

Phage WO packages cytoplasmic incompatibility genes in *Wolbachia*

Alice Brunner

`alice.brunner@ird.fr`

INRAE <https://orcid.org/0009-0005-9483-8262>

Camille Gaudiard

INRAE

Maxime Mahout

INRAE <https://orcid.org/0000-0002-2699-5582>

Blandine Trouche

INRAE <https://orcid.org/0000-0003-4307-4782>

Claire Geslin

UBO

Sarah R Bordenstein

Pennsylvania State University <https://orcid.org/0000-0001-6092-1950>

Seth R Bordenstein

Pennsylvania State University <https://orcid.org/0000-0001-7346-0954>

Julie Reveillaud

`reveillaud.j@gmail.com`

INRAE <https://orcid.org/0000-0002-2185-6583>

Research Article

Keywords: Wolbachia, Cytoplasmic incompatibility, bacteriophage, plasmid, Horizontal Gene Transfer, Metagenomics, Culex, Symbiosis, Vector control

Posted Date: May 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9836465/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Cytoplasmic incompatibility (CI), one of the most widespread reproductive manipulations in arthropods and a cornerstone of symbiont-based vector control strategies, is mediated by the phage WO genes *cifA* and *cifB* in *Wolbachia*. These genes undergo extensive horizontal transfer and rapid evolution, yet their mobilization remain poorly understood. Here, we isolated particulate WO phages from *Culex quinquefasciatus* and analyzed encapsidated phage DNA using state-of-the art metagenomic and pangenomic approaches. WO particles preferentially package prophage regions, including structural and eukaryotic association modules, and, critically, the *cif* genes. This discovery provides the first evidence that reproductive parasitism genes are physically packaged into phage particles. Coverage and assembly analyses further reveal heterogeneous packaging across multiple prophage regions, consistent with complex phage dynamics and modular genome organization. Together, our results support a model in which lytic WO phages drive the horizontal transfer and diversification of CI factors, with consequential effects on their eukaryotic hosts.

Main

Bacteria-infecting viruses (bacteriophages or phages) are among the most abundant biological entities on Earth and constitute a major reservoir of genetic diversity^{1–4}. Acting both as parasite of bacteria and vectors of horizontal gene transfer (HGT) through transduction and lysogenic conversion, phages profoundly shape bacterial genome evolution, population dynamics, and microbial community structure⁵. Moreover, phage genes and their encoded functions shape interactions with non-bacterial organisms that extend across multiple domains of life⁶.

Temperate phages establish long-term associations with their hosts by integrating as prophages, embedding viral genes that can significantly alter host biology. In addition to core viral functions involved in virus replication, prophage genomes can encode determinants that enhance bacterial fitness, including virulence factors, anti-phage defense systems, auxiliary metabolic genes and remarkably endosymbiotic traits^{7–11}. Such endosymbiotic genes confer key adaptive benefits, as illustrated by prophage-encoded genes for reproductive parasitism in phage WO of arthropod *Wolbachia*, insect parasite defense in the aphid endosymbiont *Hamiltonella defensa*, and inhibition of phagocytosis in sponge-associated bacteria^{10,12–16}.

A prominent system in which phage–bacterium–eukaryote interactions converge is the aforementioned *Wolbachia*, a maternally inherited endosymbiont that is widespread among arthropods and a major driver of host evolution^{17–19}. *Wolbachia* manipulate host reproduction through phenotypes including cytoplasmic incompatibility (CI), parthenogenesis, feminization, or male killing, thereby enhancing its own transmission and reshaping host population biology¹⁸. Because CI can rapidly invade insect populations, it has become a central mechanism underlying symbiont-based vector control strategies against mosquito-borne diseases¹⁸. CI, in which infected males cause embryo failure when mating with

uninfected females, is mediated by *cytoplasmic incompatibility factors (cif)* A and B located in a *Wolbachia* temperate bacteriophage WO^{13,20}. These genes occur within a Eukaryotic Association Module (EAM), a genomic interface bridging prokaryotic and eukaryotic domains that is proposed to facilitate interactions across domains of life, including WO's traversal of multiple membranes^{21,22}. Unfortunately, *Wolbachia* cannot be cultured, and its repeat-rich genome complicates assembly from short-read metagenomic data, leaving significant gaps in our understanding of its intrinsic genomic architecture and diversity^{23–26}.

In *Culex* mosquitoes, major vectors of medically important arboviruses including West Nile, Japanese encephalitis, and Usutu virus, there are several *Wolbachia* draft genomes (e.g., wPipPel, wPipJHB and wPipMol), but none are completely assembled into a single, closed circular contig^{23,27,28}. WO prophage regions in wPip are particularly enriched in *cif* genes, which are often present in multiple divergent copies distributed across several prophage loci^{23,25}. More than ten *cif* types have been described across *Wolbachia*^{19,29,30}, reflecting extensive diversification of these host-manipulation factors. In *Culex*, an expanded *cif* repertoire including both type I and type IV *cif* genes have been identified^{22,23}, although type I *cif* genes are the best characterized in relation to CI in this system^{13,20}. Type I CifB proteins notably contain deubiquitinase (DUB) domains that may play important role in CI induction²⁰, although their precise contribution remains not fully resolved³¹. This multiplicity, diversification and distribution of *cif* loci across WO regions point to extensive phage associated genomic rearrangements in wPip genomes^{25,32–34}. Consistent with this idea, long-read sequencing has begun to reveal extensive structural heterogeneity within WO prophage regions, suggesting dynamic phage architecture within individual *Culex* hosts^{32,33}.

Although horizontal transfer and lytic activity of phage WO have been reported between distinct *Wolbachia* strains^{35–39}, many WO regions in *Wolbachia* consist of degraded prophage relics shaped by deletions and recombination^{22,40}, complicating the identification of active phage particles and packaged genomes. Only two complete encapsidated WO genomes are currently available, limiting our understanding of phage gene packaging and functions in fully assembled genomes. Both WO genomes of wVitA-infected *Nasonia* parasitoid wasps²¹ and wSoc-infected *Allonemobius socius*⁴¹ show evidence for packaging of modules associated with lysogeny, replication, virion structure, and EAM^{21,41}. However, *cif* genes were not detected among sequences in these packaged phage particles. Despite this, their horizontal transfer has been widely hypothesized to involve phage WO activity³⁷, alongside alternative mechanisms such as transposon-mediated transfer^{33,42}.

The recent discovery of *Wolbachia* plasmids containing phage WO-like genes, insertion sequences, and in a rare instance *cif* genes^{26,43}, prompts further investigation into their possible role in HGT and pervasive recombination. Interestingly, the highly conserved pWCP plasmid in wPip was found to co-replicate with the bacterium^{44,45}. Among the 14 protein-coding genes of pWCP, seven remain of unknown function and one shows weak homology to a phage terminase, suggesting potential interactions

between plasmid and phage. Such piracy interactions could facilitate plasmid dissemination, potentially through hijacking of the phage capsid, as observed in other systems and further contribute to additional recombination^{46,47}.

Here, we investigated phage and plasmid activity in *Culex quinquefasciatus* female mosquitoes following a series of phage isolation experiments using state-of-the art metagenomic and pangenomic approaches together with qPCR. We captured WO phage particles from *Culex* and report, for the first time, the presence of reproductive manipulation genes in the packaged phage WO genome. These findings support the model that particle production of phage WO enables the direct transfer of reproductive parasitism genes and the active involvement of the WO lytic phage in the *wPip* infection worldwide.

Results

WO phage particles detected in *Culex*

We isolated viral particles from a pool of about 3,000 adult *Culex* females and sequenced the encapsidated DNA. To assess the genomic origin and distribution of the encapsidated molecules (hereafter referred as the isolate), we mapped the resulting reads to the *Wolbachia* wPipPel reference genome²³, selected for its well-annotated WO prophage regions. Read coverage showed a marked enrichment over annotated WO prophage regions (Fig. 1A), reaching ~ 1,000×, compared to ~ 10× across the rest of the *Wolbachia* chromosome. This low-level chromosomal coverage may reflect either the encapsidation of a limited subset of chromosomal genes in the pool of sequenced individuals (reflecting some generalized transduction) or background genomic material remaining after the isolation procedure. Quantitative PCR independently confirmed the observed enrichment in phage WO DNA and the efficiency of the isolation procedure. The phage marker *gpW* was enriched by approximately 3,000- to 9,000-fold relative to the *Wolbachia* chromosomal marker *wsp*, while the mosquito host gene *ace2* was undetectable at all measurement points (Fig. 1B, **Supplementary Table 1**). The plasmid marker *gp11* (pWCP) was detected at low levels, approximately 1.5- to 10-fold higher than *wsp*, which may reflect residual plasmid DNA associated with its multicopy nature or limited plasmid DNA encapsidation (**Supplementary Table 1**). At the scale of individual WO loci (WO_{Pip1–5}), read coverage across prophage regions was markedly heterogeneous (Fig. 1A).

WO phage particles carry cytoplasmic incompatibility factor genes

To characterize the gene content of the encapsidated DNA, we quantified read coverage across WO gene modules grouped by predicted functional categories, including head, connector/baseplate, tail, and tail fiber, as well as WO-associated genes, such as toxins, ankyrins, manipulative genes *cif*, and *wmk*-like genes, the latter being homologous to *Wolbachia* male-killing factors described in other hosts¹², although such a phenotype has not been reported in *Culex*. *Wolbachia* MLST genes, corresponding to

conserved chromosomal housekeeping genes commonly used for strain typing⁴⁸, located outside phage regions were included as baseline, chromosomal references (Fig. 1C).

We detected *cif* genes in the encapsidated DNA, with substantial and heterogeneous coverage across loci and gene type. Complete *cifB* type I genes reached coverage values of approximately 110×, whereas the corresponding complete *cifA* type I genes showed lower coverage, around 80×. In addition, *cifA–cifB* type IV gene pairs exhibited intermediate coverage levels, ranging from approximately 75× to 85×. Of note, short truncated *cifB* type I-related loci, including WP1291 and WP1292 located within WOPip5²³, showed the highest coverage among all *cif*-associated genes, reaching up to approximately 400×.

Because WO prophage regions in the reference genome (*wPipPel*) are highly repetitive and their number, boundaries, and synteny remain uncertain, coverage profiles can only be interpreted at the gene families and functional modules rather than at the level of individual prophage whole regions (Fig. 1, see legend). Several EAM genes other than the *cif* genes displayed elevated coverage, predominantly ankyrin repeat encoding genes (~ 90×). In contrast, genes associated with other host-interaction modules, including the Octomom-like region and *wmk*-like loci, showed only moderate coverage levels (~ 20-30×). However, their coverage remained consistently higher than that of *Wolbachia* chromosomal genes (average ~ 10×), suggesting that specific loci may be weakly represented among the encapsidated sequences. Coverage varied substantially among WO structural modules: head and connector/baseplate modules showed the highest enrichment (until ~ 420×), whereas tail and tail fiber modules were comparatively less covered (~ 60×) (Fig. 1C).

Assembly and pangenome structure of encapsidated vs. non-encapsidated WO phage genes

Assembly of viral DNA sequencing yielded 37 contigs (ranging from 1,006 bp to 34,159 bp) taxonomically assigned to *Wolbachia*, with a total of 175 predicted genes. Of note, 32 of the 37 contigs were annotated as phage-derived while the remaining ones contained genes annotated as proximal to WO prophage regions, primarily WOPip3 in *wPipPel* reference genome²³. Although some contigs showed homogeneous coverage, others exhibited substantial variability, with pronounced increases or decreases likely due to the assembly of highly repeated sequences and the nature of the pooled sequenced individuals. Figure 2A presents gene synteny and coverage profiles for contigs of interest longer than 5,000 bp. The longest contig (number #24, 34,159bp) encompassed multiple conserved WO phage core modules, including genes involved in head, tail, tail fiber, connector/baseplate, and replication or repair functions, consistent with previously described WO genomic architectures²² (Fig. 2A). Several genes belonging to these core modules displayed strikingly high read coverage, reaching up to 1,410×, such as a gene encoding a methyltransferase. This likely reflects the occurrence of highly similar or identical copies of this gene across several WO prophage regions (e.g, methyltransferase are present in up to four copies and recombinases occur in several copies in the *wPipPel* reference genome), leading to an artificial increase in coverage when reads are mapped to the reference genome.

In contrast, the eight of the additional longest contigs shown in Fig. 2A, while occasionally containing limited stretches of core phage genes, were predominantly composed of genes associated with the EAM. The presence of ankyrin repeat-containing proteins, predicted transcriptional regulators (including *wmk*-like genes) and *cif* genes within these contigs is consistent with the coverage-based analyses, confirming that EAM-associated loci detected in the isolate are assembled into genomic fragments and packaged together. Despite heterogeneous read coverage along individual contigs, the assembled fragments exhibited conserved gene order relative to prophage WO regions described in the wPipPel reference genome²³. In particular, contigs 24 and 21 showed strong synteny with regions corresponding to WOPip5 and WOPip2, respectively. Additional contigs displayed partial synteny with WOPip4 (contig 9), WOPip2 (contig 29) and WOPip1 (contig 25), indicating that the encapsidated DNA recapitulates substantial portions of multiple WO prophage regions.

We next performed a metapangenome analysis integrating the isolate assembly with available and novel draft genome wPip genomes together with read coverage patterns to further characterize the gene content of the encapsidated DNA across strains and infer possible general transduction mechanisms. We compared these genes to those of the wPipPel reference genome (1,48 Mb; 1,403 genes) along with its annotated WO prophage regions (WOPip1–5), pWCP plasmid from Reveillaud et al.²⁶ (absent from wPipPel), and a newly assembled wPipSLAB genome (Illumina sequencing, 131 contigs including a pWCP contig, 1,34 Mb, 1,281 genes), corresponding to the mosquito strain used for phage isolation (Fig. 2B). Genes from all datasets clustered into 1,141 homologous gene clusters (GCs) based on amino acid sequence similarity following the anvio pangenomics workflow⁴⁹. Among these, 117 GCs were represented in the isolate assembly, corresponding to only 10% of the total number of genes. Of these, 109 GCs were shared with wPipPel, and 97 were associated with at least one WO prophage region, showing that the encapsidated gene repertoire largely derives from previously annotated WO phage sequences rather than from the *Wolbachia* core genome or plasmid. Only eight GCs were found shared with wPipSLAB and not present in wPipPel, indicating WO genes possibly specific to SLAB strain, among which the *ankM2* protein from wPipMol²⁸ (**Supplementary Table 2**). In contrast, eleven GCs shared between WO regions of wPipPel and wPipSLAB genome were absent from the isolate assembly (Fig. 2B). These included genes annotated as a virulence-associated factor (*RhuM*), a *ParE* toxin, and an ankyrin repeat-containing protein, all of which showed low coverage values (**Supplementary Table 2**), suggesting that a subset of WO associated genes is not encapsidated. In addition, twelve GCs present in wPipPel but absent from both the isolate and wPipSLAB appeared to be strain-specific, further highlighting gene content variability among wPip lineages (**Supplementary Table 2**).

To link gene presence to their abundance in the isolate sample, we mapped isolate reads onto the wPipSLAB genome and retrieved cumulative coverage values for 1,114 GCs, visualized as a layer in the pangenome (Fig. 2B). Summed coverage values ranged from zero to 2,499.9x, with the maximum value corresponding to a protease gene belonging to the phage head module. Genes with coverage values close to zero predominantly corresponded to *Wolbachia* core genomes genes, whereas WO phage-associated genes displayed markedly higher coverage levels. Together, these results show that highly

abundant sequences in the isolate are overwhelmingly derived from WO phage regions, confirming the specificity of phage particle encapsidation.

A. Visualization of the nine longest contigs assembled from encapsidated DNA (out of 37 total contigs), showing gene organization and associated read coverage. Gene are annotated according to the WO prophage annotation described by Bordenstein and Bordenstein (2022)²².

Coverage scales are not identical across contigs and were adjusted individually for visualization purposes. **B.** Pangenome representation showing gene cluster distribution across the encapsidated DNA isolate sequenced in this study and across *Wolbachia* reference genomes and mobile genetic elements. Read coverage from the isolate is mapped onto the reconstructed wPipSLAB reference genome. Additional layers include wPipPel genome, pWCP plasmid, the five annotated WO prophage regions of wPipPel, and the newly reconstructed wPipSLAB genome, providing genomic context for the gene clusters detected in the isolate. A Venn diagram summarizes shared and unique gene clusters among the different annotated genome sequences.

Discussion

This study provides the first genomic evidence of WO phage particles harboring reproductive parasitism genes *cifA* and *cifB* in *Wolbachia* infecting *Culex*, supporting the hypothesis that lytic production of WO phage drives a capacity to horizontal transfer. This evidence is the culmination of several attempts to understand the nature of phage WO biology and its evolutionary roles. For example, phylogenetic analyses revealed incongruences between the phylogenies of *Wolbachia*, phage WO, and the arthropod hosts, suggesting the existence of horizontal transfers likely mediated by WO phages^{35–37,50–52}. These approaches have also strongly suggested WO in the movement of cytoplasmic incompatibility genes^{29,30,34,53}. Electron microscopy observations demonstrated icosahedral viral particles in *Culex* mosquitoes infected with *Wolbachia*^{54,55}. However, no study had previously characterized the DNA encapsidated in viral particles in the *Culex* system, leaving unresolved whether these genes could indeed be vectored by WO phages^{21,41}.

Despite the contig sequences obtained from phage WO enrichments, reconstructing a single genome of WO remains challenging, possibly due to biological and technical reasons such as a diversity of encapsidated WO molecules co-occurring in the sequencing library. The coverage profiles observed along the contigs specify that some WO genes are represented at variable copy numbers within the encapsidated DNA, such as methyltransferase genes, which are recurrently found across multiple WO prophage regions^{22,23}. This variability may be explained by the organization of the wPip genome, which contains several distinct WO prophage regions. In the reference genome wPipPel, five WO prophage regions are described (WOPip1 to WOPip5)²³. These regions are known to be challenging to assemble due to the presence of numerous repetitive sequences. They share many conserved genes while also displaying variations in synteny and gene content. Notably, only the prophages WOPip4 and WOPip5 contain the full set of genes required to form a complete viral particle, including the tail and tail fiber

genes. In this context, if several prophages can excise independently from the bacterial chromosome, the encapsidated DNA could originate from different combinations of prophage regions, resulting in coverage variation across specific genes.

Interestingly, several ankyrin repeat-containing genes were detected among the sequences enriched in the phage isolation. Ankyrin proteins are frequently associated with WO prophages^{23,56} and as part of the EAM²¹, which encodes proteins predicted to mediate interactions with the eukaryotic host. In other systems, phages enriched in ankyrin proteins, sometimes referred to “Ankyphages” have been shown to promote bacterial persistence in eukaryotic hosts by secreting ankyrin proteins (ANKp) that inhibit phagocytosis^{10,16}. Similarly, ankyrin rich WO phage could contribute to interactions between *Wolbachia* and its arthropod host.

This differential coverage may also be explained by the presence of satellite phages. In many bacterial systems, incomplete prophages can exploit structural proteins produced by complete phages to assemble their own viral particles. These satellite elements hijack the structural proteins and assembly mechanisms of helper phages to ensure their propagation^{57,58}. In the case of wPip, the prophages described as WOPip1, WOPip2 and WOPip3 in the reference genome, which lack certain essential structural genes, could correspond to such satellite phages and exploit WOPip4 and WOPip5 as helper phages for the encapsidation of their DNA.

The frequency of the transition from the prophage state to the particle form is uncertain and difficult to resolve, and the mechanisms triggering this induction remain unknown. Attempts to induce phage production through environmental stressors such as temperature shifts or mitomycin C exposure did not lead to increased production of viral particles under our experimental conditions, in contrast to observations reported for temperature changes in other systems such as *Nasonia vitripennis*⁵⁹. In many bacterial systems, prophage excision is triggered by the SOS response, a cellular response to DNA damage that primarily involves the RecA and LexA proteins^{60,61}. However, although the *recA* gene is present and highly expressed in *Wolbachia*, the genomes of this bacterium lack the *lexA* gene, suggesting that the canonical SOS pathway is not functional⁶². This feature, combined with the obligate intracellular lifestyle of *Wolbachia* and its reduced genome, may explain the difficulty in experimentally inducing WO phage production.

Although we were able to isolate a substantial quantity of viral particles from the sample analyzed here, subsequent attempts to isolate phages from other mosquito pools did not reproduce this result. It is therefore possible that only a subset of infections produce phages at a given time, suggesting that WO phage-mediated horizontal transfer may occur relatively infrequently. Nevertheless, even rare events may have significant long-term evolutionary consequences by contributing to the diversification of genes carried by WO prophages, particularly the *cif* genes involved in cytoplasmic incompatibility. Until now, the main mechanisms proposed to explain the diversification of these genes included horizontal transfer via mobile elements such as insertion sequences^{42,53,63,64} as well as amplification and diversification

processes within small genomic islands of phage WO recently described in *Culex*³³. These genomic islands contain recombination and mobility associated genes, including transposases and DNA repair factors such as *MutL*, which together create a recombinogenic genomic environment capable of promoting duplication, rearrangement, and diversification of *cif* loci. Thus, *cif* gene exchange and diversification is likely dynamic whereby WO phage particles underpin horizontal transfers between *Wolbachia* strains while genomic regions mobilize and replicate *cif* and gene neighbors within *Wolbachia*.

Finally, no interaction between the wPip plasmid pWCP and the particulate phages was detected here. We did not observe any signal of plasmid encapsidation, as the coverage of plasmid genes remained too low to support such a hypothesis. However, the absence of plasmid encapsidation observed here does not exclude the possibility of interactions between these two elements. In other systems, certain plasmids can hijack phage capsids to promote their horizontal transfer under specific phage induction conditions^{46,47}. Further work will be necessary to determine whether similar interactions may occur in *Wolbachia*.

In conclusion, this study provides the first experimental demonstration of particulate WO phages in *Wolbachia* infecting *Culex* and reveals the encapsidation of genes involved in host reproductive manipulation. These results shed new light on the mechanisms that can promote the diversification and dissemination of genes associated with cytoplasmic incompatibility. However, several major questions remain open, particularly regarding the mechanisms triggering phage induction, the genomic organization of encapsidated DNA, and the potential ability of these viral particles to infect new bacterial cells. Addressing these questions will be essential to better understand the role of WO phages in the evolution and dynamics of *Wolbachia* infections worldwide.

Methods

Mosquito strain and rearing

We maintained the *Culex quinquefasciatus* SLAB strain (originating from the San Joaquin Valley, California, USA, 1960s; ~96th laboratory generation) at 27°C, 70% relative humidity, under a 12 h:12 h light–dark cycle. We provided mosquitoes *ad libitum* with 10% sugar solution. Females received monthly blood meals from canaries (*Serinus canaria*) restrained in mosquito cages. We reared larvae in trays (30 × 20 × 6 cm) containing 1 L of osmosed water at a density of ~ 300 larvae per tray and fed them with a diet composed of Tetramin fish food and rabbit pellets (1:2). We collected adult females within 36 h post-emergence, prior to any blood feeding. Sampling was performed across multiple generations to obtain over 3,000 individuals, and stored at – 20°C until processing.

Viral particle isolation

We isolated viral particles following the protocol detailed in⁶⁵, with minor modifications to the initial steps. Briefly, we homogenized 2 g of pooled adult mosquitoes (**Supplementary Table 3**) on ice in 13.5

ml of cold SM buffer using a tissue homogenizer. We added 4 ml of 5 M NaCl and 2 µl of RNase A (10 mg/ml), adjusted the volume to 20 ml with SM buffer, and incubated the suspension for 1 h at 4°C with gentle agitation. We centrifuged the homogenate at 13,000g for 10 min at 4°C, and transferred the supernatant to a new tube. We added polyethylene glycol 6000 to a final concentration of 20%, and incubated samples overnight at 4°C before proceeding with the downstream steps of the original protocol. We homogenized the final particle suspension and prepared aliquots (≤ 200 µl) into 1.5 ml tubes to test DNase treatment conditions in triplicate

To remove non-encapsidated DNA, we treated isolated viral particles with DNase I (#M0303S, New England Biolabs). We evaluated DNase efficiency using increasing incubation times ranging from 20 to 60 min during the first isolation (Isolation 1; 1× DNase time kinetics). Additional experiments assessed DNase I activity on free lambda phage DNA and tested enzyme concentration kinetics across two independent isolations. Detailed protocols are provided for Isolation 2 and 3 in the **Supplementary Table 3**.

We extracted DNA using the QIAamp MiniElute Virus Spin Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions and measured DNA concentration using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Invitrogen™).

We used quantitative PCR (qPCR) to assess the relative abundance of target genes (**Supplementary Table 4**), including the mosquito *ace2* gene, the *Wolbachia* chromosomal marker *wsp*, plasmid pWCP markers (*gp11* or *gp14*), and the WO phage head–tail joining gene *gpW* (present in four copies in the wPipPel reference genome but one copy per WO regions WO_{pip1-2-3-5}; all others present as single-copy genes). We performed reactions on a LightCycler LC480 system (Roche Diagnostics) using the SensiFAST SYBR No-ROX Kit (Bioline) under the following cycling conditions: 45 cycles of 95°C for 10 s, 58°C for 20 s, and 65°C for 20 s.

We based relative quantification on standard curves generated from serial dilutions of a synthetic plasmid containing all target genes in single copy (Eurofins Scientific), following Brunner et al⁴⁵. Arbitrary values preserving dilution factors were assigned to generate relative standard curves, which were used to convert C_p values into relative concentrations. We normalized all values to total DNA concentration and expressed those as relative gene quantities per ng of DNA.

To assess the reproducibility of phage isolation and the conditions required for WO particle production, we performed additional large-scale isolations and induction assays. These experiments did not result in consistent enrichment of WO phage DNA or detectable induction under the tested conditions, suggesting that a production of WO particles is a rare and possibly transient event. Detailed methods and results are provided in **Supplementary Note 1**, **Supplementary Note 2**, **Supplementary Fig. 1** and **Supplementary Notebook 1**.

Sequencing and data processing

We sequenced encapsidated DNA from the phage isolation 1 (Isolate) using 150 nt paired-end Illumina short-read sequencing at the MGX–Montpellier GenomiX platform. Libraries were prepared with the Illumina DNA Prep kit (formerly Nextera), including enzymatic fragmentation followed by reduced-cycle PCR for adapter and index addition.

Raw reads were trimmed with BBDuk⁶⁶ and sequencing quality was assessed with the “iu-filter-quality-minoche” anvi’o program⁶⁷.

We performed *de novo* assembly of high-quality reads with MEGAHIT v1.2.9⁶⁸ using a minimum contig length of 1,000 bp. We performed gene calling with prodigal⁶⁹ and anvi’o tool “anvi-gen-contigs-database”. We then assigned taxonomy to assembled contigs using Centrifuge⁷⁰. We performed an automatic genome binning from assemblies using the “anvi-cluster-contigs” anvi’o program⁷¹ and the CONCOCT algorithm⁷². Finally, we refined the obtain metagenomic bins and retained contigs assigned to *Wolbachia* or *Wolbachia* phage (**Supplementary Notebook 2**). Functional annotation was performed using a database of homologous sequences specifically curated for WO prophage genes⁵³.

In parallel, we reconstructed a draft genome wPipSLAB of *Wolbachia* from ovaries metagenomes of an adult female mosquito from the same SLAB strain that was used for phage isolation. We performed sequencing, processed short reads and assembled *Wolbachia* contigs *de novo* using the same workflow as described above. We used the resulting genome as a strain-matching reference for comparative and pangenome analyses (**Supplementary Notebook 2**).

Meta-pangenomic analysis

We constructed a pangenome using anvi’o (anvi-pan-genome, --use-ncbi-blast, --mcl-cluster 10)⁷¹, integrating the isolate assembly with reference genomes: *Wolbachia* wPipPel, its WO prophage regions (WOPIp1–5, treated as independent genomes), pWCP plasmid from²⁶, and the newly assembled wPipSLAB genome. Homologous genes were clustered into gene clusters (GCs) and visualized using anvi-display-pan (**Supplementary Notebook 3**).

For the metapangenomics analysis, read coverage from the phage isolate mapped onto wPipSLAB genome was computed at the gene level using gene coordinates and imported into the pangenome as an additional data layer. For GCs represented by multiple gene copies, coverage values were summed across all corresponding loci (**Supplementary Notebook 4**).

Coverage visualization analyses

We mapped high-quality isolate reads to 1) the *Wolbachia* wPipPel reference genome²³ and the pWCP plasmid²⁶ and 2) the isolate assembly using Bowtie2 v2.5.1, as part of the AliNE Nextflow pipeline v1.1.1⁷³ (**Supplementary Notebook 2**). All Bowtie alignment steps were performed using defaults

parameters in order to allow primary alignment only (avoiding multiple hits and possible secondary alignments).

We visualized the mapping profile of isolate reads against wPipPel to assess the genomic origin and distribution of our encapsidated sequences compared to the reference genome. (**Supplementary Notebook 5**). Resulting coverage profiles were then used to examine read distribution across specific genomic regions, including WO phage loci such as structural modules and eukaryotic association module (EAM) genes, as well as MLST and plasmid pWCP genes (**Supplementary Notebook 6**). Finally, isolate reads were mapped to the isolate assembly in order to evaluate coverage consistency across assembled contigs and to validate assembly completeness (**Supplementary Notebook 7**).

Data availability

The raw sequencing data for the shotgun metagenomes of short read sample are available in the European Nucleotide Archive via accession number ERP188318.

Code availability

A reproducible bioinformatics workflow including seven Supplementary Notebooks used for computational analyses is available at the URL (<https://jreveillaud.github.io/RosaBioinfo/>).

Declarations

Acknowledgments

We are grateful to Arnaud Berthomieu and the Vectopole platform (IRD, Montpellier, France) for providing the SLAB mosquito strain, maintaining the colonies, and technical support, including access to canaries for blood feeding. We thank Stéphane Blanc for insightful discussions on viral particles. We acknowledge the MGX Montpellier sequencing platform for generating the sequencing data and thank Hugues Parinello for his support. This work was supported by the European Research Council RosaLind Starting Grant “948135” to JR.

Author information

Authors and Affiliations

UMR MIVEGEC (Maladies Infectieuses et vecteurs : Écologie, Génétique, Évolution et Contrôle), University of Montpellier, INRAE, CNRS, IRD, Montpellier, France

Alice Brunner, Camille Gaudiard, Maxime Mahout, Blandine Trouche, Julie Reveillaud

Department of Biology and Department of Entomology, One Health Microbiome Center, Huck Institutes of the life sciences, Pennsylvania State University, University Park, PA, USA

Seth R. Bordenstein, Sarah R. Bordenstein

UMR Biologie et Ecologie des Ecosystèmes Marins Profonds, Université de Bretagne Occidentale,
France

Claire Geslin

Contributions

AB, JR, CGa, SRB and SRB conceived the study. AB, CGa, SRB and JR developed the methodology. AB, CGa and JR performed the experiments. AB, MM and JR curated and analyzed the data. JR, SRB and SRB obtained funding. JR supervised the study. AB, MM, CGa and JR wrote the original draft of the manuscript. AB, CGa, MM, BT, CGe, SRB, SRB, JR contributed to data interpretation, provided critical feedback, and reviewed and edited the manuscript.

Corresponding authors

Correspondence to Alice Brunner and Julie Reveillaud

Ethics declarations

Competing interests

The authors declare no competing interests.

References

1. Brüssow H (2009) The not so universal tree of life or the place of viruses in the living world. *Philos Trans R Soc Lond B Biol Sci* 364:2263–2274
2. Salmond GPC, Fineran PC (2015) A century of the phage: past, present and future. *Nat Rev Microbiol* 13:777–786
3. Paez-Espino D et al (2016) Uncovering Earth's virome. *Nature* 536:425–430
4. Paez-Espino D et al (2019) Diversity, evolution, and classification of virophages uncovered through global metagenomics. *Microbiome* 7:157
5. Perfilyev A et al (2026) Large-scale analysis of bacterial genomes reveals thousands of lytic phages. *Nat Microbiol* 11:42–52
6. Wendling CC (2023) Prophage mediated control of higher order interactions - Insights from multi-level approaches. *Curr Opin Syst Biology* 35:100469
7. Bodner K, Melkonian AL, Covert MW (2021) The Enemy of My Enemy: New Insights Regarding Bacteriophage–Mammalian Cell Interactions. *Trends Microbiol* 29:528–541
8. Waldor MK, Mekalanos JJ (1996) Lysogenic Conversion by a Filamentous Phage Encoding Cholera Toxin. *Science* 272:1910–1914

9. Brüssow H, Canchaya C, Hardt W-D (2004) Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. *Microbiol Mol Biol Rev* 68:560–602
10. Jahn MT et al (2019) A Phage Protein Aids Bacterial Symbionts in Eukaryote Immune Evasion. *Cell Host Microbe* 26:542–550e5
11. Fineran PC et al (2009) The phage abortive infection system, ToxIN, functions as a protein–RNA toxin–antitoxin pair. *Proceedings of the National Academy of Sciences* 106, 894–899
12. Perlmutter JI et al (2019) The phage gene *wmk* is a candidate for male killing by a bacterial endosymbiont. *PLoS Pathog* 15:e1007936
13. LePage DP et al (2017) Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature* 543:243–247
14. Moran NA, Degnan PH, Santos SR, Dunbar HE, Ochman H (2005) The players in a mutualistic symbiosis: Insects, bacteria, viruses, and virulence genes. *Proceedings of the National Academy of Sciences* 102, 16919–16926
15. Oliver KM, Degnan PH, Burke GR, Moran NA (2010) Facultative Symbionts in Aphids and the Horizontal Transfer of Ecologically Important Traits. *Ann Rev Entomol* 55:247–266
16. Leigh BA (2019) Cooperation among Conflict: Prophages Protect Bacteria from Phagocytosis. *Cell Host Microbe* 26:450–452
17. Porter J, Sullivan W (2023) The cellular lives of *Wolbachia*. *Nat Rev Microbiol* 21:750–766
18. Kaur R et al (2021) Living in the endosymbiotic world of *Wolbachia*: A centennial review. *Cell Host Microbe* 29:879–893
19. Shropshire JD, Leigh B, Bordenstein SR (2020) Symbiont-mediated cytoplasmic incompatibility: What have we learned in 50 years? *eLife* 9, e61989
20. Beckmann JF, Ronau JA, Hochstrasser M (2017) A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat Microbiol* 2:1–7
21. Bordenstein SR, Bordenstein SR (2016) Eukaryotic association module in phage WO genomes from *Wolbachia*. *Nat Commun* 7:13155
22. Bordenstein SR, Bordenstein SR (2022) Widespread phages of endosymbionts: Phage WO genomics and the proposed taxonomic classification of Symbioviridae. *PLoS Genet* 18:e1010227
23. Klasson L et al (2008) Genome Evolution of *Wolbachia* Strain wPip from the *Culex pipiens* Group. *Mol Biol Evol* 25:1877–1887
24. Wu M et al (2004) Phylogenomics of the Reproductive Parasite *Wolbachia pipientis* wMel: A Streamlined Genome Overrun by Mobile Genetic Elements. *PLoS Biol* 2:e69
25. Bonneau M et al (2018) *Culex pipiens* crossing type diversity is governed by an amplified and polymorphic operon of *Wolbachia*. *Nat Commun* 9:319
26. Reveillaud J et al (2019) The *Wolbachia* mobilome in *Culex pipiens* includes a putative plasmid. *Nat Commun* 10:1051

27. Salzberg SL, Puiu D, Sommer DD, Nene V, Lee NH (2009) Genome Sequence of the *Wolbachia* Endosymbiont of *Culex quinquefasciatus* JHB. *J Bacteriol* 191:1725–1725
28. Pinto SB et al (2013) Transcriptional Regulation of *Culex pipiens* Mosquitoes by *Wolbachia* Influences Cytoplasmic Incompatibility. *PLoS Pathog* 9:e1003647
29. Tan Y, Aravind L, Zhang D (2024) Genomic Underpinnings of Cytoplasmic Incompatibility: CIF Gene-Neighborhood Diversification Through Extensive Lateral Transfers and Recombination in *Wolbachia*. *Genome Biol Evol* 16:evae171
30. Amoros J et al (2025) Diversity and spread of cytoplasmic incompatibility genes among maternally inherited symbionts. *PLoS Genet* 21:e1011856
31. Horard B et al (2022) Paternal transmission of the *Wolbachia* CidB toxin underlies cytoplasmic incompatibility. *Curr Biol* 32:1319–1331e5
32. Brunner A et al (2026) Extensive Mobilome Dynamics in a Widespread Endosymbiont: Long Read Metagenomics Reveal Dimeric Plasmids and Highly Fragmented Prophages in *Wolbachia* from *Culex Pipiens*. 10.21203/rs.3.rs-8855126/v1
33. Amoros J et al (2025) Genomic Islands in *Wolbachia* Prophages Drive Amplification and Diversification of Cytoplasmic Incompatibility Genes in *Culex pipiens*. 10.03.680192 Preprint at <https://doi.org/10.1101/2025.10.03.680192> (2025)
34. Namias A et al (2025) Recombination, Truncation and Horizontal Transfer Shape the Diversity of *Wolbachia*-induced Cytoplasmic Incompatibility Patterns. *Mol Biol Evol* 42:msaf200
35. Sanogo YO, Dobson SL (2004) Molecular discrimination of *Wolbachia* in the *Culex pipiens* complex: evidence for variable bacteriophage hyperparasitism. *Insect Mol Biol* 13:365–369
36. Gavotte L et al (2007) A Survey of the Bacteriophage WO in the Endosymbiotic Bacteria *Wolbachia*. *Mol Biol Evol* 24:427–435
37. Kent BN et al (2011) Complete Bacteriophage Transfer in a Bacterial Endosymbiont (*Wolbachia*) Determined by Targeted Genome Capture. *Genome Biol Evol* 3:209–218
38. Wang GH et al (2016) Bacteriophage WO Can Mediate Horizontal Gene Transfer in Endosymbiotic *Wolbachia* Genomes. *Front Microbiol* 7
39. Vancaester E, Blaxter M (2023) Phylogenomic analysis of *Wolbachia* genomes from the Darwin Tree of Life biodiversity genomics project. *PLoS Biol* 21:e3001972
40. Metcalf JA, Jo M, Bordenstein SR, Jaenike J, Bordenstein SR (2014) Recent genome reduction of *Wolbachia* in *Drosophila recens* targets phage WO and narrows candidates for reproductive parasitism. *PeerJ* 2:e529
41. Kupritz J et al (2021) Isolation and characterization of a novel bacteriophage WO from *Allonemobius socius* crickets in Missouri. *PLoS ONE* 16:e0250051
42. Cooper BS, Vanderpool D, Conner WR, Matute DR, Turelli M (2019) *Wolbachia* Acquisition by *Drosophila yakuba*-Clade Hosts and Transfer of Incompatibility Loci Between Distantly Related *Wolbachia*. *Genetics* 212, 1399–1419

43. Martinez J et al (2022) Genome sequencing and comparative analysis of *Wolbachia* strain wAlbA reveals *Wolbachia*-associated plasmids are common. PLoS Genet 18:e1010406
44. Ghousein A et al (2023) pWCP is a widely distributed and highly conserved *Wolbachia* plasmid in *Culex pipiens* and *Culex quinquefasciatus* mosquitoes worldwide. ISME COMMUN 3:1–5
45. Brunner A et al (2025) *Wolbachia* and its pWCP plasmid show differential dynamics during the development of *Culex* mosquitoes. Microbiol Spectr. <https://journals.asm.org/doi/10.1128/spectrum.00046-25>
46. Lossouarn J et al (2015) Ménage à trois': a selfish genetic element uses a virus to propagate within Thermotogales. Environ Microbiol 17:3278–3288
47. Jin M, Rouxel O, Quintin N, Geslin C (2025) Molecular piracy in deep-sea hydrothermal vent: phage-plasmid interactions revealed by phage-FISH in *Marinitoga piezophila*. Appl Environ Microbiol 91:e02306–e02324
48. Baldo L et al (2006) Multilocus Sequence Typing System for the Endosymbiont *Wolbachia pipientis*. Appl Environ Microbiol 72:7098–7110
49. Eren AM et al (2021) Community-led, integrated, reproducible multi-omics with anvi'o. Nat Microbiol 6:3–6
50. Bordenstein SR, Wernegreen JJ (2004) Bacteriophage Flux in Endosymbionts (*Wolbachia*): Infection Frequency, Lateral Transfer, and Recombination Rates. Mol Biol Evol 21:1981–1991
51. Gavotte L et al (2004) Diversity, distribution and specificity of WO phage infection in *Wolbachia* of four insect species. Insect Mol Biol 13:147–153
52. Chauvatcharin N, Ahantarig A, Baimai V, Kittayapong P (2006) Bacteriophage WO-B and *Wolbachia* in natural mosquito hosts: infection incidence, transmission mode and relative density. Mol Ecol 15:2451–2461
53. Amoros J et al (2025) Diversity and spread of cytoplasmic incompatibility genes among maternally inherited symbionts. PLoS Genet 21:e1011856
54. Wright JD, Sjostrand FS, Portaro JK, Barr AR (1978) The ultrastructure of the rickettsia-like microorganism *Wolbachia pipientis* and Associated virus-like bodies in the mosquito *Culex pipiens*. J Ultrastruct Res 63:79–85
55. Sanogo YO, Dobson SL (2006) WO bacteriophage transcription in *Wolbachia*-infected *Culex pipiens*. Insect Biochem Mol Biol 36:80–85
56. Klasson L, Kambris Z, Cook PE, Walker T, Sinkins SP (2009) Horizontal gene transfer between *Wolbachia* and the mosquito *Aedes aegypti*. BMC Genomics 10:33
57. Alqurainy N et al (2023) A widespread family of phage-inducible chromosomal islands only steals bacteriophage tails to spread in nature. Cell Host Microbe 31:69–82e5
58. Penadés JR, Seed KD, Chen J, Bikard D, Rocha E (2025) P. C. Genetics, ecology and evolution of phage satellites. Nat Rev Microbiol 23:410–422

59. Bordenstein SR, Bordenstein SR (2011) Temperature Affects the Tripartite Interactions between Bacteriophage WO, *Wolbachia*, and Cytoplasmic Incompatibility. PLoS ONE 6:e29106
60. Roberts JW, Roberts CW, Craig NL (1978) *Escherichia coli* recA gene product inactivates phage lambda repressor. *Proceedings of the National Academy of Sciences* 75, 4714–4718
61. Maslowska KH, Makiela-Dzbenska K, Fijalkowska IJ (2019) The SOS system: A complex and tightly regulated response to DNA damage. *Environ Mol Mutagen* 60:368–384
62. Fallon AM (2021) DNA recombination and repair in *Wolbachia*: RecA and related proteins. *Mol Genet Genomics* 296:437–456
63. Madhav M, Parry R, Morgan JAT, James P, Asgari S (2020) *Wolbachia* Endosymbiont of the Horn Fly (*Haematobia irritans irritans*): a Supergroup A Strain with Multiple Horizontally Acquired Cytoplasmic Incompatibility Genes. *Appl Environ Microbiol* 86:e02589–e02519
64. Oswalt K, Oladipupo S, Mendez L, Gillespie JJ, Beckmann J (2025) Origins of cin: lateral gene transfer of cytoplasmic incompatibility nuclease operon to *Orientia tsutsugamushi*. Preprint at <https://www.biorxiv.org/content/10.1101/2025.03.04.641471v1>
65. Bordenstein SR (2024) Isolation of phage WO particles from *Wolbachia*-infected arthropods. In: Fallon AM (ed) *Wolbachia*. Humana, pp 337–348
66. Bushnell B (2014) *BBMap: A Fast, Accurate, Splice-Aware Aligner*. <https://www.osti.gov/biblio/1241166>
67. Eren AM, Vineis JH, Morrison HG, Sogin ML (2013) A Filtering Method to Generate High Quality Short Reads Using Illumina Paired-End Technology. PLoS ONE 8:e66643
68. Li D et al (2016) MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102:3–11
69. Hyatt D et al (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119
70. Kim D, Song L, Breitwieser FP, Salzberg SL (2016) Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res* 26:1721–1729
71. Eren AM et al (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319
72. Alneberg J et al (2014) Binning metagenomic contigs by coverage and composition. *Nat Methods* 11:1144–1146
73. Dainat J, AliNe: (2025) A Flexible and Efficient Nextflow Pipeline for Read Alignment. *J Open Source Softw* 10:7545

Figures

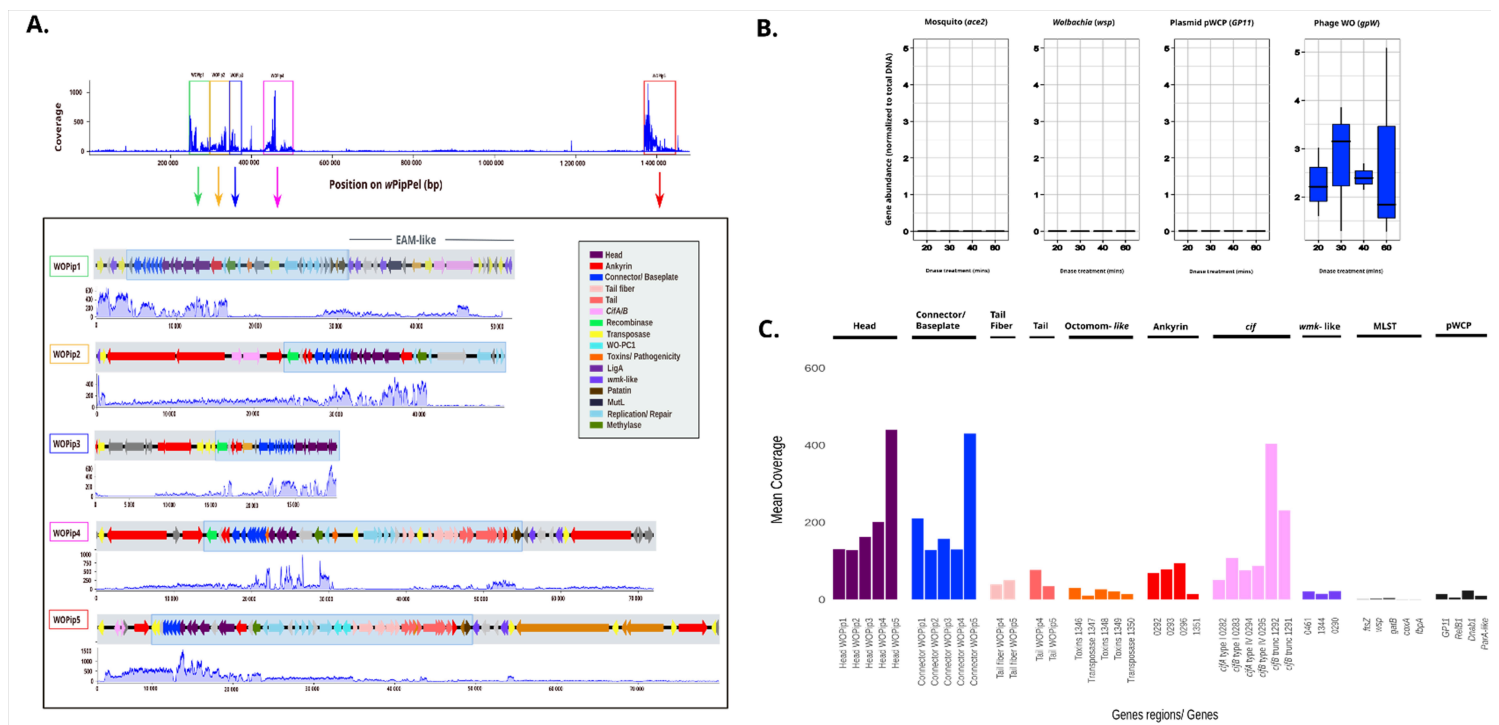


Figure 1

A. Read coverage of isolate sequencing data mapped to the *Wolbachia* wPipPel reference genome, showing the distribution of encapsidated sequences across the genome. Coverage is shown at the genome-wide scale (top) and with a focused view on the five WOPIp regions of wPipPel (bottom). Gene annotations are shown according to Bordenstein and Bordenstein²². **B.** Quantification by qPCR of *Wolbachia* (*wsp*), mosquito host (*ace2*), plasmid pWCP (*GP11*) and WO phage (*gpw*) DNA following phage isolation. **C.** Average read coverage of isolate sequencing data across selected genomic regions of interest, including WO prophage regions, ankyrin, octomom-like, *wmk*-like and CI associated phage genes, *Wolbachia* MLST genes from wPipPel genome, and genes from pWCP plasmid. Of note, each read was assigned to a single genomic position as detailed in Methods; however, reads mapping identical sequences are necessarily distributed across loci, such that coverage reflects the combined representation of shared sequence pools rather than locus-specific enrichment. For example, the truncated *cifB* type I-related locus DUB domain WP1292 is 100% identical to part of a full-length *cifB* type I gene (WP0283), resulting in a partitioning of reads between these loci. Despite the repetitive nature of this locus and independently of any genomic positioning, this domain is nevertheless overrepresented in the isolate.

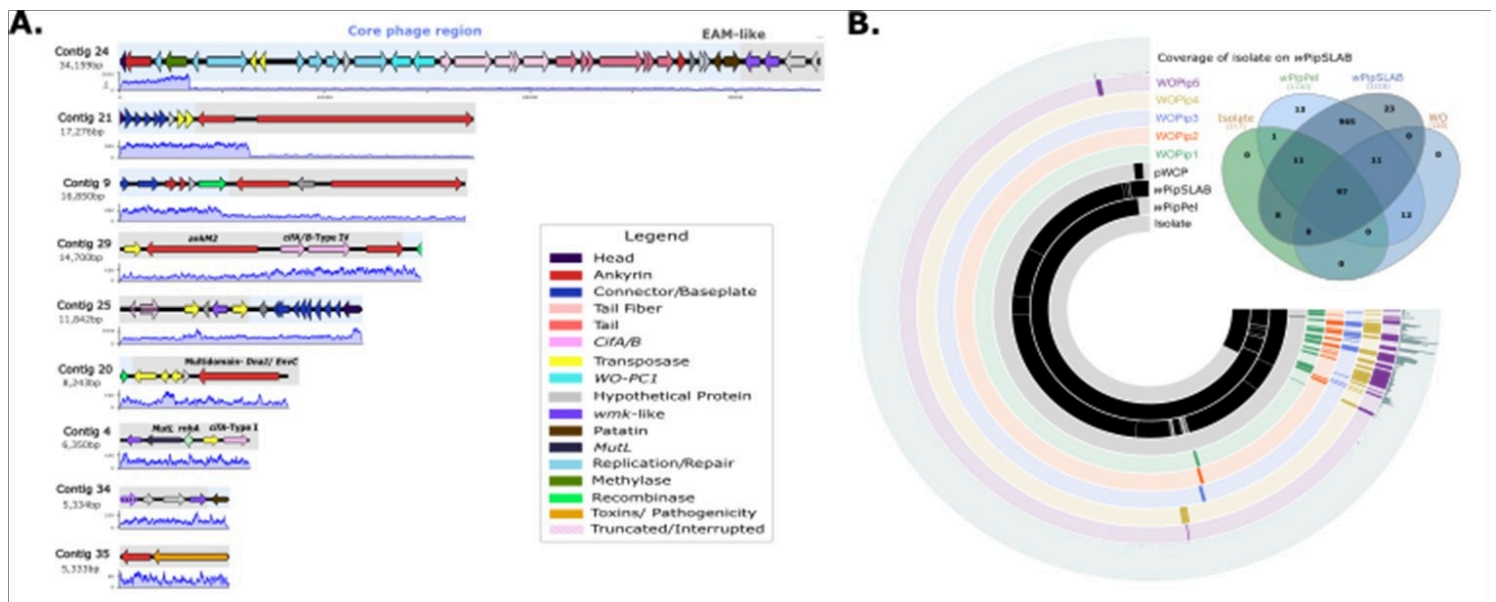


Figure 2

Assembly structure and pangenomic context of encapsidated WO phage sequences.

A. Visualization of the nine longest contigs assembled from encapsidated DNA (out of 37 total contigs), showing gene organization and associated read coverage. Gene are annotated according to the WO prophage annotation described by Bordenstein and Bordenstein (2022)²².

Coverage scales are not identical across contigs and were adjusted individually for visualization purposes. **B.** Pangenome representation showing gene cluster distribution across the encapsidated DNA isolate sequenced in this study and across *Wolbachia* reference genomes and mobile genetic elements. Read coverage from the isolate is mapped onto the reconstructed wPipSLAB reference genome. Additional layers include wPipPel genome, pWCP plasmid, the five annotated WO prophage regions of wPipPel, and the newly reconstructed wPipSLAB genome, providing genomic context for the gene clusters detected in the isolate. A Venn diagram summarizes shared and unique gene clusters among the different annotated genome sequences.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementary19052026Submitted.docx](#)
- [Supplementarytable1.xlsx](#)
- [Supplementarytable2.xlsx](#)
- [Supplementarytable3.xlsx](#)
- [Supplementarytable4.xlsx](#)