Zoological Systematics, 41(2): 123–132 (April 2016), DOI: 10.11865/zs.201610

ORIGINAL ARTICLE

General methods to obtain and analyze the complete mitochondrial genome of aphid species: *Eriosoma lanigerum* (Hemiptera: Aphididae) as an example

Yuan Wang[1, 2], Liyun Jiang[2], Yongjie Liu[1], Jing Chen[2*], Gexia Qiao[2*]

[1]*Shaanxi International Travel Healthcare Center, Shaanxi Entry-Exit Inspection and Quarantine Bureau, Xi'an 710068, China*
[2]*Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China*
[*]*Corresponding authors, E-mail: qiaogx@ioz.ac.cn; chenjing@ioz.ac.cn*

Abstract Insect mitochondrial (mt) genomes are of great interest in researching on molecular evolution, phylogenetics and population genetics. Aphidoidea have about more than 5 000 known species including some agricultural, forestry and horticultural pests. However, only nine complete mt genomes have been previously released in this insect group since the first one, *Schizaphis graminum* (Hemiptera: Aphididae: Aphidinae) in 2004. Herein, we present the complete mitochondrial genome of *Eriosoma lanigerum*, the first species from the subfamily Eriosomatinae. This study takes this species as an example, combing the research during past 12 years, and concludes the general methods to obtain, analyze and annotate the complete mitochondrial genome of aphid species.

Key words Mitochondrial genome, *Eriosoma lanigerum*, methods, aphids.

1 Introduction

The mitochondrial (mt) genome has become the most widely used genomic resource for systematic and comparative genomics entomology for the past 30 years from the first insect mt genome released (Clary & Wolstenholme, 1985). Though other sources of "-omics" data are expanding rapidly, such as chromosome (nucleus) sequences, mt genomes are still very cheaper in spending and are far lower demanding of high quality DNA templates relatively. Therefore, more and more insect groups employed the whole mt genomes to do related studies instead of partial sequences (Dowton *et al.*, 2009; Ma *et al.*, 2012; Nelson *et al.*, 2012).

Aphidinea (Aphidoidea) belongs to Hemiptera, containing three families: Aphididae, Adelgidae and Phylloxeridae. This insect group includes more than 5 000 species worldwide. Aphididae, including 25 subfamilies, is the highest species diversity (Remaudière & Remaudière, 1997; Blackman & Eastop, 2000; Favret, 2016). The research history about complete mt genomes of aphids species, only 12 years, started from *Schizaphis graminum* (Hemiptera: Aphididae: Aphidinae) in 2004 (Thao *et al.*, 2004). Table 1 summarized all the complete mt genomes of aphids from the first report to the present. All of them are from the family Aphididae, no one complete mt genome was reported from the other two families: Phylloxeridae and Adelgidae, though a nearly complete mt genome (without *nad5* gene) from *Viteusvi tifoliae* (Phylloxeridae) was submitted directly to GenBank (DQ021446). Aphidinae has the highest species richness in Aphididae, and most of the complete mt genomes are from this subfamily. Mt genomes of other subfamilies are rare, or even still no one. Thus, there are a lot of mt genome sequences of aphids need to be obtained.

Woolly apple aphid, *Eriosoma lanigerum* (Hausmann), can be a major economic problem to apple growers in most

urn:lsid:zoobank.org:pub:474E3C50-E14E-4157-A130-0CDC83883C36
Received 5 March 2016, accepted 12 April 2016
Executive editor: Fuqiang Chen

parts of the world: they leave on roots, trunk or branches, often causing deformation and cancer-like swellings of bark (Blackman & Eastop, 2000). Herein, we present the complete mitochondrial genome of *E. lanigerum*, the first species from the aphid subfamily Eriosomatinae. This study takes this species as an example, combing the research during the past 12 years, puts forward and concludes the general methods to obtain and analyze the complete mt genome of aphid species.

Table 1. The complete mt genomes of aphid species.

Category	Species	Length (bp)	GenBank No.	Reference
Aphidinae	*Schizaphis graminum*	15 721	NC_006158	Thao *et al.*, 2004
	Acyrthosiphon pisum	16 971	NC_011594	IAGC, 2010
	Diuraphis noxia	15 784	NC_022727	Zhang *et al.*, 2014
	Sitobion avenae	15 180	NC_024683	Zhang *et al.*, 2016b
	Cavariella salicicola	16 317	NC_022682	Wang *et al.*, 2013
	Aphis gossypii	15 869	NC_024581	Zhang *et al.*, 2016a
Greenideinae	*Cervaphis quercus*	15 272	NC_024926	Wang *et al.*, 2014
Mindarinae	*Mindarus keteleerifoliae*	15 199	KP722576	Wang *et al.*, 2015a
Hormaphidinae	*Hormaphis betulae*	15 088	KT875793	Li *et al.*, 2015
Eriosomatinae	*Eriosoma lanigerum*	15 640	KP722582	This study

2 Materials and methods

2.1 Experimental sample

Commonly, individuals of aphids were collected in the host of the field. Here the specimens of *E. Lanigerum* are taken from the apple trees in Linzhi, Tibet of China in 2004. The specimens for molecular research were preserved in 95% ethanol and stored at −20°C before the DNA extraction. Specimen examination should be conducted under microscopes (such as: Leica DM2500, Wetzlar, Germany) based on external morphology following the keys in authoritative monographs, and with reference to the original morphological descriptions and authoritatively identified specimens. In this study, all samples and voucher specimens of *E. Lanigerum* were deposited in the National Zoological Museum of China at the Institute of Zoology, Chinese Academy of Sciences, Beijing, China (NZMC).

2.2 DNA extraction, amplification and sequencing

We suggested that total DNA was extracted from single aphid preserved in 95% ethanol using the DNeasy Tissue Kit (QIAGEN, Hilden, Germany) following the protocols of manufacturer. Most times, we used only one individual for DNA extraction; however, 3–5 individuals from the same colony were used if one individual did not offer enough DNA.

Short and long PCR reactions were used to amplify the whole mt genome of aphid species (Fig. 1). Short PCRs were amplified using universal primers (some modification), and long PCRs were amplified using specific primers according to the fragments by short PCRs. The primers were designed according to the conserved regions by the program Primer Premier 5.0. The *E. lanigerum* was regarded as an example. All the primers used in this study were synthesized by Invitrogen Biotech (Beijing, China) and are same to the primers for amplifying the mt genome of *M. keteleerifoliae* (Wang *et al.*, 2015a, the sequences and positions of primers). Short PCRs (sequence length <1.5 k) were performed with *Taq* DNA polymerase (TransGen Biotech, Beijing, China) at the following settings: 95°C for 3 min; 35 cycles of 95°C for 1 min, 48–55°C (depending on primer pairs) for 1 min and 72°C for 2 min. A final extension step of 10 min at 72°C was added after cycling. Long PCRs (sequence length >1.5 k) were carried out using high quality polymerase, such as High Fidelity (HiFi) *Taq*DNA polymerase (TransGen Biotech, Beijing, China) (Wang *et al.*, 2013 and this study), or LA Taq™ (TaKaRa Co., Dalian, China) (Zhang *et al.*, 2014, 2016b) under the following cycling conditions: 2 min at 92°C, 10 cycles (10 s at 92°C, 30 s at 50–55°C (depending on primer pairs), and 4–8 min at 68°C), 20 cycles (10 s at 92°C, 30 s at 50–55°C, and 4–8 min at 68°C with additional 30 s per cycle), and a final prolonged elongation of 10 min at 72°C. The annealing temperatures used in amplifying long fragments including control region and possible repeat region are 48–52°C,

others are 52–55°C. PCR products were examined by electrophoresis on a 1% agarose gel and purified using *EasyPure* PCR purification Kit (TransGen Biotech, Beijing, China). All short PCR fragments were sequenced directly in both strands. Long PCR fragments were cloned into pMD19-T sequencing vector (TaKaRa, Dalian, China), and sequenced using the primer walking strategy. Sequencing reactions were performed by BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and run on an ABI 3730 automated sequencer (Applied Biosystems).

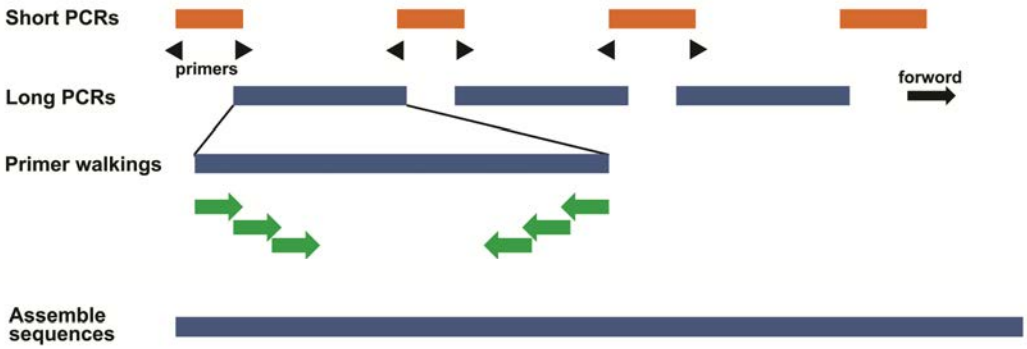

Figure 1. Procedures of sequencing one complete mt genome of aphid species.

2.3 Mt genome annotation and analyses

Except for sequencing methods, accurate annotations are necessary for mt genomes analyses (Fig. 2). Sequences were assembled using SeqMan (DNAStar Inc., Madison, WI, USA). Sequence annotation was performed using the BLAST tools in NCBI web site (http://blast.ncbi.nlm.nih.gov/Blast). The tRNAs were predicted by tRNAscan-SE Search Server v.1.21 (Lowe and Eddy, 1997) with default settings. Some tRNA genes that could not be found by tRNAscan-SE were identified by comparison with other aphids and manually edited. The 13 protein coding genes (PCGs) and two ribosomal RNA genes were identified by sequence similarity with all published aphid mt genomes (Table 1). The control regions or putative repeat regions were examined for regions of potential inverted repeats or palindromes by using the Mfold web

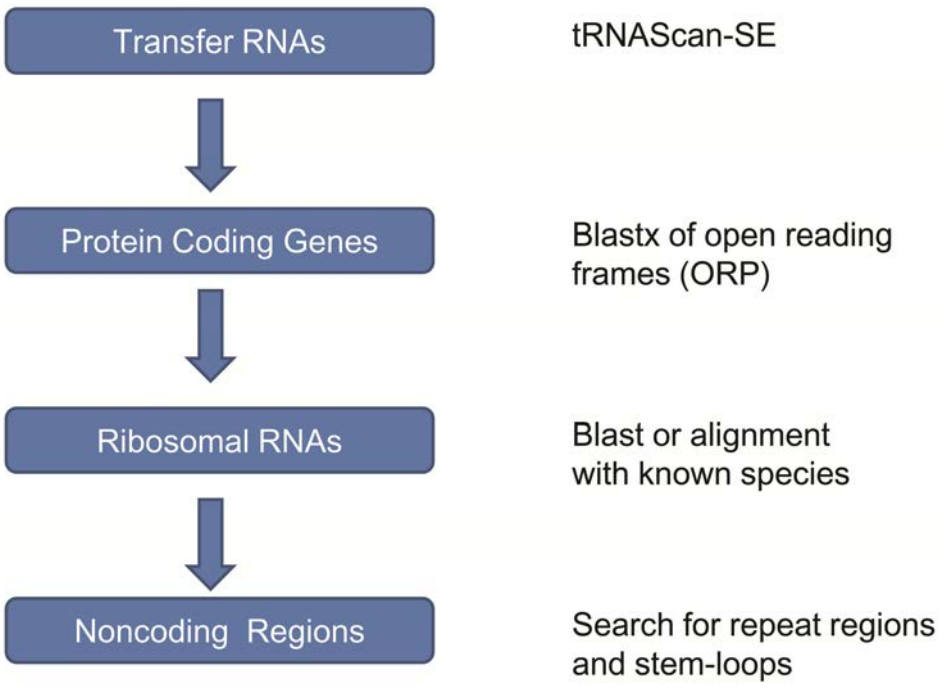

Figure 2. Steps of annotation one complete mt genome of aphid species.

server (http://www.bioinfo.rpi.edu/applications/mfold/) (Zuker, 2003).

The nucleotide sequences of PCGs were translated with the invertebrate mt genome genetic code. A+T content and codon usage of PCGs were calculated using MEGA version 6.05 (Tamura *et al*., 2013). Strand asymmetry was calculated using the formulas AT skew = (A−T)/(A+T) and GC skew = (G−C)/(G+C) (Perna & Kocher, 1995) for the strand encoding the majority of the PCGs.

2.4 Phylogenetic analysis

To construct the phylogenetic tree of aphids, concatenated PCGs were widely used (Wang *et al.*, 2013, 2015a). Here we also recommend this method because it is now the most common and suitable method for phylogenetic analysis of insects (Cameron, 2014). The multiple alignments of the 13 PCG nucleotide sequences of the 11 aphid mt genomes, including 10 species from Aphididae (Table 1) and one Phylloxeridae species as the outgroup, were done by the MEGA version 6.05 (Tamura *et al.*, 2013) and then manually proofread. Then, alignments of individual genes by codon were concatenated after deleting the stop codons.

Table 2. Organization of the *Eriosoma lanigerum* mt genome.

Gene	Strand	Position	Anticodon	Size（bp）	Start codon	Stop codon	Intergenic nucleotides*
cox1	J	1–1 531		1 531	ATA	T	0
tRNA-Leu	J	1 532–1 599	TAA	68			3
cox2	J	1 603–2 274		672	ATA	TAA	7
tRNA-Lys	J	2 282–2 353	CTT	72			2
tRNA-Asp	J	2 356–2 416	GTC	61			18
atp8	J	2 435–2 587		153	ATA	TAA	-20
atp6	J	2 568–3 221		654	ATT	TAA	-1
cox3	J	3 221–4 006		786	ATG	TAA	38
tRNA-Gly	J	4 045–4 113	TCC	69			-3
nad3	J	4 111–4 467		357	ATA	TAA	0
tRNA-Ala	J	4 468–4 532	TGC	65			-1
tRNA-Arg	J	4 532–4 592	TCG	61			0
tRNA-Asn	J	4 593–4 656	GTT	64			-1
tRNA-Ser	J	4 656–4 716	TCT	61			5
tRNA-Glu	J	4 722–4 783	TTC	62			10
tRNA-Phe	N	4 794–4 862	GAA	69			0
nad5	N	4 863–6 533		1 671	ATA	TAA	53
tRNA-His	N	6 587–6 655	GTG	69			5
nad4	N	6 661–7 969		1 309	ATA	T	3
nad4L	N	7 973–8 263		291	ATA	TAA	1
tRNA-Thr	J	8 265–8 327	TGT	63			8
tRNA-Pro	N	8 336–8 404	TGG	69			7
nad6	J	8 412–8 906		495	ATT	TAA	0
cob	J	8 907–10 022		1 116	ATG	TAA	23
tRNA-Ser	J	10 046–10 109	TGA	64			10
nad1	N	10 120–11 055		936	ATA	TAA	3
tRNA-Leu	N	11 059–11 123	TAG	65			0
rrnL	N	11 124–12 441		1 318			0
tRNA-Val	N	12 442–12 504	TAC	63			17
rrnS	N	12 522–13 297		776			0
control region		13 298–14 100		803			0
tRNA-Ile	J	14 101–14 164	GAT	64			3
tRNA-Gln	N	14 168–14 238	TTG	71			130
tRNA-Met	J	14 369–14 433	CAT	65			0
nad2	J	14 434–15 414		981	ATA	TAA	4
tRNA-Trp	J	15 419–15 482	TCA	64			-8
tRNA-Cys	N	15 475–15 540	GCA	66			23
tRNA-Tyr	N	15 564–15 632	GTA	69			8

* Indicates the intergenic spacer, negatives indicate the nucleotide number of gene overlap.

Maximum likelihood (ML) and Bayesian inference (BI) analyses were employed to do the phylogenetic analyses. ML tree could be run by using PHYML 3.0 (Guindon & Gascuel, 2003) or RAxML (Stamatakis *et al.*, 2008), and BI by MrBayes version 3.1.2 (Ronquist & Huelsenbeck, 2003). JModelTest 3.7 was used to select appropriate nucleotide substitution model for the entire matrix (Posada, 2008). The suitable model was chosen by the Bayesian information criterion (BIC) (Schwarz, 1978), and GTR+I+G was the best model for this study. ML analyses were also obtained from JModelTest under the optimal substitution model, and model parameter values were estimated during the analyses. Nodal support among branches was evaluated by bootstrap analysis with 100 replicates. In Bayesian inference, two independent runs were carried out, each with 10 000 000 generations and four chains. Each chain was sampled every 1 000 generations with a burn-in of 25%. Trees inferred prior to stationary were discarded as burn-in, and the remaining trees were used to construct a 50% majority-rule consensus tree with posterior probabilities (PP).

3 Results

3.1 General features of the mitochondrial genome of *E. lanigerum*

The *E. lanigerum* mt genome is a double-stranded circular molecule of 15 640 bp in length (Fig. 3) (GenBank accession number: KP722582). The *E. lanigerum* mt genome presented the 37 genes in atypical insect mt genome, including 13 PCGs, 22 tRNA genes, and two rRNA genes (Table 2). Twenty-three genes were transcribed on the majority

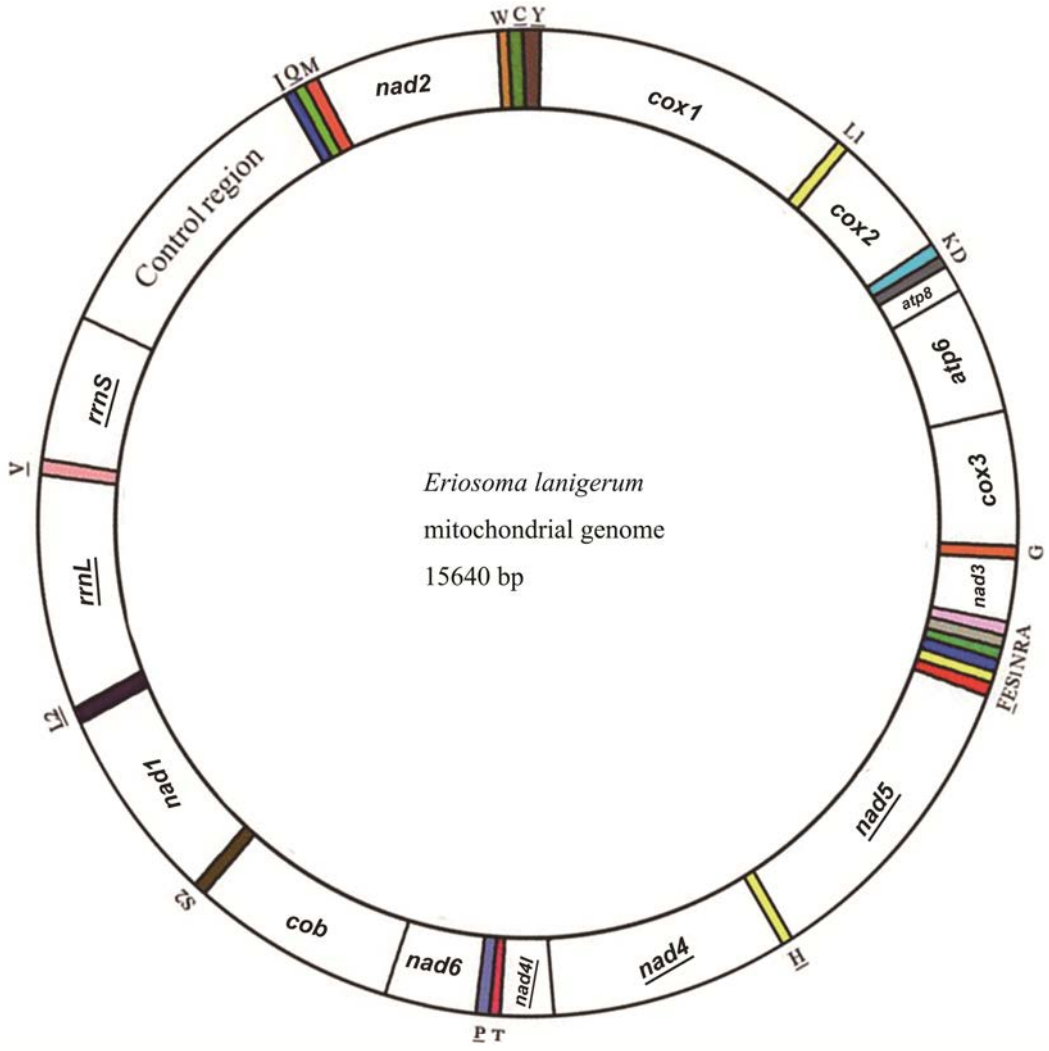

Figure 3. Circular map of the *Eriosoma lanigerum*mt genome. Gene names not underlined indicate the direction of transcription in the major strand, and underlined names mean the direction of transcription in the minor strand.

strand (J-strand), whereas the others were oriented on the minority strand (N-strand). Gene overlaps were found at 9 gene junctions; the longest overlap (20 bp) existed between *atp6* (ATP synthase F0 subunit 6) and *atp8* (ATP synthase F0 subunit 8). In addition to the control region, there were 198 nucleotides dispersed in 22 intergenic spacers, ranging from one to 130 bp. The longest spacer sequence was located between *tRNA-Gln* and *tRNA-Met*.

The nucleotide composition of the *E. lanigerum*mt genome was heavily A+T biased, with 84.7% A+T content (Table 3) measured from the J-strand. The total J-strand PCGs were much less T biased (AT-skew = –0.062) than the N-strand (AT-skew = –0.233). The skew statistics of the total PCGs indicated that the J-strand PCGs were CG-skewed and consisted of nearly equal A and T, while the N-strand PCGs were CG- and AT-skewed.

Table 3. Nucleotide composition of the *Eriosoma lanigerum* mt genome.

	A%	G%	C%	T%	A+T%	AT-skew	GC-skew	Length (bp)
Whole genome	44.9	5.6	9.7	39.8	84.7	0.060	-0.271	15 640
Protein-coding genes	54.1	5.0	9.7	31.2	85.3	0.269	-0.323	4 207
First codon	51.7	4.1	9.3	34.9	86.6	0.193	-0.383	1 403
Second codon	57.8	4.5	6.9	30.8	88.6	0.304	-0.213	1 402
Third codon	52.9	6.3	12.9	27.9	80.8	0.310	-0.346	1 402
Protein-coding genes-J	38.9	6.5	10.5	44.0	83.0	-0.062	-0.233	6 745
First codon	46.1	4.0	6.2	43.7	89.7	0.027	-0.212	2 249
Second codon	40.4	9.4	12.2	37.9	78.4	0.032	-0.128	2 248
Third codon	30.2	6.1	13.1	50.5	80.7	-0.251	-0.363	2 248
Protein-coding genes-N	54.1	5.0	9.7	31.2	85.3	0.269	-0.323	4 207
First codon	51.7	4.1	9.3	34.9	86.6	0.193	-0.383	1 403
Second codon	57.8	4.5	6.9	30.8	88.6	0.304	-0.213	1 402
Third codon	52.9	6.3	12.9	27.9	80.8	0.310	-0.346	1 402
tRNA genes	45.8	5.8	7.9	40.5	86.3	0.061	-0.152	1 446
rRNA genes	45.2	5.0	10.1	39.8	85.0	0.064	-0.338	2 094
Control region	40.1	3.7	8.2	47.9	88.0	-0.089	-0.375	803

In the *E. Lanigerum* mt genome (Table 2), all 22 typical animal tRNA genes with length from 61 to 72 bp were found. The typical clover-leaf structure was predicted in only 21 of the 22 mitochondrial tRNAs since the *tRNA-Ser(AGN)* gene included a DHU replacement loop instead of the typical DHU arm, and could not form a stem-loop structure in the TΨC arm (Fig. 4).

The boundaries of rRNA genes were implemented from the alignment with other aphid species (Thao *et al.*, 2004; Wang *et al.*, 2013, 2014). The *rrnL* of *E. lanigerum* was located between *tRNA-Leu(CUN)* and *tRNA-Val*, and *rrnS* resided between *tRNA-Val* and the control region (Fig. 3). The large ribosomal gene (*rrnL*) of *E. Lanigerum* was 1 318 bp, and the small ribosomal gene (*rrnS*) was 776 bp.

The *E. lanigerum*mt genome includes one large non-coding region, identified as the control region. The control region was rich in A+T (88.0%) and located between *rrnS* and *tRNA-Ile.* The control region of *E. Lanigerum* included three parts: the AT-rich region, the stem-loop structure, and the PolyT stretch regions.

3.2 Phylogenetic analyses

The newly sequenced *E. lanigerum*was combined with the mt genome sequences of nine aphid species for the phylogenetic analysis. The phylogenetic trees generated from ML analyses and Bayesian inferences showed similar topologies (Fig. 5). The monophyly of five subfamilies was recovered in different analyses and well supported. Within the Aphidinae subfamily, the monophyly of Aphidini and Macrosiphini had statistically high values, and is similar to the traditional taxonomic views based on morphology (Heie, 1992). *E. lanigerum* was clustered with *C. quercus*, and *H. betulae* was positioned most basally within the clade of Aphididae (Fig. 5)

4 Discussion

4.1 General methods to analyze one complete mitochondrial genome of aphid species

Here, we concluded the general methods to obtain and analysis one complete mt genome of aphid species with four

main steps.

First, to find all the genes including tRNAs, PCGs and rRNAs with different software and settings, and obtain the secondary structure of tRNAs (Fig. 4). Second, to annotate the control region and possible repeat region (Wang *et al.*, 2013; Zhang *et al.*, 2014) with special serial repeat sequences. Third, to compare with other species and find some other interesting structure or region of the genome, such as the tRNA isomerism of *tRNA-Gly* (Wang *et al.*, 2015a). In the last, to combine with other species and do some phylogenetic analyses in different categories (Wang *et al.*, 2013, 2015b; Li *et al.*, 2015).

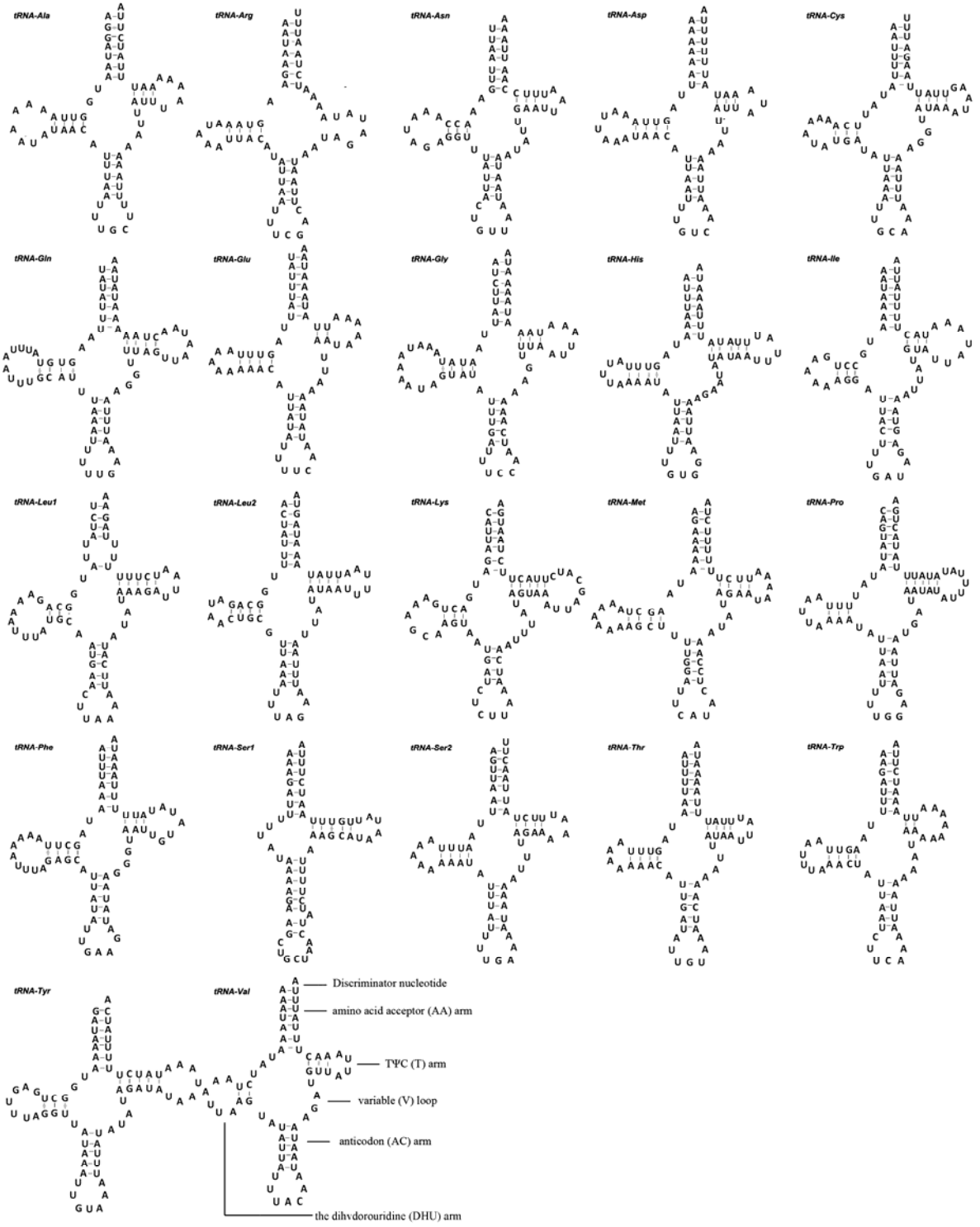

Figure 4. The secondary structure of the 22 transfer RNAs (tRNAs) in the *Eriosoma lanigerum* mt genome.

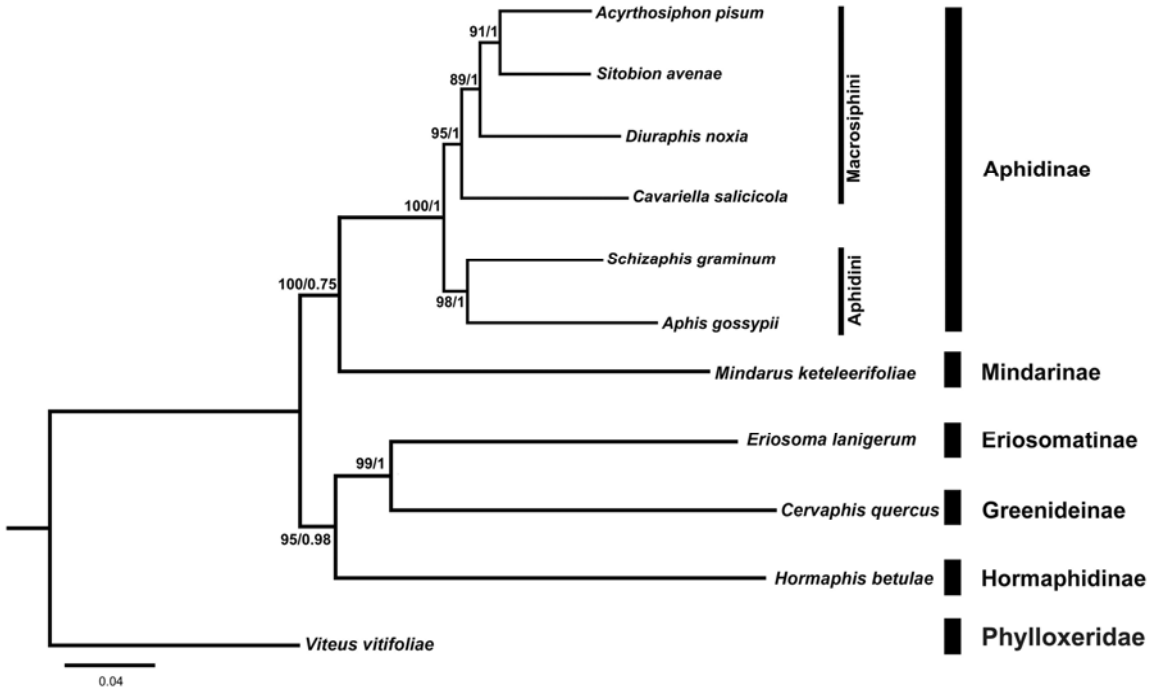

Figure 5. Bayesian inference (BI) and Maximum likelihood (ML) phylogenetic tree inferred from 11 aphids mt genome sequences. The support values on the nodes are the bootstrap (BS) values and the Bayesian posterior probabilities (BPP).

4.2 The mitochondrial genomes of aphid species

Though most of the features about aphid mt genomes are similar to the ancestral mt genome of insects, *Drosophila yakuba*, there are a few differences in mt genomes of aphids species in the following discussion.

The length variation among mt genomes of aphid species was minimal in PCGs, tRNAs, and the rRNAs, but obvious in the intergenic spacers, such as control regions and repeat regions (Thao *et al.*, 2004; IAGC, 2010; Wang *et al.*, 2013; Zhang *et al.*, 2014). The lengths of the control regions in aphid mt genomes are variable, as shown by the control region of *E. lanigerum* with a length of 803 bp. However, the length of *A. pisum* (1 336 bp), and *C. Salicicola* (1 137 bp), are the only two control regions more than 1k bp. Four species are similar length: *Mindarus keteleerifoliae* (687 bp), *C. quercus* (657 bp), *S. graminum* (682 bp) and *D. noxia* (664 bp). The only 509 bp of *H. betulae* is now the smallest control region of aphid species. These differences are because of their various structural patterns. Two species of Aphidinae contained tandem repeat sequences: *A. pisum* and *C. salicicola* (Wang *et al.*, 2013). However, the control regions of other aphid species without repeat sequences also have the conserved structural pattern found in *E. lanigerum* with three parts: the AT-rich region, the stem-loop structure, and the PolyT stretch regions, which were proposed as a widespread feature in Aphididae (Wang *et al.*, 2014, 2015a).The repeat region between *trnE* and *trnF* is also an interesting feature of the aphid mt genomes. This unique region contains serially variable numbers of tandem repeats, but this repeat region was not found in the mt genome of *E. Lanigerum* and *H. betulae* (Li *et al.*, 2015). This new evidence implies again that the repeat region may be typical for Aphidinae but not of other subfamilies in Aphididae (Wang *et al.*, 2014). So, it was proposed that the repeat region within Aphidinae is lineage specific and occurred from independent evolutionary events.

The typical clover-leaf structure was predicted in only 21 of the 22 mitochondrial tRNAs since the *tRNA-Ser(AGN)* gene contained a DHU replacement loop instead of the typical DHU arm, and could not form a stem-loop structure in the TΨC arm (Fig. 4). In many arthropod mt genomes, this is a common feature (Wolstenholme, 1992).

5 Conclusions

This paper takes the complete mt genome of *E. lanigerum* as an example to conclude the methods of sequencing (Fig. 1) and annotating (Fig. 2) one complete mt genome of aphid species. In the future studies, larger taxon sampling sizes

could follow the methods to open new view of the research on mt genomes.

Funding This work was supported by the National Natural Sciences Foundation of China (31430078, 31561163002), the External Cooperation Program of Chinese Academy of Sciences (152111KYSB20130012), and the Ministry of Science and Technology of the People's Republic of China (2011FY120200, SQ2012FY4910058).

Acknowledgments We thank Xiaolei Huang for collecting the specimens of *E. lanigerum*, and Ms. Fendi Yang for preparing aphid slides.

References

Blackman, R.L., Eastop, V.F. 2000. Aphids on the world's crops -an Identification and Information Guide, 2nd edn. John Wiley & Sons, Chichester. UK. 476 pp.

Cameron, S.L. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annual Review of Entomology*, 59: 95–117.

Clary, D.O., Wolstenholme, D.R. 1985. The mitochondrial DNA molecular of *Drosophila yakuba*: Nucleotide sequence, gene organization and genetic code. *Journal of Molecular Evolution*, 22: 252–271.

Dowton, M., Cameron, S.L., Austin, A.D., Whiting, M.F. 2009. Phylogenetic approaches for the analysis of mitochondrial genome sequence data in the Hymenoptera – a lineage with both rapidly and slowly evolving mitochondrial genomes. *Molecular Phylogenetics & Evolution*, 52: 512–519.

Favret, C. 2016. Aphid species file. Version 5.0/5.0. Available from http://Aphid.SpeciesFile.org (accessed 2 March 2016).

Guindon, S., Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, 52: 696–704.

Heie, O.E. 1992. The Aphidoidea (Hemiptera) of Fennoscandia and Denmark. IV Family Aphididae: Part 1 of the tribe Macrosiphini of subfamily Aphidinae. *In*: Kristensen, N.P., Michelsen V. (eds). *Fauna Entomologica Scandinavica Volume 25*. E.J. Brill/ Scandinavian Science Press Ltd. Leiden, New York, Kobenhavn, Köln. 188pp.

Li, Y.Q., Jing, C., Qiao, G.X. 2015. Complete mitochondrial genome of the aphid *Hormaphis betulae* (Mordvilko) (Hemiptera: Aphididae: Hormaphidinae). *Mitochondrial DNA*. doi:10.3109/19401736.2015.1118071.

Lowe, T.M., Eddy, S.R. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, 25: 955–964.

Ma, C., Yang, P.C., Jiang, F., Chapuis, M.P., Shall, Y., Sword, G.A., Kang, L. 2012. Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. *Molecular Ecology*, 21: 4 344–4 358.

Nelson, L.A., Lambkin, C.L., Batterham, P., Wallman, J.F., Dowton, M., Whiting, M.F., Yeates, D.K., Cameron, S.L. 2012. Beyond barcoding: genomic approaches to molecular diagnostics in blowflies (Diptera: Calliphoridae). *Gene*, 511: 131–142.

Perna, N.T., Kocher, T.D. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *Journal of Molecular Evolution*, 41: 353–358.

Posada, D. 2008. jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, 25: 1 253–1 256.

Remaudière, G., Remaudière, M. 1997. *Catalogue of the World's Aphididae. Homoptera Aphidoidea*. Institut National de la Recherche Agronomique. Paris, France. 473pp.

Ronquist, F., Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19: 1 572–1 574.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics*, 6: 461–464.

Stamatakis, A., Hoover, P., Rougemont, J. 2008. A rapid bootstrap algorithm for the RAxML Web-servers. *Systematic Biology*, 57: 758–771.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30: 2 725–2 729.

Thao, M.L., Baumann, L., Baumann, P. 2004. Organization of the mitochondrial genomes of whiteflies, aphids, and psyllids (Hemiptera, Sternorrhyncha). *BMC Evolutionary Biology*, 4: 25.

The International Aphid Genomics Consortium (IAGC). 2010. Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biology*, 8: e1000313.

Wang, Y., Chen, J., Jiang, L.Y., Qiao, G.X. 2015a. The complete mitochondrial genome of *Mindarus keteleerifoliae* (Insecta: Hemiptera: Aphididae) and comparison with other Aphididae insects. *International Journal of Molecular Sciences*, 16(12): 30 091–30 102.

Wang, Y., Chen, J., Jiang, L.Y., Qiao, G.X. 2015b. Hemipteran mitochondrial genomes: features, structures and implications for phylogeny. *International Journal of Molecular Sciences*, 16(6): 12 382–12 404.

Wang, Y., Huang, X.L., Qiao, G.X. 2013. Comparative analysis of mitochondrial genomes of five aphid species (Hemiptera: Aphididae) and phylogenetic implications. *PLoS One*, 8: e77511.

Wang, Y., Huang, X.L., Qiao, G.X. 2014. The complete mitochondrial genome of *Cervaphis quercus* (Insecta: Hemiptera: Aphididae: Greenideinae). *Insect Science*, 21: 278–290.

Wolstenholme, D.R. 1992. Animal mitochondrial DNA: structure and evolution. *International Review of Cytology*, 141: 173–216.

Zhang, B., Ma, C., Edwards, O., Fuller, S., Kang, L. 2014. The mitochondrial genome of the Russian wheat aphid *Diuraphis noxia*: Large repetitive sequences between *trnE* and *trnF* in aphids. *Gene*, 533: 253–260.

Zhang, S., Luo, J., Wang, C., Lv, L., Li, C., Jiang, W., Cui, J., Rajput, L.B. 2016a. Complete mitochondrial genome of *Aphis gossypii* Glover (Hemiptera: Aphididae). *Mitochondrial DNA*, 27(2): 854–845. doi: 10.3109/19401736.2014.919474.

Zhang, B., Zheng, J., Liang, L., Fuller, S., Ma, C.S. 2016b. The complete mitochondrial genome of *Sitobion avenae* (Hemiptera: Aphididae). *Mitochondrial DNA*, 27(2): 945–946. doi: 10.3109/19401736.2014.926498.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, 31: 3406–3415.