



Complete Genome Sequence of *Spiroplasma corruscae* EC-1^T (DSM 19793), a Bacterium Isolated from a Lampyrid Beetle (*Ellychnia corrusca*)

Yi-Ming Tsai (蔡乙鳴), Wen-Sui Lo (羅文穗),  Chih-Horng Kuo (郭志鴻)

Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan

ABSTRACT *Spiroplasma corruscae* EC-1^T (DSM 19793) was isolated from the gut of a lampyrid beetle (*Ellychnia corrusca*) collected in Beltsville, MD, USA, in 1983. Here, we report the complete genome sequence of this bacterium to facilitate the investigation of its biology and the comparative genomics among *Spiroplasma* species.

Spiroplasma corruscae is known to be associated with firefly beetles (Coleoptera: Lampyridae) and tabanid flies (Diptera: Tabanidae) in North America (1). EC-1^T was isolated from the gut of a lampyrid beetle (*Ellychnia corrusca*) collected in Beltsville, MD, USA, in 1983 and was designated the representative of group XIV within the genus. To facilitate future investigation on the biology of this bacterium, as well as to improve the taxon sampling of available *Spiroplasma* sequences for comparative genomics and evolutionary studies (2), we determined the complete genome sequence of *S. corruscae* EC-1^T.

The strain was acquired from the German Collection of Microorganisms and Cell Cultures (catalog number DSM 19793). The freeze-dried sample was processed according to the manufacturer's instruction and cultured in the M1D medium (3) prior to DNA extraction using the Wizard Genomic DNA purification kit (Promega, USA). PCR and Sanger sequencing were performed to verify that the 16S rRNA gene sequence matched the reference record (GenBank accession number AY189128) (4).

The procedures for genome sequencing, assembly, and annotation were based on those described in our previous studies (5–14). Briefly, we utilized the Illumina MiSeq platform to obtain 301-bp sequencing reads from one paired-end library with approximately 550-fold coverage. The initial *de novo* assembly was performed using Velvet version 1.2.10 (15). Subsequently, PAGIT version 1 (16) was used to assist an iterative process for improving the assembly. For each iteration, the raw reads were mapped to the assembly using the Burrows-Wheeler Alignment (BWA) tool version 0.7.12 (17), programmatically checked using the mpileup program in SAMtools package version 1.2 (18), and visually inspected using Integrative Genomics Viewer (IGV) version 2.3.57 (19). Polymorphic sites and gaps were corrected based on the mapped reads. The process was repeated until the complete genome sequence was obtained. The programs RNAmmer (20), tRNAscan-SE (21), and Prodigal (22) were used for gene prediction. The gene names and product descriptions were first annotated based on the homologous genes in other *Spiroplasma* genomes (5–14) as identified by OrthoMCL (23). Subsequent manual curation was based on the information obtain from the BlastKOALA tool (24) and BLASTp (25) searches against the NCBI nonredundant database (26). Putative clustered regularly interspaced short palindromic repeats (CRISPRs) were identified using CRISPRFinder (27).

Received 31 July 2017 Accepted 7 August 2017 Published 14 September 2017

Citation Tsai Y-M, Lo W-S, Kuo C-H. 2017. Complete genome sequence of *Spiroplasma corruscae* EC-1^T (DSM 19793), a bacterium isolated from a lampyrid beetle (*Ellychnia corrusca*). Genome Announc 5:e00964-17. <https://doi.org/10.1128/genomeA.00964-17>.

Copyright © 2017 Tsai 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 Chih-Horng Kuo, chk@gate.sinica.edu.tw.

The complete genome sequence of *Spiroplasma corruscae* EC-1^T consists of one circular chromosome (1,175,400 bp; 25.4% G+C) and one plasmid (29,239 bp; 23.9% G+C). The first version of annotation includes one set of 16S-23S-5S rRNA genes, 29 tRNA genes (covering all 20 amino acids), 1,039 protein-coding genes, and one pseudogene. No putative plectroviral sequence or CRISPR element was found.

Accession number(s). The complete genome sequence of *Spiroplasma corruscae* EC-1^T has been deposited at DDBJ/EMBL/GenBank under the accession numbers [CP022535](#) (chromosome) and [CP022536](#) (plasmid).

ACKNOWLEDGMENTS

The bacterial strain was imported under permit number 103-B-001 issued by the Council of Agriculture of Taiwan. The Sanger sequencing service and the Illumina sequencing library preparation service were provided by the Genomic Technology Core (Institute of Plant and Microbial Biology, Academia Sinica). The Illumina MiSeq sequencing service was provided by the DNA Sequencing Core (Institute of Molecular Biology, Academia Sinica).

The funding for this project was provided by the Institute of Plant and Microbial Biology at Academia Sinica and the Ministry of Science and Technology of Taiwan (NSC 101-2621-B-001-004-MY3 and MOST 104-2311-B-001-019) to C.-H.K. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

- Hackett KJ, Whitcomb RF, French FE, Tully JG, Gasparich GE, Rose DL, Carle P, Bové JM, Henegar RB, Clark TB, Konai M, Clark EA, Williamson DL. 1996. *Spiroplasma corruscae* sp. nov., from a firefly beetle (Coleoptera: Lampyridae) and tabanid flies (Diptera: Tabanidae). *Int J Syst Bacteriol* 46:947–950. <https://doi.org/10.1099/00207713-46-4-947>.
- Lo W-S, Huang Y-Y, Kuo C-H. 2016. Winding paths to simplicity: genome evolution in facultative insect symbionts. *FEMS Microbiol Rev* 40:855–874. <https://doi.org/10.1093/femsre/fuw028>.
- Whitcomb RF, Tully JG, McCawley P, Rose DL. 1982. Application of the growth inhibition test to *Spiroplasma* taxonomy. *Int J Syst Bacteriol* 32:387–394. <https://doi.org/10.1099/00207713-32-4-387>.
- Gasparich GE, Whitcomb RF, Dodge D, French FE, Glass J, Williamson DL. 2004. The genus *Spiroplasma* and its non-helical descendants: phylogenetic classification, correlation with phenotype and roots of the *Mycoplasma mycoides* clade. *Int J Syst Evol Microbiol* 54:893–918. <https://doi.org/10.1099/ijs.0.02688-0>.
- Lo W-S, Chen L-L, Chung W-C, Gasparich GE, Kuo C-H. 2013. Comparative genome analysis of *Spiroplasma melliferum* IPMB4A, a honeybee-associated bacterium. *BMC Genomics* 14:22. <https://doi.org/10.1186/1471-2164-14-22>.
- Ku C, Lo W-S, Chen L-L, Kuo C-H. 2013. Complete genomes of two dipteran-associated spiroplasmas provided insights into the origin, dynamics, and impacts of viral invasion in *Spiroplasma*. *Genome Biol Evol* 5:1151–1164. <https://doi.org/10.1093/gbe/evt084>.
- Lo W-S, Ku C, Chen L-L, Chang T-H, Kuo C-H. 2013. Comparison of metabolic capacities and inference of gene content evolution in mosquito-associated *Spiroplasma diminutum* and *S. taiwanense*. *Genome Biol Evol* 5:1512–1523. <https://doi.org/10.1093/gbe/evt108>.
- Ku C, Lo W-S, Chen L-L, Kuo C-H. 2014. Complete genome sequence of *Spiroplasma apis* B31^T (ATCC 33834), a bacterium associated with May disease of honeybees (*Apis mellifera*). *Genome Announc* 2(1):e01151-13. <https://doi.org/10.1128/genomeA.01151-13>.
- Chang T-H, Lo W-S, Ku C, Chen L-L, Kuo C-H. 2014. Molecular evolution of the substrate utilization strategies and putative virulence factors in mosquito-associated *Spiroplasma* species. *Genome Biol Evol* 6:500–509. <https://doi.org/10.1093/gbe/evu033>.
- Paredes JC, Herren JK, Schüpfer F, Marin R, Claverol S, Kuo C-H, Lemaître B, Béven L. 2015. Genome sequence of the *Drosophila melanogaster* male-killing *Spiroplasma* strain MSRO endosymbiont. *mBio* 6:e02437-14. <https://doi.org/10.1128/mBio.02437-14>.
- Lo W-S, Gasparich GE, Kuo C-H. 2015. Found and lost: the fates of horizontally acquired genes in arthropod-symbiotic *Spiroplasma*. *Genome Biol Evol* 7:2458–2472. <https://doi.org/10.1093/gbe/evv160>.
- Lo W-S, Lai Y-C, Lien Y-W, Wang T-H, Kuo C-H. 2015. Complete genome sequence of *Spiroplasma litorale* TN-1^T (DSM 21781), a bacterium isolated from a green-eyed horsefly (*Tabanus nigrovittatus*). *Genome Announc* 3(5):e01116-15. <https://doi.org/10.1128/genomeA.01116-15>.
- Lo W-S, Liu P-Y, Kuo C-H. 2015. Complete genome sequence of *Spiroplasma cantharicola* CC-1^T (DSM 21588), a bacterium isolated from soldier beetle (*Cantharis carolinus*). *Genome Announc* 3(5):e01253-15. <https://doi.org/10.1128/genomeA.01253-15>.
- Lo W-S, Gasparich GE, Kuo C-H. 2016. Complete genome sequence of *Spiroplasma tunicum* Tab4c^T, a bacterium isolated from horse flies (*Haematopota* sp.). *Genome Announc* 4(5):e01010-16. <https://doi.org/10.1128/genomeA.01010-16>.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
- Swain MT, Tsai IJ, Assefa SA, Newbold C, Berriman M, Otto TD. 2012. A post-assembly genome-improvement toolkit (PAGIT) to obtain annotated genomes from contigs. *Nat Protoc* 7:1260–1284. <https://doi.org/10.1038/nprot.2012.068>.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* 29:24–26. <https://doi.org/10.1038/nbt.1754>.
- Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. <https://doi.org/10.1093/nar/gkm160>.
- Lowe TM, Eddy SR. 1997. TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964.
- Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site

- identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
23. Li L, Stoeckert CJ, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13:2178–2189. <https://doi.org/10.1101/gr.1224503>.
 24. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 428:726–731. <https://doi.org/10.1016/j.jmb.2015.11.006>.
 25. 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>.
 26. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2016. GenBank. *Nucleic Acids Res* 44:D67–D72. <https://doi.org/10.1093/nar/gkv1276>.
 27. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35:W52–W57. <https://doi.org/10.1093/nar/gkm360>.