genomeA
Journals.ASM.org

Complete Genome Sequence of "*Candidatus* Sulcia muelleri" ML, an Obligate Nutritional Symbiont of Maize Leafhopper (*Dalbulus maidis*)

Hsing-Hua Chang,[a] Shu-Ting Cho,[a] Maria C. Canale,[b,c] Sam T. Mugford,[c] João R. S. Lopes,[b] Saskia A. Hogenhout,[c] Chih-Horng Kuo[a,d,e]

Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan[a]; Department of Entomology and Acarology, ESALQ/University of São Paulo, Piracicaba, SP, Brazil[b]; Department of Cell and Developmental Biology, John Innes Centre, Norwich, United Kingdom[c]; Molecular and Biological Agricultural Sciences Program, Taiwan International Graduate Program, National Chung Hsing University and Academia Sinica, Taipei, Taiwan[d]; Biotechnology Center, National Chung Hsing University, Taichung, Taiwan[e]

"*Candidatus* Sulcia muelleri" is a symbiont of sap-feeding insects in the suborder *Auchenorrhyncha*. The strain "*Ca.* Sulcia muelleri" ML is associated with the maize leafhopper (*Dalbulus maidis*), collected in Brazil, which is a disease vector that affects corn production. Here, we report the complete genome sequence of this bacterium.

Received 9 December 2014 **Accepted** 15 December 2014 **Published** 29 January 2015

Citation Chang H-H, Cho S-T, Canale MC, Mugford ST, Lopes JRS, Hogenhout SA, Kuo C-H. 2015. Complete genome sequence of "*Candidatus* Sulcia muelleri" ML, an obligate nutritional symbiont of maize leafhopper (*Dalbulus maidis*). Genome Announc 3(1):e01483-14. doi:10.1128/genomeA.01483-14.

Address correspondence to Chih-Horng Kuo, chk@gate.sinica.edu.tw.

The bacterium "*Candidatus* Sulcia muelleri" is a symbiont of sap-feeding insects in the suborder *Auchenorrhyncha* (1). This bacterium reaches high abundance inside specialized host cells (i.e., bacteriocytes) and provides its hosts with essential amino acids to supplement the nutritional deficiency of plant sap (2–7). The "*Ca.* Sulcia muelleri" genomes characterized to date have a size range of 191 to 277 kb and varied in their gene content for the biosynthesis of seven or eight essential amino acids. For the remaining two or three essential amino acids, their insect hosts have acquired lineage-specific secondary symbionts (4, 6).

To facilitate future comparative and evolutionary analyses, we determined the complete genome sequence of "*Ca.* Sulcia muelleri" strain ML from maize leafhopper (*Dalbulus maidis*). This insect is a vector for three important phytopathogens (i.e., maize bushy stunt phytoplasma, corn stunt spiroplasma, and maize rayado fino virus) that affect corn production in tropical/subtropical America (8). The colony of laboratory-reared insects used in this study was established using specimens collected in Jardinópolis (Sao Paulo State, Brazil; 20.912931 S and 47.896399 W) in 2009. The total DNA of individual insects was extracted using DNeasy blood and tissue kit (Qiagen); 39 samples were pooled for the preparation of Illumina sequencing library.

The procedure for genome sequencing, assembly, and annotation is based on that described in our previous studies (9–15). Briefly, the Illumina MiSeq platform was used to generate 251-bp reads from one paired-end library (~358 bp insert, 3,506,770 reads). The *de novo* assembly was performed using Velvet version 1.2.10 (16) with the following parameters; *k*, 141; scaffolding, no; exp_cov, auto; cov_cutoff, 10; max_coverage, 160; and min_contig_lgth, 2,000. The putative "*Ca.* Sulcia muelleri" ML contigs were identified by BLASTN (17) searches against the "*Ca.* Sulcia muelleri" ALF genome (6). The initial assembly was iteratively improved by mapping the raw reads to the contigs using the Burrows-Wheeler Aligner (BWA) version 0.6.2 (18), programmatically checked using the MPILEUP program in SAMTOOLS package version 0.1.18 (19), and visually inspected using IGV version 2.1.24 (20). All gaps were filled by using reads overhanging at the contig margins. The programs RNAmmer (21), tRNAscan-SE (22), and PRODIGAL (23) were used for gene prediction. For each protein-coding gene, the gene name and product description were initially annotated based on the homologous genes in "*Ca.* Sulcia muelleri" strains ALF (6) and GWSS (2) as identified by OrthoMCL (24). Subsequently, BLASTP (17) searches against the NCBI nonredundant (nr) protein database (25) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (26) were used to assist manual curation of the annotation.

The complete genome of "*Ca.* Sulcia muelleri" ML contains one circular chromosome that is 190,405 bp in size with a G+C content of 24.1%. The first version of annotation includes one set of 16S-23S-5S rRNA genes, 29 tRNA genes (covering all 20 amino acids), and 187 protein-coding genes.

Nucleotide sequence accession number. The complete genome sequence of "*Ca.* Sulcia muelleri" ML has been deposited at DDBJ/EMBL/GenBank under the accession number CP010105.

ACKNOWLEDGMENTS

We thank Roberto Carvalho (Dow AgroSciences Experimental Station, Jardinópolis, SP, Brazil) for collecting the leafhoppers and Willis Oliveira for help with maintenance of leafhopper colonies.

This research was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) (grants BB/K002848/1 and BB/J004553/1 to S.A.H.), the São Paulo Research Foundation (FAPESP) (grant 2012/50217-3 to J.R.S.L.), an FAPESP postdoctoral internship (grant 2013/26982-4 to M.C.C.), and the Institute of Plant and Microbial Biology at Academia Sinica and the National Science Council of Taiwan (grant NSC 101-2621-B-001-004-MY3 to C.-H.K).

REFERENCES

1. **Moran NA, Tran P, Gerardo NM.** 2005. Symbiosis and insect diversification: an ancient symbiont of sap-feeding insects from the bacterial phylum *Bacteroidetes*. Appl Environ Microbiol 71:8802–8810. http://dx.doi.org/10.1128/AEM.71.12.8802-8810.2005.

2. **McCutcheon JP, Moran NA.** 2007. Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. Proc Natl Acad Sci U S A **104:**19392–19397. http://dx.doi.org/10.1073/pnas.0708855104.
3. **McCutcheon JP, McDonald BR, Moran NA.** 2009. Convergent evolution of metabolic roles in bacterial co-symbionts of insects. Proc Natl Acad Sci U. S. A. **106:**15394–15399. http://dx.doi.org/10.1073/pnas.0906424106.
4. **McCutcheon JP, Moran NA.** 2010. Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. Genome Biol Evol **2:**708–718. http://dx.doi.org/10.1093/gbe/evq055.
5. **Woyke T, Tighe D, Mavromatis K, Clum A, Copeland A, Schackwitz W, Lapidus A, Wu D, McCutcheon JP, McDonald BR, Moran NA, Bristow J, Cheng J-F.** 2010. One bacterial cell, one complete genome. PLoS One **5:**e10314. http://dx.doi.org/10.1371/journal.pone.0010314.
6. **Bennett GM, Moran NA.** 2013. Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a phloem-feeding insect. Genome Biol Evol **5:**1675–1688. http://dx.doi.org/10.1093/gbe/evt118.
7. **Bennett GM, McCutcheon JP, MacDonald BR, Romanovicz D, Moran NA.** 2014. Differential genome evolution between companion symbionts in an insect-bacterial symbiosis. MBio **5:**e01697-14. http://dx.doi.org/10.1128/mBio.01697-14.
8. **Tsai JH.** 2008. Corn leafhopper, *Dalbulus maidis* (Delong and Wolcott) (*Hemiptera*: *Cicadellidae*), p 1072–1074. *In* Capinera JL (ed), Encyclopedia of entomology. Springer Verlag, Dordrecht, the Netherlands.
9. **Chung W-C, Chen L-L, Lo W-S, Lin C-P, Kuo C-H.** 2013. Comparative analysis of the peanut witches'-broom *Phytoplasma* genome reveals horizontal transfer of potential mobile units and effectors. PLoS One **8:**e62770. http://dx.doi.org/10.1371/journal.pone.0062770.
10. **Chung W-C, Chen L-L, Lo W-S, Kuo P-A, Tu J, Kuo C-H.** 2013. Complete genome sequence of *Serratia marcescens* WW4. Genome Announc **1**(2):e00126–13. http://dx.doi.org/10.1128/genomeA.00126-13.
11. **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. http://dx.doi.org/10.1093/gbe/evt084.
12. **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. http://dx.doi.org/10.1093/gbe/evt108.
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. http://dx.doi.org/10.1093/gbe/evu033.
14. **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:**e01151-13. http://dx.doi.org/10.1128/genomeA.01151-13.
15. **Lo W-S, Chen H, Chen C-Y, Kuo C-H.** 2014. Complete genome sequence of *Vibrio vulnificus* 93U204, a bacterium isolated from diseased tilapia in Taiwan. Genome Announc **2**(1):e01005–14. http://dx.doi.org/10.1128/genomeA.01005-14.
16. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res **18:**821–829. http://dx.doi.org/10.1101/gr.074492.107.
17. **Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.** 2009. BLAST+: architecture and applications. BMC Bioinformatics **10:**421. http://dx.doi.org/10.1186/1471-2105-10-421.
18. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. BioInformatics **25:**1754–1760. http://dx.doi.org/10.1093/bioinformatics/btp324.
19. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.** 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:**2078–2079. http://dx.doi.org/10.1093/bioinformatics/btp352.
20. **Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP.** 2011. Integrative genomics viewer. Nat Biotechnol **29:**24–26. http://dx.doi.org/10.1038/nbt.1754.
21. **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. http://dx.doi.org/10.1093/nar/gkm160.
22. **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. http://dx.doi.org/10.1093/nar/25.5.0955.
23. **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. http://dx.doi.org/10.1186/1471-2105-11-119.
24. **Li L, Stoeckert CJ, Roos DS.** 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res **13:**2178–2189. http://dx.doi.org/10.1101/gr.1224503.
25. **Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW.** 2012. GenBank. Nucleic Acids Res **40:**D48–D53. http://dx.doi.org/10.1093/nar/gkr1202.
26. **Kanehisa M, Goto S.** 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res **28:**27–30. http://dx.doi.org/10.1093/nar/28.1.27.