Naturalis Repository

The complete mitochondrial genome of Taeniogonalos taihorina (Bischoff) (Hymenoptera: Trigonalyidae) reveals a novel gene rearrangement pattern in the Hymenoptera

Qiu-Ling Wu, Qian Li, Yun Gu, Bao-Cai Shi, Cees van Achterberg, Shu-Jun Wei,
Xue-Xin Chen

Downloaded from:
http://dx.doi.org/10.1016/j.gene.2014.04.003

Article 25fa Dutch Copyright Act (DCA) - End User Rights

This publication is distributed under the terms of Article 25fa of the Dutch Copyright Act (Auteurswet) with consent from the author. Dutch law entitles the maker of a short scientific work funded either wholly or partially by Dutch public funds to make that work publicly available following a reasonable period after the work was first published, provided that reference is made to the source of the first publication of the work.

This publication is distributed under the Naturalis Biodiversity Center 'Taverne implementation' programme. In this programme, research output of Naturalis researchers and collection managers that complies with the legal requirements of Article 25fa of the Dutch Copyright Act is distributed online and free of barriers in the Naturalis institutional repository. Research output is distributed six months after its first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and copyrights owner(s) of this work. Any use of the publication other than authorized under this license or copyright law is prohibited.

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the department of Collection Information know, stating your reasons. In case of a legitimate complaint, Collection Information will make the material inaccessible. Please contact us through email: collectie.informatie@naturalis.nl. We will contact you as soon as possible.

Gene 543 (2014) 76–84

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

The complete mitochondrial genome of *Taeniogonalos taihorina* (Bischoff) (Hymenoptera: Trigonalyidae) reveals a novel gene rearrangement pattern in the Hymenoptera

Qiu-Ling Wu [a,b], Qian Li [c], Yun Gu [b], Bao-Cai Shi [a], Cees van Achterberg [d], Shu-Jun Wei [a,*], Xue-Xin Chen [c,**]

[a] *Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China*
[b] *College of Agronomy and Plant Protection, Qingdao Agricultural University, Qingdao 266109, China*
[c] *Institute of Insect Sciences, Zhejiang University, Hangzhou 310058, China*
[d] *Terrestrial Zoology, Naturalis Biodiversity Center, 2300 RA Leiden, Netherlands*

ARTICLE INFO

Article history:
Received 27 September 2013
Received in revised form 14 November 2013
Accepted 1 April 2014
Available online 5 April 2014

Keywords:
Mitochondrial genome
Gene rearrangement
Secondary structure
Intergenic spacer
Taeniogonalos taihorina

ABSTRACT

The family Trigonalyidae is considered to be one of the most basal lineages in the suborder Apocrita of Hymenoptera. Here, we determine the first complete mitochondrial genome of the Trigonalyidae, from the species *Taeniogonalos taihorina* (Bischoff, 1914). This mitochondrial genome is 15,927 bp long, with a high A + T-content of 84.60%. It contains all of the 37 typical animal mitochondrial genes and an A + T-rich region. The orders and directions of all genes are different from those of previously reported hymenopteran mitochondrial genomes. Eight tRNA genes, three protein-coding genes and the A + T-rich region were rearranged, with the dominant gene rearrangement events being translocation and local inversion. The arrangements of three tRNA clusters, *trnY–trnM–trnI–trnQ*, *trnW–trnL2–trnC*, and *trnH–trnA–trnR–trnN–trnS–trnE–trnF*, and the position of the *cox1* gene, are novel to the Hymenoptera, even the insects. Six long intergenic spacers are present in the genome. The secondary structures of the RNA genes are normal, except for *trnS2*, in which the D-stem pairing is absent.

1. Introduction

The insect mitochondrial genome is usually a double-stranded circular molecule, with a length ranging from 14 to 20 kb (Boyce et al., 1989), containing 13 protein-coding genes, 22 transfer RNA (tRNA) genes, two ribosomal RNA (rRNA) genes and an A + T-rich region (Boore, 1999). The mitochondrial genome shows maternal inheritance, lack of extensive recombination, and a high mutation rate relative to the nuclear genome (Ballard and Whitlock, 2004). Because of this, many mitochondrial genes are frequently used in studies of population genetic structure and phylogenetic inference (Simon et al., 1994, 2006). Recently, a rapid increase in complete mitochondrial genome sequences in public databases has made it possible to use the whole genome in evolutionary studies (Boore et al., 1998; Dowton et al., 2009a, 2009b; Wei et al., 2010a, 2010b). Notably, higher-level phylogenetic relationships among or within some insect orders can be robustly reconstructed by employing the complete mitochondrial genomes, such as holometabolous insects (Wei et al., 2010b), Neuropterida (Cameron et al., 2009), Orthoptera (Zhou et al., 2010), Hemiptera (Hua et al., 2008, 2009), Coleoptera (Haran et al., 2013; Kim et al., 2009), Diptera (Cameron et al., 2007; Negrisolo et al., 2004) and Hymenoptera (Cameron et al., 2008; Castro and Dowton, 2007; Dowton et al., 2009a; Gotzek et al., 2010; Wei et al., 2010b).

Gene rearrangements are usually conserved within major lineages (Boore, 1999), but may be highly rearranged in certain groups. Gene rearrangement events may serve as useful phylogenetic markers and models for evolutionary studies. Frequent gene rearrangement has been observed in apocritan Hymenoptera from broad examinations of gene segments and whole genome sequences.

The family Trigonalyidae has been proposed to be one of the most basal lineages in the suborder Apocrita of Hymenoptera (Dowton and Austin, 1994; Dowton et al., 1997; Klopfstein et al., 2013; Ronquist et al., 1999). At present, no complete mitochondrial genome has been sequenced from members of this family. Here, we determine the complete mitochondrial genome sequence of Trigonalyidae from the species *Taeniogonalos taihorina* (Bischoff). We then characterize the

---

Abbreviations: atp6, ATP synthase F0 subunit 6; *atp8*, ATP synthase F0 subunit 8; *cob*, cytochrome b; *cox1*, cytochrome c oxidase subunit 1; *cox2*, cytochrome c oxidase subunit 2; *cox3*, cytochrome c oxidase subunit 3; *nad1*, NADH dehydrogenase subunit 1; *nad2*, NADH dehydrogenase subunit 2; *nad3*, NADH dehydrogenase subunit 3; *nad4*, NADH dehydrogenase subunit 4; *nad4l*, NADH dehydrogenase subunit 4l; *nad5*, NADH dehydrogenase subunit 5; *nad6*, NADH dehydrogenase subunit 6; *rrnL*, mitochondrial 16s ribosomal RNA; *rrnS*, mitochondrial 12s ribosomal RNA; *trnX*, transfer RNA genes with isotype given in single-letter symbols of amino acid codes.

[*] Correspondence to: S.J. Wei, Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, 9 Shuguanghuayuan Middle Road, Haidian District, Beijing 100097, China.

[**] Correspondence to: X.X. Chen, Institute of Insect Sciences, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, China.

E-mail addresses: shujun268@163.com (S.-J. Wei), xxchen@zju.edu.cn (X.-X. Chen).

http://dx.doi.org/10.1016/j.gene.2014.04.003
0378-1119/

Table 1
Primers used in the study.

Region	Primer position	Product length (bp)	Primer pairs
cox2–nad3	3400–5731	2331	ATTGGACATCAATGATATTGA[a] TTAGGGTCAAATCCRCAYTC[a]
cox3–cox1	5470–6101	631	GCAGCTGCYTGATAYTGRCA[a] TAAACTTCAGGGTGACCAAAAAATCA[b]
cox1–nad4	5744–10,046	4302	GATGAACAGTTTATCCTCCTT[b] CACAATAATTGCATCATTAC[b]
nad4–rrnL	9837–14,012	4175	RAAACDGGWGCYTCWACATGAGC[b] ATACCTTAGGGATAACAGCATAA[b]
rrnL–rrnS	13,858–15,702	1844	CCGGTTTGAACTCARATCATGTAA[a] GTGCCAGCAGYYGCGGTTANAC[a]
rrnS–cox3	15,169–3174	3932	TGACGGGCTATTTGTACTTAT[b] AAGGTCAAGGTCTTACACTTACTA[b]

[a] Primers from Simon et al., 2006.
[b] Primers designed in this study.

mitochondrial genome and compare it with others across the Hymenoptera. This work provides the first report of a novel gene rearrangement pattern for both tRNA and protein-coding genes.

2. Materials and methods

2.1. Sample origin and DNA extraction

An adult female specimen of *T. taihorina* from Fengyang Mountain (Longquan City, Zhejiang Province, China) was used for the extraction. It was stored at −80 °C in 100% ethanol prior to DNA extraction. Total genomic DNA was extracted using the DNeasy tissue kit (Qiagen, Hilden, Germany) following the manufacturer's protocols, from the single adult. The remnant voucher specimens were kept in the Integrated Pest Management Laboratory, Beijing Academy of Agriculture and Forestry Sciences.

2.2. Primer design, PCR amplification and sequencing

The entire genome of *T. taihorina* was amplified in six fragments (Table 1). Partial fragments of *cox2–nad3, cox3–cox1,* and *rrnL–rrnS* were amplified using a range of universal insect mitochondrial primers from Simon et al. (2006). Species-specific primers were designed using Primer Premier 5.0, based on sequenced fragments to amplify others to bridge the gap of different fragments. Six fragments of 631–4302 base pairs (bp) were amplified, covering the entire mitochondrial genome. Polymerase chain reaction (PCR) was carried out using Takara LA Taq and Takara Hot Start LA Taq (Takara Biomedical, Japan), using the recommended conditions and components. All the PCR products were sent to TSINGKE Company (Beijing, China) for sequencing, after purification in both directions using PCR primers and designed walking primers.

2.3. Genome annotation and analysis

The transfer RNA (tRNA) genes were initially identified using the tRNAscan-SE search server (Lowe and Eddy, 1997), setting the parameters so that the source was Mito/Chloromast, and the genetic code was the Invertebrate Mito genetic code. Twenty of the 22 typical animal mitochondrial tRNA genes were identified. Two tRNA genes that were not

Table 2
Annotation of the *Taeniogonalos taihorina* mitochondrial genome.

Gene	Strand	Region	Length (bp)	Anti/start codon	Stop codon	Intergenic nucleotides
trnY	+	1–64	64	GTA		5
trnM	+	70–135	66	CAT		1
trnI	+	137–202	66	GAT		4
trnQ	+	207–271	65	TTG		45
nad2	+	317–1322	1006	ATT	T	0
trnW	+	1323–1387	65	TCA		0
trnL2	+	1388–1452	65	TAA		4
trnC	+	1457–1525	69	GCA		38
cox2	+	1564–2238	675	ATA	TAA	53
atp8	+	2292–2447	156	ATT	TAA	4
atp6	+	2452–3119	668	ATG	TA	0
cox3	+	3120–3911	792	ATG	TAA	7
trnG	+	3919–3982	64	TCC		47
nad3	+	4030–4383	354	ATT	TAA	0
A + T-rich region		4384–5304	921			0
trnD	−	5305–5371	67	GTC		−1
cox1	+	5371–6925	1555	ATA	T	−1
trnH	−	6925–6994	70	GTG		5
trnA	+	7000–7064	65	TGC		−1
trnR	+	7064–7130	67	TCG		−6
trnN	+	7125–7189	65	GTT		−2
trnS2	+	7188–7249	62	TCT		52
trnE	+	7302–7367	66	TTC		−2
trnF	−	7366–7430	65	GAA		−3
nad5	−	7428–9093	1666	ATT	T	58
nad4	−	9152–10,483	1332	ATA	TAA	1
nad4l	−	10,485–10,773	289	ATT	T	1
trnT	+	10,775–10,837	63	TGT		−1
trnP	−	10,837–10,900	64	TGG		0
cob	−	10,901–12,040	1140	ATG	TAA	0
nad6	−	12,041–12,594	554	ATG	TA	1
trnK	−	12,596–12,664	69	TTT		0
trnS1	+	12,665–12,731	67	TGA		8
nad1	−	12,740–13,686	947	ATT	TA	1
trnL1	−	13,688–13,755	68	TAG		0
rrnL	−	13,756–15,044	1289			0
trnV	−	15,045–15,106	62	TAC		−1
rrnS	−	15,106–15,927	822			

"+" indicates the gene coded on major strand, "−" indicates the gene coded on minor strand.

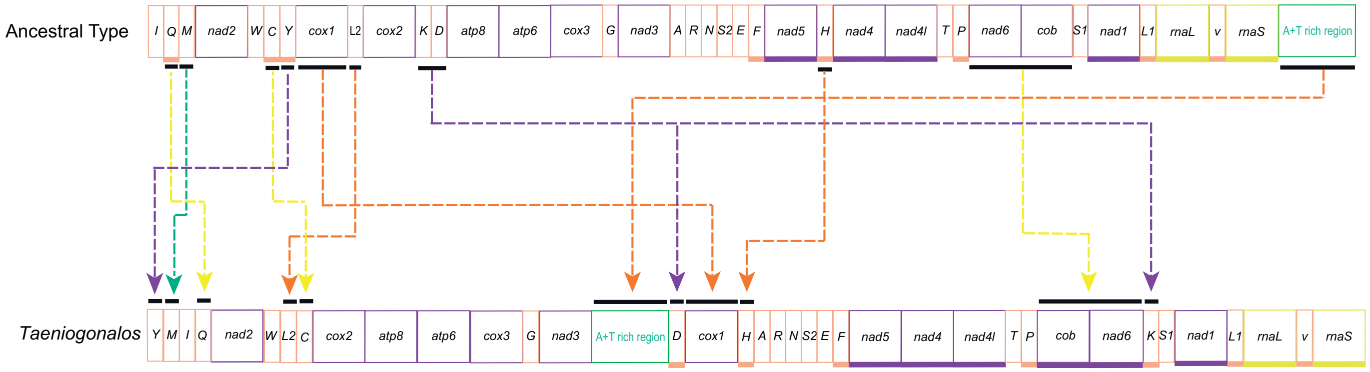

Fig. 1. Gene rearrangement in the *Taeniogonalos taihorina* mitochondrial genome. The red dashed line shows translocation of the genes. The yellow dashed line shows remote inversion of the genes. The green dashed line shows local inversion of the genes. The blue dashed line shows shuffling of the genes.

found by tRNAscan-SE were confirmed by sequence alignment in MEGA5 (Tamura et al., 2011). Protein-coding genes (PCGs) and rRNA genes were identified using BLAST searches in GenBank, and subsequently by alignment with genes of other hymenopteran insects for the initiation and termination codons.

The Relative Synonymous Codon Usage (RSCU) of all protein-coding genes was calculated by MEGA5 (Tamura et al., 2011). The GC and AT asymmetries, called GC-skew and AT-skew, were measured such that AT-skew = (A% − T%) / (A% + T%) and GC-skew = (G% − C%) / (G% + C%).

All tRNA secondary structures were predicted using the tRNAscan-SE search server (Lowe and Eddy, 1997), except for *trnR* and *trnS2*, which were predicted manually. The secondary structures of the large and small rRNAs (*rrnL* and *rrnS*) were inferred by aligning to the models predicted for *Drosophila virilis* (Cannone et al., 2002) and *Grapholita molesta* (Gong et al., 2012).

3. Results and discussion

3.1. Genome features and base composition

The complete mitochondrial genome of *T. taihorina* is 15,927 bp long and contains all of the 37 typical animal mitochondrial genes, including 13 protein-coding genes, two rRNA and 22 tRNA genes, and an A + T-rich region (Table 2, Fig. 1).

This complete mitochondrial genome is highly economized in genome size, as seen in other animals (Boore, 1999). A total of 18 bp of overlapping nucleotides exist between neighboring genes in nine locations, with 1–6 bp of overlapping sequence. A total of 335 bp of intergenic nucleotides also exist in 18 locations, with an intergenic spacer length of 1–58 bp. Ten pairs of genes are directly adjacent, without intergenic or overlapping nucleotides.

The A + T-content of the *T. taihorina* mitochondrial genome is 84.6%, which is similar to other sequenced hymenopteran species. There are more T and G nucleotides than A and C nucleotides in the major strand (Table 3).

3.2. Transfer RNA genes

The lengths of the 22 tRNA genes range from 62 to 70 bp (Table 2). Among the 22 tRNA genes, eight are encoded on the minor strand. The sizes of the variable loop and D-loop affect the tRNA length. In the *T. taihorina* mitochondrial genome, most D-loops are 4–8 bp long. All tRNA genes have a typical clover-leaf structure, except for *trnS2*. The D-stem pairings in the DHU arm are absent in *trnS2* (Fig. 2), which has also been reported in other insects (Beard et al., 1993; Crozier and

Table 3
Base composition of the hymenopteran mitochondrial genomes.

Species	Whole genome							All protein-coding genes						
	T%	C%	A%	G%	A+T%	AT skew	GC skew	T%	C%	A%	G%	A+T%	AT skew	GC skew
Taeniogonalos taihorina	43.51	6.77	41.09	8.63	84.60	−0.0286	0.1208	46.27	8.79	36.19	8.75	82.46	−0.1222	−0.0023
Bombus hypocrita	42.56	9.68	42.77	4.98	85.33	0.0025	−0.3206	46.71	8.21	37.31	7.77	84.02	−0.1119	−0.0275
Spathius agrili	45.12	6.55	38.88	9.45	84.00	−0.0743	0.1813	47.08	7.99	35.00	9.93	82.08	−0.1472	0.1083
Solenopsis richteri	38.55	16.59	38.39	6.47	76.95	−0.0021	−0.4389	42.84	14.37	31.23	11.57	74.07	−0.1567	−0.1080
Solenopsis geminata	37.94	16.95	38.60	6.51	76.54	0.0086	−0.4450	42.26	14.86	31.32	11.56	73.58	−0.1487	−0.1249
Schlettererius cinctipes	39.47	11.22	42.76	6.52	82.23	0.0400	−0.2649	43.96	9.41	37.64	8.95	81.61	−0.0774	−0.0251
Philanthus triangulum	39.21	10.22	44.43	6.14	83.64	0.0624	−0.2494	45.78	8.58	36.98	8.66	82.75	−0.1063	0.0046
Orussus occidentalis	37.46	15.55	38.75	8.23	76.21	0.0169	−0.3078	41.97	13.70	32.20	12.13	74.17	−0.1317	−0.0608
Nasonia giraulti	38.09	10.03	43.87	8.01	81.96	0.0705	−0.1120	44.90	8.59	36.61	9.90	81.52	−0.1017	0.0709
Enicospilus sp.	43.44	8.73	41.75	6.08	85.19	−0.0198	−0.1789	46.73	7.73	37.29	8.26	84.02	−0.1124	0.0332
Cephus cinctus	39.57	11.59	42.39	6.44	81.95	0.0344	−0.2856	43.99	10.71	35.05	10.25	79.05	−0.1131	−0.0220
Cotesia vestalis	47.46	5.78	39.70	7.06	87.16	−0.0890	0.0997	50.12	6.46	35.60	7.82	85.71	−0.1694	0.0952
Trissolcus basalis	40.49	9.80	43.65	6.01	84.13	0.0376	−0.2397	45.70	9.12	36.55	8.63	82.25	−0.1112	−0.0276
Perga condei	33.80	14.60	42.80	8.80	76.60	0.1175	−0.2479	43.20	11.60	33.30	12.00	76.50	−0.1294	0.0169
Vanhornia eucnemidarum	36.00	14.80	42.20	7.10	78.20	0.0793	−0.3516	42.70	11.70	35.50	10.00	78.20	−0.0921	−0.0783
Evania appendigaster	37.90	15.00	39.90	7.20	77.80	0.0257	−0.3513	42.70	13.20	31.80	12.30	74.50	−0.1463	−0.0353
Diadegma semiclausum	41.50	9.60	42.10	6.70	83.60	0.0072	−0.1779	46.90	8.20	36.80	8.10	83.70	−0.1207	−0.0061
Abispa ephippium	39.50	14.60	39.10	6.70	78.60	−0.0051	−0.3709	43.50	11.20	35.20	10.10	78.70	−0.1055	−0.0516
Polistes humilis	41.10	10.70	42.30	5.90	83.40	0.0144	−0.2892	46.60	8.50	36.80	8.10	83.40	−0.1175	−0.0241
Apis mellifera	41.20	10.50	42.10	6.30	83.30	0.0108	−0.2500	46.10	8.50	37.20	8.20	83.30	−0.1068	−0.0180
Bombus ignitus	42.30	9.40	42.80	5.60	85.10	0.0059	−0.2533	47.50	7.50	37.60	7.40	85.10	−0.1163	−0.0067
Melipona bicolor	42.50	8.50	43.80	5.20	86.30	0.0151	−0.2409	48.00	6.90	38.40	6.80	86.40	−0.1111	−0.0073

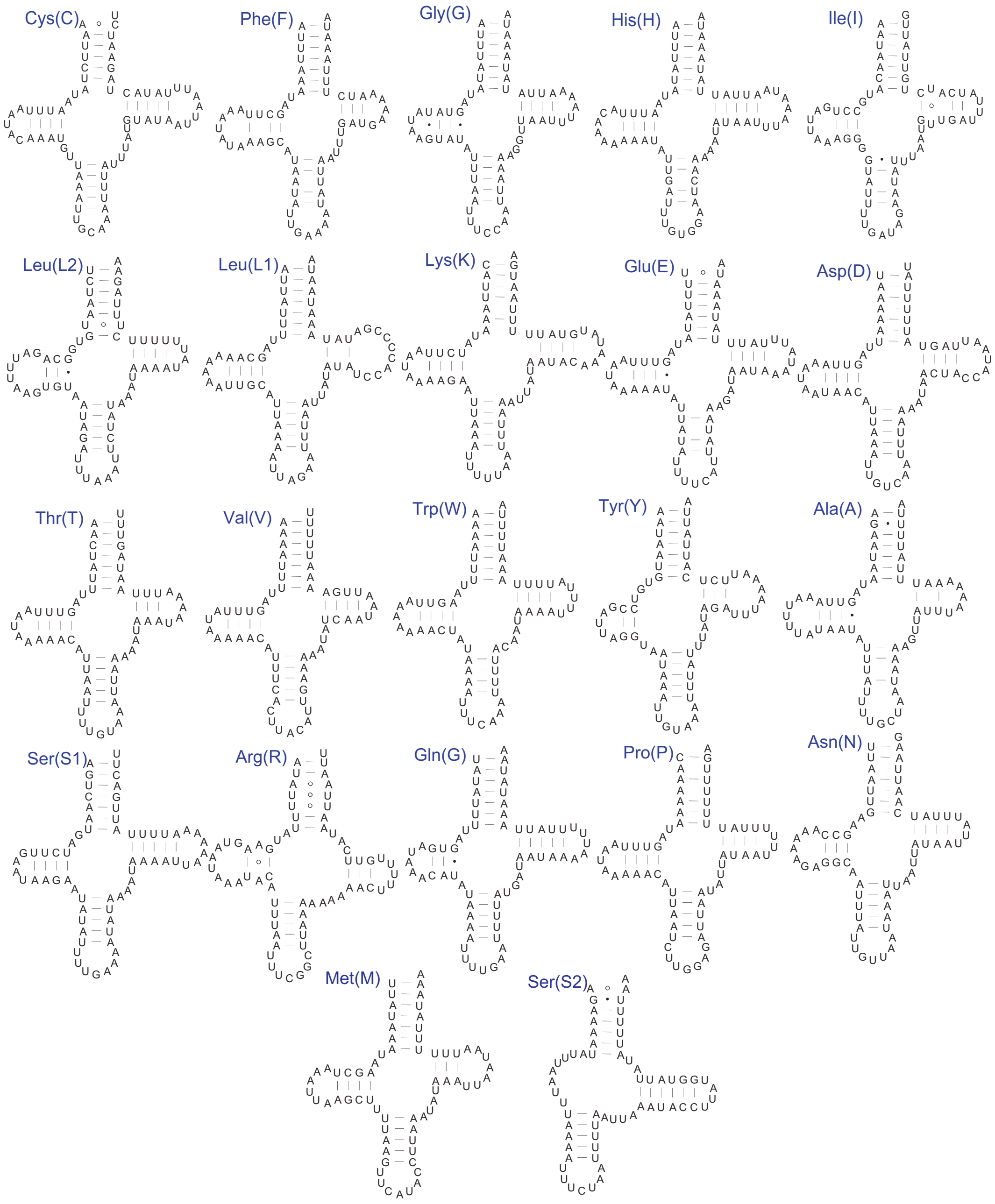

Fig. 2. Predicted secondary structures for the 22 tRNA genes of *Taeniogonalos taihorina*. Base-pairing is indicated as follows: Watson–Crick pairs by lines, wobble GU pairs by dots and other non-canonical pairs by circles.

Crozier, 1993; Shao and Stephen, 2003; Sheffield et al., 2008; Wei et al., 2010c). The *trnR* and *trnS2* genes could not be found and identified using the tRNAscan-SE website. We determined their exact boundaries by aligning with homologs identified in other hymenopteran insects. In total, 18 mismatches exist in all the stem structures of *T. taihorina*, including three A–A pairs, nine G–U pairs, five U–U pairs and one A–C pair. The number of mismatches is therefore normal in the tRNA secondary structures of *T. taihorina* compared with that of other insects.

The *trnS2* and *trnK* genes use abnormal anticodons TCT and TTT, which have been found to be correlated with frequent gene rearrangement events (Wei et al., 2009).

3.3. Ribosomal RNA genes

The position and length of the *rrnL* and *rrnS* genes are conserved. The length of the *rrnL* gene is 1289 bp, with 87.66% A + T-content, while the length of the *rrnS* gene is 822 bp with 89.42% A + T-content. The sizes of the *rrnL* and *rrnS* genes are normal. The A + T-content of the rRNA genes is comparable to homolog genes in other hymenopteran insects.

We identified the secondary structures of both *rrnL* and *rrnS* according to homolog genes from *D. virilis* (Cannone et al., 2002) and *G. molesta* (Gong et al., 2012). The 49 helices found in *Evania appendigaster* are also present in *T. taihorina* (Wei et al., 2010c). H837 usually forms a long stem structure with a small loop in the terminal (Gillespie et al., 2006), but it forms a longer stem and a larger loop in *T. taihorina* (Fig. 3). The helix H991 shows helical length and loop size variability among insect orders, and even among taxa within families of Odonata (Misof and Fleck, 2003). The predicted structures of H2347 and H2520 are variable (Buckley et al., 2000), but in *T. taihorina* they are more similar to those from *E. appendigaster* (Wei et al., 2010c) than those from other insects (Misof and Fleck, 2003).

The *rrnS* secondary structure of *T. taihorina* is characterized by 30 helices (Fig. 4), a feature also shared by *Apis mellifera* (Gillespie et al., 2006). Helix 47 in *T. taihorina* forms two loops that are similar to *E. appendigaster*, but differ from *A. mellifera*. H673 is well conserved in *T. taihorina*, as in other insects, such as *D. virilis* (Schnare et al., 1996) and *Diadegma semiclausum* (Wei et al., 2009).

3.4. Protein-coding genes

The total length of protein-coding genes is 11,134 bp, accounting for 69.91% of the total length of the *T. taihorina* mitochondrial genome. The size of each protein-coding gene is similar to that of their orthologs in other insects. The genes with the highest A + T-content in the hymenopteran mitochondrial genomes are usually *nad6* or *atp8*. In the *T. taihorina* mitochondrial genome, the highest A + T-content is 95.50% in *atp8*, while the lowest is 74.6% in the *cox1* gene (Table 4). The start codons of all protein-coding genes are ATN, including three ATA, six ATT and four ATG, which is typical in the metazoan mitochondrial genomes (Wolstenholme, 1992). The *atp6*, *nad6* and *nad1* genes use an incomplete stop codon, TA, and the *nad2*, *cox1*, *nad5* and *nad4l* genes use T. The other protein-coding genes all use TAA as the stop codon. The incomplete stop codon is common in invertebrate mitochondrial genomes (Crozier and Crozier, 1993; Masta and Boore, 2004).

The codon usage reflects the A and T bias in the *T. taihorina* mitochondrial genome (Table 5). Leu, Ile and Phe are the most commonly used amino acids in the protein-coding genes of *T. taihorina*. UUA is the most frequently used codon in Leu. The RSCU values show a biased use of A and T nucleotides in *T. taihorina* (Table 5).

3.5. Non-coding regions

One of the most interesting phenomena in the *T. taihorina* mitochondrial genome is the presence of seven major non-coding regions of more

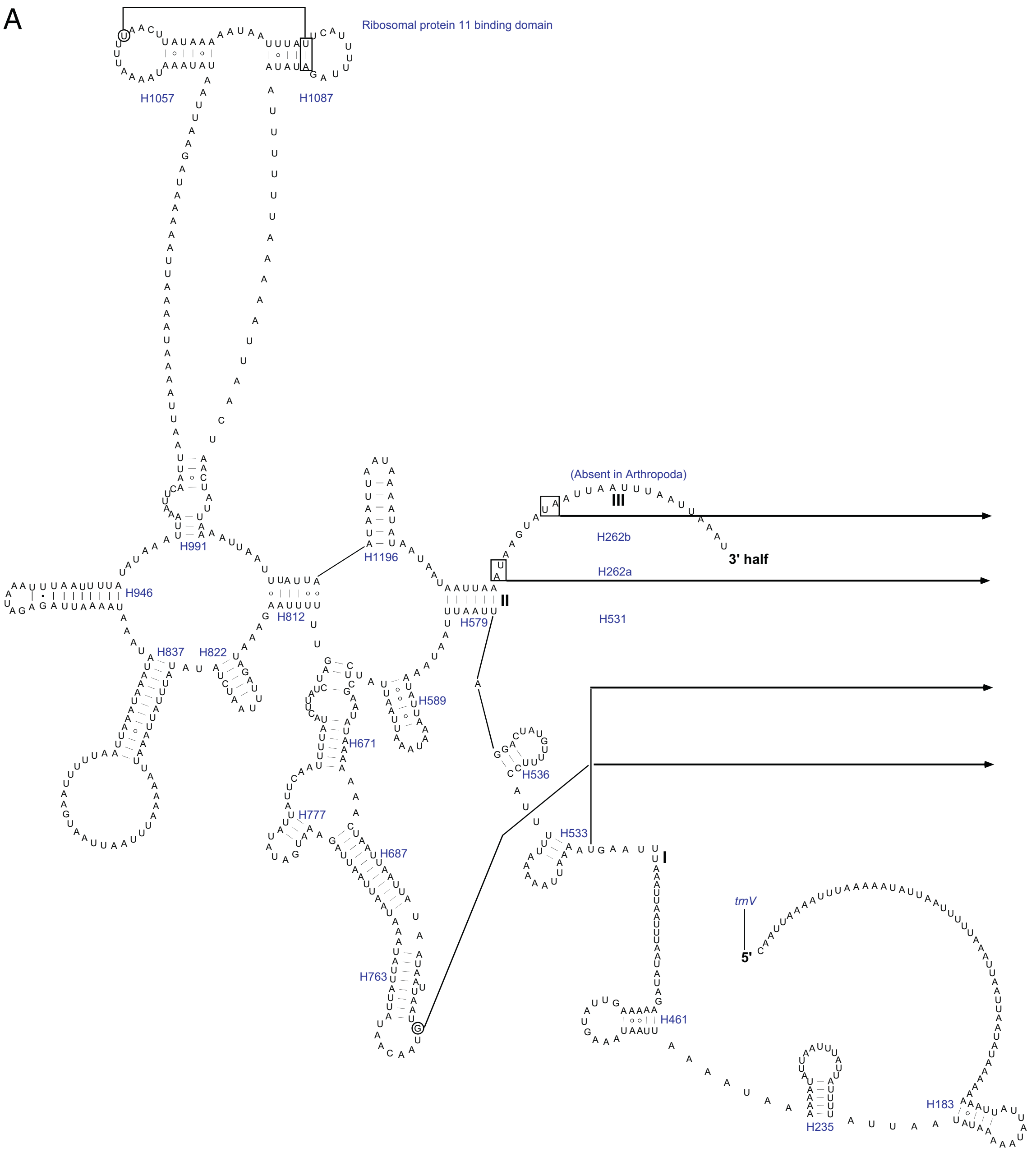

Fig. 3. Predicted *rrnL* secondary structure in the *Taeniogonalos taihorina* mitochondrial genome. A: 5′ half of *rrnL*; B: 3′ half of *rrnL*. Tertiary interactions and base triples are shown connected by continuous lines. Symbols are as in Fig. 2.

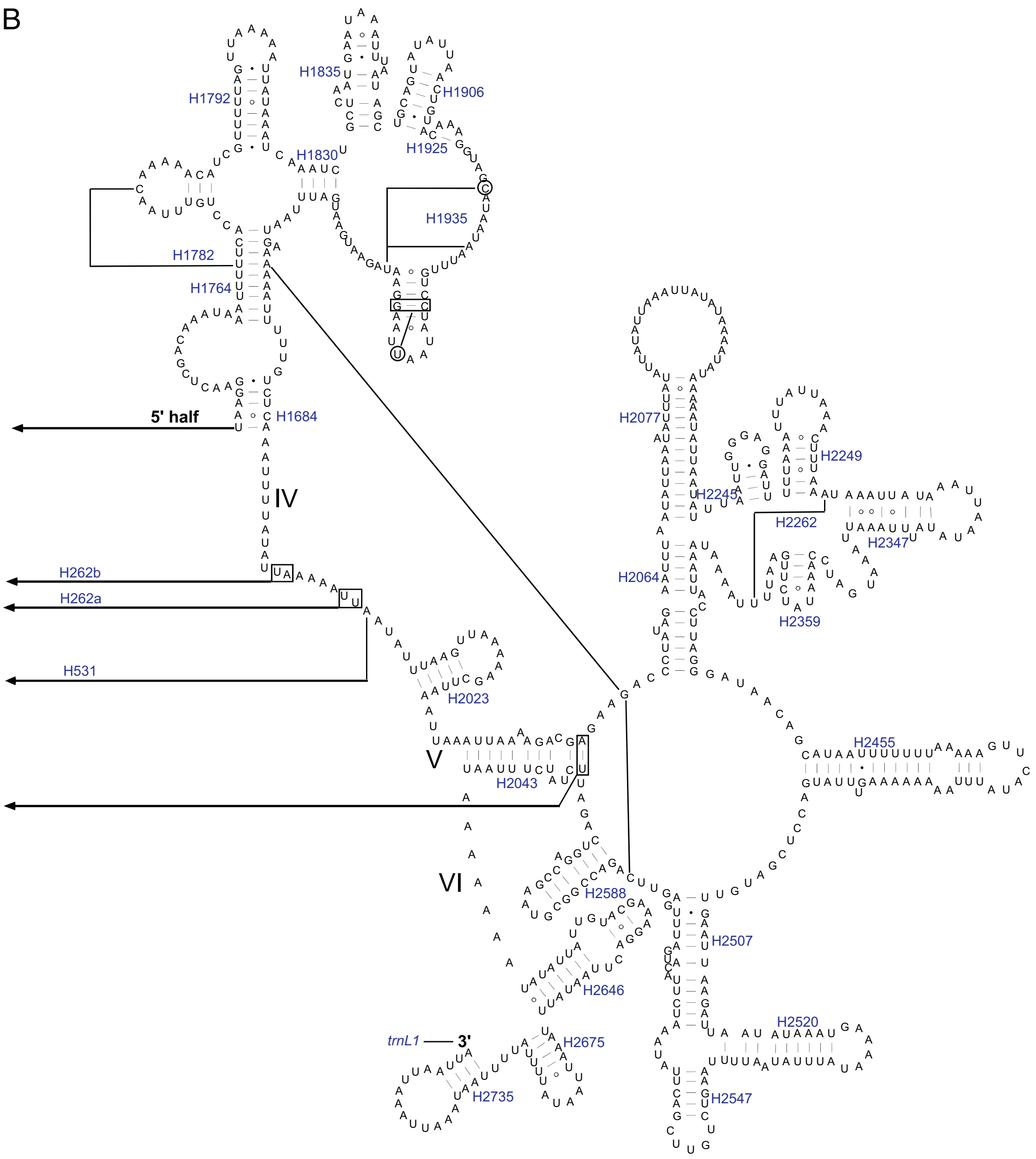

Fig. 3 (*continued*).

than 20 bp, named Spacers 1–7. Spacer 1 consists of 45 bp between *trnQ* and *nad2*. Spacer 2 is 38 bp between *trnC* and *cox2*. Spacer 3 is 53 bp between *cox2* and *atp8*. Spacer 4 is 47 bp between *trnG* and *nad3*. Spacer 5 is 52 bp between *trnS2* and *trnE*. Spacer 6 is 58 bp between *trnF* and *nad5*. Spacer 7 is 921 bp between *nad3* and *trnD*.

In insect mitochondrial genomes, the longest non-coding region is usually the A + T-rich region, located between the *rrnS* gene and the *trnI*–*trnQ*–*trnM* cluster, though exceptions exist, such as in Hymenoptera (Wei et al., 2009, 2010b). The length of the *T. taihorina* A + T-rich region is 921 bp, with an A + T-content of 92.62%, which is not located between the *rrnS* gene and the *trnI*–*trnQ*–*trnM* cluster. Long intergenic spacers have been identified in several insect mitochondrial genomes (Wei et al., 2009). The reason for the appearance of a long intergenic spacer might be the rearrangement of genes. In Hymenoptera, the tRNA rearrangement is normal, so we suggest that spacers 1, 2, 3, 5 and 6 may be the products of gene rearrangements, and have no function in translation or transcription. Spacer 7 is proposed to be the A + T-rich region because of its high A + T-content (92.62%). This region, characterized by five elements (Boore, 1999; Zhang and Hewitt, 1997), is believed to function in the initiation of replication and control of transcription. In *T. taihorina*, we found some of the elements commonly present in most insect A + T-rich regions, such as four repeats of a 58-bp long sequence, a Poly T stretch, a $(TA(A))_n$-like stretch, and a TATA motif, which may function in the initiation of genome replication (Zhang and Hewitt, 1997).

3.6. Gene rearrangement

The most remarkable characteristic of the *T. taihorina* mitochondrial genome is the novel gene arrangement pattern. Eleven genes, including eight tRNA genes, three protein-coding genes, and the A + T-rich region, were rearranged in this mitochondrial genome (Fig. 1).

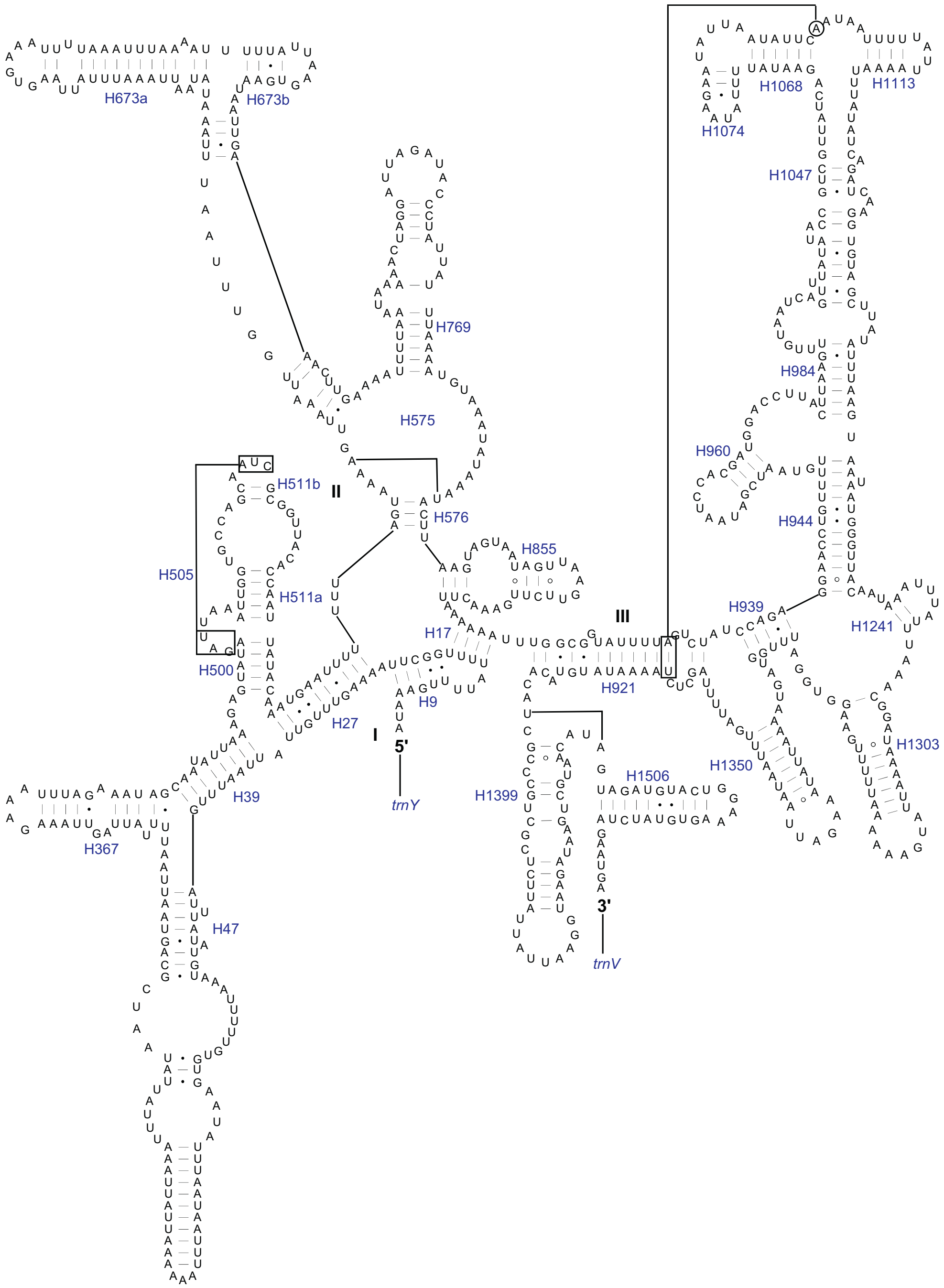

Fig. 4. Predicted *rrnS* secondary structure in the *Taeniogonalos taihorina* mitochondrial genome. Symbols are as in Figs. 2 and 3.

Rearrangement of tRNA genes is common in the hymenopteran mitochondrial genomes, especially those in tRNA clusters, such as in the junctions of the A + T-rich region—*nad2*, *nad2–cox1*, *cox2–atp8* and *nad3–nad5* (Dowton and Austin, 1999; Dowton et al., 2003, 2009b). Gene rearrangement events have been classified into translocations, local inversions (inverted in the local position), gene shuffling (local translocation) and remote inversions (translocated and inverted) (Dowton et al., 2003). Here, the *trnQ* and *trnC* genes were inverted locally compared with the putative ancestral arrangement of insects (Clary and Wolstenholme, 1985). The *trnY*, *trnD*, and *trnK* genes were remotely inverted. The local inversion might be explained by combination (Dowton et al., 2003), whereas the remote inversion might be explained by two recombination events. The *trnM* gene was shuffled upstream of the *trnI–trnQ–trnM* cluster, which might be explained by the TDRL model (tandem duplication followed by random loss) (San Mauro et al., 2006), with evidence of a 45-bp remnant of nucleotides between *trnQ* and *nad2*. The *trnL2* gene was translocated into a position between the *trnW* and *trnC* genes, whereas the *trnH* gene was translocated upstream of the *trnA–trnR–trnN–trnS2–trnE–trnF* cluster. Since the protein-coding gene *cox1* was rearranged out of its original location, the rearrangement of *trnL2* could also be classified as shuffling with *trnW* and *trnC*, because these genes would be directly adjacent after the *cox1* rearrangement. Therefore, a TDRL model could be used to explain the *trnL2* gene rearrangement, as evidenced by a 53-bp remnant of nucleotides in its original location. For the *trnH* gene, the translocation distance includes five tRNA genes and the protein-coding gene *nad5*, which is moderately long. Thus, the *trnH* rearrangement could be explained either by the illicit primer model (Dowton and Austin, 1999;

Table 4
Base composition of protein-coding and rRNA genes in the *Taeniogonalos taihorina* mitochondrial genome.

Gene	T%	C%	A%	G%	A + T%	AT skew	GC skew
nad2	50.6	5.2	39.2	5.1	89.8	−0.1269	−0.0097
cox2	44.9	8.7	36.6	9.8	81.5	−0.1018	0.0595
atp8	53.8	3.2	41.7	1.3	95.5	−0.1267	−0.4222
atp6	47.0	7.5	36.2	9.3	83.2	−0.1298	0.1071
cox3	47.0	8.3	33.6	11.1	80.6	−0.1663	0.1443
nad3	51.4	6.2	33.9	8.5	85.3	−0.2052	0.1565
cox1	44.6	9.7	30.0	15.7	74.6	−0.1957	0.2362
nad5	43.6	9.8	40.6	5.9	84.2	−0.0356	−0.2484
nad4	46.2	9.1	38.1	6.5	84.3	−0.0961	−0.1667
nad4l	49.1	6.6	38.4	5.9	87.5	−0.1223	−0.056
cob	45.5	10.8	34.3	9.4	79.8	−0.1404	−0.0693
nad6	47.8	9.0	37.2	6.0	85.0	−0.1247	−0.2000
nad1	44.9	9.9	36.0	9.2	80.9	−0.1100	−0.0366
rrnL	41.85	4.99	47.57	5.60	89.42	0.0640	0.0576
rrnS	39.80	5.59	47.87	6.75	87.66	0.0920	0.0940

Lunt and Hyman, 1997) or the TDRL model. A 58-bp remnant of nucleotides in its original location could provide evidence for the TDRL model. Among the tRNA gene arrangements, the patterns of the *trnY*–*trnM*–*trnI*–*trnQ* genes, the *trnW*–*trnL2*–*trnC* genes, and the *trnH*–*trnA*–*trnR*–*trnN*–*trnS*–*trnE*–*trnF* genes are novel to the Hymenoptera.

The rearrangement of protein-coding genes has been shown in three *Nasonia* (Chalcidoidea: Pteromalidae) species (Oliveira et al., 2008), two species from the family Agaonidae (Chalcidoidea) (Xiao et al., 2011), one species from the Ichneumonidae (Dowton et al., 2009b), one species from the Microgastrinae (Ichneumonoidea: Braconidae) (Wei et al., 2010a) and one species from the Bethylidae (Aculeata: Chrysidoidea). In the *T. taihorina* mitochondrial genome, the *cox1* gene was translocated upstream of the *trnA*–*trnR*–*trnN*–*trnS2*–*trnE*–*trnF* cluster. In the reported hymenopteran species, the position and direction of *cox1* were changed in *Philotrypesis pilosa*, *Philotrypesis* sp. (Xiao et al., 2011) and *Nasonia* spp. (Oliveira et al., 2008), with a pattern of *cox2*–*trnL2*–*cox1*. In the *T. taihorina* mitochondrial genome, although the *cox1* gene has changed its relative location, it is still encoded on the majority strand. In insects, gene rearrangement of *cox1* is rare. At present, a change of the relative position of the *cox1* gene has been shown in the Thysanoptera (Shao and Stephen, 2003; Yan et al., 2012) and Phthiraptera (Covacin et al., 2006; Shao et al., 2001). However, the arrangement pattern of the *nad3*–A + T-rich region–*cox1*–*trnH* cluster found here is novel in the insects. In the *T. taihorina* mitochondrial genome, the *cob* and *nad6* genes were inverted locally. The inversion of the *cob*–*nad6* region was also reported in *Venturia canescens* (Dowton et al., 2009b) and *Cotesia vestalis* (Wei et al., 2010b) in the Hymenoptera.

The A + T-rich region is usually located in the downstream of *rrnS*. The frequent rearrangement and duplication of the A + T-rich regions were found in thrips (Shao and Stephen, 2003; Yan et al., 2012), louse (Cameron et al., 2011; Covacin et al., 2006; Shao et al., 2001), and booklice (Wei et al., 2012). The rearrangement of the A + T-rich region was also found in the Protura (Chen et al., 2011). In the Hymenoptera, although the boundary tRNA genes of the A + T-rich region was frequently rearranged, the relative position of the A + T-rich region downstream of the *rrnS* is conserved (Wei et al., 2010b). In the newly sequenced *T. taihorina* mitochondrial genome, the A + T-rich region was translocated to the downstream of *nad3*, far from the *rrnS*, which is novel to the Hymenoptera.

Table 5
Codon usage in the *Taeniogonalos taihorina* mitochondrial genome.

AA	Codon	No.	RSCU	AA	Codon	No.	RSCU	AA	Codon	No.	RSCU
Leu	UUA	484	5.31	Tyr	UAU	179	1.77	Ala	GCU	33	1.76
	UUG	21	0.23		UAC	23	0.23		GCC	2	0.11
	CUU	31	0.34	Lys	AAA	148	1.89		GCA	38	2.03
	CUC	0	0		AAG	9	0.11		GCG	2	0.11
	CUA	11	0.12	Gly	GGU	50	1.45	Glu	GAA	65	1.86
Ile	AUU	444	1.91		GGC	4	0.12		GAG	5	0.14
	AUC	20	0.09		GGA	63	1.83	His	CAU	56	1.96
Pro	CCU	67	2.5		GGG	21	0.61		CAC	1	0.04
	CCC	1	0.04	Val	GUU	71	2.41	Asp	GAU	49	1.88
	CCA	39	1.46		GUC	1	0.03		GAC	3	0.12
Ser	UCU	93	2.11		GUA	42	1.42	Arg	CGU	12	1.14
	UCC	10	0.23		GUG	4	0.14		CGC	1	0.1
	UCA	131	2.97	Ser	AGU	22	0.5		CGA	27	2.57
	UCG	3	0.07		AGC	1	0.02		CGG	2	0.19
Asn	AAU	227	1.87		AGA	83	1.88	Gln	CAA	40	1.9
	AAC	16	0.13		AGG	10	0.23		CAG	2	0.1
Thr	ACU	58	1.46	Phe	UUU	374	1.91	Cys	UGU	32	1.88
	ACC	5	0.13		UUC	18	0.09		UGC	2	0.12
	ACA	93	2.34	Trp	UGA	72	1.78	Met	AUA	352	1.91
	ACG	3	0.08		UGG	9	0.22		AUG	17	0.09

RSCU: Relative synonymous codon usage; AA: Amino acid; No.: Number.

Conflict of interest

There is no conflict of interest in this paper.

Acknowledgments

We thank Sheng-Long Liu and Prof. Jun-Hua He (Zhejiang University) for providing specimen. Funding for this study was provided jointly by the National Science Foundation of China (31101661, 30970384), State Key Program of National Natural Science Foundation of China (31230068), the NSFC Innovative Research Groups (31021003), 973 Program (2013CB127600), the Beijing New Star Program on Science and Technology (2010B027) and the Beijing Excellent Talents Program (2010D002020000010).

References

Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Molecular Ecology 13, 729–744.

Beard, C.B., Hamm, D.M., Collins, F.H., 1993. The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. Insect Molecular Biology 2, 103–124.

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

Boore, J.L., Lavrov, D.V., Brown, W.M., 1998. Gene translocation links insects and crustaceans. Nature 392, 667–668.

Boyce, T.M., Zwick, M.E., Aquadro, C.F., 1989. Mitochondrial DNA in the bark weevils: Size, structure and heteroplasmy. Genetics 123, 825–836.

Buckley, T.R., Simon, C., Flook, P.K., Misof, B., 2000. Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. Insect Molecular Biology 9, 565–580.

Cameron, S.L., Lambkin, C.L., Barker, S.C., Whiting, M.F., 2007. A mitochondrial genome phylogeny of Diptera: Whole genome sequence data accurately resolve relationships over broad timescales with high precision. Systematic Entomology 32, 40–59.

Cameron, S.L., Dowton, M., Castro, L.R., Ruberu, K., Whiting, M.F., Austin, A.D., Diement, K., Stevens, J., 2008. Mitochondrial genome organization and phylogeny of two vespid wasps. Genome 51, 800–808.

Cameron, S.L., Sullivan, J., Song, H., Miller, K.B., Whiting, M.F., 2009. A mitochondrial genome phylogeny of the Neuropterida (lace-wings, alderflies and snakeflies) and their relationship to the other holometabolous insect orders. Zoologica Scripta 38, 575–590.

Cameron, S., Yoshizawa, K., Mizukoshi, A., Whiting, M., Johnson, K., 2011. Mitochondrial genome deletions and minicircles are common in lice (Insecta: Phthiraptera). BMC Genomics 12, 394.

Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Muller, K.M., 2002. The comparative RNA web (CRW) site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics 3, 1471–2105.

Castro, L.R., Dowton, M., 2007. Mitochondrial genomes in the Hymenoptera and their utility as phylogenetic markers. Systematic Entomology 32, 60–69.

Chen, W.J., Bu, Y., Carapelli, A., Dallai, R., Li, S., Yin, W.Y., Luan, Y.X., 2011. The mitochondrial genome of *Sinentomon erythranum* (Arthropoda: Hexapoda: Protura): An example of highly divergent evolution. BMC Evolutionary Biology 11, 246.

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

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

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

Dowton, M., Austin, A.D., 1994. Molecular phylogeny of the insect order Hymenoptera: Apocritan relationships. Proceedings of the National Academy of Sciences 91, 9911–9915.

Dowton, M., Austin, A.D., 1999. Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. Molecular Biology and Evolution 16, 298–309.

Dowton, M., Austin, A., Dillon, N., Bartowsky, E., 1997. Molecular phylogeny of the apocritan wasps: The Proctotrupomorpha and Evaniomorpha. Systematic Entomology 22, 245–255.

Dowton, M., Castro, L.R., Campbell, S.L., Bargon, S.D., Austin, A.D., 2003. Frequent mitochondrial gene rearrangements at the Hymenopteran *nad3–nad5* junction. Journal of Molecular Evolution 56, 517–526.

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. Molecular Phylogenetics and Evolution 52, 512–519.

Dowton, M., Cameron, S.L., Dowavic, J.I., Austin, A.D., Whiting, M.F., 2009b. 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.

Gillespie, J.J., Johnston, J.S., Cannone, J.J., Gutell, R.R., 2006. Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): Structure, organization, and retrotransposable elements. Insect Molecular Biology 15, 657–686.

Gong, Y.J., Shi, B.C., Kang, Z.J., Zhang, F., Wei, S.J., 2012. The complete mitochondrial genome of the oriental fruit moth *Grapholita molesta* (Busck) (Lepidoptera: Tortricidae). Molecular Biology Reports 39, 2893–2900.

Gotzek, D., Clarke, J., Shoemaker, D., 2010. Mitochondrial genome evolution in fire ants (Hymenoptera: Formicidae). BMC Evolutionary Biology 10, 300.

Haran, J., Timmermans, M.J., Vogler, A.P., 2013. Mitogenome sequences stabilize the phylogenetics of weevils (Curculionoidea) and establish the monophyly of larval ectophagy. Molecular Phylogenetics and Evolution 67, 156–166.

Hua, J.M., Li, M., Dong, P.Z., Cui, Y., Xie, Q., Bu, W.J., 2008. Comparative and phylogenomic studies on the mitochondrial genomes of Pentatomomorpha (Insecta: Hemiptera: Heteroptera). BMC Genomics 9, 610.

Hua, J.M., Li, M., Dong, P.Z., Cui, Y., Xie, Q., Bu, W.J., 2009. Phylogenetic analysis of the true water bugs (Insecta: Hemiptera: Heteroptera: Nepomorpha): Evidence from mitochondrial genomes. BMC Evolutionary Biology 9, 134.

Kim, K.-G., Hong, M.Y., Kim, M.J., Im, H.H., Kim, M.I., Bae, C.H., Seo, S.J., Lee, S.H., Kim, I., 2009. Complete mitochondrial genome sequence of the yellow-spotted long-horned beetle *Psacothea hilaris* (Coleoptera: Cerambycidae) and phylogenetic analysis among coleopteran insects. Molecules and Cells 27, 429–441.

Klopfstein, S., Vilhelmsen, L., Heraty, J.M., Sharkey, M., Ronquist, F., 2013. The hymenopteran tree of life: Evidence from protein-coding genes and objectively aligned ribosomal data. PLoS ONE 8, e69344.

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.

Lunt, D.H., Hyman, B.C., 1997. Animal mitochondrial DNA recombination. Nature 387, 247.

Masta, S.E., Boore, J.L., 2004. The complete mitochondrial genome sequence of the spider *Habronattus oregonensis* reveals rearranged and extremely truncated tRNAs. Molecular Biology and Evolution 21, 893–902.

Misof, B., Fleck, G., 2003. Comparative analysis of mt LSU rRNA secondary structures of odonates: Structural variability and phylogenetic signal. Insect Molecular Biology 12, 535–548.

Negrisolo, E., Minelli, A., Valle, G., 2004. The mitochondrial genome of the house centipede *Scutigera* and the monophyly versus paraphyly of myriapods. Molecular Biology and Evolution 21, 770–780.

Oliveira, D.C.S.G., Raychoudhury, R., Lavrov, D.V., Werren, J.H., 2008. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). Molecular Biology and Evolution 25, 2167–2180.

Ronquist, F., Rasnitsyn, A.P., Roy, A., Eriksson, K., Lindgren, M., 1999. Phylogeny of the Hymenoptera: A cladistic reanalysis of Rasnitsyn's (1988) data. Zoologica Scripta 28, 13–50.

San Mauro, D., Gower, D.J., Zardoya, R., Wilkinson, M., 2006. A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome. Molecular Biology and Evolution 23, 227–234.

Schnare, M.N., Damberger, S.H., Gray, M.W., Gutell, R.R., 1996. Comprehensive comparison of structural characteristics in eukaryotic cytoplasmic large subunit (23S-like) ribosomal RNA. Journal of Molecular Biology 256, 701–719.

Shao, R., Stephen, C.B., 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.

Shao, R., Campbell, N.J.H., Barker, S.C., 2001. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). Molecular Biology and Evolution 18, 858–865.

Sheffield, N.C., Song, H., Cameron, S.L., Whiting, M.F., 2008. A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. Molecular Biology and Evolution 25, 2499–2509.

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.

Simon, C., Buckley, T.R., Frati, F., Stewart, J.B., Beckenbach, A.T., 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 and Systematics 37, 545–579.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28, 2731–2739.

Wei, S.J., Shi, M., He, J.H., Sharkey, M.J., Chen, X.X., 2009. The complete mitochondrial genome of *Diadegma semiclausum* (Hymenoptera: Ichneumonidae) indicates extensive independent evolutionary events. Genome 52, 308–319.

Wei, S.J., Shi, M., Chen, X.X., Sharkey, M.J., van Achterberg, C., Ye, G.Y., He, J.H., 2010a. New views on strand asymmetry in insect mitochondrial genomes. PLoS ONE 5, e12708–1780.

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

Wei, S.J., Tang, P., Zheng, L.H., Shi, M., Chen, X.X., 2010c. 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.

Wei, D.D., Shao, R., Yuan, M.L., Dou, W., Barker, S.C., Wang, J.J., 2012. The multipartite mitochondrial genome of *Liposcelis bostrychophila*: Insights into the evolution of mitochondrial genomes in bilateral animals. PLoS ONE 7, e33973.

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

Xiao, J.H., Jia, J.G., Murphy, R.W., Huang, D.W., 2011. Rapid evolution of the mitochondrial genome in chalcidoid wasps (Hymenoptera: Chalcidoidea) driven by parasitic lifestyles. PLoS ONE 6, e26645.

Yan, D., Tang, Y., Xue, X., Wang, M., Liu, F., Fan, J., 2012. The complete mitochondrial genome sequence of the western flower thrips *Frankliniella occidentalis* (Thysanoptera: Thripidae) contains triplicate putative control regions. Gene 506, 117–124.

Zhang, D.X., Hewitt, G.M., 1997. Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. Biochemical Systematics and Ecology 25, 99–120.

Zhou, Z.J., Ye, H.Y., Huang, Y., Shi, F.M., 2010. The phylogeny of Orthoptera inferred from mtDNA and description of *Elimaea cheni* (Tettigoniidae: Phaneropterinae) mitogenome. Journal of Genetics and Genomics 37, 315–324.