Article

Chinese Science Bulletin

Entomology

September 2012 Vol.57 No.27: 3578–3585
doi: 10.1007/s11434-012-5271-3

Complete sequence and gene organization of the mitochondrial genome of *Batocera lineolata* Chevrolat (Coleoptera: Cerambycidae)

WANG ChengYe, FENG Ying* & CHEN XiaoMing

Key Laboratory of Cultivating and Utilization of Resource Insects of State Forestry Administration, Research Institute of Resource Insects, Chinese Academy of Forestry, Kunming 650224, China

Received December 26, 2011; accepted April 19, 2012

Batocera lineolata Chevrolat, 1852 is an important forest pest that is found mainly in China, Vietnam, India, and Japan. The complete mitochondrial genome (mitogenome) sequence of *B. lineolata* was determined by long polymerase chain reaction (PCR) and conserved primer walking approaches. The results showed that the entire mitogenome is 15418 bp long with 74.48% A+T content. The positions and arrangement of the 37 genes encoded by the genome are identical to the mitogenomes of two other longhorn beetles for which the complete gene content and arrangement are publicly available. All protein-coding genes start with the ATN codon that is a typical initiation codon in insects. All transfer RNAs (tRNAs) were predicted to form the standard clover-leaf structure, except for $tRNA^{Ser}(AGN)$, which lacks the dihydrouridine (DHU) arm. The most unusual feature that was found was the use of TCT as the $tRNA^{Ser}(AGN)$ anticodon instead of the GCT that is used in most other arthropods. The lack of tandem repeat motif in the 735 bp long A+T-rich region was another unusual feature of the *B. lineolata* mitogenome. The short, highly conserved polythymidine stretch that was previously described in the Orthoptera and Diptera orders was also present in the A+T-rich region of the *B. lineolata* (order Coleoptera) mitogenome. The sequence and annotation of the mitogenome of *B. lineolata* will provide further insights into the diversity and evolution of the Cerambycidae family of long-horned beetles. The *B. lineolata* mitogenome sequence has been deposited in GenBank (http://www.ncbi.nlm.nih.gov/Genbank/) under accession number JN986793.

Cerambycidae, mitogenome, evolution

Citation: Wang C Y, Feng Y, Chen X M. Complete sequence and gene organization of the mitochondrial genome of *Batocera lineolata* Chevrolat (Coleoptera: Cerambycidae). Chin Sci Bull, 2012, 57: 3578–3585, doi: 10.1007/s11434-012-5271-3

Mitochondrial genes have been widely used in phylogenetics, phylogeography, and molecular diagnostics studies [1–3]. In particular, full-length mitochondrial genome (mitogenome) information has become an important tool for the study of genome architecture, population genetics, primer design, and molecular evolution [4–7]. Furthermore, mitogenomic sequences of the arthropods have provided important information for pest control. For example, insecticide resistance in an arthropod pest (*Tetranychus urticae*) was reported to be controlled by the mitochondrial DNA (mtDNA) [8].

Cerambycidae is a large family in the Coleoptera order. Most of the species in the Cerambycidae family are forestry pests. To date, more than 25000 species of Cerambycidae have been described worldwide [9]. Despite the large taxonomic diversity within this family, information about the Cerambycidae mitogenome is still limited, and currently only two complete mitogenomes of the Cerambycidae species are available in GenBank (NC_013070 and NC_008221). Given the important role of the mitogenome in population studies and pest control, more studies on the mitogenomes of the Cerambycidae species are needed.

The Cerambycidae *Batocera lineolata* Chevrolat is an important forestry pest which usually attacks *Juglans sigillata*, *Castanea mollissima*, *Eriobotrya japonica*, *Ficus carica*, *Sapium sebiferum*, and *Citrus reticulate*. This long-

*Corresponding author (email: yingf@hotmail.com)

horn beetle is found mainly in China, the Korean Peninsula, Vietnam, India, and Japan (http://www.lamiinae.org/index.php?pg=fgrp&id=13065&lg=en). To date, this widely distributed species has received less attention in molecular biology. To understand more about this beetle, detailed researches on population genetics, phylogeography and other relevant areas are required.

In this study, we sequenced the entire mitogenome of *B. lineolata* and analyzed the nucleotide organization and major characteristics of the mitogenome. The genome is 15418 bp long and harbors 13 protein-coding genes (PCGs), 22 tRNA genes, two ribosomal RNA (rRNA) genes and an A+T-rich region. The gene arrangement in the *B. lineolata* mitogenome is identical to the ancestral insect mitogenome arrangement [10]. The sequence and annotation of the mitogenome of *B. lineolata* will be an important addition to the continued efforts in studying Coleoptera mitogenome architecture, phylogeography, and phylogenetics.

1 Materials and methods

1.1 Sample and DNA extraction

An adult *B. lineolata* was collected in Kunming, Yunnan Province, China, on 16 July, 2010. The freshly collected material was preserved immediately in 100% ethanol and stored in a −20°C refrigerator before genomic DNA extraction. Total genomic DNA was extracted with the Wizard™ Genomic DNA Purification Kit (Promega, Madison, WI, USA), in accordance with the manufacturer's instructions.

1.2 PCR amplification and sequencing

To sequence the complete mitogenome, long PCR primers and some short PCR primers were designed based on multiple sequence alignments of all the available complete Coleoptera mitogenomes using ClustalX1.8 [11] and the Primer Premier 5.0 software. Primer sequence information can be obtained from the authors on request. Long PCRs were performed using TaKaRa LA *Taq* polymerase with the following cycling parameters: initial denaturing for 5 min at 95°C; followed by 30 cycles at 95°C for 50 s, 50°C for 50 s, 68°C for 2.5 min; and a final extension step of 68°C for 10 min. Short fragments were amplified with TaKaRa *Taq* polymerase: initial denaturing for 5 min at 94°C; followed by 35 cycles at 94°C for 1 min, 45–53°C for 1 min, 72°C for 2 min; and a final extension step of 72°C for 10 min. The PCR products were detected via electrophoresis in 1.5% agarose gel, and purified using the QIAquick PCR Purification Kit (Qiagen, USA). The sequencing reaction was performed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). Electrophoresis of the purified sequencing product was performed on an ABI-3730 DNA Analyzer (Applied Biosystems).

1.3 Gene identification and tRNA structures

Sequences with overlapping fragments were assembled with the neighboring fragments using the SeqMan program included in the Lasergene software package (DNAStar Inc., Madison, WI, USA). Via comparisons of the DNA or amino acid sequences with the homologous regions of known full-length insect mitogenome sequences, 13 protein-coding genes (PCGs), two rRNA genes and the A+T-rich region were determined using the MEGA 5.0 software [12]. The nucleotide sequences of the PCGs were translated using the invertebrate mtDNA genetic code. The tRNA gene analysis was conducted using the tRNAscan-SE software v.1.21 [13], and the predicted folding of the tRNA sequences was further confirmed by visual inspection.

1.4 Sequence analysis

The A+T-content of the whole genome was calculated via the EditSeq program included in the Lasergene software package (DNAStar Inc.). The nucleotide composition at each codon position of the PCGs and codon usage were calculated in the MEGA 5.0 software [12]. Gene overlaps and intergenic-space sequences were hand-counted.

Nucleotide composition skew was calculated for the PCGs and for the whole genome in the EditSeq program included in the Lasergene software package (DNAStar Inc.) using the formula proposed by Perna and Kocher [14]: $\text{GC-skew} = (\text{G} - \text{C})/(\text{G} + \text{C})$ and $\text{AT-skew} = (\text{A} - \text{T})/(\text{A} + \text{T})$, where C, G, A and T are the frequencies of the four bases.

2 Results and discussion

2.1 General features of *Batocera lineolata* mitogenome

The complete mtDNA sequence of *B. lineolata* is 15418 bp long, consisting of two rRNAs (*srRNA* and *lrRNA*), 22 tRNAs, 13 PCGs (*ATP6*, *ATP8*, *COI-III*, *ND1-6*, *ND4L*, *CytB*), and one major non-coding A+T-rich region (Figure 1). As is the case in many insect mitogenomes, the major strand (J strand) codes for most of the genes (nine PCGs and 14 tRNAs), and the remaining genes are coded in the minor strand (N strand) (four PCGs, eight tRNAs and two rRNA genes), as shown in Figure 1. The gene order and orientation are identical to the most common type that has been suggested as the ancestral type for insects [10,15].

All the genes are closely assembled in the genome and only five intergenic spacers are observed. These intergenic spacers are 51 bp in total (excluding the A+T rich region), with their individual sizes ranging from 1 to 25 bp. In addition, there are a total of 47 bp overlapping sequences in 12 overlaps between different genes; the lengths of the overlaps range from 1 to 8 bp (Table 1).

A comparative analysis of the *B. lineolata* mitogenome with the two other published Cerambycidae mitogenomes of

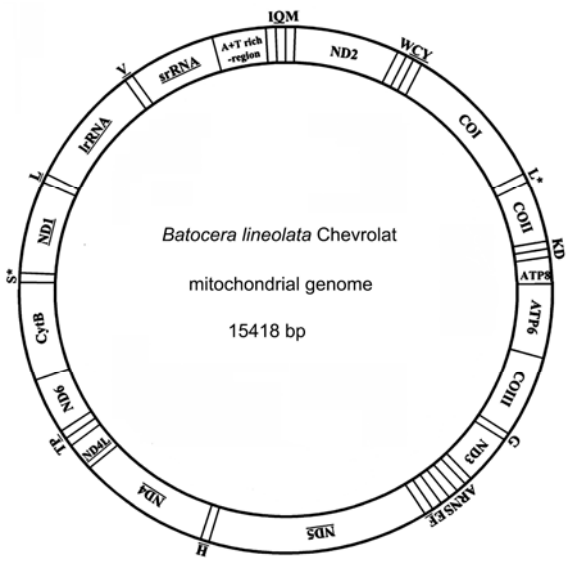


Figure 1 Circular map of the *Batocera lineolata* mitogenome. COI, COII and COIII are the subunits of cytochrome oxidase; CytB is cytochrome B; ATP6 and ATP8 are subunits 6 and 8 of F_0 ATPase; ND1–6 are the components of NADH dehydrogenase. The tRNAs are denoted using one-letter symbols that are consistent with the IUPAC-IUB single letter amino acid codes. Gene names that are not underlined indicate clockwise transcription; underlined gene names indicate counter-clockwise transcription. L* and S* denote $tRNA^{Leu}(UUR)$ and $tRNA^{Ser}(UCN)$ respectively.

Psacothea hilaris and *Anoplophora glabripennis*, showed that the mitogenomes exhibited highly conserved architectures including the genome content, gene order, nucleotide composition, codon usage, as well as the predicted amino acid composition of the PCGs [16,17]. The *B. lineolata* mitogenome sequence has been deposited in GenBank (http://www.ncbi.nlm.nih.gov/Genbank/) under accession number JN986793.

2.2 Protein-coding genes

All 13 PCGs use the standard ATN (Met) start codon (Table 1). *ATP6*, *ND4L*, and *ND6* use the common stop codon TAA; *ATP8*, *CytB*, and *ND1* use the TAG stop codon; *ND2*, *COI*, *COII*, *ND3*, *ND5*, and *ND4* are terminated with a single T residue; and *COIII* terminates with TA (Table 1). Similar termination signals have been found in a number of other insect mitogenomes [18]. For example, in *Coreana raphaelis*, a single T residue has been deemed the stop codon for *COI*, *COII*, *ND5* and *CytB*, and the TA dinucleotide has been deemed the stop codon for *ATP6*, *ND4*, *ND4L*, and *ND6* [19]; similarly, in *Hyphantria cunea*, a single T was considered to be the stop codon for *COI*, *COII* and *ND4*, and TA was the stop codon for *ATP6* [4]. Incomplete stop codons also exist in vertebrate mitogenomes, including the human mitogenome [20]. It has been demonstrated that incomplete stop codons could produce functional stop codons in polycistronic transcription cleavage and polyadenylation processes [21].

2.3 Transfer RNA and ribosomal RNA genes

Twenty-two tRNA genes (one specific for each of the amino acid and two for leucine and serine) were identified within the *B. lineolata* mitogenome. The tRNAs are interspersed throughout the mitogenome and range in length from 63 to 69 bp (Table 1). All tRNAs, except $tRNA^{Ser}(AGN)$, were predicted to fold into typical cloverleaf secondary structures (Figure 2). The unusual $tRNA^{Ser}(AGN)$ lacks dihydrouridine (DHU) arm, which is replaced by a simple loop (Figure 2). This incomplete $tRNA^{Ser}$ structure has been detected in other animals [2], including those in insect groups [4,19,22–24].

A total of 35 non-Watson-Crick base pairs were observed in the *B. lineolata* tRNAs; 29 are G-U pairs, which form a weak hydrogen bond in the tRNAs. The remaining six pairs are atypical: one U-U mismatch in $tRNA^{Arg}$; a U-U and U-C in $tRNA^{Ile}$; two U-U pairs in $tRNA^{Leu}(CUN)$; and one A-G pair in $tRNA^{Trp}$ (Figure 2). It has been demonstrated that the mismatched pairs found in the tRNAs can be corrected through RNA-editing mechanisms [25].

The *lrRNA* and *srRNA* genes in the *B. lineolata* mitogenome are 1289 and 812 bp in length, respectively. As has been observed in other insects [5,10], *lrRNA* is located between $tRNA^{Leu}(CUN)$ and $tRNA^{Val}$, and *srRNA* is between $tRNA^{Val}$ and the A+T-rich region (Figure 1). The A+T content of *lrRNA* and *srRNA* is 75.02% and 77.46%, respectively.

2.4 A+T-rich region

The A+T-rich region of the *B. lineolata* mitogenome is located between the *srRNA* and $tRNA^{Ile}$ genes (Table 1). The length of this region is 735 bp, much shorter than the length of the A+T-rich regions in the other two published Cerambycidae species *Psacothea hilaris* (1189 bp) and *Anoplophora glabripennis* (1114 bp) [16,17]. Like the two other Cerambycidae mitogenomes, the A+T-rich region also exhibits the highest A+T content (85.71%) in the *B. lineolata* mitogenome. Unlike some insect species [26,27], the 735 bp A+T-rich region does not have large tandem repetitive sequences; however, it does have some microsatellite-like repeats (for example, $(T)_{15}$, $(AT)_8$, and $(A)_7$). Especially, the poly-T stretch (15 bp) has been suggested to function as a possible recognition site for the initiation of replication of the minor strand of mtDNA [5,28]. The A+T-rich region in insect mitogenomes is equivalent to the control region of the vertebrate mitogenomes. This region has been shown to harbor the origin sites for the transcription and replication for both strands of insect mitogenomes [7,15,28].

2.5 Nucleotide composition and codon usage

The A+T content of the whole *B. lineolata* mitogenome is 74.48%, showing an obvious AT mutation bias [29], as

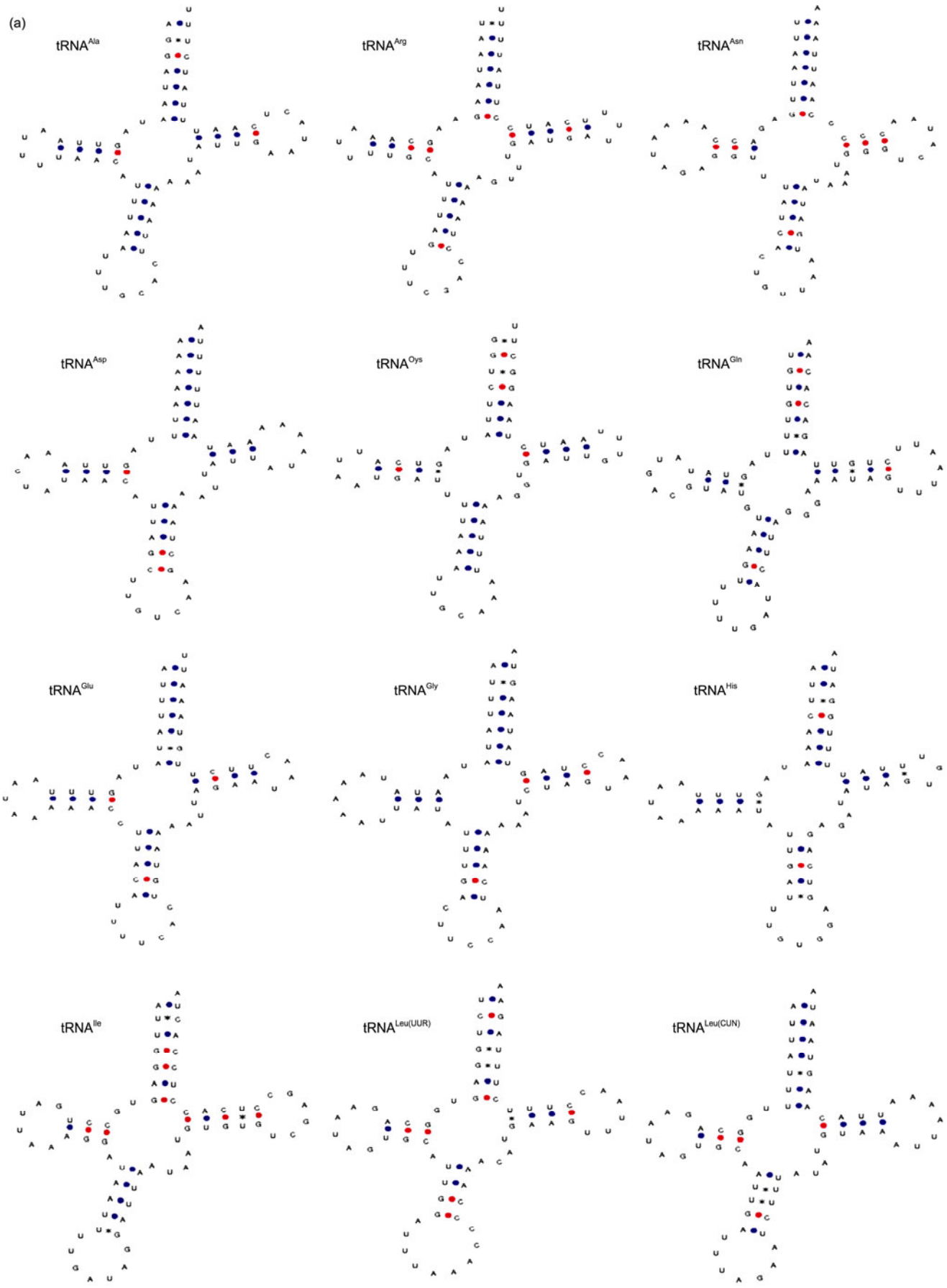
Table 1 Details of the *Batocera lineolata* mitogenome [a)]

Gene	Direction	Nucleotide number	Size (bp)	Anticodon	Start codon	Stop codon
tRNA^{Ile}	F	1–68	68	30–32 GAT	–	–
tRNA^{Gln}	R	70–138	69	106–108 TTG	–	–
tRNA^{Met}	F	138–206	69	168–170 CAT	–	–
ND2	F	207–1215	1009	–	ATC(M)	T(AA)[#]
tRNA^{Trp}	F	1216–1284	69	1246–1248 TCA	–	–
tRNA^{Cys}	R	1277–1339	63	1308–1310 GCA	–	–
tRNA^{Tyr}	R	1340–1406	67	1373–1375 GTA	–	–
COI	F	1399–2941	1543	–	ATT(M)	T(AA)[#]
tRNA^{Leu}(*UUR*)	F	2942–3006	65	2971–2973 TAA	–	–
COII	F	3007–3694	688	–	ATT(M)	T(AA)[#]
tRNA^{Lys}	F	3695–3764	70	3725–3727 TTT	–	–
tRNA^{Asp}	F	3764–3827	64	3794–3796 GTC	–	–
ATP8	F	3828–3983	156	–	ATT(M)	TAG
ATP6	F	3977–4651	675	–	ATG(M)	TAA
COIII	F	4651–5438	789	–	ATG(M)	TA(A)[#]
tRNA^{Gly}	F	5438–5501	64	5468–5470 TCC	–	–
ND3	F	5502–5853	352	–	ATC(M)	T(AA)[#]
tRNA^{Ala}	F	5854–5918	65	5883–5885 TGC	–	–
tRNA^{Arg}	F	5915–5981	67	5946–5948 TCG	–	–
tRNA^{Asn}	F	5979–6046	68	6011–6013 GTT	–	–
tRNA^{Ser}(*AGN*)	F	6047–6113	67	6083–6086 TCT	–	–
tRNA^{Glu}	F	6114–6178	65	6144–6146 TTC	–	–
tRNA^{Phe}	R	6177–6240	64	6208–6210 GAA	–	–
ND5	R	6241–7960	1720	–	ATA(M)	T(AA)[#]
tRNA^{His}	R	7958–8020	63	7988–7990 GTG	–	–
ND4	R	8021–9353	1333	–	ATG(M)	T(AA)[#]
ND4L	R	9347–9634	288	–	ATG(M)	TAA
tRNA^{Thr}	F	9637–9701	65	9667–9669 TGT	–	–
tRNA^{Pro}	R	9702–9766	65	9735–9737 TGG	–	–
ND6	F	9769–10272	504	–	ATT(M)	TAA
CytB	F	10272–11414	1143	–	ATG(M)	TAG
tRNA^{Ser}(*UCN*)	F	11413–11480	68	11442–11444 TGA	–	–
ND1	R	11502–12428	927	–	ATT(M)	TAG
tRNA^{Leu}(*CUN*)	R	12454–12519	66	12488–12490 TAG	–	–
lrRNA	R	12519–13807	1289	–	–	–
tRNA^{Val}	R	13803–13871	69	13839–13841 TAC	–	–
srRNA	R	13872–14683	812	–	–	–
A+T-rich region	R	14684–15419	735	–	–	–

a) Abbreviations of the tRNA genes follow the IUPAC-IUB three letter code; #, the TAA stop codon is completed by adding A residues to the 3' end of the mRNA.

observed in the other two publicly available Cerambycidae mitogenmes [16,17]. The AT-skew value for the major strand is 0.041, indicating the occurrence of more As than Ts in this strand. For the 13 PCGs, the mean value for the A+T content is 73.3%, with a strong A+T bias. The A+T content at the third codon position (85.3%) is higher than at the first (67.5%) and second positions (67.1%), indicating that the third codon position is the one most susceptible to AT mutation bias [29].

The relative synonymous codon usage in the *B. lineolata*


(a)
tRNA^Ala
tRNA^Arg
tRNA^Asn
tRNA^Asp
tRNA^Oys
tRNA^Gln
tRNA^Glu
tRNA^Gly
tRNA^His
tRNA^Ile
tRNA^Leu(UUR)
tRNA^Leu(CUN)

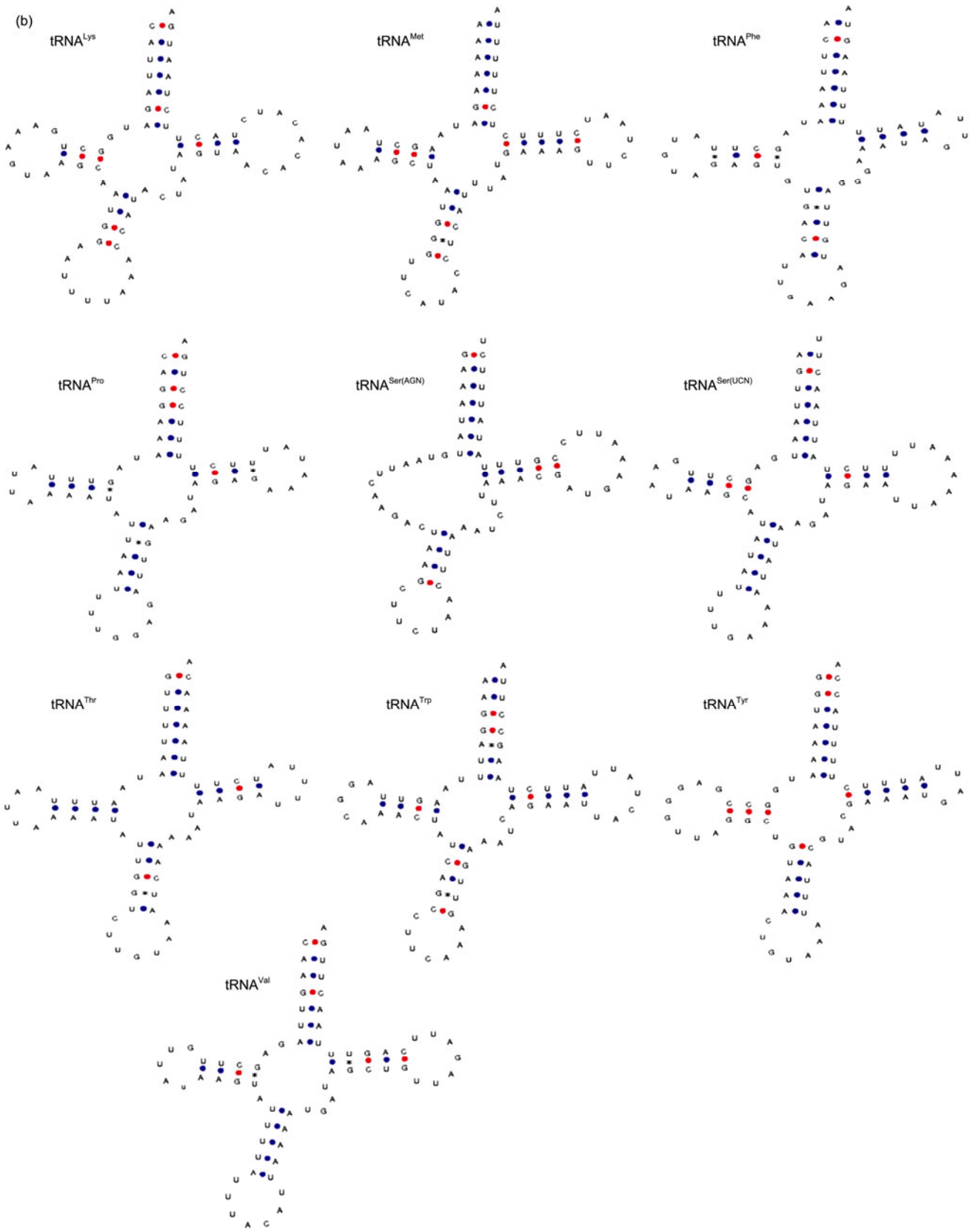


Figure 2 Predicted clover-leaf secondary structure for the 22 tRNA genes in the *Batocera lineolata* mitogenome. The tRNAs are labeled with their corresponding amino acids. Filled circles (●) indicate Watson-Crick base-pairing, with red color indicate a G-C pair and blue color an A-U pair. Asterisks (*) indicate non-Watson-Crick base-paring. The tRNA arms (clockwise from the top) are the amino acid acceptor (AA) arm, the TψC (T) arm, the anticodon (AC) arm and the dihydrouridine (DHU) arm.

Table 2 Codon number and relative synonymous codon usage in the *Batocera lineolata* mitochondrial protein coding genes[a)]

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	284	1.66	UCU(S)	128	2.91	UAU(Y)	139	1.71	UGU(C)	25	1.72
UUC(F)	58	0.34	UCC(S)	20	0.45	UAC(Y)	24	0.29	UGC(C)	4	0.28
UUA(L)	393	3.95	UCA(S)	56	1.27	UAA(*)	0	0	UGA(W)	94	1.88
UUG(L)	50	0.50	UCG(S)	2	0.05	UAG(*)	0	0	UGG(W)	6	0.12
CUU(L)	79	0.79	CCU(P)	57	1.75	CAU(H)	60	1.64	CGU(R)	16	1.08
CUC(L)	14	0.14	CCC(P)	37	1.14	CAC(H)	13	0.36	CGC(R)	3	0.20
CUA(L)	57	0.57	CCA(P)	30	0.92	CAA(Q)	65	1.81	CGA(R)	36	2.44
CUG(L)	4	0.04	CCG(P)	6	0.18	CAG(Q)	7	0.19	CGG(R)	4	0.27
AUU(I)	337	1.75	ACU(T)	95	2.07	AAU(N)	158	1.77	AGU(S)	30	0.68
AUC(I)	48	0.25	ACC(T)	28	0.61	AAC(N)	21	0.23	AGC(S)	7	0.16
AUA(M)	182	1.78	ACA(T)	60	1.30	AAA(K)	97	1.66	AGA(S)	92	2.09
AUG(M)	23	0.22	ACG(T)	1	0.02	AAG(K)	20	0.34	AGG(S)	17	0.39
GUU(V)	86	1.76	GCU(A)	81	2.06	GAU(D)	60	1.69	GGU(G)	57	1.10
GUC(V)	15	0.31	GCC(A)	26	0.66	GAC(D)	11	0.31	GGC(G)	14	0.27
GUA(V)	83	1.70	GCA(A)	47	1.20	GAA(E)	65	1.57	GGA(G)	106	2.05
GUG(V)	11	0.23	GCG(A)	3	0.08	GAG(E)	18	0.43	GGG(G)	30	0.58

a) RSCU, relative synonymous codon usage. The letters in brackets are the single-letter amino acid codes. Total number of codons, excluding the stop codons (*)=3700.

mitochondrial PCGs was investigated and the results are summarized in Table 2. The four most frequently used codons, TTA (leucine, Leu), ATT (isoleucine, Ile), TTT (phenylalanine, Phe), and ATA (methionine, Met), account for 32.3% of all the codons in the *B. lineolata* mitogenome. These four codons are composed of A or T nucleotides only, indicating that the strong AT mutation bias could obviously influence the codon usage [30,31].

The total number of non-stop codons in the PCGs is 3700. Among all the amino acids encoded by the 13 PCGs, Leu (16.14%), Ile (10.41%), Ser (9.51%), and Phe (9.24%) are the four most abundant amino acids; three of them are encoded by the AT-rich codons (see above), suggesting that the AT bias has affected the amino acid composition of the proteins encoded by the mitochondrial genes [32,33].

The authors thank Dr. Zhong-Bao Zhao for helpful suggestions in the preparation of the manuscript. This work was supported by the National Ministry of Forestry of China (2009-4-37), and the Science and Technology Committee of Yunnan Province, China (2011FB141). The authors declare that there are no conflicts of interest.

1 Kocher T D, Thomas W K, Meyer A, et al. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. Proc Natl Acad Sci USA, 1989, 86: 6196–6200
2 Wolstenholme D R. Animal mitochondrial DNA: Structure and evolution. Int Rev Cytol, 1992, 141: 173–216
3 Simmons R B, Weller S J. Utility and evolution of cytochrome *b* in insects. Mol Phylogenet Evol, 2001, 20: 196–210
4 Liao F, Wang L, Wu S, et al. The complete mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). Int J Biol Sci, 2010, 6: 172–186
5 Kim M I, Baek J Y, Kim M J, et al. Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects. Mol Cells, 2009, 28: 347–363
6 Lee E M, Hong M Y, Kim M I, et al. The complete mitogenome sequences of the palaeopteran insects *Ephemera orientalis* (Ephemeroptera: Ephemeridae) and *Davidius lunatus* (Odonata: Gomphidae). Genome, 2009, 52: 810–817
7 Yukuhiro K, Sezutsu H, Itoh M, et al. Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkmoth, *Bombyx mandarina*, and its close relative, the domesticated silkmoth, *Bombyx mori*. Mol Biol Evol, 2002, 19: 1385–1389
8 Van Leeuwen T, Vanholme B, Van Pottelberge S, et al. Mitochondrial heteroplasmy and the evolution of insecticide resistance: Non-Mendelian inheritance in action. Proc Natl Acad Sci USA, 2008, 105: 5980–5985
9 Sama G, Buse J, Orbach E, et al. A new catalogue of the Cerambycidae (Coleoptera) of Israel with notes on their distribution and host plants. Mun Ent Zool, 2010, 5: 1–55
10 Boore J L, Lavrov D V, and Brown W M. Gene translocation links insects and crustaceans. Nature, 1998, 392: 667–668
11 Thompson J D, Gibson T J, Plewniak F, et al. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res, 1997, 25: 4876–4882
12 Tamura K, Peterson D, Peterson N, et al. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol, 2011, 28: 2731–2739
13 Lowe T M, Eddy S R. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res, 1997, 25: 955–964
14 Perna N T, Kocher T D. Unequal base frequencies and the estimation of substitution rates. Mol Biol Evol, 1995, 12: 359–361
15 Taanman J W. The mitochondrial genome: Structure, transcription, translation and replication. Biochim Biophys Acta, 1999, 1410: 103–123

16 An Y, Huang X, Hu X, et al. The complete mitochondrial genome of the longhorned beetle *Anoplophora glabripennis*. 2006, RefSeq accession no. NC_008221
17 Kim K G, Hong M Y, Kim M J, et al. Complete mitochondrial genome sequence of the yellow-spotted long-horned beetle *Psacothea hilaris* (Coleoptera: Cerambycidae) and phylogenetic analysis among coleopteran insects. Mol Cell, 2009, 27: 429–441
18 Yin H, Zhi Y, Jiang H, et al. The complete mitochondrial genome of *Gomphocerus tibetanus* Uvarov, 1935 (Orthoptera: Acrididae: Gomphocerinae). Gene, 2012, 494: 214–218
19 Kim I, Lee E M, Seol K Y, et al. The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae). Insect Mol Biol, 2006, 15: 217–225
20 Andrews R M, Kubacka I, Chinnery P F, et al. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet, 1999, 23: 147
21 Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria. Nature, 1981, 290: 470–474
22 Salvato P, Simonato M, Battisti A, et al. The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae). BMC Genomics, 2008, 9: 331
23 Yang F, Du Y Z, Wang L P, et al. The complete mitochondrial genome of the leafminer *Liriomyza sativae* (Diptera: Agromyzidae): Great difference in the A+T-rich region compared to *Liriomyza trifolii*. Gene, 2011, 485: 7–15
24 Li H, Gao J, Liu H, et al. The architecture and complete sequence of mitochondrial genome of an assassin bug *Agriosphodrus dohrni* (Hemiptera: Reduviidae). Int J Biol Sci, 2011, 7: 792–804
25 Lavrov D V, Brown W M, Boore J L. A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. Proc Natl Acad Sci USA, 2000, 97: 13738–13742
26 Bae J S, Kim I, Sohn H D, et al. The mitochondrial genome of the firefly, *Pyrocoelia rufa*: Complete DNA sequence, genome organization, and phylogenetic analysis with other insects. Mol Phylogenet Evol, 2004, 32: 978–985
27 Wei S J, Tang P, Zheng L H, et al. The complete mitochondrial genome of *Evania appendigaster* (Hymenoptera: Evaniidae) has low A+T content and a long intergenic spacer between *atp8* and *atp6*. Mol Biol Rep, 2010, 37: 1931–1942
28 Saito S, Tamura K, Aotsuka T. Replication origin of mitochondrial DNA in insects. Genetics, 2005, 171: 1695–1705
29 Eyre-Walker A. Differentiating between selection and mutation bias. Genetics, 1997, 147: 1983–1987
30 Rao Y, Wu G, Wang Z, et al. Mutation bias is the driving force of codon usage in the *Gallus gallus* genome. DNA Res, 2011, 18: 499–512
31 Powell J R, Moriyama E N. Evolution of codon usage bias in *Drosophila*. Proc Natl Acad Sci USA, 1997, 94: 7784–7790
32 Min X J, Hickey D A. DNA asymmetric strand bias affects the amino acid composition of mitochondrial proteins. DNA Res, 2007, 14: 201–206
33 Foster P G, Jermiin L S, Hickey D A. Nucleotide composition bias affects amino acid content in proteins coded by animal mitochondria. J Mol Evol, 1997, 44: 282–288