Frontiers in Bioscience Landmark

Front. Biosci. (Landmark Ed) **2022**; 27(5): 154
https://doi.org/10.31083/j.fbl2705154

Original Research

The Mitogenome of *Aleuroclava Psidii* (Singh, 1931) (Hemiptera: Aleyrodidae) and Increased Number of Mitochondrial Gene Rearrangements in Whiteflies

Nan Song[1], Heng Zhang[1], Run-e Bai[1], Hao-guang Meng[1,*]

[1]College of Plant Protection, Henan Agricultural University, 450002 Zhengzhou, Henan, China
*Correspondence: menghaoguang@henau.edu.cn (Hao-guang Meng)
Academic Editor: Graham Pawelec
Submitted: 23 February 2022 Revised: 22 April 2022 Accepted: 29 April 2022 Published: 12 May 2022

Abstract

Background: In this study, the entire mitochondrial genome (mitogenome) of *Aleuroclava psidii* (Singh, 1931) (Hemiptera: Aleyrodidae) was sequenced. The species *A. psidii* is currently classified in the subfamily Aleyrodinae. This mitogenome is the first representative from the genus *Aleuroclava*. **Methods**: Next-generation sequencing was used to obtain the molecular data. We conducted phylogenetic analyses with 18 existing mitogenomes of whiteflies and three outgroups of psyllids, under the Maximum likelihood and Bayesian inference criteria. **Results**: The arrangement of genes differed between the mitogenome of *A. psidii* and the putative ancestral insect mitogenome, and also differed from the mitogenomes of other whiteflies. Mitochondrial gene rearrangements involved the transpositions of *trnQ*, *trnY*, and the protein-coding gene *nad1*. Most hemipteran mitogenomes have the same mitochondrial gene order as that inferred to be ancestral for insects. However, there are an increased number of gene rearrangements in the mitogenomes of whiteflies. Phylogenetic reconstructions supported Aleurodicinae and Aleyrodinae as being monophyletic. **Conclusions**: Comparison of the gene order of mitogenomes revealed a clade-specific evolutionary trend in whiteflies. This study demonstrates the potential of using structural rearrangements to resolve major phylogenetic relationships within Aleyrodidae.

Keywords: Aleyrodidae; *Aleuroclava psidii*; mitochondrial DNA; rearrangement

1. Introduction

Mitochondrial genome (mitogenome) sequences and structure have been widely used to investigate the phylogenetic relationships of insects [1–3]. The insect mitogenome is a small and generally circular, double-stranded molecule with a genome size ranging from 15,000 bp to 18,000 bp [1]. Typically, the insect mitogenome contains 37 genes encoding 13 protein-coding genes, 22 transfer RNA genes, and two ribosomal RNA genes. A large non-coding region is often present between the small ribosomal RNA gene and the *trnI* gene. This region usually has a higher A+T content and contains elements for transcription and/or replication of the mitogenome [1,4–6], and is thus also called the AT-rich region or control region. The entire mitogenome of insects is compact and the intergenic spacer sequences are usually very short.

The mitochondrial gene order in insects generally remains stable among most groups. The gene arrangement of the fruit fly *Drosophila yakuba* Burla, 1954 is considered to be ancestral for insects [7,8]. The large majority of insect species therefore have an identical arrangement of mitochondrial genes as *D. yakuba* [8]. Nevertheless, discrete lineages have been discovered that exhibit many more mitochondrial gene rearrangements than other groups, such as the parasitic wasps in Hymenoptera [9,10], some thrips in Thysanoptera [7,11], and the lice in Phthiraptera [12,13]. Most groups within Hemiptera have a mitogenome organization that is identical to the proposed ancestor, with just several exceptional gene rearrangements reported in the suborders Auchenorrhyncha [14,15] and Heteroptera [16–19]. In contrast to other hemipterans, however, all of the published mitogenomes of whiteflies show an extraordinary number of rearrangements. Moreover, the gene arrangements vary between different groups of whitefly species. Despite the fact that some aspects of whitefly mitogenomes are unusual, there is still no description of the characteristics of these genomes in a phylogenetic framework.

Whiteflies are a group of insects classified in the family Aleyrodidae of the suborder Sternorrhyncha within Hemiptera. The family Aleyrodidae includes many serious agricultural pests, such as *Bemisia tabaci* (Gennadius, 1889). This is an economically important invasive species that causes considerable damage to agricultural crops. In addition to directly feeding on plants, some whitefly species can also transmit plant viruses, with one study suggesting that 114 virus species are transmitted by whiteflies [20].

Mitochondria are key eukaryotic organelles involved in metabolism, apoptosis, disease and aging in animals [21–25]. While mitochondrial oxidative phosphorylation is the major source of ATP for the cell [21], each of these processes may be associated with the biological characteristics of some insects. Sequencing and characterization of complete mitogenomes provides a key to address the evolution

of mitochondria. To date (January, 2022), the mitogenome sequences of only 18 whitefly species have been reported in GenBank. The limited data has so far limited efforts to reconstruct a comprehensive phylogeny of this insect group and thus better understand the evolution of mitogenomes.

In the present study, we provide an additional mitogenome sequence from the family Aleyrodidae, namely *Aleuroclava psidii* (Singh, 1931). This is the first complete mitogenome sequence of the genus *Aleuroclava*. The two main objectives of this study were: (1) to investigate the phylogenetic relationships within Aleyrodidae using mitogenome sequences, and (2) to assess the phylogenetic utility of mitochondrial gene rearrangements for this group.

2. Materials and Methods

2.1 Specimen and DNA Extraction

Specimens of *A. psidii* (about 50 adults) were collected from photinia (*Photinia* × *fraseri* Dress) in an open field near Zhoukou City, Henan province, China (geospatial coordinates: 33.778°N, 114.859°E). All specimens were obtained from the same natural population. Morphological identification of the specimens was undertaken with reference to Yan and Bai [26] and Guastella *et al*. [27]. The specimens were preserved in absolute ethanol and stored at –80 °C.

Genomic DNA was extracted from a pool of 40 specimens using the TIANamp Micro DNA Kit (Tiangen Biotech Co., ltd., Beijing, China). The DNA quality and concentration were examined with a nucleic acid protein analyzer (Quawell Technology Inc., San Jose, CA, USA).

2.2 Next Generation Sequencing

Genomic DNA was sonicated to 300 bp using a Covaris S220 focused-ultrasonicator (Covaris Inc., Woburn, MA, USA), according to the Illumina protocol. Library generation for Illumina Hiseq sequencing was carried out using the Illumina TruSeqTM DNA Sample Prep Kit (Illumina, San Diego, CA, USA). The resulting fragment library was sequenced on an Illumina HiSeq2500 platform, with a strategy of 150 paired-end sequencing. Data quality control was performed using the NGS QC Toolkit v2.3.3 (New Delhi, India) [28], under the default settings. Adapters, ploy-N, and low-quality reads were removed from raw data. The high-quality reads (avg. Q20 >90%, and avg. Q30 >80%) were used to assemble the mitochondrial scaffold.

2.3 Mitogenome Assembly, Annotation and Analysis

Two assemblers were used for mitogenome assembly, namely GetOrganelle v1.7.5.2 (Kunming, China) [29] and Geneious R11 (Auckland, New Zealand) [30]. For assembly with GetOrganelle [29], we used the GetOrganelle animal database (-F animal_mt) to identify, filter, and assemble target-associated reads. For assembly using Geneious [30], we applied the "map to reference strategy", with the pre-sequenced mitochondrial *cox1* gene sequence as the reference. The parameter settings were selected as described by Yang *et al*. [31].

The newly sequenced mitogenome of *A. psidii* as well as mitogenome sequences obtained from GenBank were annotated in the MITOS webserver [32] (http://mitos2.bioinf.uni-leipzig.de/index.py, Leipzig, Germany) to obtain uniform annotations. The RefSeq 63 Metazoa and the Invertebrate genetic code were selected as job settings. Mitogenome structure images for all whitefly species were generated using OGDRAW (Potsdam-Golm, Germany) [33]. Gene boundaries for protein-coding genes were refined by alignment with closely related whiteflies (Table 1). Sequences of protein-coding genes were separately aligned using TranslatorX (London, UK) [34]. A detailed description of parameter settings for TranslatorX [34] is available in the section on Sequence alignment. The secondary structures of 22 tRNA genes were predicted by MITOS [32]. The gene boundaries of two rRNA genes were refined by alignment against published sequences. The rRNA sequences were aligned separately using MAFFT version 7 (Osaka, Japan) [35]. The secondary structures of rRNA genes were inferred by reference to *D. yakuba* [36]. The annotated mitogenome sequence was submitted to GenBank under the accession number OM362911.

The nucleotide composition of mitogenome sequences was computed with MEGA X (Hachioji, Japan) [37]. AT and GC-skew values were measured using the following formula: AT-skew = (A–T)/(A+T) and GC-skew = (G–C)/(G+C) [38]. Non-synonymous (*dN*) and synonymous (*dS*) substitution rates for protein-coding genes were estimated using PAML 4.9 (London, UK) [39].

Comparison of the gene order in whitefly mitogenomes was performed in CREx (Leipzig, Germany) [40] (http://pacosy.informatik.uni-leipzig.de/crex/form), with the common intervals parameter for distance measurement. In addition, mitogenome structures were mapped onto the sequence-based phylogeny to explore the potential phylogenetic information from mitochondrial gene rearrangements.

2.4 Sequence Alignment

Each of the 37 mitochondrial genes was aligned separately. Protein-coding genes were aligned using codon-based multiple alignments under the MAFFT algorithm [35] implemented in TranslatorX [34], with the Invertebrate Mitochondrial genetic code. Poorly aligned sites from the protein alignment before back-translate to nucleotides were removed with Gblocks v. 0.91 (Barcelona, Spain) [41]. The options used for less stringent selection were: allowing smaller final blocks, allowing gap positions within the final blocks, and allowing less strict flanking positions. Ribosomal and transfer RNA genes were aligned individually in the MAFFT server [35] and with the "E-INS-i" iterative refinement method. Alignments were concatenated with FASconCAT-G_v1.04 (Bonn, Germany) [42]. The partition

Table 1. Taxa included in this study.

Superfamily	Family	Subfamily	Species	GenBank accession
Aleyrodoidea	Aleyrodidae	Aleurodicinae	*Aleurodicus dugesii*	AY521251
Aleyrodoidea	Aleyrodidae	Aleurodicinae	*Aleurodicus dispersus*	KR063274
Aleyrodoidea	Aleyrodidae	Aleyrodinae	***Aleuroclava psidii***	OM362911
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Aleurocanthus spiniferus*	KJ437166
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Aleurocanthus camelliae*	KU761949
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Aleurochiton aceris*	AY572538
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*aff. Aleurochiton* sp.	MH999477
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Aleyrodes shizuokensis*	MT880225
Aleyrodoidea	Aleyrodidae	N/A	Aleyrodidae sp.	MK645228
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Bemisia tabaci*	AY521259
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Bemisia afer*	KF734668
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Bemisia emiliae*	KX714967
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Bemisia* sp.	KX714968
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Neomaskellia andropogonis*	AY572539
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Pealius mori*	LR877884
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Pealius machili*	MT015588
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Singhiella simplex*	LR877885
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Tetraleurodes acaciae*	AY521262
Aleyrodoidea	Aleyrodidae	Aleyrodinae	*Trialeurodes vaporariorum*	AY521265
Psylloidea	Liviidae	N/A	*Diaphorina citri*	KU647697
Psylloidea	Psyllidae	Acizziinae	*Acizzia uncatoides*	MG989217
Psylloidea	Psyllidae	Psyllinae	*Arytainilla spartiophila*	MG989220

Note: Bold indicates the species sequenced in this study.

option "-l" was selected to simultaneously generate a gene partition output file for the concatenated supermatrix. This partition file can be used directly or with slight adjustment in the subsequent Maximum likelihood analysis.

2.5 Phylogenetic Analysis

Three types of concatenated datasets were used in the phylogenetic analysis: (1) PCG_nt, nucleotide alignment including 13 protein-coding genes; (2) PCG_aa, amino acid alignment including 13 protein-coding genes; and (3) PCGRNA, nucleotide alignment including 13 protein-coding genes, two rRNA genes and 22 tRNA genes.

Phylogenetic trees were constructed using Maximum likelihood (ML) and Bayesian inference (BI) approaches. ML analyses were performed using IQ-TREE 1.6.10 (Vienna, Austria) [43]. Models for ML analyses (**Supplementary Table 1**) were chosen for each concatenated dataset using ModelFinder (Canberra, Australia) [44] implemented in IQ-TREE 1.6.10 [43]. For the nucleotide dataset of protein-coding genes, data blocks were predefined by gene and codon. For the amino acid dataset of protein-coding genes, data blocks were predefined by gene alone. The 22 tRNA genes were defined as a single partition due to very short gene length of each tRNA gene (60~70 bp), while two rRNA genes were defined as separate partitions. We used the option of allowing partitions to have different speeds (-spp). Nodal support values (BS) were evaluated using 1000 ultrafast bootstrap replicates [45]. BI analyses were conducted using PhyloBayes MPI on XSEDE (1.8c) (Montréal, Québec, Canada) [46], as implemented in the CIPRES Science Gateway [47]. The CAT-GTR model was employed for the nucleotide datasets PCG_nt and PCGRNA, while the CAT-mtZOA model was used for the amino acid dataset PCG_aa. All BI analyses involved two chains, with a total length of 10,000 cycles. Constant sites were removed. Convergence between the two chains was evaluated by examining the difference in frequency for all of their bipartitions (maxdiff <0.1). Trees were collected after the initial burn-in (10%) was discarded. A majority-rule consensus tree was computed using the program bpcomp implemented in PhyloBayes [46]. Branch support was assessed by clade posterior probabilities (PP).

3. Results

3.1 Data Quality and Mitogenome Assembly

A total of 197,079,804 clean read-pairs were obtained from the sequence library. GetOrganelle produced a 15,876 bp contig, while Geneious generated a 14,774 bp contig. For the mapping assembly conducted by Geneious, 28,046 of 197,079,804 reads were assembled, and the "Pairwise % Identity" reached 99.6%. The mean coverage of 14,774 bases was 286.2, with a standard deviation of 63.9. The minimum coverage was 2 and this occurred at the start site of assembly, while the maximum coverage was 448.

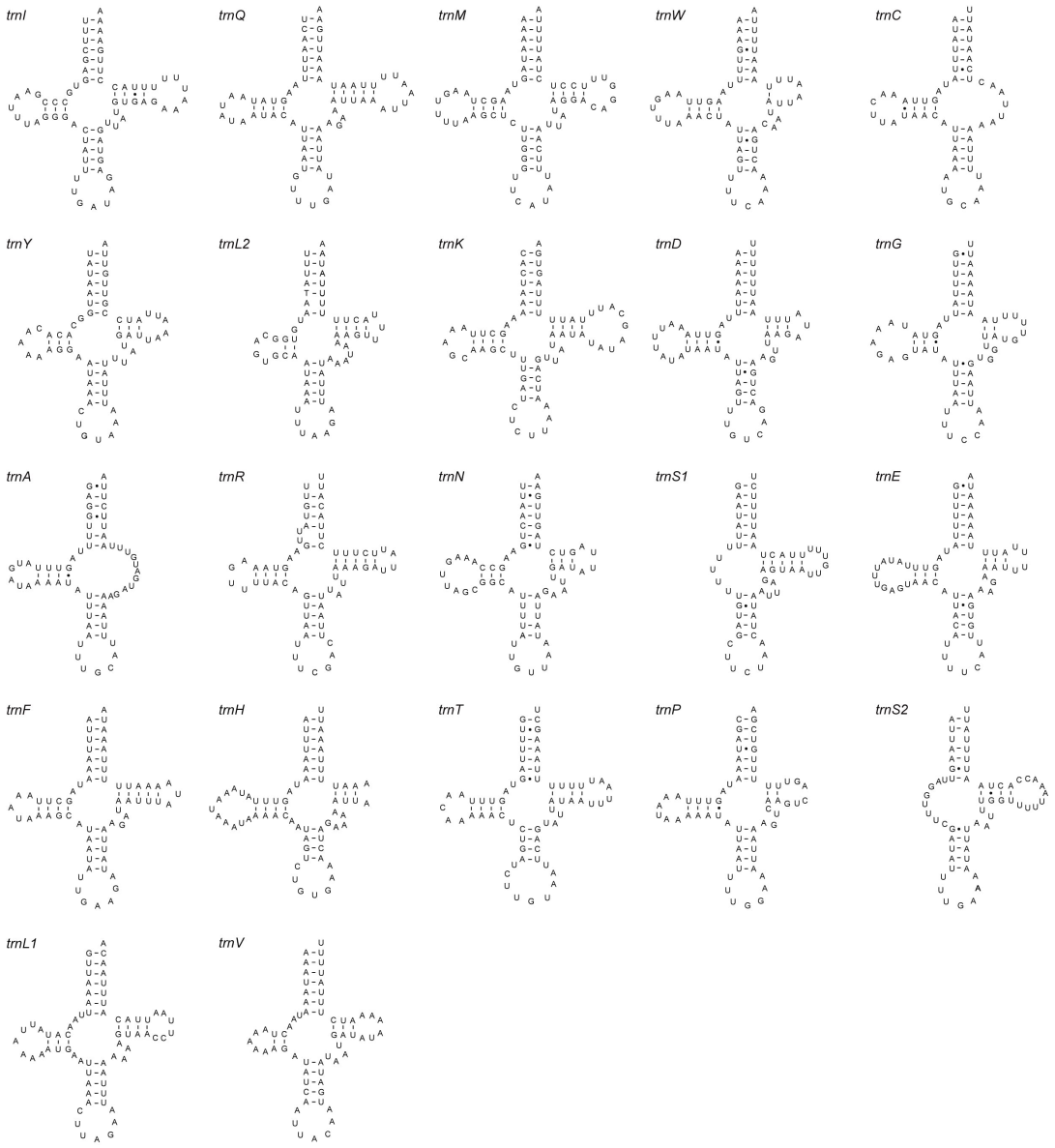

Fig. 1. The secondary structures of tRNA genes inferred for *Aleuroclava psidii.*

3.2 Characteristics of the Aleuroclava psidii Mitogenome

Alignment showed that the Geneious assembly was missing a region compared to that obtained from GetOrganelle. After annotation through MITOS, this missing region was found to span the majority of the *rrnS* gene and the first control region. With the exception of this missing region, Geneious presented an identical sequence to that from GetOrganelle. Therefore, the GetOrganelle assembly was used as the preferred result.

The *A. psidii* mitogenome contained the complete set of 37 genes usually present in insect mitogenomes. In addition, duplication of the tRNA gene cluster -*trnQ_trnI_trnM* was found in the control region. The mitochondrial gene order of *A. psidii* appeared to be unique among whiteflies. Notably, the relative position of *nad1*, which was nested within the control region, differed from that seen in all other whitefly mitogenomes. Moreover, the adjoining *trnI* and -*trnQ*, and the adjoining *trnC* and *trnY* exchanged positions with each other. The entire mitogenome had an AT content of 78.6%. On the major strand, the AT- and GC-skew values were –0.239 (higher T than A content) and 0.243 (higher G than C content), respectively.

A. psidii was unique in that additional start codons were used for protein-coding genes: GTG for *atp6*, GTT for *atp8*, and TTG for *nad2*. The remaining protein-coding genes used ATN (ATG, ATT and ATA) as the initiation codon. Eleven of the 13 protein-coding genes terminated with TAA. The abbreviated stop codon T was found for *cox1* and *nad1*. Estimation of the evolutionary rates of protein-coding genes showed that the mitogenome of *A. psidii* evolved at a low rate (dN = 0.3117) relative to other whitefly species.

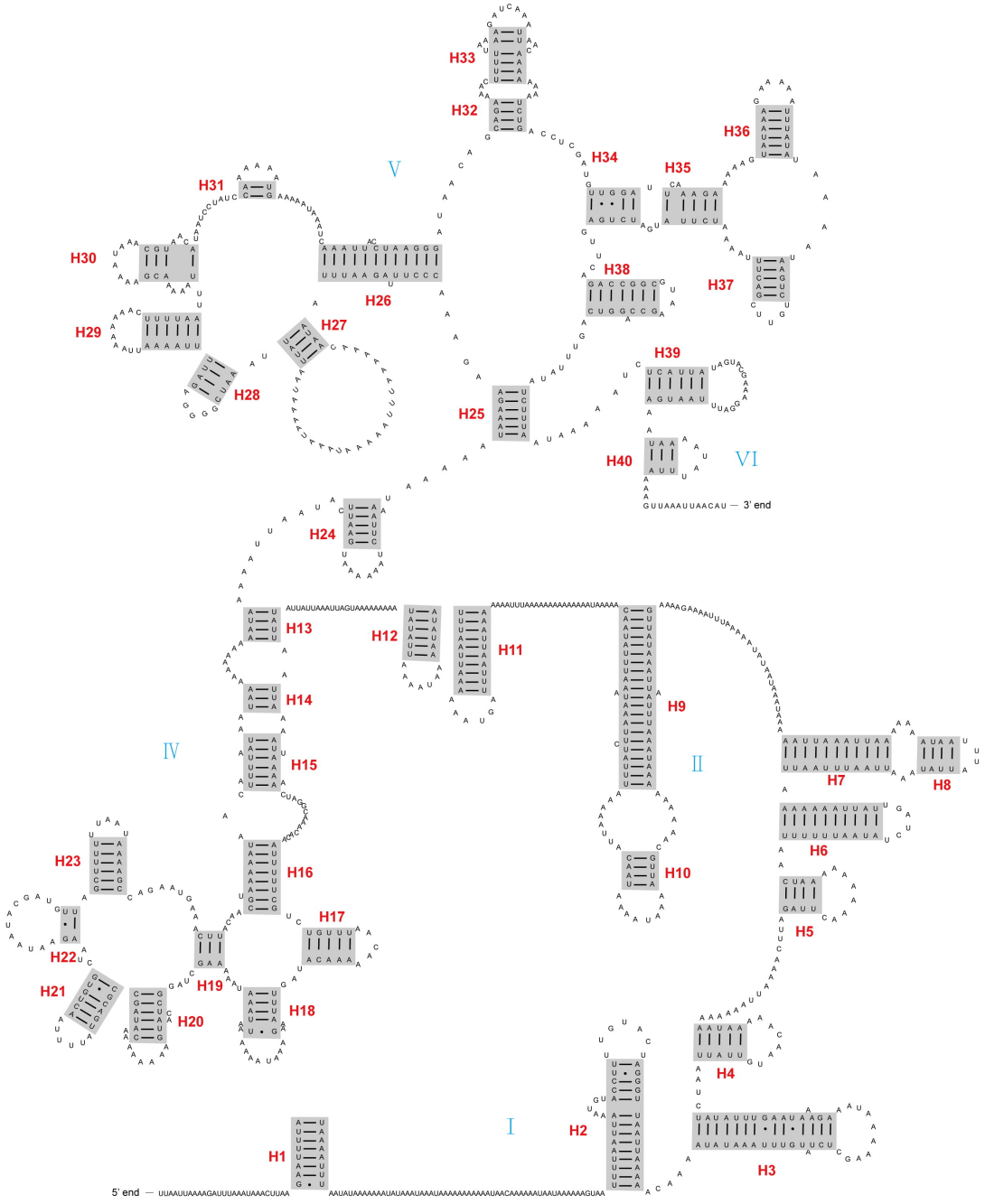

Fig. 2. The secondary structure of *rrnL* inferred for *Aleuroclava psidii*. Blue roman numbers denote the domains, while red numbers denote the helices.

The program MITOS was able to locate 20 tRNA genes. *trnS1* and *trnL1* were identified by alignments with other closely related species. Secondary structures were predicted for all tRNAs (Fig. 1). A standard cloverleaf structure was inferred for 18 tRNA genes. *trnC* and *trnA* had TΨC arm variable loops instead of the typical TΨC arm. *trnS1* and *trnS2* lacked the DHU arm. Unpaired DHU arms of these tRNA genes were also found in other insects [1,14–18].

The *rrnL* and *rrnS* genes were located on the minor strand, with lengths of 1191 bp and 748 bp, respectively. Both these genes had very similar secondary structures to those of *D. yakuba*. The *rrnL* gene of *A. psidii* contained five domains (I–II, and IV–VI) with 40 helices (Fig. 2). The *rrnS* gene had three domains (I, II, and III) composed of 27 helices (Fig. 3).

3.3 Phylogenetic Relationships in Aleyrodidae

The topologies resulting from ML and BI analyses were relatively concordant for the three concatenated datasets. Both Aleurodicinae and Aleyrodinae were monophyletic in all of these topologies with robust support (BS

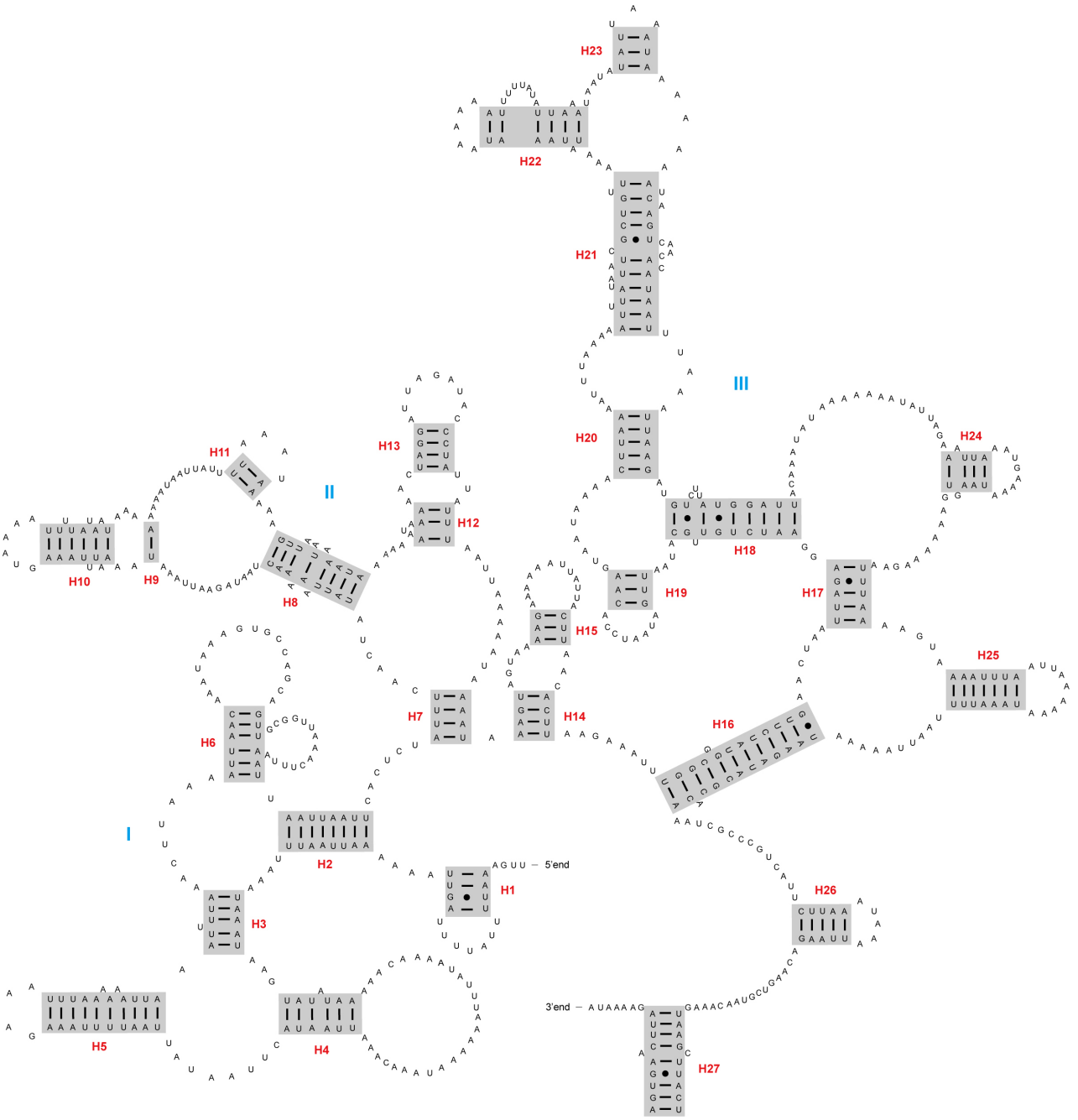

Fig. 3. The secondary structure of *rrnS* inferred for *Aleuroclava psidii*. Blue roman numbers denote the domains, while red numbers denote the helices.

= 100, PP = 1). BI analyses of the datasets PCG_aa and PCG_nt resulted in an identical tree topology (Fig. 4, and **Supplementary Figs. 1–3**).

The main difference between the analyses was the placement of *Trialeurodes vaporariorum* and *Neomaskellia andropogonis*. On all BI trees, *T. vaporariorum* was recovered as the sister to all other Aleyrodinae. *N. andropogonis* was the sister to a clade comprising *Aleurochiton* and *Pealius*. In contrast, the phylogenetic placements of *T. vaporariorum* and *N. andropogonis* varied across ML analyses.

A sister relationship between *Singhiella* and *Bemisia* was recovered in all BI analyses (PP = 0.99 or 1) and in ML analysis of PCG_aa (BS = 100). At the genus level, the monophyly of *Aleurodicus*, *Pealius*, *Aleurocanthus* and *Bemisia* was supported in most analyses. The newly sequenced *A. psidii* was sister to *Aleyrodes shizuokensis* on the trees from PCG_aa and PCG_nt (BS >73, PP = 1). On ML and BI trees from the PCGRNA dataset, *A. psidii* was closely related to *Aleyrodes shizuokensis*.

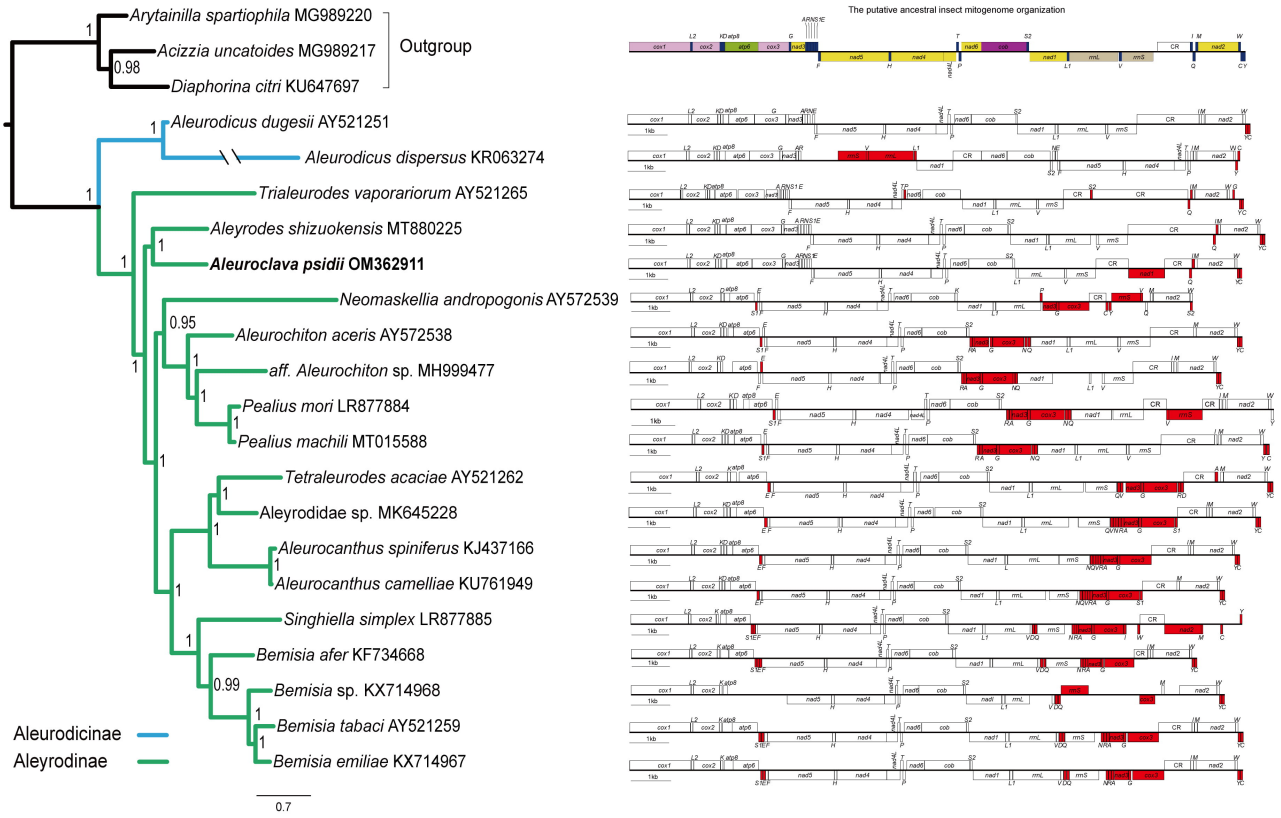

Fig. 4. Phylogenetic tree inferred from the dataset PCG_aa using PhyloBayes under the model CAT-mtZOA (left), and comparisons of the gene order of whitefly mitogenomes (right). In the phylogenetic tree, the numbers around nodes indicate posterior probabilities. The scale bar represents substitutions/site. Bold denotes the newly sequenced species. The branch of *Aleurodicus dispersus* is depicted as half of its original length. Red indicates the mitochondrial gene rearrangements found in the whitefly species examined.

3.4 Gene Rearrangements in Whitefly Mitogenomes

The gene order in the mitogenomes of three psyllid species is identical to that of the ancestral insect mitogenome (Fig. 4). One of these (*Arytainilla spartiophila*) was selected to make pairwise comparisons with the mitogenomes of whiteflies using CREx analyses. All whitefly mitogenomes investigated in this study had experienced gene rearrangements. The newly sequenced *A. psidii* had three rearrangements, involving transpositions of *trnQ*, *trnY*, and the protein-coding gene *nad1*. Notably, the nested position of *nad1* within the control region was very different to that observed in other whiteflies. At least one gene rearrangement was detected in other whitefly mitogenomes. The gene order in *Aleurodicus dugesii* was identical to that of *Arytainilla spartiophila*, except for the altered position of *trnY*. In contrast, the gene order was highly rearranged among representatives of two clades. One clade included *Aleurochiton* and *Pealius*, and the other contained *Singhiella* and *Bemisia.* Based on CREx analyses, the distinct gene order involving altered positions of protein-coding genes in Aleyrodidae sp., *Aleurocanthus spiniferus*, *Aleurocanthus camelliae*, *S. simplex*, *B. afer*, *B. tabaci* and *B. emiliae* may have evolved from the ancestral gene order following a tandem duplication random loss event.

4. Discussion

4.1 Phylogeny Inferred from the Mitogenome Sequence Data

Despite the significant economic importance and worldwide distribution of many well-known pests belonging to the family Aleyrodidae, few studies have investigated the phylogeny of this group. Here, we obtained the complete mitogenome sequence of *A. psidii*, and investigated phylogenetic relationships within Aleyrodidae. These results demonstrate the power of mitogenome sequences for establishing the relationships of whiteflies. The present mitogenome data clearly resolved Aleyrodidae into two monophyletic groups, namely the subfamilies Aleurodicinae and Aleyrodinae. All genera with more than two exemplars included were monophyletic, with the exception of *Aleurochiton*. The close relationship of *Singhiella* to *Bemisia* was confirmed, with both forming a sister group of a clade comprising *Tetraleurodes* and *Aleurocanthus*. This arrangement concurs with an earlier study of mitogenomes [48].

The phylogenetic incongruence observed here was mainly due to the unstable positions of *T. vaporariorum* and *N. andropogonis.* The long branches leading to *T. vaporariorum* and *N. andropogonis* highlight the rapid evolutionary rate observed in both species compared to other white-

flies. Moreover, one of the *Aleurodicus* species, namely *Aleurodicus dispersus* displayed the longest branches on all ML trees. Evolutionary rate analysis indicated the substitution rates for *Aleurodicus dispersus*, *T. vaporariorum* and *N. andropogonis* were higher than those of all other whitefly species (Table 2). In addition to estimating the evolutionary rate from the concatenated 13 protein-coding genes, we also calculated the evolutionary rate of each of the protein-coding genes in mitogenomes (**Supplementary Table 2**). The results showed that *Aleurodicus dispersus*, *T. vaporariorum*, *N. andropogonis* and *S. simplex* had higher non-synonymous substitution rates. In particular, *Aleurodicus dispersus* had the highest *dN* values for the *cob*, *nad1*, *nad2*, *nad4*, *nad4l*, *nad5* and *nad6* genes. Therefore, the different positions of *T. vaporariorum* and *N. andropogonis* were likely to be the result of long-branch attraction, especially in the analyses using the homogeneous model.

Table 2. Estimation of the evolutionary rates of protein-coding genes of whiteflies by PAML.

Species	*dN*	*dS*
Aleuroclava psidii	0.3117	4.8711
Pealius machili	0.3139	5.0625
Bemisia tabaci	0.3154	5.0261
Bemisia sp.	0.3158	5.0574
Bemisia emiliae	0.3183	4.9536
Aleyrodes shizuokensis	0.3191	4.6099
Aleurochiton aceris	0.3193	5.2542
Pealius mori	0.3226	4.8338
Aleurodicus dugesii	0.3279	5.0654
Aleurocanthus camelliae	0.3352	4.7860
Aleurocanthus spiniferus	0.3383	4.8478
Aleyrodidae sp.	0.3442	4.4731
aff Aleurochiton sp.	0.3455	4.9009
Bemisia afer	0.3491	4.3909
Tetraleurodes acaciae	0.3538	4.5784
Singhiella simplex	0.3701	4.2424
Neomaskellia andropogonis	0.3793	5.1452
Trialeurodes vaporariorum	0.3829	3.9358
Aleurodicus dispersus	0.5338	5.1293

Note: *dN*, non-synonymous substitution rates; *dS*, synonymous substitution rates.

To further investigate the effect of long-branch attraction, we excluded the long branched *Aleurodicus dispersus* and reran ML analysis based on the PCGRNA dataset. This resulted in the long-branched *N. andropogonis* being in a different part of the tree and having a relatively close relationship to the clade including *Aleurochiton* and *Pealius* (**Supplementary Fig. 4**). This analysis confirms the potential attraction between *Aleurodicus dispersus* and *N. andropogonis*.

In contrast to ML analyses, the phylogenetic positions of *T. vaporariorum* and *N. andropogonis* were stable in the Bayesian inferences under the heterogeneous model. *T. vaporariorum* was retrieved as the first diverging clade in all BI analyses, while *N. andropogonis* was consistently placed as sister to a clade comprising *Aleurochiton* and *Pealius*. This result demonstrates the power of the heterogeneous model in suppressing long-branch attraction artefacts when using mitogenome sequence data to construct whitefly phylogeny.

4.2 Gene Rearrangements as Phylogenetic Characters for the Phylogeny of Whiteflies

Previous studies have suggested that mitochondrial gene rearrangements can be useful for resolving relationships within some taxonomic groups of insects [12,49–55]. However, mitochondrial gene rearrangements are rare in Hemiptera. Nevertheless, the mitogenomes of single hemipteran groups often exhibit peculiar evolution that deserves careful examination. In this regard, whiteflies are among the most notable Hemiptera taxa, due to several interesting features. Most strikingly, several prior analyses of mitogenome sequences have reported that whitefly mitogenomes often have higher substitution rates than those of other hemipterans [56,57]. In fact, the evolutionary rate of whitefly mitogenome sequences is too high to be used in some phylogenetic analyses [58]. In addition, whiteflies show an unpaired degree of mitochondrial gene rearrangements. Previous studies suggested that high rates of mitochondrial DNA substitution and of mitochondrial gene rearrangement were positively correlated in some insect groups [59,60]. In the present work, we mapped the gene orders of mitogenomes onto the Aleyrodidae phylogeny based on analysis of amino acid sequences (Fig. 4). Comparisons of the gene order of mitogenomes evidenced the clade-specific evolutionary trend in whiteflies.

The mitogenome structures shown in Fig. 4 illustrate the evolutionary divergence patterns of mitochondrial gene rearrangements that frequently occur in the family Aleyrodidae. Originally, the *trnW_-trnY_-trnC* organization was retained in the ancestral mitogenome. However, in the new mitogenome of *A. psidii*, the *-trnY* gene was located between the *trnW* and *-trnC* genes. This pattern was shared by most whiteflies included in the present study. In addition, the tRNA genes located within the tRNA gene clusters and originally situated between the *nad3* and *nad5* genes and between the control region and *nad2* genes can often swap positions in some taxa of Aleyrodidae. In particular, rearrangement of the tRNA gene cluster originally situated between the *nad3* and *nad5* genes can serve as a synapomorphy for several clades within the group. Besides the rearrangements of the tRNA genes mentioned above, the position of *nad1* also changed in the mitogenome of *A. psidii.* This protein-coding gene was originally situated between the *trnS2* and *-trnL1* genes in the ancestral mitogenome. However, in the *A. psidii* mitogenome the *nad1* gene was located between two control regions. This pattern was distinct from all other whitefly mitogenomes. Assembly from

two programs (Geneious and GetOrganelle) resulted in the same position for the *nad1* gene. Nevertheless, further sequencing of mitogenomes from the genus *Aleuroclava* is needed to confirm this result.

Aleurodicus dugesii displayed the closest gene order to the ancestral mitogenome, with only a single transposition of the *-trnY* gene detected in the *Aleurodicus dugesii* mitogenome. The *trnW_-trnY_-trnC* arrangement was generally seen across the entire family Aleyrodidae. Reverse transposition of the *-trnL1_-rrnL_-trnV_-rrnS* gene cluster to the unique gene arrangement *rrnS_trnV_rrnL_trnL1* was found in *Aleurodicus dispersus*.

Similar to *Aleurodicus dugesii*, several early-diverging species within Aleyrodinae also had a small cluster of gene rearrangements, namely *T. vaporariorum*, *Aleyrodes shizuokensis* and *A. psidii*. The *trnQ_trnI_trnM* gene order was shared by *T. vaporariorum*, *Aleyrodes shizuokensis* and *A. psidii*, which may indicate a close relationship for these three species. In comparison, the remaining Aleyrodinae whiteflies had more extensively rearranged mitogenomes. In addition to the frequent rearrangements of tRNA genes, some of the gene arrangements involved protein-coding genes and/or rRNA genes. For *N. andropogonis*, the gene cluster c*ox3-trnG-nad3* originally located between *atp6* and *trnA* and encoded on the major strand was translocated to a position between *trnP* and the control region, as well as being encoded on the minor strand. Furthermore, the *rrnS* gene was reversed from the typical minor strand to the major strand. Relative to *D. yakuba*, we identified an inversion in four species of *Aleurochiton* and *Pealius*, namely *-trnR_-trnA_-nad3_-trnG_-cox3_-trnN_-trnQ* located between *trnS2* and *-nad1*. This could serve as a potential synapomorphy for the clade containing the two genera. The reverse position of *trnS1* shared by *N. andropogonis*, *Aleurochiton* and *Pealius* may indicate a close relationship of *N. andropogonis* to *Aleurochiton* and *Pealius*.

The gene cluster *-nad3_-trnG_-cox3* situated between *-rrnS* and the control region may be used as a character in support of the clade including *T. acaciae*, Aleyrodidae sp. and two species of *Aleurocanthus*. The *-trnQ_-trnV* cluster located between *-rrns* and *-nad3* together with the *-trnR_-trnD* cluster between *cox3* and the control region in *T. acaciae*, as well as the *-trnN_-trnQ_-trnV_-trnR_-trnA* cluster between *-rrnS* and *-nad3* in Aleyrodidae sp. may be autapomorphic for these two species. The pattern of *-trnN_-trnQ_-trnV_-trnR_-trnA* located between *rrnS* and *nad3* may be a synapomorphy for the genus *Aleurocanthus*.

The gene cluster *-trnS1_-trnE_-trnF* lying between *atp6* and *-nad5* and the gene cluster of *-trnV_-trnD_-trnQ_-rrnS_-trnN_-trnR_-trnA_-nad3_-trnG_-cox3* located between *-rrnL* and the control region were shared by *S. simplex* and three species of *Bemisia* (*B. afer*, *B. tabaci* and *B. emiliae*). The transposition of *-trnI* and *-trnW* may be autapomorphic for *S. simplex*.

5. Conclusions

The mitogenome of *A. psidii* is highly rearranged and has many distinguishing characters. It has two main non-coding regions and a genome structure that is different to all other whitefly species studied to date. Our phylogenetic analyses demonstrate the potential to resolve major phylogenetic relationships within Aleyrodidae by using structural rearrangements such as those described above. Despite this, we acknowledge that taxon sampling in this work was still limited by the mitogenome sequences available for the group. Future research should aim to sequence more whitefly species from the genera *Aleurodicus*, *Aleuroclava* and *Neomaskellia* in order to confirm the pattern of gene rearrangements detected in this study.

Abbreviations

mitogenome, mitochondrial genome; mtDNA, Mitochondrial DNA; PCG, protein-coding gene; rRNA, ribosomal RNA; tRNA, transfer RNA; *atp6* and *atp8*, ATP synthase subunits 6, and 8; *cob*, Apocytochrome b gene; *cox1-3*, Cytochrome c oxidase subunits 1–3 genes; *nad1-6* and *nad4L*, NADH dehydrogenase subunits 1–6, and 4 L genes; *rrnL*, large subunit ribosomal RNA; *rrnS*, small subunit ribosomal RNA; nt, nucleotide; aa, amino acid; *dN*, non-synonymous substitutions; *dS*, synonymous substitutions; ML, Maximum likelihood; BI, Bayesian inference; BS, bootstrap; PP, posterior probabilities.

Author Contributions

NS, RB and HM designed the research study. NS and HZ performed the research. RB provided the specimens and performed morphological identification. NS and HZ analyzed the data. NS, HZ and HM wrote the manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript.

Ethics Approval and Consent to Participate

This study was carried out in full compliance with the laws of the People's Republic of China. No permissions were required for insect samples collections. The study species is not included in the 'List of Protected Animals in China'.

Acknowledgment

We thank three anonymous referees for constructive comments that improved this paper.

Funding

This study was supported by funds from the National Natural Science Foundation of China (U1904104), Science and Technology Innovation Fund of Henan Agricultural University (KJCX2019A10).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/j.fbl2705154.

References

[1] Cameron SL. Insect mitochondrial genomics: implications for evolution and phylogeny. Annual Review of Entomology. 2014; 59: 95–117.

[2] Cameron SL, Beckenbach AT, Dowton MP, Whiting MF. Evidence from mitochondrial genomics on interordinal relationships in insects. Arthropod Systematics & Phylogeny. 2006; 64: 27–34.

[3] Gillett CPDT, Crampton-Platt A, Timmermans MJTN, Jordal BH, Emerson BC, Vogler AP. Bulk de novo mitogenome assembly from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionoidea). Molecular Biology and Evolution. 2014; 31: 2223–2237.

[4] Goddard JM, Wolstenholme DR. Origin and direction of replication in mitochondrial DNA molecules from the genus *Drosophila*. Nucleic Acids Research. 1980; 8: 741–757.

[5] Taanman JW. The mitochondrial genome: structure, transcription, translation and replication. Biochimica et Biophysica Acta (BBA)-Bioenergetics. 1999; 1410: 103–123.

[6] Tracy RL, Stern DB. Mitochondrial transcription initiation: promoter structures and RNA polymerases. Current Genetics. 1995; 28: 205–216.

[7] Shao R, Barker SC. The Highly Rearranged Mitochondrial Genome of the Plague Thrips, *Thrips imaginis* (Insecta: Thysanoptera): convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes. Molecular Biology and Evolution. 2003; 20: 362–370.

[8] Clary DO, Wolstenholme DR. The mitochondrial DNA molecule of *Drosophila yakuba*: Nucleotide sequence, gene organization, and genetic code. Journal of Molecular Evolution. 1985; 22: 252–271.

[9] Oliveira DC, Raychoudhury R, Lavrov DV, Werren JH. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). Molecular Biology and Evolution. 2008; 25: 2167–2180.

[10] Tang P, Zhu J, Zheng B, Wei S, Sharkey M, Chen X, *et al*. Mitochondrial phylogenomics of the Hymenoptera. Molecular Phylogenetics and Evolution. 2019; 131: 8–18.

[11] Tyagi K, Chakraborty R, Cameron SL, Sweet AD, Chandra K, Kumar V. Rearrangement and evolution of mitochondrial genomes in Thysanoptera (Insecta). Scientific Reports. 2020; 10: 695.

[12] Covacin C, Shao R, Cameron S, Barker SC. Extraordinary number of gene rearrangements in the mitochondrial genomes of lice (Phthiraptera: Insecta). Insect Molecular Biology. 2006; 15: 63–68.

[13] Cameron SL, Johnson KP, Whiting MF. The mitochondrial genome of the screamer louse *Bothriometopus* (Phthiraptera: Ischnocera): effects of extensive gene rearrangements on the evolution of the genome. Journal of Molecular Evolution. 2007; 65: 589–604.

[14] Song N, Liang A. Complete mitochondrial genome of the small brown planthopper, *Laodelphax striatellus* (Delphacidae: Hemiptera), with a novel gene order. Zoological Science. 2009; 26: 851–860.

[15] Zhang K, Zhu W, Rong X, Zhang Y, Ding X, Liu J, *et al*. The complete mitochondrial genomes of two rice planthoppers, *Nilaparvata lugens* and *Laodelphax striatellus*: conserved genome rearrangement in Delphacidae and discovery of new characteristics of atp8 and tRNA genes. BMC Genomics. 2013; 14: 417.

[16] Chen Z, Liu Y, Wu Y, Song F, Cai W, Li H. Novel tRNA gene rearrangements in the mitochondrial genome of *Camarochiloides weiweii* (Hemiptera: Pachynomidae). International Journal of Biological Macromolecules. 2020; 165: 1738–1744.

[17] Du Y, Dietrich CH, Dai W. Complete mitochondrial genome of *Macrosteles quadrimaculatus* (Matsumura) (Hemiptera: Cicadellidae: Deltocephalinae) with a shared tRNA rearrangement and its phylogenetic implications. International Journal of Biological Macromolecules. 2019; 122: 1027–1034.

[18] Li H, Liu H, Shi A, Štys P, Zhou X, Cai W. The complete mitochondrial genome and novel gene arrangement of the unique-headed bug *Stenopirates* sp. (Hemiptera: Enicocephalidae). PLoS ONE. 2012; 7: e29419.

[19] Liu Y, Li H, Song F, Zhao Y, Wilson J, Cai W. Higher-level phylogeny and evolutionary history of Pentatomomorpha (Hemiptera: Heteroptera) inferred from mitochondrial genome sequences. Systematic Entomology. 2019; 44: 810–819.

[20] Jones DR. Plant viruses transmitted by whiteflies. European Journal of Plant Pathology. 2003; 109: 195–219.

[21] Boore JL. Animal mitochondrial genomes. Nucleic Acids Research. 1999; 27: 1767–1780.

[22] Brand MD. Regulation analysis of energy metabolism. Journal of Experimental Biology. 1997; 200: 193–202.

[23] Graeber MB, Müller U. Recent developments in the molecular genetics of mitochondrial disorders. Journal of the Neurological Sciences. 1998; 153: 251–263.

[24] Wei YH. Oxidative stress and mitochondrial DNA mutations in human aging. Proceedings of the Society for Experimental Biology and Medicine. 1998; 217: 53–63.

[25] Kroemer G, Dallaporta B, Resche-Rigon M. The mitochondrial death/life regulator in apoptosis and necrosis. Annual Review of Physiology. 1998; 60: 619–642.

[26] Yan F, Bai R. Whitefly fauna of China. Henan Science and Technology Press: Zhengzhou, China. 2017.

[27] Guastella D, Tajebe LS, Rapisarda C, Evans G, Fovo FP, Legg JP. First record of *Aleuroclava psidii* (Singh) and *Aleurotrachelus tuberculatus* Singh (Hemiptera: Aleyrodidae) in East Africa. African Entomology. 2014; 22: 437–440.

[28] Patel RK, Jain M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS ONE. 2012; 7: e30619.

[29] Jin J, Yu W, Yang J, Song Y, de Pamphilis CW, Yi T, *et al*. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biology. 2020; 21: 241.

[30] Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, *et al*. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012; 28: 1647–1649.

[31] Yang W, Zhang Y, Feng S, Liu L, Li Z. The first complete mitochondrial genome of the Japanese beetle *Popillia japonica* (Coleoptera: Scarabaeidae) and its phylogenetic implications for the superfamily Scarabaeoidea. International Journal of Biological Macromolecules. 2018; 118: 1406–1413.

[32] Bernt M, Bleidorn C, Braband A, Dambach J, Donath A, Fritzsch G, *et al*. A comprehensive analysis of bilaterian mitochondrial genomes and phylogeny. Molecular Phylogenetics and Evolution. 2013; 69: 352–364.

[33] Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Research. 2019; 47: W59–W64.

[34] Abascal F, Zardoya R, Telford MJ. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Research. 2010; 38: W7–W13.
[35] Katoh K, Standley DM. MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution. 2013; 30: 772–780.
[36] Clary DO, Wolstenholme DR. The ribosomal RNA genes of *Drosophila* mitochondrial DNA. Nucleic Acids Research. 1985; 13: 4029–4045.
[37] Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution. 2018; 35: 1547–1549.
[38] Reyes A, Gissi C, Pesole G, Saccone C. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. Molecular Biology and Evolution. 1998; 15: 957–966.
[39] Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution. 2007; 24: 1586–1591.
[40] Bernt M, Merkle D, Ramsch K, Fritzsch G, Perseke M, Bernhard D, *et al*. CREx: inferring genomic rearrangements based on common intervals. Bioinformatics. 2007; 23: 2957–2958.
[41] Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic Biology. 2007; 56: 564–577.
[42] Kück P, Longo GC. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. Frontiers in Zoology. 2014; 11: 81.
[43] Nguyen L, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution. 2015; 32: 268–274.
[44] Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods. 2017; 14: 587–589.
[45] Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. Molecular Biology and Evolution. 2013; 30: 1188–1195.
[46] Lartillot N, Rodrigue N, Stubbs D, Richer J. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Systematic Biology. 2013; 62: 611–615.
[47] Miller MA, Pfeiffer W, Schwartz T. The CIPRES science gateway. Proceedings of the 2011 TeraGrid Conference on Extreme Digital Discovery - TG '11. 2011.
[48] Chen Z, Mu L, Wang J, Du Y. Complete mitochondrial genome of the citrus spiny whitefly *Aleurocanthus spiniferus* (Quaintance) (Hemiptera: Aleyrodidae): implications for the phylogeny of whiteflies. PLoS ONE. 2016; 11: e0161385.
[49] Huang Y, Liu Y, Zhu X, Xin Z, Zhang H, Zhang D, *et al*. Comparative mitochondrial genome analysis of *Grammodes geometrica* and other noctuid insects reveals conserved mitochondrial genome organization and phylogeny. International Journal of Biological Macromolecules. 2019; 125: 1257–1265.
[50] Li H, Gao J, Liu H, Cai W. Progress in the researches on insect mitochondrial genome and analysis of gene order. Science Foundation in China. 2009; 17: 39–45.
[51] Liu Q, Xin Z, Zhu X, Chai X, Zhao X, Zhou C, *et al*. A transfer RNA gene rearrangement in the lepidopteran mitochondrial genome. Biochemical and Biophysical Research Communications. 2017; 489: 149–154.
[52] Wei S, Shi M, Sharkey MJ, van Achterberg C, Chen X. Comparative mitogenomics of Braconidae (Insecta: Hymenoptera) and the phylogenetic utility of mitochondrial genomes with special reference to Holometabolous insects. BMC Genomics. 2010; 11: 371.
[53] Feng Z, Wu Y, Yang C, Gu X, Wilson JJ, Li H, *et al*. Evolution of tRNA gene rearrangement in the mitochondrial genome of ichneumonoid wasps (Hymenoptera: Ichneumonoidea). International Journal of Biological Macromolecules. 2020; 164: 540–547.
[54] Mao M, Gibson T, Dowton M. Evolutionary dynamics of the mitochondrial genome in the Evaniomorpha (Hymenoptera) - a Group with an intermediate rate of gene rearrangement. Genome Biology and Evolution. 2014; 6: 1862–1874.
[55] Dowton M, Castro LR, Austin AD. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome 'morphology'. Invertebrate Systematics. 2002; 16: 345–356.
[56] Li H, Leavengood JM, Chapman EG, Burkhardt D, Song F, Jiang P, *et al*. Mitochondrial phylogenomics of Hemiptera reveals adaptive innovations driving the diversification of true bugs. Proceedings of the Royal Society B: Biological Sciences. 2017; 284: 20171223.
[57] Song N, Zhang H, Zhao T. Insights into the phylogeny of Hemiptera from increased mitogenomic taxon sampling. Molecular Phylogenetics and Evolution. 2019; 137: 236–249.
[58] Li H, Shao R, Song N, Song F, Jiang P, Li Z, *et al*. Higher-level phylogeny of paraneopteran insects inferred from mitochondrial genome sequences. Scientific Reports. 2015; 5: 8527.
[59] Xu W, Jameson D, Tang B, Higgs PG. The Relationship between the Rate of Molecular Evolution and the Rate of Genome Rearrangement in Animal Mitochondrial Genomes. Journal of Molecular Evolution. 2006; 63: 375–392.
[60] Shao R, Dowton M, Murrell A, Barker SC. Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. Molecular Biology and Evolution. 2003; 20: 1612–1619.