



Complete mitochondrial genome sequence of *Cheirotonus jansonii* (Coleoptera: Scarabaeidae)

L.L. Shao¹, D.Y. Huang^{1,2}, X.Y. Sun³, J.S. Hao^{1,3}, C.H. Cheng¹, W. Zhang¹ and Q. Yang³

¹College of Life Sciences, Anhui Normal University, Wuhu, China

²College of Forestry, Jiangxi Environmental Engineering Vocational College, Ganzhou, China

³Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing, China

Corresponding authors: J.S. Hao / Q. Yang

E-mail: jshaonigpas@sina.com / qunyang@nigpas.ac.cn

Genet. Mol. Res. 13 (1): 1047-1058 (2014)

Received January 14, 2013

Accepted July 15, 2013

Published February 20, 2014

DOI <http://dx.doi.org/10.4238/2014.February.20.6>

ABSTRACT. We sequenced the complete mitochondrial genome (mitogenome) of *Cheirotonus jansonii* (Coleoptera: Scarabaeidae), an endangered insect species from Southeast Asia. This long legged scarab is widely collected and reared for sale, although it is rare and protected in the wild. The circular genome is 17,249 bp long and contains a typical gene complement: 13 protein-coding genes, 2 rRNA genes, 22 putative tRNA genes, and a non-coding AT-rich region. Its gene order and arrangement are identical to the common type found in most insect mitogenomes. As with all other sequenced coleopteran species, a 5-bp long TAGTA motif was detected in the intergenic space sequence located between *trnS*(UCN) and *nad1*. The atypical *cox1* start codon is AAC, and the putative initiation codon for the *atp8* gene appears to be GTC, instead of the frequently found ATN. By sequence comparison, the 2590-bp long non-coding AT-rich region is the second longest among the coleopterans, with two tandem repeat regions: one is 10 copies of an 88-bp sequence and the other is 2 copies

of a 153-bp sequence. Additionally, the A+T content (64%) of the 13 protein-coding genes is the lowest among all sequenced coleopteran species. This newly sequenced genome aids in our understanding of the comparative biology of the mitogenomes of coleopteran species and supplies important data for the conservation of this species.

Key words: Mitochondrial genome; Coleoptera; Scarabaeidae; *Cheirotonus jansoni*

INTRODUCTION

Metazoan mitochondrial DNA (mtDNA) occurs as a double-strand, circular molecule, ranging in size approximately 14–42 kb that encodes 13 PCGs, 2 rRNA genes, and 22 tRNA genes (Boore, 1999). Additionally, it contains the AT-rich region of insects, which serves as the source of length variation in the genome (Inohira et al., 1997), and includes information essential for the initiation of transcription and DNA replication (Brown et al., 1986; Shadel and Clayton, 1997; Saito et al., 2005). DNA-sequence data from the mitogenome are being used with increasing frequency to estimate phylogenetic relationships between animal taxa. Besides, mtDNA exhibits several properties that make it a valuable tool in the study of population genetics, molecular evolution, and even conservation genetics, as a result of its relatively simpler genetic structure, maternal inheritance, and higher rate of evolution compared with nuclear DNA (Moritz et al., 1987).

The insect order Coleoptera contains over 380,000 described species divided into four suborders: Adephaga, Archostemata, Myxophaga, and Polyphaga (Lawrence and Newton Jr., 1982). Present-day diversity among the four suborders is highly skewed toward the Polyphaga. Up to the present, the complete or nearly complete nucleotide sequences of coleopteran mtDNA have been determined in nearly 60 species, but this number is relatively small considering the size and diversity of the group. *Cheirotonus jansoni*, a member of the coleopterans (subfamily Euchirinae, family Scarabaeidae, suborder Polyphaga) is mainly distributed throughout Southeast Asia, with its body length as long as 65 mm and body width about 35 mm. Owing to its large body size, beautiful color and rareness, it has attracted many people to collect them illegally for sale and other purposes, and now it has been listed as one of the Class two national protected species in China. Therefore, it seems essential to gather genetic information about this species to determine its genetic diversity, as well as for its demographic population management (Wang et al., 2001). In this respect, full-length mitogenome data are one of the most important sources of this knowledge.

In this study, we determined the complete mitogenome sequence of *C. jansoni* and analyzed it by comparison with other mitogenomes of insects, especially the coleopterans. This newly sequenced mitogenome of Scarabaeidae is expected to be valuable in enhancing our understanding of the genetic entities of this species, as well as to clarify the phylogeny, molecular evolution and comparative genetics of the mitogenome of Scarabaeidae.

MATERIAL AND METHODS

Sample collection and DNA extraction

An adult *C. jansoni* specimen was collected from Jiu Lianshan Mountain, Jiangxi Prov-

ince, China, in August 2009. The fresh material was preserved in 100% ethanol solution and stored at -20°C. DNA was extracted using a DNeasy tissue kit in accordance with the manufacturer protocol (Sangon, China). Only the muscle attached to the thorax was utilized for DNA extraction.

Primer design, polymerase chain reaction (PCR), and sequencing

To sequence the entire mitogenome of *C. jansoni*, a range of Coleoptera mitochondrial primers from Timmermans et al. (2010) and a pair of universal insect mitochondrial primers from Simon et al. (1994) were used. Species-specific primers were then designed on the basis of sequenced fragments to bridge the gaps. All in all, four short fragments smaller than 800 bp and nine long fragments of 1.5-3.8 kb were amplified covering the whole mitochondrial genome. PCR of short fragments with ExTaq DNA polymerase (TaKaRa, China) was carried out under the following conditions: initial denaturation for 5 min at 95°C, followed by 35 cycles of 30 s at 94°C (denaturation), 40 s at 50°-55°C (annealing), and 60 s at 72°C (extension), with a subsequent final 10-min extension at 72°C. Long-PCR cycling conditions using the LA Taq DNA polymerase (TaKaRa) were initial denaturation step at 95°C for 5 min, followed by 30 cycles of denaturation at 95°C for 50 s, primer annealing at 50°C for 50 s and elongation at 68°C for 150 s during the first 15 cycles, and then an additional 5 s per cycle during the last 15 cycles. The final elongation step was continued at 68°C for 10 min.

PCR products were examined, purified and sequenced directly. The sequencing primers were the same as PCR amplification primers, and a subsequent sequencing was performed by the primer-walking strategy for long-PCR products.

Genome annotation and analysis

Sequences obtained were assembled and data regarding the sequence were determined using BioEdit version 4.8.9 (Hall, 1999). Protein-coding genes and ribosomal RNA genes were identified by comparing their similarity to published insect mitochondrial sequences using ClustalX 1.8 (Thompson et al., 1997) and MEGA 5.0 (Tamura et al., 2011) softwares. Both the *rrnL*- and *rrnS*-predicted secondary structures were drawn according to models proposed for these genes in other insects (Gillespie et al., 2006; Cameron and Whiting, 2008; Chai and Du, 2012). The tRNA genes were identified using tRNAscan-SE Search Server v.1.21 (Lowe and Eddy, 1997). The tRNAs not found by tRNAscan-SE were identified through comparison of *C. jansoni* nucleotide sequence with the regions coding these tRNAs in other insects. All tRNAs were folded by hand, using tRNAscan-SE output as template when possible. The AT-rich regions were determined via the alignment of the sequences with homologous regions of known full-length insect mitogenome sequences and the tandem repeats in the AT-rich region were predicted by the Tandem Repeats Finder available online (<http://tandem.bu.edu/trf/trf.html>) (Benson, 1999). The complete mtDNA sequence of *C. jansoni* was deposited in GenBank under accession No. KC428100.

RESULTS AND DISCUSSION

Genome structure, organization, and composition

The mitogenome of *C. jansoni* was determined to be 17,249 nucleotides in length. It

codes for 13 putative proteins, two ribosomal RNAs, and 22 tRNAs (Figure 1 and Table 1). Besides, it contains a 2590-bp long non-coding AT-rich region that is the second longest one (shorter than that of *Rhopaea magnicornis*, 2827 bp) among all the coleopteran mitogenomes (Sheffield et al., 2009). The gene arrangement and order is the same as that of most insect mitogenomes (Stewart and Beckenbach, 2003; Friedrich and Muqim, 2003; Bae et al., 2004; Sheffield et al., 2008, 2009; Timmermans et al., 2010; Song et al., 2010; Wan et al., 2012; Timmermans and Vogler, 2012), and shares the proposed Pancrustacea ancestral organization (Boore et al., 1998).

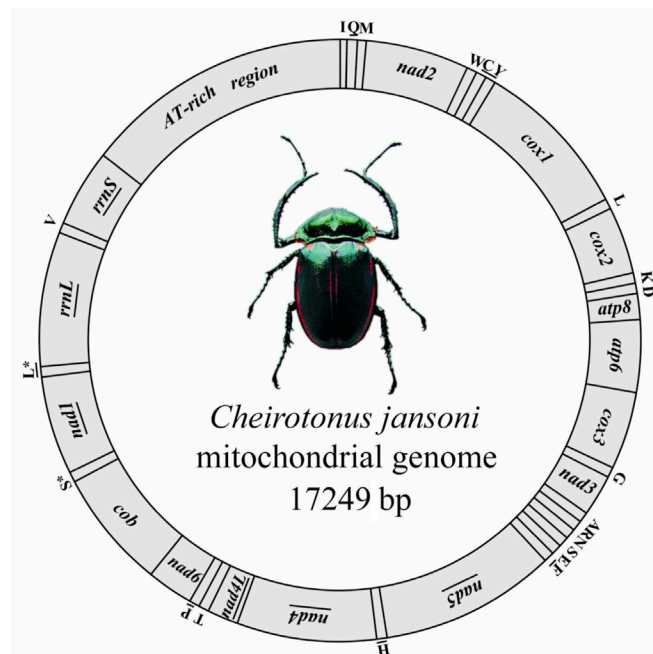


Figure 1. Map of the circular mitogenome of *Cheirotonus jansonii*. *cox1-3* stands for cytochrome oxidase subunits, *cob* for cytochrome b, *nad1-6* for NADH dehydrogenase components, and *atp6* and *atp8* for ATP synthase subunits 6 and 8 genes. The tRNAs are denoted as one-letter symbol according to the IUPAC-IUB single-letter amino acid codes. The one-letter symbols L, L*, S and S* denote *trnL*(CUN), *trnL*(UUR), *trnS*(AGN), and *trnS*(UCN), respectively. Gene names without underlines indicate the direction of transcription from left to right, and names with underlines are read from right to left.

The A+T content of the genome (65.7%) was found to fall within the range observed so far for insects, from the lowest (64.8%) in *Japyx solifugus* (Diplura; Carapelli et al., 2005) to the highest (86.7%) in *Melipona bicolor* (Hymenoptera; Silvestre D and Arias MC, unpublished data from NCBI). Thus, its G+C content (34.3%) is one of the highest among insect mitogenomes, just slightly lower than that of *Japyx solifugus* (35.2%). Furthermore, its A+T content is the lowest among the coleopterans determined, from *Lucanus mazama* (67.1%) to *Sphaerius* sp (80.8%) (data not shown). The tRNA genes, rRNA genes, non-coding sequences, and third codon position of protein-coding genes (PCGs) showed a relatively higher A+T-rich content (71.4, 71.0, 68.9, and 66.0%, respectively), while the first and second codon positions of PCGs showed a lower A+T content (61.5 and 64.5%) than that of the whole genome average (Table 2).

Table 1. Summary of mitochondrial genome of *Cheirotonus jansoni*.

Gene	Direction	Begins	Ends	Length	Spacer	Anti	Start	Stop
<i>trnI</i>	F	1	67	67	-3	GAT		
<i>trnQ</i>	R	133	65	69	-1	TTG		
<i>trnM</i>	F	133	201	69	0	CAT		
<i>nad2</i>	F	202	1209	1008	6		ATT	TAA
<i>trnW</i>	F	1216	1282	67	-8	TCA		
<i>trnC</i>	R	1337	1275	63	11	GCA		
<i>trnY</i>	R	1411	1349	63	1	GTA		
<i>cox1</i>	F	1413	2943	1531	0		AAC	T
<i>trnL</i> (UUR)	F	2944	3007	64	0	TAA		
<i>cox2</i>	F	3008	3695	688	0		ATA	T
<i>trnK</i>	F	3696	3766	71	0	CTT		
<i>trnD</i>	F	3767	3830	64	0	GTC		
<i>atp8</i>	F	3831	3986	156	-7		GTC	TAG
<i>atp6</i>	F	3980	4651	672	-1		ATG	TAA
<i>cox3</i>	F	4651	5437	787	0		ATG	T
<i>trnG</i>	F	5438	5502	65	0	TCC		
<i>nad3</i>	F	5503	5856	354	-2		ATT	TAG
<i>trnA</i>	F	5855	5918	64	0	TGC		
<i>trnR</i>	F	5919	5984	66	0	TCG		
<i>trnN</i>	F	5985	6054	70	0	GTT		
<i>trnS</i> (AGN)	F	6055	6121	67	0	TCT		
<i>trnE</i>	F	6122	6184	63	5	TTC		
<i>trnF</i>	R	6252	6190	63	0	GAA		
<i>nad5</i>	R	7966	6253	1714	-3		ATA	T
<i>trnH</i>	R	8030	7964	67	-1	GTG		
<i>nad4</i>	R	9367	8030	1338	-7		ATG	TAA
<i>nad4L</i>	R	9651	9361	291	2		ATG	TAA
<i>trnT</i>	F	9654	9716	63	-1	TGT		
<i>trnP</i>	R	9780	9716	65	1	TGG		
<i>nad6</i>	F	9782	10282	501	-1		ATA	TAA
<i>cob</i>	F	10282	11424	1143	-2		ATG	TAA
<i>trnS</i>	F	11423	11487	65	19	TGA		
<i>nad1</i>	R	12457	11507	951	0		ATT	TAA
<i>trnL</i> (CUN)	R	12519	12458	62	0	TAG		
<i>rrnL</i>	R	13813	12520	1294	0			
tRNA ^{Val}	R	13882	13814	69	0	TAC		
<i>rrnS</i>	R	14659	13883	777	0			
<i>AT-rich region</i>	R	17249	14660	2590				

Direction of the genes is presented as F for forward and R for reverse.

Table 2. Nucleotide composition for features within the *Cheirotonus jansoni* mitochondrial genome.

Genome feature	Number of positions	Proportion of nucleotides			
		T	C	A	G
All nucleotides ^a	17249	28.84	23.02	36.81	11.32
rRNA genes	2071	31.39	20.33	39.59	8.69
A+T-rich region ^a	2590	29.54	24.05	35.60	10.81
tRNA genes	1446	32.64	17.22	38.80	11.34
Unassigned nucleotides ^a	45	31.11	28.89	37.78	2.22
Protein-coding genes	11103	36.05	18.97	27.93	17.05
All codons					
First codon position ^b	3701	30.15	16.81	31.34	21.70
Second codon position ^b		44.96	20.13	19.51	15.40
Third codon position ^b		33.05	19.97	32.94	14.05
Major strand coded protein genes ^b	6819	30.53	24.95	30.94	13.58
Major strand codons	2273				
First codon position ^b		24.51	21.12	33.88	20.50
Second codon position ^b		43.29	23.36	19.84	13.51
Third codon position ^b		23.80	30.36	39.11	6.73
Minor strand coded protein genes ^b	4284	44.84	9.45	23.13	22.57
Minor strand codons					
First codon position ^b	1428	39.15	9.94	27.31	23.60
Second codon position ^b		47.62	14.99	18.98	18.42
Third codon position ^b		47.76	3.43	23.11	25.70

^aThe major coding strand sequence was used for this analysis. ^bStop codons and incomplete stop codons excluded for this analysis.

C. jansonii showed a compact mitochondrial genome, displaying a 2590-bp AT-rich non-coding region, identified by the position between the *rrnS* and *trnI-trnQ-trnM* gene cluster, which is supposed to contain the replication origin site. Another 45 non-coding nucleotides are spread over 7 regions, ranging in size from 1 to 19 bp, with the longest located between *trnS*(UCN) and *nad1* (Table 1). Similarly sized intergenic spacer sequences (16-28 bp) located between *trnS*(UCN) and *nad1* have been detected in the majority of coleopterans, although *Rhagophthalmus lufengensis* and *R. ohbai* are unusually shorter (Li et al., 2007). Furthermore, a 5-bp long motif sequence (TAGTA) has been detected in all sequenced coleopterans. Additionally, Cameron and Whiting (2008) previously detected the similar 7-bp motif (TTAGTAT) in several insect orders. Within the intergenic spacers, the 5- or 7-bp consensus sequence was suggested as the possible binding site for mtTERM, the transcription termination peptide (Taanman, 1999). Genes overlap in the mitogenome totaling 37 bp, which are spread over 12 regions, ranging from 1 to 8 bp, with the longest three being 8, 7 and 7 bp, located between *trnW* and *trnC*, between *atp8* and *atp6*, and between *nd4* and *nd4L*, respectively (Table 1).

Protein-coding genes and codon usage

Excluding the stop codons, the whole genome contains 3701 protein codons or 11,103 coding nucleotides. The TTT (Phe), TTA (Leu), ATT (Ile), and ATA (Met) are the four most frequently used, accounting for 21.26% (Table 3). The synonymous codon bias, namely the relative synonymous codon usage (RSCU) revealed that codons harboring A or T in the third position are more frequently used than those of G or C in the third position. However, the RSCU analysis showed that the values are relatively higher than those found in other detected insect species, for some codons harboring G or C in the third position (bold ones in Table 3) (Kim et al., 2009; Sheffield et al., 2009), which is congruent with the relatively low A+T content (64%) in the PCGs. Thus, considering each codon position separately, the A+T content is not so biased toward the third codon position as in some other beetles (Table 2) (Friedrich and Muqim, 2003; Stewart and Bechenbach, 2003; Cameron and Whiting, 2008; Sheffield et al., 2008, 2009; Song et al., 2010; Wan et al., 2012).

Table 3. Codon usage of the protein-coding genes of the *Cheirotonus jansonii* mitogenome.

Codon (aa)	N (RSCU)	Codon (aa)	N (RSCU)	Codon (aa)	N (RSCU)	Codon (aa)	N (RSCU)
UUU (F)	211 (1.33)	UCU (S)	72 (1.76)	UAU (Y)	111 (1.3)	UGU (C)	28 (1.3)
UUC (F)	106 (0.67)	UCC (S)	30 (0.73)	UAC (Y)	60 (0.7)	UGC (C)	15 (0.7)
UUA (L)	204 (2.05)	UCA (S)	64 (1.57)	UAA (*)	0 (0)	UGA (W)	69 (1.44)
UUG (L)	105 (1.06)	UCG (S)	14 (0.34)	UAG (*)	0 (0)	UGG (W)	27 (0.56)
CUU (L)	79 (0.8)	CCU (P)	49 (1.44)	CAU (H)	35 (0.9)	CGU (R)	13 (0.91)
CUC (L)	52 (0.52)	CCC (P)	39 (1.15)	CAC (H)	43 (1.1)	CGC (R)	9 (0.63)
CUA (L)	114 (1.15)	CCA (P)	33 (0.97)	CAA (Q)	49 (1.53)	CGA (R)	20 (1.4)
CUG (L)	42 (0.42)	CCG (P)	15 (0.44)	CAG (Q)	15 (0.47)	CGG (R)	15 (1.05)
AUU (I)	210 (1.35)	ACU (T)	57 (1.03)	AAU (N)	93 (1.19)	AGU (S)	36 (0.88)
AUC (I)	101 (0.65)	ACC (T)	70 (1.27)	AAC (N)	63 (0.81)	AGC (S)	15 (0.37)
AUA (M)	162 (1.48)	ACA (T)	81 (1.47)	AAA (K)	69 (1.3)	AGA (S)	60 (1.47)
AUG (M)	57 (0.52)	ACG (T)	13 (0.24)	AAG (K)	37 (0.7)	AGG (S)	36 (0.88)
GUU (V)	78 (1.41)	GCU (A)	72 (1.38)	GAU (D)	44 (1.26)	GGU (G)	35 (0.62)
GUC (V)	17 (0.31)	GCC (A)	65 (1.25)	GAC (D)	26 (0.74)	GGC (G)	28 (0.49)
GUA (V)	100 (1.81)	GCA (A)	58 (1.12)	GAA (E)	53 (1.38)	GGA (G)	83 (1.46)
GUG (V)	26 (0.47)	GCG (A)	13 (0.25)	GAG (E)	24 (0.62)	GGG (G)	81 (1.43)

aa = amino acid; N = frequency of codon used; RSCU = relative synonymous codon usage; *stop codon. Start codons and stop codons were excluded in total codon counts. Bold = values relatively higher than those found in other detected insect species, for some codons harboring G or C in the third position.

All PCGs except for *cox1* and *atp8* were found to have a putative ATR methionine or ATT isoleucine codons as start signals. Abnormal translational starts of *cox1* have been reported in some species of diverse insect orders, including Diptera, Lepidoptera, and Orthoptera. Generally, AAA (lysine), ATY (isoleucine), CTA (leucine), AAY (asparagine), and some other amino acids have all been proposed as possible start codons in Coleoptera as shown by Kim et al. (2009) and Sheffield et al. (2008, 2009). By aligning the sequence region encompassing *trnY* and *cox1* from all known beetle mitogenomes, the AAC (asparagine) is supposed to be the start codon for *cox1* in *C. jansoni*.

As for the *atp8*, one putative ATA initiator located 9 bp downstream of *trnD* is one of the codons generally found in insects; the other possible initiator is GTC, which is located adjacent to that downstream of *trnD*, and this codon has not been found to serve as a starter in insects. However, a putative GTT initiator was detected for the *nad1* genes in mouse and rat (Bibb et al., 1981; Gadaleta et al., 1989), and for the *cox3* gene in *Ascaris suum* (Okimoto et al., 1990); in addition, GTG was supposed to be the *nad5* initiator in *Microchorista philpotti* and *nad1* initiator in *Bittacus pilicornis* (Beckenbach, 2011). In this study, GTC was found to be the homologous site of the above species, and thus proposed to be the ATP8 initiator. Nine of the 13 PCGs use the conventional stop codon TAA (*nad2*, *atp6*, *nad1*, *nad4*, *nad4L*, *nad6*, *cob*) or TAG (*atp8*, *nad3*). The remaining four use single T as their stop codons, which are adjacent to the tRNAs. This phenomenon is characteristic of animal mitochondrial genes (Bibb et al., 1981; Clary and Wolstenholme, 1985). This single T stopper would be completed as TAA by its mRNA polyadenylation (Anderson et al., 1981).

Ribosomal and transfer RNA genes

All tRNA genes formed typical clover-leaf structures, except for *trnS*(AGN), by the dihydrouridine arm of which a simple loop could be formed (Figure 2). Such abnormalities are also observed in many metazoan *trnS*(AGN)s, including insects (Wolstenholme, 1992). Among beetles, Adephaga, Archostemata (except the species *Macrogyrus oblongus*), and Myxophaga harbor the common GCU anticodon for *trnS*(AGN); however, all polyphagan species reported harbor the uncommon UCU anticodon, which suggests that this particular anticodon may be a possible molecular synapomorphy for Polyphaga (Sheffield et al., 2008). All tRNAs have a total of 26 mismatches in the stem region: 21 U-G, 1 A-A, 1 C-C, 1 U-U, 1 A-G, and 1 U-C.

As is the case with all other insects, two rRNA genes (*rrnL* and *rrnS*) were detected in the *C. jansoni* mitogenome. They are located between *trnL* (CUN) and *trnV*, and between *trnV* and the AT-rich region, respectively (Table 1). The size and A+T content of the two rRNAs were also well within the corresponding ranges of other coleopteran mitochondrial rRNAs (data not shown). *rrnL* contained six domains (labeled I, II, III, IV, V, and VI) with 49 helices, while the *rrnS* harbored three domains (labeled I, II, III) with 33 helices (Figures 3 and 4) in the secondary structure, respectively. The morphological characteristics of both *rrnL* and *rrnS* are quite similar to their counterparts in *Apis mellifera*, *Manduca sexta*, and other insect species (Gillespie et al., 2006; Cameron and Whiting, 2008; Chai and Du, 2012).

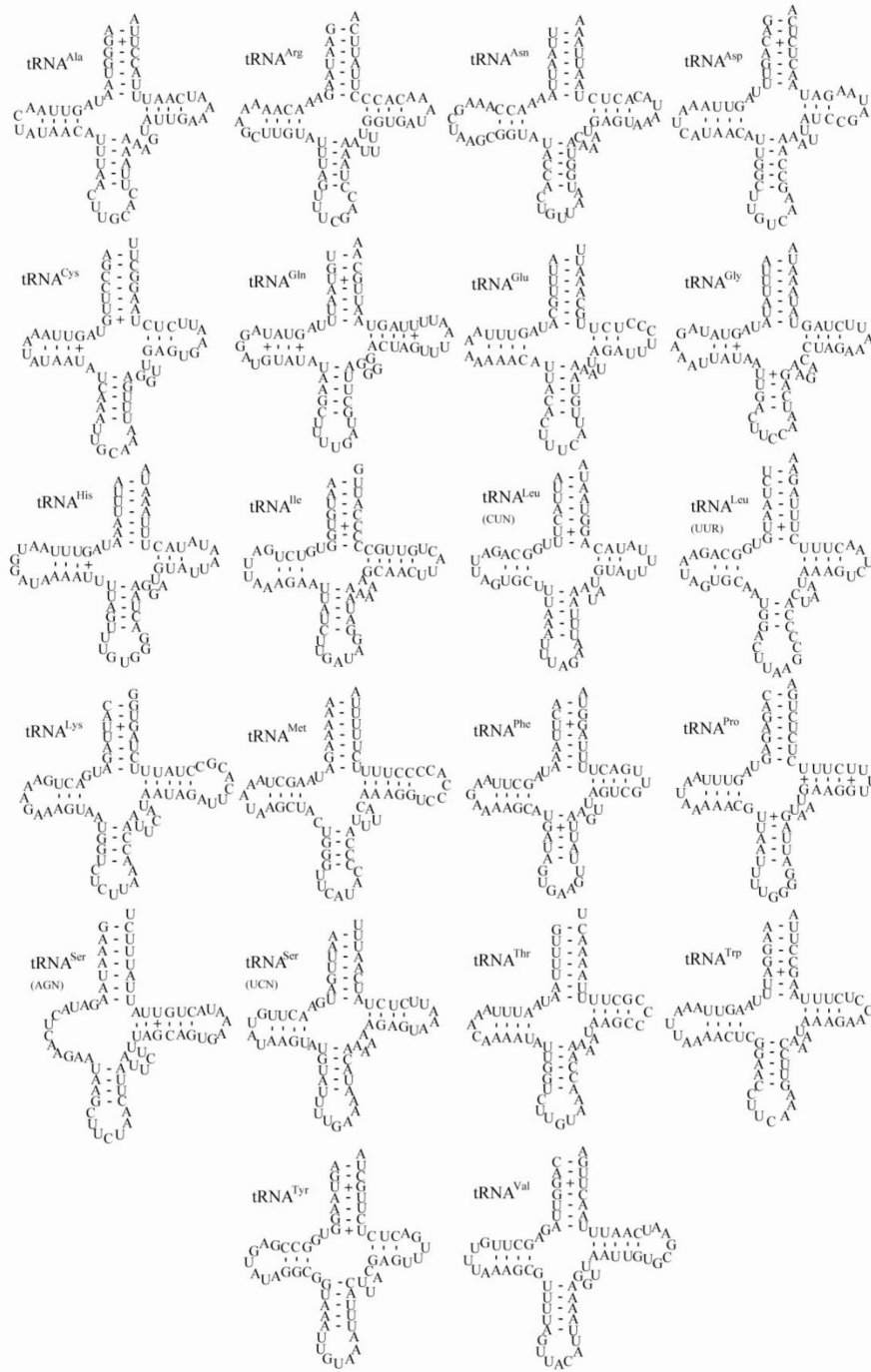


Figure 2. Predicted secondary structures of *Cheirotonus jansoni* mitochondrial tRNAs.

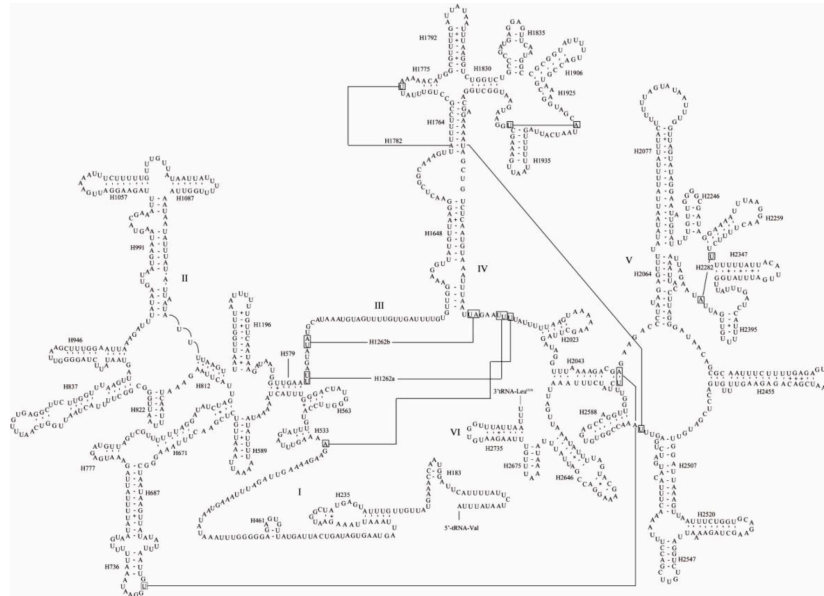


Figure 3. Predicted secondary structure of the *rrnL* gene in the *Cheirotonus jansoni* mitogenome. Tertiary interactions and base triples are shown connected by continuous lines. Base-pairing is indicated as follows: Watson-Crick pairs and wobble GU pairs by lines, other noncanonical pairs by plus signs.

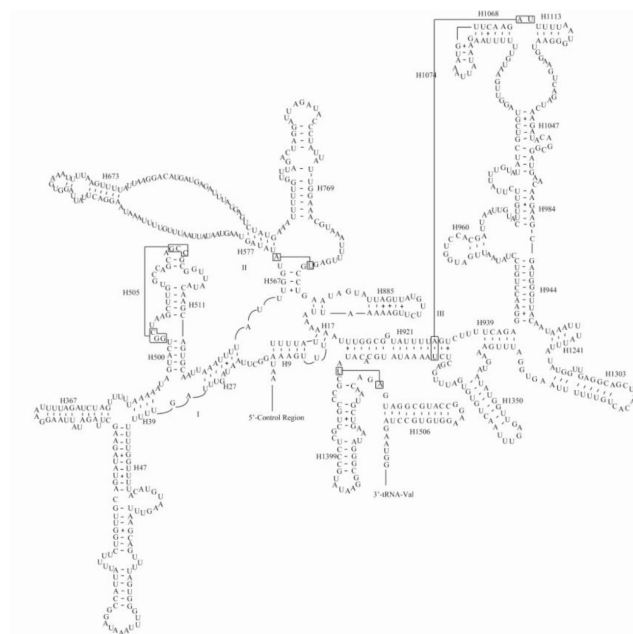


Figure 4. Predicted secondary structure of the *rrnS* gene in the *Cheirotonus jansoni* mitogenome. Tertiary interactions and base triples are shown connected by continuous lines. Base-pairing is indicated as follows: Watson-Crick pairs and wobble GU pairs by lines, other noncanonical pairs by plus signs.

AT-rich region

In general, the AT-rich region of the insect mitogenome is the longest non-coding stretch conservatively located between the *rrnS* gene and the *trnI-trnQ-trnM* gene cluster, and it has the highest AT content in the whole genome. However, the 2590-bp AT-rich region of *C. jansonii* does not possess the highest AT content throughout the whole genome, and its sequence shows little similarity compared with that of other beetles. Previous studies have also shown that size and sequence similarity vary significantly across coleopterans (Zhang and Hewitt, 1997). In addition, ten copies of a 88-bp and two copies of a 153-bp long tandem repeat abutting each other were detected (Figure 5), like those reported in other insect groups, including beetles. For example, 21 copies of a 58-bp sequence have been detected in *Trachypachus holmbergi* (Sheffield et al., 2008) and seven copies of a 57-bp sequence in *Psacotheta hilaris* (Kim et al., 2009). In addition, the A+T content of all these repeats is lower (50.5%) than that of the remaining parts of the AT-rich region (77.5%).

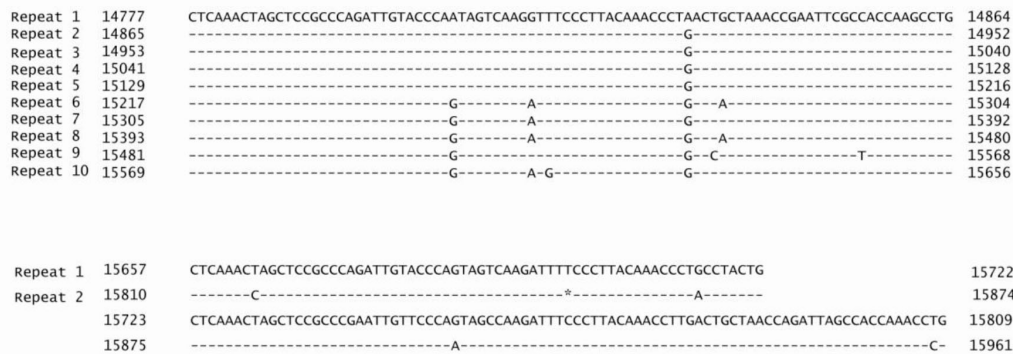


Figure 5. Alignment of tandem repeat units detected in the *Cheirotonus jansonii* AT-rich region. Dots indicate sequences identical to the repeat 1. The nucleotide position is provided at each end of the sequence with regard to the mitogenome of *Cheirotonus jansonii*.

There were two regions with inverted repeats capable of forming base-paired stem structures: one located at coordinates 15783 to 15773, pairing with the sequence at coordinates 15765 to 15755; the other located at coordinates 15935 to 15925 pairing with the sequence at coordinates 15917 to 15907. Two stem-and-loop structures were found located within the tandem repeat region as reported in *Psacotheta hilaris* (Kim et al., 2009). Unlike that observed in *Drosophila*, these 29-bp putative pairings would be quite short (Clary and Wolstenholme, 1987) with their function unknown. Besides the two tandem repeats, the majority of the AT-rich region is composed of non-repetitive sequences, including two poly-T stretches (one 17 bp, the other 13 bp), as well as several microsatellite-like TA repeats.

CONCLUSION

Our study presents the first mitogenome sequence of Euchirinae (Scarabaeidae) beetle. The results indicate that this mitogenome follows the hypothetical ancestral insect mode

without deletions or duplications. It harbors several common features that many other beetle lineages share, such as a 19-bp non-coding region between *nd1* and *trnL*(CUN), a noncanonical AAC start codon for the *cox1* gene, and an uncommon UCU anticodon for *trnS*(AGN). It also indicates that several features differ from those of other coleopterans: GTC start codon for *atp8*, the lowest A+T content of the whole genome, and relatively higher C+G content of two tandem repeats in the AT-rich region. The newly sequenced *Cheirotonus jansoni* mitogenome is expected to be useful in enhancing our understanding of the comparative biology of the coleopteran mitogenomes and supplying important data for the conservation of this species.

ACKNOWLEDGMENTS

Research supported by the CAS/SAFEA International Partnership Program for Creative Research Teams, Chinese Academy of Sciences (Grant #KZCX22YW2JC104), the Opening Funds from the State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences (Grant #104143), and the Program for Innovative Research Team in Anhui Normal University.

REFERENCES

- Anderson S, Bankier AT, Barrell BG, de Bruijn MH, et al. (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290: 457-465.
- Bae JS, Kim I, Sohn HD and Jin BR (2004). The mitochondrial genome of the firefly, *Pyrocoelia rufa*: complete DNA sequence, genome organization, and phylogenetic analysis with other insects. *Mol. Phylogenet. Evol.* 32: 978-985.
- Beckenbach AT (2011). Mitochondrial genome sequences of representatives of three families of scorpionflies (Order Mecoptera) and evolution in a major duplication of coding sequence. *Genome* 54: 368-376.
- Benson G (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27: 573-580.
- Bibb MJ, Van Etten RA, Wright CT, Walberg MW, et al. (1981). Sequence and gene organization of mouse mitochondrial DNA. *Cell* 26: 167-180.
- Boore JL (1999). Animal mitochondrial genomes. *Nucleic Acids Res.* 27: 1767-1780.
- Boore JL, Lavrov DV and Brown WM (1998). Gene translocation links insects and crustaceans. *Nature* 392: 667-668.
- Brown GG, Gadaleta G, Pepe G, Saccone C, et al. (1986). Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. *J. Mol. Biol.* 192: 503-511.
- Cameron SL and Whiting MF (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.
- Carapelli A, Nardi F, Dallai R and Boore JL (2005). Relationships Between Hexapods and Crustaceans Based on Four Mitochondrial Genes. CRC Press, Boca Raton, 295-306.
- Chai HN and Du YZ (2012). The complete mitochondrial genome of the pink stem borer, *Sesamia inferens*, in comparison with four other noctuid moths. *Int. J. Mol. Sci.* 13: 10236-10256.
- Clary DO and Wolstenholme DR (1985). The mitochondrial DNA molecular of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *J. Mol. Evol.* 22: 252-271.
- Clary DO and Wolstenholme DR (1987). *Drosophila* mitochondrial DNA: conserved sequences in the A + T-rich region and supporting evidence for a secondary structure model of the small ribosomal RNA. *J. Mol. Evol.* 25: 116-125.
- Friedrich M and Muqim N (2003). Sequence and phylogenetic analysis of the complete mitochondrial genome of the flour beetle *Tribolium castaneum*. *Mol. Phylogenet. Evol.* 26: 502-512.
- Gadaleta G, Pepe G, De Candia G, Quagliariello C, et al. (1989). The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. *J. Mol. Evol.* 28: 497-516.
- Gillespie JJ, Johnston JS, Cannone JJ and Gutell RR (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 Mol. Biol.* 15: 657-686.

- Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41: 95-98.
- Inohira K, Hara T and Matsuura ET (1997). Nucleotide sequence divergence in the A+T-rich region of mitochondrial DNA in *Drosophila simulans* and *Drosophila mauritiana*. *Mol. Biol. Evol.* 14: 814-822.
- Kim KG, Hong MY, Kim MJ, Im HH, et al. (2009). Complete mitochondrial genome sequence of the yellow-spotted longhorned beetle *Psacotha hilaris* (Coleoptera: Cerambycidae) and phylogenetic analysis among coleopteran insects. *Mol. Cells* 27: 429-441.
- Lawrence JF and Newton AF Jr (1982). Evolution and classification of beetles. *Annu. Rev. Ecol. Syst.* 13: 261-290.
- Li X, Ogoh K, Ohba N, Liang X, et al. (2007). Mitochondrial genomes of two luminous beetles, *Rhagophthalmus lufengensis* and *R. ohbai* (Arthropoda, Insecta, Coleoptera). *Gene* 392: 196-205.
- Lowe TM and Eddy SR (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955-964.
- Moritz C, Dowling TE and Brown WM (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annu. Rev. Ecol. Syst.* 18: 269-292.
- Okimoto R, Macfarlane JL and Wolstenholme DR (1990). Evidence for the frequent use of TTG as the translation initiation codon of mitochondrial protein genes in the nematodes, *Ascaris suum* and *Caenorhabditis elegans*. *Nucleic Acids Res.* 18: 6113-6118.
- Saito S, Tamura K and Aotsuka T (2005). Replication origin of mitochondrial DNA in insects. *Genetics* 171: 1695-1705.
- Shadel GS and Clayton DA (1997). Mitochondrial DNA maintenance in vertebrates. *Annu. Rev. Biochem.* 66: 409-435.
- Sheffield NC, Song H, Cameron SL and Whiting MF (2008). A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. *Mol. Biol. Evol.* 25: 2499-2509.
- Sheffield NC, Song H, Cameron SL and Whiting MF (2009). Nonstationary evolution and compositional heterogeneity in beetle mitochondrial phylogenomics. *Syst. Biol.* 58: 381-394.
- Simons C, Frati F, Beckenbach A, Crespi B, et al. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* 87: 651-701.
- Song HJ, Sheffield NC, Cameron SL and Miller KB (2010). When phylogenetic assumptions are violated: base compositional heterogeneity and among-site rate variation in beetle mitochondrial phylogenomics. *Syst. Entomol.* 35: 429-448.
- Stewart JB and Beckenbach AT (2003). Phylogenetic and genomic analysis of the complete mitochondrial DNA sequence of the spotted asparagus beetle *Crioceris duodecimpunctata*. *Mol. Phylogenet. Evol.* 26: 513-526.
- Taanman JW (1999). The mitochondrial genome: structure, transcription, translation and replication. *Biochim. Biophys. Acta* 1410: 103-123.
- Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, et al. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882.
- Timmermans MJ and Vogler AP (2012). Phylogenetically informative rearrangements in mitochondrial genomes of Coleoptera, and monophyly of aquatic elateriform beetles (Dryopoidea). *Mol. Phylogenet. Evol.* 63: 299-304.
- Timmermans MJ, Dodsworth S, Culverwell CL, Bocak L, et al. (2010). Why barcode? High-throughput multiplex sequencing of mitochondrial genomes for molecular systematics. *Nucleic Acids Res.* 38: e197.
- Wan XL, Hong MY, Liao AM and Kim MI (2012). Complete mitochondrial genome of a carabid beetle, *Damaster mirabilissimus mirabilissim* (Coleoptera: Carabidae). *Entomol. Res.* 42: 44-54.
- Wang JB, Hu CL and Xu HF (2001). Applications of mitochondrial DNA variability analysis in zoological conservation biology. *Biodiversity Sci.* 9: 181-187.
- Wolstenholme DR (1992). Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.* 141: 173-216.
- Zhang DX and Hewitt GM (1997). Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. *Biochem. Syst. Ecol.* 25: 99-120.