



Complete Reference Genome Assembly for *Commensalibacter* sp. Strain AMU001, an Acetic Acid Bacterium Isolated from the Gut of Honey Bees

Stefanos Siozios,^a Josephine Moran,^{a*} Mary Chege,^b Gregory D. D. Hurst,^a Juan C. Paredes^b

^aInstitute of Integrative Biology, Faculty of Health and Life Sciences, University of Liverpool, Liverpool, United Kingdom

^bInternational Centre of Insect Physiology and Ecology (ICIPE), Nairobi, Kenya

ABSTRACT We report here the genome sequence of a *Commensalibacter* sp. strain (AMU001) isolated from honey bees (*Apis mellifera*) from Seychelles. By combining long- and short-read sequencing technologies, we produced the first complete reference genome assembly for the *Commensalibacter* genus. We anticipate that this will aid future comparative and functional genomic studies.

Acetic acid bacteria (family *Acetobacteraceae*) represent important components of the gut microbial community in many insect taxa that feed on a sugar diet (1). *Commensalibacter* is a newly described genus within this family (type species, *Commensalibacter intestini*). It was previously identified as a major component of the *Drosophila* gut microbiota (2, 3), where it is associated with gut innate immune homeostasis (4). In honey bees, *Commensalibacter* sp. (also referred to as “Alpha 2.1”) has been reported as a core member of the gut microbiota (5, 6) observed mainly in the midgut and hindgut (7). Currently, draft genome sequences are available for three *Commensalibacter* isolates (GenBank database, October 2018) (8–10). Here, we combined Oxford Nanopore and Illumina sequencing technologies to produce the first complete and closed genome assembly of *Commensalibacter* sp. strain AMU001, which was isolated from a Seychelles population of honey bees.

A dominant *Commensalibacter* phylotype was identified in the gut microbiota from a Seychelles population of honey bees, and a cloned strain was isolated from gut homogenates on de Man-Rogosa-Sharpe (MRS) agar medium. High-molecular-weight DNA was extracted using a phenol-chloroform extraction method (11). A Nanopore sequencing library was prepared using the Rapid Sequencing (SQK-RAD004) protocol (Oxford Nanopore), with slight modifications of the manufacturer’s protocol; in total, 3 μ g of DNA were used as input, and library loading beads were omitted. Sequencing was performed on a FLO-MIN106 R9 MinION flow cell. Raw Nanopore signals were live base-called using the processing pipeline implemented in MinKNOW software version 18.01.6 (Oxford Nanopore). Illumina sequencing was performed by MicrobesNG (Birmingham, AL) using the Nextera XT library prep protocol on a MiSeq platform (Illumina, San Diego, CA, USA), and reads were adapter trimmed using Trimmomatic 0.30, with a sliding window quality cutoff of Q15 (12). A *de novo* hybrid assembly of Nanopore and Illumina reads was performed using the Unicycler pipeline version 0.4.5 under the normal mode (13). We assessed the quality of the final assembly by estimating the fraction of interrupted open reading frames (ORFs) using the Ideel method (<https://github.com/mw55309/ideel>) (14). Additionally, we checked for potential misassemblies by remapping the raw reads on the final assembly and visually inspecting for inconsistencies.

The final assembly of the *Commensalibacter* strain AMU001 consists of a circular chromosome of 2,013,417 bp (GC content, 37.9%), with an additional circular plasmid of 10,660 bp (GC content, 31.3%). The average depths of coverage of the Nanopore reads were 67 \times and 835 \times for the chromosome and the plasmid, respectively. The Illumina

Citation Siozios S, Moran J, Chege M, Hurst GDD, Paredes JC. 2019. Complete reference genome assembly for *Commensalibacter* sp. strain AMU001, an acetic acid bacterium isolated from the gut of honey bees. *Microbiol Resour Announc* 8:e01459-18. <https://doi.org/10.1128/MRA.01459-18>.

Editor Christina Cuomo, Broad Institute of MIT and Harvard University

Copyright © 2019 Siozios 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 Stefanos Siozios, siozioss@liverpool.ac.uk.

* Present address: Josephine Moran, Division of Infection, Immunity and Respiratory Medicine, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom.

Received 29 October 2018

Accepted 20 November 2018

Published 3 January 2019

data had a relatively uniform coverage of $\sim 179\times$. Genome annotation using the NCBI Prokaryotic Genome Annotation Pipeline (15) identified 1,764 protein-coding sequences (CDSs) on the main chromosome and 8 CDSs on the plasmid. The fraction of interrupted ORFs was estimated to about 2% (41 genes with length $<80\%$ of the length of their top hit in UniProt database). Finally, 4 rRNA operons and 51 tRNA genes were identified in the main chromosome.

The *Commensalibacter* sp. genome sequence reported here represents the second for a honey bee-associated *Commensalibacter* strain and the fourth genome reported for the genus. More importantly, is the first complete and closed genome sequence for the genus and thus represents a valuable reference for further comparative and functional genomic studies.

Data availability. The complete genome sequences of the *Commensalibacter* strain AMU001 main chromosome and its plasmid have been deposited in GenBank under the accession numbers [CP033087](#) and [CP033088](#), respectively (BioProject accession number [PRJNA495947](#)). The raw reads are available from the NCBI Sequence Read Archive (SRA) under the accession number [SRP166292](#).

ACKNOWLEDGMENTS

The project was supported by a Newton Funds grant from the British Council to J.C.P. and G.D.D.H. (reference 275898413) and Swiss National Science Foundation Postdoc Mobility (grants P300PA_174329 and P2ELP3_165168) to J.C.P. Illumina genome sequencing was provided by MicrobesNG (<http://www.microbesng.uk>), which is supported by the BBSRC (grant BB/L024209/1). Sampling for this study was funded by The International Fund for Agricultural Development (IFAD) grant 2000000282 under the project “Alternative livelihoods for food and income security in four Indian Ocean Island Nations and in Zanzibar, United Republic of Tanzania,” and icipe core funding provided by UK Aid from the UK Government, Swedish International Development Cooperation Agency (Sida), Swiss Agency for Development and Cooperation (SDC), Federal Democratic Republic of Ethiopia, and the Government of Kenya.

The views expressed herein do not necessarily reflect the official opinion of the donors.

We are grateful to the Seychelles Ministry of Agriculture and Fishery for their help and guidance with sampling.

REFERENCES

- Crotti E, Rizzi A, Chouaib B, Ricci I, Favia G, Alma A, Sacchi L, Bourtzis K, Mandrioli M, Cherif A, Bandi C, Daffonchio D. 2010. Acetic acid bacteria, newly emerging symbionts of insects. *Appl Environ Microbiol* 76: 6963–6970. <https://doi.org/10.1128/AEM.01336-10>.
- Roh SW, Nam Y-D, Chang H-W, Kim K-H, Kim M-S, Ryu J-H, Kim S-H, Lee W-J, Bae J-W. 2008. Phylogenetic characterization of two novel commensal bacteria involved with innate immune homeostasis in *Drosophila melanogaster*. *Appl Environ Microbiol* 74:6171–6177. <https://doi.org/10.1128/AEM.00301-08>.
- Chandler JA, Lang JM, Bhatnagar S, Eisen JA, Kopp A. 2011. Bacterial communities of diverse *Drosophila* species: ecological context of a host-microbe model system. *PLoS Genet* 7:e1002272. <https://doi.org/10.1371/journal.pgen.1002272>.
- Ryu J-H, Kim S-H, Lee H-Y, Bai JY, Nam Y-D, Bae J-W, Lee DG, Shin SC, Ha E-M, Lee W-J. 2008. Innate immune homeostasis by the homeobox gene caudal and commensal-gut mutualism in *Drosophila*. *Science* 319: 777–782. <https://doi.org/10.1126/science.1149357>.
- Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, Moran NA. 2011. A simple and distinctive microbiota associated with honey bees and bumble bees. *Mol Ecol* 20:619–628. <https://doi.org/10.1111/j.1365-294X.2010.04959.x>.
- Kwong WK, Moran NA. 2016. Gut microbial communities of social bees. *Nat Rev Microbiol* 14:374–384. <https://doi.org/10.1038/nrmicro.2016.43>.
- Anderson KE, Sheehan TH, Mott BM, Maes P, Snyder L, Schwan MR, Walton A, Jones BM, Corby-Harris V. 2013. Microbial ecology of the hive and pollination landscape: bacterial associates from floral nectar, the alimentary tract and stored food of honey bees (*Apis mellifera*). *PLoS One* 8:e83125. <https://doi.org/10.1371/journal.pone.0083125>.
- Kim E-K, Kim S-H, Nam H-J, Choi MK, Lee K-A, Choi S-H, Seo YY, You H, Kim B, Lee W-J. 2012. Draft genome sequence of *Commensalibacter intestini* A911T, a symbiotic bacterium isolated from *Drosophila melanogaster* intestine. *J Bacteriol* 194:1246. <https://doi.org/10.1128/JB.06669-11>.
- Servín-Garcidueñas LE, Sánchez-Quinto A, Martínez-Romero E. 2014. Draft genome sequence of *Commensalibacter papalotli* MX01, a symbiont identified from the guts of overwintering monarch butterflies. *Genome Announc* 2:e00128-14.
- Ellegaard KM, Engel P. 2018. New reference genome sequences for 17 bacterial strains of the honey bee gut microbiota. *Microbiol Resour Announc* 7:e00834-18.
- Quick J. 2018. Ultra-long read sequencing protocol for RAD004. <https://www.protocols.io/view/ultra-long-read-sequencing-protocol-for-rad004-mrxc57n>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- Watson M. 2018. A simple test for uncorrected insertions and deletions (indels) in bacterial genomes. *Opiniomics* <http://www.opiniomics.org/a-simple-test-for-uncorrected-insertions-and-deletions-indels-in-bacterial-genomes/>.
- 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>.