








# Complete Genome Sequence of *wBp*, the *Wolbachia* Endosymbiont of *Brugia pahangi* FR3

 Jarrett F. Lebov,<sup>a</sup> John Mattick,<sup>a</sup> Silvia Libro,<sup>b</sup> Benjamin C. Sparklin,<sup>a</sup>  Matthew Chung,<sup>a</sup>  Robin E. Bromley,<sup>a</sup> Suvarna Nadendla,<sup>a</sup> Xuechu Zhao,<sup>a</sup> Sandra Ott,<sup>a</sup> Lisa Sadzewicz,<sup>a</sup> Luke J. Tallon,<sup>a</sup> Michelle L. Michalski,<sup>c</sup>  Jeremy M. Foster,<sup>b</sup>  Julie C. Dunning Hotopp<sup>a,d,e</sup>

<sup>a</sup>Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, USA

<sup>b</sup>Division of Protein Expression and Modification, New England Biolabs, Ipswich, Massachusetts, USA

<sup>c</sup>Department of Biology and Microbiology, University of Wisconsin Oshkosh, Oshkosh, Wisconsin, USA

<sup>d</sup>Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, Maryland, USA

<sup>e</sup>Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, Maryland, USA

**ABSTRACT** Lymphatic filariasis is a devastating disease caused by filarial nematode roundworms, which contain obligate *Wolbachia* endosymbionts. Here, we assembled the genome of *wBp*, the *Wolbachia* endosymbiont of the filarial nematode *Brugia pahangi*, from Illumina, Pacific Biosciences, and Oxford Nanopore data. The complete, circular genome is 1,072,967 bp.

**B***rugia pahangi* is a zoonotic filarial nematode and sister taxon to *Brugia malayi*, a causative agent of human lymphatic filariasis. We sequenced the *Wolbachia* endosymbiont *wBp* from *B. pahangi* FR3. *B. pahangi* FR3, which was originally obtained from a green leaf monkey in Kuala Lumpur (1), is distributed by the NIAID Filariasis Research Reagent Resource Center (FR3) (2).

We assembled the *wBp* genome using multiple technologies. For short-read sequencing, genomic DNA from *B. pahangi* FR3 was acquired from BEI Resources, KAPA Hyper libraries were constructed, and 107,643,863 Illumina HiSeq 2500 paired-end 150-bp reads were generated, with quality control and trimming using FastQC v0.11.7 (3) and Trimmomatic v0.38 (4), respectively. Default parameters were used unless otherwise noted. For PacBio RS II sequencing, *B. pahangi* FR3 genomic DNA was isolated with Qiagen gravity-flow Genomic-tips from 16 male and 16 female worms, which were homogenized in Qiagen G2 buffer containing RNase A with 80 U of proteinase K and precipitated with 20  $\mu$ g of glycogen. Genomic DNA was fragmented via Covaris gTUBE centrifugation. Libraries were constructed with the SMRTbell v1.0 template preparation kit and size selected (>15 kbp) with a Sage Blue Pippin cassette, and 3,267,281 reads were sequenced with P6C4 chemistry ( $N_{50}$ , 25.4 kbp). For Oxford Nanopore Technologies (ONT) sequencing, 185 adult female worms from TRS Labs (Athens, GA, USA), which maintains worms that are used interchangeably with worms from FR3, were homogenized in liquid nitrogen, and genomic DNA was phenol-chloroform extracted, ethanol precipitated, and spooled. Three ONT SQK-RAD004 libraries were constructed, from 1.8, 0.9, and 0.18  $\mu$ g genomic DNA, and 553,429 reads, 11,638 reads, and 18,303 reads, respectively ( $N_{50}$ , 21.6 kbp, 39.3 kbp, and 48.2 kbp, respectively), were generated using an R9.4 flow cell (FLO-MIN106), replacing loading beads with water. Additionally, an SQK-RAD004 library was constructed with 4  $\mu$ g of genomic DNA, 1.5  $\mu$ l of fragmentation mixture, and 3.5  $\mu$ l of 10 mM Tris-Cl (pH 8.0) with 0.02% Triton X-100 (5), yielding 143,988 reads ( $N_{50}$ , 19.3 kbp) with an R9.4 flow cell (FLO-MIN106). All ONT reads were base called with Guppy v3.1.5.

A preliminary *wBp* contig was generated in a *de novo* assembly of all *Brugia pahangi*

**Citation** Lebov JF, Mattick J, Libro S, Sparklin BC, Chung M, Bromley RE, Nadendla S, Zhao X, Ott S, Sadzewicz L, Tallon LJ, Michalski ML, Foster JM, Dunning Hotopp JC. 2020.

Complete genome sequence of *wBp*, the *Wolbachia* endosymbiont of *Brugia pahangi* FR3. Microbiol Resour Announc 9:e00480-20. <https://doi.org/10.1128/MRA.00480-20>.

**Editor** Christina A. Cuomo, Broad Institute

**Copyright** © 2020 Lebov 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 Julie C. Dunning Hotopp, [jdhotopp@som.umaryland.edu](mailto:jdhotopp@som.umaryland.edu).

**Received** 29 April 2020

**Accepted** 2 June 2020

**Published** 2 July 2020

data using Canu v1.8 (6), with an estimated genome size of 100 Mbp (7). The *wBp* contig was identified after comparison of this assembly with the genomes of *wBm* (the *Wolbachia* endosymbiont of *B. malayi* [8]) and *B. malayi* (9) using NUCmer v3.0 (10). Using alignments against this preliminary *wBp* contig, we extracted 1,100,975 Illumina reads with BWA-MEM v0.7.17 (11) using aM options (153× depth), 6,483 PacBio reads with minimap2 v2.17 (12, 13) using ax map-pb options ( $N_{50}$ , 30.2 kbp; 87× depth), and 11,374 ONT reads with minimap2 v2.17 using ax map-ont options ( $N_{50}$ , 19.9 kbp; 92× depth). These extracted reads were assembled into a circular genome with Unicycler v0.4.8 (14). To resolve indel errors, the entire *B. pahangi* assembly (7), including the *wBp* contig, was corrected with Pilon v1.22 (mindepth, 5; K, 85; minmq, 0; minqual, 35; fix indels) (15) using all Illumina reads. The entire *B. pahangi* assembly was used because of known lateral gene transfers between *wBp* and *B. pahangi*, which confound such algorithms (16). The resulting 1,072,967-bp genome (GC content, 34.2%) was manually rotated to the *Wolbachia ori* gene (17), and 1,019 open reading frames, 34 tRNA genes, and 1 copy each of 5S, 16S, and 23S rRNA genes were identified with the IGS Prokaryotic Annotation Pipeline (18).

**Data availability.** This genome has been deposited in GenBank under accession number CP050521. The raw data have been deposited in the SRA under the accession numbers SRX4135331, SRX4135330, SRX4135329, SRX4135328, SRX4135327, SRX4135326, SRX4135325, SRX4135324, SRX4135323, SRX4135322, SRX4135321, SRX4135320, SRX4135319, SRX4135318, SRX4135317, SRX4135316, SRX4135315, SRX4135314, SRX4135313, and SRX4135312 (PacBio), SRX7658407, SRX7658383, SRX7658378, SRX7658377, SRX7658352, SRX7658349, SRX7658341, SRX7658327, SRX7658323, SRX7658322, SRX7658317, and SRR10997235 (Illumina), and SRR11565851, SRR11472020, SRR11565849, and SRR11565826 (MinION).

## ACKNOWLEDGMENTS

This project was funded by federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, under grant U19AI110820.

Genomic DNA from adult female *Brugia pahangi* strain FR3 (NR-46542) and whole adult male (NR-48906) and female (NR-48904) worms were provided by the NIH-NIAID Filariasis Research Reagent Resource Center for distribution by BEI Resources, NIAID, NIH.

## REFERENCES

- Buckley JJ, Edeson JF. 1956. On the adult morphology of *Wuchereria* sp. (*malayi*?) from a monkey (*Macaca irus*) and from cats in Malaya, and on *Wuchereria pahangi* n.sp. from a dog and a cat. *J Helminthol* 30:1–20. <https://doi.org/10.1017/S0022149X00032922>.
- Michalski ML, Griffiths KG, Williams SA, Kaplan RM, Moorhead AR. 2011. The NIH-NIAID Filariasis Research Reagent Resource Center. *PLoS Negl Trop Dis* 5:e1261. <https://doi.org/10.1371/journal.pntd.0001261>.
- Andrews S. 2013. FastQC. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- 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>.
- Jain M, Koren S, Miga KH, Quick J, Rand AC, Sasani TA, Tyson JR, Beggs AD, Dilthey AT, Fiddes IT, Malla S, Marriott H, Nieto T, O'Grady J, Olsen HE, Pedersen BS, Rhie A, Richardson H, Quinlan AR, Snutch TP, Tee L, Paten B, Phillippy AM, Simpson JT, Loman NJ, Loose M. 2018. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat Biotechnol* 36:338–345. <https://doi.org/10.1038/nbt.4060>.
- 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>.
- Mattick J, Libro S, Sparklin BC, Chung M, Bromley RE, Nadendla S, Zhao X, Ott S, Sadzewicz L, Tallon LJ, Michalski ML, Foster JM, Dunning Hotopp JC. 2020. Nearly complete genome sequence of *Brugia pahangi* FR3. *Microbiol Resour Announc* 9:e00479–20. <https://doi.org/10.1128/MRA.00479-20>.
- Foster J, Ganatra M, Kamal I, Ware J, Makarova K, Ivanova N, Bhattacharyya A, Kapatral V, Kumar S, Posfai J, Vincze T, Ingram J, Moran L, Lapidus A, Omelchenko M, Kyrpides N, Ghedin E, Wang S, Goltsman E, Joukov V, Ostrovskaya O, Tsukerman K, Mazur M, Comb D, Koonin E, Slatko B. 2005. The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. *PLoS Biol* 3:e121. <https://doi.org/10.1371/journal.pbio.0030121>.
- Tracey A, Foster JM, Paulini M, Grote A, Mattick J, Tsai Y-C, Chung M, Cotton JA, Clark TA, Geber A, Holroyd N, Korch J, Libro S, Lustigman S, Michalski ML, Rogers MB, Twaddle A, Dunning Hotopp JC, Berriman M, Ghedin E. 2020. Nearly complete genome sequence of *Brugia malayi* strain FR3. *Microbiol Resour Announc* 9:e00154–20. <https://doi.org/10.1128/MRA.00154-20>.
- Delcher AL, Salzberg SL, Phillippy AM. 2003. Using MUMmer to identify similar regions in large sequence sets. *Curr Protoc Bioinformatics Chapter 10:Unit 10.3*. <https://doi.org/10.1002/0471250953.bi1003s00>.
- 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>.
- Li H. 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* 32:2103–2110. <https://doi.org/10.1093/bioinformatics/btw152>.
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences.

- Bioinformatics 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
14. 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>.
  15. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel 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>.
  16. Hotopp JCD, Clark ME, Oliveira DCSG, Foster JM, Fischer P, Torres MCM, Giebel JD, Kumar N, Ishmael N, Wang S, Ingram J, Nene RV, Shepard J, Tomkins J, Richards S, Spiro DJ, Ghedin E, Slatko BE, Tettelin H, Werren JH. 2007. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* 317:1753–1756. <https://doi.org/10.1126/science.1142490>.
  17. Ioannidis P, Hotopp J, Sapountzis P, Siozios S, Tsiamis G, Bordenstein SR, Baldo L, Werren JH, Bourtzis K. 2007. New criteria for selecting the origin of DNA replication in *Wolbachia* and closely related bacteria. *BMC Genomics* 8:182. <https://doi.org/10.1186/1471-2164-8-182>.
  18. Galens K, Orvis J, Daugherty S, Creasy HH, Angiuoli S, White O, Wortman J, Mahurkar A, Giglio MG. 2011. The IGS standard operating procedure for automated prokaryotic annotation. *Stand Genomic Sci* 4:244–251. <https://doi.org/10.4056/sigs.1223234>.