The complete mitochondrial genome of *Platygaster robiniae* (Hymenoptera: Platygastridae): A novel tRNA secondary structure, gene rearrangements and phylogenetic implications

Lan Huang, Hui-Quan Sun, Cheng-Jin Li, Wen-Xi Zhao, Yan-Xia Yao

Key Laboratory of Forest Protection of National Forestry and Grassland Administration, Ecology and Nature Conservation Institute, Chinese Academy of Forestry, Beijing, 100091, People’s Republic of China

**A R T I C L E I N F O**

Keywords:
- *Platygaster robiniae*
- Mitochondrial genome
- Gene rearrangement

**A B S T R A C T**

*Platygaster robiniae* is economically important as a highly specific parasitoid of the invasive pest *Obolodiplosis robiniae* which was introduced into the Euro-Asia region in the last decade. Despite being a critical and specific parasitoid of the invasive pest *O. robiniae* and its use as an effective biocontrol agent, the absence of sequence information from *P. robiniae* have limited its genetic applications for pest management in forests. Mitochondrial (mt) genomes generally contain abundant nucleotide information and thus are helpful for understanding species history. Here, we sequenced the complete mt genome of *P. robiniae* using next generation sequencing, and annotated 13 protein-coding, 22 tRNA, and 2 rRNA genes and a 702 bp noncoding region. Comparative analysis indicated that this mt genome has a normal A+T content and codons use, however possessed both the expected and unique rearrangements. Ten tRNAs at four gene blocks COII-ATP8, COIII-ND3, ND3-ND5 and the A + T-rich region-ND2 were rearranged, including gene shuffles, transpositions and inversions. Notably, two genes rRNA*tor* (CO2N) and rRNA*tor(tguN)* had undergone long-range inversions, which is the first record of this rearrangement type in the superfamily Platygastridea. The D-loops of both rRNA*tor* and rRNA*tor(tguN)* were absent from the tRNA secondary structure, which has not been reported from hymenopteran previously. Phylogenetic analysis based with the maximum likelihood and Bayesian methods showed that *P. robiniae* grouped with other species of Platygastridae, and that the superfamily Platygastridea is sister to the other Proctotrupomorpha superfamilies. Our tree strongly supports the monophyly of the five superfamilies of Proctotrupomorpha. This study discovered some unique characters of *P. robiniae*, and contributes to our understanding of genome rearrangements in the order Hymenoptera.

1. Introduction

Platygastridae (Apocrita: Platygastridea) is a diverse and speciose family of parasitic Hymenoptera, consisting of approximately 1153 species in the world (Samin and Asgari, 2012). They are generally small in size (0.5–12 mm), with most species being morphologically simple compared with other parasitic wasps (Austin et al., 2004). *Platygaster robiniae* Buhl and Duso (Hymenoptera: Platygastridae) is an egg-larvae parasitoid of the black holodiploid gall midge *Obolodiplosis robiniae* (Diptera: Cecidomyiidae) (Kim et al., 2011; Yang et al., 2021) which is native to North America and has been considered a highly invasive pest insect in Europe and Asia in recent decades (Yao et al., 2015, 2020). It has been discovered in all places where its host has been found in both native areas and new regions, as first reported in 2010 in Qinhuangdao city, Hebei Province, China (Lu et al., 2010), and it was later found in 17 other provinces at 29 sites (Yang et al., 2019). Despite being a critical and specific parasitoid of the invasive pest *O. robiniae* and its use as an effective biocontrol agent, the absence of nucleotide information, population genetics and the phylogeny of *P. robiniae* have reduced the understanding of the history of its occurrence and its mechanism of population colonization and successful invasion, consequently, limiting genetic applications pest management in forests.

Mitochondrial (mt) genome sequences generally provide large and diverse datasets that contain abundant nucleotide information and thus are helpful for improving phylogenetic relationships at any taxon level (Cameron et al., 2006a; Fenn et al., 2008). Additionally, it has been

* Corresponding author.
E-mail address: yaoyx@caf.ac.cn (Y.-X. Yao).

https://doi.org/10.1016/j.ijppaw.2022.06.007

Received 13 May 2022; Received in revised form 14 June 2022; Accepted 20 June 2022

Available online 23 June 2022

2213-2244/© 2022 The Authors. Published by Elsevier Ltd on behalf of Australian Society for Parasitology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
considered a useful molecular marker for species identification and evolutionary studies because genome its features of rare recombination, maternal inheritance, conserved gene component, and high AT composition (Boore and Brown, 1999; Curole and Kocher, 1999). In insects, mt genomes are typically double-stranded circular molecules of approximately 16 kb that contain 13 protein-coding genes (PCGs), 2 ribosomal RNA (rRNA) genes, 22 transfer RNA (tRNA) genes (Boore, 1999) and contain a large noncoding region known as the A + T-rich region (in vertebrates) or control region (in invertebrates), that regulates transcription and replication (Zhang and Hewitt, 1997). With the ongoing developments of Illumina sequencing technology and bioinformatic approaches, sequence data for insect mt genomes has rapidly increased in recent years with 1000's available on NCBI databases including all orders. They are relatively conservative, most insect orders have single gene rearrangements, such as Lepidoptera and Coleoptera (Sheffield et al., 2008; Sun et al., 2020), but there are also frequent mt genome rearrangements in other orders, such as Hymenoptera (Dowton and Austin, 1999; Dowton et al., 2009b; Mao and Dowton, 2014) and Psocodea. The number of genes involved in hymenopteran mitochondrial rearrangement is large and rearrangements often independent, which has made it sometimes difficult to sequence complete hymenopteran mt genomes, resulting in fragment deletion (Mao et al., 2015).

While rearrangements of insect mt genomes have now been found in 17 orders of insects (Crozier and Crozier, 1993; Flook et al., 1995; Shao and Barker, 2003; Wang et al., 2014; Wei, 2009), the frequency, types and scales of rearrangements often differ between taxa (Chen and Du, 2016). The Hymenoptera especially the suborder Apocrita, exhibit high mt rearrangement rates, and most taxa are rearranged. The rearrangements mainly occur in tRNA genes (Wei, 2009; Wei and Chen, 2011; Wei et al., 2010b). Four types of rearrangements, translocation, inversion, shuffling and remote inversion events, are primary for hymenopterans and have been found to be present at nearly equal frequencies (Dowton and Austin, 1999; Wei, 2009). However, this phenomenon needs to be explained and may be associated with a deeper mechanism, which could be possible when we have generous species nucleotide data.

Poor representation among Hymenopteran lineages, however, has limited the application of the mt genome in evolutionary analysis, especially in Proctotrupomorpha (encompassing the superfamilies Proctotrupoidea, Cynipoidea, Diaprioidae, Mymarommatidae, Platygastroidea, and Chalcidoidea) (Mao et al., 2015; Shen et al., 2019). In total, 79 mt genomes from species of Proctotrupoidea have been reported through NCBI database search statistics, including Chalcidoidea, Proctotrupoidea, Diaprioidae, Cynipoidea and Platygastroidea which have 52, 5, 4, 7 and 11 mt genomes, respectively, yet, no complete mt genome sequences were available from Platygastroidea.

In the present study, we sequenced and annotated the mt genomes of *P. robiniae* using Illumina TruSeq and bioinformatics approaches, compared the structure of the new mt genomes with that of closely related groups, and explored rearrangement genes of Platygastroidea. Additionally, we conducted phylogenetic analyses of mt genomes within the Proctotrupoidea.

2. Materials and methods

2.1. Sampling and DNA extraction

Individuals of *P. robiniae* were collected from *Robinia pseudoacacia* L. forests in four cities in China (BJ, YC, YT, SY) in July 2017 (Table 1). Samples were preserved in 100% ethanol at −20 °C for long-term storage at the Chinese Academy of Forest (CAF). Total genomic DNA was extracted from individuals using a DNeasy tissue kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions.

2.2. Sequencing and genome assembly of mitochondrial DNA

Illumina TruSeq libraries with an average insert size of 450 bp were constructed using a TruSeq™ DNA LT Sample Prep Kit (Illumina, Inc., San Diego, CA, USA) following the manufacturer’s manuals. Clustering of the index-coded samples was performed in a cBot Cluster Generation System using a TruSeq PE Cluster Kit v3-cBot-HS (Illumina). Sequencing of the clustered flow cell was performed using the Illumina HiSeq 2500 platform (Erik et al., 2011). Prior to assembly, Illumina raw reads were filtered firstly. For each library, 6 Gb of clean data was obtained after trimming adapters and low-quality bases (quality score <20) using Adapter Removal v2 (Schubert et al., 2016). The genome was assembled using Geneious Prime 2020 (https://www.geneious.com/prime) and IDBA-UD assembler software (Peng et al., 2012). Geneious prime 2020 was used de novo assembly, preprocessed the NGS reads properly, and pruned low quality data using BBduk, then paired the trimmed data to read lists and reassemble them. IDBA-UD iterated the value of k from kmin to kmak, and gradually increased the threshold of low-depth cutting, then remove some low-depth overlapping groups and obtain longer Hk confidence overlapping groups (Ck). The missing k-mers are reconstructed by locally assembling, and the information of these missing k-mer will be transmitted to the next iteration through these overlapping groups (LCk). Finally, the overlapping groups of all outputs are used to form scaffolds by pairing the terminal read length information.

2.3. Mitochondrial genome annotation and analysis

The positions and direction of 13 PCGs, 2 rRNA genes, and 22 tRNA genes were predicted using MITOS WebServer (Matthias et al., 2013) with the following parameters: Reference = “ReSeq 89 Metazoa” and Genetic Code = “5 Invertebrate”. The secondary structures of the RNA genes were also determined using the MITOS WebServer using default settings (Tang et al., 2017). When they were not detected by this approach, we confirmed the tRNA positions by aligning with their homologous sequences from related species (GenBank:M923507, M923510, KF696669, KF696670, JN903532) (Table S1). The protein-coding and rRNA genes were initially annotated with MITOS WebServer and edited in Geneious 9.0.2 (http://www.geneious.com) through comparison to other Platygastroidea mitochondrial genomes (Table S1). The start and stop codons and length of each PCG were manually confirmed and modified. All reference mt genomes were downloaded from GenBank.

Nucleotide composition and relative synonymous codon usage (RSCU) were determined using MEGA 7.0 software (Sudhir et al., 2016). AT and GC skews were measured for the major (J) strand of each genome, using the formulae AT-skew = (A-T)/(A + T) and GC-skew = (G-C)/(G + C) (Perna and Kocher, 1995). A circular map of the complete mt genome was made using CGView (Grant and Stothard, 2008).

DnaSP6 (Rozas et al., 2003) was used to calculate the ratio of the nonsynonymous substitution rate to the synonymous substitution rate (Ka/Ks), and evolution rates for the 13 PCGs in *P. robiniae*, the evolution rates within major groups in Platygastroidea and the evolutionary rates of each mitochondrial gene.

2.4. Phylogenetic analysis

Data from the newly sequenced mt genome of *P. robiniae* and those of 48 other Proctotrupomorpha were used for phylogenetic analysis with one species from the family Ichneumonoidea (Insecta, Hymenoptera) as an outgroup (Table S1). Nucleotide sequences from each PCG and rRNA

---

**Table 1** Sequencing sample acquisition information.

| Place   | Abbreviation | Time     | Longitude and Latitude | Altitude |
|---------|--------------|----------|------------------------|----------|
| Beijing | BJ           | July 2017| 40°36'28"N,116°58'01" | 89       |
| Yinchuan| YC           | July 2017| 38°43'39"N,106°17'73" | 1076     |
| Yantai  | YT           | July 2017| 37°56'65"N,121°24'97" | 18       |
| Shenyang| SY           | July 2017| 41°77'70"N,123°43'40" | 55       |
gene were aligned individually using the MAFFT Web Server (Katoh et al., 2005), and ambiguous sites deleted manually after alignment. The dataset of 13 PCGs was used to construct phylogenetic trees. Maximum likelihood and Bayesian approaches were employed to infer phylogenetic trees. Analyses were performed using MrBayes v.3.2.5 (Ronquist and Huelsenbeck, 2003) and PhyML 3.0 (Gascuel, 2010). MrBayes v.3.2.5 was used to analyze the dataset for nucleotide substitutions with the GTR$+I+G$ model, which was selected using jModelTest 2.1.7 (David, 2008). For maximum likelihood analyses, 1000 bootstrap replicates were performed and the GTRGAMMA substitution model applied to all partitions. For Bayesian analysis, two simultaneous runs of 10,000,000 generations were conducted, sampled every 200 generations with a burn-in of 25%. Phylogenetic trees were viewed and edited in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

3. Results and discussion

3.1. Structure of the mitochondrial genome

3.1.1. Mitochondrial genome size

Two identical complete mt genomes for each sample were obtained using both Geneious Prime 2020 and IDBA-UD assembler software, and no different structures were caused by the different assembling methods. By comparing mt genome sequences from the four sites, it was found that mt genomes of the samples from YC, YT and SY were completely consistent in sequence, while the BJ sample has 1 bp more in the tRNA$^{Trp}$ gene and showed 42 base substitutions in 14 genes, control region and intergenic spacers comparing to the other three samples (Table S2). Since three of the four samples shared the same mt genomes sequence with a length of 15,348 bp, it was used to subsequently analysis and has been assigned the GenBank accession number (GenBank:OM372674).

Currently, for the available mt genomes of the Proctotrupomorpha, the length ranged from 10044 bp to 18217 bp. The length of the new sequence is within the range and contains all fragments, which is a complete mitochondrial sequence (Table S3). Altogether, the genome comprises 37 genes (22 tRNA genes, 13 PCGs (COI–III, ND1–6, ND4L, CytB, ATP6 and ATP8), 2 tRNA genes (tRNA$r$ and s$\text{rRNA}$)) and a control region (Fig. 1). Some mt genome data lacked tRNAs and protein coding genes, because in the process of gene sequencing, some tRNAs (tRNA$Ile$, tRNA$Arg$, etc.) were not sequenced or annotated successfully due to serious rearrangement, and their positions could not be determined, leading to gene fragment deletion (Samin and Asgari, 2012; Shen et al., 2019). As a result, some mt genomes had sufficient length, but the fragment number was less than 37, which is a common problem in sequencing hymenopteran mt genomes (Castro et al., 2006; Dowton, 1999; Wei et al., 2010b). Sixteen of these genes are encoded on the

Table 2

| Gene       | Length (bp) | A%   | T%   | AT% | AT-skew | C%  | G%  | CG% | GC-skew |
|------------|-------------|------|------|-----|---------|-----|-----|-----|---------|
| All gene   | 15348/15349 | 44.0/43.9 | 38.0 | 81.9 | 0.0732  | 11.5 | 6.6 | 18.1 | -0.2744/-0.2710 |
| 13-PCG     | 11151       | 43.1 | 36.4 | 79.5 | 0.0847/0.0850 | 12.9 | 7.5/7.6 | 20.5 | -0.2641/-0.2630 |
| tRNA       | 1983        | 47.3 | 41.2 | 88.5 | 0.0689  | 8.1/8.0 | 3.4/3.5 | 11.5 | -0.4053/-0.3860 |
| Control region | 702  | 42.9/42.6 | 41.5/41.3 | 84.3/83.0 | 0.0169/0.0154 | 9.8/10.0 | 5.8/6.1 | 15.7/16.1 | -0.2545/-0.2389 |

Fig. 1. Genetic map of the complete mitochondrial genome of Platygaster robiniae.
Notes: the blue arrow represents the direction of gene transcription; the black peak represents the deviation of GC%; the purple and green peaks represent the deviation in GC skew; green refers to positive skew, and purple indicates negative skew. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
Table 3
Characteristics of PCGs of mitochondrial genomes of 12 species of Platygasteroida.

| Species               | Length (bp) | A%  | T%  | AT% | AT-skew | C%  | G%  | CG% | GC-skew |
|-----------------------|-------------|-----|-----|-----|---------|-----|-----|-----|---------|
| *Platygaster robiniae*| 15348/15349 | 43.1| 36.4| 79.5| 0.0847/0.0850 | 12.9| 7.5/7.6 | 20.5| -0.2641/-0.2630 |
| *Platygaster sp. ZJUH 2016026* | 16098 | 42.3 | 40.0 | 82.3 | 0.0279 | 11.7 | 5.9 | 17.7 | -0.3310 |
| *Platygaster sp. ZJUH 2016029* | 16605 | 43.0 | 39.2 | 82.2 | 0.0468 | 11.4 | 6.4 | 17.8 | -0.2787 |
| *Ceratobaeus sp. MM-2013* | 15851 | 40.6 | 34.9 | 75.6 | 0.0757 | 15.6 | 8.9 | 24.4 | -0.2751 |
| *Habroteleia persimilis* | 17186 | 43.0 | 41.3 | 84.2 | 0.0197 | 9.8 | 6.0 | 15.8 | -0.2440 |
| *Idsa sp. MM-2013* | 15137 | 41.3 | 37.8 | 79.1 | 0.0443 | 13.6 | 7.4 | 20.9 | -0.2951 |
| *Scelio sp. ZJUH 2016028* | 16851 | 40.8 | 37.9 | 78.7 | 0.0388 | 14.4 | 6.8 | 21.3 | -0.3586 |
| *Telenomus dignus* | 14304 | 43.2 | 39.5 | 82.7 | 0.0449 | 10.2 | 7.0 | 17.3 | -0.1839 |
| *Telenomus remus* | 16014 | 43.8 | 39.1 | 83.0 | 0.0570 | 10.0 | 7.1 | 17.0 | -0.1711 |
| *Telenomus sp. ZCS-2018* | 17023 | 43.1 | 38.3 | 81.4 | 0.0586 | 11.5 | 7.1 | 18.6 | -0.2367 |
| *Trissolcus basalis* | 15768 | 43.2 | 39.2 | 82.2 | 0.0472 | 11.1 | 6.7 | 17.8 | -0.2500 |
| *Trissolcus japonicus strain CREATEJ* | 16564 | 42.6 | 38.2 | 80.8 | 0.0548 | 12.1 | 7.1 | 19.2 | -0.2616 |

Table 4
Mitochondrial genome structure of *Platygaster robiniae*.

| Gene          | Direction | Location | Length (bp) | Codon Start | Codon Stop | Intergenic Nucleotides† |
|---------------|-----------|----------|-------------|-------------|------------|------------------------|
| tRNA*leu(UUR)| R         | 1–68     | 68          |             |            |                        |
| tRNA*Ile     | R         | 75–136   | 62          |             |            |                        |
| tRNA*Gln     | R         | 135–201  | 67          |             |            |                        |
| Control region| F         | 202–903  | 702         |             |            |                        |
| ND2           | F         | 904–1900 | 997         | T-          | 0          |                        |
| rRNA*16S     | F         | 1901–1965/1901–1966 | 65/66 |            |            |                        |
| rRNA*12S     | R         | 1958–2022/1959–2023 | 65 |            |            |                        |
| rRNA*18S     | R         | 2029–2091| 63          |             |            |                        |
| COI           | F         | 2096–3628| 1533        | ATG         | TAA        | 4                      |
| tRNA*Glu     | F         | 3635–3699| 65          |             |            |                        |
| tRNA*Glu     | F         | 3702–3762| 61          |             |            |                        |
| COII          | F         | 3763–4435| 673         | ADA         | T-          | 0                      |
| rRNA*Ser     | F         | 4436–4502| 67          |             |            |                        |
| rRNA*Val     | F         | 4501–4567| 67          |             |            | -2                     |
| ATP6         | F         | 4568–4735| 168         | ATA         | TAA        | 0                      |
| ATP8         | F         | 4729–5391| 663         | ATG         | TAA        | -7                     |
| COII         | F         | 5391–6176| 786         | ATG         | TAA        | -1                     |
| rRNA*Gln     | F         | 6175–6289| 65          |             |            | -2                     |
| rRNA*Glu     | F         | 6246–6314| 69          |             |            |                        |
| tRNA*Glu     | F         | 6315–6377| 63          |             |            |                        |
| tRNA*Thr     | F         | 6378–6435| 58          |             |            |                        |
| ND3          | F         | 6442–6801| 360         | ATT         | TAA        | 6                      |
| tRNA*Arg     | R         | 6800–6864| 65          |             |            | -2                     |
| tRNA*Glu     | F         | 6879–6937| 59          |             |            | 14                     |
| tRNA*Asn     | F         | 6956–7020| 65          |             |            | 18                     |
| tRNA*Pro     | F         | 7020–7084| 65          |             |            | -1                     |
| NDS          | R         | 7085–8764| 1680        | ATT         | TAG        | 0                      |
| rRNA*Met     | F         | 8765–8830| 66          |             |            | 0                      |
| ND4           | F         | 8831–10172| 1342       | ATG         | T-          | 0                      |
| ND4L          | F         | 10166–10450| 285 |            | ATA         | TAG        | -7                     |
| rRNA*Leu     | F         | 10453–10516| 64 |            |            |                        |
| rRNA*Phe     | F         | 10517–10582| 66 |            |            |                        |
| NDS          | F         | 10599–11189| 591 | ATA         | TAA        | 16                     |
| CydB         | F         | 11193–12329| 1137 | ATG         | TAA        | 3                      |
| ND1          | R         | 12367–13302| 936 | ATA         | TAA        | 37                     |
| hrRNA         | R         | 13303–14548| 1246 |            |            | 0                      |
| rRNA*Ile     | R         | 14549–14611| 63 |            |            | 0                      |
| rRNA*Ser     | R         | 14612–15348| 737 |            |            | 0                      |

† Represents the gene interval, and the negative number represents the number of nucleotides overlapped between adjacent genes.

minor strand (N-strand), including four PCGs (ND1, ND4, ND4L, and ND5), ten tRNA genes (tRNA*Glu, tRNA*Ile, tRNA*Leu, tRNA*Met, tRNA*Gln, tRNA*His, tRNA*Phe, tRNA*Ser (UCN), tRNA*Thr, tRNA*Val), and two rRNA genes (tRNA* and sRNA), whereas the remaining 21 genes are encoded on the major strand (J-strand) in *P. robiniae*.

3.1.2. Nucleotide composition

Mitochondrial genomes are generally characterized by significant nucleotide compositional bias (Cameron, 2014; Timmermans and Vogler, 2012), and two measures of bias, non-strand specific (A + T and G + C contents) and strand specific (AT-skew and GC-skew), are used to examine its extent (Hassanin, 2006; Wei et al., 2010a). We found the *P. robiniae* mt genome to be characterized by very high A + T content (Table 2), accounting for 81.9% of the genome. High A + T content is due to the increased A content in Apocrita mt genomes (Dowton and Austin, 1997) but is common in other Hymenoptera mt genomes (Dowton, 1999; Mao and Dowton, 2014; Samin and Asgari, 2012; Shen et al., 2019). The A + T content of the control region was higher (84.3%/83%) than that of the coding regions (79.5%), which is the general pattern in insect mt genomes (Clary and Wolstenholme, 1985; Zhang and Hewitt, 1997). The *P. robiniae* mt genome had an overall positive AT-skew and negative GC-skew (Table 2), indicating no reversal
of strand asymmetry within this species although reversals have been shown in other hymenopterans (Wei et al., 2010a). Comparison of coding regions in *P. robiniae* to the other 11 species in Platygastroidea, finds that base composition of the coding region has strong AT bias (75.6%-84.2%), and each of the 12 coding regions exhibit positive AT skews (ranging from 0.0197 to 0.0847/0.0850), and negative GC-skews (ranging from -0.1711 to -0.3586) (Table 3).

### 3.1.3. Control region, intergenic spacer and overlap

The control region of insect mt genomes can show considerable variation in length (Mao and Dowton, 2014; Shen et al., 2019). Due to rearrangement of the *P. robiniae* mt genome, the control region is located between *tRNA^Gln* and *ND2*, and full length is 702 bp. The A + T content (84.3%/83.0%) of the region is higher than that of the coding region (84.2%), and each of the 12 coding regions exhibit positive AT skews (ranging from 0.0197 to 0.0847/0.0850), and negative GC-skews (ranging from -0.1711 to -0.3586) (Table 3).

### 3.1.4. Codon usage

The RSCU in the mt genome of *P. robiniae* shows a strong bias toward the usage of A and T, particularly at the third codon position. The most frequently used codon for each amino acid is NNA or NNU (Table 5). For some amino acids, the most frequently used codon is not the set that corresponds strictly to the corresponding tRNA anticodon (Table 5). The four most commonly encoded amino acids in the *P. robiniae* mt genome (with their corresponding codons), listed in order of decreasing frequency, are as follows: AAA (Lys), AUU (Ile), UUA (Phe), and AAG (Lys) (Fig. 2).

| Amino acid | Codon | Count | RSCU | Amino acid | Codon | Count | RSCU |
|------------|-------|-------|------|------------|-------|-------|------|
| Phe (F)    | UUU   | 290   | 1.53 | Tyr (Y)    | UAU   | 252   | 1.57 |
|            | UUC   | 90    | 0.47 |            | UAC   | 70    | 0.43 |
| Leu (L2)   | UUA   | 413   | 3.76 | Ser (S2)   | UCU   | 75    | 1.54 |
|            | CUA   | 48    | 0.44 | His (H)    | CAU   | 99    | 1.56 |
| Leu (L1)   | CUU   | 47    | 0.43 |            | CUG   | 84    | 0.44 |
|            | CUC   | 32    | 0.29 |            | CCA   | 28    | 0.44 |
| Ile (I)    | AUU   | 47    | 0.45 | Cys (C)    | UGG   | 8    | 0.44 |
|            | 90    | 1.55   |    |            | AAA   | 144   | 1.55 |
| Met (M)    | AUA   | 382   | 1.78 | Lys (K)    | UCG   | 42    | 0.86 |
|            | AUG   | 47    | 0.22 |            | UGA   | 4    | 0.22 |
| Val (V)    | GGU   | 32    | 1.41 | Asp (D)    | GAG   | 25    | 0.51 |
|            | 12  | 0.53 |    |            | GAC   | 16    | 0.5 |
| GUA  | 43    | 1.89 |    | Glu (E)    | GAA   | 74    | 1.49 |
| Ser (S2)   | UCU   | 75    | 1.54 |            | GAG   | 7    | 0.4 |
|            | 42  | 0.86 |    |            | UGA   | 12    | 0.5 |
| UCA  | 111   | 2.28 |    | Trp (W)    | UGA   | 59    | 1.59 |
| UGC  | 10   | 0.21 |    |            | UGG   | 15    | 0.41 |
| Pro (P)    | CUC   | 28    | 1.05 | Arg (R)    | GGU   | 11    | 1.57 |
|            | 19    | 0.71 |    |            | GCC   | 4    | 0.57 |
| GCA  | 37    | 2.13 |    |            | GCA   | 10    | 1.43 |
| Thr (T)    | ACU   | 88    | 1.68 | Ser (S1)   | AGU   | 44    | 0.9 |
|            | 42    | 0.8 |    |            | AGC   | 20    | 0.41 |
| Ala (A)    | GCC   | 11    | 1.47 | Gly (G)    | GGU   | 12    | 1.23 |
|            | 2    | 0.27 |    |            | GCC   | 2    | 0.21 |
| GCA  | 16    | 2.13 |    |            | GGA   | 22    | 2.26 |
| GCG  | 1    | 0.13 |    |            | GGG   | 3    | 0.31 |

Note: the codons underlined are those that strictly match the tRNA anticodon, and the codons in bold are those used most frequently for each amino acid.

Table 5: Statistics on codon usage of protein gene in *Platygaster robiniae* mitochondrial genome.

### 3.1.5. Protein-coding genes

The location and size of 13 PCGs were determined by comparing the mt genome of *P. robiniae* with its related species. The 13 PCGs accounted for 72.65% (11151 bp total) of the whole genome, and the AT content was 79.5% in *P. robiniae* (Table 2). Nine of the 13 PCGs are encoded by the J-strand and four by the N-strand, and the start codons of all PCGs are “ATN” in these genomes (ATG, ATA, ATT) (Table 4). The three PCGs, ND2, ND3 and ND5, use ATT as start codons, and the remaining 10 PCGs use conventional start codons, ATA or ATG. In many metazoans, numerous mitochondrial genes have incomplete termination codons (Miya et al., 2001). In the *P. robiniae* mt genome, except for ND2, ND4 and COII, use ATT as start codons, and the remaining 10 PCGs use conventional start codons TAA and TAG. The RSCU values in the mt genomes of *P. robiniae* reflect a significant bias toward A and T nucleotides which is commonly found in other species of hymenopterans (Chen et al., 2016) (Table 5).

### 3.1.6. tRNA and rRNA genes

The 22 tRNA genes in the *P. robiniae* mt genome, which is 1418 bp long, ranged from 58 bp (tRNA^{tRNA_{Leu}(CUN)}) to 69 bp (tRNA^{tRNA_{Leu}(CUN)}) in length. The A + T content and skew of the tRNAs were 89.4% and 0.0237,
respectively (Table 2). Most tRNA genes fold into a typical cloverleaf structure except tRNAAsp and tRNASer(AGN), and the two genes with D-stems are absent (Fig. 3). A missing D-stem for tRNAAsp for the Scelionidae (Mao and Dowton, 2014) has been proposed as a shared derived character (Dowton and Campbell, 2001; Poulton et al., 1993). Additionally, tRNAAsp, tRNAVal, and tRNAIle were translocated, tRNAGlu, tRNASer(AGN), and tRNAIle from between ND3 and ND5 to between COIII and ND3 (Fig. 4). Additionally, tRNAHis and tRNAAsp shuffled (switched) positions. These different types of rearrangements may occur in combination, and it is generally assumed that short-distance rearrangements are more frequent than long-distance rearrangements (Chen and Du, 2016; Mao and Dowton, 2014; Shen et al., 2019), and each of these three sets of rearrangements are short range (<1000 bp moves).

In addition, by comparing across all species of Platygastroidea (Fig. 4), we found that the srRNA-A + T-rich region is the commonest rearrangement location but that only in P. robiniae was the A + T-rich adjacent to ND2. In the other 11 species, there was one or more tRNA genes between the A + T-rich region and ND2. This may be due to the recombination, leading to the inversion of the tRNAHis, tRNAIle, and tRNAVal as well as the translocation adjacent exchange of tRNAIle, tRNAGlu, and tRNAHis (Dowton and Campbell, 2001; Poulton et al., 1993). Furthermore, by comparing the mt genome sequences, we found some unique rearrangement characteristics in the mt genome of the family Platygasteridae. The positions of Cyb-ND1 and ND1-irRNA (tRNASer(UCN), tRNALeu(CUN)) were rearranged, but tRNAGln was not rearranged, which has not been discovered in other species of Proctotrupomorph, validating that mitochondrial gene rearrangement in hymenopterans is highly diverse (Dowton et al., 2009a; Wei, 2009).

In addition, inversion is the least common type of rearrangement in insects, including local inversion and remote inversion. Remote inversion is caused by two rearrangements (Chen and Du, 2016; Wei, 2009), but it has been recorded in hymenopterans, and local inversion accounts for one-third of hymenopteran genome rearrangements (Dowton and Austin, 1999; Dowton et al., 2009b). In the P. robiniae mt genome, tRNACys and tRNAHis have moved into the junction between ND3-ND5 and COI-COI, and tRNAIle and tRNALeu(CUN) are inverted simultaneously. This remote inversion of tRNASer(UCN) and tRNALeu(CUN) was first reported in Platygastroidea. The tRNASer(UCN) inversion was found in only a few species in other superfamilies, but tRNALeu(CUN) (Fig. S1B), which is consistent with the characteristics of the insect mt genome (Wei, 2009). Among them, ND4L was positively selected (Ka/Ks > 1), while other genes were purified (Ka/Ks < 1). There were differences in the rate of mitochondrial gene evolution among the different groups. The evolution rate of hymenopteran Symphyta was similar to that of other holomorph insects, while the evolution rate of hymenopteran Apocrita was 2–3 times as fast as that of Symphyta (Dowton et al., 2009a; Wei and Chen, 2011). The comparative analysis of the evolution rate of three subfamilies of Platygastroidea showed that the evolution rate of Telenominae was the fastest, while the evolution rate of Scelioninae was the slowest (Fig. S1A).
Fig. 4. Mitochondrial genome organization of *Platygaster robiniae* and 11 species of Platygastroidea, compared with the ancestral pancrustacean mt genome organization.

Note: tRNA genes are indicated by single letter amino acid codes, L1, L2, S1 and S2 denote tRNA<sub>Leu(CUN)</sub>, tRNA<sub>Leu(UUR)</sub>, tRNA<sub>Ser(AGN)</sub> and tRNA<sub>Ser(UCN)</sub>, respectively. Genes are transcribed from left to right except those indicated by underlining. Gene movements, relative to the ancestral organization, are indicated with arrows.
inversion did not occur, suggesting that tRNA\textsubscript{Leu(CUN)} inversion was probably the special rearrangement in Platygastroidea. For three species sequenced in Platygastridae (Table S1), tRNA\textsubscript{Ser(UCN)} and tRNA\textsubscript{Leu(CUN)} rearrangements were found in P. robiniae and P. sp. ZJUH\_2016026 (Tang et al., 2019), while P. sp. ZJUH\_2016029 only has a rearrangement of the tRNA\textsubscript{Ser(UCN)} gene. The rearrangement of tRNA\textsubscript{Ser(UCN)} may be due to the large gene space between CytB and ND1, which a rearrangement hotspot (Wei, 2009). Parasitic habits have been considered an inducing factor for rearrangement (Dowton et al., 2002; Shao et al., 2001); however, it was later found that the frequency of accelerated rearrangement in hymenopterans is not consistent with the evolution of parasitic habits.

Shared gene rearrangements are considered a valuable source for deducing phylogenetic relationships (Dowton et al., 2002), and some of the rearrangements seen here have been found in other hymenopteran lineages. For example, the shuffling tRNA\textsuperscript{Asp} and tRNA\textsuperscript{Lys} reported here was identical to three known Scelionidae mt genomes (Mao and Dowton, 2014). One study found that unique rearrangements in hymenopterans are so common that only five of the 67 rearrangements identified are shared by two or more species, and only two of the five rearrangements are truly homologous (Dowton et al., 2009b). Our data compared 12 closely related species (Fig. 4), and all 12 species shared rearrangements of tRNA\textsuperscript{Apr} and tRNA\textsuperscript{Lys}, suggesting that the shuffling of tRNA\textsuperscript{Apr} and tRNA\textsuperscript{Lys} is ancestral in the Platygastroidea.

3.3. Phylogenetic analyses

Phylogenetic analyses were performed using 50 mt genomes, 12 of which were from Platygastroidea (Table S1). Diadegma semiclausum (Ichneumonoidea: Ichneumonidae) was chosen as the outgroup. The topologies of the trees generated using the two phylogenetic approaches were identical with strong support at most nodes (posterior probabilities >95% and bootstrap values >70%) (Hillis and Bull, 1993) (Fig. 5). In contrast to previous studies using hymenopteran mt genomes, the inclusion of third codon positions did not have much effect on the topology and nodal support in the current analysis, indicating that inclusion of third codon positions does not appear to be problematic when constructing the phylogeny of closely related taxa (Mao and Dowton, 2014).

The three species of Platygastridae (P. robiniae and two species of Platygastrus sp.) clustered together with strong support (pp = 1). Diaprioidea was supported as the sister group of Chalcidoidea, and this result was identical to that of previous analyses (Castro et al., 2006; Heraty et al., 2011; Mao et al., 2015). However, the placement of Cynipoidea, varied between previous analyses, suggesting Cynipoidea as a sister group to Proctotrupidae plus Diaprioidea plus Chalcidoidea (Heraty et al., 2011; Klopfstein et al., 2013; Tang et al., 2019; Vilhelmsen et al., 2010). Our results indicated that Cynipoidea is sister to Diaprioidea plus Chalcidoidea, in agreement with the results of Mao et al. (2015). Platygastroidea was well supported as monophyletic and sister to the remaining Proctotrupomorpha. Additionally, the relationship between Scelioninae
and Telenominae was shown to be obviously closer than that between Platygastrinae and Scelioninae.

As a tool for examining phylogenetics, mt gene order and tRNA secondary structures can resolve some contentious evolutionary questions (Boore and Brown, 1999; Dowton et al., 2002; Weigert et al., 2015). The gene rearrangements in Scelionidae and Platygastridae were largely different; in addition, each of the three species of *Platygaster* had different gene arrangements. This pattern likely occurs because hymenopterans exhibit a high frequency of gene rearrangement, with the gene order of each family significantly different from that of others (Chen and Du, 2016; Shen et al., 2019). Gene rearrangement events in insects contribute little to the study of phylogenetic relationships between insect orders but may be beneficial to the study of phylogenetic relationships among groups within insect orders (Cameron, 2014; Cameron et al., 2006b). In Platygastridea, the absence of tRNA\(^{\text{Val}}\) rearrangement and tRNA\(^{\text{Ser(UCN)}}\) rearrangement are differences found between Scelionidea and Platygastridea, so we believe that this rearrangement feature can distinguish the two families (Fig. 4). Incomplete mitochondrial genomes of Hymenoptera can provide important sequence information for phylogenetic studies, but in comparison, complete mt genome sequences have more utility. Different branches may share the same or different rearrangements, and research provides more comprehensive information on system development (Chen and Du, 2016; Massimiliano et al., 2014; Shen et al., 2019).

**Supplementary materials**

Fig. S1: (A): Rate of evolution of three subfamilies of Platygastridea; (B): Evolution rate of 13 protein-coding genes in the *Platygaster robiniae* mt genome. Table S1: GenBank accession numbers of published Proctotrupomorpha members and outgroup. Table S2: Comparison of mt genomes of BJ, YC, YT and SY samples. Table S3: 79 mitochondrial genomes and new sequence lengths of Proctotrupomorpha.

**Author contributions**

H.L.: Conceptualization, Methodology, Validation, Formal analysis, Data curation, Software, Visualization, Writing-original draft, Writing-
review & editing. H.Q.: Software, Visualization. C.J.: Resources. W.X.: Project administration, Funding acquisition. Y.X.: Conceptualization, Methodology, Investigation, Resources, Project administration, Funding acquisition, Writing-review & editing.

Declaration of competing interest

The authors have no conflicts of interests to declare.

Acknowledgments

We thank Mr Jun Yang (United Pesticide Industry CropScience Co. Ltd., Jinan) for the help in samplings. This work was supported by the National Key Research and Development Program of China [2016YFC1201200].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijppaw.2022.06.007.

References

Austin, A.D., Johnstone, N.F., Dowton, M.P., 2004. Systematics, evolution, and biology of scelionid and platygastrid wasps. Annu. Rev. Entomol. 50, 553–582.

Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27, 1767–1780.

Cameron, S.L., 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu. Rev. Entomol. 59, 95–117.

Cass, J.R., Cameron, S.L., Dowton, M., 2009. The evolution of strand-specific compositional bias. J. Mol. Evol. 69, 587–598.

Curole, J.P., Kocher, T.D., 1999. Mitogenomics: digging deeper with complete mitochondrial genome sequences. Trends Genet. 15, 345–356.

Dowton, M., Castro, L.R., Austin, A.D., 2002. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome organization and proliferation. Invertebr. Systat. 16, 345–366.

Dowton, M., Castro, L.R., Campbell, N.J., Bargon, S.D., Austin, A.D., 2003. Frequent mitochondrial gene rearrangements at the Hymenopteran nad3–nad5 junction. J. Mol. Evol. 56, 517–526.

Erik, B., Sverker, L., Joakim, L., Nitabach, M.N., 2011. Large scale library generation for high throughput sequencing. PLoS One 6, 231–238.

Fenn, J.D., Song, H., Cameron, S.L., Whiting, M.F., 2008. A preliminary mitochondrial genome phylogeny of Orthoptera (Insects) and approaches to maximizing phylogenetic signal found within mitochondrial genome data. Mol. Phylogenet. Evol. 49, 59–68.

Flock, P.C., Rowell, C.H.F., Gillisien, 1995. The sequence, organization, and evolution of the Locusta migratoria mitochondrial circular genome. J. Mol. Evol. 41, 50–51.

Gascuel, O., 2010. New algorithms and methods to estimate Maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Grant, J.R., Stothard, P., 2008. The CGView Server: a comparative genomics tool for circular genomes. Nucleic Acids Res. 36, 181–194.

Hassanin, A., 2006. Phylogeny of Arthropoda inferred from mitochondrial sequences: strategies for limiting the misleading effects of multiple changes in patterns and rates of substitution. Mol. Phylogenet. Evol. 38, 100–116.

Heraty, J., Ronquiot, F., Carpenter, J.M., Hawks, D., Schulmeister, M., Dowling, A.P., Murray, D., Munro, J., Wheeler, W.C., Schiff, N., 2011. Evolution of the hymenopteran megaradon. Mol. Phylogenet. Evol. 60, 73–88.

Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.

Katoh, K., Kuma, K.I., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33, 518–520.

Kim, I.K., Park, J.D., Shin, S.C., 2011. Prolonged embryonic stage and synchronized life-history of Platyrinae (Hymenoptera: Platygasteridae), a parasitoid of Opodocis robiniae (Diptera: Cecidomyiidae). Biol. Control 57, 24–30.

Klopstein, S., Vilhelmsen, L., Heraty, J.M., Sharkey, R., 2013. The Hymenoptera tree of life: evidence from protein-coding genes and objectively aligned ribosomal data. Mol. Biol. Evol. 30, 1–23.

Lu, C.K., Buhl, P.N., Duso, C., Zhao, M.Z., Zhang, J.S., Ji, Z.X., Gao, S.H., Yu, J.Y., 2010. First discovery of Platyrinae robiniarum (Hymenoptera: Platygasteridae) parasitizing the invasive Phyllobius dubius (Diptera: Cecidomyiidae), a gall maker in China. Acta Entomol. Sin. 53, 223–237.

Mao, M., Gibson, T., Dowton, M., 2015. Higher-level phylogeny of the Hymenoptera inferred from mitochondrial genomes. Mol. Phylogenet. Evol. 84, 34–43.

Massimiliano, B., Andrea, B., Antonio, S., Tomaso, P., Enrico, N., 2014. Is it an ant or a butterfly? Convergent evolution in the mitochondrial gene order of Hymenoptera and Lepidoptera. Genome Biol. Evol. 6, 27–34.

Campbell, N.J., Barker, S.C., 1999. The novel mitochondrial gene arrangement of the cattle tick, Boophilus microplus: fivefold tandem repetition of a coding region. Mol. Biol. Evol. 732–740.

Castro, L.R., Dowton, M., 2005. The position of the Hymenoptera within the holometabola as inferred from the mitochondrial genome of Perga condei (Hymenoptera: Symphyta: perigidae). Mol. Phylogenet. Evol. 34, 469–479.

Castro, L.R., Ribeiro, K., Dowton, M., 2006. Mitochondrial genomes of Vannhornia eucnemidarum (Apoidea: vannhorniidae) and prinecumididae (Apoidea: chrysididae): evidence of rearranged mitochondrial genomes within the Apocrita (Insecta: Hymenoptera). Genome 49, 752–766.

Chen, Z.T., Du, Y.Z., 2016. Rearrangement of mitochondrial genome in insects. Mol. Biol. Evol. 33, 843–851.

Clary, D.O., Wolstenholme, D.R., 1985. The mitochondrial DNA molecular of Drosophila yakuba: nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. 22, 252–271.

Cook, C.E., 2005. The complete mitochondrial genome of the stomatopod crustacean Squilla mantis. BMC Genom. 6, 105.

Crozier, R.H., Crozier, Y.C., 1993. The mitochondrial genome of the honeybee Apis mellifera: complete sequence and organization of gene. Genetica 133, 97–117.

Curto, J.P., Kocher, T.D., 1999. Mitogenomics: digging deeper with mitochondrial genomes. Trends Ecol. Evol. 14, 394–398.

David, P., 2008. ModelTest: phylogenetic model averaging. Mol. Biol. Evol. 1253–1256.

Dowton, M., 1999. Relationships among the cyclostome braconid (Hymenoptera: braconidae) subfamilies inferred from a mitochondrial DNA gene rearrangement. Mol. Phylogenet. Evol. 11, 203–207.

Dowton, M., Austin, A.D., 1997. The evolution of strand-specific compositional bias. A case study in the Hymenopteran mitochondrial 16s rna gene. Mol. Biol. Evol. 14, 109–112.

Dowton, M., Austin, A.D., 1999. Evolutionary dynamics of a mitochondrial rearrangment ‘hot spot’ in the Hymenoptera. Mol. Biol. Evol. 16, 298–309.

Dowton, M., Cameron, S.L., Austin, A.D., Whiting, M.F., 2009a. Phylogenetic approaches for the analysis of mitochondrial genome sequence data in the Hymenoptera - a lineage with both rapidly and slowly evolving mitochondrial genomes. Mol. Phylogenet. Evol. 52, 512–519.

Dowton, M., Cameron, S.L., Downavic, J.I., Austin, A.D., Whiting, M.F., 2009b. Characterization of mitochondrial DNA gene rearrangements in the Hymenoptera suggests that mitochondrial RNA gene position is selectively neutral. Mol. Biol. Evol. 1607–1617.

Dowton, M., Campbell, N.J., 2001. Intramitochondrial recombination - is it why some mitochondrial genes sleep around? Trends Ecol. Evol. 16, 269–271.
Sudhir, K., Glen, S., Koichiro, T., 2016. MEGA7: molecular evolutionary genetics analysis Version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Sun, H., Zhao, W., Lin, R., Zhou, Z., Huai, W., Yao, Y., 2020. The conserved mitochondrial genome of the jewel beetle (Coleoptera: buprestidae) and its phylogenetic implications for the suborder Polyphaga - ScienceDirect. Genomics 112, 3713–3721.

Tang, B.P., Liu, Y., Xin, Z.Z., Zhang, D.Z., Wang, Z.F., Zhu, X.Y., Wang, Y., Zhang, H.B., Zhou, C.L., Chai, X.Y., 2017. Characterisation of the complete mitochondrial genome of Helice wuana (Grapsoidea: Varunidae) and comparison with other Brachyuran crabs. Genomics 110, 221–230.

Tang, P., Zha, J.C., Zheng, B.Y., Wei, S.J., Sharkey, M., Chen, X.X., Vogler, A.P., 2019. Mitochondrial phylogenomics of the Hymenoptera. Mol. Phylogenet. Evol. 131, 8–18.

Timmermans, M.J.T.N., Vogler, A.P., 2012. Phylogenetically informative rearrangements in mitochondrial genomes of Coleoptera, and monophyly of aquatic elateriform beetles (Dryopoidea). Mol. Phylogenet. Evol. 63, 299–304.

Vilhelmsen, L., Mikó, I., Krogmann, L., 2010. Beyond the wasp-waist: structural diversity and phylogenetic significance of the mesosoma in apocritan wasps (Insecta: Hymenoptera). Zool. J. Linn. Soc. 22–194.

Yang, J., Huang, L., Li, Z.R., Sun, H.Q., Zhao, W.X., 2021. Development and preliminary application of novel genomewide SSR markers for genetic diversity analysis of an economically important bio-control agent Platygaster robiniae (Hymenoptera: Platygastridae). J. Genet. 100, 67.

Yang, J., Zhao, W., Yao, Y.X., 2019. Geographic distribution of Platygaster robiniae (Hymenoptera: Platygastridae) parasitizing an invasive insect pest Obolodiplosis robiniae (Diptera: Cecidomyiidae) in China. J. Environ. Entomol. 41, 167–172.

Yang, X.S., Xue, D.Y., Han, H.X., 2013. The complete mitochondrial genome of Iliax pantherina (Lepidoptera: geometridae), with phylogenetic utility of mitochondrial genome in the Lepidoptera. Gene 515, 349–358.

Wei, S.J., Shi, M., Sharkey, M.J., Achterberg, C.V., Chen, X.X., 2010b. Comparative mitogenomics of Braconidae (Insecta: Hymenoptera) and the phylogenetic utility of mitochondrial genomes with special reference to Holometabolous insects. BMC Genom. 11, 371.

Wei, L. Huang et al.