The mitochondrial genomes of two walnut pests, *Gastrolina depressa depressa* and *G. depressa thoracica* (Coleoptera: Chrysomelidae), and phylogenetics within Chrysomeloidea (#21999)

First submission

Editor guidance

Please submit by **8 Dec 2017** for the benefit of the authors (and your $200 publishing discount).

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Custom checks
Make sure you include the custom checks shown below, in your review.

Raw data check
Review the raw data. Download from the location described by the author.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files
Download and review all files from the materials page.

9 Figure file(s)
3 Table file(s)
1 Other file(s)

Custom checks

DNA data checks

- Have you checked the authors data deposition statement?
- Can you access the deposited data?
- Has the data been deposited correctly?
- Is the deposition information noted in the manuscript?

Structure and Criteria

Structure your review

The review form is divided into 5 sections.
Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- Data is robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.
- Speculation is welcome, but should be identified as such.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

The mitochondrial genomes of two walnut pests, *Gastrolina depressa depressa* and *G. depressa thoracica* (Coleoptera: Chrysomelidae), and phylogenetics within Chrysomeloidea

Qiqi Wang [1], **Guanghui Tang** Corresp. [1]

[1] College of Forestry, Northwest A&F University, Yangling, Shaanxi, China

Corresponding Author: Guanghui Tang
Email address: tanggh@nwsuaf.edu.cn

With more than 63,000 described extant species, Chrysomeloidea is one of seven superfamilies in series Cucujiformia. In this study, the mitochondrial genomes (mitogenomes) of walnut leaf pests, *Gastrolina depressa depressa* and *G. depressa thoracica* were determined by Sanger sequencing technology. The complete and partial mitogenomes of *G. depressa thoracica* and *G. depressa depressa* were 16,109 bp and 14,277 bp in length respectively. The genomic analyses indicated that both mitogenomes harbored typical gene content and arrangement. The formerly identified elements, 'TAGTA' between *trnSer(UCN)* and *nad2*, and 'ATGATAA' between *atp8* and *atp6*, were firstly found to be more conserved than that between *nad4l* and *nad4*, which was 'ATGTTAA' in Coleoptera excluding Polyphaga. Phylogenetic trees inferred from the amino acid and nucleotide dataset for the 13 protein-coding genes of 36 species well supported the sister relationships of Eumolpinae + Cryptocephalinae + Cassidinae, as well as the monophyly of Galerucinae+ Chrysomelinae within Chrysomelidae.

The mitochondrial genomes of two walnut pests, *Gastrolina depressa depressa* and *G. depressa thoracica* (Coleoptera: Chrysomelidae), and phylogenetics within Chrysomeloidea

Qi-qi Wang [1] & Guang-hui Tang [1,*]

[1] **College of Forestry, Northwest Agriculture and Forestry University, Yangling, 712100, China**

[*] **Correspondence: tanggh@nwafu.edu.cn**

With more than 63,000 described extant species, Chrysomeloidea is one of seven superfamilies in series Cucujiformia. In this study, the mitochondrial genomes (mitogenomes) of walnut leaf pests, *Gastrolina depressa depressa* and *G. depressa thoracica* were determined by Sanger sequencing technology. The complete and partial mitogenomes of *G. depressa thoracica* and *G. depressa depressa* were 16,109 bp and 14,277 bp in length, respectively. The genomic analyses indicated that both mitogenomes harbored typical gene content and arrangement. The formerly identified elements, 'TAGTA' between *trnSer(UCN)* and *nad2*, and 'ATGATAA' between *atp8* and *atp6*, were firstly found to be more conserved than that between *nad4l* and *nad4*, which was 'ATGTTAA' in Coleoptera excluding Polyphaga. Phylogenetic trees inferred from the amino acid and nucleotide dataset for the 13 protein-coding genes of 36 species well supported the sister relationships of Eumolpinae + Cryptocephalinae + Cassidinae, as well as the monophyly of Galerucinae+ Chrysomelinae within Chrysomelidae.

Keywords Chrysomeloidea; *Gastrolina depressa depressa*; *G. depressa thoracica*; Mitochondrial genome; Phylogenetics

Introduction

Both *Gastrolina depressa thoracica* and *G. depressa depressa* belong to Chrysomelinae (Coleoptera: Chrysomeloidea) and widely distribute in most of walnut producing areas in China. The larvae and adults of them feed on walnut leaves, seriously affecting the growth and yield of walnut. They are typical species living in the boundary between Palaearctic and oriental regions, and the main feature is the flat back, which is different from other coleopterans. Compared with the 5–7 mm body length and black prothorax of *G. depressa thoracica*, *G. depressa depressa* has the body length of 6.5–8.3 mm and a yellow prothorax, therefore they were always distinguished by body colour and length (*Chen*, *1974*). Between 25°N and 35°N, *G. depressa thoracica* and *G. depressa depressa* are frequently found and always differentiated from each other according to different altitudes, which does not always work in some regions. To date, a few studies have focused on the systematics, occurrence and spatial distribution of *G. depressa thoracica* and *G. depressa depressa* (*Chen*, *1974*; *Ge et al*., *2003*; *Fan*, *2015*), but there is no report about their mitochondrial genomes (mitogenomes) and other molecular data.

Mitogenomes have been widely regarded as useful molecular markers for phylogenetics and phylogeographics, due to several properties such as maternal inheritance, conserved gene order and orientation, low recombination rate and high mutation rate (*Avise*, *1989*; *Crozier & Crozier*, *1993*; *Jin et al*., *2004*). Furthermore, the mitogenome from metazoans is a circular double-stranded molecule, ranging in length from 14 to 36 kb, and consists of 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs), 2 ribosomal RNA genes (rRNAs) and a large non-coding region, called control region or A+T-rich region (*Wolstenholme*, *1992*; *Boore et al*., *1998*). Over the last decade, increasing attention has been drawn onto the studies of insect mitogenomes due to the ease of recovering genetic information that may be relatively useful for research in insect species identification, population genetics and molecular evolution (*Ma et al*., *2012*; *Timmermans et al*., *2014*).

Beetles harbor as much as 380, 000 descrived species and an estimated total number of about three million members, becoming the largest order and the most diverse group in the Insecta (*ØDegaard*, *2000*; *Foottit & Adler*, *2009*). The enormous series Cucujiformia, containing more than half the species of Celeoptera, has reasonably been regarded as the most highly diversified order (*Crowson*, *1960*). It is divided into two superfamilies, Curculionoidea and Chrysomeloidea, and together form the second largest lineage of Phytophaga after the order Lepidoptera (*Grimaldi & Engel*, *2005*). The phylogeny and evolution of Chrysommeloidea have attracted considerable attention of systematists, and the interrelationships among the species from this family remain many uncertain or questions. Several previous studies have evaluated its phylogenetic relationships by using morphological data (*Farrell & Sequeira*, *2004*; *Gómez-Zurita et al*., *2007*), 18S rRNA (*Hunt et al*., *2007*), 28S rDNA (*Marvaldi et al*., *2009*), partial mitochondrial genes (*rrnL* and *cox1*) (*Bocak et al*., *2014*). However, few studies deliberately focused on the relationships of families in the hyperdiverse Chrysomeloidea (*Bocak et al*., *2014*; *Li et al*., *2016*).

In the present study, the mitochondrial nucleotide sequences of *G. depressa thoracica* and *G. depressa depressa* were determined and a detailed comparative investigation into the composition of mitogenomes for Chrysomelinae was conducted. In order to understand the taxonomic status of the two insects and the phylogenetic relationships in Chrysomeloidae, the 13 PCGs (nucleotide and amino acid sequences) from the two newly sequenced insects and other 34 mitochondrial sequences downloaded from GenBank were used for the reconstruction of phylogenetic trees. The mitogenomes of *G. depressa thoracica* and *G. depressa depressa* will be useful for further studies of molecular classification and coleopteran mitogenome architecture and phylogenetics.

Materials and Methods

Sample collection

Adults of *G. depressa thoracica* and *G. depressa depressa* were originally collected from the experimental station of Northwest A&F University, Shanyang County, Shaanxi Province, China (109°88′E, 33°53′N). After identified according to key morphological characteristics, all samples were stored in 100% ethanol immediately at –20℃.

DNA extraction and PCR amplification

Total genomic DNA was extracted from an individual insect following a standard phenol-chloroform extraction (*Tamura & Aotsuka*, 1988). The mitogenomes of *G. depressa thoracica* and *G. depressa depressa* were amplified using specific primer sets (Table 1) derived from the conserved regions of a multiple alignments of sequenced coleopteran insects.

PCR amplification was conducted using TaKaRa Ex Taq™ (Takara Co., Dalian, China) in 20 μL reactions containing 14 μL distilled water, 2 μL 10×Ex Taq buffer including $MgCl_2$, 1.6 μL dNTP mixture, 0.8 μL forward and reverse primers respectively, 0.1 μL TaKaRa Ex Taq and 0.7 μL template DNA. Amplification conditions consisted of an initial denaturation at 94 °C for 2 min; followed by 35 cycles at 94 °C for 30 s, annealing temperature for 30 s, 72 °C for 1 min; and a final extension at 72 °C for 10 min. Then, PCR products (1.1–1.4 Kb in length) were identified by agarose gel electrophoresis, and positive amplifications were sequenced in both directions.

Genome assembly and annotation

Sequences were checked by blasting in NCBI (https://www.ncbi.nlm.nih.gov/) and assembled using the program SeqMan in DNAStar package v7.1 (DNAStar Inc., Madison, WI, USA). The mitogenome sequences were analyzed using MEGA v5.1. The 13 PCGs were identified using ORF Finder (available on NCBI) and the rRNAs were determined by comparison with other coleopteran sequences. The tRNAs and their cloverleaf secondary structures were identified using tRNAscan-SE Search Online Server (*Lowe & Eddy*, *1997*), and undefined tRNAs were compared with those of other species, *Atrijuglans hetaohei*, *Dastarcus helophoroides* and *Anopheles minimus* (~~*Zhang et al*., *2015*; *Hua et al*., *2016*;~~ *Wang et al*., *2016*). RNAviz v2.0 (*De Rijk et al*.,

2003) was employed to draw secondary structures of tRNAs. The base composition of nucleotide sequences was described by skewness and was measured according to the following formulas (*Perna & Kocher*, *1995*): AT-skew = (A – T) / (A + T) and GC-skew = (G – C) / (G + C). The A+T content and relative synonymous codon usage (RSCU) were calculated using MEGA v5.1.

Phylogenetic analysis

The newly sequenced mitogenomes of *G. depressa thoracica* and *G. depressa depressa* were aligned with the publicly available mitogenome sequences of 34 mitochondrial genomes downloaded from GenBank (Table S1), with *Tetraphalerus bruchi* leoptera: Archostemata: Ommatidae) and *Abax parallelepipedus* (Adephaga: Caraboidea: Carabinae) as outgroups. From the mitogenome sequences of all the retrieved species, two separate datasets were generated for phylogenetic analyses: the nucleotide sequences of 13 PCGs, and the amino acid sequences translated from their corresponding nucleotide sequences using the invertebrate genetic code. Each of nucleotide and amino acid sequences of the 13 PCGs was aligned separately using ClustalW implemented in MEGA v5.1 with default options (*Tamura et al*., *2011*). Then, Gblocks v0.91b (*Talavera & Castresana*, *2007*) was used to reduce the final alignments, leading to better trees for phylogenetic reconstruction. The best-fit substitution models were estimatied using MrModeltest v2.3 (*Nylander*, *2004*) and ProtTest v2.4 (*Abascal et al*., *2005*) for each resulting nucleotide and amino acid sequences respectively.

For the construction of phylogenetic trees, the concatenated nucleotide and amino acid sequences were analyzed using maximum likelihood (ML) and Bayesian inference (BI) respectively. For ML analyses, PhyML v3.0 (*Guindon et al*., *2010*) was used with the best-fit model for each of the 13 PCGs in the nucleotide and the animo acid datasets, and bootstrap support (BS) at each node was performed using the rapid option with 1000 replicates. For Bayesian inference, MrBayes v3.2.6 was used (*Huelsenbeck & Ronquist*, *2001*). Four simultaneous Markov chains ran for 10 million generations, sampled every 1000 generations after discarding the first 25% “burn-in” trees. Node support was assessed based on Bayesian Posterior Probabilities (BPP). The consensus trees were viewed and edited by Figtree v1.4.3 (*Rambaut*, *2009*).

Results and Discussion

Gene content and nucleotide composition

~~The mitogenomes of *G. depressa thoracica* and *G. depressa depressa* were sequenced.~~ The complete *G. depressa thoracica* mitogenome was 16,109 bp in size, consisting of 13 PCGs, 2 rRNAs and 22 tRNAs, and an A+T-rich region (Fig 1, Table 2). In the mitogenome of *G. depressa thoracica*, there were four intergenic spacers (a total length of 23 bp), ranging from 1 bp to 17 bp, and the longest intergenic spacer was located between *trnSer(AGN)* and *nad1*. Furthermore, 15 pairs of genes overlapped each other, with a length ranging from 1 to 17 bp.

The sequenced mitogenome of *G. depressa depressa* was 14,277 bp long, and was comprised of 13 PCGs, 21 tRNAs, complete *rrnL* and partial *rrnS*. The whole A+T-rich region, *trnI* and partial *rrnS* failed to be sequenced although we have t several pairs of species-specific primers. There were 3 intergenic spacers

(23 bp), ranging from 1 bp to 17 bp, and the longest intergenic spacer was detected between *trnSer(AGN)* and *nad1*. 17 pairs of genes overlapped each other, with a length ranging from 1 to 20 bp.

Two 7-bp long overlaps (ATGATAA) have been detected in *G. depressa depressa* and *G. depressa thoracica*, which were also found in many other Polyphaga insects (Fig 2). These two overlaps were located between *atp8* and *atp6* on H-strand and between *nad4l* and *nad4* on L-strand, and were thought to be translated as a bicstron (*Stewart & Beckenbach*, *2005*). Furthermore, between *trnSer(UCN)* and *nad1*, a 5 bp long motif sequence (TAGTA) was detected in two newly sequenced mitogenomes, which was present in other coleopterans (Fig 3). This consensus sequence has been thought as the possible binding site of mtTERM as its location at the end of the H-strand coding region in the circular mitogenome (*Taanman*, *1999*). The motifs between *atp8* and *atp6*, and between *trnSer(UCN)* and *nad1* were relatively conserved in four suborders by the comparison of the complete mitogenomes of 41 Polyphaga, 24 Adephaga, 2 Archostemata and 2 Myxophaga. However, the motif 'ATGATAA' between *nad4* and *nad4l* was only found in the mitogenomes of Polyphaga insects, with 'ATGTTAA' in the other three suborders, Adephaga, Archostemata and Myxophaga (Fig S1).

Protein-coding genes

In mitogenomes of *G. depressa thoracica* and *G. depressa depressa*, a total of 3731 and 3678 amino acids were encoded by 11182 bp and 11026 bp nucleotides, respectively. All genes were encoded in the same direction and arrangement. The 13 PCGs ranged from 156 bp (*atp8*) to 1722 bp (*nad5*) for *G. depressa thoracica* and from 156 bp (*atp8*) to 1728 bp (*nad5*) for *G. depressa depressa* (Table 2). Except for *nad1* which used TTG as a start codon, all PCGs followed the typical ATN rules. The inferred start codons for the PCGs in *G. depressa thoracica* were ATT for *nad2*, *cox1*, *cox2* and *nad5*, ATA for *nad3* and *nad6*, ATG for *atp6*, *cox3*, *nad4*, *nad4l* and *cob*, and ATC for *atp8*. Then the inferred start codons in *G. depressa depressa* were ATT for *nad2*, *cox1*, *cox2*, *atp8*, *nad5* and *nad6*, ATG for *atp6*, *cox3*, *nad4*, *nad4l* and *cob*, and ATC for *nad3*. The stop codons for the two mitogenomes were TAA, TAG and incomplete termination codons TA or T. The incomplete stop codons could be completed as TAA through posttranscriptional polyadenylation (*Ojala et al*., *1981*; *Boore*, *2004*).

In order to measure the nucleotide compositional behaviour of PCGs, the A+T content, AT-skew and GC-skew of the *G. depressa thoracica*, *G. depressa depressa* and three other Chrysomelidae PCGs were counted (Table 3), and the obvious nucleotide bias were observed. The base composition of the 13 PCGs showed that each codon position had a high A+T percentage (Table 3), providing the high background mutational pressure toward AT nucleotides at the third codon position (*Kim et al*., *2014*). And PCGs on different strands also showed different skew statistics. PCGs on H-strand were more TA-skewed and CG-skewed, whereas the PCGs on the L-strand had a higher frequency of T and G, which was because of a slight skew towards A on the first and second codon sites of genes on the coding strand (*Pons et al*., *2010*).

In terms of amino acid usage, the PCGs on different strands also lead to an unbalanced percentage of amino acids (Fig 4). All of them had a relatively high percentage of Leu, and the least percentage of Cys. The relative

synonymous codon usage (RSCU) also showed that the most frequently used amino acids in two newly sequenced mitogenomes were Leu, Ile, Phe and Met, while TTA for Leu, ATT for Ile, TTT for Phe and ATA for Met were the most frequent codons, as in other insect mitogenomes (Fig 5). All of these codons were comprised of A or T nucleotides, indicating a biased usage of A and T nucleotides in the PCGs, and many other coleopteran insects also revealed a similar trend (*Kim et al*., *2009*; *Du et al*., *2016*).

Transfer RNA genes

The anticodons of all the tRNAs of the two newly sequenced mitogenomes were identified in other coleopteran species. Twenty-two complete tRNAs were found in *G. depressa thoracica*. All tRNAs were between 62 bp and 71 bp in length, and fourteen of them were located on H-strand, with others on L-strand. Five tRNAs, *trnH*, *trnI*, *trnK*, *trnP* and *trnSer(AGN)*, could not be detected by tRNAscan-SE and were determined by comparing with previous coleopteran mitogenomes. Except for *trnSer(AGN)*, all tRNAs could be folded into the typical clover-leaf structure (Fig S2), containing an amino acid acceptor arm (7 bp), TΨC arm (3-5 bp), dihydorouridine (DHU) arm (3-4 bp), anticodon arm (3-5 bp) and a variable extra arm. A total of 21 tRNAs were detected in the incomplete mitogenome of *G. depressa depressa*. The length of these genes ranged from 62 bp to 70 bp. The tRNAs coded on both strands were in consistence with those in the mitogenome of *G. depressa thoracica*. Four tRNAs, *trnH*, *trnK*, *trnP* and *trnSer(AGN)*, could not be found by software and were determined through comparison with published coleopteran mitogenomes. Except for *trnSer(AGN)* without the DHU arm, all tRNAs had the typical clover-leaf structures, containing an amino acid acceptor arm (7 bp), TΨC arm (3-5 bp), DHU arm (3-5 bp), anticodon arm (3-5 bp) and a variable extra arm (Figure S2). *trnSer(AGN)* formed a simple-loop structure instead of a DHU arm, which has been observed in many metazoan mitogenomes, including insects (*Wolstenholme*, *1992*).

Phylogenetic analyses

In order to further determine the relationships within Chrysomeloidea, phylogenetic analyses based on the concatenated 13 PCGs of 36 species for the nucleotide sequences (Fig 6) and the corresponding amino acid sequences (Fig 7) were conducted by using maximum likelihood (ML) and Bayesian Inference (BI), with *Tetraphalerus bruchi* (Coleoptera: Archostemata) and *Abax parallelepipedus* (Coleoptera: Adephaga) as outgroups. The final matrix consisted of 7289 nucleotides and 2340 amino acids. Models TVM+I+G and MtREV+I+G+F for the nucleotide dataset and the amino acid dataset of 13 protein-coding genes respectively were used when constructing phylogenetic trees by using PhyML and MrBayes.

The phylogeny recovered under ML showed more resolution with strong nodal supports than that under BI. Additionally, *G. depressa thoracica* and *G. depressa depressa* formed one clade, with a high nodal support values (Bootstrap Support, BS=100).

Regardless sequencing or database errors in some taxa, Chrysomeloidea were considered to be monophyletic. Basal relationships within Chrysomeloidea remain incompletedly resolved, especially the placement of some small families Zeugophoridae, Megalopodidae, Orsodacnidae, Vesperidae and Disteniidae. Relationships of

some lineages in previous analyses (*Gómez-Zurita et al*., *2007*) were well supported, including the monophyly of Eumolpinae + Cryptocephalinae + Cassidinae and the monophyly of Galerucinae+ Chrysomelinae (*Farrell*, *1998*; *Reid*, *2000*; *Farrell and Sequeira*, *2004*; *Bocak et al*., *2014*), the latter was only observed in the phylogenetic trees derived from nucleotide dataset. Specifically, Bruchinae was sister to the clade Donaciinae + Criocerinae, which was morphologically assumpted as an evolution for adapting to monocot feeding (*Duckett et al*., *2004*). Although Lepturinae and Necydalinae were sister group in phylogenetic trees, the internal relationships of Cerambycidae were unresolved, which may because of the low taxonomic coverage and poor representation of conserved loci in the analysis. Therefore, the study of Cerambycidae will play a significnant role in the evolution of Chrysomeloidea.

Conclusions

In this study, the complete and nearly complete mitogenomes of *G. depressa thoracica* and *G. depressa depressa* were determined, which were 16,109 bp and 14,277 bp in length. The genomic analyses showed that they had the identical gene composition and arrangement, and the conserved overlaps and intergenic spacer were observed in the mitogenomes of *G. depressa thoracica* and *G. depressa depressa*. It was firstly reported that the motifs, 'TAGTA' between *trnSer(UCN)* and *nad2*, and 'ATGATAA' between *atp8* and *atp6*, were more conserved than 'ATGATAA' between *nad4* and *nad4l* in the mitogenomes of Coleopterans, which was 'ATGTTAA' in the Adephaga, Myxophaga and Archostemata. Within Chrysomeloidea, the sister relationships of Eumolpinae + Cryptocephalinae + Cassidinae, and Galerucinae + Chrysomelinae were confirmed with high nodal supports, while the interrelationship of Cerambycidae remained unclear. ~~Therefore, clearing the internal relationship within Cerambycidae is a key factor for the phylogenetic analysis of Chrysomeloidea.~~

References

Abascal F, Zardoya R, Plsada D. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21:**2104–2105

Avise JC. 1989. A role for molecular genetics in the recognition and conservation of endangered species. *Trends in Ecology & Evolution* **4:**279–281

Bocak L, Barton C, Crampton-Platt A, Chesters D, Ahrens D, Vogler AP. 2014. Building the coleoptera tree–of–life for >8000 species: composition of public dna data and fit with linnaean classification. *Systematic Entomology* **39(1):**97–110

Boore JL. 2004. Complete mitochondrial genome sequence of the polychaete annelid *Platynereis dumerilii*. *BMC Genomics* **18:**1413–1416

Boore JL, Lavrov DV, Brown WM. 1998. Gene translocation links insects and crustaceans. *Nature* **392:**667–668 (1998).

Chen SH. 1974. Key to the species and subspecies of the Chrysomelid genus Gastrolina. *Acta Entomologica Sinica* **17:**195–197

Crowson RA. 1960. The phylogeny of coleoptern. *Annual Review of Entomology* **5:**111–134

Crozier RH, Crozier YC. 1993. The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization. *Genetics* **133:**97–177

De Rijk P, Wuyts J, De Wachter R. 2003. RnaViz 2: an improved representation of RNA secondary structure. *Bioinformatics* **19:**99–300

Du C, He SL, Song XH, Liao Q, Zhang XY, Yue BS. 2016. The complete mitochondrial genome of *Epicauta chinensis* (Coleoptera: Meloidae) and phylogenetic analysis among Coleopteran insects. *Gene* **578:**274

Duckett CN, Gillespie JJ, Kjer KM. 2004. Relationships among the subfamilies of Chrysomelidae inferred from small subunit ribosomal DNA and morphology, with special emphasis on the relationship among the flea beetles and the Galerucinae. In: *New Developments in the Biology of Chrysomelidae* 3–18. The Hague: SPB Academic Publishing

Fan JZ. 2015. The occurrence and controlling to walnut pests and diseases in Shanyang country and Luonan county, Shangluo city, Shaanxi province. M.Sc. Thesis, Northwest A&F University, Yangling, Shaanxi, China

Farrell BD, Sequeira AS. 2004. Evolutionary rates in the adaptive radiation of beetles on plants. *Evolution* **58:**1984–2001

Farrell BD. 1998. "Inordinate Fondness" explained: why are there so many beetles? *Science* **281:**555–559.

Foottit RG, Adler PH. 2009. Chapter 11. Biodiversity of Coleoptera. *Insect Biodiversity: Science and Society* 265–301

Ge SQ, Yang XK, Wang SY, Cui JZ, Li WZ. 2003. A revision of the taxonomic status of three subspecies of *Gastrolina depressa* Baly (Coleoptera: Chrysomelidae, Chrysomelinae). *Acta Entomologica Sinica* **46:**512–518

Gómez-Zurita J, Hunt T, Kopliku F, Vogler AP. 2007. Recalibrated tree of leaf beetles (Chrysomelidae) indicates independent diversification of angiosperms and their insect herbivores. *PLOS ONE* **2:**e360

Grimaldi D, Engel MS. 2005. Evolution of the Insects. Cambridge: Cambridge University Press

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59(3):**307–321

Hua YQ, Ding YR,Yan ZT, Si FL, Luo QC, Chen B. 2016. The complete mitochondrial genome of *Anopheles minimus* (Diptera: Culicidae) and the phylogenetics of known anopheles mitogenomes. *Insect Science* **23:**353–365

Huelsenbeck JP, Ronquist F. 2001. Mrbayes: bayesian inference of phylogenetic trees. *Bioinformatics* **17:**754–755

Hunt, T, Bergsten J, Levkanicova Z, Papadopoulou A, St. John O, Wild R, Hammond PM, Ahrens D, Balke M, Caterino MS, Gómez-Zurita J, Ribera I, Barraclough TG, Bocakova M, Bocak L, Vogler

AP. 2007. A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science* **318:**1913–1916

Jin SB, Kim I, Sohn HD, Jin BR. 2004. The mitochondrial genome of the firefly, *Pyrocoelia rufa*: complete DNA sequence, genome organization and phylogenetic analysis with other insects. *Molecular Biology and Evolution* **32:**978–985

Kim KG, Hong MY, Kim MJ, Im HH, Kim M Ⅱ, Bae CH, Seo SJ, Lee SH, Kim I. 2009. Complete mitochondrial genome sequence of the yellow-spotted long-horned beetle *Psacothea hilaris* (Coleoptera: Cerambycidae) and phylogenetic analysis among Coleopteran insects. *Molecules and Cells* **27:**429–441

Kim MJ, Wang AR, Park JS, Kim I. 2014. Complete mitochondrial genomes of five skippers (Lepidoptera: Hesperiidae) and phylogenetic reconstruction of Lepidoptera. *Gene* **549:**97–112

Li XJ, Ou J, Wei ZM, Li YX, Tian YF. 2016. The mitogenomes of three beetles (coleoptera: polyphaga: cucujiformia): new gene rearrangement and phylogeny. *Biochemical Systematics and Ecology* **69:**101-107

Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25:**955–964

Ma C, Yang PC, Jiang F, Chapuis MP, Shali Y, Sword GA, Kang L. 2012. Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. *Molecular Ecology* **21:**4344–4358

Marvaldi AE, Duckett CN, Kjer KM, Gillespie JJ. 2009. Structural alignment of 18s and 28s rdna sequences provides insights into phylogeny of phytophaga (Coleoptera: Curculionoidea and Chrysomeloidea). *Zoologica Scripta* 38(1):63–77.

Nylander JAA. 2004. MrModeltest Version 2. Evolutionary Biology Centre. Uppsala: Uppsala University

ØDegaard F. 2000. How many species of arthropods? Erwin's estimate revised. *Biological Journal of the Linnean Society* **71:**583–597

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

Perna NT, Kocher TD. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *Journal of Molecular Evolution* **41:**353–358

Pons J, Ribera I, Bertranpetit J, Balke M. 2010. Nucleotide substitution rates for the full set of mitochondrial protein-coding genes in Coleoptera. *Molecular Biology and Evolution* **56:**796–807

Rambaut A. 2009. FigTree v1.3.1. Available at http://tree.bio.ed.ac.uk/software/ (accessed on 31 September 2017)

Reid CAM. 2000. Spilopyrinae Chapuis: a new subfamily in the Chrysomelidae and its systematic placement (Coleoptera). *Invertebrate Taxonomy* **14:**837–862

Stewart JB, Beckenbach AT. 2005. Insect mitochondrial genomics: the complete mitochondrial genome sequence of the meadow spittlebug *Philaenus spumarius* (Hemiptera: Auchenorrhyncha: Cercopoidae). *Genome* **48:**46–54

315 **Taanman JW. 1999.** The human mitochondrial genome: structure, transcription, translation and replication.
316 *Biochim Biophys Acta* **1410:**103–123
317 **Talavera G, Castresana J. 2007.** Improvement of phylogenies after removing divergent and ambiguously
318 aligned blocks from protein sequence alignments. *Systematic Biology* **56:**564–577
319 **Tamura K, Aotsuka T. 1988.** Rapid isolation method of animal mitochondrial DNA by the alkaline lysis
320 procedure. *Biochemical Genetics* **26:**815–819
321 **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011.** MEGA5: molecular evolutionary
322 genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.
323 *Molecular Biology and Evolution* **28:**2731–2739
324 **Timmermans MJ, Lees DC, Simonsen TJ. 2014.** Towards a mitogenomic phylogeny of Lepidoptera.
325 *Molecular Biology and Evolution* **7:**169–178
326 **Wang QQ, Zhang ZQ, Tang GH. 2016.** The mitochondrial genome of *Atrijuglans hetaohei* Yang
327 (Lepidoptera: Gelechioidea) and related phylogenetic analyses. *Gene* **581:**66–74
328 **Wolstenholme DR. 1992.** Animal mitochondrial DNA: structure and evolution. *Int. Rev Cytol* **141:**173–216
329 **Zhang ZQ, Wang XJ, Li RZ, Guo RJ, Zhang W, Song W, Hao CF, Wang HP, Li ML. 2015.** The
330 mitochondrial genome of *Dastarcus helophoroides* (Coleoptera: Bothrideridae) and related phylogenetic
331 analyses. *Gene* **560:** 15–24

Figure 1

Gene maps of the complete mitochondrial genome for (a)*G*. *depressa thoracica* and the incomplete mitochondrial genome for (b) *G*. *depressadepressa*.

The PCGs and rRNAs are the standard abbreviations. Each tRNA is denoted as a one-letter symbol according to the IUPAC-IUB single-letter amino acid codes. Arrows indicate coding direction.

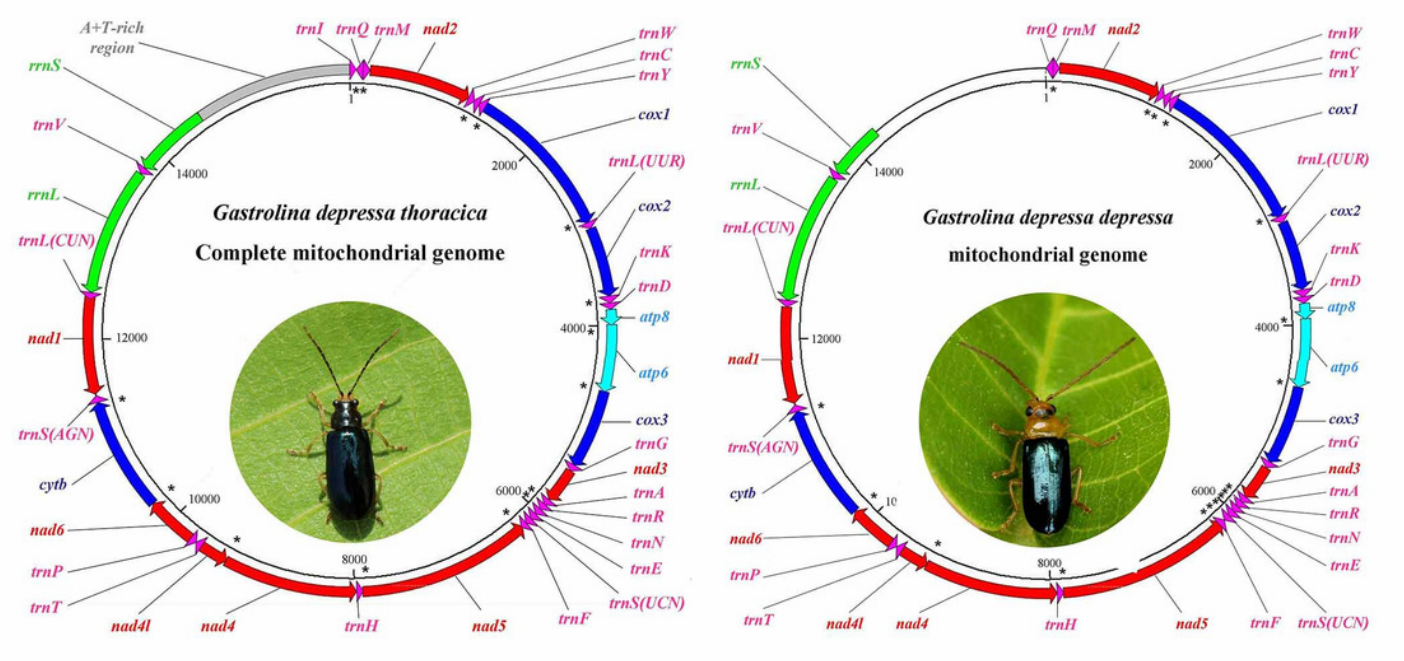
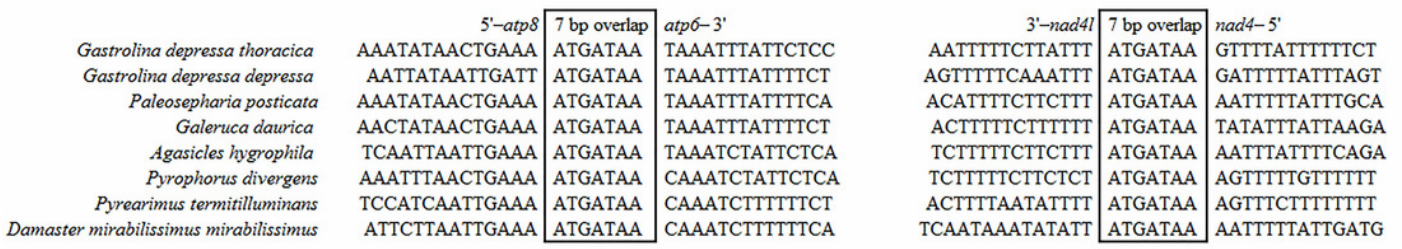


Figure 2

Sequence alignments of *atp8*/*atp6* and *nad4*/*nad 4l* of coleopteran insects.
The boxed nucleotides indicate the 7 bp conservedoverlaps (ATGATAA).



	5'–*atp8*	7 bp overlap	*atp6*– 3'	3'–*nad4l*	7 bp overlap	*nad4*– 5'
Gastrolina depressa thoracica	AAATATAACTGAAA	ATGATAA	TAAATTTATTCTCC	AATTTTTCTTATTT	ATGATAA	GTTTTATTTTTTCT
Gastrolina depressa depressa	AATTATAATTGATT	ATGATAA	TAAATTTATTTTCT	AGTTTTTCAAATTT	ATGATAA	GATTTTTATTTAGT
Paleosepharia posticata	AAATATAACTGAAA	ATGATAA	TAAATTTATTTTCA	ACATTTTCTTCTTT	ATGATAA	AATTTTTATTTGCA
Galeruca daurica	AACTATAACTGAAA	ATGATAA	TAAATTTATTTTCT	ACTTTTTCTTTTTT	ATGATAA	TATATTTATTAAGA
Agasicles hygrophila	TCAATTAATTGAAA	ATGATAA	TAAATCTATTCTCA	TCTTTTTCTTCTTT	ATGATAA	AATTTATTTTCAGA
Pyrophorus divergens	AAATTTAACTGAAA	ATGATAA	CAAATCTATTCTCA	TCTTTTTCTTCTCT	ATGATAA	AGTTTTTGTTTTTT
Pyrearimus termitilluminans	TCCATCAATTGAAA	ATGATAA	CAAATCTTTTTTCT	ACTTTTAATATTTT	ATGATAA	AGTTTCTTTTTTTT
Damaster mirabilissimus mirabilissimus	ATTCTTAATTGAAA	ATGATAA	CAAATCTTTTTTCA	TCAATAAATATATT	ATGATAA	AATTTTTATTGATG

Figure 3

Sequence alignment of the space region between *nad1* and *trnSer(UCN)* of coleopteran species. The boxed nucleotides indicate the conserved motif(TAGTA).

Figure 4

Percentages of amino acid usage in mitochondrialproteins of five species. Eachamino acid is represented by the three-letter abbreviation. Note that leucineand serine are each coded by two different genetic codons, and listedseparately.

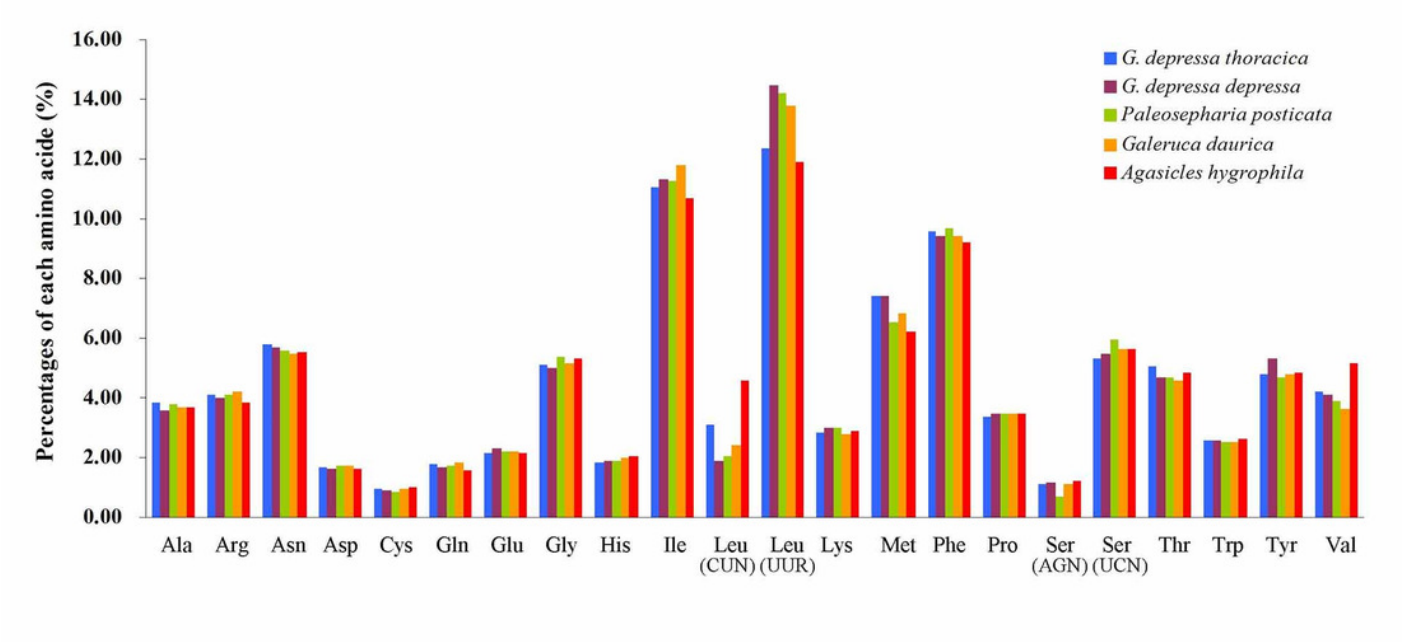


Figure 5

The mitogenome relative synonymous codon usage (RSCU)across five coleopteran insects. Codon families are provided on the x-axis.

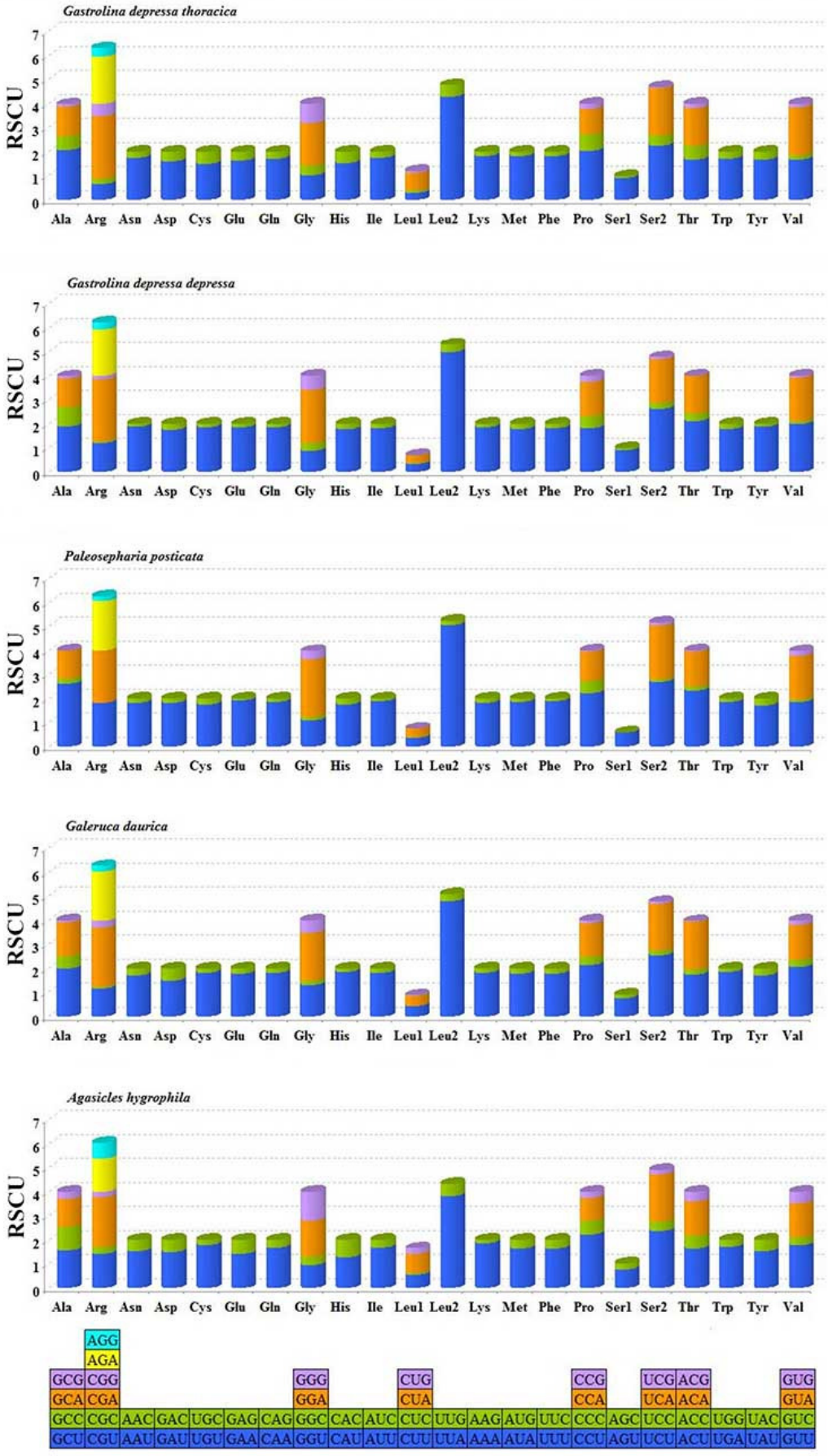

Gastrolina depressa thoracica
Gastrolina depressa depressa
Paleosepharia posticata
Galeruca daurica
Agasicles hygrophila
RSCU
0 1 2 3 4 5 6 7
Ala Arg Asn Asp Cys Glu Gln Gly His Ile Leu1 Leu2 Lys Met Phe Pro Ser1 Ser2 Thr Trp Tyr Val
AGG AGA
GCG CGG GGG CUG CCG UCG ACG GUG
GCA CGA GGA CUA CCA UCA ACA GUA
GCC CGC AAC GAC UGC GAG CAG GGC CAC AUC CUC UUG AAG AUG UUC CCC AGC UCC ACC UGG UAC GUC
GCU CGU AAU GAU UGU GAA CAA GGU CAU AUU CUU UUA AAA AUA UUU CCU AGU UCU ACU UGA UAU GUU

Figure 6

Phylogenetic tree based on the nucleotide sequence for13 protein-coding genes from the mitochondrial genomes of 36 species. Probability values (PP>60) and bootstrap values(BBP>0.7) of the branches were indicated.

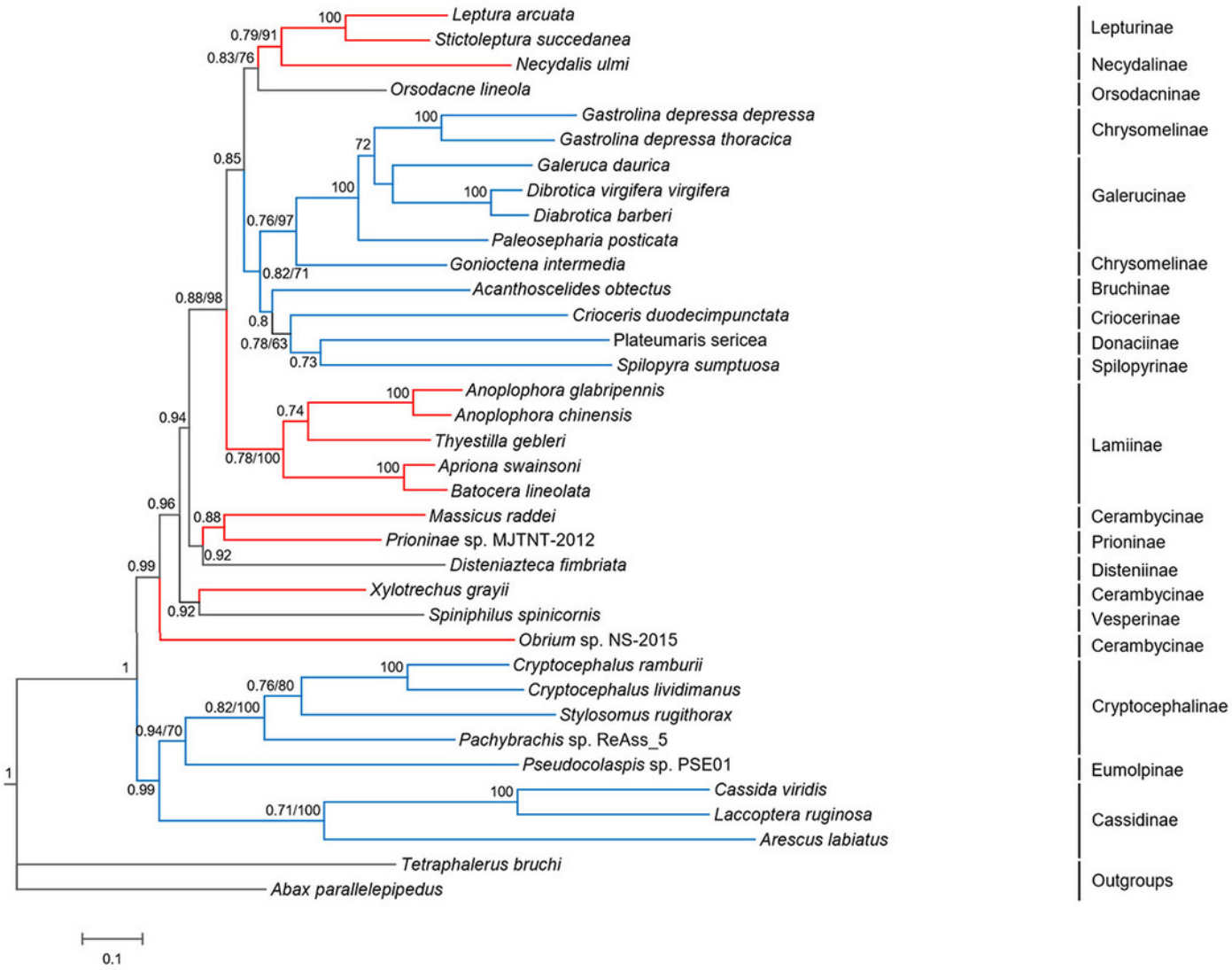


Figure 7

Phylogenetic tree based on the amino acide sequencefor 13 protein-coding genes from the mitochondrial genomes of 36 species. Probability values (PP>75) and bootstrap values(BBP>0.6) of the branches were indicated.

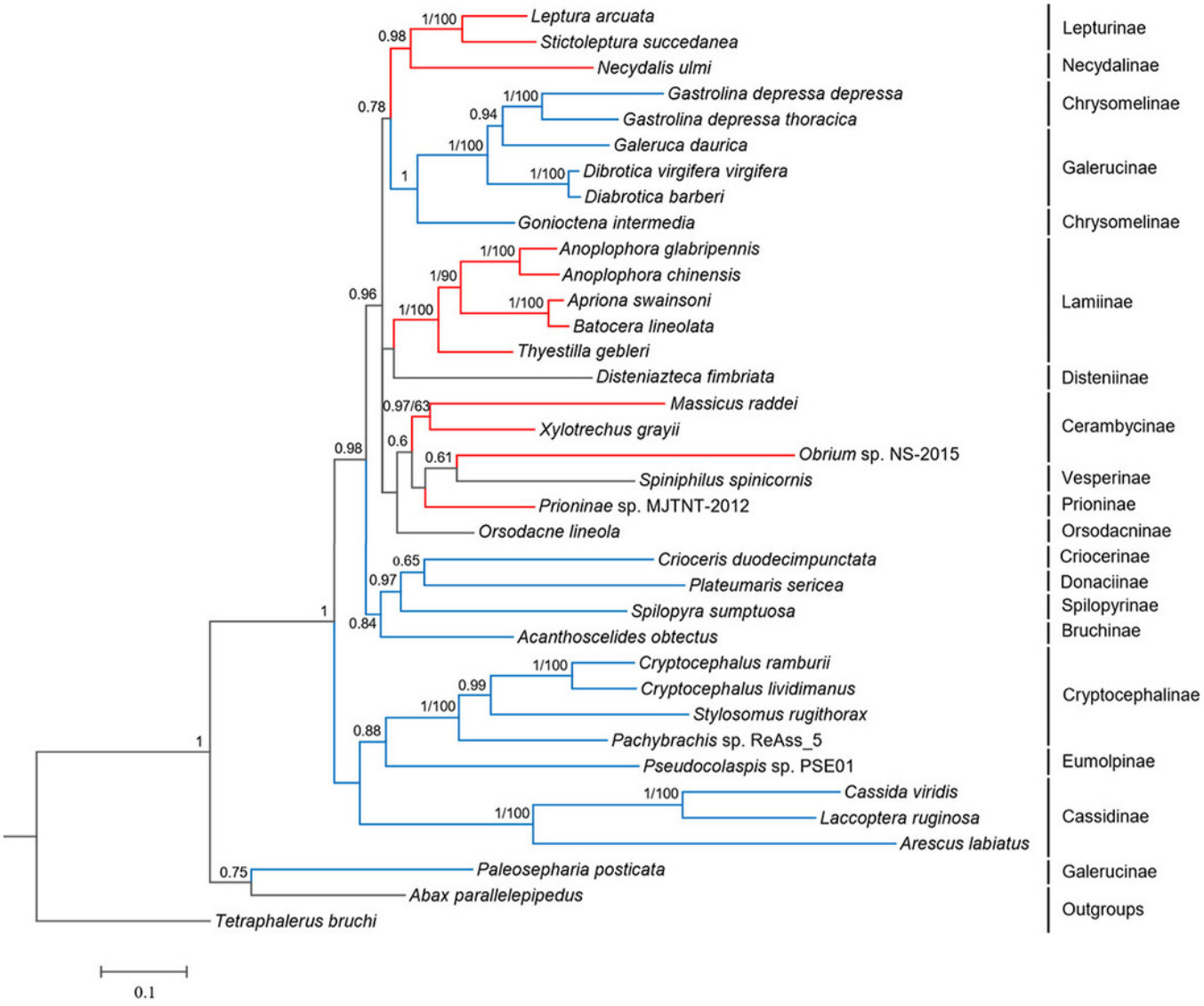


Table 1(on next page)

Primers used in this study to sequence the mitogenomesof *G. depressa thoracica* and *G. depressa depressa*.

Primer	Forward (5'→3')	Reverse (5'→3')	Tm (℃)	Length (kb)
G. depressa thoracica				
01	GCCTGAAATGAAAGGATAATTTTGATA	GCTCGGGTATCTACATCTATTC	55	2.2
02	GTTAATATAAACTCTTAACCTTCAA	CCGCAAATCTCAGAGCATTG	49	2.3
03	ACAATTGGACATCAATGATACTG	ATGACCAGCAATTATATTAG	51	1.1
04	TTAGCACATTTAGTTCCACAAGG	TATAATTAGAGCATAATTTTGAAG	50	1.9
05	TTTAATTGAAACCAAATTAGAGG	TTTTTGTCGTAATGGTC	50	4.1
06	CGCTCAGGCTGATAGCCCCA	AATCGTACTCCGTTTGATTTTGC	53	2.9
07	CGAGGTAATGTACCCCGAACCCA	GTGCCAGCAGTTGCGGTTATAC	58	2.8
08	ACCTTTATAATTGAGGTATGAAC	ATAATAGGGTATCTAATCCTAG	51	2.0
G. depressa depressa				
01	GCCTGATAAAAAGGATTATCTTGATA	TAAACTTCTGGGTGTCCAAAAAATCA	52	2.0
02	AATTGGGGGATTTGGAAATTG	CCACAAATTTCTGAACATTG	49	2.0
03	ACAATTGGACATCAATGATATTG	AGGGGCTTCTTTTTTCATAA	47	2.3
04	GCAGCTGCTTGATATTGACA	TTAGGATGGGATGGTTTGGG	54	2.2
05	TTTAATTGAAACCAAATAGAGG	GTTTGTGAGGGGGTTTTAGG	55	3.4
06	CCAGAAGAACAAATACCATG	TATCAATAGCAAATCCCCCCCA	53	2.3
07	TTCAGCAATATGAAATTTTGGATC	TTACCTTAGGGATAACAGCGTAA	53	2.4
08	CCGGTTTAAACTCAGATCATGTA	GTGCCAGCAGTTGCGGTTATAC	57	1.8

1

Table 2(on next page)

Annotations for the mitogenomes of *G.depressa thoracica* and *G. depressa depressa*.

IGS denotes the length of the intergenic spacer region, for which negative numbers indicate nucleotide overlapping between adjacent genes. H and L denote heavy and light strands, respectively. N.C. indicates non-coding sequence. Gdt and Gdd represent *G. depressa thoracica* and *G. depressa depressa* respectively.

Gene	Position		IGS/bp		Initiation/Stop Codon		Anticodon	Coding Strand
	Gdt	Gdd	Gdt	Gdd	Gdt	Gdd		
trnI	1-65	-	-	-			GAT	H
trnQ	65-133	6-74	-1	-			TTG	L
trnM	133-202	74-140	-1	-1			CAT	H
nad2	203-1209	141-1154	0	0	ATT/TA	ATT/TAA		H
trnW	1210-1273	1153-1216	0	-2			TCA	H
trnC	1266-1327	1209-1272	-8	-8			GCA	L
trnY	1328-1391	1273-1339	0	0			GTA	L
cox1	1384-2931	1332-2879	-8	-8	ATT/TAA	ATT/TAA		H
trnL(UUR)	2927-2991	2875-2939	-5	-5			TAA	H
cox2	2992-3679	2940-3627	0	0	ATT/T	ATT/T		H
trnK	3680-3750	3628-3697	0	0			TTT	H
trnD	3750-3812	3698-3761	-1	0			GTC	H
atp8	3813-3968	3762-3917	0	0	ATC/TAA	ATT/TAA		H
atp6	3962-4636	3911-4585	-7	-7	ATG/TAA	ATG/TAA		H
cox3	4636-5419	4585-5371	-1	-1	ATG/T	ATG/T		H
trnG	5420-5483	5372-5435	0	0			TCC	H
nad3	5484-5835	5436-5787	0	0	ATA/T	ATC/T		H
trnA	5836-5902	5788-5852	0	0			TGC	H
trnR	5902-5964	5852-5913	-1	-1			TCG	H
trnN	5964-6027	5911-5974	-1	-3			GTT	H
trnS(UCN)	6028-6094	5971-6034	0	-4			TCT	H
trnE	6095-6157	6033-6095	0	-2			TTC	H
trnF	6158-6220	6094-6159	0	-2			GAA	L
nad5	6204-7925	6140-7867	-17	-20	ATT/TAA	ATT/TAA		L
trnH	7923-7984	7865-7928	-2	-3			GTG	L
nad4	7985-9317	7929-9261	0	0	ATG/T	ATG/T		L
nad4l	9311-9595	9255-9539	-7	-7	ATG/TAA	ATG/TAA		L
trnT	9599-9660	9543-9606	3	3			TGT	H
trnP	9661-9723	9607-9670	0	0			TGG	L
nad6	9726-10226	9673-10143	2	2	ATA/TAA	ATT/TAA		H
cob	10226-11365	10143-11282	-1	-1	ATG/TAG	ATG/TAG		H
trnS(AGN)	11364-11430	11281-11345	-2	-2			TGA	H
nad1	11448-12398	11363-12313	17	17	TTG/TAG	TTG/TAG		L
trnL(CUN)	12400-12464	12315-12379	1	1			TAG	L
rrnL	12465-13738	12380-13661	0	0				L
trnV	13739-13806	13662-13730	0	0			TAC	L
rrnS	13807-14551	13731-14281	0	0				L
A+T-rich region	14552-16109	-	0	-				N.C.

1

Table 3(on next page)

Nucleotide compositions in the mitogenomes of *G. depressadepressa*, *G. thoracica depressa*, *Paleosephariapostivata*, *Galeruca daurica* and *Agasicles hygrophila*.

	G. depressa depressa			*G. thoracica depressa*			*Paleosepharia postivata*			*Galeruca daurica*			*Agasicles hygrophila*		
	A+T	**AT skew**	**GC skew**	**A+T**	**AT skew**	**GC skew**	**A+T**	**AT skew**	**GC skew**	**A+T**	**AT skew**	**GC skew**	**A+T**	**AT skew**	**GC skew**
Protein-coding genes	77.9	-0.143	0.018	75.2	-0.133	-0.005	78.0	-0.150	0.006	77.0	-0.141	-0.006	72.4	-0.152	-0.018
First codon position	73.3	-0.040	0.225	71.3	-0.010	0.194	72.3	-0.050	0.227	72.4	-0.025	0.188	68.9	-0.021	0.153
Second codon position	69.7	-0.384	-0.117	69.2	-0.387	-0.125	68.6	-0.391	-0.141	68.7	-0.394	-0.112	68.4	-0.395	-0.117
Third codon position	90.6	-0.040	-0.135	85.2	-0.03	-0.153	93.2	-0.049	-0.211	90.1	-0.042	-0.210	80.0	-0.058	-0.128
Protein-coding genes-H strand	76.4	-0.087	-0.106	73.3	-0.050	-0.162	76.6	-0.112	-0.110	75.1	-0.089	-0.134	70.7	-0.084	-0.145
First codon position	70.8	0.057	0.132	68.4	0.112	0.079	69.7	0.032	0.140	69.5	0.064	0.098	66.7	0.079	0.071
Second codon position	68.0	-0.361	-0.221	66.8	-0.364	-0.209	67.2	-0.361	-0.217	67.2	-0.368	-0.202	67.3	-0.70	-0.203
Third codon position	90.2	0.005	-0.470	84.8	0.066	-0.562	92.8	-0.041	-0.681	88.5	0.001	-0.557	78.1	0.025	-0.388
Protein-coding genes-L strand	80.1	-0.227	0.246	78.8	-0.246	0.286	80.4	-0.206	0.229	80.2	-0.219	0.255	75.3	-0.256	0.224
First codon position	77.0	-0.167	0.410	76.1	-0.149	0.417	76.5	-0.170	0.405	77.0	-0.153	0.380	72.5	-0.168	0.311
Second codon position	72.1	-0.431	0.053	71.6	-0.433	0.044	70.8	-0.437	-0.005	71.1	-0.443	0.051	70.1	-0.434	0.035
Third codon position	91.2	-0.116	0.429	88.7	-0.177	0.617	93.8	-0.062	0.659	92.6	-0.110	0.657	83.1	-0.183	0.414

1