



Near-Complete Genome Sequences of a *Wolbachia* Strain Isolated from *Diaphorina citri* Kuwayama (*Hemiptera: Liviidae*)

 Surendra Neupane,^a Sylvia I. Bonilla,^a Andrew M. Manalo,^a  Kirsten S. Pelz-Stelinski^a

^aEntomology and Nematology Department, Citrus Research and Education Center, University of Florida, Lake Alfred, Florida, USA

Surendra Neupane and Sylvia I. Bonilla contributed equally to this article. Author order was determined randomly.

ABSTRACT *Wolbachia* strains are one of three endosymbionts associated with the insect vector of “*Candidatus Liberibacter asiaticus*,” *Diaphorina citri* Kuwayama (*Hemiptera: Liviidae*). We report three near-complete genome sequences of samples of *Wolbachia* from *D. citri* (wDi), with sizes of 1,518,595, 1,542,468, and 1,538,523 bp.

The widespread endosymbiont *Wolbachia* is present in many insect species and can manipulate host reproduction via cytoplasmic incompatibility, male killing, and induction of parthenogenesis or feminization (1). *Diaphorina citri* Kuwayama (*Hemiptera: Liviidae*), the vector of the pathogen “*Candidatus Liberibacter asiaticus*,” associated with citrus greening disease, harbors three intracellular symbionts, including a strain of *Wolbachia*, wDi. Currently, the functional relationship of *Wolbachia* sp. strain wDi with *D. citri* and “*Ca. Liberibacter asiaticus*” is limited due to the unavailability of a genome assembly with no gaps. Here, we report three wDi genome sequences utilizing both long- and short-read sequencing methods. *Wolbachia* isolates were recovered from *D. citri* from an established laboratory culture collected in Polk County (28.0°N, 81.9°W; Lake Alfred, FL, USA). Individual psyllids were placed on sterile diet rings for 2 days prior to *Wolbachia* extraction. The psyllids were surface sterilized and immersed in 1.0 ml of Schneider’s *Drosophila* (S2) medium (catalog number 21720024; Gibco). Next, individual psyllids were homogenized and centrifuged at 100 × *g* for 5 min. The supernatants were collected and centrifuged at 400 × *g* for 5 min. The pellets were resuspended with 1.0 ml of S2 medium. The samples were centrifuged at 100 × *g* for 5 min; then, supernatants were placed in new tubes and centrifuged at 4,000 × *g* for 5 min. The pellets were resuspended in 1.0 ml of S2 medium. After isolation, an individual wDi sample was inoculated into *Drosophila* S2 cells (catalog number R69007; Gibco) and maintained in S2 medium containing 10% heat-inactivated fetal bovine serum (catalog number 16000044; Gibco), 50 units of penicillin, and 50 μg streptomycin sulfate (catalog number 15070-063; Gibco) per ml (S2 complete medium) using the method described by Dobson et al. (2). wDi-infected S2 cells were maintained using standard protocols (3). *Wolbachia* isolates were purified from a parental cell line at different passages (28, 36, and 40) to detect the generic variation of the bacteria in the host, using the protocol adopted from Rasgon et al. (4). The genomic DNA from three wDi samples were extracted using the MagAttract high-molecular-weight (HMW) DNA minikit (catalog number 67563) using the manufacturer’s protocol with little modification. The extracted DNA samples from passages 28 and 40 were cleaned up using the DNeasy PowerClean cleanup kit (catalog number 12877-50). The DNA sample from passage 36 was left unpurified. The extracted DNA was quantified using the Qubit 1 × double-stranded DNA (dsDNA) high-sensitivity (HS) assay kit (Thermo Fisher Scientific), and the quality was measured using the TapeStation genomic DNA ScreenTape (Agilent). For PacBio sequencing, a library was prepared using single-molecule real-time (SMRT) sequencing in conjunction with barcodes from the barcode adaptor kit 8A (PacBio

Citation Neupane S, Bonilla SI, Manalo AM, Pelz-Stelinski KS. 2020. Near-complete genome sequences of a *Wolbachia* strain isolated from *Diaphorina citri* Kuwayama (*Hemiptera: Liviidae*). *Microbiol Resour Announc* 9:e00560-20. <https://doi.org/10.1128/MRA.00560-20>.

Editor Irene L. G. Newton, Indiana University, Bloomington

Copyright © 2020 Neupane et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Kirsten S. Pelz-Stelinski, pelzstelinski@ufl.edu.

Received 15 May 2020

Accepted 31 July 2020

Published 27 August 2020

TABLE 1 Assembly statistics and general genome information of three wDi genome sequences

Parameter	Data for isolate:		
	wDi (passage 28)	wDi (passage 36)	wDi (passage 40)
Sequencing instrument	PacBio Sequel II	PacBio Sequel II	PacBio Sequel II
No. of polymerase reads	79,755	113,497	130,351
No. of subreads	868,919	1,298,934	1,534,487
No. of bases	4,421,085,394	6,263,217,203	7,244,262,525
Mean read length (bp)	56,224	56,019	56,435
Longest subread length (bp)	182,617	127,315	162,373
No. of CCS bases	186,807,190	257,167,133	294,137,324
No. of CCS reads	35,203	51,403	59,780
CCS coverage (×)	124.53	171.44	196.09
Assembled chromosome length (bp)	1,518,595	1,542,468	1,538,523
G+C content (%)	34.06	33.95	33.92
No. of genes (total)	1,415	1,428	1,409
No. of CDSs ^a (total)	1,376	1,387	1,368
No. of CDSs ^a (with protein)	1,251	1,249	1,238
No. of RNAs	39	41	41
No. of rRNAs	3	3	3
No. of tRNAs	32	34	34
No. of ncRNAs ^b	4	4	4
No. of pseudogenes (total)	125	138	130
CheckM completeness (%)	95.7	100	100
CheckM contamination (%)	0.43	0.43	0.43
PacBio accession no.	SRR11513698	SRR11513697	SRR11513696
Illumina accession no.	SRR11513759	SRR11513758	SRR11513757
GenBank accession no.	CP051266	CP051265	CP051264

^a CDSs, coding DNA sequences.

^b ncRNAs, noncoding RNAs.

PN 101-081-300). Sequencing of wDi was done using a PacBio SEQUEL instrument using chemistry v3.0 reagents in combination with the SMRTLink v6.0 software. For Illumina sequencing, the libraries were prepared using the NEBNext Ultra II DNA library prep kit (New England BioLabs, Ipswich, MA, USA), and genomic sequencing was performed using a MiSeq instrument (Illumina, San Diego, CA) with the 2 × 300-bp paired-end protocol using the 600-cycle v3 kit (Illumina).

The circular consensus sequences (CCS) generated from the PacBio raw reads were subjected to quality assessment and adaptor trimming using seqtk (<https://github.com/lh3/seqtk>) and Filtrlong (<https://github.com/rwick/Filtrlong>). *De novo* assembly was performed separately using Canu v1.9 (5), Unicycler v0.4.8 (6), and the hierarchical genome assembly process (HGAP) (7). These assemblers produced different-sized contigs, which were subassembled using the subassemblies parameter in Flye v2.7.1 (8) to produce the longest contig (wDi from passages 28 and 36). Canu produced the best assembly for wDi from passage 40; however, the genome sequences could not be circularized using the available tools. Paired-end Illumina reads were cleaned with Trimmomatic v0.36 (9) to remove adapters, and sequences with a Phred score of less than 30 were discarded. The trimmed Illumina reads were mapped to the corresponding PacBio chromosome using bwa v0.7.17 (10) using paired-end mode. Indel correction and polishing were done using Pilon v1.22 (11). The Pilon analyses corrected 13 single nucleotide polymorphisms (SNPs), 66 small insertions totaling 74 bases, and 1 small deletion totaling 1 base in wDi from passage 28 and 206 SNPs and 103 small insertions totaling 135 bases and 32 small deletions totaling 97 bases in wDi from passage 36. The Pilon tool did not correct any bases in wDi from passage 40. Genome annotation was conducted using the standard NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (12). Assembly statistics and general genome information for the three wDi genome sequences are represented in Table 1. Mauve (13) was used for wDi genome alignment with local colinear blocks (LCBs). Ten LCBs on each chromosome were identified and joined by connecting lines in the three genome sequences (Fig. 1).

Benchmarking Universal Single-Copy Orthologs (BUSCO) v4 using the bacteria_obd10

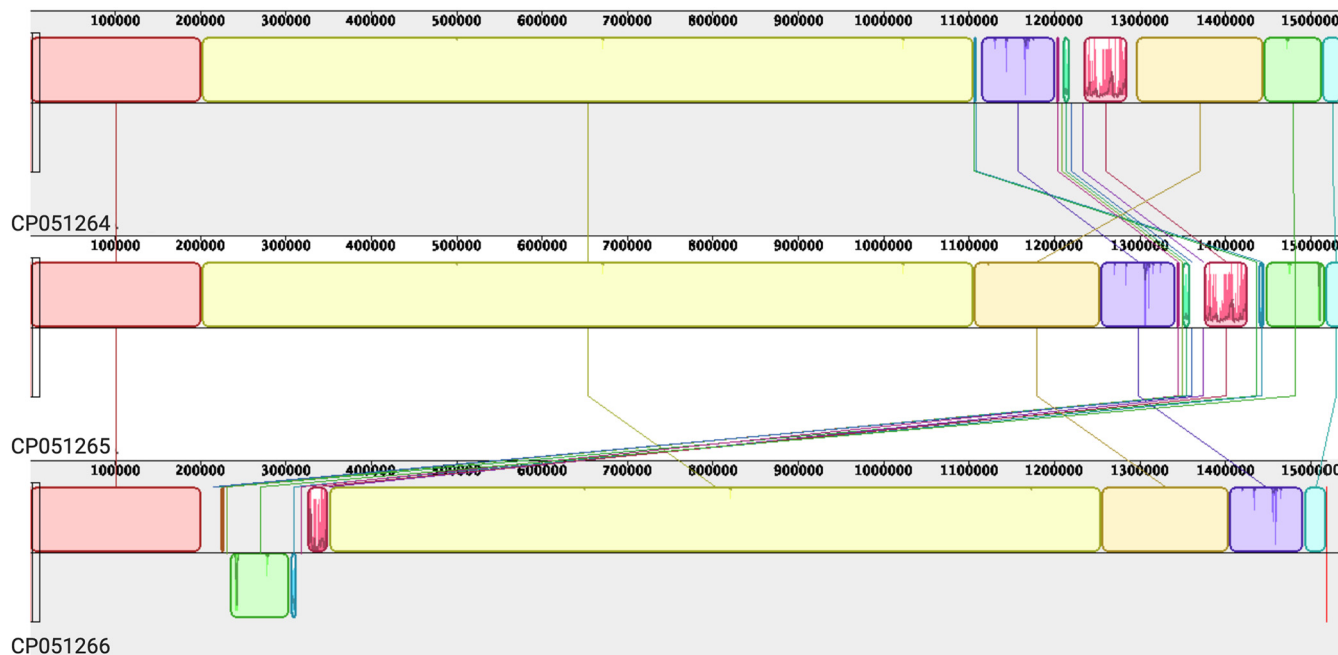


FIG 1 Mauve alignments showing chromosomal similarity among the genome assemblies under GenBank accession numbers CP051264, CP051265, and CP051266. Ten local colinear blocks (LCBs) on the chromosomes were identified and joined by connecting lines in the three genomes. Two LCBs in the assembly under accession number CP051266 are inverted, which shows reverse complement orientation.

database (14) was utilized to assess the completeness of the genomes. The completeness rates of the final assembled genomes were 80.6% (wDi sample from passage 28) and 85.5% (wDi samples from passages 36 and 40), compared with those of the reference *Wolbachia* genome sequences, wAlbB (from *Aedes albopictus*) (84.7%) (15), wMel (from *Drosophila melanogaster*) (83.1%) (16), and *Wolbachia* sp. strain wPip (from *Culex pipiens*) (86.3%) (17). In addition, CheckM analysis (18) was done to assess the completeness and contamination using 63 genomes and 368 lineage-specific markers via the Microscope platform (19) (Table 1). The genome assemblies were examined for prophage sequences using PHAGE Search Tool Enhanced Release (PHASTER; <https://phaster.ca/>) (20). PHASTER predicted variations in the three assemblies, with seven regions (range, 11.7 kb to 31.5 kb) in wDi from passage 28, six regions (range, 11.8 kb to 35.7 kb) in wDi from passage 36, and five regions (range, 11.8 kb to 35.3 kb) in wDi from passage 40.

Data availability. The three genome sequences of wDi have been deposited in GenBank under the accession numbers CP051266, CP051265, and CP051264 corresponding to BioSample accession numbers SAMN14560310, SAMN14560311, and SAMN14560312, respectively, under the BioProject accession number PRJNA603775.

ACKNOWLEDGMENTS

We thank Albert Mangual for maintaining the cell cultures and isolating wDi. We acknowledge the team at the University of Florida Interdisciplinary Center for Biotechnology Research (UF-ICBR) NextGen DNA Sequencing Center, in particular David Moraga, scientific director, for his inputs.

The funding for this work was provided to K.S.P.-S. by the Defense Advanced Research Projects Agency (DARPA), USA (award AWD05653).

REFERENCES

1. 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>.
2. Dobson SL, Marsland EJ, Veneti Z, Bourtzis K, O'Neill SL. 2002. Characterization of *Wolbachia* host cell range via the in vitro establishment of infections. *Appl Environ Microbiol* 68:656–660. <https://doi.org/10.1128/aem.68.2.656-660.2002>.
3. Baum B, Chervas L. 2008. *Drosophila* cell lines as model systems and as

- an experimental tool. *Methods Mol Biol* 2008;420:391–424. https://doi.org/10.1007/978-1-59745-583-1_25.
4. Rasgon JL, Gamston CE, Ren X. 2006. Survival of *Wolbachia pipientis* in cell-free medium. *Appl Environ Microbiol* 72:6934–6937. <https://doi.org/10.1128/AEM.01673-06>.
 5. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
 6. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
 7. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
 8. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
 9. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
 10. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* 26:589–595. <https://doi.org/10.1093/bioinformatics/btp698>.
 11. Walker BJ, Abeel T, Shea T, Priest M, Abuelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
 12. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44: 6614–6624. <https://doi.org/10.1093/nar/gkw569>.
 13. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <https://doi.org/10.1101/gr.2289704>.
 14. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.
 15. Sinha A, Li Z, Sun L, Carlow CKS. 2019. Complete genome sequence of the *Wolbachia* wAlbB endosymbiont of *Aedes albopictus*. *Genome Biol Evol* 11:706–720. <https://doi.org/10.1093/gbe/evz025>.
 16. Wu M, Sun LV, Vamathevan J, Riegler M, Deboy R, Brownlie JC, McGraw EA, Martin W, Esser C, Ahmadinejad N, Wiegand C, Madupu R, Beanan MJ, Brinkac LM, Daugherty SC, Durkin AS, Kolonay JF, Nelson WC, Mohamoud Y, Lee P, Berry K, Young MB, Utterback T, Weidman J, Nierman WC, Paulsen IT, Nelson KE, Tettelin H, O'Neill SL, Eisen JA. 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. *PLoS Biol* 2:E69. <https://doi.org/10.1371/journal.pbio.0020069>.
 17. Klasson L, Walker T, Sebahia M, Sanders MJ, Quail MA, Lord A, Sanders S, Earl J, O'Neill SL, Thomson N, Sinkins SP, Parkhill J. 2008. Genome evolution of *Wolbachia* strain wPip from the *Culex pipiens* group. *Mol Biol Evol* 25:1877–1887. <https://doi.org/10.1093/molbev/msn133>.
 18. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
 19. Vallenet D, Calteau A, Dubois M, Amours P, Bazin A, Beuvin M, Burlot L, Bussell X, Foueteau S, Gautreau G, Lajus A, Langlois J, Planel R, Roche D, Rollin J, Rouy Z, Sabatet V, Médigue C. 2020. MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Res* 48:D579–D589. <https://doi.org/10.1093/nar/gkz926>.
 20. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21. <https://doi.org/10.1093/nar/gkw387>.