Mol. Cells *27*, 429-441, April 30, 2009
DOI/10.1007/s10059-009-0064-5

Molecules and Cells

Complete Mitochondrial Genome Sequence of the Yellow-Spotted Long-Horned Beetle *Psacothea hilaris* (Coleoptera: Cerambycidae) and Phylogenetic Analysis among Coleopteran Insects

Ki-Gyoung Kim[1], Mee Yeon Hong[2], Min Jee Kim[2], Hyun Hwak Im[2], Man Il Kim[2], Chang Hwan Bae[1], Sook Jae Seo[3], Sang Hyun Lee[2], and Iksoo Kim[2,*]

We have determined the complete mitochondrial genome of the yellow-spotted long horned beetle, *Psacothea hilaris* (Coleoptera: Cerambycidae), an endangered insect species in Korea. The 15,856-bp long *P. hilaris* mitogenome harbors gene content typical of the animal mitogenome and a gene arrangement identical to the most common type found in insect mitogenomes. As with all other sequenced coleopteran species, the 5-bp long TAGTA motif was also detected in the intergenic space sequence located between tRNA^{Ser}(UCN) and ND1 of *P. hilaris*. The 1,190-bp long non-coding A+T-rich region harbors an unusual series of seven identical repeat sequences of 57-bp in length and several stretches of sequences with the potential to form stem-and-loop structures. Furthermore, it contains one tRNA^{Arg}-like sequence and one tRNA^{Lys}-like sequence. Phylogenetic analysis among available coleopteran mitogenomes using the concatenated amino acid sequences of PCGs appear to support the sister group relationship of the suborder Polyphaga to all remaining suborders, including Adephaga, Myxophaga, and Archostemata. Among the two available infraorders in Polyphaga, a monophyletic Cucujiformia was confirmed, with the placement of Cleroidea as the basal lineage for Cucujiformia. On the other hand, the infraorder Elateriformia was not identified as monophyletic, thereby indicating that Scirtoidea and Buprestoidea are the basal lineages for Cucujiformia and the remaining Elateriformia.

INTRODUCTION

The mitochondrial genomes (mitogenomes) of the majority of metazoans are circular molecules spanning approximately 16 kb in size, comprised of 13 protein-coding genes (PCGs) (ND1-6, ND4L, CytB, COI-III, ATP6 and ATP8), 22 tRNA genes, 2 rRNA genes (16S rRNA and 12S rRNA), and usually one large non-coding element which regulates the transcription and replication of the mitogenome (Wolstenholme, 1992). Owing to the high AT content of the element in insects, this is referred to as the A+T-rich region.

Mitochondrial DNA (mtDNA) exhibits several properties that make it a valuable tool in the study of population genetics, phylogenetics, molecular evolution, and even conservation genetics, due to its relatively simple genetic structure, maternal mode of inheritance (in most cases), and high rate of evolution (Avise, 1994; Moriz et al., 1987). Thus, more than 100 species with complete mitogenome sequences have been determined from a variety of insects.

These full-length mitogenome sequences have contributed significantly to our understanding of the structure of the mitogenome (Boore, 1999), gene arrangement (Shao et al., 2001), and the evolution of arthropod lineages (Hwang et al., 2001; Nardi et al., 2003). However, this list includes only 14 coleopteran species. Considering the diversity of the Coleoptera, which encompasses 350,000 described species, the existing full-length mitogenome sequence information for the Coleoptera remains rather limited. Considering, in particular, that the threat of mass extinction to present living organisms is increasing rapidly as the result of several factors (i.e., global warming, habitat destruction, and so on), it is very important to collect minimal, but meaningful amounts of genetic information from organisms that are vulnerable to future extinctions. Furthermore, considering that the re-introduction of natural or artificially-reared organisms to locally endangered or extinct populations is a frequently utilized conservation strategy, an understanding of the genetic relatedness of donor and donee populations is a critical prerequisite for such processes (Joyce and Pullin, 2004; Murata et al., 2004).

The yellow-spotted long-horned beetle, *Psacothea hilaris* (Pascoe), belongs to the insect family Cerambycidae, within the superfamily Chrysomeloidea, in the insect order Coleoptera.

[1]Biological Resources Research Department, National Institute of Biological Resources, Incheon 404-708, Korea, [2]College of Agriculture and Life Sciences, Chonnam National University, Gwangju 500-757, Korea, [3]Division of Applied Life Science (Brain Korea 21 Program), Graduate School of Gyeongsang National University, Jinju 660-701, Korea
*Correspondence: ikkim81@chonnam.ac.kr

Received December 22, 2008; revised March 3, 2009; accepted March 4, 2009; published online April 13, 2009

Keywords: coleopteran phylogeny, complete mitochondrial genome, *Psacothea hilaris*, tRNA-like sequence, yellow-spotted long-horned beetle

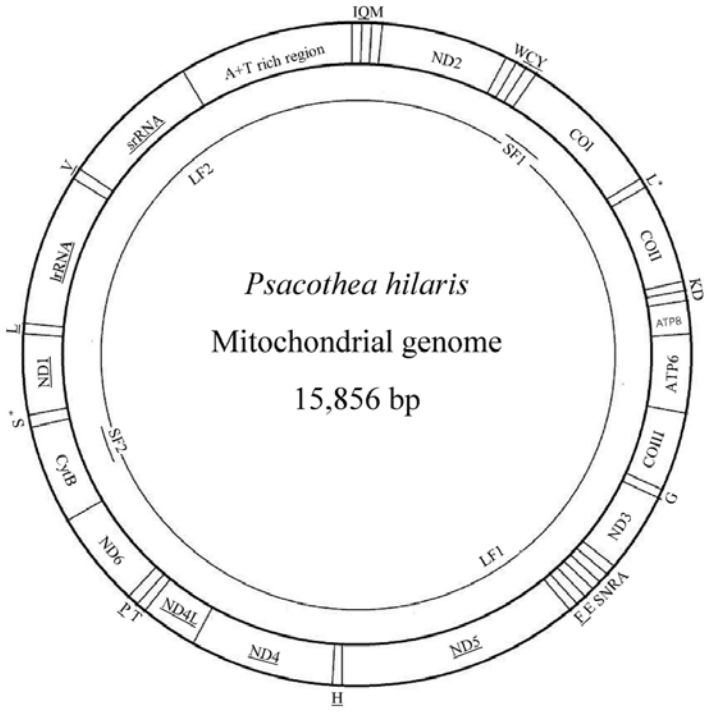

Fig. 1. Circular map of the mitogenome of *Psacothea hilaris*. The abbreviations for the genes are as follows: COI, COII and COIII refer to the cytochrome oxidase subunits, CytB refers to cytochrome *b*, and ND1-6 refer to NADH dehydrogenase components. tRNAs are denoted as one-letter symbols in accordance with the IUPAC-IUB single-letter amino acid codes. The one-letter symbols L, L*, S and S* denote $tRNA^{Leu}$(CUN), $tRNA^{Leu}$(UUR), $tRNA^{Ser}$(AGN), and $tRNA^{Ser}$(UCN), respectively. Gene names that are not underlined indicate a clockwise direction of transcription, whereas underlines indicate a counter-clockwise direction. The *P. hilaris* mitogenome was sequenced by four overlapping fragments (SF1, SF2, LF1, and LF2) shown as single lines within a circle.

This species is distributed throughout the Korean peninsula, as well as in Japan, Taiwan, and China. Currently, the species can be found only very rarely in its typical habitats, most notably the fig and mulberry trees located mostly on islands; thus, the species is listed as a second-degree endangered wild animal in Korea (Nam, 1996).

In this study, we determined the complete mitogenome sequence of *P. hilaris*. The mitogenome sequence was described by comparison with other insect mitogenomes, in particular to members of coleopteran species. Thus, the newly sequenced *P. hilaris* belonging to the Chrysomeloidea is expected to enrich our understanding of the comparative biology of the mitogenomes of coleopteran species. Furthermore, the concatenated amino acid sequences of the 13 PCGs of *P. hilaris* were utilized in order to gain insight into the phylogenetic relationships among available coleopteran suborders and superfamilies. Currently, the subordinal relationships among the four coleopteran suborders (Archostemata, Myxophaga, Adephaga, and Polyphaga) are controversial and, in fact, almost all possible relationships have already been suggested (Beutel, 1997; Beutel and Haas, 2000; Carterino et al., 2002; Kukalová-Peck and Lawrence, 1993). Within the infraorders Elateriformia and Cucujiformia, the detailed relationships among superfamilies remain poorly understood (Kukalova-Peck and Lawrence, 1993).

MATERIALS AND METHODS

DNA extraction

An adult *P. hilaris* specimen was collected from Geomun-do, Korea, in 2007. DNA was extracted using a Wizard™ Genomic DNA Purification Kit in accordance with the manufacturer's instructions (Promega, USA). Only the muscle attached to the thorax was utilized after both the thorax and abdomen were opened.

Primer design, PCR, and sequencing

In order to sequence the entire mitogenome of *P. hilaris*, ~700 bp of *P. hilaris* COI and CytB (SF1 and SF2 in Fig. 1, respectively) were first sequenced. For the amplification of COI (SF1) corresponding to the "DNA Barcode" region being utilized for global animal identification (Hebert et al., 2003), a pair of primers was designed based on the method described by Folmer et al. (1994): LCO1490, 5′-GGTCAACAAATCATAAAGATATTGG-3′, and HCO2198, 5′-TAAACTTCAGGGTGACCAAAAAATCA-3′. On the other hand, the primers for CytB (SF2) were designed via the alignment of several insect mitogenomes sequenced in their entirety: CytBF, 5′-CGAGATGTAAATTATGGATGA-3′ and CytBR, 5′-ATTAGGGATTGAACGTAGAAT-3′. These short fragments were amplified with AccuPower® PCR PreMix (Bioneer, Korea) under the following conditions: initial denaturation for 4 min at 94°C, followed by 35 cycles of 30 s at 94°C, 40 s at 50-55°C, and 60 s at 72°C, with a subsequent final 7-minute extension at 72°C. On the basis of the sequence information, two pairs of primers were designed to amplify two overlapping long fragments (LF1 and LF2) using LA Taq™ (Takara Biomedical, Japan) under the following conditions: initial denaturation for 2 min at 96°C, followed by 30 cycles of 10 s at 98°C and 15 min at 58-65°C, and a final 10-minute extension at 72°C. The primer sequences for LF1 and LF2 were: PHCOI-F, 5′-CTATCCTAGGTGCCGTAAATTTTA-3′ and PHCytB-R, 5′-TGGATCTCCTAATAAGTAGGGATT-3′ for LF1 and PHCytB-F, 5′-CAACTATTGTTCAATGAGTTTGAG-3′ and PHCOI-R, 5′-AAGGGACTAATCAATTTCCAAATC-3′ for LF2. These PCR products were then utilized in the construction of a shotgun library. In brief, DNAs were sheared into 1-5 kb fragments using Hydroshear (Gene Machine, USA), and the DNA fraction was collected using a Chromaspin TE 1000 column. The DNA fraction was then cloned into the pUC118 vector (Takara Biomedical, Japan), and each of the resultant plasmid DNAs were isolated with a Wizard *Plus* SV Minipreps DNA Purification System (Promega, USA). DNA sequencing was conducted with the ABI PRISM® BigDye® Terminator v3.1 Cycle Sequencing Kit and the ABI PRISM™ 3100 Genetic Analyzer (PE Applied Biosys-

tems, USA). All fragments were sequenced from both strands.

Sequence analysis

Individual *P. hilaris* mitochondrial genes and the A+T-rich region were determined via the alignment of the sequences with homologous regions of known full-length insect mt sequences, using CLUSTAL X software (Thompson et al., 1997). The nucleotide sequences of the PCGs were translated on the basis of the invertebrate mtDNA genetic code. The secondary structures of most of the tRNA genes were predicted using tRNAscan-SE 1.21 (Lowe and Eddy, 1997) with invertebrate codon predictors and a cove score cut off of 1; however, some [e.g., $tRNA^{Ser(AGN)}$] were drawn by hand, based on the nucleotide sequences of the tRNA genes of other insects sequenced in their entirety and edited visually, carefully considering the anticodon sequences. The entire A+T-rich region of *P. hilaris* was subjected to a search for sequences capable of forming stem-and-loop structures, using GeneBee software (Brodsky et al., 1992). The sequence data were deposited into the GenBank database under the accession no. FJ424074.

Phylogenetic analysis

In order to reconstruct the phylogenetic relationships existing among the coleopteran suborders and superfamilies, the concatenated amino acid sequence of the 13 PCGs was assessed via Bayesian inference (BI), maximum likelihood (ML), and maximum parsimony (MP) analyses. The endopterygotan Neoptera, *Polystoechotes punctatus*, *Ascaloptynx appendiculatus*, and *Corydalus cornutus* (unpublished, GenBank accession number FJ171325, FJ171324, and FJ171323, respectively) were employed as outgroups. The alignment of the amino acid sequences of the 13 individual PCGs was determined using CLUSTAL W software (Thompson et al., 1994) within BioEdit (Hall, 1999). A default gap opening/gap extension scheme was conducted, but only the conserved regions were selected using the GBlocks 0.91b program (Castresana, 2000) with the maximum number of contiguous non-conserved positions set at four. The resultant sequences were subsequently concatenated into an amino acid sequence alignment, resulting in a total of 3,232 sites in length, which is equivalent to 84% of the original length. This alignment is available upon request.

Substitution model selection was conducted via a comparison of Akaike Information Criterion (AIC) scores (Akaike, 1974), which were calculated using ProTest ver. 1.4 (Abascal et al., 2005). The mtRev-24 (Adachi and Hasegawa, 1996) +I+G+F model was selected as a model for BI and ML analyses. The BI analysis was conducted using MrBayes ver. 3.1 software (Huelsenbeck and Ronquist, 2001) under the following conditions: 1,000,000 generations, four chains (one hot chain and three cold chains), and a burn-in step of the first 10,000. The confidence values of the BI tree were expressed as the Bayesian posterior probabilities in percent (BPP). The ML analysis was conducted using PHYML (Guindon et al., 2005) under the following conditions: the proportion of invariable sites as "estimated", the number of substitution rate categories as six, the gamma distribution parameter as "estimated", and the starting tree as a BIONJ distance-based tree. MP analysis was conducted using PAUP* ver. 4.0b10 (Swofford, 2002) software via heuristic search using tree-bisection-reconnection (TBR) for the branch-swapping algorithm, steepest descent option not in effect, stepwise addition option for the starting tree, number of trees held at each step during stepwise addition for one, and initial "MaxTrees" setting at 100. Branches were collapsed if the maximum branch length was zero. Trees were evaluated via the bootstrap test (Felsenstein, 1985), with 1,000 iterations.

RESULTS AND DISCUSSION

Genome structure, organization, and composition

The entire *P. hilaris* mitogenome was PCR-amplified in four overlapping fragments (Fig. 1). It contains typical gene content: 13 PCGs, 22 tRNA genes, 2 rRNA genes, and the non-coding A+T-rich region (Table 1). The 15,856-bp long *P. hilaris* mitogenome was most similar in length to that of the within-superfamily species, *Crioceris duodecimpunctata* (15,880 bp) and *Anoplophora glabripennis* (15,774 bp) among 15 sequenced coleopteran mitogenomes (Table 2). A comparison of codon numbers across all sequenced mitogeomes showed that the average codon number in the sequenced coleopteran species was 3,703, covering a range between 3,685 in *Chaetosoma scaritides* and 3,726 in *Trachypachus holmbergi* (Table 2). Thus, the maximum difference in codon number among coleopteran species is 41 amino acids (123 bp). This, in turn, accounts for a difference of about three codons per PCG, thereby indicating that the sizes of the PCGs are relatively conserved in the coleopteran mitogenomes. On the other hand, the maximum size difference in the A+T-rich region was at least 1,153 bp, ranging in size from at least 875 bp in *Chrysochroa fulgidissima* to at least 2,028 bp in *Priasilpha obscura*. Thus, the size difference in the A+T-rich region is significantly larger than that of the PCGs. Considering that several coleopteran mitogenomes have been reported to lack a complete A+T-rich region (Sheffield et al., 2008), the actual size variations in this region may be significantly higher. This observation in coleopteran mitogenomes is strictly concordant with previous findings from other ordinal insect species (Fauron and Wolstenholme, 1980; Inohira et al., 1997; Lewis et al., 1995; Renfu et al., 2001). In fact, the A+T-rich region has been identified as the source of the length variation in the entire mitogenome, with variable numbers of nucleotides in polynucleotide runs, insertions/dele-tions, and the presence of varying copy numbers of tandem repeated elements (Zhang and Hewitt, 1997).

The orientation and gene order of all species - with the exception of *Tribolium castaneum* of the family Tenebrionidae--is identical to that of the most common type that has been hypothesized as ancestral for insects (Boore, 1999). The movement of the $tRNA^{Glu}$ to a position 3′-downstream of $tRNA^{Phe}$ resulted in an order of $tRNA^{Phe}$ and $tRNA^{Glu}$ in *T. castaneum* (Friedrich and Muquim, 2003), rather than the order $tRNA^{Glu}$ and $tRNA^{Phe}$ observed in the common type (Fig. 1). Thus, such an orientation and gene order in the species may have evolved independently after the species acquired species status, or it might be a common feature of the taxon to which the species belonged (e.g., the family Tenebrionidae). Nevertheless, considering the diversity of Coleoptera, conservation in orientation and gene order in most taxonomic groups is worthy of comment, in that many insect orders evidence either order-specific rearrangement (e.g., Lepidoptera) (Boore, 1999; Kim et al., 2006; Salvato et al., 2008) or markedly diverse gene rearrangement (e. g., Hymenoptera) (Cha et al., 2007; Crozier and Crozier, 1993; Dowton et al., 2003; Hong et al., 2008).

Genes overlap in the *P. hilaris* mitogenome in a total of 57 bp in 15 locations, and the longest overlaps are 8 bp, and are located in two places between $tRNA^{Trp}$ and $tRNA^{Cys}$, and $tRNA^{Trp}$ and COI (Table 1). The *P. hilaris* mitochondrial genes harbor a total of 35 bp of intergenic spacer sequences, which are spread over 10 regions, ranging in size from 1 to 17 bp, with the longest located between $tRNA^{Ser}$(UCN) and ND1 (Table 1). Similarly-sized intergenic spacer sequences (16-22 bp) located between $tRNA^{Ser}$(UCN) and ND1 have been detected in the majority of sequenced coleopteran insects, although that of *Rhagophthalmus lufengensis* and *R. ohbai* are unusually

Table 1. Summary of mitochondrial genome of *Psacothea hilaris*

Gene	Direction	Nucleotide number	Size (bp)	Anticodon	Codon	
					Start	Stop
$tRNA^{Ile}$	F	1-67	67	GAT (30-32)		
$tRNA^{Gln}$	R	69-137	69	TTG (105-107)		
$tRNA^{Met}$	F	137-205	69	CAT (167-169)		
ND2	F	206-1216	1011		ATT	TAA
$tRNA^{Trp}$	F	1215-1282	68	TCA (1255-1257)		
$tRNA^{Cys}$	R	1275-1336	62	GCA (1305-1307)		
$tRNA^{Tyr}$	R	1338-1402	65	GTA (1370-1372)		
COI	F	1395-2942	1548		ATT	TAA
$tRNA^{Leu*}$(UUR)	F	2938-3002	65	TAA (2967-2969)		
COII	F	3003-3690	688		ATC	T (tRNA)
$tRNA^{Lys}$	F	3691-3760	70	TTT (3721-3723)		
$tRNA^{Asp}$	F	3760-3825	66	GTC (3791-3793)		
ATPase8	F	3826-3981	156		ATT	TAG
ATPase6	F	3975-4649	675		ATG	TAA
COIII	F	4648-5437	789		ATG	TAA
$tRNA^{Gly}$	F	5440-5503	64	TCC (5470-5472)		
ND3	F	5501-5857	357		ATA	TAG
$tRNA^{Ala}$	F	5856-5920	65	TGC (5885-5887)		
$tRNA^{Arg}$	F	5917-5983	67	TCG (5948-5950)		
$tRNA^{Asn}$	F	5981-6045	64	GTT (6012-6014)		
$tRNA^{Ser}$(AGN)	F	6044-6112	68	TCT (6082-6084)		
$tRNA^{Glu}$	F	6113-6176	64	TTC (6143-6145)		
$tRNA^{Phe}$	R	6179-6242	64	GAA (6210-6212)		
ND5	R	6244-7962	1720		ATA	T (tRNA)
$tRNA^{His}$	R	7960-8022	63	GTG (7987-7989)		
ND4	R	8023-9355	1333		ATG	T (tRNA)
ND4L	R	9349-9636	288		ATG	TAA
$tRNA^{Thr}$	F	9639-9702	64	TGT (9669-9671)		
$tRNA^{Pro}$	R	9703-9766	64	TGG (9735-9737)		
ND6	F	9769-10272	504		ATT	TAA
CytB	F	10272-11411	1140		ATG	TAA
$tRNA^{Ser*}$(UCN)	F	11415-11483	69	TGA (11445-11447)		
ND1	R	11501-12448	948		ATA	TAG
$tRNA^{Leu}$(CUN)	R	12453-12517	65	TAG (12486-12488)		
lrRNA	R	12518-13789	1272			
$tRNA^{Val}$	R	13790-13858	69	TAC (13826-13828)		
srRNA	R	13859-14666	808			
A+T-rich region	R	14667-15856	1190			

Direction of the genes is presented as F for forward and R for reverse direction.
tRNA abbreviations follow the IUPAC-IUB three-letter code.

short (Fig. 2). Furthermore, a 5 bp-long motif sequence (TAGTA) has been detected in all sequenced coleopteran insects, including *P. hilaris* (Fig. 2). Within the space sequence, the 5-bp consensus sequence was suggested as the possible binding site for mtTERM, the transcription termination peptide, with the consideration that the intergenic spacer sequence can be detected at the end site of the major-strand coding region in the circular mtDNA (Taanman, 1999). Furthermore, Cameron and Whiting (2008) previously detected the 7-bp motif that extends by one-bp at both ends from the 5-bp motif from several insect orders, including all sequenced lepidopteran insects.

Another longer intergenic spacer sequence can be detected in the *Pyrocoelia rufa* mitogenome, wherein a 1,724-bp long tandem repeat unit, comprised of 12 134-bp tandem repeats plus one incomplete 116 bp repeat, is located between ND2 and $tRNA^{Trp}$ (Bae et al., 2004). The next longest intergenic

Table 2. Characteristics of the coleopteran mitogenomes sequenced in their entirety and near entirety

Taxon	Size (bp)	A+T %	No. codons[a]	PCG[b] A+T %	lrRNA		srRNA		A+T-rich region		GenBank accession no.	References
					Size (bp)	A+T %	Size (bp)	A+T %	Size (bp)	A+T %		
Polyphaga												
Elateriformia												
Buprestoidea												
Buprestidae												
Chrysochroa fulgidissima	15,592	70.0	3,718	68.8	1,302	73.6	797	71.8	875	76.7	EU826485	Hong et al. (2009)
Scirtoidea												
Scirtidae												
Cyphon sp.	15,919	75.2	3,709	72.8	1,297	80.8	781	82.1	1,043	85.2	EU877949	Sheffield et al. (2008)
Elateroidea												
Elateridae												
Pyrophorus divergens	16,120	69.4	3,707	67.5	1,269	72.8	805	72.4	1,470	75.0	EF398270	Arnoldi et al. (2007)
Phengodidae												
Rhagophthalmus lufengensis	15,982	79.6	3,697	78.1	1,270	82.2	786	83.3	1,367	86.8	DQ888607	Li et al. (2007)
Rhagophthalmus ohbai	15,704	79.1	3,692	77.5	1,277	82.4	785	83.5	1,044	85.3	AB267275	Li et al. (2007)
Lampyridae												
Pyrocoelia rufa	17,739	77.4	3,691	76.3	1,242	81.6	765	81.8	1,522	87.6	AF452048	Bae et al. (2004)
Cucujoiformia												
Cucujoidea												
Phloeostichidae												
Priasilpha obscura	16,887	76.5	3,698	73.6	1,283	81.4	788	81.3	2,028[c]	87.0	EU877952	Sheffield et al. (2008)
Chrysomeloidea												
Cerambycidae												
Anoplophora glabripennis	15,774	78.3	3,700	77.0	236	71.2	806	78.0	1,115	88.0	DQ768215	An et al. (Unpublished)
Psacothea hilaris	**15,856**	**76.7**	**3,720**	**75.9**	**1,273**	**80.0**	**808**	**80.6**	**1,190**	**78.6**	**FJ424074**	**This study**
Chrysomeloidae												
Crioceris duodecimpunctata	15,880	76.9	3,694	75.2	1,277	82.5	804	79.6	1,220	83.3	AF467886	Stewart and Beckenbach (2003)
Tenebrionoidea												
Tenebrionidae												
Tribolium castaneum	15,881	71.7	3,698	69.2	1,280	76.0	774	76.2	1,239	82.5	AJ312413	Friedrich and Muquim (2003)
Cleroidea												
Melyridae												
Chaetosoma scaritides	15,511	79.1	3,685	77.3	1,282	82.4	779	82.2	862[c]	91.0	EU877951	Sheffield et al. (2008)
Myxophaga												
Sphaerioidea												
Sphaeriusidae												
Sphaerius sp.	15,735	80.8	3,709	79.8	1,315	84.1	784	84.1	953[c]	89.6	EU877950	Sheffield et al. (2008)
Adephaga												
Caraboidea												
Trachypachidae												
Trachypachus holmbergi	15,991	79.5	3,726	78.4	1,321	81.9	788	82.5	1,119[c]	84.9	EU877954	Sheffield et al. (2008)
Archostemata												
Cupedoidea												
Ommatidae												
Tetraphalerus bruchi	15,689	70.0	3,705	65.8	1,322	68.1	791	64.2	1,000	78.4	EU877953	Sheffield et al. (2008)

[a]Termination codons were excluded in total codon count.
[b]Protein coding genes.
[c]Incomplete A+T-rich region.

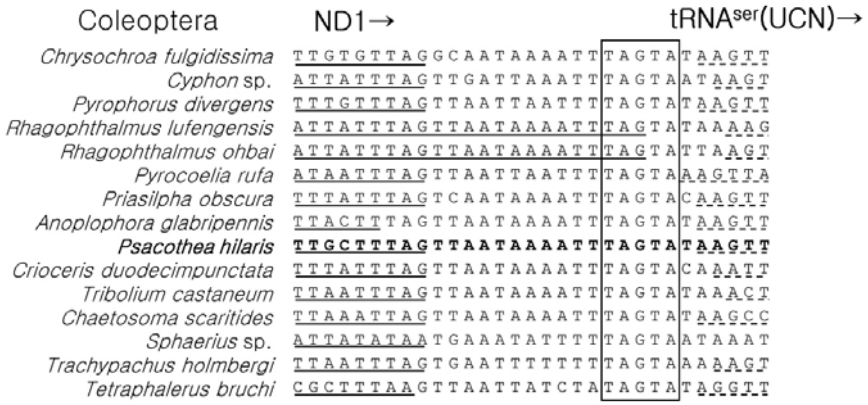

Fig. 2. Alignment of the internal spacer region located between tRNA^{Ser}(UCN) and ND1, in which their complete or nearly complete mitogenomes are available. The boxed nucleotides indicate the conserved pentanucleotide region (TAGTA) found in all sequenced coleopteran insects. Underlined and dotted nucleotides, respectively, indicate the adjacent partial sequences of ND1 and tRNA^{Ser}(UCN).

Fig. 3. Alignment of the initiation context of the COI genes of coleopteran insects, including that of *Psacothea hilaris*. The first five to seven codons are shown in uppercase letters on the right-hand side of the figure. Underlined nucleotides indicate the adjacent partial sequence of tRNA^{Tyr}. Arrows indicate the transcriptional direction. Boxed nucleotides indicate the currently proposed translation initiators for the COI gene of insects. The start codon for *P. hilaris* was designated as ATT.

spacer sequences are the 203 bp-long sequence detected between tRNA^{Tyr} and COI of the *P. obscura* mitogenome and the 177-bp long sequence located between ND2 and tRNA^{Trp} of the *Cyphon* sp. mitogenome (Sheffield et al., 2008). Excluding this, the majority of intergenic space sequences in the coleopteran insects are far less than 20 bp in the majority of complete coleopteran mitogenomes (data not shown).

PCGs

All PCGs of *P. hilaris* began with typical ATN codons: three with ATA, four with ATT, five with ATG, and one with ATC (Table 1). Abnormal translational starts of COI (e.g., tetranucleotides or hexanucleotides) have been reported in several species of diverse insect orders, including Diptera, Lepidoptera, and Orthoptera. Recently, with the inclusion of several newly sequenced coleopteran mitogenomes in the current insect mitogenome data, Sheffield et al. (2008) proposed that it would be most logical to select a start codon for COI that would minimize intergenic space and gene overlap. Based on this, they suggested that AAT or AAC, both of which designate asparagine as a start codon for COI, as the first non-overlapping in-frame codon in COI, AAT or AAC, is well-conserved throughout several divergent superfamilies within the Polyphaga (Sheffield et al., 2008). In fact, our newly sequenced *P. hilaris,* which belongs to the Polyphaga, also harbors the AAC within the frame. However, the *P. hilaris* COI gene harbors the typical start codon ATT, with an 8 bp overlap with the beginning of the precedent tRNA^{Tyr} (Fig. 3). With the exception of this, no other typical start codon is available in the start region of the *P. hilaris* COI gene and the precedent tRNA^{Tyr}. Although this proposal by Sheffield et al. (2008) appears reasonable, more data may be required for a decisive conclusion, as many PCGs of insect mitogenomes evidence an overlapping nature between neighboring genes. Furthermore, the typical start codon ATT, which overlaps with the precedent tRNA^{Tyr} in the *P. hilaris* COI gene, may continue to function if not hampered by an unknown mechanism.

10 of the 13 PCGs harbor a complete termination codon (seven TAA and three TAG), but the remaining three harbor the incomplete termination codons T: abutting (COII) or overlapping (ND5 and ND4) the neighboring tRNAs (Table 1). The most common interpretation of this phenomenon holds that TAA termini are created via post-transcriptional polyadenylation, thereby resulting in the functional complete stop codon, TAA (Anderson et al., 1981; Ojala et al., 1981).

The genome-wise A+T bias is also reflected in the codon usage of the *P. hilaris* mitogenome (Table 3). Relatively synonymous codon frequencies (RSCU) have demonstrated that codons with A or T in the third position were always overused as compared to other synonymous codons. For example, the codons GTC (Val) and GTG (Val) were utilized only 10 times for valine, representing a total RSCU of 0.27, but the synonymous codons GTT and GTA for valine were profoundly overused, representing a total RSCU of 4.94 (Table 3). Furthermore, the codons TTT (Phe), TTA (Leu), ATT (Ile), and ATA (Met) are

Table 3. Codon usage in the *Psacothea hilaris* protein coding genes

a. a.	Codon	%	n	RSCU	a. a.	Codon	%	n	RSCU
Phe (F)	UUU	8.5	315	1.82	Tyr (Y)	UAU	3.96	147	1.8
	UUC	0.86	32	0.18		UAC	0.43	16	0.2
Leu*(L)	UUA	11.92	442	4.36	Stop (*)	UAA	-	-	-
	UUG	0.65	24	0.24		UAG	-	-	-
	CUU	2.29	85	0.84	His (H)	CAU	1.75	65	1.86
	CUC	0.19	7	0.07		CAC	0.13	5	0.14
	CUA	1.16	43	0.42	Gln (Q)	CAA	1.73	64	1.76
	CUG	0.19	7	0.07		CAG	0.24	9	0.24
Ile (I)	AUU	10.09	374	1.93	Asn (N)	AAU	4.45	165	1.82
	AUC	0.35	13	0.07		AAC	0.43	16	0.18
Met (M)	AUA	5.23	194	1.83	Lys (K)	AAA	2.99	111	1.79
	AUG	0.49	18	0.17		AAG	0.35	13	0.21
Val (V)	GUU	2.89	107	2.22	Asp (D)	GAU	1.7	63	1.91
	GUC	0.16	6	0.12		GAC	0.08	3	0.09
	GUA	2.05	76	1.57	Glu (E)	GAA	2.18	81	1.86
	GUG	0.11	4	0.08		GAG	0.16	6	0.14
Ser (S)	UCU	2.75	102	1.95	Cys (C)	UGU	1.0	37	1.85
	UCC	0.3	11	0.21		UGC	0.08	3	0.15
	UCA	2.54	94	1.8	Trp (W)	UGA	2.35	87	1.9
	UCG	0.05	2	0.04		UGG	0.13	5	0.1
Ser* (S)	AGU	0.49	18	0.56	Arg (R)	CGU	0.38	14	1.03
	AGC	0.03	1	0.03		CGC	-	-	-
	AGA	2.7	100	3.07		CGA	1.02	38	2.76
	AGG	0.3	11	0.34		CGG	0.08	3	0.22
Thr (T)	ACU	2.27	84	1.92	Pro (P)	CCU	1.89	70	2.17
	ACC	0.35	13	0.3		CCC	0.57	21	0.66
	ACA	1.91	71	1.62		CCA	0.97	36	1.11
	ACG	0.19	7	0.16		CCG	0.05	2	0.06
Ala (A)	GCU	2.29	85	2.04	Gly (G)	GGU	1.43	53	1.06
	GCC	0.7	26	0.63		GGC	0.19	7	0.14
	GCA	1.46	54	1.3		GGA	3.26	121	2.41
	GCG	0.03	1	0.03		GGG	0.54	20	0.4

A total of 3,720 codons for *P. hilaris* were analyzed, excluding the termination codons.
a. a., amino acid; RSCU, relative synonymous codon usage; n = frequency of each codon.
L, L*, S, and S* indicate $tRNA^{Leu}$(CUN), $tRNA^{Leu}$(UUR), $tRNA^{Ser}$(AGN), and $tRNA^{Ser}$(UCN), respectively.

the four most frequently used codons in the *P. hilaris* PCGs, accounting for 35.74% (Table 3). These codons are all comprised of A or T nucleotides, which is indicative of the biased usage of A and T nucleotides in the *P. hilaris* PCGs. A comparative analysis among coleopteran insects also revealed a similar pattern (data not shown).

tRNA and rRNA genes

The *P. hilaris* mitogenome has a typical set of 22 tRNA genes that are interspersed between the rRNA and the PCGs (Fig. 4). All *P. hilaris* tRNAs can be folded into the typical clover-leaf structure, with the exception of $tRNA^{Ser}$(AGN), the dihydrouridine (DHU) arm of which forms a simple loop (Fig. 4). Such abnormalities are also observed in many of the $tRNA^{Ser}$(AGN) of metazoan mitogenomes, including insects (Wolstenholme, 1992). The *P. hilaris* tRNAs harbor a total of eight mismatches in the stem region (Table 4): three each of U-U and A-A, one A-C, and one C-U. Thus, six of them definitely involve A-T nucleotide substitutions, thus signifying the biased usage of A and T nucleotides in the *P. hilaris* tRNAs. This number is somewhat higher than those observed with other available coleopteran insects, in which it ranged between 3-5 (*R. ohbai*, *R. lufengensis*, *P. divergens*, and *P. rufa*), but was much lower than that noted in *C. fulgidissima*, which harbors a total of 17 mismatches. The *P. hilaris* tRNAs range in size from 62 ($tRNA^{Cys}$) to 70 bp ($tRNA^{Lys}$) (Table 4). All *P. hilaris* tRNAs invariably harbor 7 bp sequences in the amino-acyl stem, 5 bp in the anticodon stem, and 7 bp in the anticodon loop, but other portions of the tRNAs are variable, particularly within the DHU and TΨC loops. As is the case with all other insect mitogenome sequences, two rRNA genes were detected in *P. hilaris*. These were located between $tRNA^{Leu}$ (CUN) and $tRNA^{Val}$, and between $tRNA^{Val}$ and the A+T-rich region, respectively (Table 1). The size and the A+T content of the two rRNAs are well

Table 4. Size of each region of *Psacothea hilaris* tRNAs

Region	I	Q	M	W	C	Y	L*	K	D	G	A	R	N	S	E	F	H	T	P	S*	L	V
	67	69	69	68	62	65	65	70	66	64	65	67	64	68	64	64	63	64	64	69	65	69
Amino-acyl stem	7	7	7	7	7	7	7^1	7	7	7	7	7^1	7	7	7	7	7	7	7	7	7	7
DHU stem	3	4	4	4	3	3	3	3	4	3	4	4	3	-	4	4	4	4	4	4	3	4
Anticodon stem	5	5^1	5^1	5	5	5	5^1	5^1	5	5	5	5	5	5^1	5	5	5	5	5	5	5^1	5
TΨC stem	3	5	5	5	3	4	4	4	4	4	5	4	3	5	4	4	4	4	4	5	5	5
DHU loop	6	5	5	5	4	7	6	7	5	7	4	3	7	12	5	5	5	5	5	7	4	7
Anticodon loop	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
Variable loop	5	4	4	4	4	4	4	5	4	4	4	5	5	9	4	4	4	4	4	4	5	4
TΨC loop	9	7	7	7	11	5	6	9	5	4	4	5	5	3	3	4	3	4	5	7	4	7

Transfer RNA genes are labeled by one-letter symbol according to the IUPAC-IUB single little amino acid codes.
L, L*, S and S* indicate $tRNA^{Leu}$(CUN), $tRNA^{Leu}$(UUR), $tRNA^{Ser}$(AGN), and $tRNA^{Ser}$(UCN), respectively.
Superscripts indicate number of mismatch in each region.

within the range of the respective genes detected in the coleopteran mitochondrial rRNAs (Table 1).

A+T-rich region

A large non-coding region of 1,190 nucleotides is located between srRNA and $tRNA^{Ile}$. This region is A+T-rich (78.6%) and, in fact, harbors the highest A+T content levels among any of the regions of the *P. hilaris* mitogenome (Table 2). The *P. hilaris* A+T-rich region harbors a tandem repeat comprised of 57 bp-long seven identical copies, with a total length of 399 bp, beginning from nucleotide number 14,676 bp with regard to the *P. hilaris* mitogenome (Fig. 5). The repeat sequence is comprised of 65% A+T (37 bp) and 35% G+C nucleotides (20 bp). Thus, the A+T content is somewhat lower than that of the entire A+T-rich region (78.6%). A BLAST search to detect any relationship of the sequence to other organisms or sequences proved relatively unsuccessful. The matched sequences originated largely from bacterial and human chromosomal sequences, and most were too short (> 15 bp), and evidenced low identity, although our findings may provide some signal as to the relevant relationships.

The presence of tandem repeats in the mitochondrial A+T-rich region has been frequently reported in other insects. The 747-bp long *B. mandarina* A+T-rich region harbors a tandem triplication of a 126-bp fragment composed of identical first and second copies and one nucleotide-substituted third copy (Yukuhiro et al., 2002). In the case of *Locusta migratoria*, the 875-bp A+T-rich region harbors two G+C-rich repetitive sequences, composed of 155 bp and 146 bp with a sequence homology between them of 89.7%. Tandem repeats have often been detected in other insect A+T-rich regions, but it was only recently that such repeats have been reported in the coleopteran mitogenome. Sheffield et al. (2008) reported six coleopteran mitogenomes, and three of them harbor such repeats. For example, the *T. holmbergi* A+T-rich region harbors 21 almost identical copies of a 58-bp sequence, and one copy that lacks the final 10 bp (Sheffield et al., 2008). Interestingly, the first 10 tandemly repeated copies and the second 11 tandemly repeated copies are separated by 278 bp of typical sequences of the A+T-rich region containing higher A+T content in the *T. holmbergi* A+T-rich region. Also, the *P. obscura* A+T-rich region harbors six almost identical copies of a 132-bp sequence, and one copy lacking the last 79 bp. Finally, *C. scaritides* also harbors three ~96 bp-long similar tandemly repeated copies. Considering that these three coleopteran mitogenomes are incomplete in the A+T-rich region, the magnitudes of the tandem repeat sequence in those species remain unknown. With the exception of those repeats, the majority of the *P. hilaris* A+T-rich region is composed of non-repetitive sequences, but includes a 12 bp-long poly-T stretch and a 13 bp-long poly-A stretch, as well as several microsatellite-like TA repeats scattered throughout the region.

The *P. hilaris* A+T-rich region harbors several stretches of sequences with the potential to form stem-and-loop structures, formed by stems with perfect matches of varying nucleotide pairs and loops of various sizes (Fig. 6). It has been previously suggested that the sequence flanking the stem-and-loop structure is highly conserved among several insect orders, in contrast to the primary sequence divergence in the stem-and-loop structure: "TATA" consensus sequences at the 5′ end and "GAA(A)T" consensus sequences at the 3′ end (Schultheis et al., 2002; Zhang et al., 1995). The presence of such structures in the conserved flanking sequences has been suggested to be of secondary strand-replication origin (Zhang et al., 1995). Thus, such stem-and-loop structures, along with the flanking sequence in the A+T-rich region, have been detected in a variety of insect orders, including the Orthoptera, Lepidoptera, Diptera, Plecoptera, and Hymenoptera (Brehm et al., 2001; Cha et al., 2007; Crozier and Crozier, 1993; Hong et al., 2008; Inohira et al., 1997; Monforte et al., 1993; Rand and Harrison, 1989; Schultheis et al., 2002; Zhang et al., 1995). Nevertheless, *P. hilaris* harbors no structures containing such conserved flanking sequences at both ends of the stem-and-loop structures. If it must be present for a functional role in the A+T-rich region, the *P. hilaris* mitogenome may have evolved to harbor some alternative flanking sequence forms.

One of the more unusual features of the *P. hilaris* mitogenome is the presence of two tRNA-like structures: one $tRNA^{Arg}$-like sequence and one $tRNA^{Lys}$-like sequence, both of which are encoded in the major strand (Fig. 7). The detection of the proper anticodon TCG in the case of the $tRNA^{Arg}$-like sequence (Fig. 7A) or TTT in the case of one $tRNA^{Lys}$-like sequence (Fig. 7B), in addition to the formation of clover-leaf structures, suggests that they may be functional, but the large number of mismatches present in the stem regions cast doubt on the functionality of the structures. The 58% sequence identity noted between the regular $tRNA^{Arg}$ and $tRNA^{Arg}$-like sequence and the 52% sequence identity between the regular $tRNA^{Lys}$ and $tRNA^{Lys}$-like sequence are indicative of the presence of substantial sequence divergence between the regular tRNAs and tRNA-like sequences. One possible mechanism by which the tRNA-like sequences can be incorporated in the A+T-rich region may be an illicit priming of a

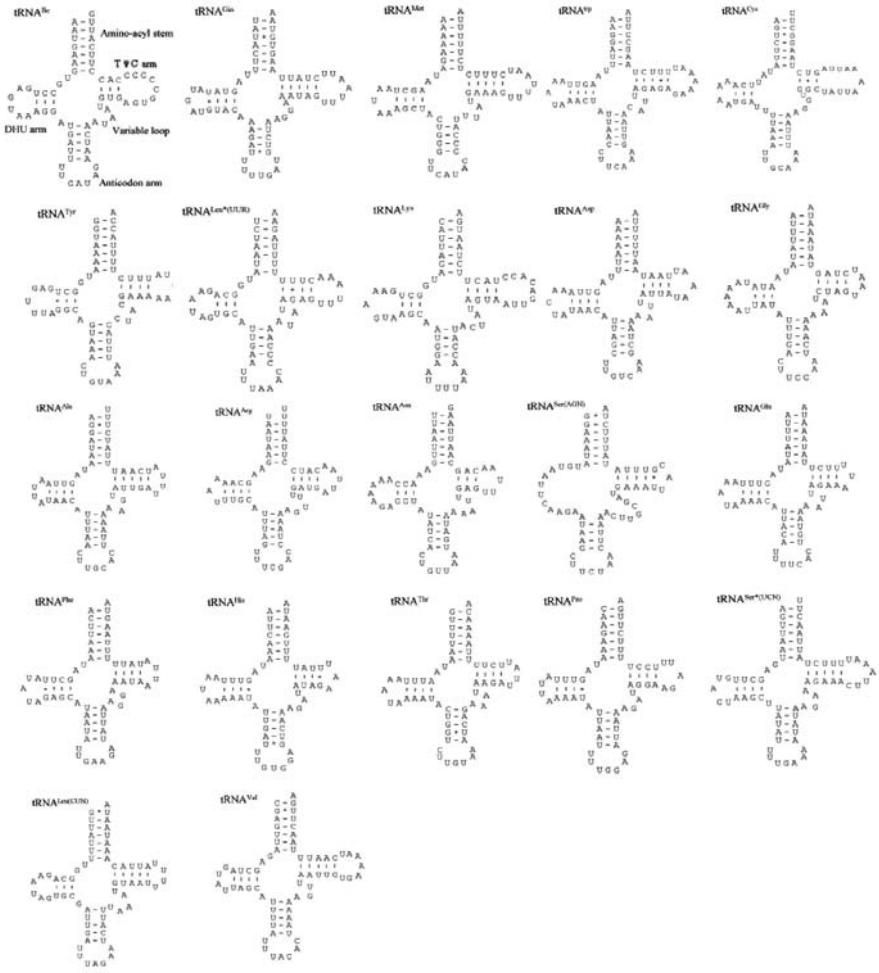

Fig. 4. Predicted secondary clover-leaf structures for the 22 tRNA genes of *Psacothea hilaris*. The tRNAs are labeled with the abbreviations of their corresponding amino acids. Nucleotide sequences from 5′ to 3′ are indicated for $tRNA^{Ile}$. Dashes (-) indicate Watson-Crick base-pairing, and centered asterisks (*) indicate G-U base-pairing. Arms of tRNAs (clockwise from top) indicate the amino acid acceptor (AA) arm, TΨC (T) arm, the anticodon (AC) arm, and the dihydro-uridine (DHU) arm.

Repeat 1	14676	TTAAAATTTCATCAGCTAGCTCCGCTATATAAAATCGCCTACCTTTAAAATTTCCCC	14733
Repeat 2	14734	.....	14790
Repeat 3	14791	.....	14847
Repeat 4	14848	.....	14904
Repeat 5	14905	.....	14961
Repeat 6	14962	.....	15018
Repeat 7	15019	.....	15075

Fig. 5. Alignment of tandem repeat units detected in the *Psacothea hilaris* A+T-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 *P. hilaris*.

tRNA molecule within the A+T-rich region, and the subsequent failure to cleave the tRNA primer from the nascent DNA strand, thereby ultimately resulting in the incorporation of a tRNA gene into the mitogenome (Cantatore et al., 1987). The divergence of the tRNA-like structure from the corresponding regular tRNAs and the presence of several mismatches detected in the stem region may be explained in terms of the evolutionary nature of the A+T-rich region. The presence of tRNA-like sequences within the A+T-rich region has also been reported in the hymenopteran species, *Apis mellifera* (Coizier and Crozier, 1993), *Bombus ignitus* (Cha et al., 2007), *B. hypocrita sapporoensis* (Hong et al., 2008), a lepidopteran species, *Antheraea yamamai* (Kim et al., 2009), and a coleopteran species, *C. fulgidissima* (Hong et al., 2009). Thus, it appears that an incorporation of tRNAs in the region is a frequent event in insects, although further extensive investigation into this phenomenon may be required.

Phylogenetic analysis

Data from coleopteran mitogenomes, including that of *P. hilaris,* have allowed us to construct the phylogenetic relationships among four coleopteran suborders (Archostemata, Myxophaga, Adephaga, and Polyphaga) and among the two polyphagan infraorders, Elateriformia and Cucujiformia, comprising seven superfamilies using the concatenated amino acid sequences of 13 PCGs (Fig. 8). Among almost all possible relationships of the four coleopteran suborders, two of the most widely discussed were the placement of Polyphaga to the sister group of Myxophaga (Fig. 8A; Beutel, 1997; Beutel and Haas, 2000) and the placement of Polyphaga to the sister group of all remaining suborders (Fig. 8B; Kukalová-Peck and Lawrence, 1993; Lawrence and Newton, 1982). Our phylogenetic analysis among species of the four suborders positioned Polyphaga as a sister group to the remaining three suborders (Figs. 8D and 8E). The bootstrapping values at the node supporting the Archostemata,

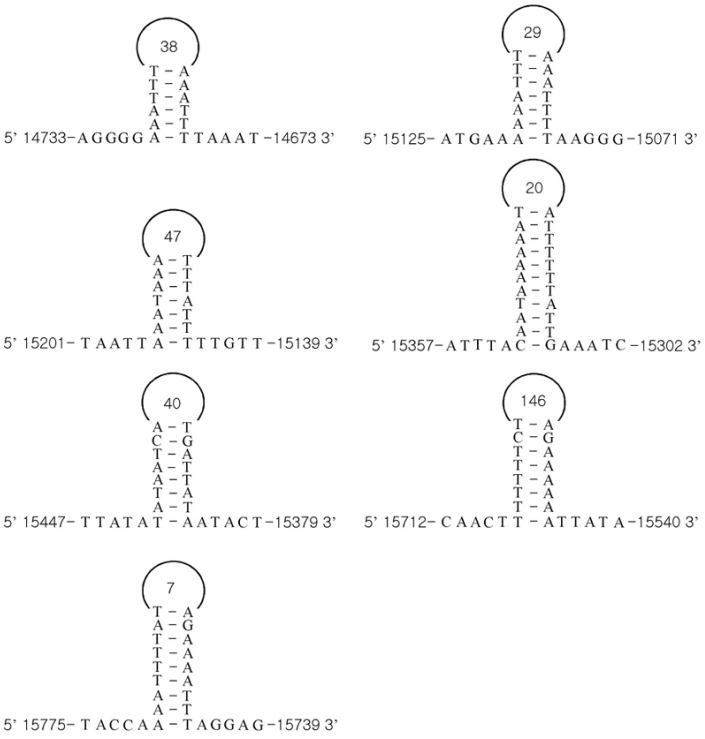

Fig. 6. Stem-and-loop structures in the *Psacothea hilaris* A+T-rich region. Nucleotide positions with regard to the *P. hilaris* mitogenome are presented at each end of the sequence. Numbers in the loop indicate the number of nucleotides forming the loop.

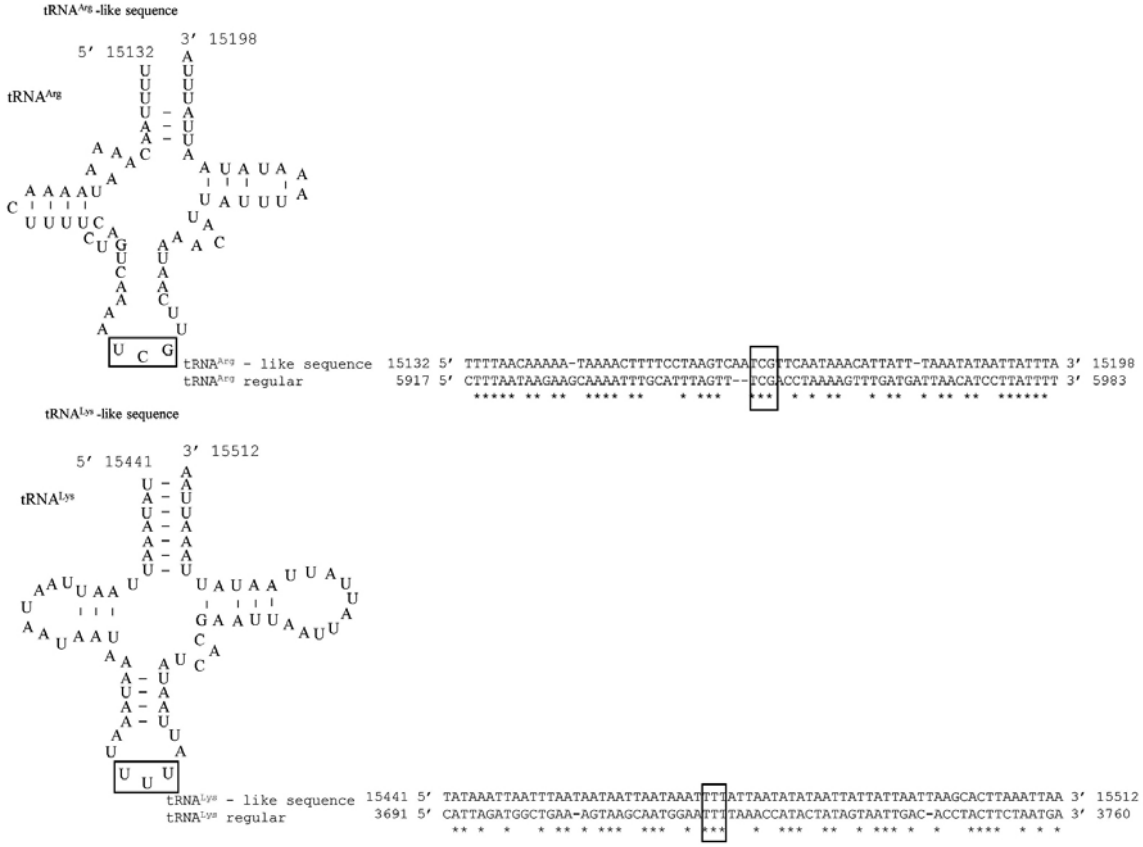

Fig. 7. Predicted secondary clover-leaf structure of (A) the $tRNA^{Arg}$ (UUR)-like sequence and (B) the $tRNA^{Lys}$-like sequence detected in the *Psacothea hilaris* A+T-rich region. Alignments with the corresponding regular tRNA sequences are also provided. The boxed nucleotides indicate the anticodon, which designates the corresponding tRNA. The asterisks indicate consensus sequences in the alignment.

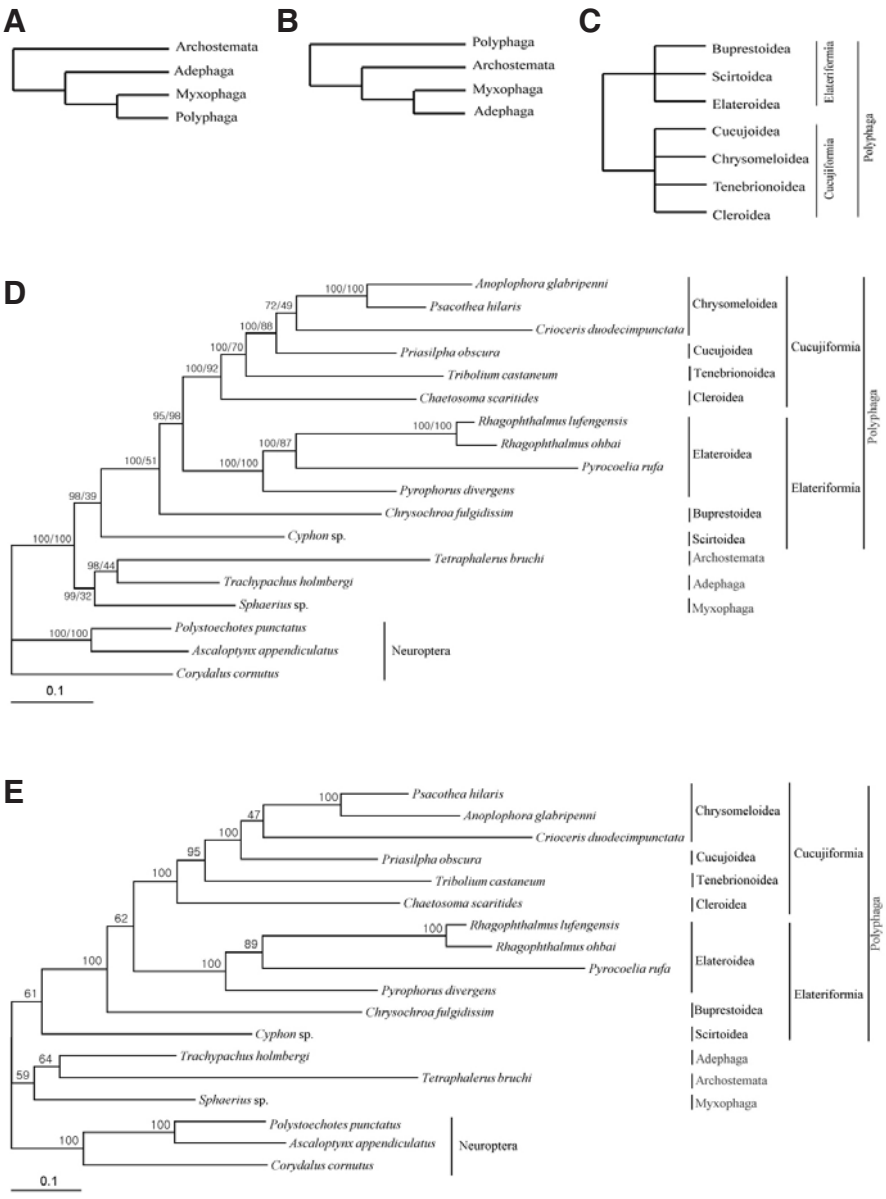

Fig. 8. Phylogeny of suborders of Coleoptera and polyphagan superfamilies. (A) One current hypothesis among coleopteran suborders (Beutel, 1997; Beutel and Haas, 2000). (B) Another current hypothesis among coleopteran suborders (Kukalová-Peck and Lawrence, 1993; Lawrence and Newton, 1982). (C) Current hypothesis of polyphagan superfamily relationships (Kukalova-Peck and Lawrence, 1993). (D) Bayesian Inference phylogram of Coleoptera obtained with amino acid sequences of concatenated 13 PCGs. The endopterygotan Neoptera, *Polystoechotes punctatus*, *Ascaloptynx appendiculatus*, and *Corydalus cornutus* (unpublished, GenBank accession numbers FJ171325, FJ171324, and FJ 171323, respectively) were utilized as outgroups. Numbers at each node specify BPP by BI analysis (first value) and bootstrap percentages of 1,000 pseudoreplicates by MP analysis (second value). The scale bar indicates the number of substitutions per site. (E) Phylogenetic tree of Coleoptera obtained with amino acid sequences of 13 concatenated PCGs via the ML method using PHYML. The endopterygotan Neoptera, *P. punctatus*, *A. appendiculatus*, and *C. cornutus* (unpublished, GenBank accession numbers FJ171325, FJ171324, and FJ171323, respectively) were used as outgroups. Numbers at each node specify bootstrap percentages of 100 pseudoreplicates. The scale bar indicates the num ber of substitutions per site.

Myxophaga, and Adephaga as a monophyletic group were quite low in our MP analysis (32%; Fig. 8D) and in the ML analysis (59%; Fig. 8E). However, the BPP acquired via BI was quite high, at 99% (Fig. 8D), supporting the hypothesis that Polyphaga is the sister group of all remaining suborders (Kukalová-Peck and Lawrence, 1993; Lawrence and Newton, 1982). Thus, our phylogenetic tree is more likely to support the sister group relationships between Polyphaga and all remaining suborders. From the perspective of morphology, the wing structure and the loss of the cervical sclerites were those primarily considered for the hypothesis that Polyphaga is the sister group to the remaining beetle suborders (Kukalová-Peck and Lawrence, 1993; Lawrence and Newton, 1982). Nevertheless, a recent molecular analysis using the 18S rRNA gene has suggested that Polyphaga is the sister group to Adephaga, thus resulting in the following topology: {[(Adephaga + Polyphaga) + Myxophaga] + Archostemata} (Caterino et al., 2002). Furthermore, a recent comprehensive phylogeny of Coleoptera also evidenced similar topology, in that Polyphaga is the sister group to Adephaga, with Myxophaga and Archostemata positioned at their base (Hunt et al., 2007). As more mitogenome sequences become available, further conclusive results can be obtained.

The suborder Polyphaga is subdivided into four infraorders: Elateriformia, Cucujiformia, Scarabaeiformia, and Bostrichiformia. Among them, full-length mitogenome sequences are available from the species of several superfamilies belonging to the Elateriformia (Buprestoidea, Scirtoidea, and Elateroidea) and Cucujiformia (Cucujoidea, Chrysomeloidea, Tenebroionoidea, and Cleroidea), including our *P. hilaris* within the Chrysomeloidea, but not from the Scarabaeiformia and Bostrichiformia. Although several lines of evidence have suggested that the infraorder Elateriformia and the infraorder Cucujiformia are respectively monophyletic (Fig. 8C; Beutel, 1995; Caterino et al., 2005; Lawrence, 1982) the relationships among superfamilies within the Elateriformia and Cucujiformia are not currently understood in great detail (Fig. 8C; Kukalova-Peck and Lawrence, 1993). Our phylogenetic analysis among superfamilies showed that the Cucujiformia is a strong monophyletic group, according to the results

of BI (100% BPP), MP (92% support) (Fig. 8D), and ML (100% support) (Fig. 8E). This result is consistent with a recent comprehensive phylogeny of Coleoptera, in that the Cucujiformia formed a strong monophyletic group, in addition to several other lines of evidence (Hunt et al., 2007). Within the Chrysomeloidea, *Anoplophora glabripennis* (unpublished, GenBank accession number DQ768215) and our *P. hilaris* belonging to the same family formed a strong group in all analyses (Figs. 8D and 8E). Among superfamilies within the Cucujiformia, our trees consistently provided a monophyletic Cucujiformia, evidencing a topology of {[(Chrysomeloidea + Cucujoidea) + Tenebroionoidea] + Cleroidea} (Figs. 8D and 8E). Thus, Cleroidea is positioned as the basal lineage of the remaining superfamilies. The placement of Cleroidea as a basal lineage among the superfamilies was also reported in a recent study using 18S and 28S rDNA sequences, wherein the Cleroidea, represented by the Cleridae and Melyridae, are positioned as the most basal lineage among the Cucujiformia superfamilies, although two among seven Cucujoidea families were unexpectedly grouped together with the Cleroidea (Marvaldi et al., 2008). Except for the splitting of Cucujoidea phylogenetic relationships among the superfamilies of Cucujiformia, as previously reported by Marvaldi et al. (2008), evidence the topology similar to that observed in this study.

In contrast to the typical phylogenetic hypothesis (Fig. 8C), the superfamilies of Elateriformia did not form a monophyletic group (Figs. 8D and 8E). Rather, all three analyses consistently placed the Scirtoidea and Buprestoidea represented by each one species at the bottommost and next basal lineages for the remaining Elateriformia, represented by four species of Elateroidea and all Cucujiformia (Figs. 8D and 8E). The divergence of Buprestoidea and the resultant lack of monophylety among Elateriformia was previously suggested by the complete mitogenome sequences of Coleoptera, including the Buprestoidea represented by *C. fulgidissima* (Hong et al., 2009), although this result contradicts the traditional view (Lawrence, 1982) as well as the results of recent molecular studies (Caterino et al., 2005). The separation of Scirtoidea from Elateriformia was suggested by the results of several studies. For example, Bocakova et al. (2007) found substantial distance and non-monophylety of Scirtoidea from the remaining Elateriformia in the molecular phylogenetic analysis of Elateriformia. Furthermore, a recent comprehensive phylogeny of Coleoptera also identified Scirtoidea as the most basal lineage for Polyphaga, rather than including it within the Elateriformia (Hunt et al., 2007), thereby demonstrating the divergent relationship of Scirtoidea from the remaining Elateriformia. Collectively, our phylogenetic analysis using the amino acid sequences of concatenated PCGs among the available coleopteran mitogenomes supports the notion of a possible sister group relationship of Polyphaga to all remaining suborders and a monophyletic Cucujiformia, with Cleroidea as the basal lineage. Additionally, we did not find adequate support for the notion of a monophyletic Elateriformia, identifying the Buprestoidea and Scirtoidea as the basal lineages of Cucujiformia and the remaining Elateriformia. Considering that the available mitogenome sequences of Coleoptera are somewhat limited, further decisive conclusions can be drawn as more complete coleopteran mitogenome sequences become available.

ACKNOWLEDGMENT

This work was fully supported by a 2007 grant from the National Institute of Biological Resources (Research title: The genetic evaluation of important biological resources) (No. 074-1800-1844-304).

REFERENCES

Abascal, F., Zardoya, R., and Posada, D. (2005). ProTest: selection of best-fit models of protein evolution. Bioinformatics *21*, 2104-2105.

Adachi, J., and Hasegawa, M. (1996). Model of amino acid substitution in proteins encoded by mitochondrial DNA. J. Mol. Evol. *42*, 459-468.

Akaike, H. (1974). A new look at the statistical model identification. IEEE Trans. Autom. Contr. *19*, 716-723.

Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijin, M.H.L., Droujn, A.R.J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., et al. (1981). Sequence and organization of the human mitochondrial genome. Nature *290*, 457-465.

Arnoldi, F.G., Ogoh, K., Ohmiya, Y., and Viviani, V.R. (2007). Mitochondrial genome sequence of the Brazilian luminescent click beetle *Pyrophorus divergens* (Coleoptera: Elateridae): mitochondrial genes utility to investigate the evolutionary history of Coleoptera and its bioluminescence. Gene *405*, 1-9.

Avise, J.C. (1994). Molecular markers, natural history and evolution (New York: Champman & Hall).

Bae, J.S., Kim, I., Sohn, H.D., and Jin, B.R. (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.

Beutel, R.G. (1995). Phylogenetic analysis of Elateriformia (Coleoptera: Polyphaga) based on larval characters. J. Zool. Syst. Evol. Res. *33*, 145-171.

Beutel, R.G. (1997). Uber phylogenese und evolution der Coleoptera (Insecta), insbesondere der Adephaga. Verh. Naturwiss. Ver. Hamburg *31*, 1-164.

Beutel, R., and Haas, F. (2000). Phylogenetic relationships of the suborders of Coleoptera (Insecta). Cladistics *16*, 103-141.

Bocakova, M., Bocak, L., Hunt, T., Teräväinen, M., and Vogler, A.P. (2007). Molecular phylogenetics of Elateriformia (Coleoptera): evolution of bioluminescence and neoteny. Cladistics *23*, 477-496.

Boore, J.L. (1999). Animal mitochondrial genomes. Nucleic Acids Res. *27*, 1767-1780.

Brehm, A., Harris, D.J., Hernandez, M., Cabrera, V.M., Larruga, J.M., Pinto, F.M., and Gonzalez, A.M. (2001). Structure and evolution of the mitochondrial DNA complete control region in the *Drosophila subobscura* subgroup. Insect Mol. Biol. *10*, 573-578.

Brodsky, L.I., Vasiliev, A.V., Kalaidzidis, Y.L., Osipov, Y.S., Tatuzov, A.R.L., and Feranchuk, S.I. (1992). GeneBee: the program package for biopolymer structure analysis. Dimacs *8*, 127-139.

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.

Cantatore, P., Gadaleta, M.N., Roberti, M., Saccone, C., and Wilson, A.C. (1987). Duplication and remodeling of tRNA genes during the evolutionary rearrangement of mitochondrial genomes. Nature *329*, 853-855.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic tool. Curr. Opin. Genet. Dev. *8*, 668-674.

Caterino, M.S., Shull, V.L., Hammond, P.M., and Vogler, A.P. (2002). Basal relationships of Coleoptera inferred from 18S rDNA sequences. Zool. Scr. *31*, 41-49.

Caterino, M.S., Hunt, T., and Vogler, A.P. (2005). On the constitution and phylogeny of Staphyliniformia (Insecta: Coleoptera). Mol. Phylogenet. Evol. *34*, 655-672.

Cha, S.Y., Yoon, H.J., Lee, E.M., Yoon, M.H., Hwang, J.S., Jin, B.R., Han, Y.S., and Kim, I. (2007). The complete nucleotide sequence and gene organization of the mitochondrial genome of the bumblebee, *Bombus ignitus* (Hymenoptera: Apidae). Gene *392*, 206-220.

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

Dowton, M., Castro, L.R., Campbell, S.L., Bargon, S.D., and Austin, A.D. (2003). Frequent mitochondrial gene rearrangements at the hymenopteran nad3-nad5 junction. J. Mol. Evol. *56*, 517-526.

Fauron, C.M.R., and Wolstenholme, D.R. (1980). Extensive diversity among *Drosophila* species with respect to nucleotide sequences

within the adenine+thymine-rich region of mitochondrial DNA molecules. Nucleic Acids Res. *8*, 2439-2452.
Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. Evolution *39*, 783-791.
Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. *3*, 294-299.
Friedrich, M., and Muquim, N. (2003). Sequence and phylogenetic analysis of the complete mitochondrial genome of the flour beetle *Tribolium castanaeum*. Mol. Phylogenet. Evol. *26*, 502-512.
Guindon, S., Lethiec, F., Duroux, P., and Gascuel, O. (2005). PHYML: online-a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. *33*, W557-559.
Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. *41*, 95-98.
Hebert, P.D.A., Cywinska, A., Ball, S.L., and deWaard, J.R. (2003). Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B *270*, 313-322.
Hong, M.Y., Lee, E.M., Jo, Y.H., Park, H.C., Kim, S.R., Hwang, J.S., Jin, B.R., Kang, P.D., Kim, K.-G., Han, Y.S., et al. (2008). Complete nucleotide sequence and organization of the mito-genome of the silk moth *Caligula boisduvalii* (Lepidoptera: Saturniidae) and comparison with other lepidopteran insects. Gene *413*, 49-57.
Hong, M.Y., Jeong, H.C., Kim, M.J., Jeong, H.U., Lee, S.H., and Kim, I. (2009). Complete mitogenome sequence of the jewel beetle, *Chrysochroa fulgidissima* (Coleoptera: Buprestidae). Mitochondrial DNA. (in press) (DOI 10.1080/19401730802644978).
Huelsenbeck, J.P., and Ronquist, F. (2001). MrBayes: Bayesian inference of phylogeny. Bioinformatics *17*, 754-755.
Hunt, T., Bergsten, J., Levkanicova, Z., Papadopoulou, A., St. John, O., Wild, R., Hammond, P.M., Ahrens, D., Balke, M., Caterino, M.S., et al. (2007). A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. Science *318*, 1913-1916.
Hwang, U.W., Friedrich, M., Tautz, D., Park, C.J., and Kim, W. (2001). Mitochondrial protein phylogeny joins myriapods with chelicerates. Nature *413*, 154-157.
Inohira, K., Hara, T., and Matsuura, E.T. (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.
Joyce, D.A., and Pullin, A.S. (2004). Using genetics to inform reintroduction strategies for the chequered skipper butterfly (*Carterocephalus palaemon*, Pallas) in England. J. Insect Conserv. *8*, 69-74.
Kim, I., Lee, E.M., Seol, K.Y., Yun, E.Y., Lee, Y.B., Hwang, J.S., and Jin, B.R. (2006). The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae). Insect Mol. Biol. *15*, 217-225.
Kim, S.R., Kim, M.I., Hong, M.Y., Kim, K.Y., Kang, P.D., Hwang, J.S., Han, Y.S., Jin, B.R., and Kim, I. (2009). The complete mitogenome sequence of the Japanese oak silkmoth, *Antheraea yamamai* (Lepidoptera: Saturniidae). Mol. Biol. Rep. (in press) (DOI 10.1007/s11033-008-9393-2).
Kukalová-Peck, J., and Lawrence, J.F. (1993). Evolution of the hind wing in Coleoptera. Can. Entomol. *125*, 181-258.
Lawrence, J.F. (1982). Coleoptera. In Synopsis and Classificatiion of Living Organisms, S. Parker, ed. (New York, USA: McGraw-Hill), pp. 482-553.
Lawrence, J.F., and Newton, A.F. (1982). Evolution and classification of beetles. Annu. Rev. Ecol. Syst. *13*, 261-290.
Lewis, D.L., Farr, C.L., and Kaguni, L.S. (1995). *Drosophila melanogaster* mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons. Insect Mol. Biol. *4,* 263-278.
Li, X., Ogoh, K., Ohba, N., Liang, X., and Ohmiya, Y. (2007). Mitochondrial genomes of two luminous beetles, *Rhagophthalmus lufengensis* and *R. ohbai* (Arthropoda, Insecta, Coleoptera). Gene *392*, 196-205.
Lowe, T.M., and Eddy, S.R. (1997). tRNA-scan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. *25*, 955-964.
Marvaldi, A.E., Duckett, C.N., Kjer, K.M., and Gillespie, J.J. (2008). Structural alignment of 18S and 28S rDNA sequences provides insights into phylogeny of Phytophaga (Coleoptera: Curculionoidea and Chrysomeloidea). Zool. Scr. *38*, 63-77.
Monforte, A., Barrio, E., and Latorre, A. (1993). Characterization of the length polymorphism in the A+T-rich region of the *Drosophila ob scura* group species. J. Mol. Evol. *36*, 214-223.
Moritz, C., Dowling, T.E., and Brown, W.M. (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu. Rev. Ecol. Syst. *18*, 269-292.
Murata, K., Satou, M., Matsushima, K., Satake, S., and Yamamoto, Y. (2004). Retrospective estimation of genetic diversity of an extinct oriental white stork (*Ciconia boyciana*) population in Japan using mitochondrial specimens and implications for reintroduction programs. Conserv. Genetics *5*, 553-560.
Nam, S.H. (1996). The insects of Korea (Seoul, Korea: Kyo-Hak Publishing Co.).
Nardi, F., Carapelli, A., Dallai, R., and Frati, F. (2003). The mitochondrial genome of the olive fly *Bactrocera oleae*: two haplotypes from distant geographical locations. Insect Mol. Biol. *12*, 605-611.
Ojala, D., Montoya, J., and Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. Nature *290*, 470-474.
Rand, D.M., and Harrison, R.G. (1989). Molecular population genetics of mtDNA size variation in crickets. Genetics *121*, 551-569.
Renfu, S., Nick, J.H., Campbell, H., and Barker, S.C. (2001). Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). Mol. Biol. Evol. *18*, 858-865.
Salvato, P., Simonato, M., Battisti, A., and Negrisolo, E. (2008). The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae). BMC Genomics *9*, 331.
Schultheis, A.S., Weigt, L.A., and Hendricks, A.C. (2002). Arrangement and structural conservation of the mitochondrial control region of two species of Plecoptera: utility of tandem repeat-containing regions in studies of population genetics and evolutionary history. Insect Mol. Biol. *11*, 605-610.
Shao, R., Campbell, N.J.H., and Barker, S.C. (2001). Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). Mol. Biol. Evol. *18*, 858-865.
Sheffield, N.C., Song, H., Cameron, S.L., and Whiting, M.F. (2008). A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. Mol. Biol. Evol. *25*, 2499-2509.
Stewart, J.B., and Beckenbach, A.T. (2003). Phylogenetic and genomic analysis of the complete mitochondrial DNA sequence of the spotted asparagus beetle *Crioceris duodecimpunctata*. Mol. Phylogenet. Evol. *26*, 513-526.
Swofford, D.L. (2002). PAUP*. Phylogenetic analysis using parsimony (*and other methods) ver 4.10 (Sunderland, USA: Sinauer Associates).
Taanman, J.W. (1999). The mitochondrial genome: structure, transcription, translation and replication. Biochim. Biophys. Acta *1410*, 103-123.
Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673-4680.
Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *24*, 173-216.
Wolstenholme, D.R. (1992). Animal mitochondrial DNA: structure and evolution. Int. Rev. Cytol. *141*, 173-216.
Yukuhiro, K., Sezutsu, H., Itoh, M., Shimizu, K., and Banno, Y. (2002). Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silk moth, *Bombyx mandarina*, and its close relative, the domesticated silk moth, *Bombyx mori*. Mol. Biol. Evol. *19*, 1385-1389.
Zhang, D., Szymura, J.M., and Hewitt, G.M. (1995). Evolution and structural conservation of the control region of insect mitochondrial DNA. J. Mol. Evol. *40*, 382-391.
Zhang, D.X., and Hewitt, G.M. (1997). Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. Biochem. Syst. Evol. *25*, 99-120.