

Article

The Genome of *Arsenophonus* sp. and Its Potential Contribution in the Corn Planthopper, *Peregrinus maidis*

Yu-Hui Wang ¹, Aram Mikaelyan ¹ , Brad S. Coates ²  and Marcé Lorenzen ^{1,*} 

¹ Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC 27695, USA; ywang259@ncsu.edu (Y.-H.W.); amikael@ncsu.edu (A.M.)

² USDA-ARS, Corn Insects & Crop Genetics Research Unit, Ames, IA 50011, USA; brad.coates@ars.usda.gov

* Correspondence: marce_lorenzen@ncsu.edu; Tel.: +1-919-513-7857

Simple Summary: Microbes play important roles in the biochemistry, physiology and survival of insects. An example is *Arsenophonus* sp. in the brown planthopper (*Nilaparvata lugens*), an endosymbiotic bacterium supplementing the nutritionally poor diet of its sap-feeding host. In this study, we use the genome of *NI Arsenophonus* sp. as a reference to facilitate the genome assembly of a closely related *Arsenophonus* sp. found in the corn planthopper (*Peregrinus maidis*). Our assembly is one of the largest *Arsenophonus* genomes reported to date. We also investigated the role of *Arsenophonus* sp. in *P. maidis*, in which it appears to provide B vitamins and essential amino acids.

Abstract: The co-evolution between symbionts and their insect hosts has led to intricate functional interdependencies. Advances in DNA-sequencing technologies have not only reduced the cost of sequencing but, with the advent of highly accurate long-read methods, have also enabled facile genome assembly even using mixed genomic input, thereby allowing us to more easily assess the contribution of symbionts to their insect hosts. In this study, genomic data recently generated from *Peregrinus maidis* was used to assemble the genome of a bacterial symbiont, *Pm Arsenophonus* sp. This ~4.9-Mb assembly is one of the largest *Arsenophonus* genomes reported to date. The Benchmarking Universal Single-Copy Orthologs (BUSCO) result indicates that this *Pm Arsenophonus* assembly has a high degree of completeness, with 96% of the single-copy Enterobacterales orthologs found. The identity of the *Pm Arsenophonus* sp. was further confirmed by phylogenetic analysis. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis indicates a major contribution by *Pm Arsenophonus* sp. to the biosynthesis of B vitamins and essential amino acids in *P. maidis*, where threonine and lysine production is carried out solely by *Pm Arsenophonus* sp. This study not only provides deeper insights into the evolutionary relationships between symbionts and their insect hosts, but also adds to our understanding of insect biology, potentially guiding the development of novel pest control methods.

Keywords: genome assembly; *Arsenophonus* sp.; *Peregrinus maidis*; hemiptera



Citation: Wang, Y.-H.; Mikaelyan, A.; Coates, B.S.; Lorenzen, M. The Genome of *Arsenophonus* sp. and Its Potential Contribution in the Corn Planthopper, *Peregrinus maidis*. *Insects* **2024**, *15*, 113. <https://doi.org/10.3390/insects15020113>

Academic Editor: Nicolas N. Nègre

Received: 2 January 2024

Revised: 26 January 2024

Accepted: 31 January 2024

Published: 5 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The corn planthopper, *Peregrinus maidis*, is a notorious sap-feeding pest that targets corn and sorghum, and is known to vector the maize mosaic rhabdovirus and maize stripe tenuivirus [1]. The combination of insect damage and viral infection can significantly reduce crop yield [2]. Current management tactics mainly rely on insecticide applications, but effective control remains difficult because of the limited direct exposure of *P. maidis* under field conditions. Insecticides also have negative effects on ecosystems and human health. As such, new efficacious and sustainable methods are needed for the management of *P. maidis* and the viruses it transmits. Sequence-specificity-based RNA interference (RNAi) holds promise for the targeted control of pest species [3]. The efficacy of RNAi on *P. maidis* has been validated under laboratory conditions [4–6], but an efficient system

for delivering double-stranded RNA (dsRNA) to *P. maidis* in the field is not yet available. Innovative solutions for delivery involve the modification of symbionts to express host-specific dsRNAs [7–9], but research into the symbionts of *P. maidis* is still limited. Therefore, we leveraged the genomic sequence data recently generated for *P. maidis* in order to assemble the genome of the *Arsenophonus* sp. associated with this hemipteran. We focused on *Arsenophonus* sp. because of its demonstrated nutritional importance in another planthopper, *Nilaparvata lugens* [10], which suggests it may play a similar role in *P. maidis*. Moreover, the fact that protocols have been worked out for culturing and genetically modifying other *Arsenophonus* spp. [11,12] bodes well for the development of similar methods for *Pm Arsenophonus* sp. and supports the potential of using this symbiont to deliver dsRNA to *P. maidis*.

The genus *Arsenophonus*, first proposed in 1991 for *Arsenophonus nasoniae* [13], derives its name from the Greek “Arsen” (male) and “phonos” (slayer), reflecting the ability of *A. nasoniae* to kill the male offspring of *Nasonia vitripennis*. *Arsenophonus* spp. have been found in many insect species with diverse lifestyles [14]. They serve as primary symbionts in *Lipoptena cervi* and *Aleurodicus disperses* [15,16], and as secondary symbionts in *Aleurodicus floccissimus*, *Trialeurodes vaporariorum*, *N. lugens* and triatomine bugs [10,16,17]. Such a stable association is not observed in the honeybee, *Apis mellifera*, as individuals can acquire *Arsenophonus* sp. from the environment or nestmates [18]. Although it is rare, *Arsenophonus* spp. can act as insect-vectored plant pathogens; for example, *A. phytopathogenicus* transmitted by *Pentastiridius leporinus* causes low sugar content in sugar beets [19].

In this study, we identified *Arsenophonus*-specific scaffolds within a newly generated *P. maidis* genome assembly. The ~4.9 Mb *Pm Arsenophonus* genome assembly was surveyed for genes involved in key biosynthetic pathways, and the role *Pm Arsenophonus* sp. plays in the production of B vitamins and essential amino acids was determined. The data presented here have the potential to assist with the development of novel pest control methods, both by increasing our understanding of *P. maidis* biology, and by providing a candidate for use in symbiont-mediated RNAi.

2. Materials and Methods

2.1. Insect Rearing

Corn planthoppers were reared on 5-week-old corn (cultivar Early Sunglow, Park Seed Company, Greenwood, SC, USA) and kept in 30 cm × 30 cm × 60 cm cages covered by nylon-mesh screens (BioQuip Products Inc., Compton, CA, USA) in a room at 25 ± 1 °C, ~70% RH and at a 14:10 (light:dark) photoperiod.

2.2. Genome Sequencing, Assembly and Annotation

We extracted genomic DNA from *P. maidis* using the Mag-Bind Blood and Tissue DNA HDQ 96 Kit (Omega Bio-tek, Norcross, GA, USA). The quality of genomic DNA was assessed and submitted to Eremid Genomic Services (North Carolina Research Campus, Kannapolis, NC, USA) for library preparation using SMRTbell Express Template Prep Kit 2.0 (Pacific Biosystems, Menlo Park, CA, USA) and subsequent high-fidelity (HiFi) read generation on a Sequel IIe system (Pacific Biosystems). Raw reads were converted to fastq format using BamTools v. 2.5.1 [20], filtered to remove reads < 8 kb, and the data was transferred to the USDA–ARS SCINet high-performance computing cluster, Ceres. A histogram of canonical k-mer counts was made at k = 21 for filtered reads using Meryl [21], and was then applied as the maximum depth input for the R script, GenomeScope 2.0 [22,23]. Next, reads were assembled into contigs using hifiasm [24] with default assembly options but purging adjusted to [-I 3 (aggressive purging); -purge-max 14; -s 0.55; n-hap 2 (diploid)] on 40 cores with 768 GB DDR3 ECC RAM. Then, contigs were scaffolded using RagTag [25] on Galaxy Europe (<https://usegalaxy.eu>, accessed on 25 May 2022) [26]. RagTag uses whole-genome homology to scaffold, in which a reference genome is used to inform the order and orientation of query sequences. The genome of *Arsenophonus* sp. in *N. lugens* (accession number: GCA_000757905.1) was used as a reference to scaffold contigs [10]. Completeness of the *Pm*

Arsenophonus assembly was compared to other Arsenophonus genome assemblies (accession number: GCA_004768525.1, GCA_020268605.1, GCA_001534665.1, GCA_900343015.1, GCA_900343025.1, GCA_002287155.1 and GCA_000757905.1) using Benchmarking Universal Single-Copy Orthologs (BUSCO) [27–29] on Galaxy Europe (<https://usegalaxy.eu>, accessed on 22 June 2022) [26]. Genes were then predicted from the genome assemblies above using Glimmer [30], and annotation of the resulting genes was conducted using the BLASTx algorithm against the non-redundant protein database (E-value cutoff of 10^{-5}) through OmicsBox software (BioBam, Valencia, Spain) [31]. Then, criteria based on a previous study (E-value $< 10^{-15}$ and similarity $< 80\%$) [32] were applied to the predicted genes to roughly estimate the number of pseudogenes, and free-living bacteria *Proteus mirabilis*, *Escherichia coli* and *Pseudomonas aeruginosa* were included for comparison. Pathway analysis was then conducted on the genes predicted from the *Pm Arsenophonus* assembly and *P. maidis* transcriptome data [5] using the Kyoto Encyclopedia of Genes and Genomes (KEGG) [33] through OmicsBox.

2.3. RNA Polymerase β Subunit Gene (*rpoB*) Identification and Phylogenetic Analysis

The amino-acid sequence of *rpoB* (accession number: WP_000263098.1) was used as a query to search for *rpoB* orthologs in the predicted Arsenophonus genes exported from OmicsBox. The *rpoB* nucleotide sequences were then aligned using the MUSCLE algorithm [34] and evaluated against 24 nucleotide substitution models to find the most appropriate one for phylogenetic analysis in MEGAX [35]. The phylogenetic tree based on *rpoB* was constructed using the Maximum Likelihood method, with the best-fitting substitution model identified by the lowest Bayesian information criterion score. Tree topology was evaluated using bootstrap analysis with 1000 replications. The *rpoB* genes of *Buchnera aphidicola* and *Riesia pediculicola* were used as outgroups to root the phylogenetic tree. The predicted genes of *B. aphidicola* and *R. pediculicola* were generated from genome assemblies *B. aphidicola* JF99 (accession number: GCA_000183285.1) and *R. pediculicola* (accession number: GCA_014879315.1), and were queried for *rpoB* orthologs as described above.

3. Results and Discussion

Eleven *Arsenophonus*-specific scaffolds were identified within the *P. maidis* genomic data. Together these formed a genome assembly of ~4.9 Mb, which was then compared to data from *Arsenophonus* spp. whose genome assemblies and roles in their respective hosts were available. An exception was made for *Arsenophonus* sp. in the keeled treehopper (*Entylia carinata*), due to the closer relationship between its insect host and *P. maidis*. Our *Pm Arsenophonus* assembly is ~4.9 Mb in size, one of the largest reported to date. While the difference in genome size might be attributed to genuine biological variation among the genomes, it could also result from the different sequencing platforms. Notably, the larger *Arsenophonus* genome assemblies have been mostly generated using third-generation PacBio and/or Nanopore reads, which generate mean read lengths that are sufficient to span repetitive regions and facilitate more accurate and complete genome assemblies. In contrast, assemblies from short-read data often leave repetitive regions unassembled, resulting in gapped pseudomolecules [36]. Therefore, the comparatively larger-sized *Pm Arsenophonus* genome could be a consequence of our use of PacBio HiFi reads. *De novo* assembly of other *Arsenophonus* genomes with long-read data may alter the ranking based on genome size. The large size of the *Pm Arsenophonus* genome may contribute to it containing 423 of the 440 single-copy orthologs in the BUSCO Enterobacterales data set (Figure 1), which is among the highest compared to other *Arsenophonus* genomes. BUSCO results also suggest that some *Arsenophonus* genome assemblies may be incomplete and thus have underestimated sizes, particularly those that are not primary symbionts (*A. triatominarum*, *Arsenophonus* sp. in *A. floccissimus* and *Arsenophonus* sp. in *N. lugens*).

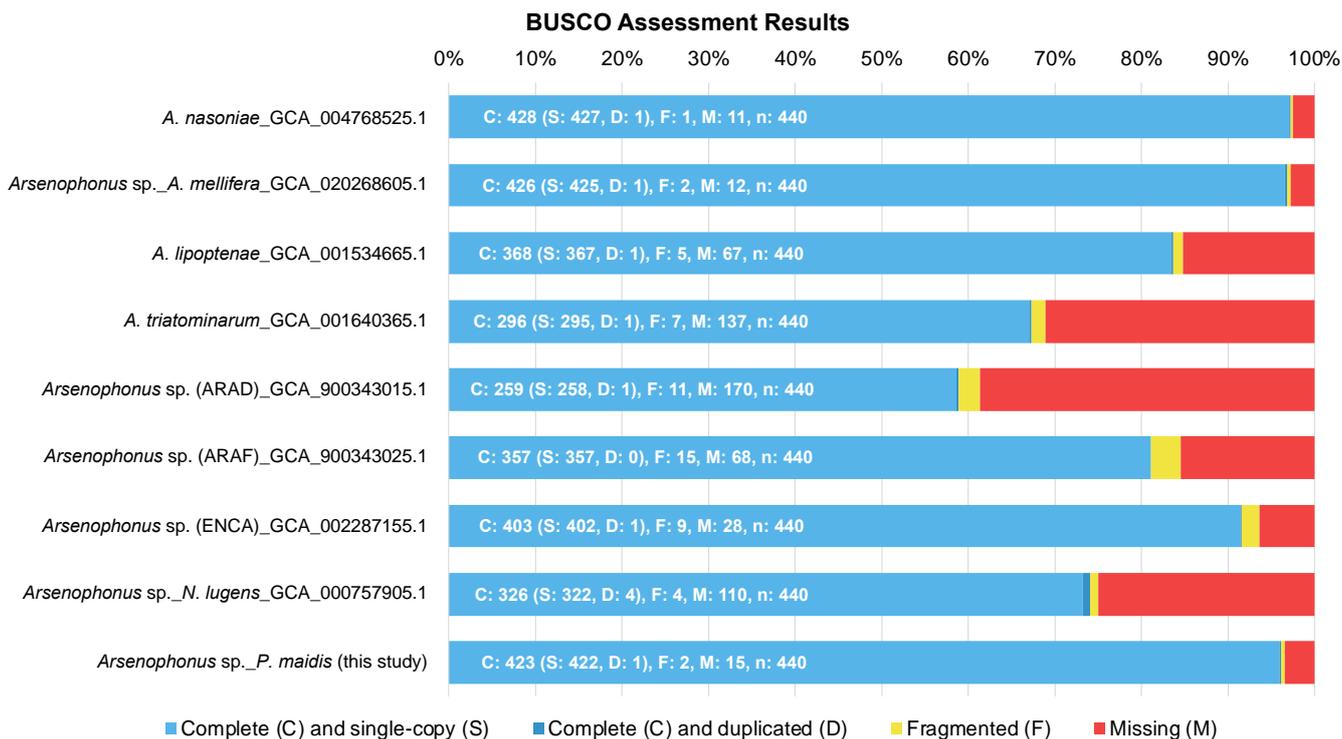


Figure 1. BUSCO assessment results. The completeness of genome assemblies was assessed using BUSCO on Galaxy Europe (<https://usegalaxy.eu>, accessed on 22 June 2022).

During symbiont evolution, characteristic changes manifest within the genomes of host-dependent bacteria, reflecting the adaptation to different symbiotic lifestyles and ongoing selection pressures. One commonly observed trend is genome size reduction, which leads to more host-restricted symbionts [37] and the accumulation of mutations. A prevalent mutational bias towards AT suggests that primary symbionts generally tend to have genomes with a lower GC content. This is reflected in our comparative genomic analysis (Table 1). Specifically, the primary symbionts *A. lipopteni* and *Arsenophonus* sp. in *A. disperses* (hereafter “*Arsenophonus* sp. (ARAD)”) have a lower GC content, with an estimated 29.4 and 32.2%, respectively. The high 41.3% GC content of the *Pm Arsenophonus* genome and the comparatively large genome size suggest that *Pm Arsenophonus* sp. is at a relatively early stage of symbiont evolution.

Pseudogenes, or putative non-functional remnants of once-active genes, serve as another indicator of evolutionary stage and degree of endosymbiotic specialization. In the process of the transition from a free-living bacterium to a primary symbiont, pseudogenes accumulate due to the changing selective pressures associated with the symbiotic lifestyle. In a free-living state, gene function is maintained to support the diverse requirements of independent life. As an organism enters into a symbiotic relationship, a subset of genes that were essential in a free-living state might become redundant, leading to the formation of pseudogenes. Once the symbiotic relationship is well established and the organism has become a primary symbiont, there could be a deletional bias whereby non-functional DNA (like pseudogenes) is removed from the genome [37,38]. This pattern may lead to relatively higher numbers of pseudogenes in organisms that are neither fully free-living nor fully adapted as primary symbionts. Based on the criteria applied, 25 pseudogenes were estimated to be present in the genomes of *P. mirabilis*, one in *E. coli* and none in *P. aeruginosa*. Among *Arsenophonus* spp., the fewest number of pseudogenes was estimated to be present in the genomes of primary symbionts *A. lipopteni* and *Arsenophonus* sp. (ARAD) (Table 1). This was in line with expectations for free-living bacteria and primary symbionts. Interestingly, the *Pm Arsenophonus* genome appears to contain 1123 pseudogenes, which

indicates that it is neither a free-living bacterium nor a primary symbiont. This fits well with the idea that *Pm Arsenophonus* sp. is at a relatively early stage of symbiont evolution.

Table 1. Comparison between *Arsenophonus* genomes.

Name	<i>Arsenophonus nasoniae</i>	<i>Arsenophonus</i> sp.	<i>Arsenophonus lipopteni</i>	<i>Arsenophonus triatominarum</i>	
Accession No.	GCA_004768525.1	GCA_020268605.1	GCA_001534665.1	GCA_001640365.1	
Host	<i>Nasonia vitripennis</i> Oxford Nanopore MinION;	<i>Apis mellifera</i> Oxford Nanopore MinION;	<i>Lipoptena fortisetosa</i>	<i>Triatoma infestans</i>	
Sequencing	PacBio RS II; Illumina MiSeq	Illumina MiSeq	Illumina	PacBio	
Genome size (bp)	4,987,107	3,639,254	836,724	3,858,720	
GC content (%)	38.1 ¹	37.7 ¹	24.9 ¹	38.3	
Pseudogene ²	1086	720	104	523	
Name	<i>Arsenophonus</i> sp. (ARAD)	<i>Arsenophonus</i> sp. (ARAF)	<i>Arsenophonus</i> sp. (ENCA)	<i>Arsenophonus</i> sp.	<i>Arsenophonus</i> sp. (this study)
Accession No.	GCA_900343015.1	GCA_900343025.1	GCA_002287155.1	GCA_000757905.1	
Host	<i>Aleurodicus disperses</i>	<i>Aleurodicus floccissimus</i>	<i>Entylia carinata</i>	<i>Nilaparvata lugens</i>	<i>Peregrinus maidis</i>
Sequencing	Illumina HiSeq 2000	Illumina HiSeq 2000	Illumina MiSeq	Illumina HiSeq 2000	PacBio
Genome size (bp)	663,125	3,001,875	3,228,533	2,953,863	4,888,380
GC content (%)	32.2 ¹	37	39.5	37.6	41.3
Pseudogene ²	29	382	430	469	1123

¹ Plasmid sequence was not included in the calculation of GC content. ² The number of pseudogenes was roughly estimated by applying the criteria based on a previous study (E value < 10⁻¹⁵ and similarity < 80%) [32] to the predicted genes.

The identity of *Pm Arsenophonus* sp. was further confirmed by phylogenetic analysis. The 16S ribosomal RNA (16S rRNA) gene, which produces the RNA component of the prokaryotic small ribosomal subunit, is commonly used for bacterial identification. The essential role of 16S rRNA results in its ubiquitous presence in almost all bacteria, making it a good molecular marker. Additionally, since the function of 16S rRNA has not changed, random changes in the sequence can be used to measure evolutionary time. Available 16S rRNA gene databases, such as Greengenes [39], Ribosomal Database Project [40] and SILVA [41], reflect the common use of the 16S rRNA gene. Despite the ease of using the 16S rRNA gene for bacterial identification, problems can arise [42]. For example, some species possess multiple copies of the 16S rRNA gene, and the sequences can differ by as much as 11.6% [43], which can cause identification issues. Since *Arsenophonus* spp. also possess multiple copies of the 16S rRNA gene, we used the single-copy gene, encoding the β subunit of RNA polymerase (*rpoB*) instead for phylogenetic analysis and the identification of *Arsenophonus* spp. (Figure 2). The *rpoB* gene was first used in 1997 [44], and it provides better phylogenetic resolution than the 16S rRNA gene at both the species and subspecies levels [43]. Although it was not investigated further, likely due to the lower completeness of the *A. triatominarum* genome assembly (Figure 1, 137 of 440 BUSCO, 31.1% predicted as missing), its *rpoB* could not be identified, and therefore was excluded from phylogenetic analysis. Regardless, the resultant phylogenetic tree provides strong support for the bacterial species (with bootstrap values \geq 97%), and groups primary symbionts into a distinct clade. The tree topology also reflects the relatedness of their insect hosts.

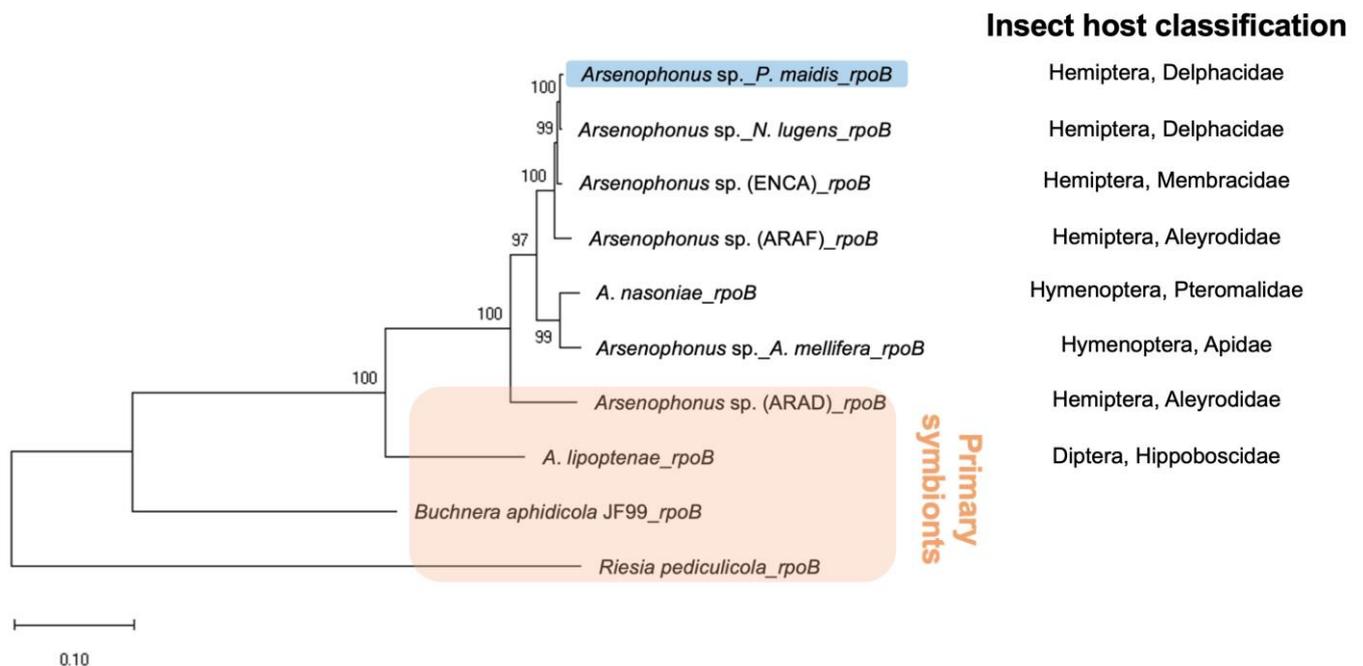


Figure 2. Maximum likelihood-based phylogenetic tree of *Arsenophonus* spp. estimated from a consensus *rpoB* nucleotide sequence alignment. The proportion (%) of 1000 bootstrap replicates supporting each node is indicated. The *rpoB* gene of *Pm Arsenophonus sp.* is highlighted in blue, while those identified in primary symbionts are highlighted in orange. The scale bar represents an evolutionary distance of 0.10 nucleotide substitutions per site.

To gain insights into the role *Pm Arsenophonus sp.* plays in *P. maidis*, we conducted KEGG pathway analysis on the *Arsenophonus* genome and previously reported *P. maidis* transcriptome [5]. Since some *Arsenophonus* spp. are reported to provide their insect hosts with B vitamins (Table 2) [10,15,16,45], we specifically interrogated components of the biosynthetic pathways for vitamins B1, B2, B5, B6, B7 and B9. Our KEGG pathway analysis suggests that *Pm Arsenophonus sp.* complements the biosynthetic pathways for vitamins B1, B2, B6, B7 and B9 in *P. maidis* (Figure 3). The nine enzymes for vitamin B7 production are encoded solely within *Pm Arsenophonus sp.*, suggesting that the insect is dependent on the bacterium for B7 provisioning. We also examined components of biosynthetic pathways for essential amino acids, since symbionts in hemipterans are known for supplying essential amino acids to their insect hosts. KEGG analysis suggests that *Pm Arsenophonus sp.* provides components for the production of isoleucine, leucine, valine, threonine and lysine (Figure 4). The sole encoding of enzymes for threonine and lysine biosynthesis within the *Pm Arsenophonus* genome further emphasizes their symbiotic relationship. While such contributions of *Arsenophonus* spp. to essential amino acid provisioning in insect hosts have not been observed in previous studies (Table 2), our findings point to a potential novel aspect of the involvement of *Arsenophonus sp.* in the metabolism of *P. maidis*, meriting further investigation.

Table 2. Comparison among *Arsenophonus* spp.

Name	Host	Role and Function ¹	Transmission	Reference
<i>Arsenophonus phytopathogenicus</i>	<i>Pentastiridius leporinus</i>	Plant pathogen of sugar beets vectored by <i>P. leporinus</i>	Maternal and horizontal (major) transmission by infecting the same sugar beet	[19]
<i>Arsenophonus nasoniae</i>	<i>Nasonia vitripennis</i> <i>Nasonia longicornis</i>	Reproductive parasite Killing sons, 80% of sons die	Maternal and horizontal transmission by infecting the same pupa host	[46]
<i>Arsenophonus</i> sp.	<i>Apis mellifera</i>	Providing B vitamins (B2, B6, B7, B9)	Horizontal transmission by social interactions (trophallaxis and/or general contact) and environmental acquisition	[18,45]
<i>Arsenophonus lipopteni</i>	<i>Lipoptena cervi</i>	Primary endosymbiont Providing B vitamins (B2, B6, B7)	Unclear	[15]
<i>Arsenophonus triatominarum</i>	triatomine bugs	Secondary endosymbiont No apparent effect on host fitness or reproduction	Vertical transmission (transovarially)	[17]
<i>Arsenophonus</i> sp. (ARAD)	<i>Aleurodicus disperses</i>	ARAD as primary endosymbiont ARAF and ARTV as secondary endosymbiont	Unclear	[16]
<i>Arsenophonus</i> sp. (ARAF)	<i>Aleurodicus floccissimus</i>	ARAD providing cofactors and B vitamins		
<i>Arsenophonus</i> sp. (ARTV)	<i>Trialeurodes vaporariorum</i>	(B1, B2, B6, B7, B9)		
<i>Arsenophonus</i> sp.	<i>Nilaparvata lugens</i>	Secondary endosymbiont Providing B vitamins	Maternal transmission	[10,47]

¹ B vitamins in bold indicate that *Arsenophonus* sp. possesses all the genes in the biosynthetic pathway; B vitamins with an underline indicate that genes from both the *Arsenophonus* sp. and the host are required to complete the biosynthetic pathway; the remainder are not specified in references.

Whitten et al. [7] demonstrated the feasibility of symbiont-mediated RNAi in the kissing bug (*Rhodnius prolixus*) and the western flower thrips (*Frankliniella occidentalis*). They found that targeted insect genes were significantly knocked down by feeding the insects dsRNA-expressing symbiotic bacteria. Furthermore, the modified symbionts were able to successfully compete with the gut microbiota. While there is more to learn, the important role *Pm Arsenophonus* sp. plays in *P. maidis* and the availability of protocols for culturing and genetically modifying *Arsenophonus* spp. support *Pm Arsenophonus* sp. as an excellent candidate for use in symbiont-mediated RNAi.

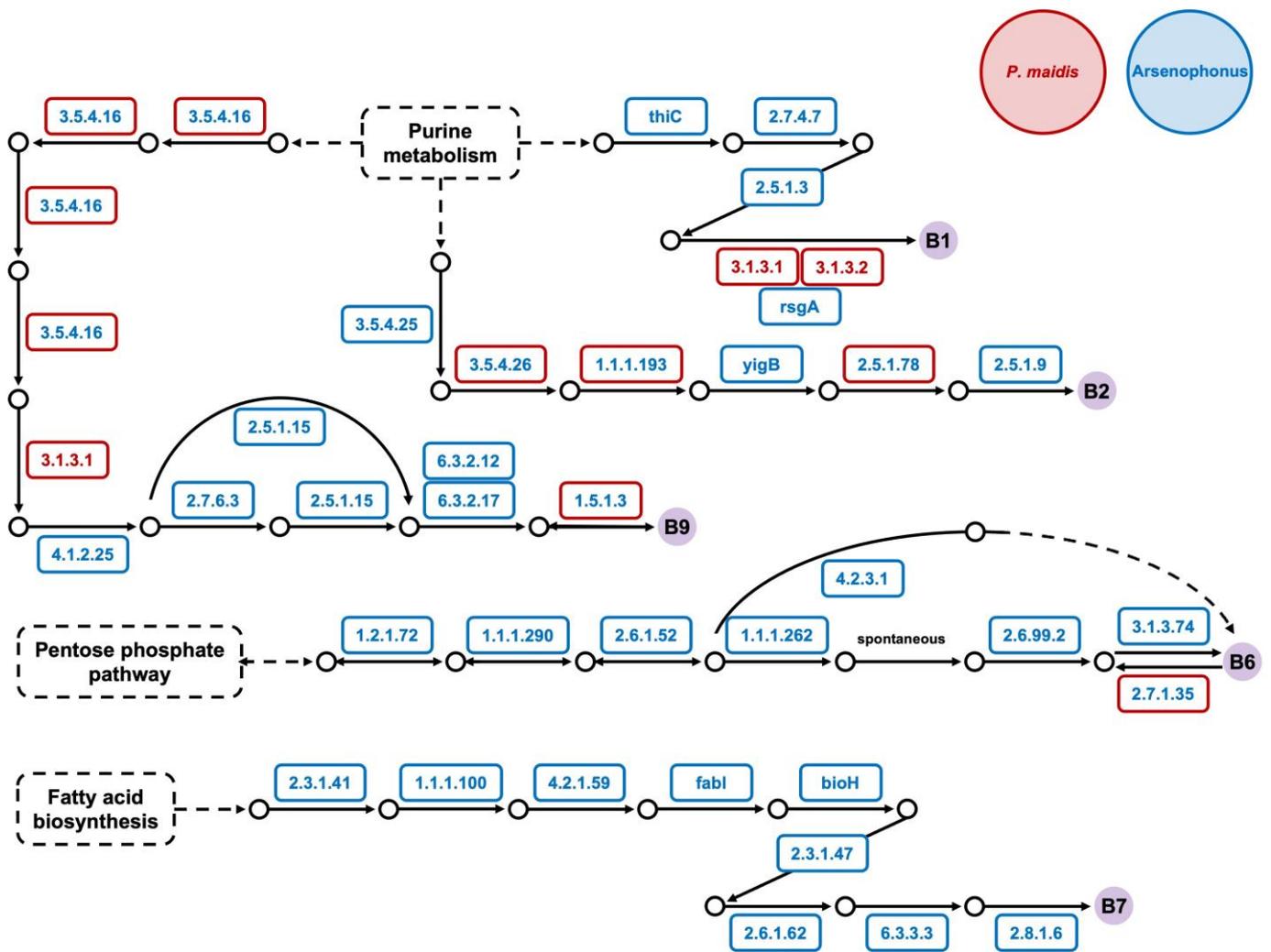


Figure 3. The biosynthetic pathway of B vitamins completed by *Arsenophonus* sp. and *P. maidis*. The KEGG pathway analysis was conducted, and enzyme commission numbers (EC numbers) are shown above. The EC numbers in red font within red boxes represent enzymes that were found in the *P. maidis* transcriptome; those in blue font within blue boxes represent enzymes that were found in the *Pm Arsenophonus* genome; and those in blue font within red boxes indicate enzymes that were found in both. B vitamins are shown in solid purple circles.

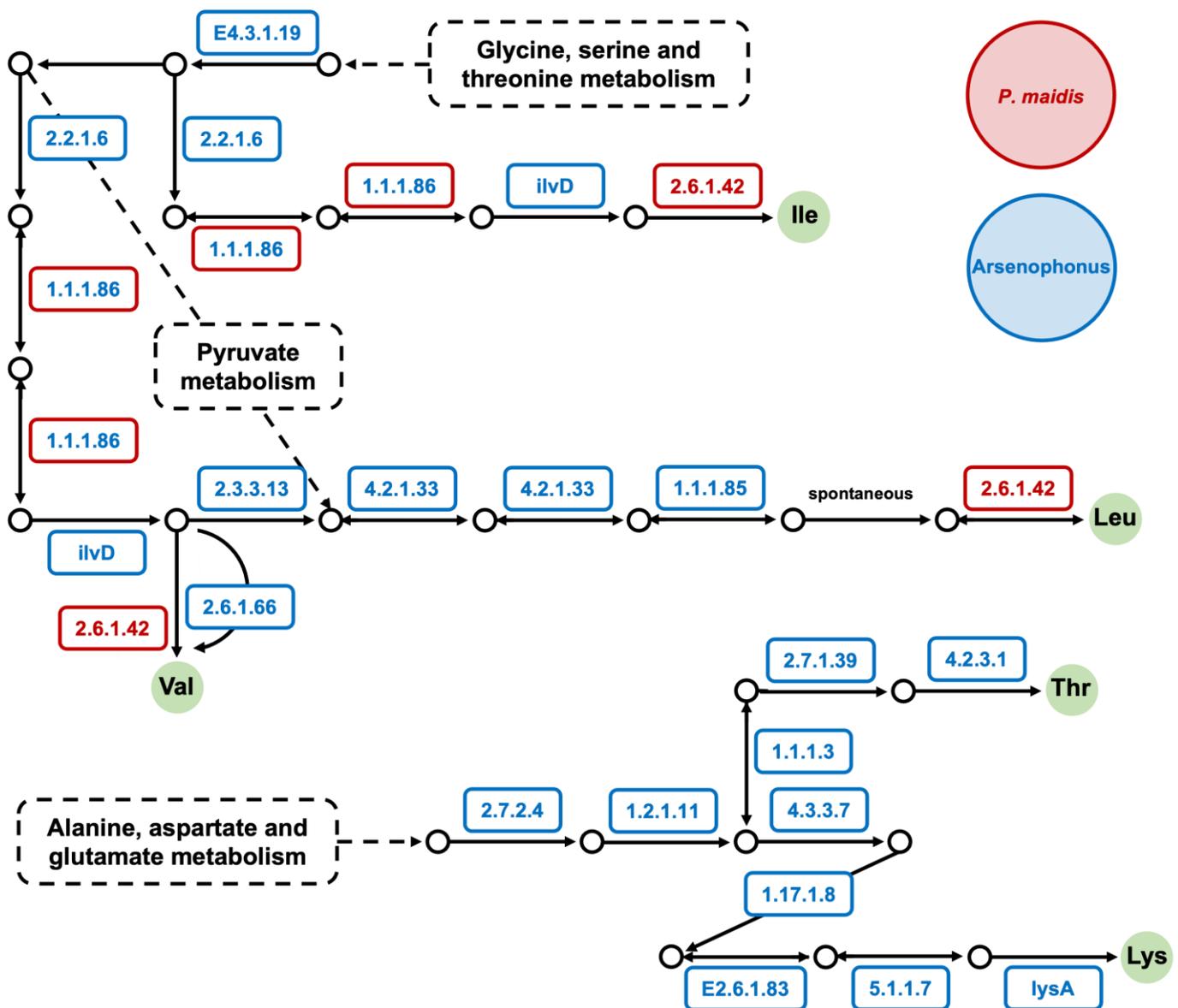


Figure 4. The biosynthetic pathway of essential amino acids completed by *Arsenophonus* sp. and *P. maidis*. The KEGG pathway analysis was conducted, and EC numbers are shown above. The EC numbers in red font within red boxes represent enzymes that were found in the *P. maidis* transcriptome; those in blue font within blue boxes represent enzymes that were found in the *Pm Arsenophonus* genome; and those in blue font within red boxes indicate enzymes that were found in both. Essential amino acids are shown in solid green circles.

Author Contributions: Conceptualization, Y.-H.W. and M.L.; methodology, Y.-H.W., A.M., B.S.C. and M.L.; validation, Y.-H.W.; investigation, Y.-H.W.; resources, B.S.C. and M.L.; writing—original draft preparation, Y.-H.W.; writing—review and editing, Y.-H.W., A.M., B.S.C. and M.L.; visualization, Y.-H.W.; supervision, M.L.; funding acquisition, M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Defense Advanced Research Projects Agency (DARPA) Insect Allies Program.

Data Availability Statement: The data are available at NCBI with the BioProject ID PRJNA1060725.

Acknowledgments: This research was performed in cooperation with the United States Department of Agriculture (USDA), Agricultural Research Service (ARS) (CRIS Project 5030-22000-019-00D)

“Ecologically-based Management of Arthropods in the Maize Agroecosystem”. This research used resources provided by the SCINet project of the USDA–ARS, project number 0500-00093-001-00-D. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA–ARS is an equal opportunity employer and provider. Y.-H.W. is supported by a Taiwan Government Fellowship for Studying Abroad. Work by A.M. is supported by a National Institute of Food and Agriculture Hatch Project (accession number 1019324). The authors acknowledge the support of the Freiburg Galaxy Team: Person X and Björn Grüning, Bioinformatics, University of Freiburg (Germany), funded by the Collaborative Research Centre 992 Medical Epigenetics (DFG grant SFB 992/1 2012) and the German Federal Ministry of Education and Research BMBF grant 031 A538A de.NBI-RBC.

Conflicts of Interest: North Carolina State University, Department of Entomology and Plant Pathology, was part of a team supporting DARPA’s Insect Allies Program. The views, opinions and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. The authors declare no competing interests.

References

1. Singh, B.U.; Seetharama, N. Host plant interactions of the corn planthopper, *Peregrinus maidis* Ashm. (Homoptera: Delphacidae) in maize and sorghum agroecosystems. *Arthropod Plant Interact.* **2008**, *2*, 163–196. [[CrossRef](#)]
2. Autrey, L. Maize mosaic virus and other maize virus diseases in the islands of the western Indian Ocean. In Proceedings of the International Maize Virus Diseases Colloquium and Workshop, Wooster, OH, USA, 2–6 August 1982.
3. Gundersen, D.E.; Adrianos, S.L.; Allen, M.L.; Becnel, J.J.; Chen, Y.; Choi, M.Y.; Estep, A.; Evans, J.D.; Garczynski, S.F.; Geib, S.M.; et al. Arthropod genomics research in the United States Department of Agriculture-Agricultural Research Service: Applications of RNA interference and CRISPR gene editing technologies in pest control. *Trends Entomol.* **2017**, *13*, 109–137.
4. Yao, J.; Rotenberg, D.; Afsharifar, A.; Barandoc-Alviar, K.; Whitfield, A.E. Development of RNAi methods for *Peregrinus maidis*, the corn planthopper. *PLoS ONE* **2013**, *8*, e70243. [[CrossRef](#)]
5. Wang, Y.-H.; Klobasa, W.; Chu, F.-C.; Huot, O.; Whitfield, A.E.; Lorenzen, M. Structural and functional insights into the ATP-binding cassette transporter family in the corn planthopper, *Peregrinus maidis*. *Insect Mol. Biol.* **2023**, *32*, 412–423. [[CrossRef](#)] [[PubMed](#)]
6. Xavier, C.A.D.; Tyson, C.; Kerner, L.M.; Whitfield, A.E. Exportin 1 is required for the reproduction and maize mosaic virus accumulation in its insect vector *Peregrinus maidis*. *bioRxiv* **2023**. [[CrossRef](#)]
7. Whitten, M.M.; Facey, P.D.; Del Sol, R.; Fernández-Martínez, L.T.; Evans, M.C.; Mitchell, J.J.; Bodger, O.G.; Dyson, P.J. Symbiont-mediated RNA interference in insects. *Proc. R. Soc. B Biol. Sci.* **2016**, *283*, 20160042. [[CrossRef](#)] [[PubMed](#)]
8. Leonard, S.P.; Powell, J.E.; Perutka, J.; Geng, P.; Heckmann, L.C.; Horak, R.D.; Davies, B.W.; Ellington, A.D.; Barrick, J.E.; Moran, N.A. Engineered symbionts activate honey bee immunity and limit pathogens. *Science* **2020**, *367*, 573–576. [[CrossRef](#)]
9. Elston, K.M.; Maeda, G.P.; Perreau, J.; Barrick, J.E. Addressing the challenges of symbiont-mediated RNAi in aphids. *PeerJ* **2023**, *11*, e14961. [[CrossRef](#)]
10. Xue, J.; Zhou, X.; Zhang, C.X.; Yu, L.L.; Fan, H.W.; Wang, Z.; Xu, H.-J.; Xi, Y.; Zhu, Z.-R.; Zhou, W.-W.; et al. Genomes of the rice pest brown planthopper and its endosymbionts reveal complex complementary contributions for host adaptation. *Genome Biol.* **2014**, *15*, 521. [[CrossRef](#)]
11. Dale, C.; Beeton, M.; Harbison, C.; Jones, T.; Pontes, M. Isolation, pure culture, and characterization of “*Candidatus Arsenophonus arthropodicus*”, an intracellular secondary endosymbiont from the hippoboscid louse fly *Pseudolynchia canariensis*. *Appl. Environ. Microbiol.* **2006**, *72*, 2997–3004. [[CrossRef](#)]
12. Nadal-Jimenez, P.; Griffin, J.S.; Davies, L.; Frost, C.L.; Marcello, M.; Hurst, G.D. Genetic manipulation allows in vivo tracking of the life cycle of the son-killer symbiont, *Arsenophonus nasoniae*, and reveals patterns of host invasion, tropism and pathology. *Environ. Microbiol.* **2019**, *21*, 3172–3182. [[CrossRef](#)]
13. Gherna, R.L.; Werren, J.H.; Weisburg, W.; Côté, R.; Woese, C.R.; Mandelco, L.; Brenner, D.J. *Arsenophonus nasoniae* gen. nov., sp. nov., the causative agent of the son-killer trait in the parasitic wasp *Nasonia vitripennis*. *Int. J. Syst. Evol. Microbiol.* **1991**, *41*, 563–565. [[CrossRef](#)]
14. Nováková, E.; Hypša, V.; Moran, N.A. *Arsenophonus*, an emerging clade of intracellular symbionts with a broad host distribution. *BMC Microbiol.* **2009**, *9*, 143. [[CrossRef](#)] [[PubMed](#)]
15. Nováková, E.; Hypša, V.; Nguyen, P.; Husník, F.; Darby, A.C. Genome sequence of *Candidatus Arsenophonus lipopteni*, the exclusive symbiont of a blood sucking fly *Lipoptena cervi* (Diptera: Hippoboscidae). *Stand. Genom. Sci.* **2016**, *11*, 72. [[CrossRef](#)] [[PubMed](#)]
16. Santos-Garcia, D.; Juravel, K.; Freilich, S.; Zchori-Fein, E.; Latorre, A.; Moya, A.; Morin, S.; Silva, F.J. To B or not to B: Comparative genomics suggests *Arsenophonus* as a source of B vitamins in whiteflies. *Front. Microbiol.* **2018**, *9*, 2254. [[CrossRef](#)] [[PubMed](#)]

17. Wilkes, T.E.; Duron, O.; Darby, A.C.; Hypša, V.; Nováková, E.; Hurst, G.D. The Genus *Arsenophonus*. In *Manipulative Tenants: Bacteria Associated with Arthropods*; Zchori-Fein, E., Bourtzis, K., Eds.; CRC Press: Boca Raton, FL, USA, 2011; pp. 225–244.
18. Drew, G.C.; Budge, G.E.; Frost, C.L.; Neumann, P.; Siozios, S.; Yañez, O.; Hurst, G.D. Transitions in symbiosis: Evidence for environmental acquisition and social transmission within a clade of heritable symbionts. *ISME J.* **2021**, *15*, 2956–2968. [[CrossRef](#)] [[PubMed](#)]
19. Bressan, A. Emergence and evolution of *Arsenophonus* bacteria as insect-vectorized plant pathogens. *Infect. Genet. Evol.* **2014**, *22*, 81–90. [[CrossRef](#)]
20. Barnett, D.W.; Garrison, E.K.; Quinlan, A.R.; Strömberg, M.P.; Marth, G.T. BamTools: A C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* **2011**, *27*, 1691–1692. [[CrossRef](#)]
21. Miller, J.R.; Delcher, A.L.; Koren, S.; Venter, E.; Walenz, B.P.; Brownley, A.; Johnson, J.; Li, K.; Mobarry, C.; Sutton, G. Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics* **2008**, *24*, 2818–2824. [[CrossRef](#)]
22. Vurture, G.W.; Sedlazeck, F.J.; Nattestad, M.; Underwood, C.J.; Fang, H.; Gurtowski, J.; Schatz, M.C. GenomeScope: Fast reference-free genome profiling from short reads. *Bioinformatics* **2017**, *33*, 2202–2204. [[CrossRef](#)]
23. Ranallo-Benavidez, T.R.; Jaron, K.S.; Schatz, M.C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **2020**, *11*, 1432. [[CrossRef](#)] [[PubMed](#)]
24. Cheng, H.; Concepcion, G.T.; Feng, X.; Zhang, H.; Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **2021**, *18*, 170–175. [[CrossRef](#)] [[PubMed](#)]
25. Alonge, M.; Lebeigle, L.; Kirsche, M.; Jenike, K.; Ou, S.; Aganezov, S.; Wang, X.; Lippman, Z.B.; Schatz, M.C.; Soyk, S. Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* **2022**, *23*, 258. [[CrossRef](#)] [[PubMed](#)]
26. Giardine, B.; Riemer, C.; Hardison, R.C.; Burhans, R.; Elnitski, L.; Shah, P.; Zhang, Y.; Blankenberg, D.; Albert, I.; Taylor, J.; et al. Galaxy: A platform for interactive large-scale genome analysis. *Genome Res.* **2005**, *15*, 1451–1455. Available online: <http://www.genome.org/cgi/doi/10.1101/gr.4086505> (accessed on 22 June 2022). [[CrossRef](#)] [[PubMed](#)]
27. 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](#)] [[PubMed](#)]
28. Seppey, M.; Manni, M.; Zdobnov, E.M. BUSCO: Assessing genome assembly and annotation completeness. In *Gene Prediction*; Kollmar, M., Ed.; Humana: New York, NY, USA, 2019; pp. 227–245. Available online: https://link.springer.com/protocol/10.1007/978-1-4939-9173-0_14 (accessed on 22 June 2022).
29. Manni, M.; Berkeley, M.R.; Seppey, M.; Simão, F.A.; Zdobnov, E.M. BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* **2021**, *38*, 4647–4654. [[CrossRef](#)] [[PubMed](#)]
30. Delcher, A.L.; Harmon, D.; Kasif, S.; White, O.; Salzberg, S.L. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **1999**, *27*, 4636–4641. [[CrossRef](#)] [[PubMed](#)]
31. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
32. Lerat, E.; Ochman, H. Ψ-Φ: Exploring the outer limits of bacterial pseudogenes. *Genome Res.* **2004**, *14*, 2273–2278. [[CrossRef](#)]
33. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [[CrossRef](#)]
34. Edgar, R.C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* **2004**, *5*, 113. [[CrossRef](#)]
35. Kumar, S.; Stecher, G.; Li, M.; Niyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
36. De Maio, N.; Shaw, L.P.; Hubbard, A.; George, S.; Sanderson, N.D.; Swann, J.; Wick, R.; AbuOun, M.; Stubberfield, E.; Hoosdally, S.J.; et al. Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes. *Microb. Genom.* **2019**, *5*, e000294. [[CrossRef](#)]
37. Lo, W.S.; Huang, Y.Y.; Kuo, C.H. Winding paths to simplicity: Genome evolution in facultative insect symbionts. *FEMS Microbiol. Rev.* **2016**, *40*, 855–874. [[CrossRef](#)]
38. Mira, A.; Ochman, H.; Moran, N.A. Deletional bias and the evolution of bacterial genomes. *Trends Genet.* **2001**, *17*, 589–596. [[CrossRef](#)]
39. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072. [[CrossRef](#)]
40. Cole, J.R.; Wang, Q.; Fish, J.A.; Chai, B.; McGarrell, D.M.; Sun, Y.; Brown, C.T.; Porras-Alfaro, A.; Kuske, C.R.; Tiedje, J.M. Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **2014**, *42*, 633–642. [[CrossRef](#)]
41. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, 590–596. [[CrossRef](#)]
42. Janda, J.M.; Abbott, S.L. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. *J. Clin. Microbiol.* **2007**, *45*, 2761–2764. [[CrossRef](#)]
43. Case, R.J.; Boucher, Y.; Dahllöf, I.; Holmström, C.; Doolittle, W.F.; Kjelleberg, S. Use of 16S rRNA and *rpoB* genes as molecular markers for microbial ecology studies. *Appl. Environ. Microbiol.* **2007**, *73*, 278–288. [[CrossRef](#)]

44. Mollet, C.; Drancourt, M.; Raoult, D. *rpoB* sequence analysis as a novel basis for bacterial identification. *Mol. Microbiol.* **1997**, *26*, 1005–1011. [[CrossRef](#)]
45. Nadal-Jimenez, P.; Siozios, S.; Frost, C.L.; Court, R.; Chrostek, E.; Drew, G.C.; Evans, J.D.; Hawthorne, D.J.; Burritt, J.B.; Hurst, G.D. *Arsenophonus apicola* sp. nov. isolated from the honeybee *Apis mellifera*. *Int. J. Syst. Evol. Microbiol.* **2022**, *72*, 005469. [[CrossRef](#)]
46. Dittmer, J.; Van Opstal, E.J.; Shropshire, J.D.; Bordenstein, S.R.; Hurst, G.D.; Brucker, R.M. Disentangling a holobiont–recent advances and perspectives in *Nasonia* wasps. *Front. Microbiol.* **2016**, *7*, 1478. [[CrossRef](#)] [[PubMed](#)]
47. Qu, L.Y.; Lou, Y.H.; Fan, H.W.; Ye, Y.X.; Huang, H.J.; Hu, M.Q.; Zhu, Y.-N.; Zhang, C.X. Two endosymbiotic bacteria, *Wolbachia* and *Arsenophonus*, in the brown planthopper *Nilaparvata lugens*. *Symbiosis* **2013**, *61*, 47–53. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.