Article

Complete De Novo Assembly of *Wolbachia* Endosymbiont of *Frankliniella intonsa*

Zhijun Zhang [1,*] , Jiahui Zhang [1,2], Qizhang Chen [1], Jianyun He [1], Xiaowei Li [1] , Yunsheng Wang [2,*] and Yaobin Lu [1]

1   State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-Products, Institute of Plant Protection and Microbiology, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China; 15096100996zjh@gmail.com (J.Z.); 17300906030@163.com (Q.C.); hjy12242021@163.com (J.H.); lixiaowei1005@163.com (X.L.); lvyb@zaas.ac.cn (Y.L.)
2   Hunan Provincial Key Laboratory for Biology and Control of Plant Diseases and Insect Pests, College of Plant Protection, Hunan Agricultural University, Changsha 410128, China
*   Correspondence: zhangzj@zaas.ac.cn (Z.Z.); wangys@hunau.edu.cn (Y.W.)

Abstract: As an endosymbiont, *Wolbachia* exerts significant effects on the host, including on reproduction, immunity, and metabolism. However, the study of *Wolbachia* in Thysanopteran insects, such as flower thrips *Frankliniella intonsa*, remains limited. Here, we assembled a gap-free looped genome assembly of *Wolbachia* strain *w*FI in a length of 1,463,884 bp (GC content 33.80%), using Nanopore long reads and Illumina short reads. The annotation of *w*FI identified a total of 1838 protein-coding genes (including 85 pseudogenes), 3 ribosomal RNAs (rRNAs), 35 transfer RNAs (tRNAs), and 1 transfer-messenger RNA (tmRNA). Beyond this basic description, we identified mobile genetic elements, such as prophage and insertion sequences (ISs), which make up 17% of the entire *w*FI genome, as well as genes involved in riboflavin and biotin synthesis and metabolism. This research lays the foundation for understanding the nutritional mutualism between *Wolbachia* and flower thrips. It also serves as a valuable resource for future studies delving into the intricate interactions between *Wolbachia* and its host.

Keywords: *Wolbachia*; *Frankliniella intonsa*; symbiosis; phylogeny; comparative genomics

1. Introduction

Wolbachia, gram-negative bacteria, were first discovered in the reproductive tissue of *Culex pipiens* in 1924 and officially named *Wolbachia pipientis* in 1936 [1,2]. These matrilineally inherited endosymbionts are widely distributed among arthropods and nematodes [3]. *Wolbachia* are primarily transmitted in their host by vertical and horizontal modes [4]. Vertical transmission, via cytoplasmic transfer from the mother's germ cell to the offspring, is the primary mode of transmission, while horizontal transmission occurs between different hosts.

The wide distribution of *Wolbachia* can be attributed to its ability to be efficiently transmitted maternally and to manipulate host reproduction in a way that benefits infected females [5]. *Wolbachia* play a critical role in regulating the reproductive phenotype of their host through various mechanisms, including cytoplasmic incompatibility (CI), parthenogenesis induction (PI), feminization (FEM), and male killing (MK). In arthropods, CI is the most common phenotype induced by *Wolbachia*, resulting in embryonic death when uninfected females mate with infected males or males and females carrying incompatible *Wolbachia*. The *cifA* and *cifB* genes associated with phage WO are of significant importance in the study of CI, with their homologs classified as Types I-V [6]. In addition, *Wolbachia* can induce parthenogenesis in haplodiploid insects, leading to the production of female rather than male offspring from unfertilized eggs. Furthermore, *Wolbachia* influences the morphological expression of female characteristics in male offspring and can selectively eliminate developing *Wolbachia*-infected males thereby consequently altering the male–female ratio.

In *Drosophila melanogaster*, the presence of *Wolbachia* has been shown to confer resistance to RNA viruses such as *Drosophila* C virus, Nora virus, and Flock House virus, the first two of which are natural pathogens of *D. melanogaster* [7]. Conversely, when *w*MelPop from *D. melanogaster* was introduced into *Aedes aegypti* mosquitoes, it resulted in a reduction in the lifespan of adult female mosquitoes [8]. In addition, *Wolbachia* has the ability to infect thrips and regulate their reproductive phenotype. For example, it mediates parthenogenesis in *Franklinothrips vespiformis* and induces CI in *Pezothrips kellyanus* [9,10]. Comparative genomic studies have identified several genes associated with CI modification and rescue functions [11]. Remarkable transcriptional differences have been observed in the *cifA* and *cifB* genes during host development, and these genes may be degraded or lost in *Wolbachia* strains that no longer induce CI [12]. While most *Wolbachia* infections in arthropods are considered facultative mutualists, meaning that the host can survive and reproduce without the presence of *Wolbachia*, there are cases where *Wolbachia* act as obligate mutualists. For example, bed bugs rely on *Wolbachia* to provide essential B vitamins for their normal development and reproduction. Removing *Wolbachia* from bed bugs with antibiotics can lead to impaired growth and sterility [13]. Additionally, genome analysis of *Wolbachia* in planthoppers has revealed the presence of intact biotin and riboflavin biosynthetic operons. This suggests that *Wolbachia* have the ability to synthesize biotin and riboflavin, potentially enhancing the reproductive capacity of the host [14].

Initially, the classification of *Wolbachia* was primarily based on the 16S rRNA and *wsp* genes [15,16], which encode *Wolbachia* surface proteins. However, 16S rRNA, with its high conservation and slow evolutionary rate, cannot effectively distinguish closely related strains. In addition, the high recombination and strong diversifying selection of the *wsp* gene has led to potential confusion in strain clustering and introduced bias. To address these issues, MLST was introduced as an effective tool for *Wolbachia* classification [17]. The combination of alleles from five conserved genes (*ftsZ*, *gatB*, *coxA*, *hcpA*, and *fbpA*) in *Wolbachia* was used as the MLST to assign a sequence type (ST) to the *Wolbachia* strain. Each ST represented a unique genetic allelic profile characterizing a particular strain of *Wolbachia*. The relevant information on the five conserved genes and their locus, as well as details of the STs, can be accessed via the PubMLST database [18].

Wolbachia strains from different hosts typically belong to different supergroups. *Wolbachia* of supergroups A and B are known to infect arthropods exclusively, with supergroup A being predominantly found in Diptera and Hymenoptera, whereas supergroup B is mainly found in Lepidoptera [19]. Conversely, supergroups C, D, and J are exclusively associated with *Wolbachia* infecting filarial nematodes, as supported by several studies [20–22]. In addition, supergroup F *Wolbachia* have been found to infect both arthropods and nematodes [23]. Meanwhile, supergroup L *Wolbachia* have been reported to exclusively infect plant-parasitic nematodes [24,25]. Notably, supergroups G and R are no longer considered to be separate entities, with the former considered to be recombinants of supergroup A and B, while the latter is considered to be part of supergroup A [26,27]. Other supergroups of *Wolbachia* are predominantly found in arthropods [28–32]. Supergroup A and B *Wolbachia* have been observed to infect and coexist within the same arthropod hosts, highlighting the irreversible separation of supergroups. Similarly, supergroup B and K have also been reported to coexist within the same host [31,33].

Wolbachia has previously been detected in flower thrips, *F. intonsa*, and the housekeeping genes (*gatB*, *coxA*, *hcpA*, *ftsZ*, and *fbpA*) have been cloned for MLST, resulting in ST code 397 [34]. However, despite this previous work, there has been limited investigation of the relevant genomic information of *Wolbachia* in *F. intonsa* (*w*FI). This study presents the first complete genome of *Wolbachia* in *F. intonsa*, which will serve as a valuable resource for future investigations into the functional role of this endosymbiont in flower thrips.

2. Results

2.1. Assembly and Annotation

We have obtained the genome assembly of the *Wolbachia* strains that infect *F. intonsa*. The genome of *Wolbachia* strain *w*FI was 1,463,884 bp in size and had a GC content of 33.80%. A total of 1877 genes were identified, including 1838 protein-coding genes with 85 candidate pseudogenes, 3 rRNA genes, 35 tRNA genes, and 1 tmRNA (Table 1). The BUSCO completeness scores of the *w*FI genome indicated that it contained 183 complete and single-copy BUSCO groups, 3 complete and duplicated BUSCO groups, 2 fragmented BUSCO groups, and 31 missing BUSCO groups. The BUSCO completeness of the *w*FI genome score of 85.0% was comparable to that of the other two *Wolbachia* complete genomes belonging to supergroup B (Figure S1). The genomic features of *w*FI, including coding sequences (CDS), tRNA, rRNA, and others, were visualized using the CGView Server (Figure 1 and Figure S2). The presence of the irregular GC skew in *w*FI, meant that there was no obvious GC skew to identify a possible origin of replication, which also suggested that *w*FI may have undergone frequent rearrangements. This abnormal GC skew pattern has also been observed in other *Wolbachia* genomes [35–39]. However, it has been observed that *Wolbachia* with reduced genomes exhibit a strong GC skew [40].

Table 1. Key characteristics of *Wolbachia* genome in *F. intonsa* (*w*FI).

| Attribute	*w*FI (GCA_029856955.1)
Size (bp)	1,463,884
Total genes	1877
Pseudogenes *	85
GC (%)	33.80
Proteins	1838
rRNAs	3
tRNAs	35
tmRNA	1
BUSCO Score	85.0%

* Candidate pseudogenes were mentioned in the log file rather than the result file.
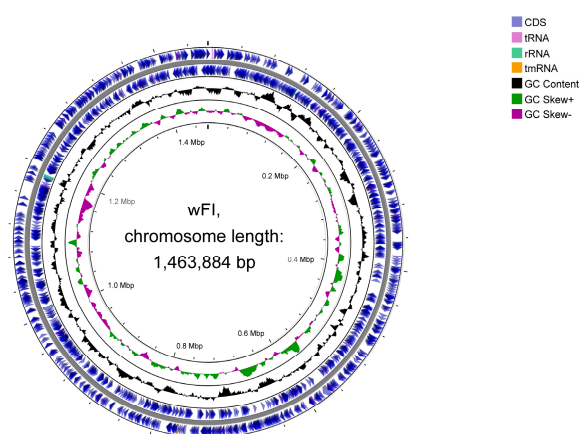
Figure 1. Circular map of the *w*FI genome. The outermost tracks 1 and 2 represent the positions of the CDS, tRNA, rRNA, and tmRNA genes on the positive and negative strands. Tracks 3 and 4 tracks show the GC content and GC skew, respectively.

A total of 1564 protein-coding genes with Clusters of Orthologous Genes (COG) functional annotations were identified in the *w*FI genome by searching against the eggNOG database v5.0.2 [41]. Of these, 713 (45.58%) genes were classified under the functional category of replication, recombination, and repair (Figure 2A, Table S1A). A total of 486 genes were categorized into 32 different Gene Ontology (GO) terms (Figure 2B). These GO terms

were further grouped into three main categories: Biological Process, Cellular Component, and Molecular Function. The Biological Process category consisted of 19 distinct elements with 434 genes. The Cellular Component category included 2 elements, with a total of 401 genes assigned to them. Within the Molecular Function there were 11 elements, with 415 genes assigned to them. Notably, the number of genes associated with cellular process, metabolic process, and the cellular anatomical entity GO term was approximately 400. A total of 540 genes were mapped to 28 different Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Figure 2C). Among these pathways, the Metabolism pathway stood out with 309 genes, representing a significantly higher number of coding genes compared to other pathways. Within the metabolism pathway, several subcategories were identified. These included energy metabolism with 96 genes, metabolism of cofactors and vitamins with 75 genes, and carbohydrate metabolism with 64 genes. Furthermore, the genetic information processing pathway included a total of 172 genes, of which 83 genes were associated with translation, 56 genes were involved in replication and repair processes, 34 genes were associated with folding, sorting and degradation, and 3 genes were involved in transcription.
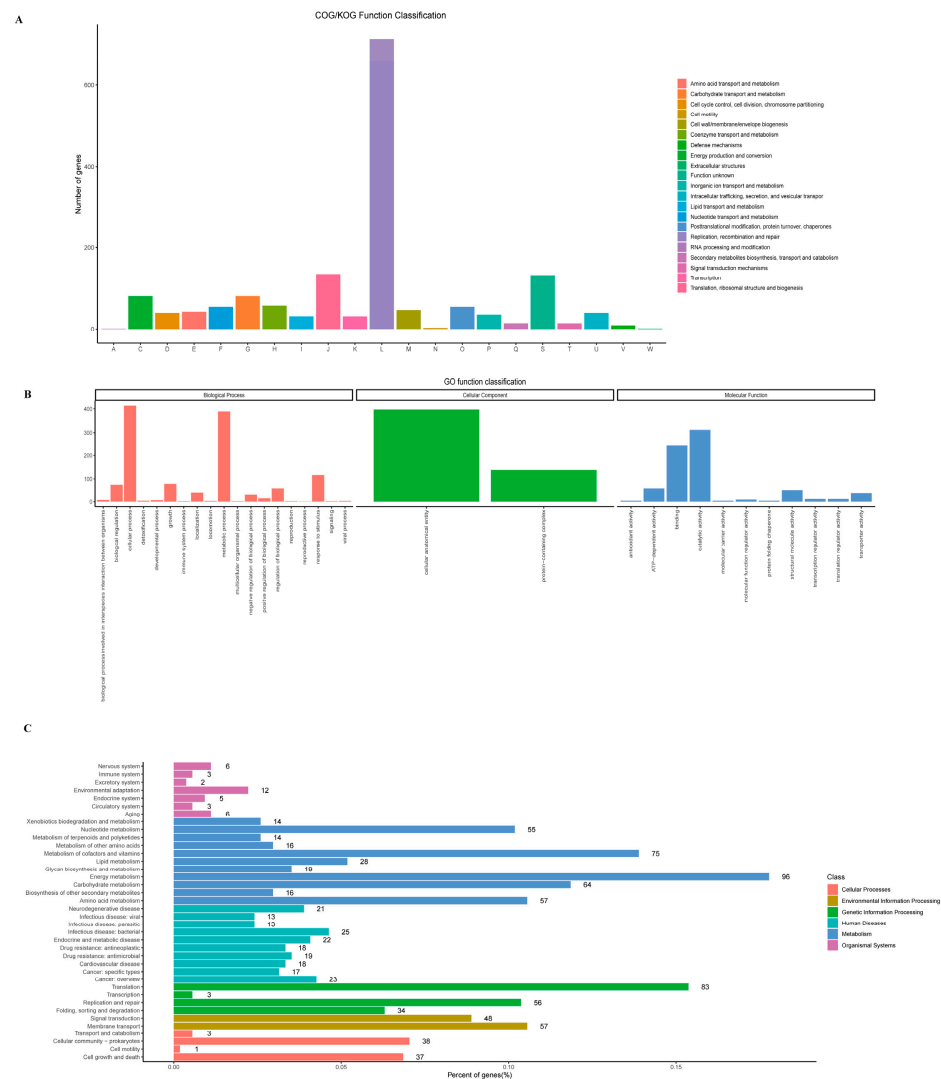
Figure 2. The functional genes annotation of the *w*FI genome. (**A**) COG functional annotation classification statistics. (**B**) GO function classification annotation. (**C**) KEGG pathway classification annotation.

Genome-wide pathway annotation based on the eggNOG database v5.0.2 revealed that the *w*FI genome lacked key genes involved in biotin synthesis, including *bioA*, *bioD*, *bioC*, *bioH*, *bioF*, and *bioB*. However, it contained the *bioY* gene (ko: K03524), which is responsible for biotin transport, and the *birA* gene (ko: K03523), which is involved in the synthesis of essential metabolites using biotin as a precursor. Moreover, we found that, while *w*FI encodes riboflavin synthesis operons, it was not a compact operon but scattered throughout the *Wolbachia* genome, as *ribA*, *ribD*, *ribB*, *ribH*, *ribE*, and *ribF* (Figure S3, Table S1B). The hypothetical phosphatase responsible for coverting 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione5′-phosphate into 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione was absent in Figure S3, which may be due to the fact that it remains uncharacterized in most organisms [42].

Pfam annotation of the *w*FI genome revealed 1352 protein-coding genes containing at least one Pfam domain (Table S2). The most abundant Pfam profile identified was the endonuclease of the DDE superfamily DDE_3 (PF13358). In addition, the domain related to mobile genetic elements, such as DDE_Tnp_1 (PF01609), DDE_Tnp_1_3 (PF13612), DDE_Tnp_1_5 (PF13737), DDE_Tnp_1_IS240 (PF13610), and HTH_Tnp_IS630 (PF01710), were found abundantly in *w*FI genome, which are transposase domains. Moreover, the retroviral integrase domain rve (PF00665), rve_2 (PF13333), rve_3 (PF13683), reverse transcriptase domain RVT_1 (PF00078), and Group II intron reverse transcriptase domain GIIM (PF08388) were identified in the *w*FI genome. Members of the IS3 family have shown remarkable similarity to integrase [43]; this similarity was also shown in *w*FI. The results of the Pfam and IS annotation (Tables S2 and S3) revealed that most of the proteins containing retroviral integrase domains were identified as IS3 family members. The Group II intron encoded proteins with reverse transcriptase, maturase, and endonuclease activities. It was observed that introns could mediate genomic rearrangements in the *Wolbachia* genome [44]. This suggested that intron-mediated genomic rearrangement events might be present in *w*FI.

2.2. Insertion Sequences, Ankyrin Repeat, Type IV Secretory System, and Prophage Genes

ISs are a critical component of bacterial genomes and consist of inverted repeats and transposase-encoding sequences. They are the simplest mobile genetic elements and play a crucial role in bacterial evolution. To date, approximately 20 families of IS have been identified [45]. *Wolbachia* genomes are known to contain numerous ISs, accounting for approximately 10% of the genome, and play a critical role in *Wolbachia* evolution [46,47]. In this study, 587 open reading frames (ORFs) associated with ISs belonging to 15 IS families were identified in the *w*FI genome. Among them, 80 were considered as putative complete ORFs, 410 as putative partial ORFs, and 97 remained uncategorized ORFs. Notably, the number of IS-associated ORFs in *w*FI was higher compared to other supergroup B *Wolbachia* genomes like *w*Di (GCA_019355355.1) and *w*AlbB (GCA_004171285.1) (Table S3). The ISs had a median size of 357 bp, ranging from 93 to 1389 bp, and a total of 254.637 bp, accounting for around 17% of the entire *w*FI genome (Table S3A). Among the ISs, the largest family identified was IS 630 with 400 copies (Table S3B).

The Ankyrin repeat (ANK) protein family has been shown to play a critical role in mediating protein-protein interactions and is of particular interest in the context of host-*Wolbachia* interactions. Examination of Pfam protein domains within the *w*FI genome revealed the presence of 38 proteins containing at least one ANK domain, comprising 61 ANK domains with diverse functions (Table S2).

The Type IV secretion system (T4SS) is a pivotal mechanism used by *Wolbachia* to transfer DNA and/or proteins to eukaryotic cells, which is crucial for successful host infection and proliferation [48,49]. Two operons associated with T4SS have been identified in the *Wolbachia* genome. The first operon comprises *virB8*, *virB9*, *virB10*, *virB11*, *virD4*, and the downstream *wspB* locus, while the second operon includes *virB3*, *virB4*, *virB6*, and several open reading frames (*orf1* to *orf4*) [50]. By combining the prokka, pfam, and eggNOG annotation results and then matching them with NR database, the *w*FI genome

revealed the presence of 13 genes associated with T4SSs (Table S4). These genes were organized into two operons in the *w*FI genome. Furthermore, three duplicated genes, *virB4-2* (*AJACLIMF_00502*), *virB8* (*AJACLIMF_01223*), and *virB9* (*AJACLIMF_01675*), were found scattered elsewhere in the genome.

The majority of the sequenced *Wolbachia* genomes contain the prophage WO, except for those involved in obligate mutualistic relationships [51]. Prophage genes, which are dynamic elements that mediate gene transfer, play a critical role in the evolution and adaptation of the *Wolbachia* genome [52,53]. Notably, these genes are widely distributed throughout the *Wolbachia* genome. The *w*FI genome does not support the notion that each *Wolbachia* genome contains at least one intact prophage WO and usually several degenerate, independently acquired WO prophages [54]. In the *w*FI genome, PHASTER analysis revealed an incomplete prophage region (position: 955,379–965,628) measuring 10.2 kb and consisting of 12 proteins. The incomplete prophage region in *w*FI consisted of two attachment sites (attL and attR), four transposases, a lytic transglycolase, an endonuclease, a transferase, a phage portal protein, one ParE toxin of the type II toxin-antitoxin system, one gpw residue protein, and two uncharacterized proteins. The prophage region of *w*FI, which was incomplete, lacked the modules responsible for encoding head, baseplate, and tail-associated proteins. As a result, it was considered cryptic, meaning that it no longer possessed the ability to form virions and lyse host cells. Additionally, through functional annotation, we identified several prophage-related genes outside of the incomplete prophage region. These genes encode products including phage integrase (*AJACLIMF_00069*, *AJACLIMF_01338*), phage gp6-like head-tail connector protein (*AJACLIMF_01293*), phage tail tube protein (*AJACLIMF_01290*), phage portal protein (*AJACLIMF_01198*, *AJACLIMF_01350*, *AJACLIMF_01351*, *AJACLIMF_01719*), and phage terminase large subunit (*AJACLIMF_01354*) (Figure S2). Prophage was integrated into the *w*FI genome, it could be threatened by selective pressure to lose the gene components required for infection and/or genome replication and scattered on the *w*FI genome as prophage remnants.

2.3. Orthology Analysis and Identification of a Core Proteome across Completed Wolbachia Genomes

Using Orthofinder v2.5, we analyzed the orthology relationships between the *w*FI genome and 25 complete *Wolbachia* genomes (Table S5). Our analysis revealed 1786 orthogroups of 32,733 proteins (Table S6). Among these orthogroups, 604 were shared by all *Wolbachia* genomes, of which 502 were single-copy orthogroups (Table S6A). Furthermore, orthology analysis identified 79 orthogroups containing 379 proteins that were unique to each individual *Wolbachia* genome analyzed. Additionally, 739 genes were not assigned to any orthogroup (Table S6B). For *w*FI, we assigned 1783 (97.0%) of its protein-coding genes to 973 orthogroups, of which 191 (10.4%) were assigned to 21 species-specific orthogroups, whereas 55 protein-coding genes were not assigned to any orthogroup (Table S6C). The assignment rate for protein-coding genes in other *Wolbachia* genomes was higher than 91.4%, the *w*Cle, which had 150 (11.5%) protein-coding genes that were not assigned to any orthogroup (Table S6C). The core proteome, consisting of 604 common orthogroups that included all proteins present in the analyzed genomes, was significantly higher than the number of orthogroups among other species (Figure 3).

A total of 604 common orthogroups were identified, consisting of 16,224 genes. Among the top 10 enriched GO terms, the ATP metabolic process stood out with 460 genes involved, while regulation of translation and regulation of cellular amide metabolic process had the fewest genes, with 361 each (Figure 4A). As for the top 10 enriched KEGG pathways, Protein export showed the highest number of enriched genes (419), closely followed by Mismatch repair (378) and DNA replication (353) (Figure 4B). In addition, a total of 21 *w*FI-specific orthogroups, containing 191 genes were identified, with the majority of these genes being ISs and belonging to COG L genes. During the enrichment analysis, we

found that no genes related to any GO terms were enriched, and none of the genes could be mapped to any specific pathways.
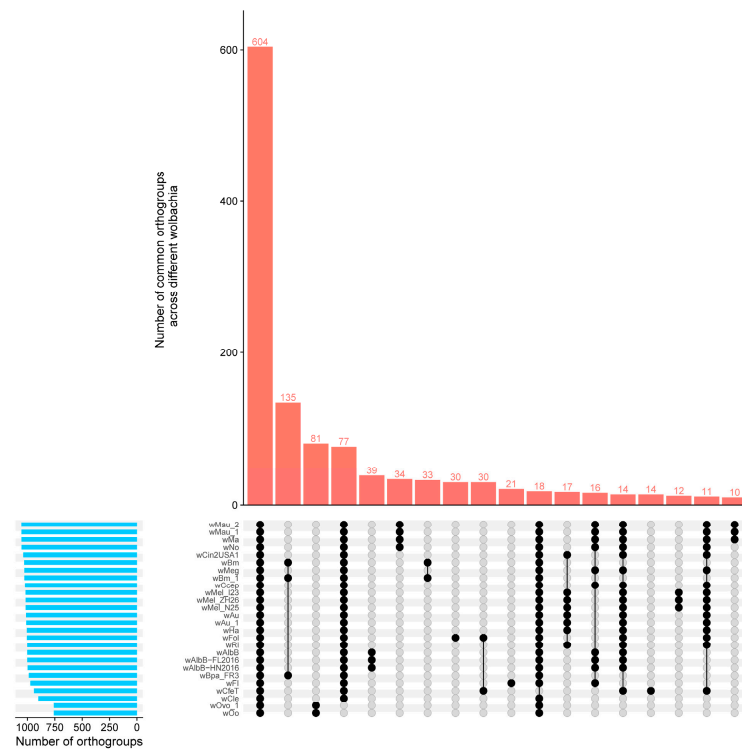
Figure 3. Analysis of protein orthology in complete genomes of *Wolbachia* belonging to supergroups A to F. The blue bar represents the number of orthogroups for the different *Wolbachia* strain. The red bar represents the number of common orthogroups across different *Wolbachia* strains, the first red bar represents 604 orthogroups across the 26 *Wolbachia* strains, the tenth red bar represents 21 *w*FI-specific orthogroups. Black dots indicate the presence, and gray dots indicate the absence of orthogroups in each *Wolbachia* strain.
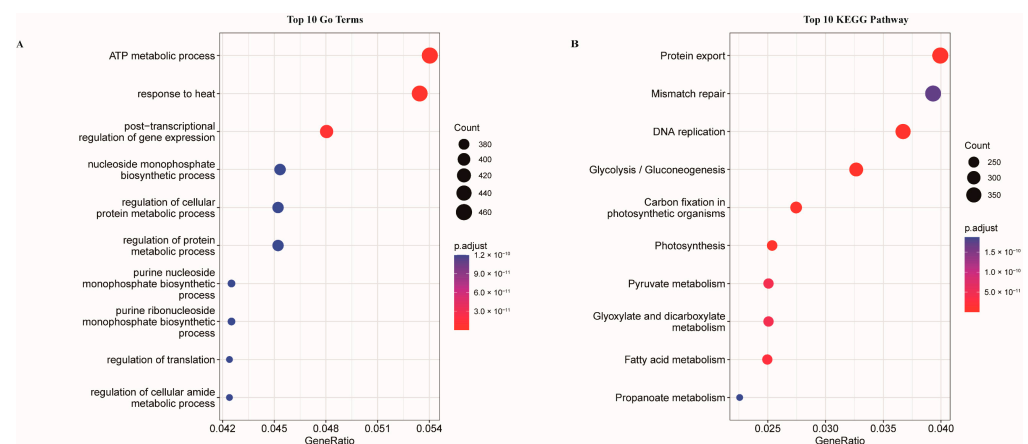

Figure 4. The GO and KEGG enrichment of the common orthogroups. (**A**) GO enrichment. (**B**) KEGG enrichment.

2.4. wFI Belong to Supergroup B

The phylogenetic analysis, based on the concatenated sequences of allele of the MLST locus, revealed that *w*FI (ST-397) formed a distinct clade with other ST codes belonging to supergroup B, indicating its classification within this supergroup (Figure 5). To eliminate the interference caused by abnormal *Wolbachia* genomes, we performed a thorough

examination of the GC content and genome size. As a result, we identified eight candidate abnormal *Wolbachia* genomes (Figure S4). Further investigation in the NCBI led us to exclude three of the eight candidate abnormal *Wolbachia* genomes. Moreover, two of the eight candidate abnormal *Wolbachia* genomes with an unknown host were also excluded (Table S7A). This resulted in a total of 76 complete genomes, 22 chromosome-scale genomes, 48 scaffold-scale genomes, and 1218 contig-scale genomes (Table S7B). The majority of the 1364 *Wolbachia* genomes belonged to strains infecting the hosts *D. melanogaster* and *D. simulans*, both of which are classified under supergroup A. Using these genomes and *w*FI, we constructed a large-scale NJ phylogenetic tree, which showed that *w*FI clustered with the *Wolbachia* genome belonging to supergroup B (Figure S5). The *Wolbachia* from different supergroups can infect the same host [55]. *Wolbachia* infecting *A. aegypti*, *A. albopictus*, *Dactylopius coccus*, *D. simulans*, and *Nasonia vitripennis* could be categorized into two supergroups, A and B (Figure S5), highlighting the distinct and irreversible segregation of supergroups as evidenced by the co-infection of the same host species by *Wolbachia* belonging to different supergroups. We would like to acknowledge the presence of supergroup classification unknowns, namely *w*Pcr (GCA_918697765.1), *w*Pch (GCA_918342435.1), and *w*Cas (GCA_918231965.1), which infect *Phyllotreta cruciferae*, *Psylliodes chrysocephala*, and *Ceutorhynchus assimilis*, respectively. Notably, *w*Pcr and *w*Pch clustered with supergroup A, whereas *w*Cas clustered with supergroup B, indicating their potential supergroup classification (Figure S5). We used concatenated protein sequences of single-copy genes from 26 complete *Wolbachia* genomes to construct a maximum-likelihood phylogenetic tree using IQ-TREE v2.2.0.3. The resulting tree indicated that the *w*FI formed a clade with the supergroup B *Wolbachia* strains (Figure 6).
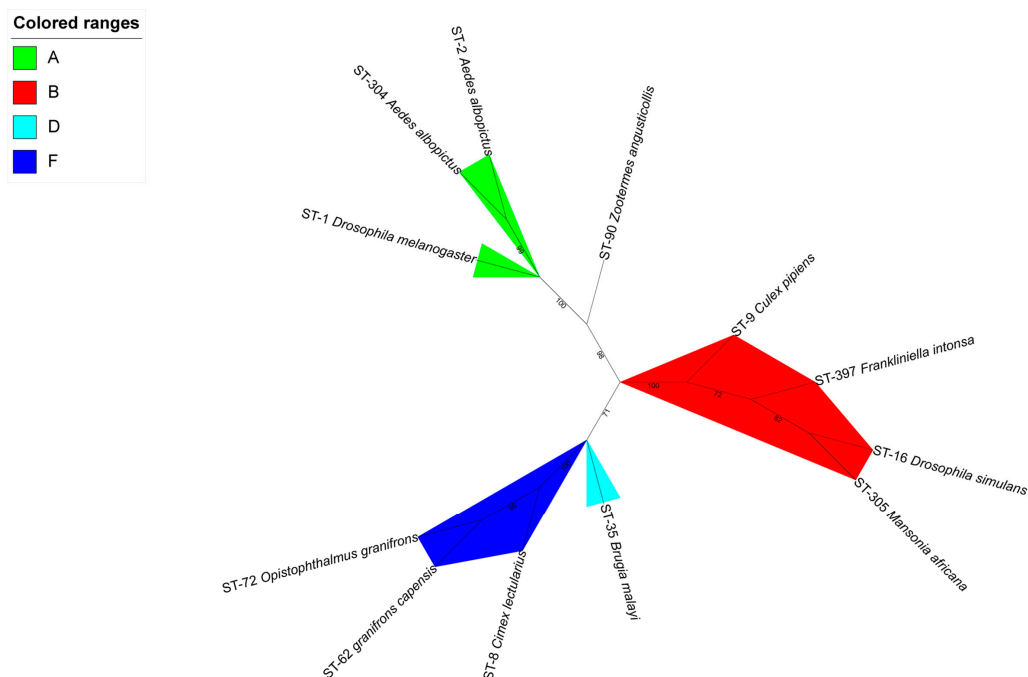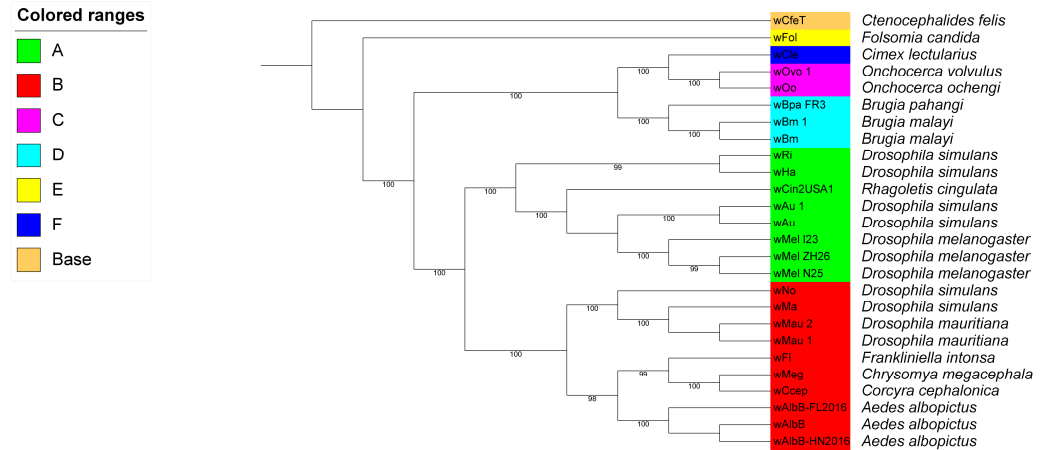

Figure 5. Maximum-likelihood phylogenetic tree of *Wolbachia* MLST from supergroups A, B, D, F, and H. The tree was generated using IQ-TREE v2.2.0.3 based on five MLST genes using ultrafast bootstrap mode with 5000 iterations. The amino acid substitution model used was GTR + F + G4. The *Wolbachia* supergroups are color-coded and shown in colored ranges.

In prokaryotes, a commonly used threshold for defining species is a 95% average nucleotide identity (ANI), and an ANI greater than 95% usually indicates that organisms belong to the same species. Conversely, the ANI between different species is typically less than 83% [56]. Notably, the ANI between *w*CfeT and other *Wolbachia* genomes was approximately 78% (Figure 7), indicating that *w*CfeT may have ancestral origins in most

other *Wolbachia lineages* [57]. The ANI between *w*FI and *Wolbachia* in supergroup B was approximately 95% (Figure 7), providing compelling evidence that *w*FI belongs to this particular supergroup, which is typically found in arthropods [28].

Figure 6. Maximum-likelihood phylogeny of the complete genomes of *Wolbachia* strains belonging to supergroups A to F, with *w*CfeT set as the outgroup. The tree was constructed using IQ-TREE v2.2.0.3 with ultrafast bootstrap mode, 5000 iterations. Branch support was estimated using a Shimodaira–Hasegawa (SH)-like approximate likelihood ratio test with 1000 replicates. The amino acid substitution model used was FLAVI + F + I + I + R4. Bootstrap values > 95% are shown at each node. The *Wolbachia* supergroups are color coded and shown in colored ranges.
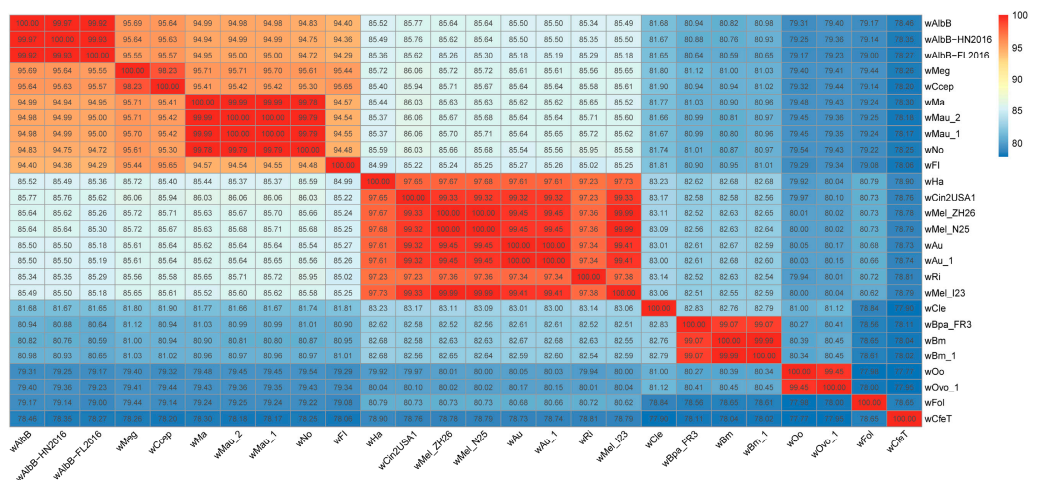
Figure 7. Heat map showing the ANI between 26 complete genomes of *Wolbachia* strains belonging to supergroups A to F. ANI were calculated using fastANI v1.33 with default parameters and visualized using the R package pheatmap v1.0.12.

3. Discussion

ISs have been shown to cause large-scale rearrangements in the genome of *Escherichia coli* [58]. In the case of *w*FI, ISs represented a significant proportion, accounting for 17% of the entire genome. Notably, a considerable number of COG L genes, involved in DNA replication, recombination, and repair, were associated with IS elements (Figure S2). The prevalence of ISs in *w*FI suggested that ISs may have caused rearrangements in the *w*FI genome, potentially contributing to the abnormal GG skew observed.

Biotin is a part of the vitamin B family and is not normally synthesized by insects. In many cases, insects rely on microbial symbionts to meet their vitamin B requirements [59]. Studies have shown that *w*Cle (GCA_000829315.1) possesses a complete biotin pathway for vitamin B7, which is essential for the growth and reproduction of its host, the bed

bug [60]. Additionally, a complete vitamin B7 synthesis operon was discovered in *w*Nfla (GCA_001675695.1) and *w*Nleu (GCA_001675715.1) [61], while the complete biotin and riboflavin biosynthesis operons were identified in *w*Lug (GCA_007115045.1) and *w*Stri (GCA_001637495.1) [14]. A comprehensive genomic survey investigating the B vitamin synthetic capabilities of various insect-associated *Wolbachia* strains revealed that the riboflavin synthesis pathway was the only one highly conserved pathway [62]. According to the EggNOG annotation results, the biotin synthesis pathway was incomplete in *w*FI. However, despite this deficiency, the *w*FI genome contained the genes *bioY* and *birA*, suggesting that *w*FI could potentially interact with its host or other endosymbionts to provide biotin to the host. The *w*FI possessed a complete pathway of riboflavin, suggesting that *w*FI could have provided riboflavin to the host as a mutualist (Figure S3).

The large-scale NJ tree was able to classify several *Wolbachia* genomes whose supergroup classification was previously unknown. Specifically, *w*Pcr and *w*Pch clustered with supergroup A, whereas *w*Cas clustered with supergroup B (Figure S5). Notably, although WOLB1015 (GCA_902646985.2) and *w*MoviF (GCA_023661065.1) belong to supergroup F and are both presented in the insect *Melophagus ovinus*, they did not cluster together in the same clade in the NJ tree and had an ANI of approximately 99%. In addition, the *w*Chem (GCA_014771645.1) from supergroup T was found to be closely related to *Wolbachia* strains supergroups F, S and D (Figure S5), which is consistent with previous research [63]. The mashtree v1.2.0 [64] allows the rapid clustering of *Wolbachia* genomes into NJ trees, which can provide a first indication of the supergroup classification of unknown *Wolbachia* genomes, and help to identify misclassifications or relationships between supergroups. The classification results obtained using mashtree v1.2.0 [64], *w*FI belong to supergroup B, which is consistent with the results of a previous MLST analysis [34].

*w*FI belong to supergroup B by using mashtree v1.2.0 [64] is consistent of previous MLST analysis result [34].

4. Materials and Methods

4.1. Genome Assembly

In 2017, adults of *F. intonsa* were collected from *Capsicum annuum* L. plants in Jiaxing, Zhejiang, China (30.75° N, 120.79° E). These insects were subsequently reared on fresh cowpea (*Vigna unguiculata* ssp. *Sesquipedalis*) under controlled conditions at 25 ± 1 °C with a 16 h light cycle. A total of approximately 200 adult *F. intonsa*, representing a mixture of ages, were selected from the laboratory population and subjected to decontamination procedures. Briefly, the adult *F. intonsa* were immersed in 1% sodium hypochlorite solution for 5 min, followed by rinsing with sterile water, then immersion in 70% ethanol, and finally another rinse with sterile water. Subsequently, the samples were rapidly frozen using liquid nitrogen and preserved at −80 °C. The *F. intonsa* genomic DNA was extracted and purified using the QIAGEN DNA tissue kit (QIAGEN 69506, Hilden, Germany), and prepared for sequencing libraries following the manufacturer's guidelines for sequencing technology (Nextomics Biosciences Co., Ltd., Wuhan, China). Long DNA fragments were sequenced on the Oxford Nanopore PromethION platform, and short-read sequencing was performed on the Illumina NovaSeq 6000 platform.

The *F. intonsa* genome assembly was performed using NextDenovo v2.5.0 [65] with clean Nanopore reads. The assembled sequences were further corrected using NextPolish v1.4.0 [66] with long sequencing reads and Illumina short reads in two iterations, following default parameters. A careful analysis of the *F. intonsa* genome was performed to identify any potential bacterial contamination. During the investigation, a particular looped contig raised suspicion. To gain a clearer understanding, manual binning was carried out using Anvi'o v7 [67], which definitively confirmed that the looped contig corresponded to the *Wolbachia* endosymbiont of *F. intonsa*, specifically identified as *w*FI.

4.2. Genome Annotations and Assessments

The identification and annotation of genetic elements are critical for understanding the functional properties of a genome. Here, we employed several bioinformatics tools to identify protein-coding genes as well as non-coding RNA components, including 5S rRNA, 16S rRNA, 23S rRNA, tRNA, and tmRNA. Specifically, we used prokka v1.14.6 [68] to identify these genetic elements with the default parameters. It is important to note that information regarding candidate pseudogenes was mentioned in the log file rather in the result file. To evaluate the completeness of genome, we employed Benchmarking Universal Single-Copy Orthologs (BUSCO) v5.4.3, based on the Proteobacteria_odb10 database [69]. Functional annotation of protein-coding genes was performed by searching against the eggNOG database v5.0.2 [41] using eggNOG-Mapper v2.1.9 [70]. The annotation was performed using the—tax_scope Bacteria parameter to focus on bacterial annotations specifically. Additionally, Pfam domains were annotated using the pfam_scan.pl v1.6 script to search against Pfam database v35.0 [71]. In addition, we employed the PHASTER web server [72] to identify prophage regions within the genome. Furthermore, the ISsaga web server [73] was utilized to identify IS elements presented in the genome. Finally, we used the CGview Server [74] to represent the IS elements, ANK, T4SS, and prophage regions in a circle map.

4.3. Comparative Genomics Analysis

The complete assembly of *w*FI was compared to 25 other complete genomes, as listed in Table S5. All genomes were re-annotated using prokka v1.14.6 [68] with the default parameters. Functional annotation of the protein-coding genes was conducted by searching against the eggNOG database v5.0.2 [41] using eggNOG-Mapper v2.1.9 [70] with the -tax_scope Bacteria parameter. The ANI between *w*FI and 25 complete genomes was calculated using fastANI v1.33 [56]. Orthogroups shared between wFI and 25 complete genomes were identified using Orthofinder v2.5.4 [75], and the common orthogroups across multiple genomes were visualized as UpSet plots [76] using the R package ComplexUpset v1.3.3 [77]. Enrichment analysis for common orthogroups was performed using clusterProfiler V4.6.2 [78].

4.4. Phylogenetic Analysis

A total of 12 STs, derived from *Wolbachia* strains belonging to supergroups A, B, D, F, and H, were analyzed (Table S8). To construct a robust phylogenetic tree, concatenated allelic sequences of the five MLST genes were retrieved from the PubMLST database [18] and aligned using Muscle v5.1 [79]. Poorly aligned regions were then trimmed using Trimal v1.4.1 [80]. The resulting alignment was then subjected to maximum-likelihood analysis using IQ-TREE v2.2.0.3 [81] with ultrafast bootstrap mode and 5000 iterations. The resulting tree was used to classify the *w*FI supergroup. Finally, the maximum-likelihood phylogenetic tree was visualized using ITOL v6 [82].

The NCBI dataset command line tools were used to obtain all *Wolbachia* genome sequences from the GenBank database. Scatter plots based on GC content and genome size were used to identify candidate abnormal *Wolbachia* genomes, and then candidate abnormal *Wolbachia* genomes were further investigated by searching NCBI to filter out abnormal *Wolbachia* genomes. The mashtree v1.2.0 [64] was used to construct a large-scale NJ phylogenetic tree of the normal *Wolbachia* genomes (Table S7B). The ANI between *w*FI and other *Wolbachia* genomes was calculated using fastANI v1.33 [56], and the results were visualized using the Interactive Tree Of Life (ITOL) v6 [82].

We selected *w*FI and 25 complete *Wolbachia* genomes from the A–F supergroups (Table S5). Multiple sequence alignments for single-copy genes were performed using Muscle v5.1 [79]. Poorly aligned regions were further trimmed using Trimal v1.4.1 [80]. We used ModelFinder [83] to obtain the best amino acid substitution model based on Bayesian Information Criteria. The phylogenetic tree was constructed using IQ-TREE v2.2.0.3 [81] with ultrafast bootstrap mode and 5000 iterations; *w*CfeT was used as

an outgroup. Branch support was estimated using a Shimodaira–Hasegawa (SH)-like approximate likelihood ratio test with 1000 replicates. Finally, the maximum-likelihood phylogenetic tree was visualized using the ITOL v6 [82]. We also used the R package pheatmap v1.0.12 [84] to display the ANI between 26 complete *Wolbachia* genomes with a heat map.

5. Conclusions

We have successfully provided the complete assembly of the *Wolbachia* genome infecting *F. intonsa*, which belongs to supergroup B. Our genomic analyses have revealed several essential features, including mobile genetic elements such as prophage and ISs, as well as genes related to ANK, T4SS, riboflavin, and biotin synthesis and metabolism. The discovery of a complete riboflavin pathway in *w*FI suggested possibilities for meeting the riboflavin needs of the host. Additionally, the identification of an incomplete biotin synthesis pathway indicated potential interactions between *w*FI and the host or other endosymbionts to meet the host's biotin needs. The first complete genome of the *Wolbachia* endosymbiont of *F. intonsa* provides a valuable resource for future investigations of *Wolbachia*–host interactions, comparative genomics of *Wolbachia*, and phylogenetic relationships between different supergroups of *Wolbachia*.

References

1. Hertig, M.; Wolbach, S.B. Studies on Rickettsia-Like Micro-Organisms in Insects. *J. Med. Res.* **1924**, *44*, 329–374.7. [PubMed]
2. Hertig, M. The Rickettsia, *Wolbachia Pipientis* (Gen. et Sp.n.) and Associated Inclusions of the Mosquito. *Culex Pipiens. Parasitol.* **1936**, *28*, 453–486. [CrossRef]
3. Stouthamer, R.; Breeuwer, J.A.J.; Hurst, G.D.D. *Wolbachia Pipientis*: Microbial Manipulator of Arthropod Reproduction. *Annu. Rev. Microbiol.* **1999**, *53*, 71–102. [CrossRef] [PubMed]
4. Werren, J.H. Biology of *Wolbachia*. *Annu. Rev. Entomol.* **1997**, *42*, 587–609. [CrossRef] [PubMed]
5. Serbus, L.R.; Casper-Lindley, C.; Landmann, F.; Sullivan, W. The Genetics and Cell Biology of *Wolbachia*-Host Interactions. *Annu. Rev. Genet.* **2008**, *42*, 683–707. [CrossRef] [PubMed]
6. Martinez, J.; Klasson, L.; Welch, J.J.; Jiggins, F.M. Life and Death of Selfish Genes: Comparative Genomics Reveals the Dynamic Evolution of Cytoplasmic Incompatibility. *Mol. Biol. Evol.* **2021**, *38*, 2–15. [CrossRef] [PubMed]
7. Teixeira, L.; Ferreira, Á.; Ashburner, M. The Bacterial Symbiont Wolbachia Induces Resistance to RNA Viral Infections in Drosophila Melanogaster. *PLoS Biol.* **2008**, *6*, e1000002. [CrossRef] [PubMed]
8. McMeniman, C.J.; Lane, R.V.; Cass, B.N.; Fong, A.W.C.; Sidhu, M.; Wang, Y.-F.; O'Neill, S.L. Stable Introduction of a Life-Shortening *Wolbachia* Infection into the Mosquito *Aedes aegypti*. *Science* **2009**, *323*, 141–144. [CrossRef]
9. Arakaki, N.; Miyoshi, T.; Noda, H. *Wolbachia*–Mediated Parthenogenesis in the Predatory Thrips *Franklinothrips Vespiformis* (Thysanoptera: Insecta). *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **2001**, *268*, 1011–1016. [CrossRef]

10. Nguyen, D.T.; Morrow, J.L.; Spooner-Hart, R.N.; Riegler, M. Independent Cytoplasmic Incompatibility Induced by *Cardinium* and *Wolbachia* Maintains Endosymbiont Coinfections in Haplodiploid Thrips Populations. *Evolution* **2017**, *71*, 995–1008. [CrossRef]

11. Baião, G.C.; Janice, J.; Galinou, M.; Klasson, L. Comparative Genomics Reveals Factors Associated with Phenotypic Expression of *Wolbachia*. *Genome Biol Evol.* **2021**, *13*, evab111. [CrossRef] [PubMed]

12. Lindsey, A.R.I.; Rice, D.W.; Bordenstein, S.R.; Brooks, A.W.; Bordenstein, S.R.; Newton, I.L.G. Evolutionary Genetics of Cytoplasmic Incompatibility Genes *CifA* and *CifB* in Prophage WO of *Wolbachia*. *Genome Biol. Evol.* **2018**, *10*, 434–451. [CrossRef]

13. Hosokawa, T.; Koga, R.; Kikuchi, Y.; Meng, X.-Y.; Fukatsu, T. *Wolbachia* as a Bacteriocyte-Associated Nutritional Mutualist. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 769–774. [CrossRef] [PubMed]

14. Ju, J.-F.; Bing, X.-L.; Zhao, D.-S.; Guo, Y.; Xi, Z.; Hoffmann, A.A.; Zhang, K.-J.; Huang, H.-J.; Gong, J.-T.; Zhang, X.; et al. *Wolbachia* Supplement Biotin and Riboflavin to Enhance Reproduction in Planthoppers. *ISME. J.* **2020**, *14*, 676–687. [CrossRef] [PubMed]

15. O'Neill, S.L.; Giordano, R.; Colbert, A.M.; Karr, T.L.; Robertson, H.M. 16S RRNA Phylogenetic Analysis of the Bacterial Endosymbionts Associated with Cytoplasmic Incompatibility in Insects. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 2699–2702. [CrossRef] [PubMed]

16. Zhou, W.; Rousset, F.; O'Neill, S. Phylogeny and PCR–Based Classification of *Wolbachia* Strains Using Wsp Gene Sequences. *Proc. R. Soc. Lond. B Biol. Sci.* **1998**, *265*, 509–515. [CrossRef]

17. Baldo, L.; Dunning Hotopp, J.C.; Jolley, K.A.; Bordenstein, S.R.; Biber, S.A.; Choudhury, R.R.; Hayashi, C.; Maiden, M.C.J.; Tettelin, H.; Werren, J.H. Multilocus Sequence Typing System for the Endosymbiont Wolbachia Pipientis. *Appl. Environ. Microbiol.* **2006**, *72*, 7098–7110. [CrossRef] [PubMed]

18. Jolley, K.A.; Bray, J.E.; Maiden, M.C.J. Open-Access Bacterial Population Genomics: BIGSdb Software, the PubMLST.Org Website and Their Applications. *Wellcome Open Res.* **2018**, *3*, 124. [CrossRef]

19. Vancaester, E.; Blaxter, M. Phylogenomic Analysis of *Wolbachia* Genomes from the Darwin Tree of Life Biodiversity Genomics Project. *PLoS Biol.* **2023**, *21*, e3001972. [CrossRef]

20. Bandi, C.; Anderson, T.J.C.; Genchi, C.; Blaxter, M.L. Phylogeny of *Wolbachia* in Filarial Nematodes. *Proc. R. Soc. Lond. B Biol. Sci.* **1998**, *265*, 2407–2413. [CrossRef]

21. Casiraghi, M.; Bain, O.; Guerrero, R.; Martin, C.; Pocacqua, V.; Gardner, S.L.; Franceschi, A.; Bandi, C. Mapping the Presence of *Wolbachia Pipientis* on the Phylogeny of Filarial Nematodes: Evidence for Symbiont Loss during Evolution. *Int. J. Parasitol.* **2004**, *34*, 191–203. [CrossRef] [PubMed]

22. Lefoulon, E.; Bain, O.; Makepeace, B.L.; d'Haese, C.; Uni, S.; Martin, C.; Gavotte, L. Breakdown of Coevolution between Symbiotic Bacteria *Wolbachia* and Their Filarial Hosts. *PeerJ* **2016**, *4*, e1840. [CrossRef] [PubMed]

23. Mahmood, S.; Nováková, E.; Martinů, J.; Sychra, O.; Hypša, V. Supergroup F *Wolbachia* with Extremely Reduced Genome: Transition to Obligate Insect Symbionts. *Microbiome* **2023**, *11*, 22. [CrossRef] [PubMed]

24. Brown, A.M.V.; Wasala, S.K.; Howe, D.K.; Peetz, A.B.; Zasada, I.A.; Denver, D.R. Genomic Evidence for Plant-Parasitic Nematodes as the Earliest *Wolbachia* Hosts. *Sci. Rep.* **2016**, *6*, 34955. [CrossRef] [PubMed]

25. Haegeman, A.; Vanholme, B.; Jacob, J.; Vandekerckhove, T.T.M.; Claeys, M.; Borgonie, G.; Gheysen, G. An Endosymbiotic Bacterium in a Plant-Parasitic Nematode: Member of a New *Wolbachia* Supergroup. *Int. J. Parasitol.* **2009**, *39*, 1045–1054. [CrossRef] [PubMed]

26. Baldo, L.; Werren, J.H. Revisiting *Wolbachia* Supergroup Typing Based on WSP: Spurious Lineages and Discordance with MLST. *Curr. Microbiol.* **2007**, *55*, 81–87. [CrossRef] [PubMed]

27. Gerth, M. Classification of *Wolbachia* (Alphaproteobacteria, Rickettsiales): No Evidence for a Distinct Supergroup in Cave Spiders 2016. *bioRxiv* **2016**, bioRxiv:046169.

28. Lo, N.; Casiraghi, M.; Salati, E.; Bazzocchi, C.; Bandi, C. How Many *Wolbachia* Supergroups Exist? *Mol. Biol. Evol.* **2002**, *19*, 341–346. [CrossRef]

29. Lo, N.; Paraskevopoulos, C.; Bourtzis, K.; O'Neill, S.L.; Werren, J.H.; Bordenstein, S.R.; Bandi, C. Taxonomic Status of the Intracellular Bacterium *Wolbachia Pipientis*. *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 654–657. [CrossRef]

30. Bordenstein, S.; Rosengaus, R.B. Discovery of a Novel *Wolbachia* Supergroup in Isoptera. *Curr. Microbiol.* **2005**, *51*, 393–398. [CrossRef]

31. Ros, V.I.D.; Fleming, V.M.; Feil, E.J.; Breeuwer, J.A.J. How Diverse Is the Genus *Wolbachia*? Multiple-Gene Sequencing Reveals a Putatively New *Wolbachia* Supergroup Recovered from Spider Mites (Acari: Tetranychidae). *Appl. Environ. Microbiol.* **2009**, *75*, 1036–1043. [CrossRef] [PubMed]

32. Glowska, E.; Dragun-Damian, A.; Dabert, M.; Gerth, M. New *Wolbachia* Supergroups Detected in Quill Mites (Acari: Syringophilidae). *Infect. Genet. Evol.* **2015**, *30*, 140–146. [CrossRef] [PubMed]

33. Ellegaard, K.M.; Klasson, L.; Näslund, K.; Bourtzis, K.; Andersson, S.G.E. Comparative Genomics of *Wolbachia* and the Bacterial Species Concept. *PLoS Genet.* **2013**, *9*, e1003381. [CrossRef] [PubMed]

34. Heng, L.; YaTing, C.; YanKai, Z.; XianMing, Y.; JingTao, S.; XiaoYue, H. *Wolbachia* infection and its relationship with mtDNA diversity in the flower thrips, *Frankliniella intonsa* (Thysanoptera: Thripidae). *Acta Entomol. Sinica* **2015**, *58*, 750–760.

35. Wu, M.; Sun, L.V.; Vamathevan, J.; Riegler, M.; Deboy, R.; Brownlie, J.C.; McGraw, E.A.; Martin, W.; Esser, C.; Ahmadinejad, N.; et al. Phylogenomics of the Reproductive Parasite *Wolbachia* Pipientis WMel: A Streamlined Genome Overrun by Mobile Genetic Elements. *PLoS Biol.* **2004**, *2*, e69. [CrossRef] [PubMed]

36. Klasson, L.; Walker, T.; Sebaihia, M.; Sanders, M.J.; Quail, M.A.; Lord, A.; Sanders, S.; Earl, J.; O'Neill, S.L.; Thomson, N.; et al. Genome Evolution of *Wolbachia* Strain WPip from the Culex Pipiens Group. *Mol. Biol. Evol.* **2008**, *25*, 1877–1887. [CrossRef] [PubMed]

37. Penz, T.; Schmitz-Esser, S.; Kelly, S.E.; Cass, B.N.; Müller, A.; Woyke, T.; Malfatti, S.A.; Hunter, M.S.; Horn, M. Comparative Genomics Suggests an Independent Origin of Cytoplasmic Incompatibility in Cardinium Hertigii. *PLOS Genet.* **2012**, *8*, e1003012. [CrossRef] [PubMed]

38. Neupane, S.; Bonilla, S.I.; Manalo, A.M.; Pelz-Stelinski, K.S. Complete de Novo Assembly of *Wolbachia* Endosymbiont of *Diaphorina Citri* Kuwayama (Hemiptera: Liviidae) Using Long-Read Genome Sequencing. *Sci. Rep.* **2022**, *12*, 125. [CrossRef]

39. Muro, T.; Hikida, H.; Fujii, T.; Kiuchi, T.; Katsuma, S. Two Complete Genomes of Male-Killing *Wolbachia* Infecting Ostrinia Moth Species Illuminate Their Evolutionary Dynamics and Association with Hosts. *Microb. Ecol.* **2023**. [CrossRef]

40. Comandatore, F.; Cordaux, R.; Bandi, C.; Blaxter, M.; Darby, A.; Makepeace, B.L.; Montagna, M.; Sassera, D. Supergroup C *Wolbachia*, Mutualist Symbionts of Filarial Nematodes, Have a Distinct Genome Structure. *Open Biol.* **2015**, *5*, 150099. [CrossRef]

41. Huerta-Cepas, J.; Szklarczyk, D.; Heller, D.; Hernández-Plaza, A.; Forslund, S.K.; Cook, H.; Mende, D.R.; Letunic, I.; Rattei, T.; Jensen, L.J.; et al. EggNOG 5.0: A Hierarchical, Functionally and Phylogenetically Annotated Orthology Resource Based on 5090 Organisms and 2502 Viruses. *Nucleic Acids Res.* **2019**, *47*, D309–D314. [CrossRef] [PubMed]

42. García-Angulo, V.A. Overlapping Riboflavin Supply Pathways in Bacteria. *Crit Rev Microbiol.* **2017**, *43*, 196–209. [CrossRef]

43. Fayet, O.; Ramond, P.; Polard, P.; Prère, M.F.; Chandler, M. Functional Similarities between Retroviruses and the IS3 Family of Bacterial Insertion Sequences? *Mol. Microbiol.* **1990**, *4*, 1771–1777. [CrossRef]

44. Leclercq, S.; Giraud, I.; Cordaux, R. Remarkable Abundance and Evolution of Mobile Group II Introns in *Wolbachia* Bacterial Endosymbionts. *Mol. Biol. Evol.* **2011**, *28*, 685–697. [CrossRef] [PubMed]

45. Chandler, M.; Mahillon, J. Insertion Sequences Revisited. In *Mobile DNA II*; Craig, N.L., Craigie, R., Gellert, M., Lambowitz, A.M., Eds.; ASM Press: Washington, DC, USA, 2007; pp. 303–366. ISBN 978-1-68367-415-3.

46. Klasson, L.; Westberg, J.; Sapountzis, P.; Näslund, K.; Lutnaes, Y.; Darby, A.C.; Veneti, Z.; Chen, L.; Braig, H.R.; Garrett, R.; et al. The Mosaic Genome Structure of the *Wolbachia w*Ri Strain Infecting *Drosophila Simulans*. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5725–5730. [CrossRef] [PubMed]

47. Cerveau, N.; Leclercq, S.; Leroy, E.; Bouchon, D.; Cordaux, R. Short- and Long-Term Evolutionary Dynamics of Bacterial Insertion Sequences: Insights from *Wolbachia* Endosymbionts. *Genome Biol. Evol.* **2011**, *3*, 1175–1186. [CrossRef] [PubMed]

48. Pichon, S.; Bouchon, D.; Cordaux, R.; Chen, L.; Garrett, R.A.; Grève, P. Conservation of the Type IV Secretion System throughout *Wolbachia* Evolution. *Biochem. Biophys. Res. Commun.* **2009**, *385*, 557–562. [CrossRef] [PubMed]

49. Sinha, A.; Li, Z.; Sun, L.; Carlow, C.K.S. Complete Genome Sequence of the *Wolbachia w*AlbB Endosymbiont of *Aedes Albopictus*. *Genome Biol. Evol.* **2019**, *11*, 706–720. [CrossRef]

50. Rancès, E.; Voronin, D.; Tran-Van, V.; Mavingui, P. Genetic and Functional Characterization of the Type IV Secretion System in *Wolbachia*. *J. Bacteriol.* **2008**, *190*, 5020–5030. [CrossRef]

51. Bordenstein, S.R.; Bordenstein, S.R. Eukaryotic Association Module in Phage WO Genomes from *Wolbachia*. *Nat. Commun.* **2016**, *7*, 13155. [CrossRef]

52. Masui, S.; Kuroiwa, H.; Sasaki, T.; Inui, M.; Kuroiwa, T.; Ishikawa, H. Bacteriophage WO and Virus-like Particles in *Wolbachia*, an Endosymbiont of Arthropods. *Biochem. Biophys. Res. Commun.* **2001**, *283*, 1099–1104. [CrossRef] [PubMed]

53. Bordenstein, S.R. Bacteriophage Flux in Endosymbionts (*Wolbachia*): Infection Frequency, Lateral Transfer, and Recombination Rates. *Mol. Biol. Evol.* **2004**, *21*, 1981–1991. [CrossRef] [PubMed]

54. Kent, B.N.; Funkhouser, L.J.; Setia, S.; Bordenstein, S.R. Evolutionary Genomics of a Temperate Bacteriophage in an Obligate Intracellular Bacteria (*Wolbachia*). *PLoS ONE* **2011**, *6*, e24984. [CrossRef] [PubMed]

55. Jamnongluk, W.; Kittayapong, P.; Baimai, V.; O'Neill, S.L. *Wolbachia* Infections of Tephritid Fruit Flies: Molecular Evidence for Five Distinct Strains in a Single Host Species. *Curr Microbiol.* **2002**, *45*, 255–260. [CrossRef] [PubMed]

56. Jain, C.; Rodriguez-R, L.M.; Phillippy, A.M.; Konstantinidis, K.T.; Aluru, S. High Throughput ANI Analysis of 90K Prokaryotic Genomes Reveals Clear Species Boundaries. *Nat. Commun.* **2018**, *9*, 1–8. [CrossRef] [PubMed]

57. Vasconcelos, E.J.R.; Billeter, S.A.; Jett, L.A.; Meinersmann, R.J.; Barr, M.C.; Diniz, P.P.V.P.; Oakley, B.B. Assessing Cat Flea Microbiomes in Northern and Southern California by 16S RRNA Next-Generation Sequencing. *Vector-Borne Zoonotic Dis.* **2018**, *18*, 491–499. [CrossRef] [PubMed]

58. Lee, H.; Doak, T.G.; Popodi, E.; Foster, P.L.; Tang, H. Insertion Sequence-Caused Large-Scale Rearrangements in the Genome of Escherichia Coli. *Nucleic Acids Res.* **2016**, *44*, 7109–7119. [CrossRef]

59. Douglas, A.E. The B Vitamin Nutrition of Insects: The Contributions of Diet, Microbiome and Horizontally Acquired Genes. *Curr. Opin. Insect Sci.* **2017**, *23*, 65–69. [CrossRef]

60. Nikoh, N.; Hosokawa, T.; Moriyama, M.; Oshima, K.; Hattori, M.; Fukatsu, T. Evolutionary Origin of Insect–*Wolbachia* Nutritional Mutualism. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 10257–10262. [CrossRef]

61. Gerth, M.; Bleidorn, C. Comparative Genomics Provides a Timeframe for *Wolbachia* Evolution and Exposes a Recent Biotin Synthesis Operon Transfer. *Nat. Microbiol.* **2016**, *2*, 16241. [CrossRef]

62. Moriyama, M.; Nikoh, N.; Hosokawa, T.; Fukatsu, T. Riboflavin Provisioning Underlies *Wolbachia*'s Fitness Contribution to Its Insect Host. *mBio* **2015**, *6*, e01732-15. [CrossRef] [PubMed]

63. Laidoudi, Y.; Levasseur, A.; Medkour, H.; Maaloum, M.; Ben Khedher, M.; Sambou, M.; Bassene, H.; Davoust, B.; Fenollar, F.; Raoult, D.; et al. An Earliest Endosymbiont, *Wolbachia Massiliensis* Sp. Nov., Strain PL13 from the Bed Bug (*Cimex Hemipterus*), Type Strain of a New Supergroup T. *Int. J. Mol. Sci.* **2020**, *21*, 8064. [CrossRef] [PubMed]

64. Katz, L.S.; Griswold, T.; Morrison, S.S.; Caravas, J.A.; Zhang, S.; Bakker, H.C.; Deng, X.; Carleton, H.A. Mashtree: A Rapid Comparison of Whole Genome Sequence Files. *J. Open Source Softw.* **2019**, *4*, 1762. [CrossRef] [PubMed]

65. Hu, J.; Wang, Z.; Sun, Z.; Hu, B.; Ayoola, A.O.; Liang, F.; Li, J.; Sandoval, J.R.; Cooper, D.N.; Ye, K.; et al. An Efficient Error Correction and Accurate Assembly Tool for Noisy Long Reads. *bioRxiv* **2023**. [CrossRef]

66. Hu, J.; Fan, J.; Sun, Z.; Liu, S. NextPolish: A Fast and Efficient Genome Polishing Tool for Long-Read Assembly. *Bioinformatics* **2020**, *36*, 2253–2255. [CrossRef] [PubMed]

67. Eren, A.M.; Esen, Ö.C.; Quince, C.; Vineis, J.H.; Morrison, H.G.; Sogin, M.L.; Delmont, T.O. Anvi'o: An Advanced Analysis and Visualization Platform for 'omics Data. *PeerJ* **2015**, *3*, e1319. [CrossRef] [PubMed]

68. Seemann, T. Prokka: Rapid Prokaryotic Genome Annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef] [PubMed]

69. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [CrossRef]

70. Cantalapiedra, C.P.; Hernández-Plaza, A.; Letunic, I.; Bork, P.; Huerta-Cepas, J. EggNOG-Mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol. Biol. Evol.* **2021**, *38*, 5825–5829. [CrossRef]

71. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.L.; Tosatto, S.C.E.; Paladin, L.; Raj, S.; Richardson, L.J.; et al. Pfam: The Protein Families Database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412–D419. [CrossRef]

72. Arndt, D.; Grant, J.R.; Marcu, A.; Sajed, T.; Pon, A.; Liang, Y.; Wishart, D.S. PHASTER: A Better, Faster Version of the PHAST Phage Search Tool. *Nucleic Acids Res.* **2016**, *44*, W16–W21. [CrossRef] [PubMed]

73. Varani, A.M.; Siguier, P.; Gourbeyre, E.; Charneau, V.; Chandler, M. ISsaga Is an Ensemble of Web-Based Methods for High Throughput Identification and Semi-Automatic Annotation of Insertion Sequences in Prokaryotic Genomes. *Genome Biol.* **2011**, *12*, R30. [CrossRef] [PubMed]

74. Grant, J.R.; Stothard, P. The CGView Server: A Comparative Genomics Tool for Circular Genomes. *Nucleic Acids Res.* **2008**, *36*, W181–W184. [CrossRef] [PubMed]

75. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic Orthology Inference for Comparative Genomics. *Genome Biol.* **2019**, *20*, 238. [CrossRef] [PubMed]

76. Lex, A.; Gehlenborg, N.; Strobelt, H.; Vuillemot, R.; Pfister, H. UpSet: Visualization of Intersecting Sets. *IEEE Trans. Vis. Comput. Graphics* **2014**, *20*, 1983–1992. [CrossRef] [PubMed]

77. Krassowski, M. ComplexUpset: Create Complex UpSet Plots Using "ggplot2" Components. R Package Version 1.3.3 2021. Available online: http://doi.org/10.5281/zenodo.3700590 (accessed on 15 July 2023).

78. Yu, G.; Wang, L.-G.; Han, Y.; He, Q.-Y. ClusterProfiler: An R Package for Comparing Biological Themes among Gene Clusters. *OMICS A J. Integr. Biol.* **2012**, *16*, 284–287. [CrossRef] [PubMed]

79. Edgar, R.C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef] [PubMed]

80. Capella-Gutiérrez, S.; Silla-Martínez, J.M.; Gabaldón, T. TrimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [CrossRef]

81. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; von Haeseler, A.; Lanfear, R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [CrossRef]

82. Letunic, I.; Bork, P. Interactive Tree Of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [CrossRef]

83. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermiin, L.S. ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates. *Nat. Methods* **2017**, *14*, 587–589. [CrossRef] [PubMed]

84. Kolde, R. Pheatmap: Pretty Heatmaps. R Package Version 1.0.12 2019. Available online: https://cran.r-project.org/web/packages/pheatmap/index.html (accessed on 15 July 2023).