Draft Genome Assembly of a *Wolbachia* Endosymbiont of *Plutella australiana*

Christopher M. Ward, Simon W. Baxter

School of Biological Sciences, University of Adelaide, Adelaide, Australia

ABSTRACT *Wolbachia* spp. are endosymbiotic bacteria that infect around 50% of arthropods and cause a broad range of effects, including manipulating host reproduction. Here, we present the annotated draft genome assembly of *Wolbachia* strain wAus, which infects *Plutella australiana*, a cryptic ally of the major *Brassica* pest *Plutella xylostella* (diamondback moth).

Address correspondence to Simon W. Baxter, simon.baxter@adelaide.edu.au.

Plutella australiana (Lepidoptera: Plutellidae) is endemic to Australia and morphologically cryptic with the worldwide *Brassica* pest, *Plutella xylostella* (1, 2). Despite the ability to hybridize under laboratory conditions, substantial variation across several traits has been documented between these two species, including prevalence of *Wolbachia* infection (3). *Wolbachia* spp. are a diverse group of intracellular bacteria that infect arthropods and nematodes and often act as reproductive parasites on hosts to promote their own transmission (4). Infecting mosquitos with specific *Wolbachia* strains has been used as a nontraditional method for blocking vector-borne diseases, demonstrating useful applications for these symbionts (5). In Australia, *Wolbachia* infection occurs in only 1.5% of *P. xylostella* moths, yet appears fixed among *P. australiana*. Whole-genome short-read sequencing of a whole *P. australiana* male moth facilitated the identification and genome assembly of a *Wolbachia* endosymbiont we named wAus.

Plutella australiana paired-end short reads (2 × 150 bp) were mapped to the wPip (GenBank accession no. NC_010981) and wMel (GenBank accession no. NC_002978) reference genomes using BWA-MEM (6) to separate putative *Wolbachia* reads from the host and other contaminants. The two resulting BAM files were converted to fastq using BEDTools (7) and concatenated, and duplicate sequences removed. This recovered 1,119,295 reads, of which 1,081,300 (96.61% of total reads) were properly paired (mapQ ≥ 5). The concatenated paired-end short reads were then assembled using Velvet version 1.2.10 (8), with a k-mer of 65. The wAus draft assembly has a total length of 1,158,805 bp across 95 contigs (N_{50} value, 19,935 bp), the largest of which is 72,415 bp. The NCBI Prokaryotic Genome Annotation Pipeline (https://www.ncbi.nlm.nih.gov/genome/annotation_prok/) identified annotations for 1,040 protein-coding genes, 43 pseudogenes, 34 tRNAs, 4 noncoding RNAs (ncRNAs), and 3 rRNAs (5S, 16S, and 23S). To test for non-*Wolbachia* bacterial sequence contamination in the assembly, contigs were divided into 1-kbp fragments and queried against a Kraken database (9) built from all complete bacterial references in RefSeq (https://www.ncbi.nlm.nih.gov/refseq/). Most fragments were classified as *Wolbachia* (98.7%), yet 1.3% reported no bacterial homology and were subsequently subjected to a BLAST search against the NCBI Genome database (https://www.ncbi.nlm.nih.gov/genome/) using Geneious version 10.1.3 (10). This failed to match known sequences, suggesting these regions may be specific to wAus.

Phylogenetic reconstruction using multilocus sequence typing genes (*coxA*, *gatB*, *ftsZ*, *fbpA*, and *hcpA*) (11) placed wAus into supergroup B, with 100% bootstrap support. Based on these five genes, wAus was most similar to the *Wolbachia* endosymbiont

of *Culex quinquefasciatus* (wPip) and significantly different from other *Wolbachia* known to infect *Plutella* species (12). Recently, two genes causing cytoplasmic incompatibility in the *Wolbachia* strain wMel were identified as *cifA* (WD0631) and *cifB* (WD0632) (13); however, orthologs were absent from the wAus assembly and the *P. australiana* genome. Nevertheless, this draft genome provides an opportunity to investigate reproductive phenotypes associated with wAus infection, which may have future applications for biological control.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. MRWX00000000. The version described in this paper is the first version.

REFERENCES

1. Furlong MJ, Wright DJ, Dosdall LM. 2013. Diamondback moth ecology and management: problems, progress, and prospects. Annu Rev Entomol 58:517–541. https://doi.org/10.1146/annurev-ento-120811-153605.
2. Landry J-F, Hebert P. 2013. *Plutella australiana* [Lepidoptera, Plutellidae], an overlooked diamondback moth revealed by DNA barcodes. ZooKeys 327:43–63.
3. Perry KD, Baker GJ, Powis KJ, Kent JK, Ward CM, Baxter SW. 2017. A global *Brassica* pest and a sympatric cryptic ally, *Plutella australiana* (Lepidoptera: Plutellidae), show strong divergence despite the capacity to hybridize. bioRxiv. https://doi.org/10.1101/189266.
4. Werren JH, Baldo L, Clark ME. 2008. *Wolbachia*: master manipulators of invertebrate biology. Nat Rev Microbiol 6:741–751. https://doi.org/10.1038/nrmicro1969.
5. Schmidt TL, Barton NH, Rašić G, Turley AP, Montgomery BL, Iturbe-Ormaetxe I, Cook PE, Ryan PA, Ritchie SA, Hoffmann AA, O'Neill SL, Turelli M. 2017. Local introduction and heterogeneous spatial spread of dengue-suppressing *Wolbachia* through an urban population of *Aedes aegypti*. PLoS Biol 15:e2001894. https://doi.org/10.1371/journal.pbio.2001894.
6. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv arXiv:1303.3997v2.
7. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033.
8. Zerbino DR. 2010. Using the Velvet *de novo* assembler for short-read sequencing technologies. Curr Protoc Bioinformatics Chapter 11:Unit 11.5. https://doi.org/10.1002/0471250953.bi1105s31.
9. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol 15:R46. https://doi.org/10.1186/gb-2014-15-3-r46.
10. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199.
11. Baldo L, Dunning Hotopp JC, Jolley KA, Bordenstein SR, Biber SA, Choudhury RR, Hayashi C, Maiden MCJ, Tettelin H, Werren JH. 2006. Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. Appl Environ Microbiol 72:7098–7110. https://doi.org/10.1128/AEM.00731-06.
12. Delgado AM, Cook JM. 2009. Effects of a sex-ratio distorting endosymbiont on mtDNA variation in a global insect pest. BMC Evol Biol 9:49. https://doi.org/10.1186/1471-2148-9-49.
13. LePage DP, Metcalf JA, Bordenstein SR, On J, Perlmutter JI, Shropshire JD, Layton EM, Funkhouser-Jones LJ, Beckmann JF, Bordenstein SR. 2017. Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. Nature 543:243–247. https://doi.org/10.1038/nature21391.