Zoological Systematics, 39(2): 209–223 (April 2014), DOI: 10.11865/zs20140202

ORIGINAL ARTICLE

Sequencing and analysis of the complete mitochondrial genome of the giant dobsonfly *Acanthacorydalis orientalis* (McLachlan) (Insecta: Megaloptera: Corydalidae)

Yi-Ran Wang, Yu-Yu Wang, Wen-Cheng Chang, Xing-Yue Liu*, Ding Yang

Department of Entomology, China Agricultural University, Yuanmingyuan West Road 2, Beijing 100193, China
**Corresponding author, E-mail: liuxingyue@cau.edu.cn*

Abstract The complete mitochondrial genome of *Acanthacorydalis orientalis* (McLachlan) was determined and analyzed (GenBank accession number: KF840564). This paper represents the first mitochondrial (mt) genome of the dobsonfly genus *Acanthacorydalis*. The mt genome is a typical circular DNA of 15 753 bp composed of 37 genes with an A+T content of 76.7%. It has an ancestral gene arrangement of the insect mt genomes. Eleven of the 13 PCGs start with codon ATT and ATG, while several exceptions such as ATA and TTG respectively for *atp*8 and *nad*1 are also present. Five protein-coding genes end with a single T, while others have a termination codon of TAA or TAG. Most tRNAs are folded into the typical clover-leaf structure except for the *trnS*1 whose dihydrouridine arm was a simple loop. The secondary structure of *rrnl* consists of five structural domains and 50 helices, while the *rrns* includes three domains and 34 helices. The control region has a stretches of Ts with a length of 22 bp but lacks obvious tandem repeat region. Both Bayesian inference and maximum likelihood (ML) analyses, based on all 13 protein-coding genes and two rRNA genes of the mt genomes, confirm the monophyly of Corydalinae and suggest that *Acanthacorydalis*, together with *Corydalus*, which is an endemic dobsonfly genus from the New World, belong to a monophyletic lineage.

Key words *Acanthacorydalis orientalis*, Corydalidae, Megaloptera, mitochondrial genome.

1 Introduction

The mitochondrion in most eukaryotes is one of the important organelle, which is related to the energy metabolism, senescence and apoptosis of cells as well as disease (Nass & Nass, 1963). The mt genomes are widely used in the studies of phylogenetics and evolutionary biology due to some typical traits in most animals such as the stable gene content, relatively conserved gene arrangement, maternal inheritance and infrequent recombination (Wilson *et al.*, 1985; Wolstenholme, 1992). The insect mt genome has been considered to be a universal marker for species identification, phylogenetic inference and phylogeography (Avise, 1994; Wei & Chen, 2011).

The dobsonfly genus *Acanthacorydalis* van der Weele belongs to the subfamily Corydalinae (Megaloptera: Corydalidae). This genus is endemic to Asia, comprising eight species, six of which are recorded in China (Yang & Liu, 2010). The adults of *Acanthacorydalis* are among the largest and most bizarre-appearing living insects, and the males are noted for their disproportionately large mandibles (Liu *et al.*, 2005). The *Acanthacorydalis* larvae are exclusively aquatic

urn:lsid:zoobank.org:pub:52D32326-2E14-484F-B4CC-908E99FD4C95
Received 14 November 2013, accepted 3 April 2014.

and sensitive to the change of the water quality. *Acanthacorydalis orientalis* (McLachlan) is a Chinese endemic species and widely distributed from Southwestern to Northern China. This species is one of the well known megalopterans in China due to its large body-size and enlarged male mandibles. The adults of *A. orientalis* can be distinguished from other *Acanthacorydalis* species by the head with three yellowish spots before and behind ocellar triangle (Liu *et al.*, 2005). Traditionally, the *A. orientalis* larvae are used as food and medicinal material in some local regions of Southwestern China (Cao & Liu, 2013).

Hitherto, only five complete mt genomes of Megaloptera have been determined, including two dodsonfly species, *Corydalus cornutus* (L.) (Beckenbach & Stewart, 2009) and *Protohermes concolorus* Yang & Yang (Corydalidae: Corydalinae) (Hua *et al.*, 2009), two fishfly species, *Neochauliodes punctatolosus* Liu & Yang (Corydalidae: Chauliodinae) (Wang *et al.*, 2012) and *Neochauliodes bowringi* (McLachlan) (Li *et al.*, 2013), and one alderfly species, *Sialis hamata* Ross (Sialidae) (Cameron *et al.*, 2009). In this study, we determined the complete mt genome of *A. orientalis*, which represents the first mt genome of the genus *Acanthacorydalis*. We analyzed the genomic organization and arrangement, codon usage, and the secondary structure of tRNAs and rRNAs, and we also compared the mt genomes of *A. orientalis* with other sequenced mt genomes of Megaloptera.

2 Material and methods

2.1 Specimens

An adult specimen of *A. orientalis* was collected from Mt. Emei, Sichuan Province, China in July 2011. The specimen was preserved in 95% ethanol at -20℃. The voucher specimen was identified to be *A. orientalis* by Xing-Yue Liu, and it is a female adult, having typical specific characters, i.e. the head with three large yellow markings with of *A. orientalis*.

2.2 DNA extraction

We extracted the total genomic DNA from the muscle tissue of thorax of the specimen by TIANamp Genomic DNA kit (TIANGEN BIOTECH (Beijing) CO., LTD.) following manufacturer recommendations. The extracted DNA was then also stored at -20℃.

2.3 PCR amplification and sequencing

A series of universal mitochondrial primers were used for the insect mtDNA amplification of overlapping PCR (Simon *et al.*, 2006) (Table 1). According to the bridge gaps between some sequenced fragments, we designed remaining species-specific primers. The polymerase chain reaction was carried out with several cycling conditions including 30 s initial denaturation at 95℃, 40 cycles of 10 s denaturation at 95℃, 50 s annealing at 48℃, 2 min elongation based on the size of amplicons at 65℃ with a standard of 60 s/kb, and a final elongation step for 10 min at 65℃, finally holding at 10℃. Following the steps above we electrophoresed the PCR products in 1% agarose gel and make further accurate judgment on the quality of products by spectrophotometry. All fragments were sequenced in both directions using the BigDye Terminator Sequencing Kit (Applied Biosystems) and the ABI 3730XL Genetic Analyzer (PE Applied Biosystems, San Francisco, CA, USA) with two vector-specific primers and internal primers for primer walking.

2.4 Sequence assembly, annotation and bioinformatics analysis

Sequence assembly was done by using ContigExpress. tRNAs were identified by tRNAscan-SE Search Server v. 1.21 (Lowe & Eddy, 1997), while PCGs and rRNAs were identified by alignment with relevant genes of other species of Megaloptera with invertebrate mitochondrial genetic codes. The nucleotide composition and codon usage were analyzed by MEGA 5.0 (Tamura *et al.*, 2011). The secondary structure of tRNAs was generated by tRNAscan as stated, while that of rRNAs were predicted through the relevant genes from other species of insects (Cannone *et al.*, 2002; Gillespie *et al.*, 2006; Cameron & Whiting, 2008; Buckley *et al.*, 2000; Winterton *et al.*, 2007).

Table 1. Primer sequences used in this study.

No. fragment*	Primer ID	Nucleotide sequence (5'-3')	Reference
1	F17(SPB-14197)	GTAAAYCTACTTTGTTACGACTT	Simon *et al.*, 2006
	R17(SPB-14745)	GTGCCAGCAAYCGCGGTTATAC	Simon *et al.*, 2006
2	ACA-F	AGGGTATCTAATCCTAGTTT	Present study
	ACA-R	GAGGCTAAAGCTTGAGCTAGAAAAT	Present study
3	FL-34	GCCTGAAAAAGGGTTACTTTGATA	Simon *et al.*, 2006
	TR1284	ACARCTTTGAAGGYTAWTAGTTT	Simon *et al.*, 2006
4	ACA2.F	GTTATAACTGCTTTAATTAC	Present study
	ACA2.R	CAGAGTAGCTATGTTCCATT	Present study
5	F21(SPB-1709)	AATTGGTGGTTTTGGAAATTG	Simon *et al.*, 2006
	R21(SPB-2776)	GGTAATCAGAGTATCGACG	Simon *et al.*, 2006
6	F22(SPB-2756)	ACATTTTTTCCTCAACATTT	Simon *et al.*, 2006
	R22(SPB-3389)	TATTCATATCTTCAATATCATTGATG	Simon *et al.*, 2006
7	F01(SPA-2756)	ACATTTTTTCCTCAACATTT	Simon *et al.*, 2006
	R01(SPA-3665)	CCACAAATTTCTGAACACTG	Simon *et al.*, 2006
8	TF3399	ACAATTGGTCAYCAATGATAYTG	Simon *et al.*, 2006
	TR3796	ACTATTAGATGGTTTAAGAG	Simon *et al.*, 2006
9	TF3790	CATTAGATGACTGAAAGCAAGTA	Simon *et al.*, 2006
	TR4552	ATGTCCWGCAATYATATTWGC	Simon *et al.*, 2006
10	TF4463	TTTGCCCATCTWGTWCCNCAAGG	Simon *et al.*, 2006
	TR4908	CGAGTTAYATCTCGTCATCATTG	Simon *et al.*, 2006
11	F05(SPA-4792)	GTAGATGCAAGCCCTTGACC	Simon *et al.*, 2006
	R05(SPA-5731)	ATTGGATCAAATCCACATTC	Simon *et al.*, 2006
12	TF5470	GCAGCTGCYTGATAYTGRCA	Simon *et al.*, 2006
	TR6384	TATATTTAGAGYATRAYAYTGAAG	Simon *et al.*, 2006
13	TF5747	CCATTTGAATGTGGRTTTGAYCC	Simon *et al.*, 2006
	TR7211	TTAAGGCTTTAYTATTTATRTGYGC	Simon *et al.*, 2006
14	F08(SPA-7077)	TTAAATCCTTTGAGTAAAATCC	Simon *et al.*, 2006
	R08(SPA-7793)	TTAGGTTGAGATGGTTTAGG	Simon *et al.*, 2006
15	TF-J7572	AAAGGGAATTTGAGCTCTTTTWGT	Simon *et al.*, 2006
	TR-N8487	TATCAGSTAATATRGCWGCTCC	Simon *et al.*, 2006
16	ACA3.F	TTCCATGCTGACTATAAGAG	Present study
	ACA3.R	AAGAAGAGTTGGATTGTTAT	Present study
17	F10(SPA-8641)	CCAGAAGAACATAGCCCATG	Simon *et al.*, 2006
	R10(SPA-9629)	GTTTGTGAAGGTGTGTTGGG	Simon *et al.*, 2006
18	ACA5.F	TTATTACCTGAAAAACTCAA	Present study
	ACA5.R	TGGGGATTATAGAAAGGATG	Present study
19	F11(SPA-9648)	TCCCAACACACCTTCACAAAC	Simon *et al.*, 2006
	R11(SPA-11010)	TATCAACAGCAAATCCTCCTCA	Simon *et al.*, 2006
20	F23(SPC-10621)	CTCATACTGATGAAATTTTGGTTC	Simon *et al.*, 2006
	R23(SPC-11526)	TTCTACTGGTCGTGCTCCAATTCA	Simon *et al.*, 2006
21	F12(SPB-11335)	CATATTCAACCAGAATGATA	Simon *et al.*, 2006
	R12(SPB-12067)	AATCGTTCTCCATTTGATTTTGC	Simon *et al.*, 2006
22	F13(SPB-11876)	CGAGGTAAAGTACCACGTACTCA	Simon *et al.*, 2006
	R13(SPB-12595)	GTTGGATTTCTAACTTTATTRGARCG	Simon *et al.*, 2006
23	F14(SPB-12261)	TACCTCATAAGAAATAGTTTGAGC	Simon *et al.*, 2006
	R14(SPB-13000)	TTACCTTAGGGATAACAGCGTAA	Simon *et al.*, 2006
24	F15(SPB-12888)	CCGGTCTGAACTCAGATCATGTA	Simon *et al.*, 2006
	R15(SPB-13889)	ATTTATTGTACCTTTTGTATCAG	Simon *et al.*, 2006
25	F16(SPB-13342)	CCTTTGCACAGTCAAAATACTGC	Simon *et al.*, 2006
	R16(SPB-14220)	TTATGCACACATCGCCCGTC	Simon *et al.*, 2006

* The orientations is shown in Fig. 1.

2.5 Phylogenetic analysis

Five megalopteran species with available mt genomes were selected as ingroups. *Thyridosmylus langii* (McLachlan) (Neuroptera: Osmylidae) and *Mongoloraphidia harmandi* (Navás) (Raphidioptera: Raphidiidae) were selected as outgroups (Table 2).

Table 2. Taxa used in the present phylogenetic analysis.

Order	Family/Subfamily	Species	Accession number
	Corydalidae/Corydalinae	*Acanthacorydalis orientalis*	KF840564
	Corydalidae/Corydalinae	*Corydalus cornutus*	NC_011276
Megaloptera	Corydalidae/Corydalinae	*Protohermes concolorus*	NC_011524
	Corydalidae/Chauliodinae	*Neochauliodes punctatolosus*	NC_018772
	Sialidae	*Sialis hamata*	NC_013256
Neuroptera	Osmylidae	*Thyridosmylus langii*	NC_021415
Raphidioptera	Raphidiidae	*Mongoloraphidia harmandi*	NC_013251

The sequences of all 13 protein-coding genes and two rRNA genes were used, and the dataset was generated after alignment in MEGA 5.0 (Tamura *et al.*, 2011). A Bayesian inference was performed in MrBayes Version 3.1.2 (Ronquist & Huelsenbeck, 2003) with the GTR+I+G model estimated by Modeltest 3.7 (Posada & Crandall, 1998). Two simultaneous runs of 2 000 000 generations were conducted. The dataset was sampled every 200 generations with a burnin of 25%. A maximum likelihood analysis was conducted with PHYML online web server (Guindon & Gascuel, 2003; Guindon *et al*., 2005). In the ML algorithms, bootstrap values (BP) (Felsenstein, 1985) were calculated with 100 replicates.

3 Results and discussion

3.1 Genome organization and structure

The complete mt genome of *A. orientalis* is a typically circular, double-stranded molecule (GenBank accession number: KF840564) of 15 753 bp, which is similar in length with the other four sequenced megalopteran mt genomes, ranging from 15 608 bp to 15 851 bp (Beckenbach & Stewart, 2009; Wang *et al.*, 2012; Hua *et al.*, 2009; Cameron *et al.*, 2009). The mt genome of *A. orientalis* ranked the second largest one among the known mt genomes of Megaloptera. The mt genome of *A. orientalis* consists of 37 genes, including 22 tRNAs, 13 protein-coding genes, two rRNAs and a large A+T rich non-coding region as the putative control region (Fig. 1, Table 3). Considering the gene order, all of the 37 genes besides the non-coding region shows a conserved arrangement pattern, being consistent with that of *Drosophila yakuba* (Clary & Woletenholme, 1985), which has been hypothesized to be ancestral for insects (Boore, 1999). 23 genes are encoded on the majority strand (J-strand) while the other 14 genes are on the minority strand (N-strand). Gene overlaps were observed at 15 locations and involved a total of 45 bp. Two gene pairs including *atp*8-*atp*6 and *nad*4-*nad*4*l* overlap 7 nucleotides, also shared the sequence ATGNTAA which has been reported in many other insect mt genomes with a little difference appeared as ATGATAA and ATGTTAA respectively (Stewart & Beckenbach, 2005; Hua *et al.*, 2009). Furthermore, in spite of the difference in sequence, the longest overlap reaching 8 bp exited in two gene pairs including *trnW-trnC* and *trnY-cox*1, same to that in other three megalopteran genera, i.e. *Corydalus*, *Protohermes* and *Sialis*. In the mt genome of *Neochauliodes*, there are also two overlapping regions in *trnW-trnC* and *trnY-cox*1, with length of 9 bp and 14 bp respectively. Besides the large non-coding region, nine small non-coding intergenic spacers were found in this mt genome, with a range from 1–23 in size.

3.2 Protein-coding genes

Most of the 13 PCGs in the *A. orientalis* mt genome use the regular triplet start codons ATT and ATG but with exceptions in *atp*8 and *nad*1, which start with codons ATA and TTG, respectively. This phenomenon can also be observed in other Megaloptera species with some variations appeared in *atp8* which was reported to be ATT or ATC. Particularly, the *nad*1 of the *S. hamata* mt genome starts ATG. The termination codons are commonly TAA (*nad*2, *atp*8, *atp*6, *nad*4*l*,

*nad*6, *cob*, *nad*1) or TAG (*nad*3 only) for 8 of the 13 PCGs while the remaining five PCGs were stopped with T or TA, which has been found in many insect mt genomes and was considered to be completed via post-transcriptional polyadenylation (Ojala *et al.*, 1981). The common stop codons TAA or TAG could always be found overlapping several nucleotides within the down-stream tRNAs, which was presumed to act as "backup" to prevent translational read through if the transcripts were not properly cleaved (Boore, 2006).

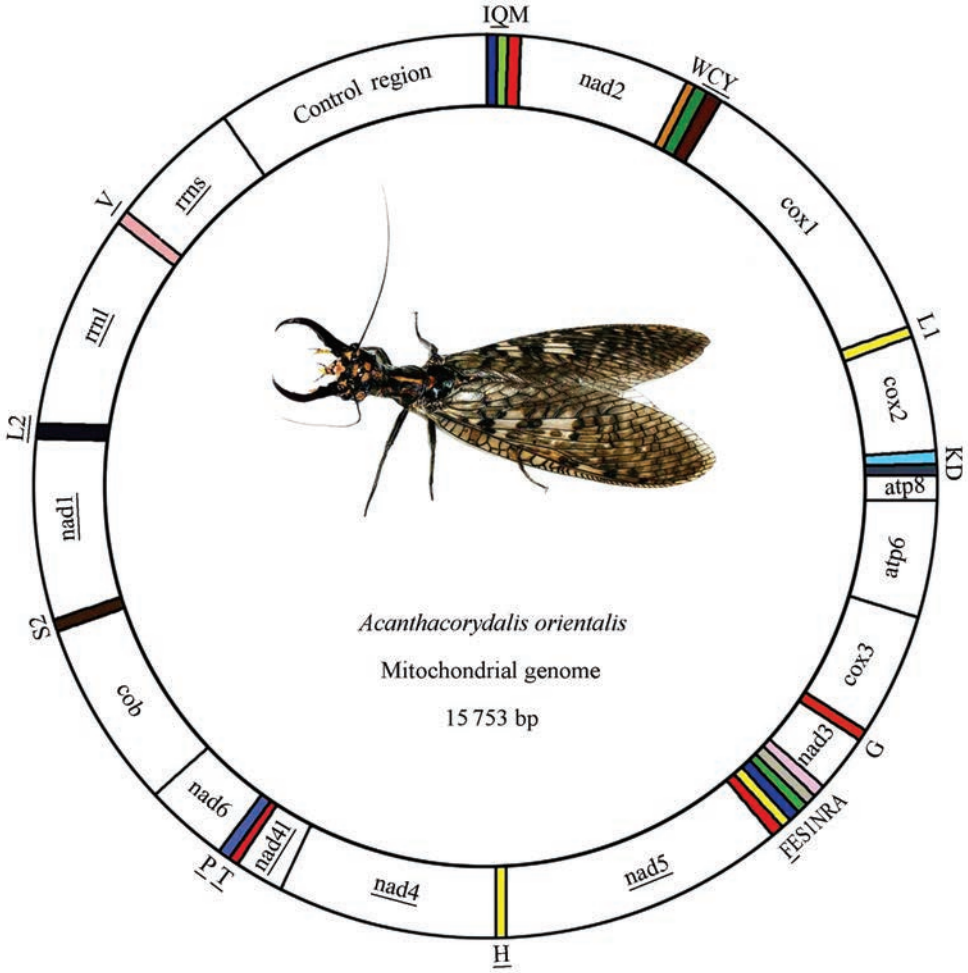

Fig. 1. Mitochondrial genome map of *Acanthacorydalis orientalis*. The tRNAs are denoted by the color blocks and are labeled according to the IUPACIUB single-letter amino acid codes. Gene name without underline indicates the direction of transcription from left to right, and with underline indicates right to left. Overlapping lines within the circle denote PCR fragments amplified used for cloning and sequencing.

Generally, the PCGs in the five sequenced mt genomes of Megaloptera do not greatly differed from each other (Table 4). Particularly, the start and stop codons showed remarkable consistency in *A. orientalis* and *C. cornutus* with an exception of *atp*8 which started with ATA and ATT respectively (Beckenbach & Stewart, 2009). This may provide evidence that the lineage of *Acanthacorydalis* and the lineage including *Corydalus* have close relationships, which was also supported by the intergeneric phylogeny of Corydalinae based on the morphological data (Contreras-Ramos, 2004, 2011). Among the PCGs, *atp*6, *nad*5, *nad*4 and *nad*4*l* are all of the same start and stop codons in these five megalopteran species, which may reflect the conservatism and stability of these four genes (Beckenbach & Stewart, 2009; Wang *et al.*, 2012; Hua *et al.*, 2009; Cameron *et al.*, 2009).

3.3 Transfer RNAs

All of the 22 typical tRNAs detected in the arthropod mt genomes were found in the *A. orientalis* mt genome with length ranging from 62 to 71 bp. Fourteen genes are encoded by the J-strand and the remains are encoded by the N-strand. Most tRNAs could be folded into the typical clover-leaf structure (Fig. 2), while the *trnS*1 is an exception for its DHU arm forms a simple loop same to other metazoan mt genomes (Wolstenholme, 1992).

Table 3. Organization of the *Acanthacorydalis orientalis* mt genome.

Gene	Direction	Location	Size (bp)	Anticodon[1)]	Start codon	Stop codon	Intergenic nucleotides (bp)[2)]
trnI	F	1–64	64	30–32 GAT			
trnQ	R	88–156	69	124–126 TTG			23
trnM	F	157–225	69	187–189 CAT			0
*nad*2	F	226–1245	1020		ATT	TAA	0
trnW	F	1244–1309	66	1274–1276 TCA			-2
trnC	R	1302–1365	64	1334–1336 GCA			-8
trnY	R	1365–1428	64	1396–1398 GTA			-1
*cox*1	F	1421–2960	1540		ATT	T-	-8
*trnL*1	F	2962–3026	65	2991–2993 TAA			1
*cox*2	F	3029–3710	682		ATG	T-	2
trnK	F	3712–3782	71	3742–3744 CTT			1
trnD	F	3782–3849	68	3811–3813 GTC			-1
*atp*8	F	3850–4008	159		ATA	TAA	0
*atp*6	F	4002–4676	675		ATG	TAA	-7
*cox*3	F	4676–5462	787		ATG	T-	-1
trnG	F	5463–5525	63	5493–5495 TCC			0
*nad*3	F	5526–5879	354		ATT	TAG	0
trnA	F	5878–5940	63	5907–5909 TGC			-2
trnR	F	5963–6025	63	5992–5994 TCG			22
trnN	F	6025–6090	66	6056–6058 GTT			-1
*trnS*1	F	6090–6158	69	6116–6118 GCT			-1
trnE	F	6158–6222	65	6188–6190 TTC			-1
trnF	R	6221–6283	63	6251–6253 GAA			-2
*nad*5	R	6284–8009	1726		ATT	T-	0
trnH	R	8010–8073	64	8041–8043 GTG			0
*nad*4	R	8074–9412	1339		ATG	T-	0
*nad*4*l*	R	9406–9699	294		ATG	TAA	-7
trnT	F	9702–9764	63	9732–9734 TGT			2
trnP	R	9765–9829	65	9797–9799 TGG			0
*nad*6	F	9832–10344	513		ATT	TAA	2
cob	F	10344–11480	1137		ATG	TAA	-1
*trn S*2	F	11479–11545	67	11508–11510TGA			-2
*nad*1	R	11558–12511	954		TTG	TAA	12
*trn L*2	R	12513–12574	62	12543–12545TAG			1
rrnl	R	12575–13882	1308				0
trnV	R	13883–13953	71	13918–13920 TAC			0
rrns	R	13954–14738	785				0
Control region		14739–15753	1015				0

1) Numbers indicate the anticodon position of each tRNA gene in the whole genome.
2) Negative numbers indicate overlapping nucleotides between adjacent genes.

3.4 Ribosomal RNAs

The length of *rrnl* and *rrns* of *A. orientalis* were sequenced to be 1 308 bp and 785 bp respectively. The boundaries of the rRNAs were assumed to extend to the boundaries of flanking genes without any obvious start or stop codons (Boore,

2001, 2006). Hence, by alignments with homologous sequences in other Megaloptera mt genomes, the *rrnl* are located between *trnL*2 and *trnV*, while the *rrns* are between *trnV* and the control region (Boore & Brown, 2000; Shao *et al.*, 2001; Wang & Lavrov, 2007).

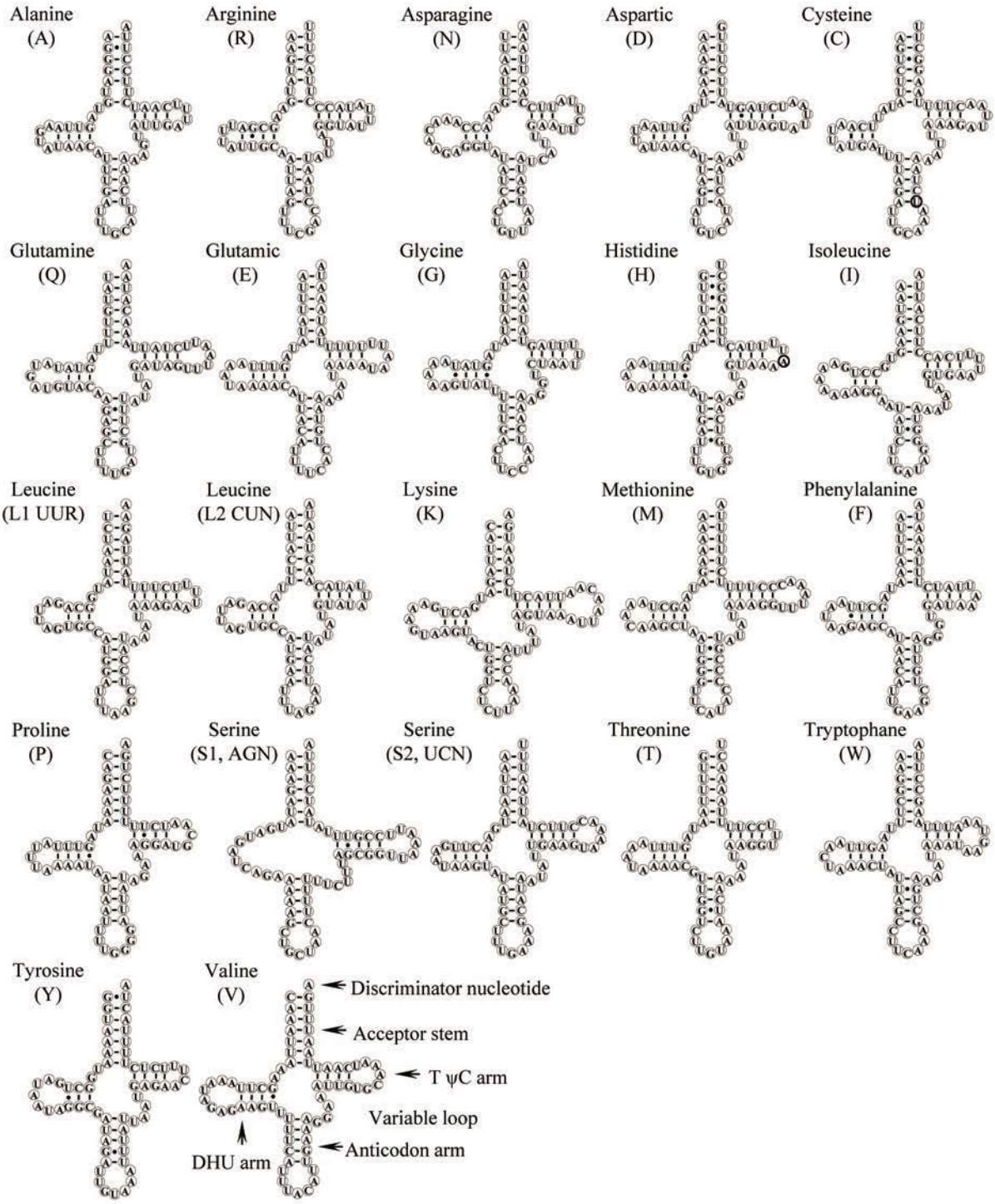

Fig. 2. Inferred secondary structure of 22 tRNAs of the *Acanthacorydalis orientalis* mt genome. The tRNAs are labeled with the abbreviations of their corresponding amino acids. Dash (-) indicates Watson-Crick bonds and dot (·) indicates GU bonds.

According to the published rRNAs secondary structures of an owlfly species *Libelloides macaionius* (Negrisolo *et al.*, 2011) and a fishfly species *N. punctatolosus* (Wang *et al.*, 2012), we inferred the secondary structures of *rrnl* and *rrns* of *A. orientalis*. The large ribosomal RNA was detected as a structure consists of five structural domains (I-II, IV-VI) and 50 helices with domain III absent, which is a typical characteristic in arthropods (Cannone *et al.*, 2002). The structure of the small ribosomal RNA in *A. orientalis* included three domains and 34 helices, which is in accordance with that of other species of Neuropterida (Figs 3–4).

Table 4. Start and stop codons of five megalopteran mt genomes.

Gene	Start codon					Stop codon				
	A	C	N	P	S	A	C	N	P	S
*nad*2	ATT	ATT	ATT	ATT	ATT	TAA	TAA	TAA	TGT	TAA
*cox*1	ATT	ATT	ATC	ATT	ATT	T-	T-	T-	TAA	T-
*cox*2	ATG	ATG	ATG	ATG	ATG	T-	T-	TAA	T-	T-
*atp*8	ATA	ATT	ATC	ATT	ATC	TAA	TAA	TAA	TAA	TAA
*atp*6	ATG	ATG	ATG	ATG	ATG	TAA	TAA	TAA	TAA	TAA
*cox*3	ATG	ATG	ATG	ATG	ATG	T-	T-	TAA	T-	TAA
*nad*3	ATT	ATT	ATT	ATT	ATT	TAG	TAG	TAG	T-	TAA
*nad*5	ATT	ATT	ATA	ATT	ATT	T-	T-	T-	T-	T-
*nad*4	ATG	ATG	ATG	ATG	ATG	T-	T-	T-	T-	T-
*nad*4*l*	ATG	ATG	ATG	ATG	ATG	TAA	TAA	TAA	TAA	TAA
*nad*6	ATT	ATT	ATA	ATT	ATT	TAA	TAA	TAA	TAA	TAA
cob	ATG	ATG	ATG	ATG	ATG	TAA	TAA	TAA	T-	T-
*nad*1	TTG	TTG	TTG	TTG	ATG	TAA	TAA	TAA	TAA	TAA

Note. A, *Acanthacorydalis orientalis*; C, *Corydalus cornutus*; N, *Neochauliodes punctatolosus*; P, *Protohermes concolorus*; S. *Sialis hamata*.

3.5 Non-coding regions

There were 10 non-coding regions detected in the *A. orientalis* mt genome with locations among tRNAs, PCGs and rRNAs. Most small regions are shorter than 20 bp, which is same to that in most arthropod mt genomes (Cook, 2005). The largest region with a size of 1015 bp is the so-called control region. There were two special small non-coding regions reaching length longer than 20 bp, which are located between *trnI* and *trnQ* as well as *trnA* and *trnR* with a size of 23 bp and 22 bp, respectively, and both two regions are rich in A+T content.

As for the largest non-coding region flanked by *rrns* and *trnI* (Fig. 1), we usually designate it as the putative control region. The size of the control region in the *A. orientalis* mt genome still ranks the second among all the sequenced mt genomes of Megaloptera (Beckenbach & Stewart, 2009; Wang *et al.*, 2012; Hua *et al.*, 2009; Cameron *et al.*, 2009). We found the high A+T content of 88.8% in the control region of the *A. orientalis* mt genome. It is necessary to point out that most insect mitochondrial control regions are composed of more than 85% of A+T and the highest is detected to be 96% so far (Zhang & Hewitt, 1997).

It has been reported that the control region contains four different motifs in some arthropods including tandem repeated sequences, a long sequence of Ts, a sub-region of even higher A+T content and the stem-loop structures, which contribute a lot to the role of the control region in the regulation of transcription and control of DNA replication (Cook, 2005; Clayton, 1982; Clayton, 1991). However, we did not find any obvious tandem repeat sequences in *A. orientalis* but stretches of Ts with a size of 22 bp, which is much longer than that in other megalopteran control regions.

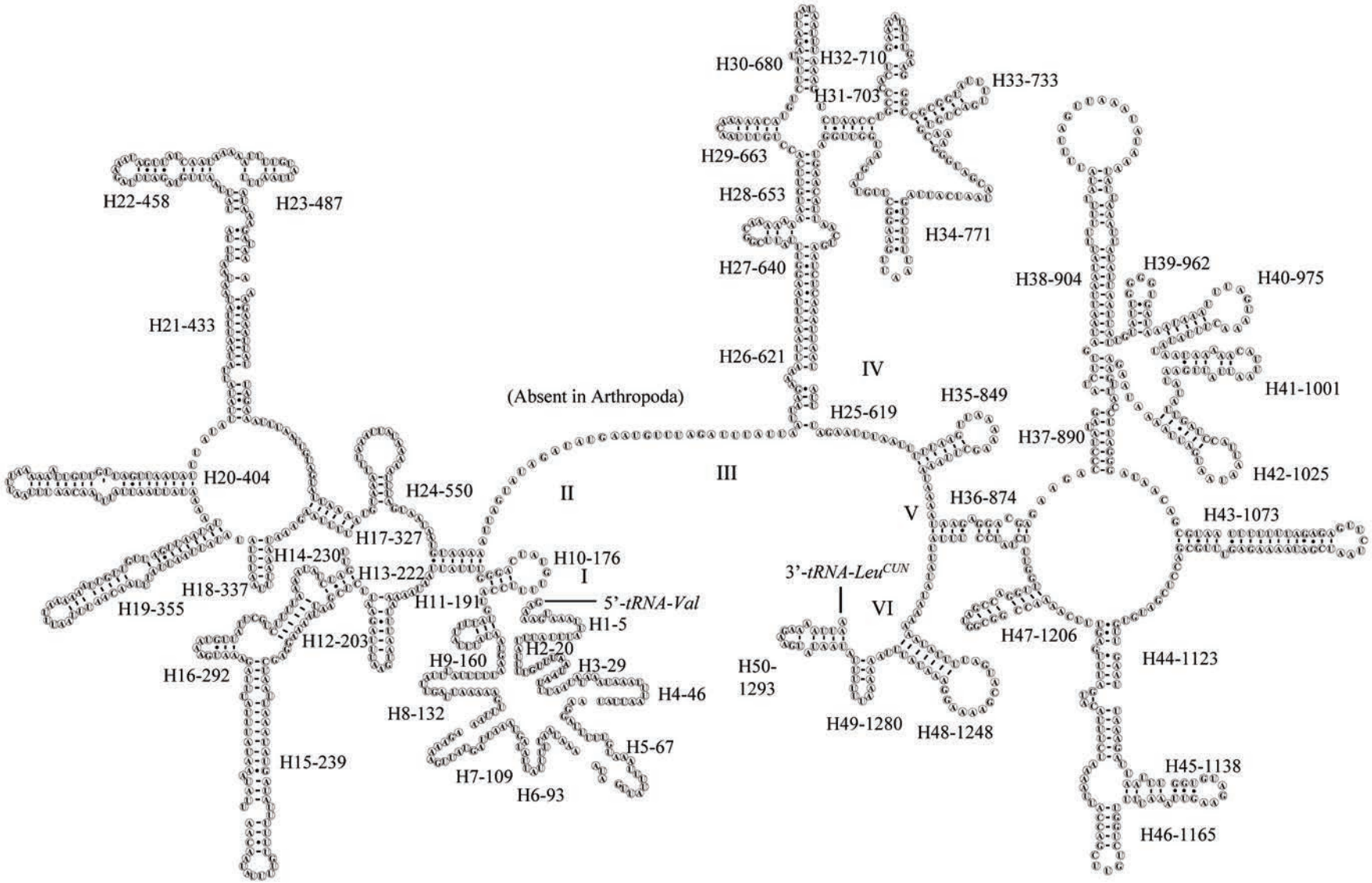

Fig. 3. Predicted secondary structure of the *rrnl* in the *Acanthacorydalis orientalis* mt genome. Roman numerals denote the conserved domain structure. Dash (-) indicates Watson-Crick base pairing and dot (•) indicates G-U base pairing.

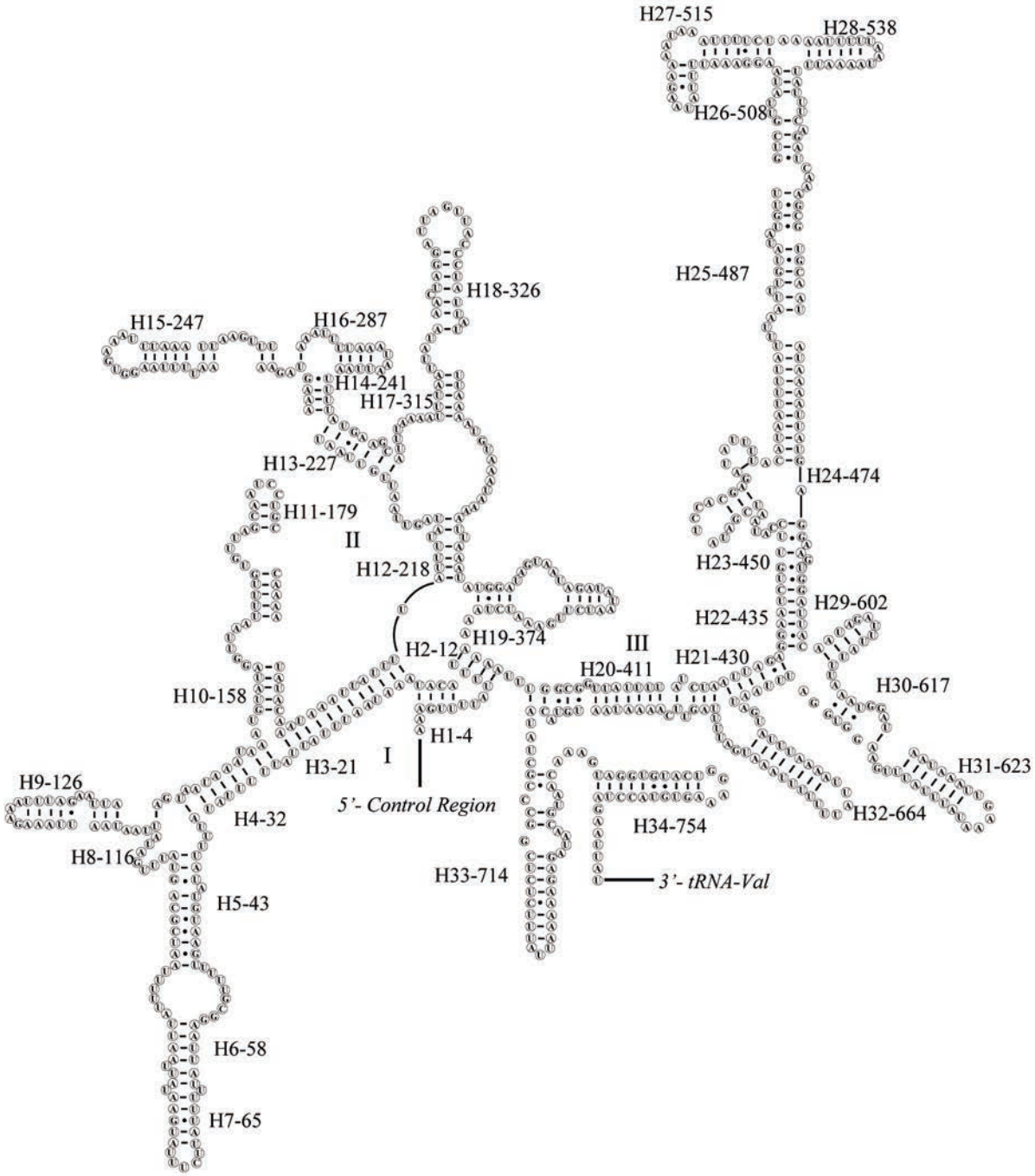

Fig. 4. Predicted secondary structure of the *rrns* in the *A. orientalis* mt genome. Roman numerals denote the conserved domain structure. Dash (-) indicates Watson-Crick base pairing and dot (•) indicates G-U base pairing.

3.6 Nucleotide composition and codon usage

The proportion of the base composition is unbalanced and presents a typical bias towards A and T with the content of 76.7% in the whole *A. orientalis* mt genome, which is the same as many other holometabolous insects (Hassanin *et al.*, 2005; Hassanin, 2006). It also appeared at a normal standard of A+T content in PCGs (74.9%), tRNAs (76.1%), rRNAs (81.1%) as well as the control region (88.8%). Besides the A+T content, the character of the nucleotide composition in the mt genomes can also be measured with AT- and GC-skew. In most metazoan mt genomes, this bias usually favors A and C which means positive AT-skew and negative GC-skew (Wei & Chen, 2011). But in the *A. orientalis* mt genome, it is

negative both in AT-skew (-0.009) and GC-skew (-0.216). Nevertheless, the content of A (38.0%) was not far from that of T (38.7%), therefore, we still consider it to be a normal base composition. In addition, both PCGs and rRNAs show negative AT-skew and positive GC-skew. The control region was detected as negative AT- and GC-skew which was contrary to tRNAs shown as positive AT- and GC-skew. We also measured the base composition of PCGs in both J-strand and N-strand, and the A+T content is 73.1% and 77.7% respectively while the A+T content of the third codon position is higher than the first and second codon positions (Table 5). The value of GC-skew is associated with replication orientation while that of AT-skew varies with gene direction, replication and codon positions. The strand bias of base composition may contribute to genome-wide studies of replication and transcription mechanisms (Wei *et al.*, 2010).

Table 5. Nucleotide composition of the *Acanthacorydalis orientalis* mt genome.

Feature	A%	T%	C%	G%	A+T%	AT-skew	GC-skew
Whole genome	38.0	38.7	14.1	9.1	76.7	-0.009	-0.216
Protein-coding genes	31.5	43.4	12.4	12.7	74.9	-0.159	0.012
First codon position	32.2	36	12.5	19.5	68.2	-0.056	0.219
Second codon position	19.8	47	18.7	14.6	66.8	-0.407	-0.123
Third codon position	42.5	47	6.1	4.1	89.5	-0.050	-0.196
PCGs J-strand	31.4	41.7	15.5	11.5	73.1	-0.141	-0.148
First codon position	32.6	32	16	19	64.6	0.009	0.086
Second codon position	20.3	45	21.5	13.2	65.3	-0.378	-0.239
Third codon position	41.1	48	9	2.2	89.1	-0.077	-0.607
PCGs N-strand	31.7	46	7.5	14.8	77.7	-0.184	0.327
First codon position	31.4	41	7	20.4	72.4	-0.133	0.489
Second codon position	18.9	50	14.2	16.7	68.9	-0.451	0.081
Third codon position	44.7	47	1.4	7.2	91.7	-0.025	0.674
tRNA genes	38.5	37.6	10.2	13.7	76.1	0.012	0.146
tRNA genes J-strand	38.7	37.3	11.5	12.5	76.0	0.018	0.042
tRNA genes N-strand	38.1	38.1	7.9	15.9	76.2	0.000	0.336
rRNA genes	40.1	41.0	6.3	12.6	81.1	-0.011	0.333
Control region	42.4	46.4	7.2	4.0	88.8	-0.045	-0.286

Analysis of nucleotide composition in the codon usage also reflected the A and T bias. A and T present preponderant compared to G and C. Moreover, NNG are the most infrequently used codons while NNU are the most frequently used codons. A considerable proportion of codons use NNU among all PCGs except for Leu (UUR), Met (AUN), Gln (CAN), Lys (AAN), Glu (GAN), Trp (UGN), Arg (CGN) and Gly (GGN) using codons ended with adenine. Measured from the protein genes of N-strand, NNU was still used more than NNA, but the two are the same in the J-strand. In addition, AT-rich codons were widely used such as TTT-Phe, TTA-Leu, ATT-Ile, ATA-Met, TAT-Tyr, AAT-Asn, AAA-Lys in order (Table 6).

3.7 Phylogeny

The Bayesian inference and ML algorithm generated fully resolved trees with the same topology (Fig. 5). Within Megaloptera, the monophyly of Corydalidae, as well as the monophyly of Corydalinae, were corroborated, being consistent with the previous studies on the higher phylogeny of Megaloptera (Wang *et al.*, 2012). Nonetheless, our present analysis mainly aims to unravel the phylogenetic status of the genus *Acanthacorydalis*. Based on the result, within Corydalinae, *Protohermes* is assigned to be the sister of a monophyletic clade composed of *Acanthacorydalis* and *Corydalus*. In the latest phylogenetic study on the intergeneric relationships within Corydalinae suggested that *Protohermes* together with *Neurhermes* are sister group of the remaining Corydalinae genera except *Choloroniella* (Contreras-Ramos, 2004, 2011). Our result based on the mt genomic data generally conforms to the phylogeny recovered from the morphological data. Due to lacking of majority of Corydalinae genera, *Acanthacorydalis* and *Corydalus*, both of which have sexual dimorphism of adult mandibles, are assigned to be sister groups, which however need further clarification by adding the missing genera of Corydalinae in future mt phylogenomic studies.

Table 6. Codon usage of the *Acanthacorydalis orientalis* mt genome.

Amino acid	Codon	N	RSCU	N+	RSCU	N-	RSCU
Phe (F)	UUU	**297**	**1.77**	**171**	**1.64**	**126**	**2**
	UUC	38	0.23	38	0.36	0	0
Leu (L)	UUA	**445**	**4.45**	**217**	**3.85**	**228**	**5.22**
	UUG	28	0.28	5	0.09	23	0.53
	CUU	**70**	**0.7**	**61**	**1.08**	**9**	**0.21**
	CUC	6	0.06	6	0.11	0	0
	CUA	49	0.49	48	0.85	1	0.02
	CUG	2	0.02	1	0.02	1	0.02
Ile (I)	AUU	**365**	**1.88**	**249**	**1.85**	**116**	**1.93**
	AUC	24	0.12	20	0.15	4	0.07
Met (M)	AUA	**231**	**1.84**	**112**	**1.88**	**119**	**1.8**
	AUG	20	0.16	7	0.12	13	0.2
Val(V)	GUU	**85**	**1.93**	40	1.7	**45**	**2.2**
	GUC	11	0.25	9	0.38	2	0.1
	GUA	75	1.7	**44**	**1.87**	31	1.51
	GUG	5	0.11	1	0.04	4	0.2
Ser (S)	UCU	**122**	**2.86**	**75**	**3.03**	**47**	**2.63**
	UCC	10	0.23	9	0.36	1	0.06
	UCA	87	2.04	61	2.46	26	1.45
	UCG	2	0.05	1	0.04	1	0.06
Pro (P)	CCU	**74**	**2.24**	**53**	**2.14**	**21**	**2.55**
	CCC	13	0.39	12	0.48	1	0.12
	CCA	40	1.21	30	1.21	10	1.21
	CCG	5	0.15	4	0.16	1	0.12
Thr (T)	ACU	**75**	**1.79**	**56**	**1.66**	**19**	**2.3**
	ACC	18	0.43	16	0.47	2	0.24
	ACA	74	1.76	63	1.87	11	1.33
	ACG	1	0.02	0	0	1	0.12
Ala (A)	GCU	**96**	**2.19**	**51**	**1.82**	**45**	**2.86**
	GCC	23	0.53	23	0.82	0	0
	GCA	53	1.21	37	1.32	16	1.02
	GCG	3	0.07	1	0.04	2	0.13
Tyr (Y)	UAU	**137**	**1.75**	**57**	**1.5**	**80**	**1.98**
	UAC	20	0.25	19	0.5	1	0.02
Stop (*)	UAA	7	1.75	5	1.67	2	2
	UAG	1	0.25	1	0.33	0	0
His (H)	CAU	**67**	**1.72**	**54**	**1.69**	**13**	**1.86**
	CAC	11	0.28	10	0.31	1	0.14
Gln (Q)	CAA	**68**	**1.86**	**51**	**2**	**17**	**1.55**
	CAG	5	0.14	0	0	5	0.45
Asn (N)	AAU	**164**	**1.81**	**109**	**1.74**	**55**	**1.96**
	AAC	17	0.19	16	0.26	1	0.04
Lys (K)	AAA	**69**	**1.55**	**44**	**1.87**	**25**	**1.19**
	AAG	20	0.45	3	0.13	17	0.81
Asp (D)	GAU	**53**	**1.58**	**34**	**1.45**	**19**	**1.9**
	GAC	14	0.42	13	0.55	1	0.1
Glu (E)	GAA	**74**	**1.78**	**47**	**1.92**	**27**	**1.59**
	GAG	9	0.22	2	0.08	7	0.41
Cys (C)	UGU	**33**	**1.89**	**12**	**1.85**	**21**	**1.91**
	UGC	2	0.11	1	0.15	1	0.09

Table 6 (continued)

Amino acid	Codon	N	RSCU	N+	RSCU	N-	RSCU
Trp (W)	UGA	**98**	**1.88**	**66**	**1.94**	**32**	**1.78**
	UGG	6	0.12	2	0.06	4	0.22
Arg (R)	CGU	20	1.43	12	1.33	**8**	**1.6**
	CGC	1	0.07	1	0.11	0	0
	CGA	**30**	**2.14**	**22**	**2.44**	**8**	**1.6**
	CGG	5	0.36	1	0.11	4	0.8
Ser (S)	AGU	38	0.89	20	0.81	18	1.01
	AGC	10	0.23	8	0.32	2	0.11
	AGA	71	1.67	**24**	**0.97**	**47**	**2.63**
	AGG	1	0.02	0	0	1	0.06
Gly (G)	GGU	66	1.16	38	1.14	28	1.19
	GGC	8	0.14	5	0.15	3	0.13
	GGA	**112**	**1.97**	**69**	**2.08**	**43**	**1.83**
	GGG	41	0.72	21	0.63	20	0.85

Note. N, total number in all proteins; N+, total number in J-strand; N-, total number in N-strand; RSCU, relative synonymous codon usage. Codon bolded: most commonly used codon for the amino acid. Codon underlined: cognate codon of tRNA for each amino acid.

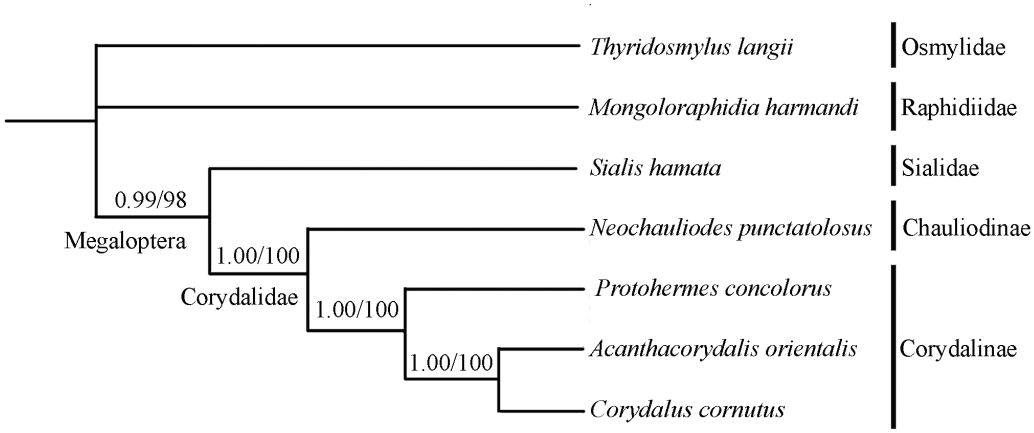

Fig. 5. Phylogenetic relationships among the sequenced Megaloptera insects. Numbers at the nodes are Bayesian posterior probabilities (left) and ML bootstrap values (right).

Funding This research was supported by the National Natural Science Foundation of China (31322051, 31320103902) and the Foundation for the Author of National Excellent Doctoral Dissertation of PR China (201178).

References

Avise, J. C. 1994. Molecular Markers, Natural History and Evolution. *Chapman and Hall*, New York.

Beckenbach, A. T. and Stewart, J. B. 2009. Insect mitochondrial genomics 3: the complete mitochondrial genome sequences of representatives from two neuropteroid orders: a dobsonfly (order Megaloptera) and a giant lacewing and an owlfly (order Neuroptera). *Genome*, 52: 31–38.

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

Boore, J. L. 2001. Complete mitochondrial genome sequence of the polychaete annelid. *Molecular Biology and Evolution*, 18: 1413–1416.

Boore, J. L. 2006. The complete sequence of the mitochondrial genome of *Nautilus macromphalus* (Mollusca: Cephalopoda). *BMC Genomics*, 7:182.

Boore, J. L. and Brown, W. M. 2000, Mitochondrial genomes of *Galathealinum*, *Helobdella*, and *Platynereis*: sequence and gene arrangement comparisons indicate that Pogonophora is not a phylum and Annelida and Arthropoda are not sister taxa. *Molecular Biology and Evolution*, 17(1): 87–106.

Buckley, T., Simon, C., Flook, P. and 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. and Whiting, M. F. 2008. The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta* (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. *Gene*, 408: 112–123.

Cameron, S. L., Sullivan, J., Song, H, Miller, K. B. and 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.

Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y, Feng, B, Lin, N, Madabusi, L. V., Müller, K. M., Pande, N., Shang, Z, Yu, N and Gutell, R. R. 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: 15.

Cao, C-Q and Liu, X-Y 2013. Description of the final-instar larva and pupa of *Acanthacorydalis orientalis* (McLachlan, 1899) (Megaloptera: Corydalidae) with some life history notes. *Zootaxa*, 3691: 145–152.

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

Clayton, D. A. 1982. Replication of animal mitochondrial DNA. *Cell*, 28: 693–705.

Clayton, D. A. 1991. Replication and transcription of vertebrate mitochondrial DNA. *Annual Review of Cell Biology*, 7:453–478.

Contreras-Ramos, A. 2004. Is the family Corydalidae (Neuropterida, Megaloptera) a monophylum. *Denisia*, 13: 135–140.

Contreras-Ramos, A. 2011. Phylogenetic review of dobsonflies of the subfamily Corydalinae and the genus *Corydalus* Latreille (Megaloptera: Corydalidae). *Zootaxa*, 2862: 1–38.

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

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39: 783–791.

Gillespie, J., Johnston, J., Cannone, J. and Gutell, R. 2006. Characteristics of the nuclear (18S, 5.8 S, 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.

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

Guindon, S., Lethiec, F., Duroux, P. and Gascuel, O. 2005. PHYML Online–a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Research*, 33: W557–W559.

Hassanin, A., Léger, N. and Deutsch, J. 2005. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of Metazoa, and consequences for phylogenetic inferences. *Systematic Biology*, 54: 277–298.

Hassanin, A. 2006. Phylogeny of Arthropoda inferred from mitochondrial sequences: strategies for limiting the misleading effects of multiple changes in pattern and rates of substitution. *Molecular Phylogenetics and Evolution*, 38: 100–116.

Hua, J-M, Li, M, Dong, P, Xie, Q and Bu, W-J 2009. The mitochondrial genome of *Protohermes concolorus* Yang *et* Yang, 1988 (Insecta: Megaloptera: Corydalidae). *Molecular Biology Reports*, 36: 1757–1765.

Li, Q, Wei, S-J, Shi, M and Chen, X-X 2013. Complete mitochondrial genome of *Neochauliodes bowringi* (MacLachlan)(Megaloptera: Corydalidae). *Mitochondrial DNA*: DOI: 10.3109/19401736.2013.803542.

Liu, X-Y, Yang, D, Ge, S-Q and Yang, X-K 2005. Phylogenetic review of the Chinese species of *Acanthacorydalis* (Megaloptera, Corydalidae). *Zoologica Scripta*, 34: 373–387.

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

Nass, M. M. K. and Nass, S. 1963. Intramitochondrial fibers with DNA characteristics. *Journal of Cell Biology*, 19 (3): 593–611.

Negrisolo, E., Babbucci, M. and Patarnello, T. 2011. The mitochondrial genome of the ascalaphid owlfly *Libelloides macaronius* and comparative evolutionary mitochondriomics of neuropterid insects. *BMC Genomics*, 12: 221.

Ojala, D., Montoya, J. and Attardi, G. 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature*, 290: 470–474.

Posada, D. and Crandall, K. A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics*, 14: 817–818.

Ronquist, F. and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19: 1572–1574.

Shao, R, Campbell, N. J., Schmidt, E. R. and Barker, S. C. 2001. Increased rate of gene rearrangement in the mitochondrial genomes of three orders of hemipteroid insects. *Molecular Biology and Evolution*, 18(9): 1828–1832.

Simon, C., Buckley, T. R., Frati, F., Stewart, J. B. and 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, Evolution, and Systematics*, 37: 545–579.

Stewart, J. B. and Beckenbach, A. T. 2005. Insect mitochondrial genomics: the complete mitochondrial genome sequence of the meadow spittlebug *Philaenus spumarius* (Hemiptera: Auchenorrhyncha: Cercopoidae). *Genome*, 48(1):46–54.

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

Wang, X and Lavrov, D. V. 2007. Mitochondrial genome of the homoscleromorph *Oscarella carmela* (Porifera, Demospongiae) reveals unexpected complexity in the common ancestor of sponges and other animals. *Molecular Biology and Evolution*, 24(2):363–373.

Wang, Y-Y, Liu, X-Y, Shaun, L. W. and Yang, D 2012. The first mitochondrial genome for the fishfly subfamily Chauliodinae and implications for the higher phylogeny of Megaloptera. *PLoS ONE*, 7 (10): e47302.

Wei, S-J, Shi, M, Chen, X-X, Sharkey, M. J., van Achterberg, C., Ye, G-Y and He, J-H 2010. New views on strand asymmetry in insect mitochondrial genomes. *PLoS ONE*, 5 (9): e12708.

Wei, S-J and Chen, X-X 2011. Progress in research on the comparative mitogenomics of insects. *Chinese Journal of Applied Entomology*, 48(6): 1573–1585.

Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U., Helm-Bychowski, K. M., Higuchi, R. G., Palumbi, S. R., Prager, E. M. and Sage, R. D. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society*, 26 (4): 375–400.

Winterton, S. L., Wiegmann, B. M. and Schlinger, E. I. 2007. Phylogeny and Bayesian divergence time estimations of small-headed flies (Diptera: Acroceridae) using multiple molecular markers. *Molecular Phylogenetics and Evolution*, 43: 808–832.

Wolstenholme, D. R. 1992. Animal mitochondrial DNA: structure and evolution. *International Review of Cytology: A Survey of Cell Biology*, 141(1): 173–216.

Yang, D and Liu, X-Y 2010. Fauna Sinica Insecta Vol.51 Megaloptera. Science Press, Beijing. 457 pp, XIV pl. (Monograph, in Chinese).

Zhang, D-X and Hewitt, G. M. 1997. Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. *Biochemical Systematics and Ecology*, 25 (2): 99–120.