Title: Draft genome sequence of *Spiroplasma platyhelix* ATCC 51748, isolated from a dragonfly

Emily A. Green [a], Jonathan L. Klassen [a]#

[a] Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut, USA

All questions and correspondence should be addressed to Jonathan L. Klassen, jonathan.klassen@uconn.edu

ABSTRACT

Spiroplasma platyhelix is a helical bacterium belonging to the class Mollicutes. First isolated from a *Pachydiplax longipennis* dragonfly, it has the smallest reported *Spiroplasma* genome size of 740 kbp. Here we report the genome sequence of *S. platyhelix* ATCC 51748.

ANNOUNCEMENT

Spiroplasma platyhelix is a species of Mollicute bacteria that was first isolated from a *Pachydiplax longipennis* dragonfly (1). It has the smallest genome size of all known *Spiroplasma* species, bacteria that infect insects, animals, and plants, acting as pathogens or mutualists (1). Bacteria related to *S. platyhelix* have been found in several fungus-growing ants (2, 3) and field crickets (NCBI: JQ768460).

S. platyhelix ATCC 51748 was purchased from the ATCC and grown in SP-4 medium (4) at 30 °C for 7 days in a 15 mL screw-top Falcon tube. Genomic DNA was extracted using the Epicentre MasterPure Complete DNA and RNA Purification kit following the DNA purification protocol for cell samples. The DNA concentration was determined using a Qubit high-sensitivity assay. Libraries were prepared from the same DNA extract for whole genome shotgun sequencing using both 2x150 bp and 2x250 bp Nextera library kits, following the manufacturer's protocol, and sequenced on an Illumina MiSeq. Adapters were removed using Trimmomatic v0.39 (5), and data were quality checked using FastQC v0.11.8 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). A Nanopore MinION long-read sequencing library was prepared from this extract following the Native barcoding genomic DNA (with EXP-NBD104, EXP-NBD114, and SQK-LSK109) protocol (vNBE_9065_v109_revV_14Aug2019). Modifications included using a NEBNext® Companion Module for Oxford Nanopore Technologies Ligation Sequencing kit for ligation, omitting

shearing or size-selection, and not using long fragment buffer during adapter ligation and clean-up. The library was sequenced on a MK1B MinION device using R9.4 flowcells and using MinKNOW v19.12.2 to run MiKNOW core v3.6.0 and Guppy v3.2.8 for basecalling, demultiplexing, and barcode trimming. Reads were assembled using SPAdes v3.13.0 (6) and annotated using Prokka (7). Genome completeness was checked against the BUSCO Tenericutes database (8), and average nucleotide identities (ANIs) were calculated using the ANI Calculator (https://www.ezbiocloud.net/tools/ani) (9). All analyses used default parameters.

In total, 332,773 and 326,607 reads were obtained for the 2x150 bp and 2x250 bp libraries, respectively, and 2,976 reads with an N50 of 2,924 for the Nanopore library. After SPAdes assembly, contaminant contigs < 500 bp were removed using QUAST v5.0.2 (10). The assembled genome is composed of 3 contigs, with an N50 of 407,984, a genome size of 739,863 bp and a 26.7% GC content. The genome is 98.7% complete according to the BUSCO analysis. Prokka annotated 1,349 protein-coding genes, one rRNA operon, and 27 tRNAs encoding for every amino acid except tryptophan (although a tryptophan ligase gene was detected). This genome size is consistent with the 740–780 kbp value previously estimated using pulsed-field gel electrophoresis (1). *S. platyhelix* has an ANI 77.3% when compared to the fungus-growing ant metagenome-assembled genome (MAG) EntAcro10 (11). The EntAcro10 MAG is 100 kbp larger than that of *S. platyhelix*, and both strains possess genes for the transport and metabolism of arginine and glycerol, consistent with mutualistic insect-microbe relationships (11). However, *S. platyhelix* lacks the CRIPSR-cas1 locus encoded by the EntAcro10 MAG. The *S. platyhelix* genome will facilitate future studies of Mollicute evolution and the mechanisms of insect and plant symbioses.

DATA AVAILABILITY

All data is available on NCBI under BioProject: PRJNA608455. Raw sequencing reads are deposited in SRA under accession numbers: SRX7849346 (2x250), SRX7849347 (2x150), and SRX7849348 (Nanopore). The assembled genome is deposited under accession number JAAVVK000000000.

AKNOWLEDGEMENTS

We thank the UConn Microbial Analysis, Resources, and Services (MARS) facility for Illumina MiSeq genome library preparations and sequencing, and Kathleen Kyle and Bailey Randolf for Nanopore library preparation and sequencing. Funding for this project was provided by NSF IOS-1656475. The funders had no role in the study design, data processing/collection or work for publication.

REFERENCES

1. Williamson DL, Adams JR, Whitcomb RF, Tully JG, Carle P, Konai M, Bové JM, Henegar RB. 1997. *Spiroplasma platyhelix* sp. nov., a new Mollicute with unusual morphology and genome size from the dragonfly *Pachydiplax longipennis*. Int J Syst Bacteriol 47:763–766.
2. Sapountzis P, Nash DR, Schiøtt M, Boomsma JJ. 2019. The evolution of abdominal microbiomes in fungus-growing ants. Mol Ecol 28:879–899.
3. Russell JA, Moreau CS, Goldman-Huertas B, Fujiwara M, Lohman DJ, Pierce NE. 2009. Bacterial gut symbionts are tightly linked with the evolution of herbivory in ants. Proc Natl Acad Sci 106:21236–21241.
4. Tully JG, Rose DL, Whitcomb RF, Wenzel RP. 1979. Enhanced isolation of *Mycoplasma pneumoniae* from throat washings with a newly modified culture medium. J Infect Dis 139:478–482.
5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.
6. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477.
7. Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.
8. Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: Assessing Genome Assembly and Annotation Completeness, p. 227–245. *In* Kollmar, M (ed.), Gene Prediction: Methods

and Protocols. Springer New York, New York, NY.

9. Yoon S-H, Ha S-m, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 110:1281–1286.

10. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

11. Sapountzis P, Zhukova M, Shik JZ, Schiøtt M, Boomsma JJ. 2018. Reconstructing the functions of endosymbiotic mollicutes in fungus-growing ants. eLife 7:e39209.