C) Ultrathin sections of the same bacteriome under transmission electron microscopy (TEM) showing "*Ca.* Nardonella" rod-shaped cells outside (B) and within (C) the host cell.

number SRR12633329 [11]) with POLCA (MaSuRCA v4.0.1) (12, 13). During the different assembly, circularization, and polishing steps, the genome quality was assessed using BUSCO v4.1.4 (14) with the Gammaproteobacteria database. The final genome was automatically annotated using GenBank with PGAP r2021-01-09.build5126 (Table 1) (15).

Genome comparison with the closest genome (RefSeq accession number NZ_AP018161 [2]), using ACT (Artemis v18.1.0 [16]), revealed that the gene encoding the isoleucine tRNA ligase (*ileS*) was complete in our genome, while containing a 1-nucleotide frameshift at position 820. This difference demonstrates the importance of sequencing the same streamlined bacterial endocytobiont from different host populations, as genome reduction through random genetic mutations combined with a maternal transmission bottleneck can result in genomic differences within the same endosymbiont species.

Data availability. The assembly has been deposited in GenBank under accession number CP069383 and BioProject accession number PRJNA699994. The version described

TABLE 1 "*Candidatus* Nardonella dryophthoridicola" strain NardRF long-read and genomic summary features

	Data for:	
Feature	**Metagenome**	**Strain NardRF**
Long-read features		
No. of reads	3,474,690	836,116
Mean read length (bp)	2,021	2,018
Longest read (bp)	114,533	88,252
Shortest read (bp)	500	500
N_{50} (bp)	3,035	2,991
Genome features		
Size (bp)	NA[a]	200,313
GC content (%)	NA	15.33
No. of genes	NA	231
No. of CDS[b]	NA	199
No. of RNAs	NA	32
No. of ribosomal operons	NA	1

[a] NA, not applicable.
[b] CDS, coding DNA sequences.

in this paper is the first version, CP069383.1. The Oxford Nanopore reads used for the assembly of "*Ca*. Nardonella dryophthoridicola" have been deposited under SRA accession number SRR14598013.

REFERENCES

1. Kuriwada T, Hosokawa T, Kumano N, Shiromoto K, Haraguchi D, Fukatsu T. 2010. Biological role of Nardonella endosymbiont in its weevil host. PLoS One 5:e13101. https://doi.org/10.1371/journal.pone.0013101.
2. Anbutsu H, Moriyama M, Nikoh N, Hosokawa T, Futahashi R, Tanahashi M, Meng X-Y, Kuriwada T, Mori N, Oshima K, Hattori M, Fujie M, Satoh N, Maeda T, Shigenobu S, Koga R, Fukatsu T. 2017. Small genome symbiont underlies cuticle hardness in beetles. Proc Natl Acad Sci U S A 114:E8382–E8391. https://doi.org/10.1073/pnas.1712857114.
3. Vigneron A, Masson F, Vallier A, Balmand S, Rey M, Vincent-Monégat C, Aksoy E, Aubailly-Giraud E, Zaidman-Rémy A, Heddi A. 2014. Insects recycle endosymbionts when the benefit is over. Curr Biol 24:2267–2273. https://doi.org/10.1016/j.cub.2014.07.065.
4. Maire J, Chouaia B, Zaidman-Rémy A, Heddi A. 2020. Endosymbiosis morphological reorganization during metamorphosis diverges in weevils. Commun Integr Biol 13:184–188. https://doi.org/10.1080/19420889.2020.1840707.
5. Maire J, Parisot N, Ferrarini MG, Vallier A, Gillet B, Hughes S, Balmand S, Vincent-Monégat C, Zaidman-Rémy A, Heddi A. 2020. Spatial and morphological reorganization of endosymbiosis during metamorphosis accommodates adult metabolic requirements in a weevil. Proc Natl Acad Sci U S A 117:19347–19358. https://doi.org/10.1073/pnas.2007151117.
6. Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. Genome Biol 20:129. https://doi.org/10.1186/s13059-019-1727-y.
7. Li H. 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. Bioinformatics 32:2103–2110. https://doi.org/10.1093/bioinformatics/btw152.
8. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191.
9. 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.
10. Hunt M, De Silva N, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. https://doi.org/10.1186/s13059-015-0849-0.
11. Hazzouri KM, Sudalaimuthuasari N, Kundu B, Nelson D, Al-Deeb MA, Le Mansour A, Spencer JJ, Desplan C, Amiri KMA. 2020. The genome of pest *Rhynchophorus ferrugineus* reveals gene families important at the plant-beetle interface. Commun Biol 3:323. https://doi.org/10.1038/s42003-020-1060-8.
12. 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.
13. Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. Bioinformatics 29:2669–2677. https://doi.org/10.1093/bioinformatics/btt476.
14. Waterhouse RM, Seppey M, Simao FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. Mol Biol Evol 35:543–548. https://doi.org/10.1093/molbev/msx319.
15. 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.
16. Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. 2005. ACT: the Artemis comparison tool. Bioinformatics 21:3422–3423. https://doi.org/10.1093/bioinformatics/bti553.