

RESEARCH ARTICLE

All 37 Mitochondrial Genes of Aphid *Aphis craccivora* Obtained from Transcriptome Sequencing: Implications for the Evolution of Aphids

Nan Song¹*, Hao Zhang²*, Hu Li³, Wanzhi Cai³

1 College of Plant Protection, Henan Agricultural University, Zhengzhou, China, **2** Henan Vocational and Technological College of Communication, Zhengzhou, China, **3** Department of Entomology, China Agricultural University, Beijing, China

© These authors contributed equally to this work.

* songnan@henau.edu.cn



CrossMark
click for updates

OPEN ACCESS

Citation: Song N, Zhang H, Li H, Cai W (2016) All 37 Mitochondrial Genes of Aphid *Aphis craccivora* Obtained from Transcriptome Sequencing: Implications for the Evolution of Aphids. PLoS ONE 11(6): e0157857. doi:10.1371/journal.pone.0157857

Editor: Xiao-Yue Hong, Nanjing Agricultural University, CHINA

Received: October 19, 2015

Accepted: June 6, 2016

Published: June 17, 2016

Copyright: © 2016 Song et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This research is supported by grants from the National Natural Science Foundation of China (No. 31402002), and Key scientific research projects of Henan Province, China (No. 14B210036, No. 16A210029).

Competing Interests: The authors have declared that no competing interests exist.

Abstract

The availability of mitochondrial genome data for Aphididae, one of the economically important insect pest families, in public databases is limited. The advent of next generation sequencing technology provides the potential to generate mitochondrial genome data for many species timely and cost-effectively. In this report, we used transcriptome sequencing technology to determine all the 37 mitochondrial genes of the cowpea aphid, *Aphis craccivora*. This method avoids the necessity of finding suitable primers for long PCRs or primer-walking amplicons, and is proved to be effective in obtaining the whole set of mitochondrial gene data for insects with difficulty in sequencing mitochondrial genome by PCR-based strategies. Phylogenetic analyses of aphid mitochondrial genome data show clustering based on tribe level, and strongly support the monophyly of the family Aphididae. Within the monophyletic Aphidini, three samples from *Aphis* grouped together. In another major clade of Aphididae, *Pterocomma pilosum* was recovered as a potential sister-group of *Cavariella salicicola*, as part of Macrosiphini.

Introduction

The rapid expansion of genomic resources and explosion of new genome sequencing technologies are revolutionizing the research of phylogenetics. Researchers are now able to obtain phylogenomic data by fast and cost-effective means. RNA-seq (RNA sequencing), also called whole transcriptome shotgun sequencing, is a technology that uses the capabilities of next generation sequencing to reveal a snapshot of RNA presence, which is often used to generate sequence data for structural and functional studies. These data resource are important for many fields of organelle research [1]. Simultaneously, transcriptome sequencing can provide a relatively large proportion of sampled cDNA clones corresponding to mitochondrial transcripts. Thus, it is possible to mine mitochondrial genes as a by-product of transcriptome sequencing.

The mitochondrion is a fundamental eukaryotic organelle, which has a central role in energy transduction via electron transport coupled with oxidative phosphorylation to produce ATP [2]. Mitochondrial DNA (mtDNA) encodes a limited number of RNAs and proteins essential for formation of a functional mitochondrion [3]. In general, the animal mtDNA is a small (15–20 kb) genome containing 37 genes, which can be divided into three categories: the protein-coding genes (PCGs) encode 13 protein subunits of the enzymes of oxidative phosphorylation, the two rRNAs of the mitochondrial ribosome (*rnl* and *rns*), and the 22 tRNAs necessary for the translation of the proteins encoded by mtDNA [4]. The mtDNA exhibits some properties, like simple genetic structure, maternal inheritance, and very high genome copy numbers, that make it relatively easy obtained and widely used in the study of phylogenetics, molecular evolution, and conservation genetics [4]. Previous studies have shown that mitochondrial genomes (mitogenomes) are reliable markers for phylogeny reconstruction among diverse insect groups [5]. In particular, the mitochondrial cytochrome oxidase subunit 1 (*cox1*) gene has been used as a DNA barcode to identify species, for the purpose of controlling and monitoring insects that pose the greatest threats to crops as diverse as wheat, barley, beans and potatoes. This has further increased the rate of mitochondrial gene sequencing [6].

Compared with conventional approaches [7], transcriptome sequencing can overcome the difficulty of designing a mass of species-specific primers and long-PCR amplifications. Smith (2013) advocated that RNA-seq data are an excellent and untapped resource for investigating many aspects of organelle function and evolution, and he especially emphasized that researchers do not overlook the mitochondrial-derived sequences in the RNA-seq data [1]. Typically, the transcriptome datasets produced by RNA sequencing include all the mitochondrial PCGs and most of the rRNA genes [8]. In comparison, tRNA genes are less covered. This might be because yields of mRNA are much higher than tRNA in RNA-seq, due to the excision of endonucleases. Whereas, it is believed that this situation can be improved by greater sequencing depth [7].

The cowpea aphid, *Aphis craccivora*, is an important insect pest that can cause significant losses in cereal crops and locust trees in China. Some plant pathogenic viruses, such as soybean mosaic potyvirus (SoyMV) and peanut stripe virus (PStV), have been reported to be transmitted by this aphid [9, 10]. These transmitted viruses usually inflict more dramatic damage on crops than direct injury caused by feeding of *A. craccivora*. The *A. craccivora* is classified into the family Aphididae (Hemiptera: Aphidoidea), which comprised of approximate 5,000 known species worldwide. Most aphids have intricate life cycles, of which cyclical parthenogenesis is characteristic of these insects. Many aphid species are monophagous, while others feed on hundreds of plant species across many families. Sometimes plant host alternation is related to the polyphagous aphids. To understand the evolution of insect-plant associations and to further control aphids to reduce the damage of commercial crops, the construction of reliable insect phylogenies is of primary importance. Different systematics and phylogenetics of aphids have been proposed based on morphological characters [11, 12] and molecular sequences [13–16]. Generally, three families have been recognized to constitute the superfamily Aphidoidea: the viviparous Aphididae (that are the so-called true aphids), and the oviparous Adelgidae and Phylloxeridae (both are more properly classified as aphid-like insects, because of their no cauda or cornicles) [13, 15, 17]. In the study by Nováková et al. (2013), the monophyletic Aphidini was recovered by the aphid symbiont *Buchnera* gene sequences [16]. Within the Aphidini, the monophyly of the genus *Aphis* and the sister-group relationship between *Rhopalosiphum padi* and *Schizaphis graminum* were strongly supported [16]. However, the phylogenetic relationships within the whole Aphididae are not resolved with confidence. Compared with previous molecular studies based on partial sequence fragment [13–16], mitogenome include much more phylogenetic information. Yet fewer mitogenomic data are available for aphids. At

present, only seven complete (*Cervaphis quercus*, *Sitobion avenae*, *Aphis gossypii*, *Cavariella salicicola*, *Diuraphis noxia*, *Acyrtosiphon pisum* and *Schizaphis graminum*) and three partial (*Aphis glycines*, *Daktulosphaira vitifoliae* and *Pterocomma pilosum*) mitogenomes in the Aphidoidea have been published in GenBank [18–23]. Although Wang et al. (2013) presented the first comparative analysis of mitogenomes of aphids and a phylogenetic reconstructions based on 11 mitochondrial PCGs, their taxon sample are very sparse (only six aphids included in their analysis) [19]. More recently, Wang et al. (2015) sequenced the complete mitogenome of aphid *Mindarus keteleerifoliae* [24] (however, to date, this data can't be available in GenBank, and thus this species is not included in our study). Their phylogenetic analysis based on 13 PCG nucleotide sequences from eight aphids recovered Aphidini (included only two species) and Macrosiphini (included no *Pterocomma*) in a monophyletic Aphidinae [24]. With conventional PCR and Sanger sequencing method, it is difficult to achieve the aphid mitogenome sequences, owing to the complex secondary structure, higher A + T composition and large tandem repeat regions [19].

In this study, all thirty seven mitochondrial genes of *A. craccivora* were obtained from transcriptome sequencing. In addition, combined with other hemipteran mitogenome sequences, the phylogenetic position of *A. craccivora* within Aphididae was investigated. The present study demonstrates the effectiveness of determining the whole mitochondrial gene data from transcriptome sequencing for those species with difficulty in sequencing mitogenome by conventional PCR methodology.

Materials and Methods

Ethics Statement

No specific permits were required for the insect specimens collected for this study in China. These specimens were collected in the residential garden of the author. The field studies did not involve endangered or protected species. *A. craccivora* is a common aphid species in China and is not included in the “List of Protected Animals in China”.

Sample collection and RNA sequencing

Samples of *A. craccivora* (about 300 heads) were collected from locust trees on May 2015 in Zhengzhou of China (the geospatial coordinates: 34.723°N, 113.635°E). Total RNA was extracted from 35 to 50 winged adult individuals by Trizol reagent (Invitrogen, CA, USA) following the manufacturer's procedure. The total RNA quantity and purity were determined using a Bioanalyzer 2100. Sequencing libraries were constructed using Illumina TruSeq™ RNA Sample Preparation Kit (Illumina, San Diego, CA, USA). The results of RNA quantity and purity checking were the following: the concentration was around 3.15 ug/ul, A260/280 was 2.09, A260/230 was 1.52, total content was 300 ug, 28S/18S was 1.1, and RIN was 8.0. Prior to sequencing, two samples were prepared for repeat checking. The best run was used for further analysis. The RNA transcript was sequenced on an Illumina (Solexa) GAII sequencing machine in Shanghai OE Biotech CO., LTD. The sequencing depth was set to 4 Gb Raw data for each sample.

Transcripts assembly and mitochondrial gene identification

The transcripts were assembled *de novo* using the Trinity method [25], and the results were inputted into BioEdit version 7.0.5.3 [26] to build a local BLAST to search genes using published aphid mitogenomes (mainly based on the *A. glycines* and *A. gossypii*) as bait sequences. [Table 1](#) lists the analyzed aphid species, the taxon status and the GenBank accession numbers

Table 1. Taxonomic information and GenBank accession numbers for the taxa included in this study.

Higher Taxon	Superfamily	Family	Subfamily	Tribe	Species	Accession number	References
Cimicomorpha	Cercopoidea	Aphrophoridae			<i>Philaenus spumarius</i>	AY630340	[27]
	Cercopoidea	Cercopidae			<i>Abidama producta</i>	NC_015799	[28]
Fulgoromorpha	Fulgoroidea	Delphacidae			<i>Laodelphax striatella</i>	JX880068	[29]
	Fulgoroidea	Fulgoridae			<i>Lycorma delicatula</i>	EU909203	[30]
	Fulgoroidea	Flatidae			<i>Geisha distinctissima</i>	FJ230961	[31]
Sternorrhyncha	Psylloidea	Psyllidae			<i>Pachyphylla venusta</i>	NC_006157	[18]
	Aphidoidea	Aphididae	Aphidinae	Aphidini	<i>Schizaphis graminum</i>	AY531391	[18]
	Aphidoidea	Aphididae	Aphidinae	Aphidini	<i>Rhopalosiphum padi</i>	Unpublished data	This study
	Aphidoidea	Aphididae	Aphidinae	Aphidini	<i>Aphis craccivora</i>	KT889380	This study
	Aphidoidea	Aphididae	Aphidinae	Aphidini	<i>Aphis glycines</i>	KC840675	[19]
	Aphidoidea	Aphididae	Aphidinae	Aphidini	<i>Aphis gossypii</i>	NC_024581	[20]
	Aphidoidea	Aphididae	Aphidinae	Macrosiphini	<i>Acyrtosiphon pisum</i>	FJ411411	Moran et al, Unpublished
	Aphidoidea	Aphididae	Aphidinae	Macrosiphini	<i>Diuraphis noxia</i>	NC_022727	[21]
	Aphidoidea	Aphididae	Aphidinae	Macrosiphini	<i>Sitobion avenae</i>	NC_024683	[22]
	Aphidoidea	Aphididae	Aphidinae	Macrosiphini	<i>Cavariella salicicola</i>	NC_022682	[19]
	Aphidoidea	Aphididae	Pterocommatinae	Pterocommatini	<i>Pterocomma pilosum</i>	KC840676	[19]
	Aphidoidea	Greenideidae			<i>Cervaphis quercus</i>	NC_024926	[23]
	Aphidoidea	Phylloxeridae			<i>Daktulosphaira vitifoliae</i>	DQ021446	Baumann and Baumann, Unpublished

doi:10.1371/journal.pone.0157857.t001

[27–31]. The best hit sequences were retrieved from transcript data. Then, the whole retrieved sequences were aligned against all the published aphid mitochondrial sequences to identify the gene boundaries. The clover-leaf secondary structures of tRNA genes predicted by tRNAscan-SE server [32] are presented in S1 Fig. New mitochondrial DNA sequences obtained in this study were deposited in GenBank under accession number of KT889380.

Sequence alignment and characteristics analyses

In total, eighteen mitogenome data included twelve aphids, and comprised the ingroup (10 species from Aphididae, each one from Greenideidae and Phylloxeridae, respectively), and six other homopteran insects, and selected as outgroup (one taxa from Psylloidea, 2 from Cercopoidea, and 3 from Fulgoroidea). Each of the 37 mitochondrial genes were aligned separately for further analyses. For PCGs, firstly stop codons were excluded. Subsequently, each was aligned based on the invertebrate mitochondrial genetic code with Perl script TransAlign [33]. Both the mitochondrial tRNA and rRNA genes were aligned with reference to the conserved secondary structure. Every tRNA gene was aligned manually. Each of the two rRNAs was aligned by the R-Coffee web server [34]. Finally, all alignments were concatenated in a single matrix using FASconCAT_v1.0 [35].

Nucleotide composition of these sequences was calculated using MEGA 6 [36]. Sequence potential saturation was assessed using the index of substitution saturation (*I_{ss}*) of Xia et al.

(2003) [37] implemented in the DAMBE 5 [38]. To detect nucleotide homogeneity across taxa, the chi-square test was performed for the concatenated datasets using PAUP* 4.0b10 [39]. Estimates of nonsynonymous (dN) and synonymous (dS) substitution rates of concatenated protein-coding genes were obtained by Yang and Nielsen (2000) method [40] using the program yn00 as implemented in PAML 4.9 [41]. The software SPSS 16.0 was used to perform one-way ANOVA analyses in order to test for significant differences of substitution rates between aphid lineages. For 12 aphid species, three groups were designed as the priori independent variable. Of which five species from Aphidini made up the group 0, four from Macrosiphini made up the group 1, and the remaining three aphids made up the group 2 (Table 2). The values of dN , dS and dN/dS were set to be the dependent variables for each test, respectively. The post hoc multiple comparisons were conducted using the method of the Least-significant difference (LSD), and the significance level was set to be 0.05.

Phylogenetic analyses

Tree searches were conducted on the combined dataset using both Maximum likelihood (ML) and Bayesian inference (BI). Before undertaking ML analyses, PartitionFinder was employed to infer the optimal partitioning strategy [42], meanwhile the best-fitting model was selected for each partition using the Bayesian Information Criterion (BIC). The data blocks were defined by gene types (each genes of 13 PCGs as independent blocks, while both tRNA and rRNA as two blocks) and by codon positions, in total 41 blocks were utilized. The partition schemes and best-fitting models selected are presented in S1 Table.

ML searches were carried out using the partition schemes and the selected models described above with RAxML as implemented in the CIPRES Portal [43]. Support for nodes was assessed with the fast bootstrap method using 1000 non-parametric bootstrap inferences. The impact of outgroup on the phylogeny were assessed by RAxML analyses based on the combined datasets with reduced taxa (removing partial or entire outgroups). The parameter settings are as those described above.

The BI analyses were conducted using PhyloBayes with a parallel version (pb_mpi1.5a) [44, 45] as implemented on a HP server with twenty-four CPU and 64 G memory. The GTR-CAT model was used for nucleotide analyses. Two chains were run, and started from a random topology. The Maximum “maxdiff” value to be accepted was set as 0.1.

Table 2. The non-synonymous and synonymous nucleotide substitutions calculated for each taxa.

Species	Group	dN	dS	dN/dS
<i>Acyrtosiphon pisum</i>	0	0.0396	2.3374	0.0169
<i>Diuraphis noxia</i>	0	0.0447	2.8293	0.0158
<i>Sitobion avenae</i>	0	0.0405	2.4343	0.0166
<i>Cavariella salicicola</i>	0	0.0472	2.4823	0.019
<i>Aphis craccivora</i>	1	0.0408	2.3948	0.017
<i>Aphis gossypii</i>	1	0.0415	2.3349	0.0178
<i>Aphis glycines</i>	1	0.0421	2.1178	0.0199
<i>Rhopalosiphum padi</i>	1	0.0416	2.4355	0.0171
<i>Schizaphis graminum</i>	1	0.042	2.7876	0.0151
<i>Pterocomma pilosum</i>	2	0.0482	2.0899	0.0231
<i>Cervaphis quercus</i>	2	0.0831	4.6348	0.0179
<i>Daktulosphaira vitifoliae</i>	2	0.1016	4.7822	0.0212

dN , non-synonymous nucleotide substitutions; dS , synonymous nucleotide substitutions.

doi:10.1371/journal.pone.0157857.t002

Results

Mitochondrial gene sequences from transcriptome sequencing data

In total, the sequenced 13,788 nucleotides of *A. craccivora* contained all the 37 mitochondrial genes typically present in insect mitogenomes. Of them, eleven complete PCGs (i.e., *atp6*, *atp8*, *cox1*, *cox2*, *cox3*, *cytb*, *nad1*, *nad3*, *nad4*, *nad4l* and *nad6*) were identified. There were 341 bp nucleotides missed for 3' end of *nad2* on major strand compared to other aphid mitochondrial genomes, and 75 bp for 5' end of *nad5* on minor strand. Compared with closely related species (e.g., *A. glycines* and *A. gossypii*), the 3' end of *rrnS* gene determined from transcriptome sequencing of *A. craccivora* lacked 118 bp sequences on the minor strand. For the *rrnL* gene, there were 24 bp or 23 bp nucleotides missed in 5' end on minor strand compared with *A. glycines* or *A. gossypii*. For the 22tRNA genes, each had a similar gene length to the released aphid mitogenomes.

All the newly determined genes exhibited strong AT nucleotide bias, such that A + T frequencies of PCGs were 82.77%, rRNAs were 84.69% and tRNAs were 84.71%. The results of the substitution saturation tests showed that the value of substitution saturation index for the combined dataset ($Iss = 0.4853$) was significantly lower than the critical values ($Iss.cSym = 0.8494$ or $Iss.cAsym = 0.6575$). This indicated that the combined data suitable for further phylogenetic analysis. The chi-square test of homogeneity of base frequencies across taxa indicated that there was significant heterogeneity among taxa for the combined dataset ($p < 0.05$).

Protein-coding genes

Compared with the aphid *A. glycines* using all assembled contigs of *A. craccivora*, homology between each PCG gene from two species were shown as following: *atp6* 89%, *atp8* 85%, *cox1* 93%, *cox2* 95%, *cox3* 91%, *cytb* 92%, *nad1* 95%, *nad2* 91%, *nad3* 91%, *nad4* 95%, *nad4l* 97%, and *nad6* 89%. Because the partial mitogenome of *A. glycines* contained no *nad5* gene, there was no blast of this one. With regard to another closely related aphid *A. gossypii*, homology between each PCG gene from two species were similar to comparisons between *A. craccivora* and *A. glycines* (i.e., *atp6* 91%, *atp8* 90%, *cox1* 93%, *cox2* 95%, *cox3* 92%, *cytb* 92%, *nad1* 94%, *nad2* 91%, *nad3* 92%, *nad4* 94%, *nad4l* 95%, *nad5* 94%, and *nad6* 92%). Analyses of polymorphic sites between *A. craccivora* and two close aphid isolates were provided in [Table 3](#). The results showed that the polymorphic sites detected between *A. craccivora* and *A. glycines* or *A. gossypii* were usually more than those between *A. glycines* and *A. gossypii*. This might indicate that the relationship between *A. glycines* and *A. gossypii* was closer to *A. craccivora*.

Of the eleven complete PCGs, five (*atp6*, *atp8*, *nad1*, *nad3* and *nad6*) used ATT as start codons, whereas *cox1*, *cox2*, *nad4* and *nad4l* used ATA, and *cox3* and *cytb* used ATG. The partial *nad2* gene had ATT to be start codons. All the PCG genes used TAA as stop codons except for *cox1* and *nad4*. The genes *cox1* and *nad4* ended with incomplete stop codons (T or TA). Because of the missing 3' end of *nad2* and missing 5' end of *nad5*, there were no stop codons or start codons to be found for them.

Transfer RNA genes

The standard 22 tRNA genes were found in the transcriptome sequencing data of *A. craccivora* which ranged from 62 bp (*trnD*, *trnG*, *trnS-AGN*, *trnT* and *trnV*) to 73 bp (*trnK*) in size. Of them, all tRNAs could be folded into typical cloverleaf structure except for *trnS-AGN* ([S1 Fig](#)). *trnS-AGN* lacks the dihydrouridine (DHU) arm, as in many other insect species [[46](#)].

Table 3. Analyses of polymorphic sites among *Aphis craccivora*, *Aphis glycines* and *Aphis gossypii*.

Gene regions	Total length (bp) ^a			Variable sites (bp)			Total number of Synonymous changes (bp)			Total number of Replacement changes (bp)		
	<i>craccivora</i>	<i>glycines</i>	<i>gossypii</i>	<i>craccivora</i>	<i>craccivora</i>	<i>glycines</i>	<i>craccivora</i>	<i>craccivora</i>	<i>glycines</i>	<i>craccivora</i>	<i>craccivora</i>	<i>glycines</i>
				vs <i>glycines</i>	vs <i>gossypii</i>	vs <i>gossypii</i>	vs <i>glycines</i>	vs <i>gossypii</i>	vs <i>gossypii</i>	vs <i>glycines</i>	vs <i>gossypii</i>	vs <i>gossypii</i>
<i>atp6</i>	648	648	648	68	54	42	47	41	31	21	13	11
<i>atp8</i>	153	153	153	22	16	13	11	8	9	11	8	4
<i>cox1</i>	1530	1530	1530	103	102	76	97	98	70	6	4	6
<i>cox2</i>	669	669	669	27	31	22	23	28	19	4	3	3
<i>cox3</i>	783	783	783	66	58	40	51	47	34	15	11	6
<i>cytb</i>	1113	1113	1113	79	87	49	65	75	42	14	12	7
<i>nad1</i>	927	927	927	45	51	38	32	38	30	13	13	8
<i>nad2</i>	975	975	975	57	55	71	34	29	41	23	26	30
<i>nad3</i>	351	351	351	35	30	21	21	17	15	14	13	6
<i>nad4</i>	1308	1308	1308	64	60	41	55	51	32	9	9	9
<i>nad4l</i>	288	288	288	6	12	8	6	11	7	0	1	1
<i>nad5</i>	1650		1650		94			61			33	
<i>nad6</i>	486	486	486	45	51	24	24	37	13	21	14	11

^a Total length indicates sequences without stop codons.

craccivora, *Aphis craccivora*; *glycines*, *Aphis glycines*; *gossypii*, *Aphis gossypii*.

doi:10.1371/journal.pone.0157857.t003

Phylogenetic analyses

The newly obtained full mitochondrial gene data were included in phylogenetic analyses along with other available aphid mitochondrial data. The two trees resulting from the ML and BI analyses had a similar topology (Fig 1). The only discrepancy was in the basal interfamilial relationships within Aphidoidea. ML analysis recovered the Phylloxeridae as the first branch, and the Greenideidae as the next. However, Bayesian analysis retrieved these two families as sister group with low statistical support (posterior probability 0.69). In both ML and BI trees, one notable aspect was the distinctive branch lengths seen between outgroup and ingroup taxa. In particular, outgroups *P. venusta* and *L. striatella* had longer branches in comparison to ingroup aphids. To investigate the potential effect of fast evolving outgroup on the tree topology, we successively removed the long-branched *P. venusta*, *L. striatella*, *Geisha distinctissima* and *Lycorma delicatula* or all six outgroup taxa to rerun tree searches (S2A and S2B Fig). The resulting ingroup relationships are identical to Fig 1.

Moreover, both ML and BI analyses showed strong support for a monophyletic Aphididae (BP = 100, PP < 0.9). However, within the family Aphididae, the subfamily Aphidinae was recovered as a paraphyletic group, with the Pterocommatinae (represented by *P. pilosum*) nested within. In addition, two tribes were strongly supported: Aphidini, and Macrosiphini including *P. pilosum*. Within the Aphidini, both trees strongly supported a monophyletic *Aphis*, which was found to be a sister taxon to the clade of (*R. padi* + *S. graminum*). Thus, two subtribes, Aphidina and Rhopalosiphina, were strongly supported (BP = 100, PP > 0.95). In the genus *Aphis*, the *A. glycines* had a closer relation to *A. gossypii* than to *A. craccivora*. This might be the fact that sequence variation (polymorphic sites) between *A. glycines* and *A. gossypii* were relatively low. For the rest of aphids, the relationships of ((*C. salicola* + *P. pilosum*) + (*D. noxia* + (*A. pisum* + *S. avenae*))) were consistently supported by all analyses.

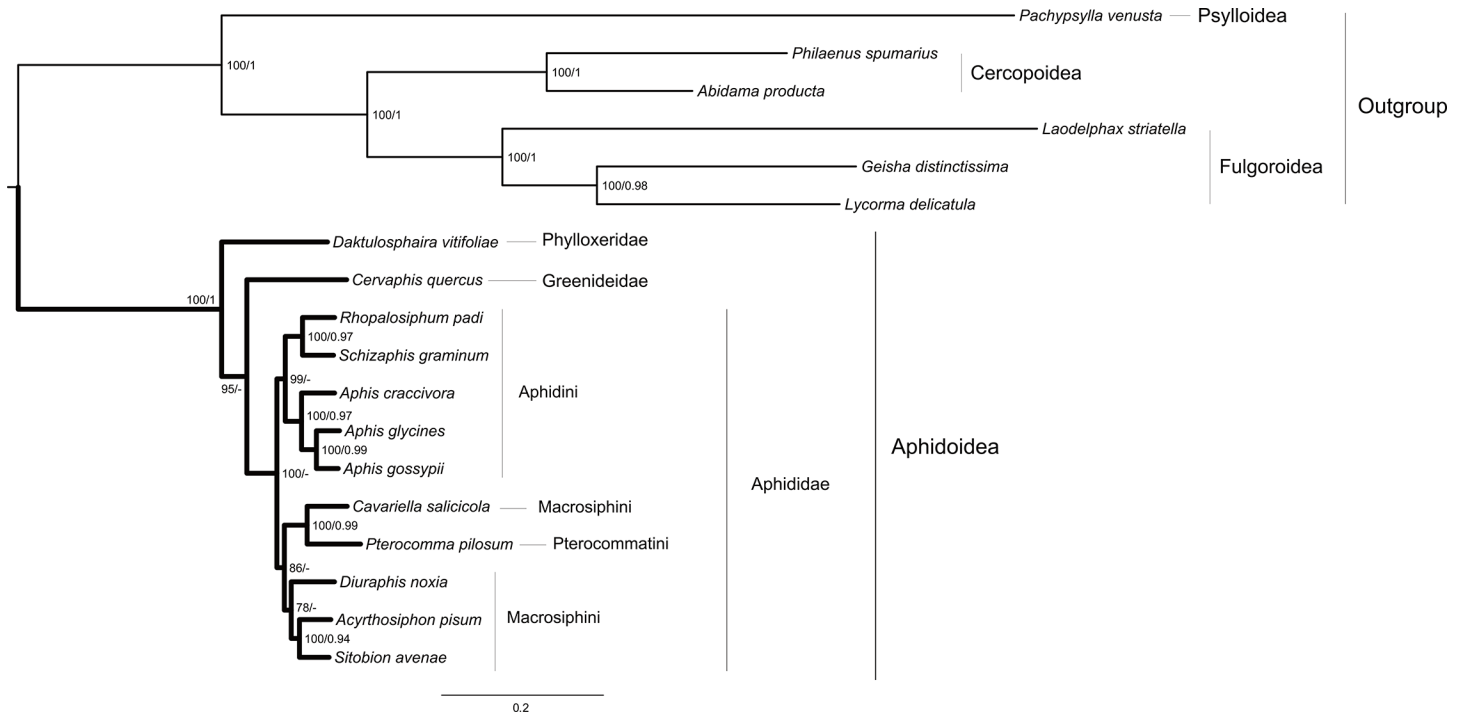


Fig 1. Maximum likelihood tree estimated from the combined dataset. Node numbers show bootstrap support values (above 70, left) and posterior probabilities (above 0.9, right). Scale bar represents substitutions/site.

doi:10.1371/journal.pone.0157857.g001

Substitution rate analyses

The substitution rate analyses showed that the newly sequenced *A. craccivora* had a similar *dN* value to other aphids from the tribe Aphidini (0.0408 ~ 0.0420) (Table 2). For five aphid species from the group of the Macrosiphini (including *P. pilosum*), *A. pisum*, *D. noxia* and *S. avenae* had a relatively lower *dN* values (0.0396 ~ 0.0447), whereas *C. salicicola* and *P. pilosum* had a higher *dN* value (namely, 0.0472 and 0.0482). These results corresponded to the two clades, which was respectively constituted by the former three aphids and the latter two ones. The rest two aphids (i.e., *C. quercus* and *D. vitifoliae*) had obviously higher *dN* values than other ones. Similarly, all aphids had the *dS* values ranging from 2.0899 to 2.8293 except for the *C. quercus* and *D. vitifoliae*, and the two latter had higher *dS* values than 4.6. The statistical analyses revealed no significant differences of *dN* or *dS* values among the defined groups ($P > 0.05$). However, there were significant differences for *dN/dS* values between the group comprised by the species of Macrosiphini and the group including *P. pilosum*, *C. quercus* and *D. vitifoliae* (comparison between group 0 and group 2, $P = 0.030$), and between the group of Aphidini and the group including the *P. pilosum*, *C. quercus* and *D. vitifoliae* (comparison between group 1 and group 2, $P = 0.036$). This result basically confirmed the recovered phylogeny of aphids at the tribe level.

Discussion

Next generation sequencing for complete mitochondrial gene data

Traditionally, complete mitogenome data has been generated by PCR-based strategies, which can be readily confounded by degraded DNA templates, PCR amplicons and sequencing conditions, species-specific primers, and complex genome structure and organization, etc. Thus, it

may be time consuming and costly. In contrast, next generation sequencing can overcome these difficulties and allows determining full mitochondrial gene data more effectively, in particular when it is becoming relatively cheaper already today. This study demonstrates the usefulness of transcriptome sequencing for obtaining complete mitogenome data. However, the obvious drawback of this approach is the inability to find out gene arrangement and to achieve the mitochondrial control region directly from transcript data. In addition, the number of mitochondrial genes detected by transcriptome sequencing varied depending on the different expression level with different sampling [8]. Sampling includes various complex factors, for example, different insect species, developmental stages, sexes and biotypes. A study of the brown planthopper *Nilaparvata lugens* revealed that long wing forms have higher expression levels of genes involved in respiration and energy metabolism compared to short wing forms [47]. They contributed this to the fact that long wing forms required more energy than short wing forms for flight. In this study, we chanced to sample the winged forms of aphid *A. craccivora* for transcriptome sequencing. The mitochondria is an important cell organelle responsible for energy metabolism. Thus, transcripts with abundant mitochondrial genes were obtained, and the entire set of 37 mitochondrial genes could be determined from *A. craccivora* transcript. This is only a speculation, the fact needs to be verified by further experiment. In general, transcriptome sequencing can provide the full set of mitochondrial PCGs and partial RNA genes, which are in fulfillment of the need of systematic research based on mitogenome sequences.

Phylogeny

Mitogenome data are often employed for phylogenetic analyses and are useful to study the intra-order relationships of insects [48–50]. Nevertheless, attempts to reconstruct the higher mitogenomic phylogeny have often failed [51, 52]. The rapid rate of mitogenome evolution limits the resolving power of this type of molecule marker for deep phylogeny reconstruction. In the current study, our phylogenetic analyses suggest that the full mitochondrial gene data may be appropriate for providing fine resolution of evolutionary relationships for the insect family Aphididae.

Outgroup selection is important for phylogenetic inference [53–58]. According to previous studies on Hemiptera phylogeny [30, 59–62], we used a comprehensive taxon sampling from other homopteran lineages which are closely related to aphids. Although some outgroup taxa exhibited obviously long-branch lengths due to higher sequence evolutionary rate, taxa-excluding analyses demonstrated that the resultant tree topology was not affected by long branches (S2A and S2B Fig). Therefore, the current outgroup choice is appropriate.

Two different inference methods under different evolutionary models resulted in congruent tree topology within the superfamily Aphidoidea (Fig 1). The families Phylloxeridae and Greenideidae were successively recovered as the early diverging lineages in Aphidoidea. This arrangement is also supported by the symbiont *Buchnera* DNA sequences [16]. The monophyletic family Aphididae and tribe Aphidini were retrieved. However, the monophyly of the subfamily Aphidinae was not supported due to the nested position of Pterocommatinae. Classifications of Aphidinae have been controversial. Until the latter half of the 20th century, most taxonomists concurred that the Aphidinae included three major groupings: “pterocommatines” (*Pterocomma*, etc.), “aphidines” (*Aphis*, *Rhopalosiphum*, *Schizaphis*, etc.), and “macrosiphines” (*Acyrtosiphon*, *Diuraphis*, *Sitobion*, etc) [14, 63]. However, the relationships among these lineages were not fully resolved. Remaudière and Remaudière (1997) grouped aphidines and macrosiphines together, both of which comprised the subfamily Aphidinae [64]. And they classified pterocommatines as the independent subfamily Pterocommatinae. But the

affinities of Pterocommatinae to Aphidinae were not clear. In contrast, von Dohlen and Moran (2000) strongly supported a monophyletic group of (pterocommatines + aphidines + macrosiphines) [65]. Shaposhnikov et al. (1998) placed Pterocommatini as a sister group to the clade (Aphidini + Macrosiphini) [66]. This hypothesis was confirmed by the study of Ortiz-Rivas and Martínez-Torres (2010), based on the combined analysis of nuclear and mitochondrial sequences [15]. However, von Dohlen et al. (2006) recovered a sister-group of *Pterocomma* plus *Cavariella* [14]. And they thought that there seemed to be no morphological synapomorphies to support the grouping of (Aphidini + Macrosiphini). By contrast, they demonstrated that the distance between stigmal pores on abdominal segments could be considered as a key character to support the relationship of Pterocommatini with Macrosiphini [14]. Our mitogenomic analyses recovered the relationships of ((Pterocommatini + Macrosiphini) + Aphidini). In fact, this result had been obtained using single gene locus EF1 α by Ortiz-Rivas and Martínez-Torres (2010) [15] (see Fig 2 in their paper). Another prior mitogenomic phylogeny by Wang et al. (2013) also recovered the sister relationship of *C. salicicola* to *P. pilosum*, and they advocated that pterocommatines should be transferred into Macrosiphini [19]. This point has also been confirmed by the phylogenetic analyses in the present study.

Besides *C. salicicola*, all the other *macrosiphines* clustered in a clade, and the sister-group of (*A. pisum* + *S. avenae*) were strongly supported (BP = 100, PP = 0.94). Close relationship between these two aphid species has been reported in previous molecular studies [14, 15]. Within the tribe Aphidini, two subtribes were recognized: Aphidina (*Aphis*, *Toxoptera*, etc.) and Rhopalosiphina (*Rhopalosiphum*, *Schizaphis*, etc.) [67]. In our analyses, this two subtribes were consistently recovered with strong nodal support (BP = 100, PP = 0.97). In addition, our data strongly support a sister group relationship between *R. padi* and *S. graminu*, as suggested by *Buchnera* DNA sequences [16]. The newly determined mitochondrial gene data of *A. craccivora* fell within the genus *Aphis*, and clustered with the published mitogenomes of *A. glycines* and *A. gossypii*. Thus, this result validates the approach of obtaining mitogenome sequence from transcriptome sequencing.

Implications for aphid host alternation and pest control

For the lineage Aphidinae, the strong support found for the relationships among its tribes and subtribes, and especially about the position of *Pterocomma*, makes us to be able to discuss some issues on the evolution of life cycles and some strategies of pest control of this group under the framework of reconstructed phylogeny.

The evolution of seasonal host alternation and complex life cycle led insects from Aphidinae to be explosively diversified in Tertiary [68]. In particular, presence of winged males during the autumn migration makes the Aphidinae distinguished from other aphid lineages, and implies a separate origin in this subfamily [69]. These characteristics also enable Aphidinae to be better adapted to the modern, seasonal, north-temperate climate and vegetation [70]. Thus, investigating origins of host alternation of aphids will benefit research on aphid host range and on aphid pest control. In the present study, three aphid tribes were included in the family Aphididae: Pterocommatini, Macrosiphini, and Aphidini. The reconstructed aphid phylogeny from mitogenomic data strongly supported *P. pilosum* (Pterocommatini: *Pterocomma*) embedded within Macrosiphini as the sister to *C. salicicola*. This is congruent with the result from von Dohlen et al. (2006) [14]. Pterocommatini have simple life cycles of non-host-alternating on hosts in the Salicaceae. Whereas, both Macrosiphini and Aphidini comprised rich species with or without host-alternating life cycles on diversified host plants of Rosaceae, Asteraceae, Poaceae, etc. Based on our phylogenetic analysis, two conclusions can be drawn on the origins of host alternation of Aphididae. First, the inferred basal position of *P. pilosum* (Pterocommatini)

in Macrosiphini from mitogenomic data is in agreement with the view that the common ancestor of Aphidinae had a simple, non-host-alternating life cycle on a woody host [14]. Second, the similar phylogenetic relationships of Aphididae found here to that from von Dohlen et al. (2006) [14] further support the idea of the existence of several independent origins of host alternating life cycles in this group of insects.

A solid phylogeny is critical for inferences not only of the evolution of host-plant associations in Aphididae, but also for pest control. For example, based on comparative genomic data, researchers can predict a drug effect on some pests when molecular targets of a compound are known or suspected [71]. The similar phylogenomic pipeline can be used to study the mechanisms behind the action of insecticides on aphid pest. Although obtaining an aphid phylogeny largely concurrent with some previous studies, this study is preliminary due to still limited taxon sampling. Future studies need to include more mitogenomes from this important agricultural insect pest group. Especially, with the explosion of new genome sequencing technologies, researchers should explore large phylogenomic data from the rapid expansion of genomic resources for mitochondrial sequences to reconstruct a reliable phylogenetic relationship of aphids.

Supporting Information

S1 Fig. The secondary structures of mitochondrial tRNAs predicted from *Aphis craccivora*. (TIF)

S2 Fig. Maximum likelihood trees estimated from the combined datasets with reduced taxa. **A)** Removing long-branched outgroups: *Pachypsylla venusta*, *Laodelphax striatella*, *Geisha distinctissima* and *Lycorma delicatula*. **B)** Removing all outgroups. Node numbers show bootstrap support values (above 70). Scale bar represents substitutions/site. For S2A Fig, the internal branches between outgroup and ingroup were halved to more clearly illustrate the tree.
(EPS)

S1 Table. The partition schemes and best-fitting models selected.
(XLSX)

Author Contributions

Conceived and designed the experiments: NS. Performed the experiments: NS. Analyzed the data: NS HZ. Contributed reagents/materials/analysis tools: NS HZ HL WC. Wrote the paper: NS HZ HL WC.

References

1. Smith DR. RNA-Seq data: a goldmine for organelle research. *Brief Funct Genomics*. 2013; 12: 454–456. doi: [10.1093/bfgp/els066](https://doi.org/10.1093/bfgp/els066) PMID: [23334532](https://pubmed.ncbi.nlm.nih.gov/23334532/)
2. Saraste M. Oxidative phosphorylation at the fin de siecle. *Science*. 1999; 283: 1488–1493. PMID: [10066163](https://pubmed.ncbi.nlm.nih.gov/10066163/)
3. Tzagoloff A, Myers AM. Genetics of mitochondrial biogenesis. *Annu Rev Biochem*. 1986; 55: 249–285. PMID: [2427014](https://pubmed.ncbi.nlm.nih.gov/2427014/)
4. Boore JL. Animal mitochondrial genomes. *Nucleic Acids Res*. 1999; 27: 1767–1780. PMID: [10101183](https://pubmed.ncbi.nlm.nih.gov/10101183/)
5. Cameron SL. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annu Rev Entomol*. 2014; 59: 95–117. doi: [10.1146/annurev-ento-011613-162007](https://doi.org/10.1146/annurev-ento-011613-162007) PMID: [24160435](https://pubmed.ncbi.nlm.nih.gov/24160435/)
6. Ratnasingham S, Hebert PDN. BOLD: the barcode of life data system. *Mol Ecol Notes*. 2007; 7: 355–64. PMID: [18784790](https://pubmed.ncbi.nlm.nih.gov/18784790/)

7. Cameron SL. How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. *Syst Entomol.* 2014; 39: 400–411.
8. Nabholz B, Jarvis ED, Ellegren H. Obtaining mtDNA genomes from next-generation transcriptome sequencing: a case study on the basal Passerida (Aves: Passeriformes) phylogeny. *Mol Phylogenet Evol.* 2010; 7: 466–470.
9. Coceano PG, Peressini S, Bianchi GL. The role of winged aphid species in the natural transmission of soybean mosaic potyvirus to soybean in North-east Italy. *Phytopathol Mediterr* 1998; 37: 111–118
10. Chen K, Xu Z, Zhang Z, Fang X, Yan L, Higgins CM, et al. Biological characters and sequence analysis of coat protein gene of Chinese strains of peanut stripe virus (PSTV). *Chinese journal of oil crop sciences/Zhongguo nong ye ke xue yuan you liao zuo wu yan jiu suo zhu ban,* 1998; 21: 55–59, 64.
11. Heie OE. Palaeontology and phylogeny. In: Minks A.K., Harrewijn P. (Eds.), *Aphids. Their Biology, Natural Enemies and Control.* Elsevier, Amsterdam, 1987; 367–391.
12. Sano M, Akimoto SI. Morphological phylogeny of gall-forming aphids of the tribe Eriosomatini (Aphididae: Eriosomatinae). *Syst Entomol.* 2011; 36: 607–627.
13. Ortiz-Rivas B, Moya A, Martínez-Torres D. Molecular systematics of aphids (Homoptera: Aphididae): new insights from the long-wavelength opsin gene. *Mol Phylogenet Evol.* 2004; 30: 24–37. PMID: [15022755](#)
14. von Dohlen CD, Rowe CA, Heie OE. A test of morphological hypotheses for tribal and subtribal relationships of Aphidinae (Insecta: Hemiptera: Aphididae) using DNA sequences. *Mol Phylogenet Evol.* 2006; 38: 316–329. PMID: [16368250](#)
15. Ortiz-Rivas B, Martínez-Torres D. Combination of molecular data support the existence of three main lineages in the phylogeny of aphids (Hemiptera: Aphididae) and the basal position of the subfamily Lachninae. *Mol Phylogenet Evol.* 2010; 55: 305–317. doi: [10.1016/j.ympev.2009.12.005](#) PMID: [20004730](#)
16. Nováková E, Hypša V, Klein J, Footitt RG, Dohlen CDV, Moran NA. Reconstructing the phylogeny of aphids (Hemiptera: Aphididae) using DNA of the obligate symbiont *Buchnera aphidicola*. *Mol Phylogenet Evol.* 2013; 68: 42–54. doi: [10.1016/j.ympev.2013.03.016](#) PMID: [23542003](#)
17. Mc Gavin GC. *Bugs of the World.* Infobase Publishing. ISBN 0-8160-2737-4. 1993.
18. Thao MLL, Baumann L, Baumann P. Organization of the mitochondrial genomes of whiteflies, aphids, and psyllids (Hemiptera, Sternorrhyncha). *BMC Evol Biol.* 2004; 4: 25. PMID: [15291971](#)
19. Wang Y, Huang XL, Qiao GX. Comparative analysis of mitochondrial genomes of five aphid species (Hemiptera: Aphididae) and phylogenetic implications. *PLoS One.* 2013; 8: e77511. doi: [10.1371/journal.pone.0077511](#) PMID: [24147014](#)
20. Zhang S, Luo J, Wang C, Lv L, Li C, Jiang W, et al. Complete mitochondrial genome of *Aphis gossypii* Glover (Hemiptera: Aphididae). *Mitochondr. DNA.* 2014; 1–2.
21. Zhang B, Ma C, Edwards O, Fuller S, Kang L. The mitochondrial genome of the Russian wheat aphid *Diuraphis noxia*: large repetitive sequences between trnE and trnF in aphids. *Gene.* 2014; 533: 253–260. doi: [10.1016/j.gene.2013.09.064](#) PMID: [24095774](#)
22. Zhang B, Zheng J, Liang L, Fuller S, Ma CS. The complete mitochondrial genome of *Sitobion avenae* (Hemiptera: Aphididae). *Mitochondr. DNA* 2016; 27: 945–946.
23. Wang Y, Huang XL, Qiao GX. The complete mitochondrial genome of *Cervaphis quercus* (Insecta: Hemiptera: Aphididae: Greenideinae). *Insect Sci.* 2014; 21: 278–290. doi: [10.1111/1744-7917.12112](#) PMID: [24482299](#)
24. Wang Y, Chen J, Jiang LY, Qiao GX. The complete mitochondrial genome of *Mindarus keteleerifoliae* (Insecta: Hemiptera: Aphididae) and comparison with other Aphididae insects. *Int. J. Mol. Sci.* 2015; 16: 30091–30102. doi: [10.3390/ijms161226219](#) PMID: [26694371](#)
25. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* 2013; 29: 644–652.
26. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 1999; 41: 95–98.
27. Stewart JB, Beckenbach AT. Insect mitochondrial genomics: the complete mitochondrial genome sequence of the meadow spittlebug *Philaenus spumarius* (Hemiptera: Auchenorrhyncha: Cercopoidae). *Genome.* 2005; 48: 46–54. PMID: [15729396](#)
28. Liu J, Bu C, Wipfler B, Liang A. Comparative analysis of the mitochondrial genomes of *Callitettixini Spittlebugs* (Hemiptera: Cercopidae) confirms the overall high evolutionary speed of the AT-rich region but reveals the presence of short conservative elements at the Tribal Level. *PLoS One.* 2014; 9: e109140. doi: [10.1371/journal.pone.0109140](#) PMID: [25285442](#)

29. Zhang KJ, Zhu WC, Rong X, Zhang YK, Ding XL, 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. doi: [10.1186/1471-2164-14-417](https://doi.org/10.1186/1471-2164-14-417) PMID: [23799924](https://pubmed.ncbi.nlm.nih.gov/23799924/)
30. Song N, Liang A, Bu C. A molecular phylogeny of Hemiptera inferred from mitochondrial genome sequences. *PLoS One* 2012; 7: e48778. doi: [10.1371/journal.pone.0048778](https://doi.org/10.1371/journal.pone.0048778) PMID: [23144967](https://pubmed.ncbi.nlm.nih.gov/23144967/)
31. Song N, Liang A. The complete mitochondrial genome sequence of *Geisha distinctissima* (Hemiptera: Flatidae) and comparison with other hemipteran insects. *Acta Bioch Bioph Sin*. 2009; 41: 206–216.
32. Lowe TD, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997; 25: 955–964. PMID: [9023104](https://pubmed.ncbi.nlm.nih.gov/9023104/)
33. Bininda-Emonds OR. transAlign: using amino acids to facilitate the multiple alignment of protein-coding DNA sequences. *BMC Bioinformatics*. 2005; 6: 156. PMID: [15969769](https://pubmed.ncbi.nlm.nih.gov/15969769/)
34. Moretti S, Wilm A, Higgins DG, Xenarios I, Notredame C. R-Coffee: a web server for accurately aligning noncoding RNA sequences. *Nucleic Acids Res*. 2008; 36: W10–13. doi: [10.1093/nar/gkn278](https://doi.org/10.1093/nar/gkn278) PMID: [18483080](https://pubmed.ncbi.nlm.nih.gov/18483080/)
35. Kuck P, Meusemann K. FASconCAT: Convenient handling of data matrices. *Mol Phylogenet Evol*. 2010; 56: 1115–1118. doi: [10.1016/j.ympev.2010.04.024](https://doi.org/10.1016/j.ympev.2010.04.024) PMID: [20416383](https://pubmed.ncbi.nlm.nih.gov/20416383/)
36. Tamura K, Stecher G, Peterson D, Filipiński A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*. 2013; 30: 2725–2729. doi: [10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197) PMID: [24132122](https://pubmed.ncbi.nlm.nih.gov/24132122/)
37. Xia X, Xie Z, Salemi M, Chen L, Wang Y. An index of substitution saturation and its application. *Mol Phylogenet Evol*. 2003; 26: 1–7. PMID: [12470932](https://pubmed.ncbi.nlm.nih.gov/12470932/)
38. Xia XH. DAMBE5: A Comprehensive Software Package for Data Analysis in Molecular Biology and Evolution. *Mol Biol Evol*. 2013; 30: 1720–1728. doi: [10.1093/molbev/mst064](https://doi.org/10.1093/molbev/mst064) PMID: [23564938](https://pubmed.ncbi.nlm.nih.gov/23564938/)
39. Swofford DL. PAUP*. Phylogenetic analysis using Parsimony * and other methods. Version 4 beta 10. Sinauer Associates, Sunderland, Massachusetts. 2003.
40. Yang Z, Nielsen R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol*. 2000; 17: 32–43. PMID: [10666704](https://pubmed.ncbi.nlm.nih.gov/10666704/)
41. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007; 24: 1586–1591. PMID: [17483113](https://pubmed.ncbi.nlm.nih.gov/17483113/)
42. Lanfear R, Calcott B, Ho SYW, Guindon S. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol Biol Evol*. 2012; 29: 1695–701. doi: [10.1093/molbev/mss020](https://doi.org/10.1093/molbev/mss020) PMID: [22319168](https://pubmed.ncbi.nlm.nih.gov/22319168/)
43. Miller M, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees//Gateway Computing Environments Workshop (GCE). 2010; 1–8.
44. Lartillot N, Philippe H. A Bayesian mixture model for across-site heterogeneities in the amino–acid replacement process. *Mol Biol Evol*. 2004; 21: 1095–109. PMID: [15014145](https://pubmed.ncbi.nlm.nih.gov/15014145/)
45. Lartillot N, Lepage T, Blanquart S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*. 2009; 25: 2286–2288. doi: [10.1093/bioinformatics/btp368](https://doi.org/10.1093/bioinformatics/btp368) PMID: [19535536](https://pubmed.ncbi.nlm.nih.gov/19535536/)
46. Wolstenholme DR. Animal mitochondrial DNA: structure and evolution. *Int Rev Cytol*. 1992; 141: 173–216. PMID: [1452431](https://pubmed.ncbi.nlm.nih.gov/1452431/)
47. Xue J, Bao YY, Li B, Cheng YB, Peng ZY, Liu H, et al. Transcriptome analysis of the brown planthopper *Nilaparvata lugens*. *PLoS One*. 2010; 5: e14233. doi: [10.1371/journal.pone.0014233](https://doi.org/10.1371/journal.pone.0014233) PMID: [21151909](https://pubmed.ncbi.nlm.nih.gov/21151909/)
48. Cameron SL, Lambkin CL, Barker SC, Whiting MF. A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high precision. *Syst Entomol*. 2007; 32: 40–59.
49. Fenn JD, Song H, Cameron SL, Whiting MF. A preliminary mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data. *Mol Phylogenet Evol*. 2008; 49: 59–68. doi: [10.1016/j.ympev.2008.07.004](https://doi.org/10.1016/j.ympev.2008.07.004) PMID: [18672078](https://pubmed.ncbi.nlm.nih.gov/18672078/)
50. Ma C, Liu C, Yang P, Kang L. The complete mitochondrial genomes of two band-winged grasshoppers *Gastrimargus marmoratus* and *Oedaleus asiaticus*. *BMC Genomics*. 2009; 10: 156. doi: [10.1186/1471-2164-10-156](https://doi.org/10.1186/1471-2164-10-156) PMID: [19361334](https://pubmed.ncbi.nlm.nih.gov/19361334/)
51. Bernt M, Bleidorn C, Braband A, Dambach J, Donath A, Fritzsche G, et al. A comprehensive analysis of bilaterian mitochondrial genomes and phylogeny. *Mol Phylogenet Evol*. 2013; 69: 352–364. doi: [10.1016/j.ympev.2013.05.002](https://doi.org/10.1016/j.ympev.2013.05.002) PMID: [23684911](https://pubmed.ncbi.nlm.nih.gov/23684911/)
52. Simon S, Hadrys H. A comparative analysis of complete mitochondrial genomes among Hexapoda. *Mol Phylogenet Evol*. 2013; 69: 393–403. doi: [10.1016/j.ympev.2013.03.033](https://doi.org/10.1016/j.ympev.2013.03.033) PMID: [23598069](https://pubmed.ncbi.nlm.nih.gov/23598069/)
53. Wheeler WC. Nucleic acid sequence phylogeny and random outgroups. *Cladistics*. 1990; 6: 363–367.

54. Nixon KC, Carpenter JM. On outgroups. *Cladistics*. 1993; 9: 413–426.
55. Smith AB. Rooting molecular trees—problems and strategies. *Biol J Linn Soc*. 1994; 51, 279–292.
56. Stackebrandt E, Ludwig W. The importance of using outgroup reference organisms in phylogenetic studies—the *Atopobium* case. *Syst Appl Microbiol*. 1994; 17, 39–43.
57. Holland BR, Penny D, Hendy MD. Hendy. Outgroup misplacement and phylogenetic inaccuracy under a molecular clock: a simulation study. *Syst Biol*. 2003; 52: 229. PMID: [12746148](#)
58. Ware JL, Litman J, Klass KD, Spearman LA. Relationships among the major lineages of Dictyoptera: the effect of outgroup selection on dictyopteran tree topology. *Syst Entomol*. 2008; 33: 429–450.
59. von Dohlen CD, Moran NA. Molecular phylogeny of the Homoptera: a paraphyletic taxon. *J Mol Evol*. 1995; 41: 211–223. PMID: [7666451](#)
60. Campbell BC, Steffen-Campbell JD, Sorensen JT, Gill RJ. Paraphyly of Homoptera and Auchenorrhyncha inferred from 18S rDNA nucleotide sequences. *Syst Entomol*. 1995; 20: 175–194.
61. Forero D. The systematics of the Hemiptera. *Revista Colombiana de Entomología*. 2008; 34: 1–21.
62. Cryan JR, Urban JM. Higher-level phylogeny of the insect order Hemiptera: is Auchenorrhyncha really paraphyletic?. *Syst Entomol*. 2012; 37: 7–21.
63. Blackman RL, Eastop VF. *Aphids on the world's crops: An identification and information guide*, second ed. John Wiley & Sons Ltd., Chichester. 2000.
64. Remaudière G, Remaudière M. *Catalogue des Aphididae du Monde. Homoptera Aphidoidea*. INRA, Paris. 1997.
65. von Dohlen CD, Moran NA. Molecular data support a rapid radiation of aphids in the Cretaceous and multiple origins of host alternation. *Biol J Linn Soc*. 2000; 71: 689–717.
66. Shaposhnikov GKH, Kuznetsova V, Stekolshchikov A. Evolutionary tendencies and system of Aphididae. In: Nieto J.M., Dixon A.F.G. (Eds.), *Proceedings of the Aphids in Natural and Managed Ecosystems*, Universidad de Leon, Leon. Secretariado de Publicaciones, London. 1998; 481–487.
67. Ilharco FA, van Harten A. Systematics. In: Minks A.K., Harrewijn P. (Eds.), *Aphids. Their Biology, Natural Enemies and Control*, vol. 2A. Elsevier, Amsterdam. 1987; 51–77.
68. Heie OE. Recent advances in palaeoaphidology. *Acta Phytopathol Entomol Hung*. 1990; 25: 253–260.
69. Moran NA. The evolution of host–plant alternation in aphids: evidence for specialization as a dead end. *Am Nat*. 1988; 132: 681–706.
70. Dixon AFG. *Aphid Ecology. An Optimization Approach*. Chapman & Hall, London. 1998.
71. Rufener L, Keiser J, Kaminsky R, Mäser P, Nilsson D. Phylogenomics of ligand-gated ion channels predicts monepantel effect. *PLoS Pathog*. 2010; 6: e1001091. doi: [10.1371/journal.ppat.1001091](#) PMID: [20838602](#)