The complete mitochondrial genome and gene rearrangements in a gall wasp species, *Dryocosmus liui* (Hymenoptera: Cynipoidea: Cynipidae)

Cheng-Yuan Su[1], Dao-Hong Zhu[1], Yoshihisa Abe[2], Tatsuya Ide[3] and Zhiwei Liu[4]

[1] Laboratory of Insect Behavior and Evolutionary Ecology, College of Life Science and Technology, Central South University of Forestry & Technology (CSUFT), Changsha, Hunan, China
[2] Faculty of Social and Cultural Studies, Kyushu University, Fukuoka, Kyushu, Japan
[3] Department of Zoology, National Museum of Nature and Science, Tsukuba, Ibaraki, Japan
[4] Biological Sciences Department, Eastern Illinois University, Charleston, IL, United States of America

Corresponding author
Dao-Hong Zhu,
daohongzhu@yeah.net

ABSTRACT

Mitochondrial genomes (mitogenomes) have been widely used in comparative and evolutionary genomics, molecular evolution, phylogenetics, and population genetics, but very limited information is available for the family Cynipidae. In this report, we describe the mitogenome of *Dryocosmus liui* Pang, Su et Zhu, providing the first complete mitogenomic data for a cynipid gall wasp species. The mitogenome of *D. liui* is 16,819 bp in length, and contains the typical set of 37 genes. Two control regions were detected, with the second being a perfect inverted repeat of the major portion of the first. Gene rearrangements were found in transfer RNA (tRNA) genes, protein-coding genes (PCGs) and ribosomal RNA (rRNA) genes, compared with the putative ancestral mitogenome. Similar to two other Cynipidae species with mitogenome data available, *D. liui* has a novel tRNA gene cluster *trnL1–trnI–trnL2–trnW–trnM–trnQ* between *nad1* and *nad2*. Phylogenetic analysis based on sequences of PCGs and rRNA genes with *D. liui* included obtained topologies identical to previous studies supporting the a relationship of (Cynipoidea , (Platygastroidea, Proctotrupoidea)) within the monophyletic Proctotrupomorpha and (Cynipidae, Figitidae), Ibaliidae) within the Cynipoidea.

INTRODUCTION

The mitochondrial genome (mitogenome) of most animals is a double-stranded circular DNA molecule and usually includes 13 protein-coding genes (PCGs), 22 transfer RNAs (tRNAs), two ribosomal RNA (rRNA) genes, and a major non-coding sequence called the control region. The size of animal mitogenomes ranges from 13 to 20 kb (*Boore, 1999*; *Cameron, 2014*). Given its maternal inheritance, low rate of recombination, relatively high evolutionary rate and conserved gene components, mitogenomes have been widely used in comparative and evolutionary genomics, molecular evolution, phylogenetics, and

population genetics (*Simon et al., 2006*; *Dowton et al., 2009*; *Cameron, 2014*; *Mao, Gibson & Dowton, 2015*; *Feng et al., 2020*). In recent years, with the development of sequencing technology and the decreased associated costs, there has been a significant increase in the availability of whole mitogenome data, including for hymenopteran insects. Several remarkable features are observed in the hymenopteran mitogenome, including a high A + T content, large variation in substitution rates between lineages, and frequent gene rearrangement (*Dowton et al., 2009*; *Cameron, 2014*; *Mao, Gibson & Dowton, 2014*; *Kahnt et al., 2015*). Most metazoan mitogenomes have rarely experienced gene rearrangements in the course of evolution and such events usually involve only a few tRNA genes (*Cameron, 2014*). By contrast, gene rearrangements are frequently found in Hymenoptera in tRNA genes and protein-coding genes (*Mao, Gibson & Dowton, 2014*; *Mao, Gibson & Dowton, 2015*; *Kahnt et al., 2015*; *Chen et al., 2018*; *Tang et al., 2019*). Rearrangements of mitochondrial genes include transposition, inversion, and inverse transposition (*Cameron, 2014*). The gene order of the mitogenome contains phylogenetic signals, which may be informative of for phylogenetic relationship among groups (*Sankoff et al., 1992*; *Dowton et al., 2009*; *Tang et al., 2019*).

Cynipidae are a phytophagous group in the superfamily Cynipoidea (Hymenoptera), and its members are well known as gall wasps because they usually induce structurally complex galls on different plant organs, *i.e.,* flowers, leaves, buds, stems, twigs, and roots. It is the second most species-rich group of gall inducers after the gall midge family Cecidomyiidae (Diptera), with roughly 1,400 described species (*Ronquist et al., 2015*). In addition to gall formers, the Cynipidae also include phytophagous inquilines, which live inside the galls induced by another species. *Dryocosmus liui Pang, Su & Zhu (2018)* is a recently described species from Hunan Province in central south of China, forming galls on branches of *Castanopsis tibetana* Hance. The genus *Dryocosmus* Giraud is of particular interest among Cynipidae because of its wide distribution, use of host plants belonging to multiple genera in the beech family (Fagaceae), and include a species that is completely parthenogenic (*Buffington & Morita, 2009*; *Zhu et al., 2015*; *Pang, Su & Zhu, 2018*). In addition, two species of the genus are horticultural pests, *Dryocosmus kuriphilus,* the well known Oriental chestnut gall wasp is a global horticultural pest that causes serious damage to the chestnut industry in its native China as well as in invaded locations (*Rieske, 2007*; *Bosio, Gerbaudo & Piazza, 2010*; *Zhu et al., 2015*; *Martinez-Sañudo et al., 2019*), while the recently described *D. zhuili* causes significant damage to *Castanea henryi* in southern China (*Zhu et al., 2015*).

The number of complete mitogenome sequences available has increased dramatically in Hymenoptera in recent years, but for Cynipidae, only incomplete mitogenome sequences have been reported for two species, *Synergus* sp. (*Tang et al., 2019*) and *Trichagalma acutissimae* (*Xue et al., 2020*). In this study, the mitogenome of *D. liui* was sequenced and annotated, providing the first complete mitogenome sequence in the family Cynipidae. The mitogenome structure and gene rearrangements in this lineage were analyzed. In addition, the phylogeny of Proctotrupomorpha was analyzed based on available mitogenome sequences including that of *D. liui* sequenced in the present study.

MATERIALS AND METHODS

Insects and DNA extraction

Galls of *D. liui* were collected on branches of *C. tibetana* from Yanling, Hunan, China (113°93′E, 26°57′N) in March 2016 and March 2020. The galls collected were cage-reared at room temperature in the laboratory of CSUFT. Adults were preserved directly in 100% ethanol at −80 °C within 4 days after emergence until DNA extraction. The gall wasps were washed twice with sterile water to remove surface contamination before DNA extraction. Genomic DNA was extracted from each individual using SDS/proteinase K digestion and phenol–chloroform extraction methods as previously described (*Zhu et al., 2007*). The DNA was resuspended in sterile water and stored at 4 °C.

Mitogenome sequencing and assembly

Short fragments of the *cox1*, *cob*, and *rrnL* genes were amplified using universal insect mitochondrial primers designed by *Folmer et al. (1994)* and *Simon et al. (1994)* (Table S1). The PCR products were purified and sequenced directly using the Sanger method at Wuhan Icongene Co, Ltd. (Wuhan, China). Based on the sequences obtained, specific PCR primers (cox1F, cox1R; cobF, cobR and rrnLF, rrnLR) (Table S2) were subsequently designed to amplify the remaining genome by long PCR. The reaction mixture was composed of 1 μL PrimeSTAR® GXL DNA Polymerase (Takara Biomedical Technology Co., Ltd, Dalian, China), 10 μL buffer, 4 μL dNTPs, 1 μL of each primer, and 2 μL of DNA then water was added to give a total volume of 50 μL. The amplification was conducted using a C1000 touch thermal cycler (Bio-Rad, Hercules, CA, USA). The cycling conditions were 98 °C for 3 min, 35 cycles of 98 °C for 10 s, 55–60 °C for 30 s and 68 °C for 10 min. To obtain the complete mitogenome sequence of *D. liui*, the following amplification steps were performed. First, referring to the known mitogenome organization of Cynipoidea (*Mao, Gibson & Dowton, 2015*; *Tang et al., 2019*; *Xue et al., 2020*), the primer pairs cox1F/cobR, cobF/rrnLF and rrnLR/cox1R were used to obtain large fragments. A single band of about 9 kb was successfully obtained using the cox1F/cobR primer pairs, but multiple bands appeared when the other two primer pairs were used. Second, because rearrangements of mitochondrial genes are frequently found in Hymenoptera, including transposition, inversion, and inverse transposition, (*Mao, Gibson & Dowton, 2014*; *Mao, Gibson & Dowton, 2015*; *Kahnt et al., 2015*; *Chen et al., 2018*; *Tang et al., 2019*), the *D. liui* mitochondrial gene may include rearrangements. Thus, we used cobF/rrnLR and rrnLF/cox1R primer pairs to amplify the rest of the genome fragments, and were able to obtain about 4 kb and 3 kb target bands, respectively. Finally, to verify the accuracy of the results, the full-length sequence of the *D. liui* mitogenome was amplified using cox1F/rrnLR and cobR/rrnLF primer pairs, and single target bands of about 12 kb and 14 kb were obtained, respectively (Fig. S1).

Those PCR products were subsequently purified using a TaKaRa MiniBEST Agarose Gel DNA Extraction Kit Ver.4.0 (Takara Biomedical Technology Co.). The primer walking method was employed to determine the sequence for each long PCR product using an ABI 3730XL DNA sequencer (Applied Biosystems, Foster City, CA, USA) at Wuhan Icongene Co, Ltd. (for sequencing primers see Tables S2–S3). Raw sequences were assembled into contigs in ChromasPro Ver 1.33 (Technelysium Ltd., Tewantin, Australia), then checked

and corrected manually. The sequencing results of the long PCR products obtained in the third step described above were completely consistent with those in the first and second steps, validating the reliability of the amplification strategy used in this study.

Mitogenome annotation and genome feature analysis

The PCGs were identified by the ORF Finder in NCBI (http://www.ncbi.nlm.nih.gov), by specifying the invertebrate mitochondrial genetic code and were confirmed by alignment with homologous genes from closely related species in Geneious 10.1.3. The start and stop codons of some genes were corrected according to the boundaries of tRNA genes and through alignment with other hymenopteran mitogenome sequences. The tRNA genes were initially identified using tRNA-scan SE 1.21 (http://lowelab.ucsc.edu/tRNAscan-SE/), with the parameters: source = "Mito/Chloromast" and genetic code = "Invertebrate Mito". Twenty-one of the 22 typical animal mitochondrial tRNA genes were identified. The remaining one tRNA and two rRNA genes were identified by the MITOS WebServer using the invertebrate mitochondrial genetic code (http://mitos.bioinf.uni-leipzig.de/index.py). The rRNA genes were confirmed by sequence comparison with published hymenopteran mitochondrial rRNA sequences. The annotated complete mitogenome sequence of *D. liui* has been submitted to GenBank with the accession number MW368384.

The nucleotide composition of the mitogenome was calculated using MEGA 7.0. We also measured the AT- and GC-skews using the following formulas: AT-skew = $(A − T)/(A + T)$, GC-skew = $(G − C)/(G + C)$, for whole mitogenome sequences and 13 PCGs. The tandem repeats in the two control regions were predicted using the Tandem Repeats Finder available online (https://tandem.bu.edu/trf/trf.html).

Phylogenetic analyses

For the phylogenetic analyses, the mitogenome sequences of 19 species of Proctotrupomorpha, including five species of Cynipoidea, were included and additional four species of the sister clade Ichneumonomorpha were used as outgroups (Table S4). Nucleotide sequences for each of the two rRNA genes and the 13 protein-coding genes were imported into separate files using MEGA7.0.

The two rRNA genes were individually aligned using the MAFFT 7.0 online server with a Q-INS-i strategy (*Katoh & Standley, 2013*). Each protein-coding gene was aligned individually using codon-based multiple alignments and the MAFFT algorithm within the TranslatorX server. A nucleotide alignment was inferred from the amino acid alignment. All alignments were subsequently checked and corrected manually in Geneious 10.1.3 for quality control. Individual gene alignments were concatenated using SequenceMatrix prior to phylogenetic analysis. Alignments were concatenated as two datasets: (1) the PCG123+2rRNA matrix, including all three codon positions of PCGs and two rRNA genes; and (2) the PCG12+2rRNA matrix, including the first and second codon positions of PCGs and two rRNA genes. Maximum likelihood, Bayesian and PhyloBayes approaches were employed to infer phylogenetic trees using MrBayes v. 3.2.7a, IQ-Tree 2.1.1, and PhyloBayes MPI 1.7b, respectively, *via* the online CIPRES Science Gateway portal (*Miller, Pfeiffer & Schwartz, 2010*). For Bayesian analyses, the best partitioning schemes and corresponding
nucleotide substitution models were determined with PartitionFinder2 v.2.1.1 *via* the online CIPRES Science gateway portal, and two independent runs were performed with the default priors and MCMC parameters except that the MCMC runs comprised 10 million generations sampled at every 1,000 generations with 25% burn-in time. For maximum likelihood analyses, a total of 10,000 bootstrap replicates were obtained with the auto model applied to all partitions. For PhyloBayes analyses, phylogenetic trees were inferred from the datasets with the CAT + GTR model.

RESULTS AND DISCUSSION

Genome structure and nucleotide composition

The complete mitogenome of *D. liui* was sequenced, and its length was 16,819 bp (GenBank accession number: MW368384). The mitogenome contained the typical gene repertoire of 13 protein-coding genes (*nad1-6*, *nad4L*, *cox1-3*, *atp6*, *atp8*, and *cob*), two rRNA genes (*rrnL*, *rrnS*) and 22 tRNA genes; among these, 22 genes were found on the majority strand (J-strand) and 15 genes were on the minority strand(N-strand) (Fig. 1 and Table 1). A total of 660 bp of intergenic nucleotides ranging from 2 to 104 bp were found in 21 locations. The longest intergenic spacer (104 bp) was found between *nad1* and *trnL*. There were 112 overlapping nucleotides dispersed in eight locations, *i.e.*, *trnM-trnQ*, *trnQ-nad2*, *cox2-trnk*, *atp8-atp6*, *trnA-trnR*, *trnR-trnN*, *trnE-nad5*, and *nad4-nad4l*, and their length ranged from 1 to 41 bp (Table 1). The whole mitogenome had an overall base composition of 41.9% A, 41.8% T, 8.1% G, and 8.2% C, with a high AT bias of 83.7% (Table S5), as reported for most hymenopteran insects (*Chen et al., 2018*). The mitogenome has a positive AT-skew value (0.0016) and negative GC-skew value (−0.0095), biased for the A and C nucleotides (Table S5).

PCG, tRNA, and rRNA genes

In the *D. liui* mitogenome, eight of the 13 protein-coding genes were located on the majority strand, while the other five protein-coding genes were encoded by the minority strand. The average A + T content of the 13 protein-coding genes was 81.7%, and the highest was 89.8% for *nad2* (Table 1, Table S5). All PCGs started with the standard ATN codons, as in most other insect mitogenomes. There were six genes (*nad2*, *cox1*, *atp6*, *cox3*, *nad4*, and *cob*) starting with ATG, five genes (*atp8*, *nad3*, *nad5*, *nad4l*, and *nad6*) starting with ATT, and two genes (*nadl* and *cox2*) starting with ATA. Most of the PCGs terminated with a TAA stop codon, while *atp8* and *nad6* ended with a TAG codon (Table 1).

The 22 tRNA genes of *D. liui* were spread across nine clusters; 12 were on the majority strand while 10 were located on the minority strand. The tRNA genes ranged from 63 bp (*trnF*) to 79 bp (*trnK*) (Table 1). All the tRNA genes could be folded into the typical cloverleaf secondary structure, except for *trnT* which lacked a complete TψC-loop (Fig. S2). In most insects, the length of the acceptor arm (7 bp) and of the anticodon arm (stem: 5 bp, loop: 7 bp) is highly conserved across tRNAs whereas the DHU- and TψC-arms are quite variable in size (*Kahnt et al., 2015*). In the *D. liui* mitogenome, the anticodon arms of *trnA* and *trnM* were 4 bp in stems and 9 bp in loops. The stem of the DHU-arm ranged from 2 to 4 bp and the loop from 4 to 10 bp; the stem of the TψC-arm varied between 3
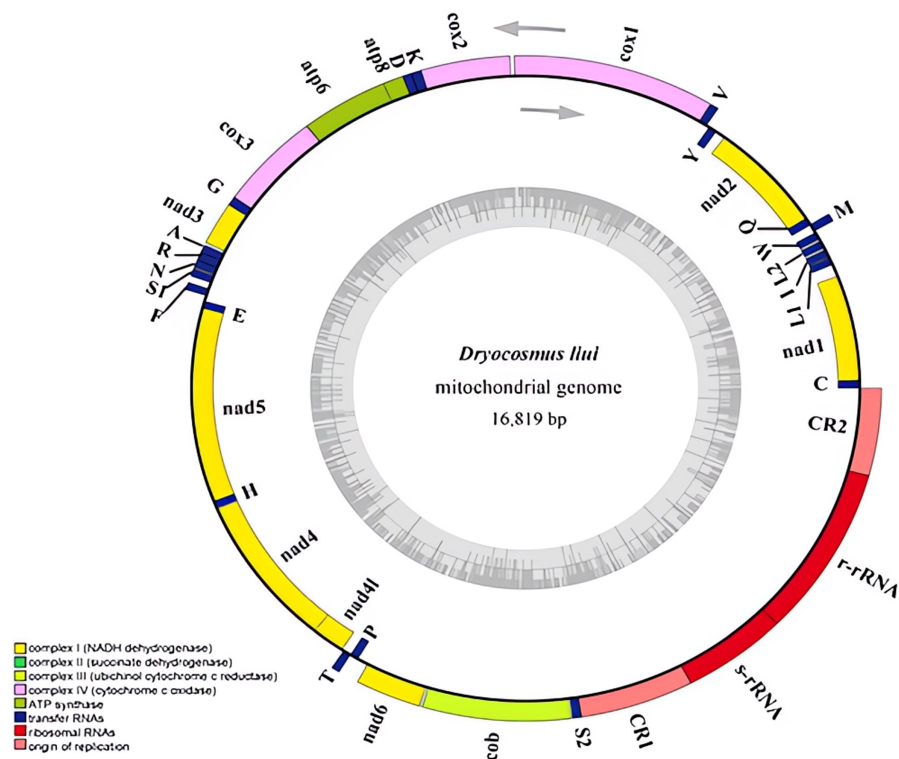
Figure 1 Mitochondrial genome of *Dryocosmus liui* sequenced in this study. Genes outside the circle are encoded by the major coding strand, and inside genes are encoded by the minor strand. Abbreviations of gene names are: atp6 and atp8 for atp-synthase subunits 6 and 8, cox1-3 for cytochrome oxidase subunits 1–3, cob for cytochrome b, nad1-6 and nad4L for NADH dehydrogenase subunits 1–6 and 4 L, l-rRNA and s-rRNA for arge and small rRNA subunits. tRNA genes are indicated with their one-letter corresponding amino acids. CR for control region. The GC content is plotted using a black sliding window.

and 5 bp; and the loop between 3 and 9 bp (but *trnk* is 16 bp). Sixteen mismatched base pairs were detected in the tRNAs of *D. liui*, all of which were G-U or U-U pairs.

The positions of the two rRNAs appear to be highly conserved across Hymenoptera (*Mao, Gibson & Dowton, 2015*). However, *D. liui* is different from most hymenopteran insects: the large and small rRNA genes(*rrnL* and *rrnS*), are next to each other between two large non-coding regions on the minority strand; the length of the *rrnS* gene was 808 bp, with an A + T content of 90.5%, while the *rrnL* gene was 1391 bp with an A + T content of 89.4% (Table 1).

Control region

Two copies of control regions (A + T rich regions) were detected in the *D. liui* mitogenome and were flanked by *trnS2* and *rrnS* (872 bp in length), and *rrnL* and *trnC* (713 bp in length), respectively (Table 1). In the first copy, three 69 bp and two 54 bp nontandem repeat units, a 122 bp A + T rich region (AT% = 94.3%) and a 129 bp A + T rich region (AT% = 90.6%) were present. The second copy was a perfect inverted repeat of the major portion of the first copy, including three 69 bp and two 54 bp repeat units,

Table 1 Annotation of the *Dryocosmus liui* mitochondrial genome.

Gene	Position (bp)	Size (bp)	Strand*	Intergenic nucleotides	Anti or Start codon	Stop codon	A + T%
trnC	1–64	64	−	2	GCA		89.1
nad1	67–984	918	−	104	ATA	TAA	80.6
trnL1	1089–1115	67	−	6	TAG		90.1
trnI	1162–1231	70	−	14	GAT		81.4
trnL2	1246–1315	70	−	14	TAA		87.1
trnW	1330–1398	69	−	13	TCA		91.3
trnM	1412–1477	66	+	−5	CAT		89.4
trnQ	1473–1541	69	−	−41	TTG		88.4
nad2	1501–2499	999	−	81	ATG	TAA	89.8
trnY	2581–2649	69	−	25	GTA		89.8
trnV	2675–2738	64	+	6	TAC		92.2
cox1	2745–4292	1548	+	31	ATG	TAA	74.7
cox2	4324–5022	699	+	−20	ATA	TAA	79.4
trnK	5003–5081	79	+	0	TTT		88.6
trnD	5082–5148	67	+	0	GTC		92.5
atp8	5149–5307	159	+	−7	ATT	TAG	87.4
atp6	5301–5975	675	+	5	ATG	TAA	82.3
cox3	5981–6772	792	+	0	ATG	TAA	77.7
trnG	6773–6844	72	+	0	TCC		91.7
nad3	6845–7195	351	+	21	ATT	TAA	85.2
trnA	7217–7281	65	+	−2	TGC		86.2
trnR	7280–7348	69	+	−1	TCG		85.5
trnN	7348–7414	67	+	7	GTT		85.1
trnS1	7422–7487	66	+	54	TCT		87.9
trnF	7542–7604	63	+	47	GAA		88.9
trnE	7652–7716	65	−	−29	TTC		95.4
nad5	7688–9307	1620	−	90	ATT	TAA	84.0
trnH	9398–9461	64	−	9	GTG		79.7
nad4	9471–10790	1320	−	−7	ATG	TAA	83.9
nad4l	10784–11056	273	−	16	ATT	TAA	83.9
trnT	11073–11137	65	+	0	TGT		87.7
trnP	11138–11207	70	−	88	TGG		82.9
nad6	11296–11805	510	+	16	ATT	TAG	89.0
cob	11823–12956	1134	+	11	ATG	TAA	77.5
trnS2	12968–13035	68	+	0	TGA		85.3
CR1	13036–13907	872		0			82.6
rrnS	13908–14715	808	+	0			90.5
rrnL	14716–16106	1391	+	0			89.4
CR2	16107–16819	713		0			80.5

Notes.
*+ indicates the gene is coded on majority strand while −indicates the gene is coded on minority strand.
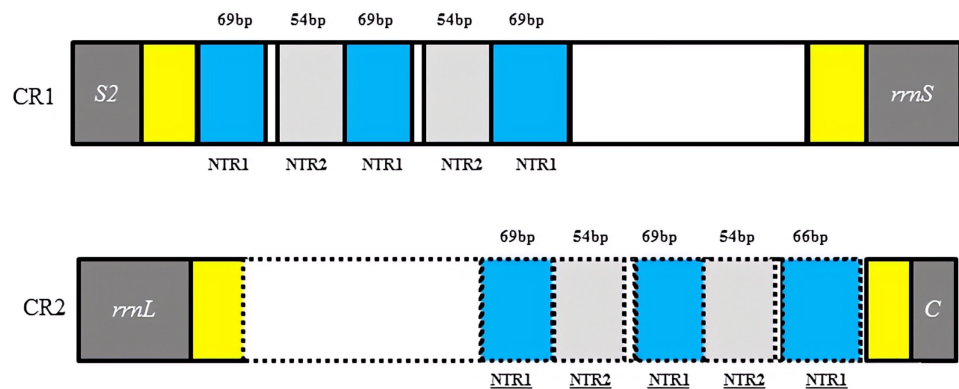
Figure 2 **Structures of the two control regions in *Dryocosmus liui* mitochondrial genome.** NTR indicates nontandem repeat. Yellow shows A + T rich regions. Within the dotted line in CR2 is the inverted, duplicated sequence of CR1.

a 65 bp A + T rich region (AT% = 95.5%), and a 29 bp A + T rich region (AT% = 82.8%) (Fig. 2). Duplicated control regions (or multiple control regions) are not universal among insect species, and are mainly found in barklice, thrips, and wasps (*Shao & Barker, 2003*; *Mao, Gibson & Dowton, 2015*; *Yoshizawa et al., 2018*; *Tyagi et al., 2020*). Among Hymenoptera, this phenomenon has been found in *Orthogonalys pulchella* (Trigonalyoidea: Trigonalyidae), *Ceraphron* sp. (Ceraphronoidea: Ceraphronidae) (*Mao, Gibson & Dowton, 2014*), *Ibalia leucospoides* (Cynipoidea: Ibaliidae) (*Mao, Gibson & Dowton, 2015*), *Alloxysta* sp. (Cynipoidea: Figitidae) (*Tang et al., 2019*), and *Aphidius gifuensis* (Ichneumonoidea: Aphidiinae) (*Feng et al., 2020*). To our knowledge, the inverted, duplicated control regions have been found only in Cynipoidea, *i.e., D. liui* (this study) and *Ibalia leucospoides* (*Mao, Gibson & Dowton, 2015*). It is proposed that the presence of two control regions may be advantageous and be maintained either through stabilizing selection or through gene conversion (*Eberhard, Wright & Bermingham, 2001*). Furthermore, a polyT stretch at the 5′ end of the control region and a [TA(A)]n-like stretch following the polyT stretch could be found in the two copies of the control region, and have been proposed as initiation sites for replication and transcription of the mitochondrial genome (*Wei et al., 2010*).

Gene arrangement

Gene rearrangements have been found not only in tRNA genes but also in protein-coding genes in several species of Hymenoptera (*Mao, Gibson & Dowton, 2014*; *Mao, Gibson & Dowton, 2015*; *Kahnt et al., 2015*; *Chen et al., 2018*; *Tang et al., 2019*). In these mitogenomes, the gene clusters between *cox2–atp8*, *nad3–nad5*, *nad2–cox1*, and control region–*nad2* have been identified as the most frequently rearranged regions (*Mao, Gibson & Dowton, 2014*; *Wei et al., 2014*; *Chen et al., 2018*). When compared with the putative ancestral pattern of the pancrustacean/hymenopteran mitogenome (*Cook, 2005*; *Mao et al., 2012*), rearrangement events of the *D. liui* mitogenome were found to be mainly on the gene clusters between *nad3–nad5*, *nad2–cox1*, and control region–*nad2*, involving PCG, tRNA, and rRNA genes, respectively (Fig. 3). There are nine tRNA gene rearrangements
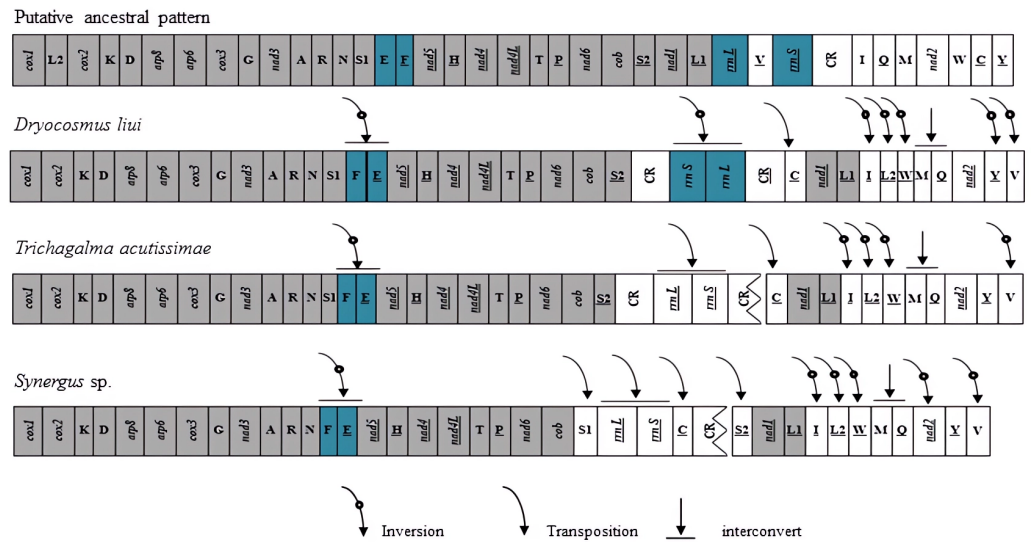
Figure 3 **Mitochondrial genome organization and gene rearrangement in Cynipidae compared with the ancestral type of the insect mitochondrial genome.** The gene order is linearized for easy view. The gene blocks with inversion are shown in dark blue, while the conserved gene blocks are showed in grey.

relative to their ancestral positions: *trnE* and *trnF* have inverted and swapped positions; *trnC* has moved to a position upstream of *nad1*; *trnV* has inverted and moved to the *nad2–cox1* gene junction; *trnQ* has swapped position with *trnM*; and *trnL2* and *trnW* have inverted and changed positions, with inverted *trnI* forming a novel tRNA gene cluster *trnL1–trnI–trnL2–trnW–trnM–trnQ* between *nad1* and *nad2*. The protein-coding gene *nad2* is inverted in *D. liui*, as was reported for *Ibalia leucospoides* (Cynipoidea: Ibaliidae) (*Mao, Gibson & Dowton, 2015*). Gene rearrangement was also found in rRNA genes in the *D. liui* mitogenome. The two rRNA genes have inverted and moved into the *cob–nad1* junction, flanked by the control region and an inverted, duplicated control region. Inverted rRNA genes have been reported in other insect orders, the Psocodea and Thysanoptera (*Shao & Barker, 2003*; *Covacin et al., 2006*; *Cameron, Johnson & Whiting, 2007*; *Cameron et al., 2011*). However, rearrangements of the rRNA genes are rare in Hymenoptera, only occurring in one species, *Megalyra* sp. (Megalyroidea) (*Mao, Gibson & Dowton, 2014*; *Tang et al., 2019*), making what is observed in the mitogenome of *D. liui* regarding the rRNA genes particularly interesting.

We also analyzed the mitochondrial gene rearrangements of two other Cynipidae species, *Trichagalma acutissimae* (MN928529, *Xue et al., 2020*) and *Synergus* sp. (MG923514, *Tang et al., 2019*) based on incomplete mitogenome sequences obtained from the Genbank. There are different mitochondrial gene rearrangement events among the three gall wasp species relative to the ancestral hymenopteran (Fig. 3), but they have four obvious common characteristics: *trnE* and *trnF* have inverted and swapped positions; *rrnS* and *rrnL* genes have moved into the *cob–nad1* junction; a novel tRNA gene cluster *trnL1–trnI–trnL2–trnW–trnM–trnQ* has been formed between *nad1* and *nad2*; and *trnV* has inverted and moved to the *nad2–cox1* gene junction. Whether these events are common among Cynipidae species
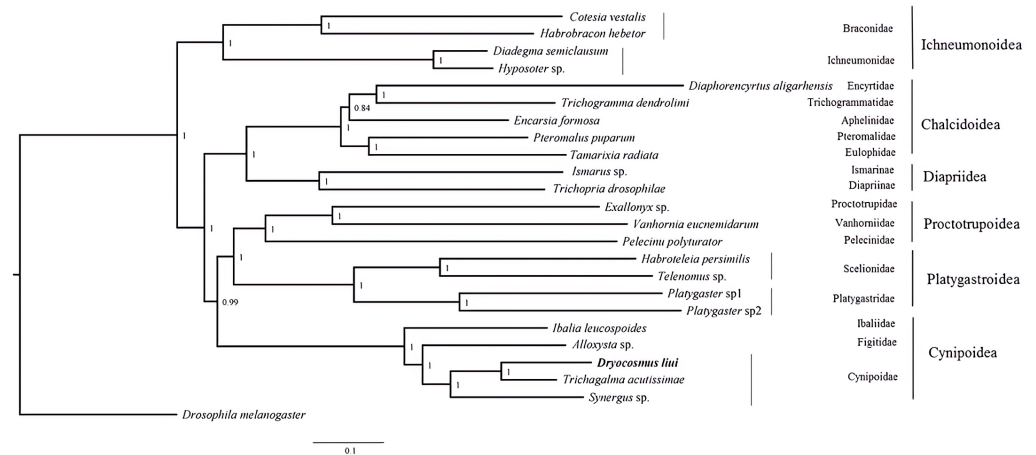
Figure 4 Bayesian analysis of the MAFFT alignment dataset were inferred from the datasets PCG123 +2rRNA. Posterior probabilities are shown at each node.

remains unclear, due to the scarcity of mitogenome sequence data for other species of the family.

Phylogenetic analysis

Phylogenetic analyses were performed using available mitogenome sequences of 19 Proctotrupomorpha. Two datasets (PCG123 + 2rRNA and PCG12 + 2rRNA) were analyzed using the MAFFT alignment approach and three phylogenetic approaches (Bayesian inference, PhyloBayes and Maximum likelihood). Different analytical approaches and datasets did not affect the topology, but did affect the level of node support (Fig. 4 and Figs. S3–S7).

Within the Proctotrupomorpha, most of the recent molecular analyses have generally recovered a sister relationship for Cynipoidea and Platygastroidea (*Castro & Dowton, 2006*; *Dowton & Austin, 2001*; *Heraty et al., 2011*; *Klopfstein et al., 2013*; *Tang et al., 2019*). In this study, the five superfamilies clustered into two branches, namely Chalcidoidea + Diaprioidea and Proctotrupoidea + Platygastroidea + Cynipoidea. It is strongly suggested that Cynipoidea are closely related to Platygastroidea and Proctotrupoidea, with high nodal support (BPP = 1) (Fig. 4). The support values were higher in the analyses using the PCG123 + 2rRNA dataset and BI phylogenetic approach, which may suggest that the inclusion of 3rd codon positions and rRNA genes for Bayesian analysis was more suitable for making inferences regarding this relationship, as discussed by *Tang et al. (2019)*. The relationship within Cynipoidea based on limited mitogenome data is consistent with multiple phylogenetic analyses based on various types of data with more extensive sampling (*Ronquist, 1995*; *Ronquist, 1999*; *Ronquist et al., 2015*; *Liljeblad & Ronquist, 1998*; *Rokas et al., 2003*; *Blaimer et al., 2020*). It would be highly desirable to sample more extensively in future projects to reconstruct Cynipoidea phylogeny based on mitogenome independently and in combination with other types of data.

CONCLUSIONS

In summary, the mitogenome contained the typical gene repertoire of 13 protein-coding genes, two rRNA genes and 22 tRNA genes. Inverted and duplicated control regions have been found which may be advantageous and be maintained either through stabilizing selection or through gene conversion. Rearrangement events of the *D. liui* mitogenome were found to be mainly on the gene clusters between *nad3–nad5*, *nad2–cox1*, and control region–*nad2*, involving PCG, tRNA, and rRNA genes, respectively. It is obvious that ,the mitochondrial genome is a powerful molecular markers for phylogenetic analyses.

Funding

This work was supported by the National Key Research and Development Program of China (No. 2016YFE0128200). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
National Key Research and Development Program of China: 2016YFE0128200.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Cheng-Yuan Su conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Dao-Hong Zhu conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Yoshihisa Abe analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Tatsuya Ide analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Zhiwei Liu analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

Data is available at GenBank: MW368384.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.15865#supplemental-information.

REFERENCES

Blaimer BB, Gotzek D, Brady SG, Buffington ML. 2020. Comprehensive phylogenomic analyses rewrite the evolution of parasitism within cynipoid wasps. *BMC Evolutionary Biology* **20**:155 DOI 10.1186/s12862-020-01716-2.

Boore JL. 1999. Animal mitochondrial genomes. *Nucleic Acids Research* **27**:1767–1780 DOI 10.1093/nar/27.8.1767.

Bosio G, Gerbaudo C, Piazza E. 2010. Dryocosmus kuriphilus Yasumatsu: an outline seven years after the first report in Piedmont (Italy). *Acta Horticulturae* **866**:341–348 DOI 10.17660/ActaHortic.2010.866.43.

Buffington ML, Morita SI. 2009. Not all oak gall wasps gall oaks: the description of *Dryocosmus rileypokei*, a new, apostate species of Cynipini from California. *Proceedings of the Entomological Society of Washington* **111**:244–253 DOI 10.4289/0013-8797-111.1.244.

Cameron SL. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annual Review of Entomology* **59**:95–117 DOI 10.1146/annurev-ento-011613-162007.

Cameron SL, Johnson KP, Whiting MF. 2007. 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* **65**:589–604 DOI 10.1007/s00239-007-9042-8.

Cameron SL, Yoshizawa K, Mizukoshi A, Whiting MF, Johnson KP. 2011. Mitochondrial genome deletions and minicircles are common in lice (Insecta: Phthiraptera). *BMC Genomics* **12**:394 DOI 10.1186/1471-2164-12-394.

Castro LR, Dowton M. 2006. Molecular analyses of the Apocrita (Insecta: Hymenoptera) suggest that the Chalcidoidea are sister to the diaprioid complex. *Invertebrate Systematics* **20**:603–614 DOI 10.1071/IS06002.

Chen L, Chen PY, Xue XF, Hua HQ, Li YX, Zhang F, Wei SJ. 2018. Extensive gene rearrangements in the mitochondrial genomes of two egg parasitoids, Trichogramma japonicum and *Trichogramma ostriniae* (Hymenoptera: Chalcidoidea: Trichogrammatidae). *Scientific Reports* **8**:7034 DOI 10.1038/s41598-018-25338-3.

Cook CE. 2005. The complete mitochondrial genome of the stomatopod crustacean *Squilla mantis*. *BMC Genomics* **6**:105 DOI 10.1186/1471-2164-6-105.

Covacin C, Shao R, Cameron S, Barker SC. 2006. Extraordinary number of gene rearrangements in the mitochondrial genomes of lice (Phthiraptera: Insecta). *Insect Molecular Biology* **15**:63–68 DOI 10.1111/j.1365-2583.2005.00608.x.

Dowton M, Austin AD. 2001. Simultaneous analysis of 16S, 28S, COI and morphology in the Hymenoptera: Apocrita–evolutionary transitions among parasitic wasps. *Biological Journal of the Linnean Society* **74**:87–111 DOI 10.1006/bijl.2001.0577.

Dowton M, Cameron SL, Dowavic JI, Austin AD, Whiting MF. 2009. Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral. *Molecular Biology and Evolution* **26**:1607–1617 DOI 10.1093/molbev/msp072.

Eberhard JR, Wright TF, Bermingham E. 2001. Duplication and concerted evolution of the mitochondrial control region in the parrot genus Amazona. *Molecular Biology and Evolution* **18**:1330–1342 DOI 10.1093/oxfordjournals.molbev.a003917.

Feng ZB, Wu YF, Yang C, Gu XH, Wilson JJ, Li H, Cai WZ, Yang HL, Song F. 2020. Evolution of tRNA gene rearrangement in the mitochondrial genome of ichneumonoid wasps (Hymenoptera: Ichneumonoidea). *International Journal of Biological Macromolecules* **164**:540–547 DOI 10.1016/j.ijbiomac.2020.07.149.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**:294–299.

Heraty J, Ronquist F, Carpenter JM, Hawks D, Schulmeister S, Dowling AP, Murray D, Munro J, Wheeler WC, Schiff N, Sharkey M. 2011. Evolution of the hymenopteran megaradiation. *Molecular Phylogenetics and Evolution* **60**:73–88 DOI 10.1016/j.ympev.2011.04.003.

Kahnt B, Gerth M, Paxton RJ, Bleidorn C, Husemann M. 2015. The complete mitochondrial genome of the endemic and highly specialized South African bee species Rediviva intermixta (Hymenoptera: Melittidae), with a comparison with other bee mitogenomes. *Biological Journal of the Linnean Society* **116**:940–953 DOI 10.1111/bij.12627.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**:772–780 DOI 10.1093/molbev/mst010.

Klopfstein S, Vilhelmsen L, Heraty JM, Sharkey M, Ronquist F. 2013. The hymenopteran tree of life: evidence from protein–coding genes and objectively aligned ribosomal data. *PLOS ONE* **8**:e69344 DOI 10.1371/journal.pone.0069344.

Liljeblad J, Ronquist F. 1998. A phylogenetic analysis of higher–level gall wasp relationships (Hymenoptera: Cynipidae). *Systematic Entomology* **23**:229–252 DOI 10.1046/j.1365-3113.1998.00053.x.

Mao M, Gibson T, Dowton M. 2014. Evolutionary dynamics of the mitochondrial genome in the evaniomorpha (hymenoptera)—a group with an intermediate rate of gene rearrangement. *Genome Biology and Evolution* **6**:1862–1874 DOI 10.1093/gbe/evu145.

Mao M, Gibson T, Dowton M. 2015. Higher-level phylogeny of the Hymenoptera inferred from mitochondrial genomes. *Molecular Phylogenetics and Evolution* **84**:34–43 DOI 10.1016/j.ympev.2014.12.009.

Mao M, Valerio A, Austin AD, Dowton M, Johnson NF. 2012. The first mitochondrial genome for the wasp superfamily Platygastroidea: the egg parasitoid *Trissolcus basalis*. *Genome* **55**:194–204 DOI 10.1139/g2012-005.

Martinez-Sañudo I, Mazzon L, Simonato M, Avtzis D, Pujade-Villar J, Faccoli M. 2019. Tracking the origin and dispersal of the Asian chestnut gall wasp *Dryocosmus kuriphilus* Yasumatsu (Hymenoptera, Cynipidae) in Europe with molecular markers. *Bulletin of Entomological Research* **109**:300–308 DOI 10.1017/S000748531800069X.

Miller M, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: *Gateway computing environments workshop (GCE)*. New Orleans, LA, USA, 1–8 DOI 10.1109/GCE.2010.5676129.

Pang Y, Su CY, Zhu DH. 2018. Discovery of a new species of Dryocosmus Giraud, 1859 (Hymenoptera: Cynipidae: Cynipini) on Castanopsis tibetana from China. *Zootaxa* **4403**:099–110 DOI 10.11646/zootaxa.4403.1.5.

Rieske LK. 2007. Success of an exotic gallmaker, Dryocosmus kuriphilus, onchestnut in the USA: a historical account. *EPPO Bulletin* **37**:172–174 DOI 10.1111/j.1365-2338.2007.01100.x.

Rokas A, Melika G, Abe Y, Nieves-Aldrey JL, Cook JM, Stone GN. 2003. Lifecycle closure, lineage sorting, and hybridization revealed in a phylogenetic analysis of European oak gallwasps (Hymenoptera: Cynipidae: Cynipini) using mitochondrial sequence data. *Molecular Phylogenetics and Evolution* **26**:36–45 DOI 10.1016/s1055-7903(02)00329-9.

Ronquist F. 1995. Phylogeny and early evolution of the Cynipoidea (Hymenoptera). *Systematic Entomology* **20**:309–335 DOI 10.1111/j.1365-3113.1995.tb00099.

Ronquist F. 1999. Phylogeny, classification and evolution of the Cynipoidea. *Zoologica Scripta* **28**:139–164 DOI 10.1046/j.1463-6409.1999.00022.x.

Ronquist F, Nieves-Aldrey JL, Buffington ML, Liu ZW, Liljeblad J, Nylander JAA. 2015. Phylogeny, evolution and classification of gall wasps: the plot thickens. *PLOS ONE* **10**:e0123301 DOI 10.1371/journal.pone.0123301.

Sankoff D, Leduc G, Antoine N, Paquin B, Lang BF, Cedergren R. 1992. Gene order comparisons for phylogenetic inference: evolution of the mitochondrial genome. *Proceedings of the National Academy of Sciences of the United States of America* **89**:6575–6579 DOI 10.1073/pnas.89.14.6575.

Shao R, Barker SC. 2003. 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* **20**:362–370 DOI 10.1093/molbev/msg045.

Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT. 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics* **37**:545–579 DOI 10.1146/annurev.ecolsys.37.091305.110018.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of The Entomological Society of America* **87**:651–701 DOI 10.1093/aesa/87.6.651.

Tang P, Zhu JC, Zheng BY, Wei SJ, Sharkey M, Chen XX, Vogler AP. 2019. Mitochondrial phylogenomics of the Hymenoptera. *Molecular Phylogenetics and Evolution* **131**:8–18 DOI 10.1016/j.ympev.2018.10.040.

Tyagi K, Chakraborty R, Cameron SL, Sweet AD, Chandra K, Kumar V. 2020. Rearrangement and evolution of mitochondrial genomes in Thysanoptera (Insecta). *Scientific Reports* **10**:695 DOI 10.1038/s41598-020-57705-4.

Wei SJ, Li Q, Achterberg K, Chen XX. 2014. Two mitochondrial genomes from the families Bethylidae and Mutillidae: independent rearrangement of protein–coding genes and higher–level phylogeny of the Hymenoptera. *Molecular Phylogenetics and Evolution* **77**:1–10 DOI 10.1016/j.ympev.2014.03.023.

Wei SJ, Tang P, Zheng LH, Shi M, Chen XX. 2010. The complete mitochondrial genome of *Evania appendigaster* (Hymenoptera: Evaniidae) has low A + T content and a long intergenic spacer between atp8 and atp6. *Molecular Biology Reports* **37**:1931–1942 DOI 10.1007/s11033-009-9640-1.

Xue S, Zhang YC, Gao SS, Lu SH, Wang JS, Zhang KP. 2020. Mitochondrial genome of *Trichagalma acutissimae* (Hymenoptera: Cynipoidea: Cynipidae) and phylogenetic analysis. *Mitochondrial DNA Part B-Resources* **5**:1073–1074 DOI 10.1080/23802359.2020.1721366.

Yoshizawa K, Johnson KP, Sweet AD, Yao I, Ferreira RL, Cameron SL. 2018. Mitochondrial phylogenomics and genome rearrangements in the barklice (Insecta: Psocodea). *Molecular Phylogenetics and Evolution* **119**:118–127 DOI 10.1016/j.ympev.2017.10.014.

Zhu DH, He YY, Fan YS, Ma MY, Peng DL. 2007. Negative evidence of parthenogenesis induction by *Wolbachia* in a gallwasp species, Dryocosmus kuriphilus. *Entomologia Experimentalis Et Applicata* **124**:279–284 DOI 10.1111/j.1570-7458.2007.00578.x.

Zhu DH, Liu ZW, Lu PF, Yang XH, Su CY, Liu P. 2015. New gall wasp species attacking chestnut trees: *Dryocosmus zhuili* n. sp. (Hymenoptera: Cynipidae) on *Castanea henryi* from Southeastern China. *Journal of Insect Science* **15**:156 DOI 10.1093/jisesa/iev118.