



# Complete Genome Sequence of *Spiroplasma floricola* 23-6<sup>T</sup> (ATCC 29989), a Bacterium Isolated from a Tulip Tree (*Liriodendron tulipifera* L.)

Yi-Ming Tsai (蔡乙鳴),<sup>a</sup> Pei-Shan Wu (吳佩珊),<sup>a</sup> Wen-Sui Lo (羅文穗),<sup>a</sup>  Chih-Horng Kuo (郭志鴻)<sup>a</sup>

<sup>a</sup>Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan

**ABSTRACT** *Spiroplasma floricola* 23-6<sup>T</sup> (ATCC 29989) was isolated from the flower surface of a tulip tree (*Liriodendron tulipifera* L.). Here, we report the complete genome sequence of this bacterium to facilitate the investigation of its biology and the comparative genomics among *Spiroplasma* species.

The strain *Spiroplasma floricola* 23-6<sup>T</sup> was isolated from a tulip tree (*Liriodendron tulipifera* L.) flower collected in the state of Maryland in the United States (1–3). This bacterium was shown to be pathogenic to larvae of the greater wax moth (*Galleria mellonella*) in an artificial infection experiment (4). To facilitate future investigation of the biology of this bacterium, as well as to improve the taxon sampling of available *Spiroplasma* sequences for comparative genomics and evolutionary studies (5), we determined its complete genome sequence.

The strain was acquired from the American Type Culture Collection (catalog number ATCC 29989). The freeze-dried sample was processed according to the manufacturer's instruction and cultured in M1D medium (6) 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 AY189131) (7).

The procedures for genome sequencing, assembly, and annotation were based on those described in our previous studies (8–19). Briefly, the Illumina platform was used to generate raw reads from one paired-end library (~283-bp insert, ~416-fold coverage) and one mate-pair library (~4,000-bp insert, ~476-fold coverage). The initial *de novo* assembly was performed using ALLPATHS-LG release 52188 (20). Subsequently, PAGIT version 1 (21) was used to assist an iterative process for improving the assembly. For each iteration, the raw reads were mapped to the assembly using BWA version 0.7.12 (22), programmatically checked using the mpileup program in the SAMtools package version 1.2 (23), and visually inspected using IGV version 2.3.57 (24). 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 (25), tRNAscan-SE (26), and Prodigal (27) were used for gene prediction. The gene names and product descriptions were first annotated based on the homologous genes in other *Spiroplasma* genomes (8–19) as identified by OrthoMCL (28). Subsequent manual curation was based on the information obtain from the BlastKOALA tool (29) and BLASTP (30) searches against the NCBI nonredundant database (31). Putative clustered regularly interspaced short palindromic repeats (CRISPRs) were identified using CRISPRFinder (32).

The complete genome sequence of *Spiroplasma floricola* 23-6<sup>T</sup> consists of one circular chromosome that is 1,284,130 bp in size with 24.3% G+C content; no plasmid was found. The first version of annotation includes one set of 16S-23S-5S rRNA genes, 29 tRNA genes (covering all 20 amino acids), 1,122 protein-coding genes, 8 pseudo-

Received 9 March 2018 Accepted 22 March 2018 Published 19 April 2018

**Citation** Tsai Y-M, Wu P-S, Lo W-S, Kuo C-H. 2018. Complete genome sequence of *Spiroplasma floricola* 23-6<sup>T</sup> (ATCC 29989), a bacterium isolated from a tulip tree (*Liriodendron tulipifera* L.). Genome Announc 6:e00302-18. <https://doi.org/10.1128/genomeA.00302-18>.

**Copyright** © 2018 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](mailto:chk@gate.sinica.edu.tw).

genes, and 1 CRISPR locus (chromosomal positions 565,140 to 565,547, containing five spacers).

**Accession number(s).** The complete genome sequence of *Spiroplasma floricola* 23-6<sup>T</sup> has been deposited at DDBJ/EMBL/GenBank under the accession number [CP025057](https://doi.org/10.1139/m78-158).

## ACKNOWLEDGMENTS

The Illumina paired-end sequencing service was provided by Yourgene Bioscience (New Taipei City, Taiwan). The Illumina mate-pair sequencing service was provided by the Genomic Technology Core Facility (Institute of Plant and Microbial Biology) and the Genomics Core Facility (Institute of Molecular Biology) at Academia Sinica (Taipei, Taiwan).

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

- Davis RE. 1978. *Spiroplasma* associated with flowers of the tulip tree (*Liriodendron tulipifera* L.). *Can J Microbiol* 24:954–959. <https://doi.org/10.1139/m78-158>.
- Davis RE. 1981. Antibiotic sensitivities in vitro of diverse *Spiroplasma* strains associated with plants and insects. *Appl Environ Microbiol* 41:329–333.
- Davis RE, Lee I-M, Worley JF. 1981. *Spiroplasma floricola*, a new species isolated from surfaces of flowers of the tulip tree, *Liriodendron tulipifera* L. *Int J Syst Bacteriol* 31:456–464. <https://doi.org/10.1099/00207713-31-4-456>.
- Eskafi FM, McCoy RE, Norris RC. 1987. Pathology of *Spiroplasma floricola* in *Galleria mellonella* larvae. *J Invertebr Pathol* 49:1–13. [https://doi.org/10.1016/0022-2011\(87\)90119-4](https://doi.org/10.1016/0022-2011(87)90119-4).
- 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<sup>T</sup> (ATCC 33834), a bacterium associated with May disease of honeybees (*Apis mellifera*). *Genome Announc* 2: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, Lemaitre 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<sup>T</sup> (DSM 21781), a bacterium isolated from a green-eyed horsefly (*Tabanus nigrovittatus*). *Genome Announc* 3: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<sup>T</sup> (DSM 21588), a bacterium isolated from soldier beetle (*Cantharis carolinus*). *Genome Announc* 3:e01253-15. <https://doi.org/10.1128/genomeA.01253-15>.
- Lo W-S, Gasparich GE, Kuo C-H. 2016. Complete genome sequence of *Spiroplasma turonicum* Tab4c<sup>T</sup>, a bacterium isolated from horse flies (*Haematopota* sp.). *Genome Announc* 4:e01010-16. <https://doi.org/10.1128/genomeA.01010-16>.
- Shen W-Y, Lo W-S, Lai Y-C, Kuo C-H. 2016. Complete genome sequence of *Spiroplasma helicoides* TABS-2<sup>T</sup> (DSM 22551), a bacterium isolated from a horsefly (*Tabanus abactor*). *Genome Announc* 4:e01201-16. <https://doi.org/10.1128/genomeA.01201-16>.
- Tsai Y-M, Lo W-S, Kuo C-H. 2017. Complete genome sequence of *Spiroplasma corruscae* EC-1<sup>T</sup> (DSM 19793), a bacterium isolated from a lampyrid beetle (*Ellychnia corrusca*). *Genome Announc* 5:e00964-17. <https://doi.org/10.1128/genomeA.00964-17>.
- Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A* 108:1513–1518. <https://doi.org/10.1073/pnas.1017351108>.
- 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, 1000 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, Stærfeldt 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>.

26. Lowe T, Eddy S. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964.
27. Hyatt D, Chen G-L, LoCascio P, Land M, Larimer F, Hauser L. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
28. 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>.
29. 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>.
30. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden T. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
31. 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>.
32. 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>.