Draft Genome Sequence of "*Candidatus* Sulcia muelleri" Strain KPTW1 from *Kolla paulula*, a Vector of *Xylella fastidiosa* Causing Pierce's Disease of Grapevine in Taiwan

Hsien-Tzung Shih,[a] Chiou-Chu Su,[b] Chung-Jan Chang,[c] Sonia Vargas,[d] Zehan Dai,[e] Jianchi Chen[e]

[a]Applied Zoology Division, Taiwan Agricultural Research Institute, Council of Agriculture, Taiwan, Republic of China
[b]Pesticide Application Division, Taiwan Agricultural Research Institute, Council of Agriculture, Taiwan, Republic of China
[c]Department of Plant Pathology, University of Georgia, Griffin, Georgia, USA
[d]Department of Biology, University of California, Merced, California, USA
[e]United States Department of Agriculture-Agriculture Research Service, Parlier, California, USA

ABSTRACT The genome of "*Candidatus* Sulcia muelleri" strain KPTW1 from *Kolla paulula*, a vector of *Xylella fastidiosa* that causes Pierce's disease (PD) of grapevine in Taiwan, was sequenced. The strain has a genome size of 253,942 bp, GC content of 22.7%, 237 predicted protein-coding genes, and 34 RNA genes.

In 2002, Pierce's disease (PD) of grapevine caused by *Xylella fastidiosa* was found in several municipalities in Taiwan (1, 2). *Kolla paulula* was confirmed to be an insect vector of *X. fastidiosa* (3, 4). The insect occurs in all vineyards and is distributed over the medium- and low-altitude areas in Taiwan (3, 5). PD is endemic in California and the southeastern United States and is reportedly transmitted by xylem-feeding leafhoppers, such as blue-green sharpshooter (BGSS; *Graphocephala atropunctata*) and glassy winged sharpshooter (GWSS; *Homalodisca vitripennis*). Studies on BGSS and GWSS have revealed the presence of several endosymbionts, including "*Candidatus* Sulcia muelleri," which is important for insect growth and development (6–8). In this study, we report the whole-genome sequence of "*Ca.* Sulcia muelleri" in *K. paulula* from Taiwan.

A colony of *K. paulula* (female/male) was established and maintained on *Commelina diffusa* under laboratory conditions (25°C, 60% RH, 13:11 light/day [L:D]) at the Taiwan Agricultural Research Institute in 2016. DNA was extracted from a single insect using a DNeasy blood and tissue kit (Qiagen, Valencia, CA). Whole-genomic DNA was amplified using an illustra GenomiPhi version 2 DNA amplification kit (GE Healthcare, Waukesha, WI) and sequenced using MiSeq (2 × 150) format (Illumina, San Diego, CA) after preparation using the Illumina TruSeq DNA PCR-free library prep kit.

A total of 28,327,432 reads (151 bp each) with average Q value above 30 were generated. *De novo* assembly using default settings (word size = 20, bubble size = 50) was performed with CLC Genomics Workbench. The top 20 largest contigs ranging from 251,844 to 12,494 bp were collected. The top 200 bp at the 5′ end of each contig were used for BLASTn search against GenBank nr database (version 225.0) containing 21 whole-genome sequences of "*Ca.* Sulcia muelleri." Three contigs hit the whole-genome sequence of "*Ca.* Sulcia muelleri" from BGSS (GenBank accession number CP008986) (6), with 100% query coverage and 93% identity. CP008986 was used for reference mapping (length fraction = 100, similarity fraction = 90) to assemble the "*Ca.* Sulcia muelleri" genome sequence from *K. paulula* using the MiSeq data with CLC Genomics Workbench. Gap closures among reference mapping contigs were first made based on information in *de novo* contigs, followed by read walking/extension. Briefly, 100 Ns were added at either end of a targeted contig. The contig was used for reference mapping with MiSeq reads (length fraction = 50, identity fraction = 100). Overlapping

reads (50 bp minimum) were identified and added for contig extension. The extended contig was used to search for overlapping with other long contigs for extension if found. The read walking procedure continued until no overlapped reads were found.

The draft genome of "*Ca.* Sulcia muelleri," designated strain KPTW1, consists of 9 contigs from 1,695 to 137,380 bp, with a total size of 253,942 bp (average nucleotide coverage, 64×) and a GC content of 22.7%. Sequence annotation was performed using the RAST server (http://rast.nmpdr.org/) (9). The draft genome was predicted to have 237 protein-coding genes and 34 RNA genes. The KPTW1 genome is 96.4% of that of the reference BGSS (244,618 bp).

Data availability. This complete genome sequence has been deposited in DDBJ/EMBL/GenBank under the accession number QWZP00000000. The MiSeq read data are available in SRA under the accession number SRR8245541. The version described in this paper is the first version, QWZP01000000.

REFERENCES

1. Su CC, Chang CM, Chang CJ, Su WY, Deng WL, Shih HT. 2013. Occurrence of Pierce's disease of grapevines and its control strategies in Taiwan. Plant Pathol Bull 22:245–258.
2. Su CC, Chang CJ, Chang CM, Shih HT, Tzeng KC, Jan FJ, Kao CW, Deng WL. 2013. Pierce's disease of grapevines in Taiwan: isolation, cultivation and pathogenicity of *Xylella fastidiosa*. J Phytopathol 161:389–396. https://doi.org/10.1111/jph.12075.
3. Shih HT, Wen YD, Fanjian CC, Chang CJ, Chang CM, Lee CY, Yao MH, Chang SC, Jan FJ, Su CC. 2013. Potential vectors of Pierce's disease in Taiwan: ecology and integrated management, p 163–175. *In* Chang CJ, Lee CY, Shih HT (ed), Proceedings of the 2013 International Symposium on Insect Vectors and Insect-borne diseases. Special Publication of TARI, Taiwan, ROC, no. 173.
4. Tuan SJ, Hu FT, Chang HY, Chang PW, Chen YH, Huang TP. 2016. *Xylella fastidiosa* transmission and life history of two Cicadellinae sharpshooters, *Kolla paulula* and *Bothrogonia ferruginea* (Hemiptera: Cicadellidae), in Taiwan. J Econ Entomol 109:1034–1040. https://doi.org/10.1093/jee/tow016.
5. Shih HT, Su CC, Feng CY, Fanjian CC, Hung WF, Huang LY. 2009. Studies on the morphology, ecology, and host range for *Kolla paulula* (Walker,

1858) Hemiptera: Membracoidea: Cicadellidae: Cicadellinae. Formos Entomol 29:353.
6. 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. https://doi.org/10.1128/mBio.01697-14.
7. McCutcheon JP, Moran NA. 2007. Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. Proc Natl Acad Sci USA 104:19392–19397. https://doi.org/10.1073/pnas.0708855104.
8. 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. https://doi.org/10.1128/AEM.71.12.8802-8810.2005.
9. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.