Complete circular genome sequence of the *Wolbachia* wAlbB endosymbiont of *Aedes albopictus*

Amit Sinha[1], Zhiru Li[1] Luo Sun, and Clotilde K. S. Carlow
New England Biolabs, Ipswich.
[1]These authors contributed equally to this work.

Abstract

Wolbachia are α-proteobacteria belonging to the order *Rickettsiales*. It is a maternally transmitted, intracellular symbiont of arthropods and nematodes and estimated to infect 40-60% of arthropod species. The tiger mosquito *Aedes albopictus* is naturally infected with *Wolbachia* strains *w*AlbA and *w*AlbB. Cell line Aa23 established from *Aedes albopictus* embryos retains only *w*AlbB and is used as a key model to study host-endosymbiont interactions. The available *w*AlbB genome with 156 scaffolds is incomplete, hampering a comprehensive analysis of the genome. We have assembled the complete circular genome of a *w*AlbB strain from the Aa23 cell line, from long-read PacBio sequencing data at 450X coverage. The assembled circular chromosome is 1,484,007 bp in size, an increase of 321 kb over the published *w*AlbB genome, making it the largest sequenced *Wolbachia* genome to date. The annotation of the genome identified 1,207 protein coding genes, 34 tRNA, 3 rRNA and 1 tmRNA loci. The long reads enabled sequencing over complex repeat regions which have been be difficult to resolve with short-read sequencing. The availability of a complete circular genome from *w*AlbB will enable further biochemical, molecular and genetic analyses on this strain and related *Wolbachia*.

Introduction

- *Wolbachia* is an intracellular alpha-proteobacteria occuring in an estimated 40-65% of insect species and parasitic filarial nematodes
- Cell line Aa23 established from *Ae. albopictus* embryos retains only *w*AlbB and is used as a key model to study host-endosymbiont interactions.
- High cell density during passaging helps to maintain a relatively stable infection rate (Figure 1A, B) and cell-free *Wolbachia* can be obtained from spent culture media (Figure 1C)
- Pacific Biosciences (PacBio) SMRT technology produces long reads, some as long as 100kb, making single and continuous assembly possible

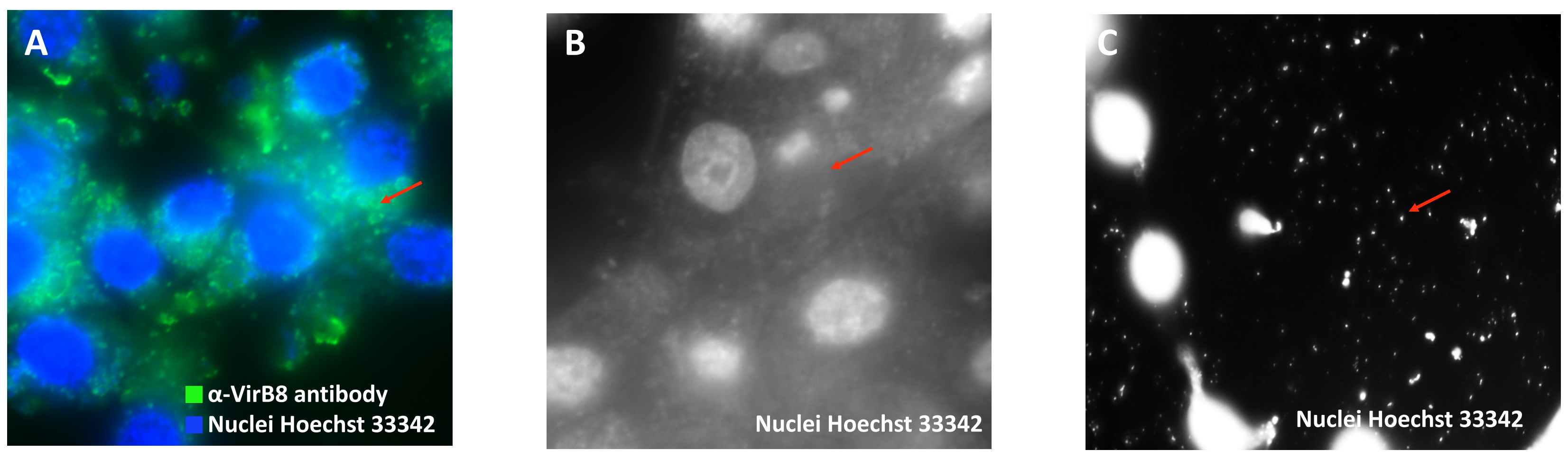
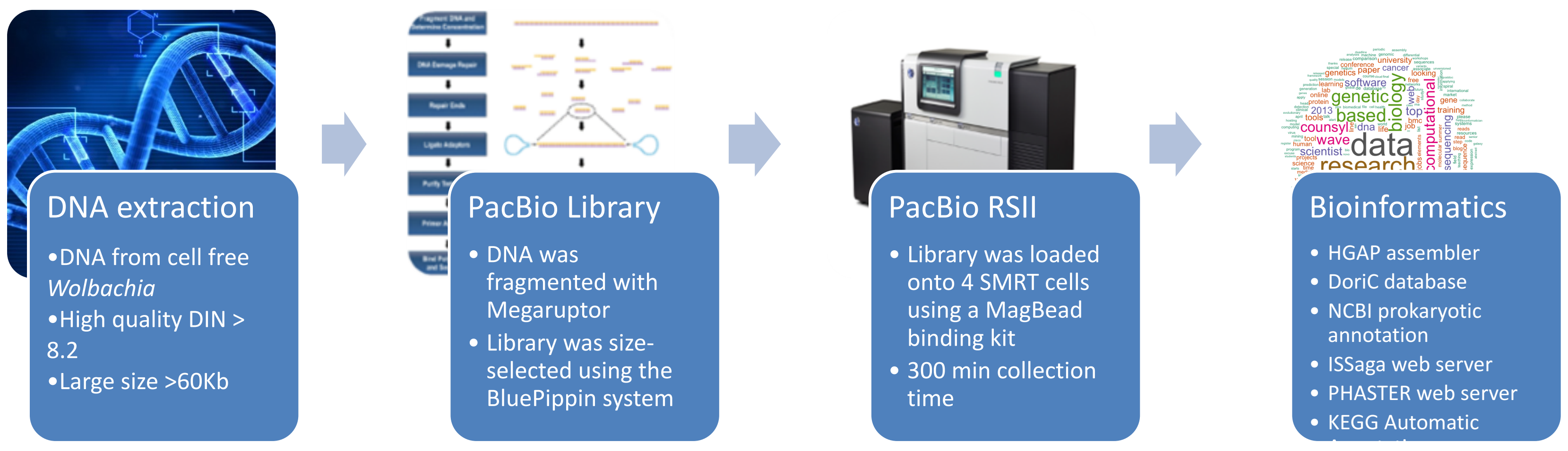


Figure 1: *Wolbachia* staining inside Aa23 cells (A, B) and in the spare medium (C). Red arrows: *Wolbachia*

Methods and Materials



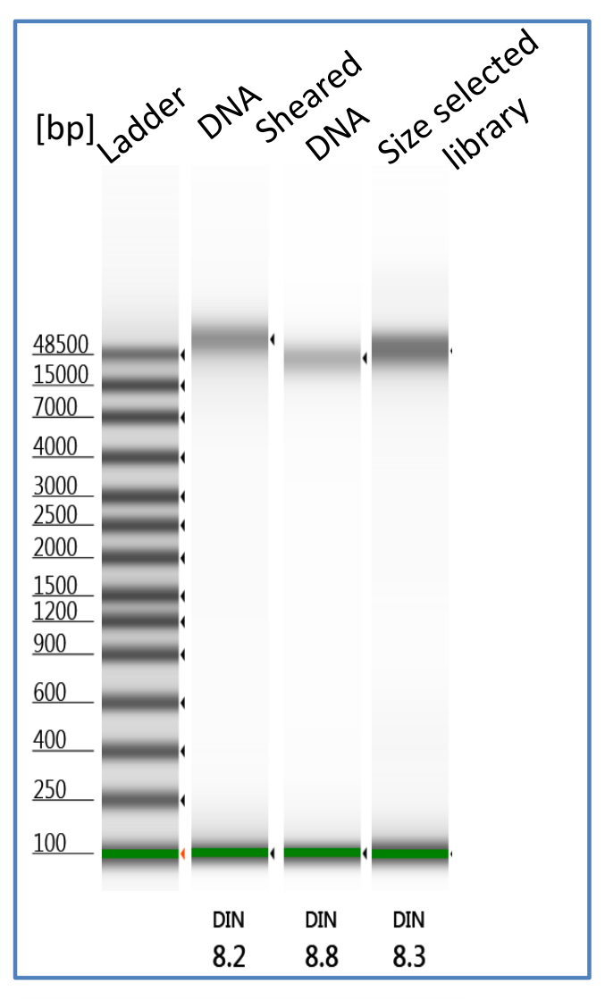
DNA extraction
- DNA from cell free *Wolbachia*
- High quality DIN > 8.2
- Large size >60Kb

PacBio Library
- DNA was fragmented with Megaruptor
- Library was size-selected using the BluePippin system

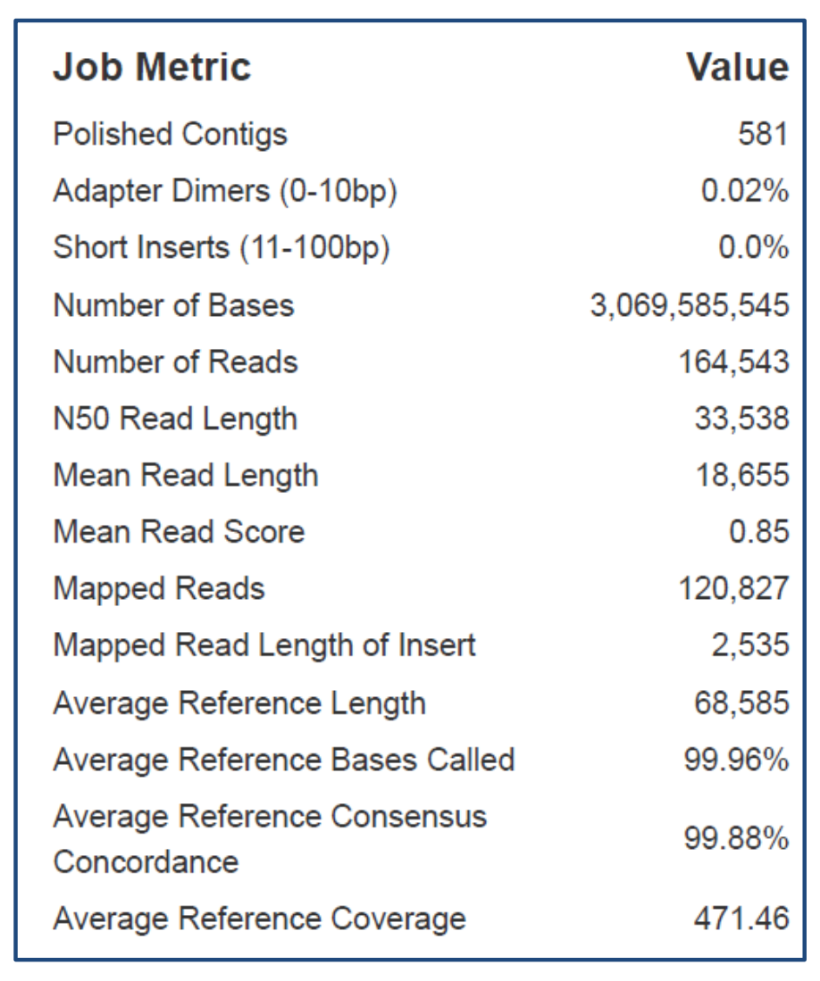
PacBio RSII
- Library was loaded onto 4 SMRT cells using a MagBead binding kit
- 300 min collection time

Bioinformatics
- HGAP assembler
- DoriC database
- NCBI prokaryotic annotation
- ISSaga web server
- PHASTER web server
- KEGG Automatic

DNA and Library quality

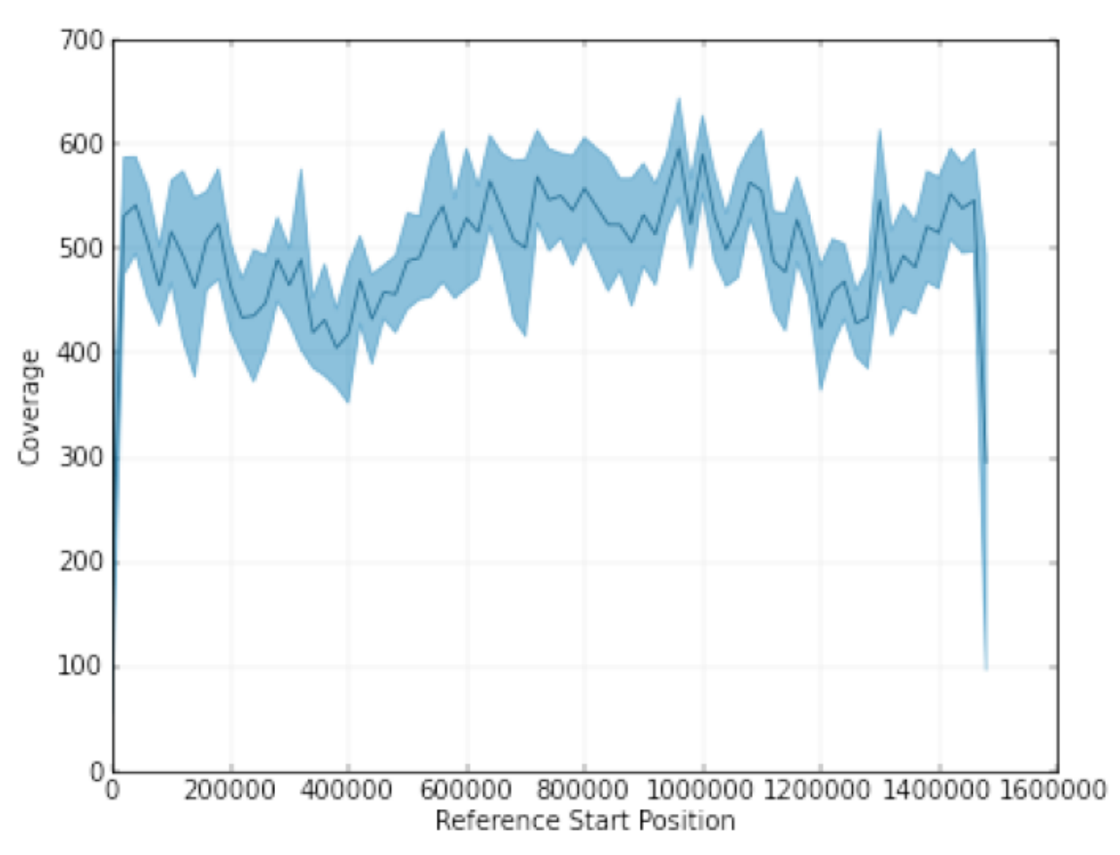


HGAP assembly output



Job Metric	Value
Polished Contigs	581
Adapter Dimers (0-10bp)	0.02%
Short Inserts (11-100bp)	0.0%
Number of Bases	3,069,585,545
Number of Reads	164,543
N50 Read Length	33,538
Mean Read Length	18,655
Mean Read Score	0.85
Mapped Reads	120,827
Mapped Read Length of Insert	2,535
Average Reference Length	68,585
Average Reference Bases Called	99.96%
Average Reference Consensus Concordance	99.88%
Average Reference Coverage	471.46

450X coverage across *w*AlbB contig



Results

Complete, circular chromosome and PCR confirmation

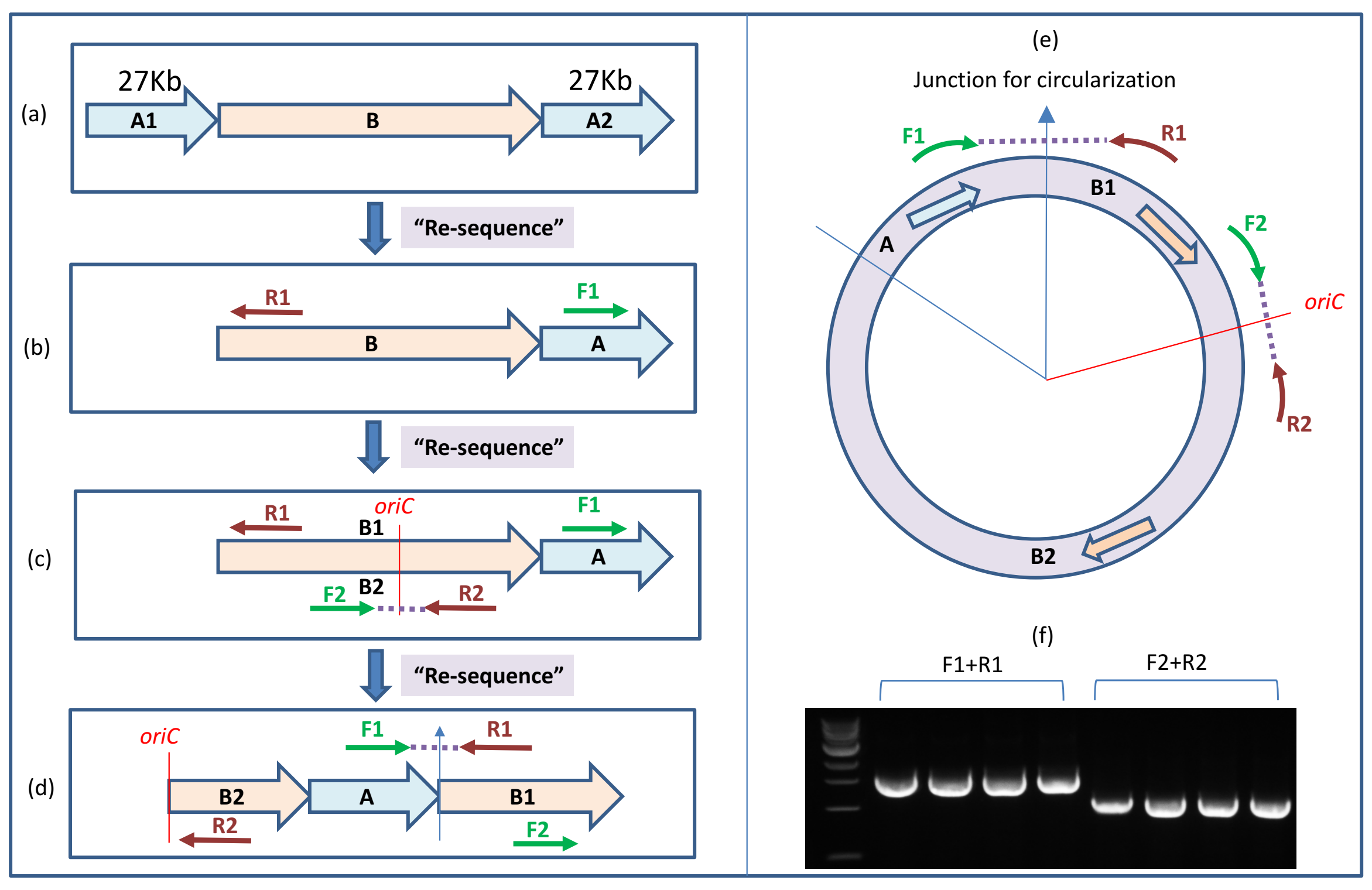


Figure 2: Pipeline used to circularize the genome and reset the origin to the *oriC* locus

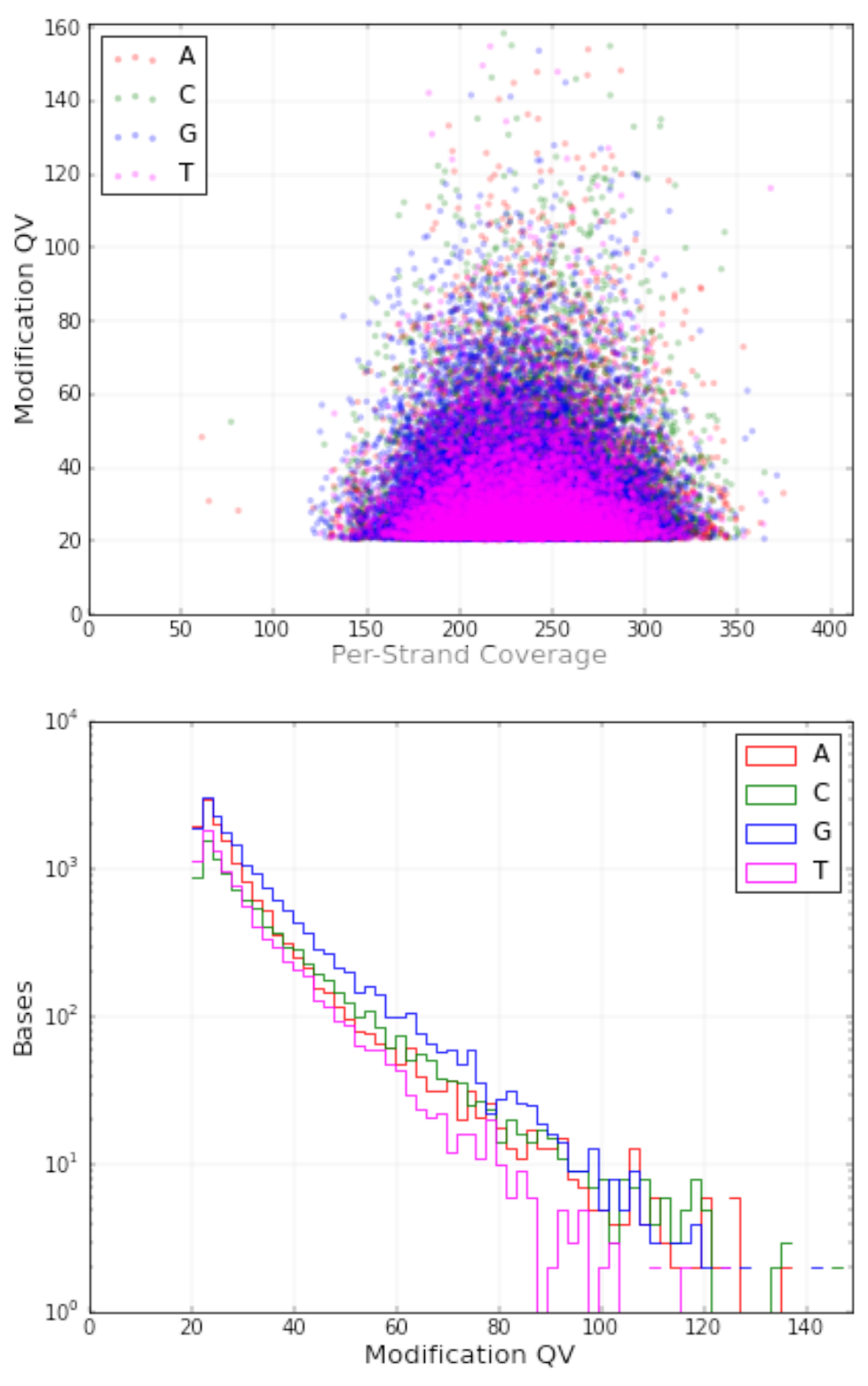


Figure 3: No detectable methylation signal

Genome analysis and annotation

- No methylation detected in *w*AlbB genome (PacBio modification detection pipeline) (Figure 3)
- Genome encodes 1,207 protein coding genes, 34 tRNA, 3 rRNA, 1 tmRNA loci and 188 pseudogenes
- Complete genome enables accurate annotation of repeat elements, potential prophages, pseudogenes
- 250 genes encode Insertion Sequence (IS) elements (Figure 4A)
    - Most abundant families: IS982 family (105 copies); IS481 family (76 copies)
- Two partial prophage regions were found (Figure 4B)
    - Prophage region 1: 16 kb, 18 genes
    - Prophage region 2: 5 kb, 8 genes
- Genome analysis of wAlbB has revealed the presence of a T4SS with 15 components organized in two operons and 4 individual genes (Figure 4B)
- Using a genome-wide pfam protein domain scan, 37 *w*AlbB genes were found to contain at least one copy of an ankyrin repeat domain, with a total of 87 copies of various ankyrin repeat containing domains (Figure 4B)
- Comparative KEEG analysis of wAlbB and other *Wolbachia* genome revealed 5 genes are absent in *w*AlbB (Table 1)

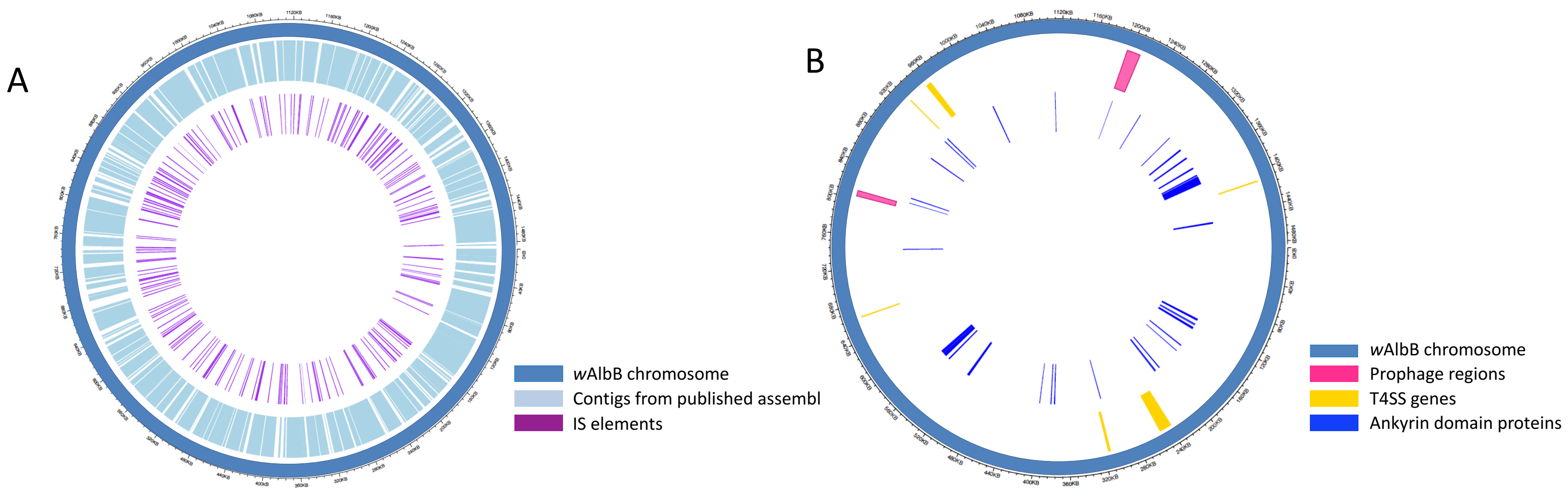


Figure 4: Circular genome map of *w*AlbB (A) Annotated IS and comparison to published *w*AlbB scaffolds (B) Annotated prophage, T4SS and ankyrin region

Table1. Five genes that are absent in *w*AlbB

Gene	wAlbB	wMel	wRi	wHa	wNo	wPip	wBm	wOo	wCle
DgkA	Absent	WD_1163	wRi_011390	wHa_09720	wNo_07140	WP0909	Absent	Absent	Absent
MPG	Absent	WD_1110	WRi_012850	wHa_09290	wNo_05480	WP0867	Wbm0254	wOo_04680	WCLE_011920
CytA	Absent	WD_0740	WRi_007360	wHa_06280	Absent	Absent	Absent	Absent	Absent
CytB	Absent	WD_0741	WRi_007350	wHa_06290	Absent	Absent	Absent	Absent	Absent
FtsI	Absent	WD_1273	WRi_012430	wHa_10600	Absent	Absent	Absent	Absent	WCLE_010110

Conclusions

- We assembled a complete circular *w*AlbB genome from the Aa23 cell line
- PacBio long reads enabled sequencing over complex repeat regions which have been be difficult to resolve with short read sequencing
- The availability of a complete circular chromosome from *w*AlbB will enable further biochemical, molecular and genetic analyses on this strain and related *Wolbachia*

References

1. Werren JH, Baldo L, Clark ME. 2008. *Wolbachia*: master manipulators of invertebrate biology. Nat Rev Microbiol 6:741-751.
2. O'Neill SL, Pettigrew MM, Sinkins SP, Braig HR, Andreadis TG, Tesh RB. 1997. In vitro cultivation of *Wolbachia pipientis* in an *Aedes albopictus* cell line. Insect Mol Biol 6:33-39.
3. Gerenday A, Fallon AM. 1996. Cell cycle parameters in *Aedes albopictus* mosquito cells. In Vitro Cell Dev Biol Anim 32:307-312.
4. Rhoads A, Au KF. 2015. PacBio Sequencing and Its Applications. Genomics Proteomics Bioinformatics 13:278-289.
5. Mavingui P, Valiente Moro C, Tran-Van V, Wisniewski-Dye F, Raquin V, Minard G, Tran FH, Voronin D, Rouy Z, Bustos P, Lozano L, Barbe V, Gonzalez V. 2012. Whole-genome sequence of *Wolbachia* strain *w*AlbB, an endosymbiont of tiger mosquito vector *Aedes albopictus*. J Bacteriol 194:1840.

Acknowledgement

We gratefully acknowledge the encouragement of Donald Comb and support of New England Biolabs .