

Article

Comparative Genomics Reveals Three Genetic Groups of the Whitefly Obligate Endosymbiont *Candidatus Portiera aleyrodidarum*

Teng Lei ^{1,†} , Ning Luo ^{2,†}, Chao Song ¹ , Junwei Yu ¹, Yuhang Zhou ¹, Xin Qi ^{1,*}  and Yinquan Liu ^{3,*} ¹ Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation, School of Life Sciences, Taizhou University, Taizhou 318000, China; leiteng@tzc.edu.cn (T.L.)² Natural Resources and Planning Bureau of Linhai City, Linhai 317000, China³ Ministry of Agriculture and Rural Affairs Key Laboratory of Agricultural Entomology, Key Laboratory of Biology of Crop Pathogens and Insects of Zhejiang Province, Institute of Insect Sciences, Zhejiang University, Hangzhou 310058, China

* Correspondence: qixin0612@tzc.edu.cn (X.Q.); yqliu@zju.edu.cn (Y.L.)

† The authors contributed equally to this work.

Simple Summary: Many obligate endosymbionts colonize an invertebrate host and are directly transferred maternally to the host's embryo. Consequently, the symbiotic bacteria from diverse host species evolve independently, reflecting the host's phylogeny. Whiteflies, which consist of thousands of species, harbor obligate endosymbionts of the *Portiera* genus. The divergence status of these bacteria, after a long history of coevolution with their hosts, remains ambiguous. In the present study, we aim to unravel the divergence of obligate endosymbionts from different whitefly species through genome comparison. Our findings indicate that these endosymbionts have diverged into at least three disparate genetic groups. Such findings underscore the divergence of whitefly obligate endosymbionts and provide a cue for investigation into the co-divergence between obligate endosymbionts and their hosts.



Citation: Lei, T.; Luo, N.; Song, C.; Yu, J.; Zhou, Y.; Qi, X.; Liu, Y. Comparative Genomics Reveals Three Genetic Groups of the Whitefly Obligate Endosymbiont *Candidatus Portiera aleyrodidarum*. *Insects* **2023**, *14*, 888. <https://doi.org/10.3390/insects14110888>

Academic Editor: Christian Stauffer

Received: 7 October 2023

Revised: 16 November 2023

Accepted: 16 November 2023

Published: 17 November 2023



Copyright: © 2023 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/>).

Abstract: Maternally inherited obligate endosymbionts codiverge with their invertebrate hosts and reflect their host's evolutionary history. Whiteflies (Hemiptera: Aleyrodidae) harbor one obligate endosymbiont, *Candidatus Portiera aleyrodidarum* (hereafter *Portiera*). *Portiera* was anciently acquired by whitefly and has been coevolving with its host ever since. Uncovering the divergence of endosymbionts provides a fundamental basis for inspecting the coevolutionary processes between the bacteria and their hosts. To illustrate the divergence of *Portiera* lineages across different whitefly species, we sequenced the *Portiera* genome from *Aleyrodes shizuokensis* and conducted a comparative analysis on the basic features and gene evolution with bacterial genomes from five whitefly genera, namely *Aleurodicus*, *Aleyrodes*, *Bemisia*, *Pealius*, and *Trialeurodes*. The results indicated that *Portiera* from *Bemisia* possessed significantly larger genomes, fewer coding sequences (CDSs), and a lower coding density. Their gene arrangement differed notably from those of other genera. The phylogeny of the nine *Portiera* lineages resembled that of their hosts. Moreover, the lineages were classified into three distinct genetic groups based on the genetic distance, one from *Aleurodicus* (Aleurodicinae), one from *Bemisia* (Aleyrodinae), and another from *Aleyrodes*, *Pealius*, and *Trialeurodes* (Aleyrodinae). Synonymous and nonsynonymous rate analyses, parity rule 2 plot analyses, neutrality plot analyses, and effective number of codons analyses supported the distinction of the three genetic groups. Our results indicated that *Portiera* from distant hosts exhibit distinct genomic contents, implying codivergence between hosts and their endosymbionts. This work will enhance our understanding of coevolution between hosts and their endosymbionts.

Keywords: comparative genomics; divergence; whitefly obligate endosymbiont; *Portiera*

1. Introduction

Many invertebrates rely on maternally inherited obligate endosymbionts for their viability and fecundity [1,2]. The symbionts colonize hosts' specialized bacteriocytes and are transferred to hosts' embryos [3,4]. In such cases, host and symbiont phylogenetic trees are often congruent, indicating cospeciation and synchronous diversification [5,6]. During long-term coevolution with hosts since their transmission from a free-living to an intracellular lifestyle, endosymbiotic bacteria have been exposed to distinct pressure compared with that of free-living ones, leading to an extraordinary genome evolution process [7,8]. Understanding the characteristics of endosymbiont genomes helps reveal the forces that shape the evolution of these bacterial associates.

Due to the intimate association between invertebrates and their obligate endosymbionts, the latter is often used as an auxiliary tool to resolve the phylogenetic relationships underlying the former. The endosymbiont topology of *Candidatus* Evansia muelleri, *Candidatus* Carsonella ruddii, and *Candidatus* Portiera aleyrodidarum was found to be concordant with the topology of their insect hosts Coleorrhyncha, Psylloidea, and Aleyrodidea [9]. Codivergence was also observed between *Diaphorina citri* and its obligate endosymbiont *Candidatus* Carsonella ruddii [10]. The molecular phylogenies inferred from aphid and *Buchnera* genes indicated significant congruence between aphids and *Buchnera* at generic and interspecific levels [11]. Nevertheless, whitefly *Bemisia tabaci* possessed *Portiera* with extensive genome rearrangement [12]. Considering the genomic instability of some *Portiera* lineages, we should be cautious when utilizing this endosymbiont to investigate the phylogeny of whiteflies and the divergence of *Portiera* warrants further studies.

Anciently acquired symbionts exhibit greater genome stability and slower sequence evolution than recently acquired ones. The former exhibit lower synonymous and non-synonymous substitution rates [13]. Some symbiont lineages from different host species present significant differences with regard to rates of gene evolution [14]. It remains unclear whether genomic differentiation has led to the species differentiation of bacterial lineages from closely related hosts. Investigations of codon usage bias in genomes can reveal phylogenetic relationships between organisms and the molecular evolution of genes [15]. This has been fully investigated in whitefly obligate endosymbionts.

Whitefly (Hemiptera: Aleyrodidae) is a model insect coexisting with its obligate endosymbiont *Portiera*. *Portiera*, which is strictly maternally inherited by whiteflies, has been codiverging with hosts since its origin [16,17]. Complete *Portiera* genomes have been described from *Aleurodicus*, *Bemisia*, *Pealius* and *Trialeurodes* whiteflies [12,14,16,18–20]. Comparative genomics revealed that *Bemisia* obligate endosymbionts have lost the ancestral genome order. The genome instability of *Bemisia* obligate endosymbiont might result from the loss of the DNA polymerase III proofreading subunit (*dnaQ*) [16]. The presence and absence of *dnaQ* has been determined in multiple *Portiera* strains. Nevertheless, the category of registered genomes is far from adequate to analyze *Portiera* genome instability and the loss of *dnaQ*. Moreover, codon bias signatures have not been revealed for these genome-stable and genome-unstable lineages.

In this work, we obtained the complete genome of a novel *Portiera* lineage from *Aleyrodes shizuokensis*. This is the first *Portiera* genome from *Aleyrodes*. We analyzed its gene evolution together with other genome-stable and genome-unstable lineages, aiming to reveal the divergence of *Portiera*.

2. Materials and Methods

2.1. Insect Samples and Genome Sequencing

Whitefly adults were collected from *Oxalis corniculata* in July 2020 in Hangzhou, China (30°18'32" N, 120°5'49" E). The specimens were identified as *Aleyrodes shizuokensis* Kuwana from their morphological traits [21] and further confirmed via by COI barcode data blasted in NCBI [22]. The insects were kept in 75% ethanol by volume at −80 °C until DNA extraction. Total genomic DNA was extracted from a single female adult using the DNeasy Blood & Tissue Kit (QIAGEN, Valencia, CA, USA) following the manufacturer's instructions

with slight modifications. In brief, the insects were homogenized in 54 µL of Buffer ATL followed by incubation with 6 µL of proteinase K at 56 °C for 6 h to lyse the tissue. Then, equivalent volumes of Buffer AL and ethanol were successively added to the sample. The mixture was centrifuged at 6000× *g* for 1 min at room temperature and washed using 200 µL of Buffer W1 and 200 µL of Buffer W2 successively. Finally, the DNA was eluted with 60 µL of Buffer AE and then stored at −20 °C for later use. The isolated DNA was used for a PCR reaction and whole-genome shotgun sequencing. The *COI* barcode was amplified using universal PCR primers of C1-J-2195 and TL2-N-3014 [23]. PCR was performed in a 20 µL reaction volume, including 1 U rTaq (Takara), 2 µL of 10 × PCR Buffer (Mg²⁺ Plus), 3.2 µL of dNTPs (2.5 mmol/L each), 0.2 µL of each primer (20 µmol/L), 12.2 µL of ddH₂O, and 2 µL of DNA template. The PCR reaction program was initialized at 94 °C for 3 min, followed by 34 cycles of 96 °C for 30 s, 52 °C for 30 s, 72 °C for 90 s, with a final extension for 10 min at 72 °C. The remaining DNA was fragmented via sonication and used to construct short-insert libraries (insert size of 150 bp) using the Illumina TruSeq Nano DNA Library Preparation Kit. Sequencing was performed on the Illumina HiSeq 4000 platform with a PE150 strategy.

2.2. Genome Assembly and Annotation

Raw data generated from the Illumina libraries were trimmed, and the adapters were filtered using Trimmomatic (version 0.39) [24]. The quality of the clean data was assessed using FastQC (version 0.11.9) (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>; last accessed 30 October 2023). To screen out candidate reads belonging to *Portiera*, clean data were mapped to *Portiera* complete genomes available from NCBI (Table 1) using Bowtie2 (version 2.3.5.1) [25]. Draft genomes and those from the same host species were ignored to reduce computational complexity. The selected reads were assembled de novo using SPAdes (version 3.14.1) [26], and the assembled contigs and scaffolds were evaluated using QUAST (version 5.0.2) [27]. To discard sequences not belonging to *Portiera*, a minimum coverage of 200 was required. One scaffold consisting of five contigs was generated. PCR reaction was used to close gaps between contigs. PCR primers and procedures are provided in Table S1. The assembled sequence of the bacterial genome was annotated using Prodigal, RNAmmer, Aragorn, SignalP and Infernal integrated in the Prokka tool [28]. The gene *dnaK* was identified and used to set the coordinates of the genome. Gene synteny between *Portiera* genomes, based on nucleotide sequences, was pairwise analyzed and plotted with the Python package MCscan (<https://github.com/tanghaibao/mcscan>; last accessed 30 October 2023) [29].

Table 1. Complete *Candidatus Portiera aleyrodidarum* genomes used in this study.

Portiera ID	Host Subfamily	Host Species	Accession Number ¹	Reference
BTB	Aleyrodinae	<i>Bemisia tabaci</i> MEAM1	CP003708	[19]
BTQ	Aleyrodinae	<i>Bemisia tabaci</i> MED	CP003835	[18]
BTZ1	Aleyrodinae	<i>Bemisia tabaci</i> Asia II 3	CP016327	[20]
BTZ3	Aleyrodinae	<i>Bemisia tabaci</i> China1	CP016343	[20]
AdSh	Aleyrodinae	<i>Aleyrodes shizuokensis</i>	GWHBOVO00000000 ²	This study
PeMo	Aleyrodinae	<i>Pealius mori</i>	LR744089	[16]
TrVa	Aleyrodinae	<i>Trialeurodes vaporariorum</i>	CP004358	[12]
AlDi	Aleurodicinae	<i>Aleurodicus dispersus</i>	LN649255	[14]
AlFl	Aleurodicinae	<i>Aleurodicus floccissimus</i>	LN734649	[14]

¹ The genomes are deposited at GenBank of NCBI unless otherwise stated. ² The genome of AdSh is deposited at the National Genomics Data Center, China National Center for Bioinformation, Chinese Academy of Sciences.

2.3. Phylogenetic Relationships

Protein sequence from the *Portiera* genomes were used as input for OrthoFinder (version 2.5.4) [30] to identify orthologous genes. Upset plots of single-copy orthologous genes were drawn by the UpSetR (version 1.4.0) [31] and plotrix (version 3.8-2;

<https://www.rdocumentation.org/packages/plotrix/versions/3.8-2>; last accessed 30 October 2023) packages in R. All single-copy orthologous genes were used for the reconstruction of phylogenetic trees. Sequences of individual genes from the *Portiera* lineages were aligned using MAFFT (version 7.490) [32], followed by eliminating poorly aligned positions using trimAL (version 1.4) [33]. Then, the aligned sequences were concatenated using phylo-tools package in R (version 0.2.2; <https://github.com/helixcn/phylo-tools>; last accessed 30 October 2023). The concatenated datasets were subjected to IQ-TREE (version 2.2.0) for model selection under AICc criterion [34,35]. A maximum likelihood (ML) phylogenetic tree was reconstructed using the inferred model with 1 000 replicates of ultrafast bootstrap in IQ-TREE [36,37]. A Bayesian inference (BI) tree was reconstructed using the inferred model in MrBayes (version 3.2) with a chain length of 10,000,000 [38]. By running BLAST against nr database in NCBI, we chose *Zymobacter palmae* (accession number AP018933) as an outgroup because 16S rRNA genes of *Z. palmae* and *Portiera* had the highest similarity.

To reconstruct the phylogenetic tree of *Portiera* hosts, whitefly mitochondrial genomes were downloaded from NCBI (Table 2). The COI gene was used to set the coordinates of the mitochondrial genomes. Then, the nucleotide sequences were aligned using MAFFT and trimmed using trimAL. Model selection and phylogenetic tree reconstruction methods resembled those of *Portiera*. *Diaphorina citri* (accession number NC_030214) [39] was used as an outgroup because of its close relationship with whiteflies [9].

Table 2. Whitefly mitochondrial genomes used in this study.

Subfamily	Species	Accession Number	Reference
Aleyrodinae	<i>Bemisia tabaci</i> MEAM1	KR559508	NA ¹
Aleyrodinae	<i>Bemisia tabaci</i> MED	JQ906700	[40]
Aleyrodinae	<i>Bemisia tabaci</i> China1	KR559506	NA ¹
Aleyrodinae	<i>Aleurocanthus camelliae</i>	KU761949	[41]
Aleyrodinae	<i>Aleurocanthus spiniferus</i>	KJ437166	[42]
Aleyrodinae	<i>Aleyrodes shizuokensis</i>	MT880225	[43]
Aleyrodinae	<i>Pealius mori</i>	LR877884	NA ¹
Aleyrodinae	<i>Tetraleurodes acaciae</i>	AY521262	[44]
Aleyrodinae	<i>Trialeurodes vaporariorum</i>	AY521265	[44]
Aleurodicinae	<i>Aleurodicus dispersus</i>	KR063274	[45]
Aleurodicinae	<i>Aleurodicus dugesii</i>	AY521251	[44]

¹ NA, reference not available; downloaded from NCBI.

2.4. Genetic Distance

The genetic distance of *Portiera* lineages was evaluated using three indicators, the average nucleotide identity (ANI), the amino acid identity (AAI) and Mash distance. The ANI and AAI values of the single-copy orthologous genes were pairwise-compared using the fastANI package in R [46] and the EzAAI package in Java [47]. Visualization was performed using the R package corrplot. Mash distance values and a Mash-based phylogenetic tree were estimated based on concatenated datasets using Mashtree (version 1.2.0) with 1 000 replicates of ultrafast bootstrap [48]. Based on the results of genetic distance analysis, the nine *Portiera* lineages were divided into three genetic groups. Group BM consisted of *Portiera* from *B. tabaci* MEAM1 (BTB), *B. tabaci* MED (BTQ), *B. tabaci* Asia II 3 (BTZ1) and *B. tabaci* China1 (BTZ3). Group AD consisted of *Portiera* from *Aleyrodes shizuokensis* (AdSh), *P. mori* (PeMo) and *T. vaporariorum* (TrVa). Group AL consisted of *Portiera* from *Aleurodicus dispersus* (AlDi) and *Aleurodicus floccissimus* (AlFi). The whiteflies' Mash distance based on mitochondrial genomic data was also estimated using Mashtree with default parameters.

2.5. Codon Usage Bias

The codon usage bias of three *Portiera* genetic groups was comparatively analyzed. The *Portiera* of the *B. tabaci* cryptic species complex presented the highest ANI and AAI values. One of them, BTZ3, was used as a reference to evaluate nucleotide substitution

rates. K_a indicated the rate of nonsynonymous nucleotide substitution per nonsynonymous site. K_s indicated the rate of synonymous substitution per synonymous site. K_a and K_s values were calculated using the KaKs_Calculator 2.0 toolkit [49]. Bias in the third codon letters of each genome or genetic group was analyzed using a parity rule 2 (PR2) bias plot. A3, T3, C3, and G3 indicated the nucleotide content of A, T, C, and G, respectively, at the third base of codons in each gene. Their values were calculated using CodonW (version 1.4.4) (<https://sourceforge.net/projects/codonw/>; last accessed 30 October 2023). In PR2 bias plot maps, $G3/(G3 + C3)$ and $A3/(A3 + T3)$ were set as the abscissa and ordinate, respectively [50]. The GC content of the first, second and third bases of codons (GC1, GC2 and GC3) for each gene was calculated in Python, and a linear regression of GC12 and GC3 was performed in R to conduct a neutrality plot analysis, where GC12 represented the average of GC1 and GC2 [51]. To reveal the relationship between nucleotide composition and codon bias, the GC content of the third synonymous position (GC3s) and effective number of codons (ENC) generated by CodonW were used to conduct ENC plot analysis. In ENC plot maps, ENC values were compared to the theoretical curve $ENC_{exp} = 2 + GC3s + 29/[GC3s^2 + (1 - GC3s)^2]$ [52].

3. Results

3.1. The Genomic Features of *Portiera*

The genome of AdSh from *A. shizuokensis* is composed of a single circular chromosome with an average coverage of $317\times$. The genome is 283,014 bp in length and contains 268 coding sequences (CDSs), three rRNAs, 34 tRNAs, one tmRNA and one RnaseP RNA. It exhibits a low GC content (23.8%) and a high coding density (91.4%) (Figure 1A). Functional *dnaQ* is observed in AdSh genome. These features are roughly similar to those of previously sequenced *Portiera* genomes except those of the *Bemisia tabaci* cryptic species complex [16]. Based on the basic genomic features of *Portiera*, the nine genomes are clustered into two groups. Genomes of BTB, BTQ, BTZ1 and BTZ3 from *B. tabaci* present lower coding density (67.6–67.9%), fewer CDSs (247–267) and larger sizes (349.1 kb–350.1 kb). Genomes from the other whiteflies present higher coding density (91.7–94.9%), more CDSs (266–279) and smaller sizes (271.2 kb–283.6 kb) (Figure 1B). Synteny analysis reveals two *Portiera* gene arrangement patterns, one from *B. tabaci* and the other from other whiteflies (Figure 1C). *Portiera* lineages from *B. tabaci* have lost the ancestral genome order as described by Santos-Garcia et al., 2020 [16].

3.2. The Phylogeny of *Portiera* and Hosts

A total of 218 orthogroups were present in all *Portiera* lineages, among which 198 were single-copy orthogroups. Eighteen orthogroups were absent in the *Portiera* of *B. tabaci*, and eight were present only in those of *B. tabaci*. Six orthogroups were present only in those of Aleurodicinae and three only in those of Aleyrodinae (Figure 2). The results indicated that the *Portiera* of *B. tabaci* differed from those of other whiteflies in gene content. And the *Portiera* of Aleurodicinae differed from those of Aleyrodinae.

BI and ML trees were obtained using *Portiera* single-copy orthologous genes. The two trees shared a common topology where *Portiera* formed four major clades. One clade consisted of AIDi and AIFi, the endosymbionts of Aleurodicinae. Other clades consisted of endosymbionts of Aleyrodinae. This topology was consistent with the tree based on two rRNA (16S and 23S) and three coding genes (*groEL*, *rpoD*, and *dnaK*) [16]. BI and ML trees were also inferred based on whiteflies' mitochondrial genomes. Similarly, *Aleurodicus dispersus* and *Aleurodicus dugesii* from Aleurodicinae formed one clade, and most Aleyrodinae species formed another huge clade. Intriguingly, *Aleyrododes shizuokensis* from Aleyrodinae was found to be closely related to Aleurodicinae clade. Despite this species and its *Portiera* endosymbiont, *Portiera*-based and mitogenome-based trees were mostly consistent in topology (Figure 3). The divergence between the two topologies implied that the evolution of endosymbionts and hosts might not be completely parallel.

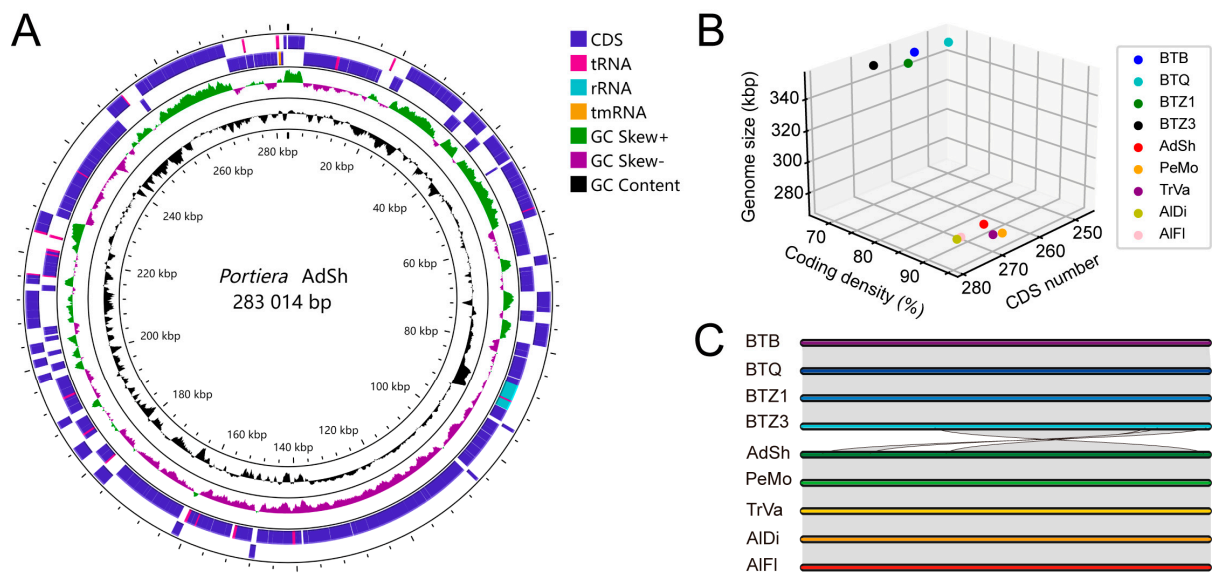


Figure 1. Genome features of *Portiera* lines. (A) Circular map of *Portiera* AdSh genome. From outer to inner: genes, GS skew, GC content, and scale bar. (B) Diagram of *Portiera* genome size, coding density and coding sequences (CDSs). BTB, BTQ, BTZ1, BTZ3, AdSh, AIDi, AIFI, PeMo and TrVa represent the *Portiera* lines from *Bemisia tabaci* MEAM1, *B. tabaci* MED, *B. tabaci* Asia II 3, *B. tabaci* China1, *Aleyrodes shizuokensis*, *Aleyrodicus disperses*, *Aleyrodicus floccissimus*, *Pealius mori* and *Trialeurodes vaporariorum*, respectively. (C) Genome synteny of *Portiera* lines. Genomic protein-coding nucleotide sequences were used to conduct synteny analyses. The synteny was pairwise-analyzed, and the results were merged in one figure.

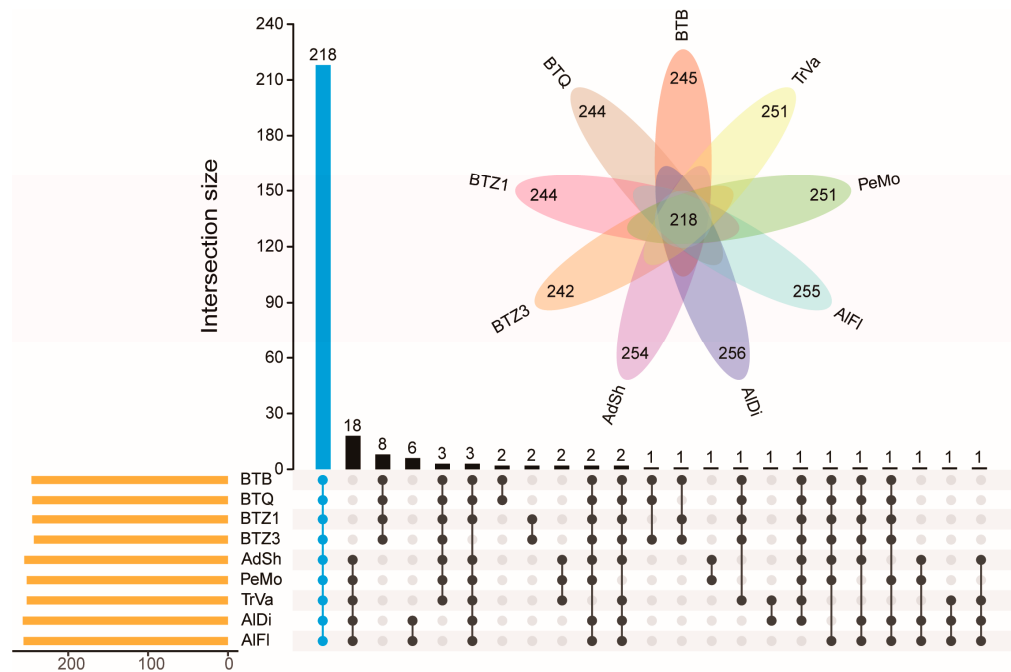


Figure 2. UpSet plot of orthogroups from *Portiera* genomes. The orthogroups present in all lineages are colored blue. Those present in partial lineages are colored black. The numbers of orthogroups in each lineage are indicated in the Venn diagram.

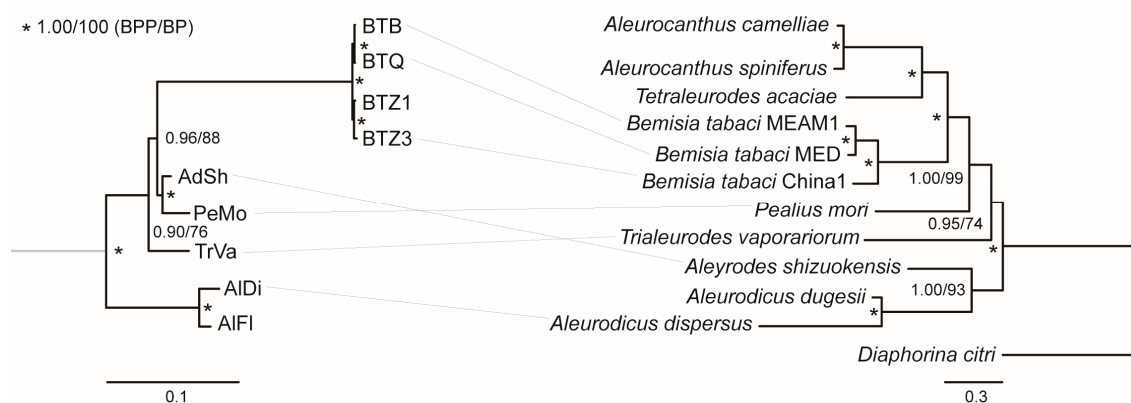


Figure 3. Bayesian inference (BI) and maximum likelihood (ML) trees inferred from *Portiera* single-copy orthologous genes (left) and whiteflies' mitochondrial genomes (right). Topologically consistent BI and ML trees are integrated. Bayesian posterior probabilities (BPP) and ultrafast bootstrap support values (BP) are given at the nodes. *Portiera* lineages and corresponding host species are connected by straight lines. *Zymobacter palmae* AP018933 is used as the outgroup for *Portiera* but is not displayed for plotting reasons.

3.3. Genetic Distance

The results of ANI and AAI analysis indicated that bacterial genome similarities were mostly congruous with their insect hosts, except for *B. tabaci* endosymbionts. Specifically, phylogenetically closely related hosts harbored endosymbionts with high ANI values (AIDi vs. AIFI, 0.98; AdSh vs. PeMo, 0.98; AdSh vs. TrVa, 0.96; PeMo vs. TrVa, 0.95) and AAI values (AIDi vs. AIFI, 0.98; AdSh vs. PeMo, 0.96; AdSh vs. TrVa, 0.93; PeMo vs. TrVa, 0.91). Between symbionts from distantly related Aleurodicinae and Aleyrodinae hosts, the ANI and AAI values were no more than 0.90 and 0.85, respectively. Additionally, *B. tabaci* cryptic species harbored endosymbionts with the highest ANI and AAI values, both over 0.99. Although *B. tabaci* belonged to Aleyrodinae, their endosymbionts shared low ANI and AAI values with other Aleyrodinae endosymbionts (ANI, 0.86–0.88; AAI, 0.79–0.82). The values were even lower than those shared by other Aleyrodinae and Aleurodicinae endosymbionts (ANI, 0.88–0.90; AAI, 0.84–0.85) (Figure 4A). The ANI and AAI of shared genes between two prokaryotes provide a robust means to assess genetic relatedness among lineages. ANI and AAI values over 95% are necessary conditions for species definition in prokaryotes [53,54]. In this study, due to the lack of other sufficient conditions for species definition, we used the criteria ANI > 95% and AAI > 90% to divide the *Portiera* lineages into three genetic groups, one from Aleurodicinae whiteflies (hereafter AL) and two from Aleyrodinae whiteflies. Aleyrodinae whiteflies harbor two distinct *Portiera* genetic groups, one from the *Bemisia* genus (hereafter BM) and the other from other genera (hereafter AD). Mash distance analysis supported the division of the three genetic groups. The within-group distance (0.0017–0.0454) was much shorter than the between-group distance (0.0889–0.1766) (Figure 4B). Additionally, the *Portiera* genomic distance was much shorter than whiteflies' mitogenomic distance, indicating a slower divergence rate of *Portiera* lineages than their hosts (Table S2).

3.4. Codon Usage Bias

3.4.1. Synonymous and Nonsynonymous Substitution Rates

The K_a and K_s values were used to test for signatures of natural selection. This index was used to assess the difference in selective pressure by *Portiera* genetic group and lineage. Taking BTZ3 as a reference, most genes in the BM group have not undergone synonymous or nonsynonymous mutations. The *B. tabaci* cryptic species diverged recently, the other whitefly species diverged a long time ago [16]. Mutations might have no time to occur or become fixed in *Portiera* from *B. tabaci* species while that happened in the other species. In

AD and AL groups, K_a values were observed to be much smaller than K_s values for most of the genes, indicating the substantial influence of purifying selection (Figure 5). Compared with AD group, AL exhibited elevated synonymous and nonsynonymous substitution rates (Table S3). The results supported the divergence of the three genetic groups we defined.

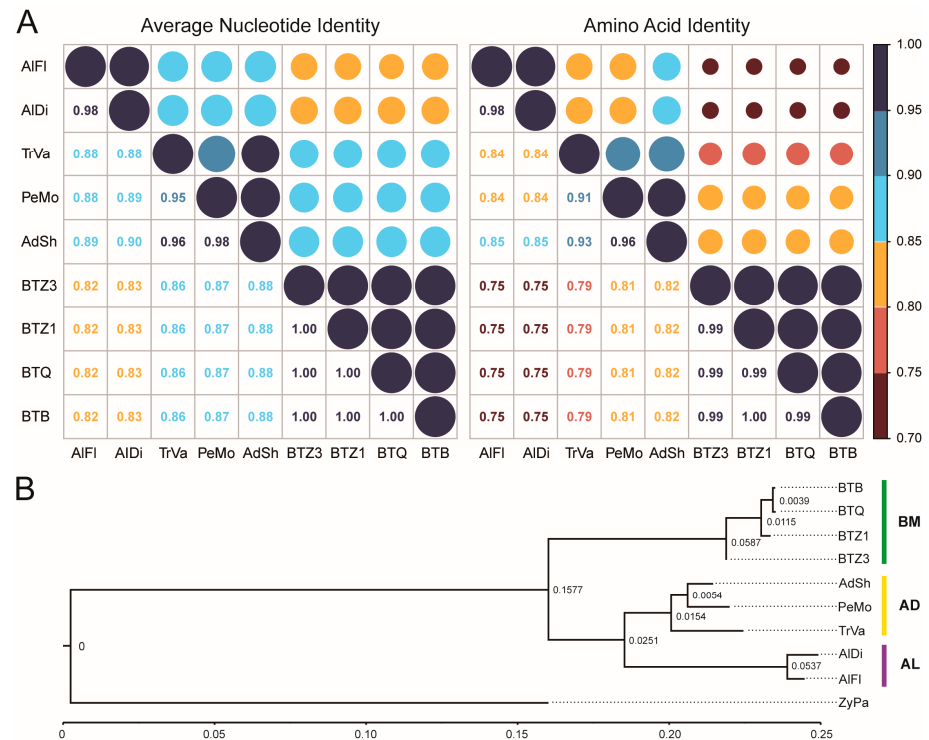


Figure 4. Genetic distance of *Portiera* lineages. **(A)** Matrix diagram of average nucleotide identity (ANI) and amino acid identity (AAI) among *Portiera* genomes. The values of ANI and AAI are shown by both scores and circles, where the values are in accordance with the area of circles and the element colors according to the legend. **(B)** Mash-based phylogenetic tree derived from nine *Portiera* lineages. *ZyPa*, *Zymobacter palmae* AP018933, is used as outgroup to root the tree. Branch lengths representing Mash distance values are indicated at the nodes.

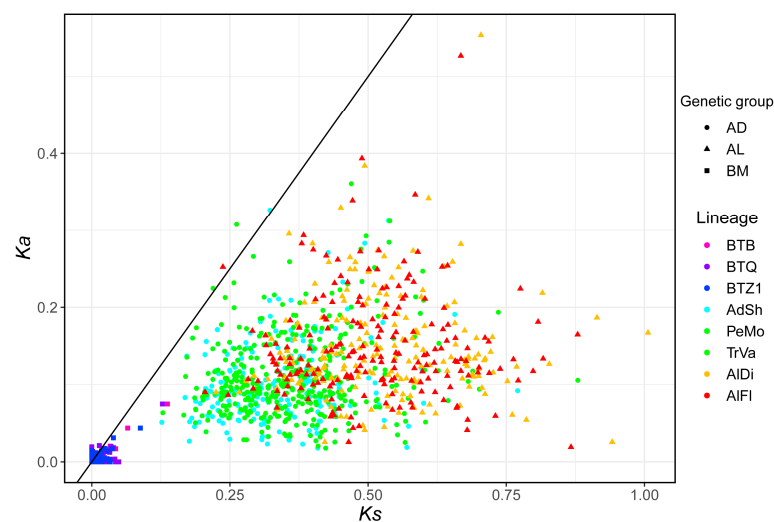


Figure 5. Distribution of K_a and K_s values in *Portiera* genomes. K_a , the rate of nonsynonymous nucleotide substitution per nonsynonymous site; K_s , the rate of synonymous substitution per synonymous site. *Portiera* K_a and K_s values are calculated with reference to BTZ3. *Portiera* lineages are distinguished by plot colors, and genetic groups are distinguished by plot shapes.

3.4.2. PR2 Bias Analysis

Bias in the third codon letters of each genetic group or genome was analyzed using PR2 bias analysis to evaluate their genetic difference. In the PR2 bias plot graphs, $G3/(G3 + C3)$ and $A3/(A3 + T3)$ were set as the abscissa and ordinate, respectively. The graphs are divided into four quadrants by the lines $G3/(G3 + C3) = 0.5$ and $A3/(A3 + T3) = 0.5$. The majority of gene points are distributed in the quadrants $G3/(G3 + C3) > 0.5$ and $A3/(A3 + T3) > 0.5$, indicating significant G and A bias at the third codon letter (Figures 6 and S1). Furthermore, the distribution patterns of genes varied among *Portiera* genetic groups. The plots of the BM group were the most concentrated, while those of the AL group were the most scattered in the graph, implying that PCGs of the BM group were stricter with the content of bases at the third position of the codons.

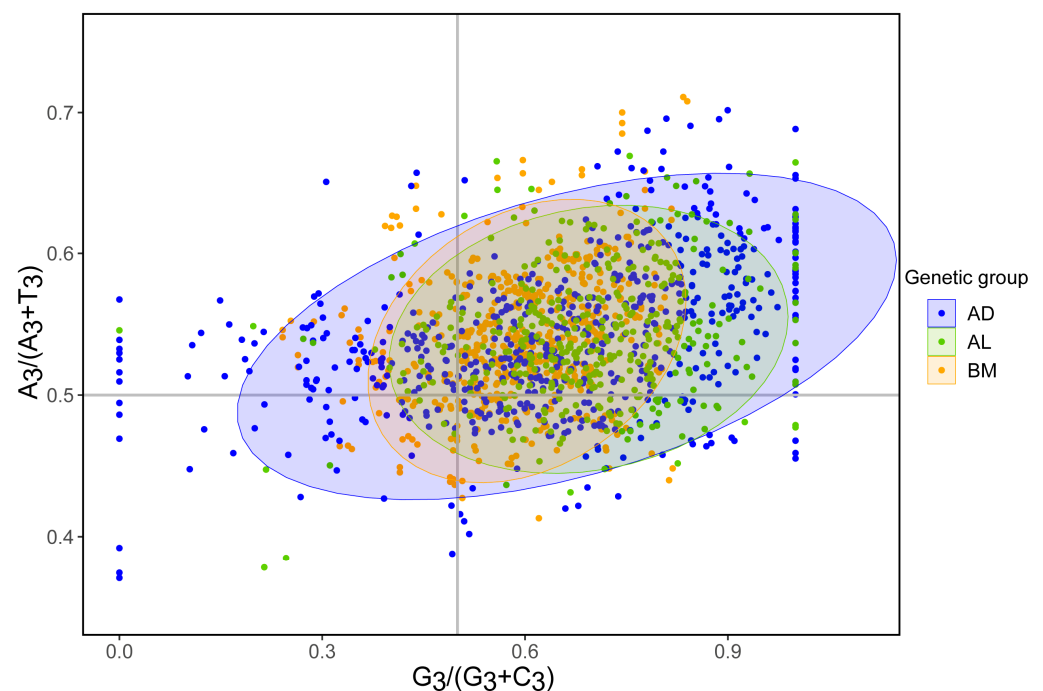


Figure 6. Parity rule 2 (PR2) bias plot analysis of *Portiera* genetic groups. A3, T3, C3, and G3 indicate the nucleotide contents of A, T, C, and G, respectively, at the third base of codons in each gene. The genes are plotted on the graphs with $G3/(G3 + C3)$ on the abscissa and $A3/(A3 + T3)$ on the ordinate. The graphs and plots are divided into four quadrants by the lines $G3/(G3 + C3) = 0.5$ and $A3/(A3 + T3) = 0.5$. The oval shadows represent the 95% confidence interval. PR2 bias plot analysis of individual lineages is provided in Figure S1.

3.4.3. Neutrality Plot Analysis

Neutrality plot analysis was performed to estimate the extent of gene directional mutation pressure against selection. The results helped to reveal the genetic diversification of *Portiera* genetic groups and lineages. Linear regression analysis between GC12 and GC3 indicated that correlation coefficients (R^2) were low in all the genetic groups. This implies that there is no significant correlation between GC12 and GC3. Thus, the influence of natural selection on codon usage bias outweighs the impact of mutation. The correlation coefficient of the BM group ($R^2 = 0.02181$) was the lowest among the three genetic groups, implying that GC12 and GC3 of BM genes were the least correlated (Figures 7 and S2). Thus, BM group suffered more selection than the AD and AL groups.

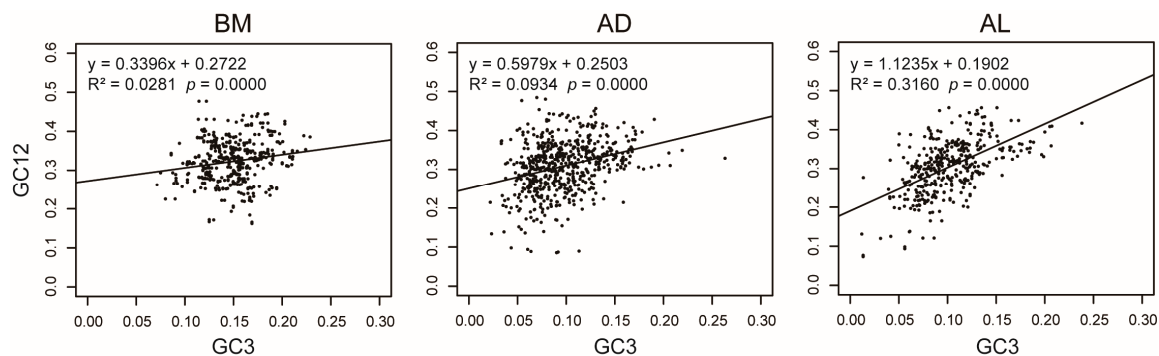


Figure 7. Neutrality plot analysis of *Portiera* genetic groups. GC3, GC content of the third base of the codons in each gene; GC12, average GC content of the first and second bases of the codons in each gene. The values of GC3 and GC12 are indicated at the abscissae and ordinates, respectively. Linear regression analysis between GC12 and GC3 was performed using R. Neutrality plot analysis of individual lineages is provided in Figure S2.

3.4.4. ENC Plot Analysis

ENC plot analysis was used to investigate the codon usage patterns across the *Portiera* genetic groups and lineages to reveal their genetic diversification. The relationship between the theoretical ENC value (ENC_{exp}) and GC3s under H_0 (no selection) was approximated by $ENC_{exp} = 2 + GC3s + 29/[GC3s^2 + (1 - GC3s)^2]$. Points beneath the curve by more than a 5% error margin imply selection pressure on the genes. In the graphs, most of the points are below the curve for all the genetic groups, indicating selection pressure on these genes (Figures 8 and S3). The BM group was found to possess the highest GC3s and ENC values. Nevertheless, there was no obvious difference between the AD and AL groups (Table S4). That was probably because the divergence among them was not large enough to be distinguished via the ENC plot analysis. *Buchnera* lineages displaying notable variation in the PR2 analysis also exhibited comparable ENC analysis results, indicating a lack of discernible differences [8].

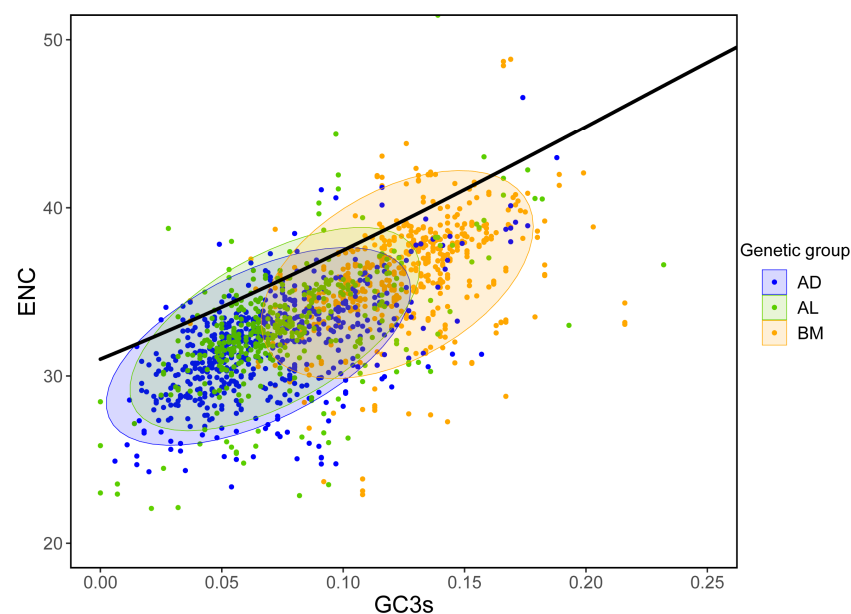


Figure 8. Effective number of codons (ENC) plot analysis of *Portiera* genetic groups. GC3s indicates the GC content at the third synonymous position. ENC_{exp} indicates the theoretical ENC value. The continuous curve $ENC_{exp} = 2 + GC3s + 29/[GC3s^2 + (1 - GC3s)^2]$ represents the relationship between ENC_{exp} and GC3s under H_0 (no selection). The oval shadows represent the 95% confidence interval. ENC plot analysis of individual lineages is provided in Figure S3.

4. Discussion

The whitefly obligate endosymbiont *Portiera* established symbiosis with its host before the origin of Aleyrodidae and has been codiverging with whiteflies ever since [16]. Because of the strict mother-to-offspring transmission, *Portiera* does not switch to other host species [55]. Thus, *Portiera* reflect both their own and their host's phylogenetic relationships [44]. In this work, the phylogenetic tree inferred from bacterial single-copy orthologous genes was consistent with their hosts' phylogenetic relationships. In contrast, many facultative endosymbionts have been demonstrated to be horizontally transmitted, leading to the same symbiont residing in multiple host species [56,57]. Furthermore, although it has not been reported, one facultative endosymbiont might be horizontally transferred to another host and then back to its former host. Transferable facultative endosymbionts coexist with multiple hosts and experience a shorter codivergence period with their present host; thus, their phylogeny cannot reflect that of their hosts [57]. Relative to facultative endosymbionts, maternally inherited obligate endosymbionts are more reliable for the phylogenetic analysis of insect hosts.

The genome evolution of endosymbionts is shaped by different selection intensities that may reflect the different ages and metabolic roles of symbiont types. Compared with the recently acquired sharpshooter symbiont *Candidatus* *Baumannia cicadellinicola*, the anciently acquired *Candidatus* *Sulcia muelleri* exhibits much greater genome stability and slower sequence evolution [13]. Intraspecific genetic variation in hosts also affects the regulation of obligate symbionts, leading to lineage-specific patterns of genome evolution [58,59]. *Portiera* is an anciently acquired endosymbiont of the whitefly and has codiverged with its host for millions of years. Our results indicated that *Portiera* genomes are stable in size, gene content and gene order, except those from *B. tabaci*. *Portiera* of *B. tabaci* formed a distinct genetic group. The genome instability of the *B. tabaci* endosymbiont might result from the absence of *dnaQ* [16]. We assembled a *Portiera* genome from the *Aleyrodes* genus, which contains *dnaQ*, and found that the general genome features were similar to those of other genome-stable *Portiera* lineages. Previous suggestions that *dnaQ* might be associated with *Portiera* genome instability were sustained.

Maternally inherited obligate endosymbionts codiverge with insect hosts. During the process of host species differentiation, it is unknown whether the symbiotic bacteria split into different species accordingly. Differences in genome size among symbiont lineages from closely related host species have been reported for *Buchnera* of aphids, *Blochmannia* of carpenter ants, and *Portiera* of whiteflies [60,61]. In this work, we compared general genomic features, genetic distance, and codon usage bias to assess the genetic variation in *Portiera*. Based on these results, three *Portiera* genetic groups, AL, AD, and BM, were defined. However, the variation was not qualified to define a new species.

Codon usage bias patterns of *Buchnera* genomes from different aphid subfamily varied. Pemphigidae possessed *Buchnera* with a more complete loss of codon bias, stronger strand asymmetry, and more reduced genome than Aphididae does [62]. The *Buchnera* codon usage pattern mostly resulted from mutation pressure [8]. In this study, mutation was also found to play a crucial role in shaping the codon usage of the three *Portiera* genetic groups we defined, while the BM group suffered more selection than the AD and AL groups. Obvious codon usage bias differences were observed among the genetic groups. Synonymous and nonsynonymous nucleotide substitution rates within genetic groups were lower than those between the genetic groups. The data in this study are necessary but not sufficient to define the genetic groups as species. However, we can tell that *Portiera* species differentiation is slower than that of their hosts.

5. Conclusions

We provided a complete genome of *Portiera*, the obligate endosymbiont of *Aleyrodes shizuokensis*. Comparative genomics was performed with eight other *Portiera* genomes. Our results revealed that these *Portiera* lineages have differentiated into at least three genetic groups, one of which resulted from the loss of stability during its evolutionary

history. Obligate endosymbionts reflect their hosts' evolutionary history. The divergence of endosymbionts should be investigated to further illustrate the speciation of their hosts.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/insects14110888/s1>, Table S1: PCR primers and procedures for completing AdSh genome; Table S2: Mash distance of *Portiera* genomes and whitefly mitochondrial genomes; Table S3: Synonymous and nonsynonymous substitution rates of *Portiera* genomes; Table S4: Codon nucleotide content and ENC values of *Portiera* genomes; Figure S1: Parity rule 2 bias plot analysis of *Portiera* lineages; Figure S2: Neutrality plot analysis of *Portiera* lineages; Figure S3: Effective number of codons plot analysis of *Portiera* lineages.

Author Contributions: Conceptualization, X.Q. and Y.L.; methodology, T.L. and N.L.; software, C.S., J.Y. and Y.Z.; validation, N.L. and C.S.; formal analysis, T.L. and N.L.; investigation, T.L. and Y.Z.; resources, Y.L.; data curation, T.L. and N.L.; writing—original draft preparation, T.L.; writing—review and editing, N.L. and Y.Z.; visualization, J.Y.; supervision, C.S.; project administration, X.Q. and Y.L.; funding acquisition, X.Q. and Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (NSFC, Grant No. 32070481), the Earmarked Fund for China Agriculture Research System (Grant No. CARS-23-C05), the Zhejiang Provincial Natural Science Foundation of China (Grant No. LY22C040003), and the Science & Technology Project of Taizhou (Grant No. 21hb04, 21nya17, 1902gy23).

Data Availability Statement: The data supporting the findings of this study are openly available from the National Genomics Data Center, China National Center for Bioinformation (CNCB-NGDC), Chinese Academy of Sciences at <https://ngdc.cncb.ac.cn> (accessed on 16 November 2023) [63]. The complete genome of *Aleyrodes shizuokensis* obligate endosymbiont *Portiera* lineage AdSh has been deposited in CNCB-NGDC under accession number GWHBOVO00000000. The associated BioProject, BioSample and GSA accession numbers are PRJCA012377, SAMC916243 and CRA008827.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Moran, N.A.; Telang, A. Bacteriocyte-associated symbionts of insects. *BioScience* **1998**, *48*, 295–304. [\[CrossRef\]](#)
2. Douglas, A.E. How multi-partner endosymbiosis function. *Nat. Rev. Microbiol.* **2016**, *14*, 731–743. [\[CrossRef\]](#)
3. Bright, M.; Bulgheresi, S. A complex journey: Transmission of microbial symbionts. *Nat. Rev. Microbiol.* **2010**, *8*, 218–230. [\[CrossRef\]](#)
4. Luan, J.-B.; Shan, H.-W.; Isermann, P.; Huang, J.-H.; Lammerding, J.; Liu, S.-S.; Douglas, A.E. Cellular and molecular remodelling of a host cell for vertical transmission of bacterial symbionts. *Proc. R. Soc. B* **2016**, *283*, 20160580. [\[CrossRef\]](#)
5. Moran, N.A.; Baumann, P. Bacterial endosymbionts in animals. *Curr. Opin. Microbiol.* **2000**, *3*, 270–275. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Moeller, A.H.; Sanders, J.G.; Sprockett, D.D.; Landers, A. Assessing co-diversification in host-associated microbiomes. *J. Evol. Biol.* **2023**, online. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Wernegreen, J.J. Genome evolution in bacterial endosymbionts of insects. *Nat. Rev. Genet.* **2002**, *3*, 850–861. [\[CrossRef\]](#)
8. Ding, H.; Gao, J.; Yang, J.; Zhang, S.; Han, S.; Yi, R.; Ye, Y.; Kan, X. Genome evolution of *Buchnera aphidicola* (Gammaproteobacteria): Insights into strand compositional asymmetry, codon usage bias, and phylogenetic implications. *Int. J. Biol. Macromol.* **2023**, *253*, 126738. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Santos-Garcia, D.; Latorre, A.; Moya, A.; Gibbs, G.; Hartung, V.; Dettner, K.; Kuechler, S.M.; Silva, F.J. Small but powerful, the primary endosymbiont of moss bugs, *Candidatus Evansia muelleri*, holds a reduced genome with large biosynthetic capabilities. *Genome Biol. Evol.* **2014**, *6*, 1875–1893. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Wang, Y.; Lu, J.; Beattie, G.A.; Islam, M.R.; Om, N.; Dao, H.T.; Nguyen, L.V.; Zaka, S.M.; Guo, J.; Tian, M.; et al. Phylogeography of *Diaphorina citri* (Hemiptera: Liviidae) and its primary endosymbiont, '*Candidatus Carsonella ruddii*': An evolutionary approach to host-endosymbiont interaction. *Pest Manag. Sci.* **2018**, *74*, 2185–2194. [\[CrossRef\]](#)
11. Liu, L.; Li, X.-Y.; Huang, X.-L.; Qiao, G.-X. Evolutionary relationships of *Pemphigus* and allied genera (Hemiptera: Aphididae: Eriosomatinae) and their primary endosymbiont, *Buchnera aphidicola*. *Insect Sci.* **2014**, *21*, 301–312. [\[CrossRef\]](#)
12. Sloan, D.B.; Moran, N.A. The evolution of genomic instability in the obligate endosymbionts of whiteflies. *Genome Biol. Evol.* **2013**, *5*, 783–793. [\[CrossRef\]](#)
13. Bennett, G.M.; McCutcheon, J.P.; MacDonald, B.R.; Romanovicz, D.; Moran, N.A. Differential genome evolution between companion symbionts in an insect-bacterial symbiosis. *mBio* **2014**, *5*, e01697-01614. [\[CrossRef\]](#)
14. Santos-Garcia, D.; Vargas-Chavez, C.; Moya, A.; Latorre, A.; Silva, F.J. Genome evolution in the primary endosymbiont of whiteflies sheds light on their divergence. *Genome Biol. Evol.* **2015**, *7*, 873–888. [\[CrossRef\]](#) [\[PubMed\]](#)

15. Parvathy, S.T.; Udayasuriyan, V.; Bhadana, V. Codon usage bias. *Mol. Biol. Rep.* **2022**, *49*, 539–565. [\[CrossRef\]](#)
16. Santos-Garcia, D.; Mestre-Rincon, N.; Ouvrard, D.; Zchori-Fein, E.; Morin, S. *Portiera* gets wild: Genome instability provides insights into the evolution of both whiteflies and their endosymbionts. *Genome Biol. Evol.* **2020**, *12*, 2107–2124. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Xu, X.-R.; Li, N.-N.; Bao, X.-Y.; Douglas, A.E.; Luan, J.-B. Patterns of host cell inheritance in the bacterial symbiosis of whiteflies. *Insect Sci.* **2020**, *27*, 938–946. [\[CrossRef\]](#)
18. Santos-Garcia, D.; Farnier, P.-A.; Beitia, F.; Zchori-Fein, E.; Vavre, F.; Mouton, L.; Moya, A.; Latorre, A.; Silva, F.J. Complete genome sequence of “*Candidatus Portiera aleyrodidarum*” BT-QVLC, an obligate symbiont that supplies amino acids and carotenoids to *Bemisia tabaci*. *J. Bacteriol.* **2012**, *194*, 6654–6655. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Sloan, D.B.; Moran, N.A. Endosymbiotic bacteria as a source of carotenoids in whiteflies. *Biol. Lett.* **2012**, *88*, 986–989. [\[CrossRef\]](#)
20. Zhu, D.-T.; Zou, C.; Ban, F.-X.; Wang, H.-L.; Wang, X.-W.; Liu, Y.-Q. Conservation of transcriptional elements in the obligate symbiont of the whitefly *Bemisia tabaci*. *PeerJ* **2019**, *7*, e7477. [\[CrossRef\]](#)
21. Chen, C.H.; Dubey, A.K.; Ko, C.C. Comparative morphological studies on two species of *Aleyrodes* (Hemiptera: Aleyrodidae). *Pan-Pac. Entomol.* **2007**, *83*, 244–254. [\[CrossRef\]](#)
22. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Simon, C.; Frati, F.; Beckenbach, A.; Crespi, B.; Liu, H.; Flook, P. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* **1994**, *87*, 651–701. [\[CrossRef\]](#)
24. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [\[CrossRef\]](#)
26. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Pribelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [\[CrossRef\]](#)
27. Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **2013**, *29*, 1072–1075. [\[CrossRef\]](#)
28. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [\[CrossRef\]](#)
29. Tang, H.; Bowers, J.E.; Wang, X.; Ming, R.; Alam, M.; Paterson, A.H. Synteny and collinearity in plant genomes. *Science* **2008**, *320*, 486–488. [\[CrossRef\]](#)
30. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **2019**, *20*, 238. [\[CrossRef\]](#)
31. Conway, J.R.; Lex, A.; Gehlenborg, N. UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics* **2017**, *33*, 2938–2940. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [\[CrossRef\]](#)
33. 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. [\[CrossRef\]](#)
34. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermini, L. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, *14*, 587–589. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Chernomor, O.; von Haeseler, A.; Minh, B.Q. Terrace aware data structure for phylogenomic inference from supermatrices. *Syst. Biol.* **2016**, *65*, 997–1008. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.; 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. [\[CrossRef\]](#)
37. Hoang, D.H.; Chernomor, O.; von Haeseler, A.; Minh, B.Q.; Vinh, L.S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **2017**, *35*, 518–522. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [\[CrossRef\]](#)
39. Wu, F.; Cen, Y.; Deng, X.; Zheng, Z.; Chen, J.; Liang, G. The complete mitochondrial genome sequence of *Diaphorina citri* (Hemiptera: Psyllidae). *Mitochondrial DNA B Resour.* **2016**, *1*, 239–240. [\[CrossRef\]](#)
40. Wang, H.-L.; Yang, J.; Boykin, L.M.; Zhao, Q.-Y.; Li, Q.; Wang, X.-W.; Liu, S.-S. The characteristics and expression profiles of the mitochondrial genome for the Mediterranean species of the *Bemisia tabaci* complex. *BMC Genom.* **2013**, *14*, 401. [\[CrossRef\]](#)
41. Chen, S.-C.; Wang, X.-Q.; Li, P.-W.; Hu, X.; Wang, J.-J.; Peng, P. The complete mitochondrial genome of *Aleurocanthus camelliae*: Insights into gene arrangement and genome organization within the family Aleyrodidae. *Int. J. Mol. Sci.* **2016**, *17*, 1843. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Chen, Z.-T.; Mu, L.-X.; Wang, J.-R.; Du, Y.-Z. Complete mitochondrial genome of the citrus spiny whitefly *Aleurocanthus spiniferus* (Quaintance) (Hemiptera: Aleyrodidae): Implications for the phylogeny of whiteflies. *PLoS ONE* **2016**, *11*, e0161385. [\[CrossRef\]](#)

43. Lei, T.; Zhong, Y.-W.; Liu, Y.-Q. Complete mitochondrial genome of the whitefly *Aleyrodes shizuokensis* Kuwana (Hemiptera: Aleyrodidae), new record from Chinese mainland. *Mitochondrial DNA Part B Resour.* **2021**, *6*, 397–398. [[CrossRef](#)]
44. Thao, M.L.; Baumann, P. Evolutionary relationships of primary prokaryotic endosymbionts of whiteflies and their hosts. *Appl. Environ. Microb.* **2004**, *70*, 3401–3406. [[CrossRef](#)] [[PubMed](#)]
45. Lu, M.-X.; Chen, Z.-T.; Yu, W.-W.; Du, Y.-Z. The complete mitochondrial genome of a spiraling whitefly, *Aleurodicus dispersus* Russell (Hemiptera: Aleyrodidae). *Mitochondrial DNA A DNA Mapp. Seq. Anal.* **2017**, *28*, 165–166. [[CrossRef](#)]
46. Jain, C.; Rodriguez, R.L.; Phillippy, A.M.; Konstantinidis, K.T.; Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **2018**, *9*, 5114. [[CrossRef](#)]
47. Kim, D.; Park, S.; Chun, J. Introducing EzAAI: A pipeline for high throughput calculations of prokaryotic average amino acid identity. *J. Microbiol.* **2021**, *59*, 476–480. [[CrossRef](#)]
48. Katz, L.S.; Griswold, T.; Morrison, S.S.; Caravas, J.A.; Zhang, S.; den 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](#)]
49. Wang, D.; Zhang, Y.; Zhang, Z.; Zhu, J.; Yu, J. KaKs_Calculator 2.0: A toolkit incorporating Gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinform.* **2010**, *8*, 77–80. [[CrossRef](#)]
50. Sueoka, N. Intrastrand parity rules of DNA base composition and usage biases of synonymous codons. *J. Mol. Evol.* **1995**, *40*, 318–325. [[CrossRef](#)]
51. Sueoka, N. Directional mutation pressure and neutral molecular evolution. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 2653–2657. [[CrossRef](#)]
52. Wright, F. The ‘effective number of codons’ used in a gene. *Gene* **1990**, *87*, 23–29. [[CrossRef](#)]
53. Konstantinidis, K.T.; Tiedje, J.M. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2567–2572. [[CrossRef](#)] [[PubMed](#)]
54. Thompson, C.C.; Chimento, L.; Edwards, R.A.; Swings, J.; Stackebrandt, E.; Thompson, F.L. Microbial genomic taxonomy. *BMC Genom.* **2013**, *14*, 913. [[CrossRef](#)]
55. Luan, J.; Sun, X.; Fei, Z.; Douglas, A.E. Maternal inheritance of a single somatic animal cell displayed by the bacteriocyte in the whitefly *Bemisia tabaci*. *Curr. Biol.* **2018**, *28*, 459–465. [[CrossRef](#)] [[PubMed](#)]
56. Boyle, L.; O’Neill, S.L.; Robertson, H.M.; Karr, T.L. Interspecific and intraspecific horizontal transfer of *Wolbachia* in *Drosophila*. *Science* **1993**, *260*, 1796–1799. [[CrossRef](#)] [[PubMed](#)]
57. Huigens, M.E.; de Almeida, R.P.; Boons, P.A.; Luck, R.F.; Stouthamer, R. Natural interspecific and intraspecific horizontal transfer of parthenogenesis-inducing *Wolbachia* in *Trichogramma* wasps. *Proc. Biol. Sci.* **2004**, *271*, 509–515. [[CrossRef](#)]
58. Chong, R.A.; Moran, N.A. Intraspecific genetic variation in hosts affects regulation of obligate heritable symbionts. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13114–13119. [[CrossRef](#)]
59. Chong, R.A.; Park, H.; Moran, N.A. Genome evolution of the obligate endosymbiont *Buchnera aphidicola*. *Mol. Biol. Evol.* **2019**, *36*, 1481–1489. [[CrossRef](#)]
60. Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature* **2000**, *407*, 81–86. [[CrossRef](#)]
61. Gil, R.; Silva, F.J.; Zientz, E.; Delmotte, F.; Gonzalez-Candelas, F.; Latorre, A.; Rausell, C.; Kamerbeek, J.; Gadau, J.; Holldobler, B.; et al. The genome sequence of *Blochmannia floridanus*: Comparative analysis of reduced genomes. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9388–9393. [[CrossRef](#)] [[PubMed](#)]
62. Rispe, C.; Delmotte, F.; van Ham, R.C.H.J.; Moya, A. Mutational and selective pressures on codon and amino acid usage in *Buchnera*, endosymbiotic bacteria of aphids. *Genome Res.* **2004**, *14*, 44–53. [[CrossRef](#)] [[PubMed](#)]
63. CNCB-NGDC members and partners. Database resources of the National Genomics Data Center, China National Center for Bioinformation in 2022. *Nucleic Acids Res.* **2022**, *50*, D27–D38. [[CrossRef](#)] [[PubMed](#)]

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.