Draft Genome Sequence of “Candidatus Hamiltonella defensa,” an Endosymbiont of the Whitefly *Bemisia tabaci*

Qiong Rao,a Shuang Wang,b Yun-Lin Su,a Xiao-Li Bing,a Shu-Sheng Liua and Xiao-Wei Wang*b

Ministry of Agriculture Key Laboratory of Agricultural Entomology, Institute of Insect Sciences, Zhejiang University, Hangzhou, China,* and BGI-Shenzhen, Shenzhen, China*b

“Candidatus Hamiltonella defensa” is a facultative endosymbiont of the whitefly *Bemisia tabaci*. Herein, we report the first draft genome sequence of “Candidatus Hamiltonella defensa” from the invasive Mediterranean cryptic species of the *B. tabaci* complex. The 1.84-Mbp genome sequence comprises 404 contigs and contains 1,806 predicted protein-coding genes.

Uncultured symbionts were isolated from the Mediterranean cryptic species of *B. tabaci*, and the whole genome of purified bacteria was amplified by multiple displacement amplification (MDA) using a Repli-g ultrafast minikit (Qiagen). Two paired-end libraries with average insert sizes of 200 bp and 2 kb were generated. Amplified DNA was sequenced using an Illumina HiSeq 2000 instrument at BGI-Shenzhen (Shenzhen, China). We obtained 240 Mb and 84 Mb of high-quality useful data from each library. The paired-end reads were de novo assembled using SOAPdenovo v1.05; and contigs were connected according to the 2-kb paired-end relationships. The coding sequences (CDSs) were predicted by Glimmer v3.0, and homologous comparison to the NCBI nonredundant database was performed by BLAST for function annotation. The rRNA and tRNA were identified using rRNAmmer and tRNAscan.

The results show that the draft genome of the *B. tabaci* symbiont “Candidatus Hamiltonella defensa” has approximately 1.84 Mbp with a G+C content of 40.3% and comprises 404 contigs (N50 11.8 kb), which can be assembled into 372 scaffolds (N50 14.2 kb).

A total of 1,806 genes and 38 tRNA genes are predicted within the genome of “Candidatus Hamiltonella defensa.” A total of 1,112 genes were mapped to pathways in the Kyoto Encyclopedia of Genes and Genomes, among which “Membrane Transport” (n = 301) has the highest percentage of genes, followed by “Replication and Repair” (n = 227), “Translation” (n = 136), “Carbohydrate Metabolism” (n = 128), “Metabolism of Cofactors and Vitamins” (n = 88), and “Amino Acid Metabolism” (n = 87). These annotations provide a valuable resource for investigation of the functions of “Candidatus Hamiltonella defensa” in the whitefly. Further genomic analyses and comparative studies will provide additional insight into this symbiont-host system.

**Nucleotide sequence accession numbers.** This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number AJLH00000000. The version described in this paper is the first version, AJLH01000000.

**ACKNOWLEDGMENTS**

This work was supported by grants from the National Basic Research and Development Program of China (project 2009CB119203) and the National Natural Science Foundation of China (projects 31171848, 31021003, and 31101437).

**REFERENCES**