JB
Journals.ASM.org

Genome Sequences of the Primary Endosymbiont "*Candidatus* Portiera aleyrodidarum" in the Whitefly *Bemisia tabaci* B and Q Biotypes

Zi-Feng Jiang,[a] Fangfang Xia,[b] Kipp W. Johnson,[a] Elizabeth Bartom,[c] Jigyasa H. Tuteja,[a] Rick Stevens,[a,b] Robert L. Grossman,[a] Marina Brumin,[d] Kevin P. White,[a] and Murad Ghanim[d]

Institute for Genomics and Systems Biology, The University of Chicago, Chicago, Illinois, USA[a]; Argonne National Laboratory, Argonne, Illinois, USA[b]; Center for Research Informatics, The University of Chicago, Chicago, Illinois, USA[c]; and The Agricultural Research Organization (ARO), Department of Entomology, Volcani Center, Bet Dagan, Israel[d]

"*Candidatus* Portiera aleyrodidarum" is the obligate primary endosymbiotic bacterium of whiteflies, including the sweet potato whitefly *Bemisia tabaci*, and provides essential nutrients to its host. Here we report two complete genome sequences of this bacterium from the B and Q biotypes of *B. tabaci*.

"*Candidatus* Portiera aleyrodidarum" is the obligate primary endosymbiotic bacterium hosted by whiteflies, including the sweet potato whitefly *Bemisia tabaci*, one of the most globally damaging insect pests in open fields and protected agricultural crops, causing annual losses estimated at 1 to 2 billion dollars. *B. tabaci* is one of the top 100 invasive species worldwide (5). Similar to other obligate bacteria living in sap-sucking insects, "*Ca*. Portiera aleyrodidarum" is thought to provide essential nutrients to whiteflies (2). *B. tabaci* is a species complex comprised of 11 high-level genetic groups that are well defined by DNA markers and at least 24 morphologically indistinguishable species (1, 4). The most predominant and damaging biotypes are B and Q, which differ considerably with regard to various fitness parameters: while B is defined by high fecundity and a wide host range, Q is known to develop higher resistance to insecticides (7). Despite these significant differences in their biology, nothing is known about the comparative genome sequences of the B and Q primary endosymbionts. Here we report the assembly and comparison of "*Ca*. Portiera aleyrodidarum" genomes from these two biotypes.

B and Q biotype whiteflies were collected in Israel, and each strain was interbred over 200 generations under standard laboratory conditions. High-quality DNA (>40 kb) was extracted from whole bodies of multiple pooled insect adults of each species. Paired-end libraries with 180-bp insertions and 5-kb mate-pair libraries were constructed and sequenced. Adaptors and low-quality reads were removed before assembly. ALLPATHS-LG (3) and Velvet (8) were used for *de novo* assembly. Based on the depth of coverage, paired-end connections (180 ± 10 bp), and 5-kb mate-pair connections, "*Ca*. Portiera aleyrodidarum" contigs were selected and linked into a closed circular molecule by customized scripts. We observed a limited number of short repeats/homopolymers that appear polymorphic and have taken the major allele in these cases as the consensus sequence. Over 1 million reads were mapped to each of the final assemblies. Assembly accuracy and the coverage (at least 25× coverage for each nucleotide) were manually curated. Annotation was done on the RAST annotation server (6).

The "*Ca*. Portiera aleyrodidarum" genome assembly is a 351-kb circular molecule. The "*Ca*. Portiera aleyrodidarum" genome from whitefly B (here referred to as WB) and the "*Ca*. Portiera aleyrodidarum" genome from whitefly Q (here referred to as WQ) share high

similarity in sequence (99%). Both genomes are highly AT biased (73%) and have 36 RNA-coding genes. A total of 277 protein-coding genes were identified in the "*Ca*. Portiera aleyrodidarum" WB genome, and 281 were identified in the "*Ca*. Portiera aleyrodidarum" WQ genome. Similar to other genome-sequenced obligate endosymbionts from sap-feeding insects, enrichment for genes involved in essential amino acid biosynthesis was observed, while genes involved in processes such as membrane transport, cell wall/capsule, or motility were not found. These results support the hypotheses that primary endosymbionts primarily supply their hosts by amino acids. Comparing "*Ca*. Portiera aleyrodidarum" genome sequences from the B and Q biotypes showed a number of regions that have diverged between the two genomes. Such divergence might play a differential role in "*Ca*. Portiera aleyrodidarum" interactions with their whitefly host.

Nucleotide sequence accession numbers. The complete chromosome sequences have been deposited in the NCBI GenBank database under accession numbers CP003867 (for "*Ca*. Portiera aleyrodidarum" WQ) and CP003868 (for "*Ca*. Portiera aleyrodidarum" WB).

ACKNOWLEDGMENTS

REFERENCES

1. **Alemandri V, et al.** 2012. Species within the *Bemisia tabaci* (Hemiptera: Aleyrodidae) complex in soybean and bean crops in Argentina. J. Econ. Entomol. **105**:48–53.
2. **Baumann P.** 2005. Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects. Annu. Rev. Microbiol. **59**:155–189.
3. **Gnerre S, et al.** 2011. High-quality draft assemblies of mammalian ge-

nomes from massively parallel sequence data. Proc. Natl. Acad. Sci. U. S. A. **108**:1513–1518.

4. **Hu J, et al.** 2011. An extensive field survey combined with a phylogenetic analysis reveals rapid and widespread invasion of two alien whiteflies in China. PLoS One **6**:e16061.

5. **Kontsedalova S, et al.** 2012. *Bemisia tabaci* biotype dynamics and resistance to insecticides in Israel during the years 2008-2010. J. Integr. Agric. **11**:312–320.

6. **Meyer F, et al.** 2008. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics **9**:386.

7. **Skaljac M, et al.** 2012. Diversity and localization of bacterial symbionts in three whitefly species (Hemiptera: Aleyrodidae) from the east coast of the Adriatic Sea. Bull. Entomol. Res. **15**:1–12.

8. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.