AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Host-Microbial Interactions | Announcement

De novo whole-genome assembly of the *Wolbachia* sp. endosymbiont from *Anastrepha fraterculus* using long- and short-read metagenomic data

Claudia A. Conte,[1] Maximo Rivarola,[2] Sergio Gonzalez,[3] Fabian H. Milla,[1] Celeste Soria,[1] María C. Giardini,[1] Diego F. Segura,[4] Alfred M. Handler,[5] Kostas Bourtzis,[6] Jiannis Ragoussis,[7] Silvia B. Lanzavecchia[1]

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT A whole-genome assembly and annotation of *Wolbachia* sp. infecting *Anastrepha fraterculus* sp. 1 were generated by a metagenomic analysis of sequencing reads from a host genome project. This study contributes to the characterization of this endosymbiotic bacterium and provides valuable insights for research on host-symbiont interactions and pest management strategies.

KEYWORDS *Wolbachia*, whole-genome assembly, fruit fly

Microorganisms, including archaea, bacteria, viruses, and fungi, play crucial roles in insect digestion, nutrition, development, and immunity (1). Particularly, studying symbiotic bacteria in insect pests opens new research avenues for developing innovative and sustainable control strategies (2). The South American fruit fly, *Anastrepha fraterculus* Wiedemann 1830 (Diptera: Tephritidae), is one of the most destructive insect pests in South and Central American countries causing significant economic losses (3). We present a *de novo* genome assembly and annotation of *Wolbachia* sp. infecting *A. fraterculus* Brazilian 1 morphotype from Argentina based on metagenomic analyses combining Oxford Nanopore Technologies (ONT) and Illumina-MGI reads from the host genome project (4).

Genomic DNA was obtained from a newly emerged female of the *A. fraterculus* line infected with the *w*AfraCast2_A *Wolbachia* strain (5), as previously described (4). Size selection was performed using the SRE-XS Kit (Pacific Biosciences, CA). Libraries were prepared with the SQK-LSK110 Kit (ONT, UK), sequenced on a PromethION R9 flow cell for 72 h, and base-called with Guppy-v6.4.6 (ONT, UK) in high-accuracy mode. Illumina-compatible libraries were prepared by Covaris shearing and NxSeq AmpFREE Low-DNA-Ligation Kit (BioSearch Technologies, CA, USA) with IDT xGen-Dual-Index-Adapters, converted to MGI format, and sequenced on a DNBSEQ-G400 (MGI Tech, USA) with 150-bp paired-end reads using cPAS chemistry. All sequencing services were provided by the McGill Center (Montreal, Canada). Quality control of ONT reads was conducted using Nanoplot-v1.44.0 (6), Porechop-v0.2.4 (7) to trim adapters, and Filtlong-v0.2.1 (8) to filter long reads. A *de novo* metagenome assembly was achieved with Flye-v2.9.5-b1801 (9) using the --meta option. Kraken2 with the StandardPF database (10) was used for contig taxonomic classification. The contig assigned to the *Wolbachia* taxon was extracted using Seqtk-v1.5 (11), and terminal overlap was verified by self-alignment with Minimap2-v2.3 (12). Illumina-MGI short reads were used for polishing using Pilon-v1.24 (13). Genome completeness and integrity were evaluated with Benchmarking Universal Single-Copy Orthologs (BUSCO)-v5.8.2 (14) and the *Rickettsia* gene data set. Pairwise average nucleotide identity (ANI) was calculated using FastANI-v1.34 (15) against the *Wolbachia* endosymbiont of *Drosophila melanogaster w*Mel (NZ_CP046925.1).

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

Address correspondence to Claudia A. Conte, conte.claudia@inta.gob.ar.

The authors declare no conflict of interest.

Received 7 April 2026
Accepted 8 April 2026
Published 23 April 2026

TABLE 1 Taxonomic identification of microorganisms in female *A. fraterculus* sp.1 ONT long reads (total raw reads = 15,810,325; high-quality reads = 15,810,276)

Microorganism group	Taxonomic classification	# of reads
Virus	*Herpesvirales*	19
	Poxviridae	13
	Baculoviridae	10
	Caudovirales	4
	dsDNA viruses, no RNA stage	3
	Total virus	49
Fungi	*Fusarium*	4,948
	Yarrowia	928
	Candida	374
	Kluyveromyces	169
	Naumovozyma	137
	Tetrapisispora	135
	Botrytis	108
	Total fungi	6,799
Bacteria	*Wolbachia*	239,236
	Flammeovirga	774
	Erwinia	280
	Streptomyces	126
	Segatella	84
	Staphylococcus	51
	Vibrio	48
	Sphingobacterium	44
	Rickettsia	38
	Acinetobacter	34
	Brevibacterium	25
	Total bacteria	240,740
Total		247,588
Contaminants	*Homo sapiens*	82,520

Structural annotation was performed with RASTtk-v1.073 (16). The *dnaA* gene position was determined from the annotation, and the sequence was rotated using SeqKit-v2.10.0 (17). PHASTER (18) was used to annotate prophage sequences within the *Wolbachia* genome.

A total of 247,588 high-quality long ONT reads were assigned to microorganisms (Table 1), and 1,363 contigs were assembled. A circular contig of 1,463,312 nucleotides was taxonomically assigned to *Wolbachia* sp. with 869× coverage. After polishing using 965,855 Illumina-MGI reads, the assembly yielded a 1,463,854 bp genome of 35.07% guanine-cytosine (GC) content and 99.1% ANI to the *wMel* reference genome. BUSCO analysis indicated a high-quality assembly with 99.5% completeness and 0.3% duplicated sequences. The annotated genome includes 1,549 predicted protein-coding sequences, 34 tRNAs, and a complete set of rRNA genes in two unlinked transcription units (16S and 23S-5S). Two complete prophage regions and 119 open reading frames (ORFs) related to insertion sequences (ISs) from IS5 and IS4 families were identified (Fig. 1). The *Wolbachia* endosymbiont of *A. fraterculus* genome provides a valuable resource for studying host-symbiont interactions enhancing pest management strategies.

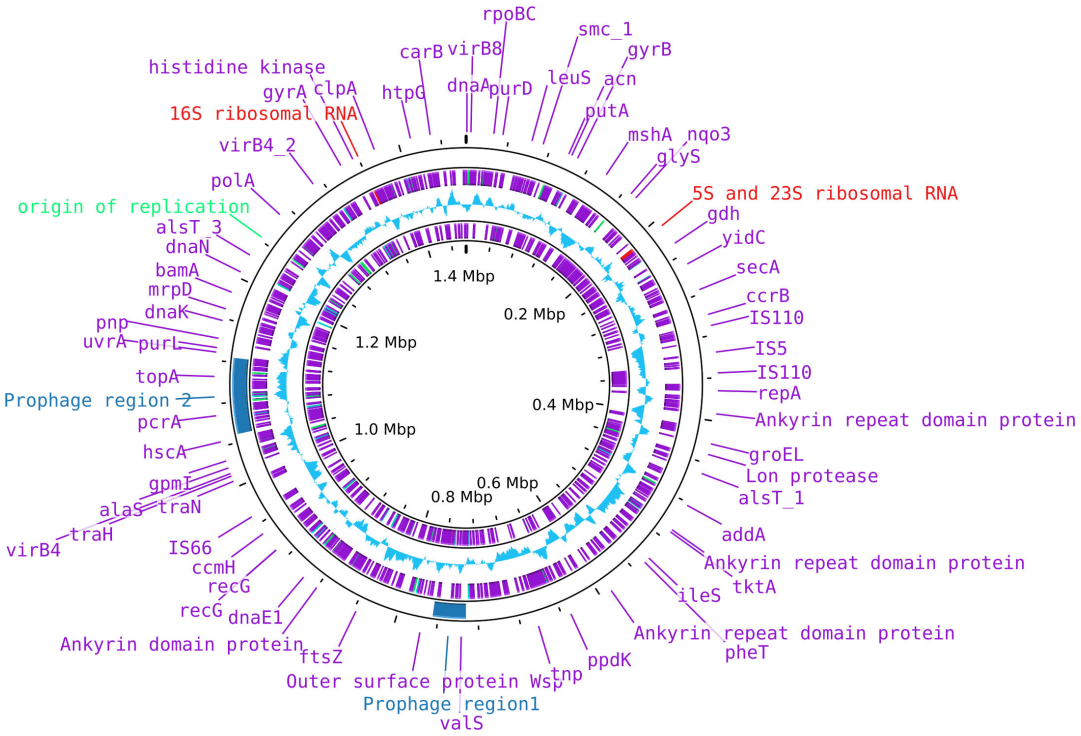

FIG 1 Representation of the genome of the *Wolbachia* strain *w*AfraCast2_A infecting *Anastrepha fraterculus* sp. 1. From outer to inner rings: complete prophage regions, coding sequences, GC content, and reverse-strand coding sequences. The origin of replication is indicated in green, and rRNA transcription units in red.

ACKNOWLEDGMENTS

This research was supported by PICT 2021 0491 and PICT 2021-0447 (Agencia Nacional de Promoción Científica y Técnica), PDI087 (Instituto Nacional de Tecnología Agropecuaria), and the International Atomic Energy Agency Research Contract no. 23402 as part of the Coordinated Research Project "Generic approach for the development of genetic sexing strains for SIT applications."

AUTHOR AFFILIATIONS

[1]Laboratorio de Genética y Genómica de Insectos, Instituto de Genética, Instituto Nacional de Tecnología Agropecuaria, Hurlingham, Buenos Aires, Argentina
[2]Apolo Biotech-CONICET, Santa Fe, Argentina
[3]Instituto de Agrobiotecnología y Biología Molecular-CONICET, Instituto Nacional de Tecnología Agropecuaria, Buenos Aires, Argentina
[4]Instituto de Genética, GV-IABIMO, INTA-CONICET, Hurlingham, Argentina
[5]U.S. Department of Agriculture, Agricultural Research Service, Center for Medical, Agricultural, and Veterinary Entomology, Gainesville, Florida, USA
[6]Department of Nuclear Sciences and Applications, Insect Pest Control Section, Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture, International Atomic Energy Agency, Vienna, Austria
[7]McGill Genome Centre, McGill University, Montreal, Canada

AUTHOR ORCIDs

Claudia A. Conte http://orcid.org/0009-0005-0233-9264

AUTHOR CONTRIBUTIONS

Claudia A. Conte, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Visualization, Writing – original draft, Writing –

review and editing | Maximo Rivarola, Data curation, Formal analysis, Methodology, Software, Supervision, Writing – original draft, Writing – review and editing | Sergio Gonzalez, Data curation, Formal analysis, Methodology, Software, Supervision, Writing – original draft, Writing – review and editing | Fabian H. Milla, Investigation, Methodology, Resources, Writing – original draft, Writing – review and editing | Celeste Soria, Investigation, Methodology, Resources, Writing – original draft, Writing – review and editing | María C. Giardini, Investigation, Methodology, Resources, Writing – original draft, Writing – review and editing | Diego F. Segura, Resources, Writing – original draft, Writing – review and editing | Alfred M. Handler, Investigation, Methodology, Writing – original draft, Writing – review and editing | Kostas Bourtzis, Investigation, Methodology, Writing – original draft, Writing – review and editing | Jiannis Ragoussis, Formal analysis, Methodology, Writing – original draft, Writing – review and editing | Silvia B. Lanzavecchia, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

The raw sequencing reads have been deposited in the Sequence Read Archive (SRA) under accession numbers SRR30230720–SRR30230729 for ONT data and SRR30233774 for Illumina-MGI reads. The genome assembly of the Wolbachia endosymbiont of A. fraterculus has been deposited in GenBank accession number CM143105.1. The Whole Genome Shotgun project has been deposited in accession number JBSXML000000000 and JBSXML010000001.

REFERENCES

1. Holt JR, Cavichiolli de Oliveira N, Medina RF, Malacrinò A, Lindsey ARI. 2024. Insect–microbe interactions and their influence on organisms and ecosystems. Ecol Evol 14:e11699. https://doi.org/10.1002/ece3.11699
2. Saikumar T, Manideep S, B T, Manoj MS, Paschapur AU, D T. 2025. Bacterial symbionts in tephritid fruit flies: biological roles and management strategies. Arthropod Plant Interact 19:46. https://doi.org/10.1007/s11829-025-10152-2
3. Lima de Brida A, Jean-BaptisteMC, Suárez L, Ovruski SM, Cancino J, Liburd OE, Mello Garcia FR. 2024. Biological control of fruit flies with emphasis on microbial control, p 127–141. In Management of Fruit Flies in the Americas. Springer International Publishing, Cham.
4. Rivarola M, Conte CA, Berube P, Chen S-H, Giardini MC, Scannapieco AC, Milla FH, Soria MC, Russo RM, Wulff JP, Djambazian HH, Pomar RR, Handler AM, Bourtzis K, Ragoussis I, Lanzavecchia SB. 2025. Chromosome-scale genome assembly of the South American fruit fly, Anastrepha fraterculus sp.1. Insect Sci, no. 33:640–664. https://doi.org/10.1111/1744-7917.70175
5. Conte CA, Segura DF, Milla FH, Augustinos A, Cladera JL, Bourtzis K, Lanzavecchia SB. 2019. *Wolbachia* infection in Argentinean populations of *Anastrepha fraterculus* sp1: preliminary evidence of sex ratio distortion by one of two strains. BMC Microbiol 19:289. https://doi.org/10.1186/s12866-019-1652-y
6. De Coster W, Rademakers R. 2023. NanoPack2: population-scale evaluation of long-read sequencing data. Bioinformatics 39:btad311. https://doi.org/10.1093/bioinformatics/btad311
7. Wick R. 2017. Porechop. GitHub Repository. GitHub. Available from: https://github.com/rrwick/Porechop
8. Wick R. 2017. Filtlong. GitHub Repository. GitHub. Available from: https://github.com/rrwick/Filtlong
9. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, Salzberg SL, Steinegger M. 2022. Metagenome analysis using the Kraken software suite. Nat Protoc 17:2815–2839. https://doi.org/10.1038/s41596-022-00738-y
10. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, Kuhn K, Yuan J, Polevikov E, Smith TPL, Pevzner PA. 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat Methods 17:1103–1110. https://doi.org/10.1038/s41592-020-00971-x
11. Shen W, Le S, Li Y, Hu F. 2016. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. PLoS ONE 11:e0163962. https://doi.org/10.1371/journal.pone.0163962
12. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191
13. 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
14. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol 38:4647–4654. https://doi.org/10.1093/molbev/msab199
15. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114. https://doi.org/10.1038/s41467-018-07641-9
16. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA 3rd, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365
17. Shen W, Sipos B, Zhao L. 2024. SeqKit2: a Swiss army knife for sequence and alignment processing. iMeta 3:iMeta https://doi.org/10.1002/imt2.191
18. 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:W16–21. https://doi.org/10.1093/nar/gkw387