

The complete mitochondrial genome of *Dysgonia stuposa* (Lepidoptera: Erebidae) and phylogenetic relationships within Noctuoidea

Yuxuan Sun^{Equal first author, 1}, Yeshu Zhu^{Equal first author, 1}, Chen Chen¹, Qunshan Zhu¹, Qianqian Zhu¹, Yanyue Zhou¹, Xiaojun Zhou¹, Peijun Zhu¹, Jun Li^{Corresp., 1}, HaiJun Zhang^{Corresp. 1}

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

Corresponding Authors: Jun Li, HaiJun Zhang

Email address: healthlicn@chnu.edu.cn, haijunzhang@163.com

To determine the *Dysgonia stuposa* mitochondrial genome (mitogenome) structure and to clarify its phylogenetic position (formerly in Lepidoptera: Erebidae), the entire mitogenome of *D. stuposa* was sequenced and annotated. The *D. stuposa* mitogenome was 15,721 bp in size, containing 37 genes (including protein-coding genes (PCGs), transfer RNA genes (tRNAs), ribosomal RNA genes (rRNAs)) usually found in lepidopteran mitogenomes. The newly sequenced mitogenome contained some common features such as the A+T biased nucleotide composition and the *cox1* gene initiated by non-canonical CGA codon, which have also been reported in other Erebidae species. In addition, like other insect mitogenomes, the *D. stuposa* mitogenome had a conserved sequence 'ATACTAA' in intergenic spacer, a motif 'ATAGA' followed by a 20 bp poly-T stretch in A+T rich region. The phylogenetic analyses revealed that *D. stuposa* reside in the Erebidae family, and suggest a novel lineages for Erebidae which contains different subfamilies: Arctiinae, Aganainae, Catocalinae, Lymantriinae and Hypeninae.

**The complete mitochondrial genome of *Dysgonia stuposa* (Lepidoptera: Erebidae) and
phylogenetic relationships within Noctuoidea**

Yuxuan Sun*, Yeshu Zhu*, Chen, Chen, Qunshan Zhu, Qianqian Zhu, Yanyue Zhou, Xiaojun
Zhou, Peijun Zhu, Jun Li#, HaiJun Zhang#

(College of Life Sciences, Huaibei Normal University, Huaibei, Anhui, China)

*These two authors contributed equally to this work.

#Corresponding authors.

Address: College of Life Sciences, Huaibei Normal University, 100 Dongshan Road, Huaibei,
Anhui, 235000 China.

Email: healthlicn@chnu.edu.cn (Jun Li), haijunzhang@163.com(HaiJun Zhang)

Abstract

To determine the *Dysgonia stuposa* mitochondrial genome (mitogenome) structure and to clarify its phylogenetic position (formerly in Lepidoptera: Erebidae), the entire mitogenome of *D. stuposa* was sequenced and annotated. The *D. stuposa* mitogenome was 15,721 bp in size, containing 37 genes (including protein-coding genes (PCGs), transfer RNA genes (tRNAs), ribosomal RNA genes (rRNAs)) usually found in lepidopteran mitogenomes. The newly sequenced mitogenome contained some common features such as the A+T biased nucleotide composition and the *cox1* gene initiated by non-canonical CGA codon, which have also been reported in other Erebidae species. In addition, like other insect mitogenomes, the *D. stuposa* mitogenome had a conserved sequence ‘ATACTAA’ in intergenic spacer, a motif ‘ATAGA’ followed by a 20 bp poly-T stretch in A+T rich region. The phylogenetic analyses revealed that *D. stuposa* reside in the Erebidae family, and suggest a novel lineages for Erebidae which contains different subfamilies: Arctiinae, Aganainae, Catocalinae, Lymantriinae and Hypeninae.

Keywords: Phylogenetic relationship; *D. stuposa*; Mitochondrial genome; Noctuoidea

Introduction

D. stuposa (Lepidoptera: Erebidae) is an important pest species. It is broadly distributed Southern and Eastern parts of Asia. Its larvae usually consume the leaves of *Phyllanthus* plants. Despite the importance of this species, knowledge is scarce on its biology and evolutionary status at molecular level. To protect from harsh environment, *D. stuposa* transform to pupal stage during winter season. Interestingly, many insect species have approximately similar morphology at their pupal stage. Thus, the identification of *D. stuposa* at pupal stage based on its

morphological characteristics is highly difficult for taxonomists and population ecologists. The molecular technologies such as DNA barcodes and PCR-RFLP methods are more credible and reliable for morphology and taxonomy (Arimoto & Iwaizum, 2014; Raupach et al., 2010). Therefore, the study to determine the sequence of *D. stuposa* mitogenome will further aid for identification, classification, genecology analyses and evolutionary study of this species.

The insect mitogenome is a closed, circular, double-stranded DNA molecule, and the length is usually 14-19 kb (Boore, 1999). Unlike nuclear genomes, mitogenomes is smaller and more easily to acquired and sequenced. It usually contains several common characteristics, such as the stability of gene composition, conservativeness of gene arrangement and maternal inheritance, that can be widely used in molecular identification, phyletic evolution, systematics and biogeography study (Wolstenholme, 1992; Wilson et al., 2010). Since the huge diversity of insects, the mitogenome analysis are highly helpful and have been widely used in the analyses of genomic evolution and phylogenetic relationships (Lu et al., 2013; Cameron, 2014).

Noctuoidea is one of the biggest superfamilies of Lepidoptera with over 42,400 described species (Nieukerken et al., 2011). Since the taxonomic history of Noctuoidea is complicated, researchers rely on molecular features instead of morphological characteristics, which are usually misleading during the classification of organisms. However, classification based on the morphological characteristics was the only option before the molecular era. At that time, it was considered that the Noctuoidea has unique characteristics such as a metathoracic tympanal organ (Miller, 1999). Later on, researchers identified 10 other morphological characteristics to subdivide it at the family level and to resolve the monophyly of the families like Arctiidae,

Nolidae, Lymantriidae and Pantheidae (Speidel & Fanger, 2010; Kitching et al., 2010). Furthermore, based on the adult and larval morphological characteristics, Lafontaine and Fibiger proposed a monophyly of the ‘trifid’ (3-branching forewing cubital vein) and ‘quadrifids’ noctuoids (4-branching forewing cubital vein) (Fibiger & Lafontaine, 2005; Lafontaine & Fibiger, 2006). The morphological based taxonomy failed to resolve classification problems especially at the family and sub-family level. Therefore, the molecular studies based on the gene sequencing can play a crucial role depth in the systematic analysis of species. Zahiri et al. reconstructed the molecular phylogenetics of Noctuoidea using the PCGs of 152 species, and proposed a new and acceptable perspective on the phylogenetics of Noctuoidea (Zahiri et al., 2011). They broke the traditional group of quadrifid noctuids, reconstructed Erebidae and Nolidae, and classified Noctuoidea into six families: Oenosandridae, Notodontidae, Erebidae, Nolidae, Euteliidae and Noctuidae. However, the morphological studies are not consistent with the molecular studies in some aspects, and the morphological support in Erebidae classification is also not enough (Minet et al., 2012). Moreover, the phylogenetic relationships between Noctuidae and Erebidae remain controversial, and further studies based on molecular levels are require to resolve noctuoid relationships (Lafontaine & Fibiger, 2006).

The complete mitogenome differs from the nuclear genome, and has been increasingly used to solve the questions with poorly support in multi-gene phylogenetic analyses such as the position of Nymphalidae under Papilionoidea (Yang et al., 2009; Cameron, 2014). Since, many species of the genus *Dysgonia* (formerly belonging to family Erebidae) have been moved to other families according to classification given by Holloway and Miller (Holloway & Miller,

2003), further status of many unordered species remains uncertain in phylogenetic relationships. In our study, we sequenced and characterized the complete mitogenome of *D. stuposa* and reconstructed the phylogenetic relationships to confirm its phylogenetic status in the Noctuoidea. Newly sequenced mitogenome revealed a novel and stable phylogenetic relationships within Erebidae, and will provide a foundation to further study the Noctuidae and Erebidae mitogenome composition features, biogeography, and phylogenetics.

Materials and Methods

Specimen collection and Genomic DNA extraction

The *D. stuposa* moths were collected from Xiangshan mountains (N33° 59' 47" , E116° 47' 42"), Huaibei, Anhui, China. Based on the morphological characteristics, the collected specimens were identified as *D. stuposa* using *Fauna Sinica* (Chen, 2003). The identification was confirmed by analyzing the DNA barcoding sequences of this species by BOLD systems. The genomic DNA of *D. stuposa* was isolated using the Extraction Kit according to the manufacturer's instructions (Sangon, China). The extracted DNA was used as template for PCR amplification.

PCR amplification and fragment sequencing

To amplify the *D. stuposa* mitogenome, long fragment PCR and universal primers were synthesized, all the primers were tested and utilized for insect mitogenome amplification (Table 1) (Sun et al., 2016). All PCR amplifications were executed using DNA Polymerase (PrimeSTAR® GXL, Takara, China). The PCR procedure was performed according to the previous protocol (Sun et al., 2016). The PCR products were determined by agarose gel with

106 TAE buffer, then sequenced by General Biosystems (General, China).

107 **Sequence assembly and annotation**

108 The complete mitogenome was assembled by the sequence results using the DNAMAN
109 software. Sequence annotation was performed by BLAST tool in NCBI website and compared
110 with the homologous sequences from other lepidopterans. To determine initiation and
111 termination codons of PCGs, sequences were aligned with other published Noctuoidea sequences
112 using ClustalX 2.0. AT skew and GC skew values were used to represent base composition of
113 nucleotide sequences and calculated using the methods given by Junqueira (Junqueira et al.,
114 2004). MEGA 5.0 software was used to analyze the relative synonymous codon usage (RSCU)
115 (Tamura et al., 2011).

116 The tRNA genes were determined by tRNAscan Search Server and the secondary structures
117 were detected using sequence alignment with appropriate anticodons which can fold into the
118 proper clover-leaf structure (Lowe & Eddy, 1997). Tandem Repeats Finder were used to analyze
119 the tandem repeats (Benson, 1999).

120 **Phylogenetic analysis**

121 To rebuild the phylogenetic relationships among Noctuoidea at superfamily level, a total of
122 29 species concatenated amino acid sequence sets of all PCGs were used to perform
123 phylogenetic analysis. All PCGs sequences were downloaded from GenBank. Additionally, the
124 mitogenome sequences of *Bombyx mori* (AY048187) and *Antheraea pernyi* (AY242996) (Liu et
125 al., 2008) were used as outgroups. The PCGs sequences were aligned using ClustalX (Thompson
126 et al., 1997). Then, two different analytical approaches including ML (Maximum Likelihood)

and BI (Bayesian Inference) were used to construct the phylogenetic relationships. The ML analysis was performed using MEGA 5.0, and the clade support was investigated with 1000 bootstrap replicates. The BI analysis used four chains MCMC with MrBayes 3.2 version program, and running for 2,000,000 generations. The phylogenetic trees were analyzed by FigTree software.

Results

Genome organization and composition

The *D. stuposa* mitogenome is a circular DNA molecule with a length of 15,721 bp (Figure 1). The size of newly sequenced mitogenome is comparable to the mitogenome described for other Noctuoidea species, which range from 15,801 bp in *Gynaephora minora* to 15,377 bp in *Agrotis ipsilon* (Table 2). The mitogenome annotation showed that it contained 37 genes and a A+T rich region, which usually present in insect mitogenome (Table 3). In addition, highly A and T biased nucleotide composition is a characteristic feature of insect mitogenomes including *D. stuposa*. Our results revealed that the nucleotide composition is highly biased towards using A and T (A=39.98%, T=40.38%, G=7.5%, C=12.14%) (Table 2). The A+T content of *D. stuposa* is 80.36%, which is comparable to previously sequenced lepidopteran species (77.84% in *Ochrogaster lunifer* - 81.49% in *G. minora*).

The AT and GC skewness were computed to determine the base composition (Junqueira et al., 2004). The *D. stuposa* mitogenome contained negative AT skewness (-0.005) and GC skewness (-0.236), indicating the presence of more Ts than As, and Cs than Gs, respectively (Table 2). These results are consistent with previous studies, as negative AT skewness has been

reported in several other insect species such as *A. plana lacteata* (-0.002), *Risoba prominens* (-0.007) and *A. ipsilon* (-0.006).

Protein-coding genes and codon usage

A total of 13 PCGs were identified from the *D. stuposa* mitogenome, with a length of 11,334 bp, and account 72% of the whole mitogenome. In insects, most of the PCGs locate on the J strand (majority strand), while some of them reside on N strand (minority strand) (Simon et al., 1994). In *D. stuposa*, nine of the thirteen PCGs (*nad2*, *cox1*, *cox2*, *atp8*, *atp6*, *cox3*, *nad3*, *nad6* and *cob*) were resided on the J-strand, while remaining PCGs (*nad5*, *nad4*, *nad4L* and *nad1*) were located on the N-strand. The ATN codon always initiates the transcription of insect PCGs, but the exception is *cox1* that utilizes CGA codon (Table 3).

To estimate the codon usage among insect species, the entire nucleotide sequences of seven Noctuoidea (belonging to four families: Erebiidae, Noctuidae, Nolidae and Notodontidae) were obtained from GenBANK (Figure 2). The results showed that Phe, Asn, Leu, Met, Tyr and Ile were the most commonly used amino acids, while Cys was found to be rarely utilized amino acid. The codon usage of amino acids was almost similar in the Noctuoidea. Furthermore, we used CDspT (codons per thousand) to illustrate the codons distribution in different species (Figure 3). The results exhibited approximately similar trends in the Noctuoidea superfamily, and the maximum CDspT value was observed for Asn and Ile.

The value of Relative Synonymous Codon Usage (RSCU) for Noctuoidea were given in Figure 4. We observed that the codons usage of same amino acid varied in different species. All the codons were appeared in the *D. stuposa* mitogenome, except for ACG and CCG codons. In

addition, some species are lack of synonymous codon with high G or C in the third codon position, such as GCG, CGC, GGC and CCG were abandoned in *A. ipsilon*.

Ribosomal RNA and transfer RNA genes

The *D. stuposa* mitogenome contains both large ribosomal gene (*rrnL*) and a small ribosomal gene (*rrnS*) with a length of 1,308 bp and 782 bp, respectively (Figure 1, Table 3). Both genes were encoded by the J-strand, and located at the same position as found in previously sequenced mitogenomes (Yang et al., 2009).

There were 22 tRNA genes in the *D. stuposa* mitogenome, ranging in size from 57 bp (*trnA*) to 71 bp (*trnK*) (Table 3). Almost all tRNAs had standard anticodons and comprised the typical clover-leaf secondary structure. The exception was *trnS1* which lacks the dihydrouridine (DHU) arm (Figure 5). Moreover, we observed base pairs mismatches in the secondary structure of tRNAs, such as A-A mismatch (*trnM*), U-U mismatch (*trnL2*, *trnS2*) and U-C mismatch (*trnA*).

Overlapping , intergenic spacer and A+T rich regions

The overlapping of genes extended the genetic information in limited genomes, and commonly found in metazoan mitogenomes (Wolstenholme, 1992). We identified nine overlapping regions with a total length of 144 bp in the mitogenome of *D. stuposa* (Table 3). The longest overlapping region of 65 bp was found between *trnH* and *nad4*. In addition, we observed a seven bp overlapping region located between *atp6* and *atp8* that has also been reported in other insects (Zhu et al., 2013).

The *D. stuposa* mitogenome had 21 spacer regions, ranged in size from 1 bp to 105 bp. The 105 bp spacer contained highly A and T content, and located between *trnA* and *trnR*. We also

observed a 22 bp spacer contained an ‘ATACTAA’ motif and located between *nad1* and *trnS2*, which was common in the species analyzed in the present study (Figure 6A).

Metazoan mitogenomes usually have a very large non-coding region, named as A+T rich region (Lv et al., 2018). It contains high A+T content and can regulate and initiate DNA transcription and replication in a mitogenome (Clayton, 1991; Fernández-Silva et al., 2003). The A+T rich region of *D. stuposa* mitogenome located between the *rrnS* and *trnM* genes with 406 bp in size (Table 3), and had highest A+T content (92.37%), and highly negative GC skewness (-0.355) (Table 2). Further, it usually contained several multiple tandem repeat elements, which are a remarkable feature of this region (Zhang & Hewitt, 1997). The A+T rich region of *D. stuposa* didn’t harbor long repeats, but we found a few short repeating sequences in this region. It has an ‘ATAGA’ motif alongwith a 20 bp poly-T repeats, a microsatellite-like (AT)₁₀ and a poly-A repeat sequence upstream of the *trnM* (Figure 6B).

Phylogenetic relationships

To determine the phylogenetic position of *D. stuposa*, we reconstructed the phylogenetic relationship among superfamily Noctuoidea. In phylogenetic analysis, mitogenome PCGs has the lowest sensitivity to analytical changes compared to that of other genes such as tRNA or rRNA genes (Yang et al., 2015). Additionally, the third codon of PCGs may influence the phylogenetic resolution and has been reported in different orders (Cameron et al., 2010; Yang et al., 2013). Here, we applied the amino acid sequences of 13 PCGs for phylogenetic analyses by using BI and ML methods. The results revealed that *D. stuposa* is closely related to *A. plana lacteata* that was well supported by both the analytical approaches (Figure 7A and 7B). In addition, the

different methods produced identical topologies with both posterior probabilities and bootstrap support for each node (Figure 7A and 7B). Hypeninae is closely related to Lymantriinae, and Lymantriinae is a subfamily of Erebidae and sister group of Arctiinae and others. The novel lineages of Erebidae contain different subfamilies: Arctiinae, Aganainae, Catocalinae, Lymantriinae and Hypeninae. In family level, the results showed a monophyletic Noctuoidea contains known families Notodontidae, Erebidae, Nolidae and Noctuidae.

Discussion

The *D. stuposa* mitogenome is a circular DNA molecule and contains 37 genes, which usually present in insect mitogenome (Figure 1). Compared with other mitogenomes, the size variation is generally due to difference in the length of their noncoding regions (intergenic spacers and A+T rich region) (Lv et al., 2018). Furthermore, the A and T biased nucleotide composition is a characteristic feature of insect mitogenomes. *D. stuposa* nucleotide composition is highly biased towards using A and T (A=39.98%, T=40.38%, G=7.5%, C=12.14%) (Table 2). However, the pattern of base composition in different insect species remains complex to the present time and still it is a matter of debate. The most primitive pterygote insect, *Parafronurus youi* mitogenome contains only 57% A+T content, lowest among all the sequenced insect species (Zhang et al., 2008). A+T content increased in insects throughout the evolutionary period, and the most advanced form of insects (holometabolous insects) constitute higher A+T content (Gotzek et al., 2010). The AT and GC skewness were determined to reflect the base composition in another side. The negative AT skewness (-0.005) and GC skewness (-0.236) of *D. stuposa* mitogenome, indicating the presence of more Ts than As, and Cs than Gs, and consistent with

previous several other insect species.

We identified 13 PCGs in *D. stuposa* mitogenome, with a length of 11,334 bp. Nine PCGs were resided on the J-strand, four PCGs were located on the N-strand. Additionally, *coxI* utilizes CGA codon to initiate the transcription (Table 3). The utilize of untypical initiation codon for *coxI* is a common feature across insects (Liu et al., 2014; Dai et al., 2016). The codon usage was used to estimate the amino acids of PCGs and CDspT was used to illustrate the codons distribution. We select seven Noctuoidea sequences belonging to four families and want to find some similarities and differences. The results showed that the codon usage and CDspT both exhibited approximately similar trends in the Noctuoidea. Asn and Ile were the most commonly used amino acids and contained the maximum CDspT value. In addition, *D. stuposa* mitogenome lacks ACG and CCG codons (Figure 4). Besides, some species are lack of synonymous codon with high G or C in the third codon position, such as GCG, CGC, GGC and CCG were abandoned in *A. ipsilon*. The rare presence or absence of high G and C content codons usually occur in various insect species, and are considered to be a remarkable feature for them (Yu et al., 2017; Li et al., 2018).

D. stuposa contains 22 tRNA genes, and most of them can formed the typical clover-leaf secondary structure. *trnSI* is the exception which lacks the dihydrouridine (DHU) arm. The unnormal feature of *trnSI* is common across mitogenome of insects (Lavrov et al., 2000; Zhang et al., 2013). Meanwhile, the secondary structure of *D. stuposa* tRNAs also contains many mismatches. These mismatches usually happened in the stem structure of tRNAs, and can be corrected by RNA-editing processes in many insect mitogenomes (Lavrov et al., 2000). In

overlapping and intergenic spacer regions, *D. stuposa* has some common and conserved features with other insects (Cameron & Whiting, 2008), such as a seven bp overlapping region located between *atp6* and *atp8*, an ‘ATACTAA’ motif located between *nad1* and *trnS2*. The 406 bp sized A+T rich region of *D. stuposa* had highest A+T content and a few features like an ‘ATAGA’ motif, a 20 bp poly-T repeats and so on (Figure 6B). The poly-T stretch varies in different species (Dai et al., 2015), but ‘ATAGA’ motif is conserved in insects (Cameron & Whiting, 2008). Besides, the A+T rich region usually contains extra tRNA-like structure which harbored correct anticodon structure and clover-leaf structure (Gotzek et al., 2010). Interestingly, *D. stuposa* do not contain this structure, that let us wonder the extra tRNA-like structure is necessary or not, and what function is it?

Phylogenetic relationships were established using ML and BI analytical approaches. These results exhibited that *D. stuposa* belongs to the Erebiidae family and Catocalinae subfamily that is consistent with the newly reported classification of Erebiidae (Zahiri et al., 2012). The Erebiidae is the largest and complicated noctuoid family, however its monophyly still has not been completely confirmed even some researchers has made efforts to resolve this problem. Additionally, a few species of the *Dysgonia* genus have been reclassified into other families (Holloway & Miller, 2003) that further cause complications in phylogenetic analysis. The present study clarified the taxonomic status of *D. stuposa*, which will help to understand the classification of *Dysgonia*. Within Erebiidae, the monophyly between subfamilies suggested that Lymantriinae is a subfamily and sisters group of Arctiinae and others, and has been strongly supported by Zahiri et al., 2011 (Figure 7A and 7B). We have also found that Noctuidae is a

stable system with high bootstrap values in each node except in Noctuinae and Hadeninae. In fact, Hadeninae and Noctuinae consist a “core” Noctuinae with other tribes like Caradrinini, Noctuini, and so on (Regier et al., 2016). However, enough mitogenome sequences are not available to support this stable relationship.

Within the family level, the phylogenetic reconstruction of Noctuoidea has been proposed by using different approaches and resulted different family groups. Kitching and Rawlins based on the apomorphic character, classified Noctuoidea into three fundamental lineages: Oenosandridae, Doidae+Notodontidae, and the quadrid families (including Lymantriidae, Arctiidae, Nolidae, Noctuidae, etc) (Kitching et al., 1998). Later on, Fibiger and Lafontaine suggested a new classification with five families: Oenosandridae, Micronoctuidae, Doidae, Notodontidae and Noctuidae (Fibiger & Lafontaine, 2005). However, these families still have unnatural groups and some species of quadrid families remains difficult to classify. Zahiri et al., reconstructed the Noctuoidea by using the molecular evolutionary methods and gene sequences (mitochondrial gene and nuclear genes). The authors suggested a stable family level classification, which contains six major lineages and interpret as families: Notodontidae, Erebidae, Nolidae, Oenosandridae, Euteliidae and Noctuidae (Zahiri et al., 2011). In the present study, we observed that the superfamily can be classify into 4 families viz., Notodontidae, Erebidae, Nolidae and Noctuidae. Since the data is limited on the complete mitogenome sequences of species belonging to Oenosandridae and Euteliidae in the public repository NCBI, our results are consistent with the previous phylogenetic hypothesis that was proposed by Zahiri et al., 2011. Overall, we inferred from our study, more mitogenomes studies on Noctuoidea are

required to better analyze and understand the phylogenetic relationships among them.

Conclusion

In summary, we analyzed the structure and composition of *D. stuposa* mitogenome and clarified the taxonomic status of this species in Lepidoptera. The *D. stuposa* mitogenome was a circular DNA molecule that contained some characteristic features such as the A+T biased nucleotide composition and the *cox1* gene initiated by non-canonical CGA codon, a conserved sequence ‘ATACTAA’ in a intergenic spacer, a motif ‘ATAGA’ followed by a 20 bp poly-T stretch in A+T rich region. The visible and invisible features of this mitogenome will provide more molecular information for the improvement of various research fields. Further, the phylogenetic analyses revealed that *D. stuposa* reside in the Erebidae family, and it also exhibited that a novel and robust lineages of Erebidae that contain subfamilies: Arctiinae, Aganainae, Catocalinae, Lymantriinae, Hypeninae.

Funding

This work was supported by the National Undergraduate Training Program for Innovation and Entrepreneurship [Grant No. 201810373111] and Youth Project of Anhui Natural Science Foundation [1908085QC94].

Competing Interests

The authors declare that they have no competing interests.

316

317 Reference

318 Arimoto M, Iwaizum R. 2014. Identification of Japanese *Lymantria* species (Lepidoptera:
319 *Lymantriidae*) based on PCR-RFLP analysis of mitochondrial DNA. *Applied Entomology*
320 *and Zoology* 49:159-169.

321 Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences, *Nucleic acids*
322 *research* 27:573-580.

323 Boore JL. 1999. Animal mitochondrial genomes. *Nucleic Acids Research* 27:1767-1780.

324 Cameron SL. 2014. Insect Mitochondrial Genomics: Implications for Evolution and Phylogeny.
325 *Annual Review of Entomology* 59: 95-117.

326 Cameron SL, Lambkin CL, Barker SC, et al. 2010. A mitochondrial genome phylogeny of
327 *Diptera*: whole genome sequence data accurately resolve relationships over broad
328 timescales with high precision. *Systematic Entomology* 32:40-59.

329 Cameron SL, Whiting MF. 2008. The complete mitochondrial genome of the tobacco hornworm,
330 *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial
331 gene variability within butterflies and moths. *Gene* 408:0-123.

332 Chen YX. 2003. *Fauna Sinica*: Vol. 16. Insecta, Lepidoptera, Noctuidae (Chinese Edition).
333 Beijing: Science Press, 2-3.

334 Clayton DA. 1991. Replication and Transcription of Vertebrate Mitochondrial DNA. *Annu Rev*
335 *Cell Bio* 17:453-478.

336 Dai LS, Qian C, Zhang CF, et al. 2015. Characterization of the Complete Mitochondrial Genome

of *Cerura menciata* and Comparison with Other Lepidopteran Insects. *PLoS ONE* 10:e0132951.

Dai LS, Zhu BJ, Qian C, et al. 2016. The complete mitochondrial genome of the diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae). *Mitochondrial Dna. Part A* 27:1512-1513.

Fernández-Silva P, Enriquez JA, Montoya J. 2003. Replication and transcription of mammalian mitochondrial DNA. *Exp. Physiol.* 88:41-55.

Fibiger M, Lafontaine JD. 2005. A review of the higher classification of the Noctuoidea (Lepidoptera) with special reference to the Holarctic fauna. *Esperiana* 11:7-92.

Gotzek D, Clarke J, Shoemaker DW. 2010. Mitochondrial genome evolution in fire ants (Hymenoptera: Formicidae). *Bmc Evolutionary Biology* 10:300.

Holloway JD, Miller SE. 2003. The composition, generic placement and host-plant relationships of the joviana-group in the *Parallelia* generic complex (Lepidoptera : Noctuidae, Catocalinae). *Invertebrate Systematics* 17:111-128.

Junqueira ACM, Lessinger AC, Torres TT, et al. 2004. The mitochondrial genome of the blowfly *Chrysomya chloropyga* (Diptera: Calliphoridae). *Gene* 339:7-15.

Kitching IJ, Rawlins JE, Kristensen NP. 1998. The Noctuidae. Handbook of Zoology, vol. IV Arthropoda: Insecta. Lepidoptera, Moths and Butterflies, vol. 1: Evolution, Systematics and Biogeography. Berlin: De Gruyter, 355-402.

Lafontaine JD, Fibiger M. 2006. Revised higher classification of the Noctuoidea (Lepidoptera). *Canadian Entomologist* 138:610-635.

358 Lavrov DV, Brown WM, Boore JL. 2000. A novel type of RNA editing occurs in the
359 mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proceedings of the National*
360 *Academy of Sciences* 97:13738-13742.

361 Li J, Zhang Y, Hu K, et al. 2018. Mitochondrial genome characteristics of two Sphingidae
362 insects (*Psilogramma increta*, and, *Macroglossum stellatarum*) and implications for their
363 phylogeny. *International Journal of Biological Macromolecules* :S0141813018304422.

364 Liu QN, Zhu BJ, Dai LS, et al. 2014. The complete mitochondrial genome of the common
365 cutworm, *Spodoptera litura* (Lepidoptera: Noctuidade). *Mitochondrial DNA* :1-2.

366 Liu Y, Li Y, Pan M, et al. 2008. The complete mitochondrial genome of the Chinese oak
367 silkworm, *Antheraea pernyi* (Lepidoptera: Saturniidae), *Acta Biochim. Biophys. Sin.* 40:693-
368 703.

369 Lowe TM, Eddy SR. 1997. tRNAscan-SE: A Program for Improved Detection of Transfer RNA
370 Genes in Genomic Sequence. *Nucleic Acids Research* 25:955-64.

371 Lu HF, Su TJ, Luo AR, et al. 2013. Characterization of the Complete Mitochondrion Genome of
372 Diurnal Moth *Amata emma* (Butler) (Lepidoptera: Erebidae) and Its Phylogenetic
373 Implications. *PLoS ONE* 8:, e72410.

374 Lv C, Li Q, Kong L, et al. 2018. Comparative analyses of the complete mitochondrial genomes
375 of *Dosinia* clams and their phylogenetic position within Veneridae. *PLoS ONE* 13:
376 e0196466.

377 Miller JS. 1991. Cladistics and classification of the Notodontidae (Lepidoptera, Noctuoidea)
378 based on larval and adult morphology. *Bulletin of the American Museum of Natural History*

204:1-230.

Minet J, Barbut J, Lalanne-Cassou B. 2012. Les Noctuelles: classification et clé de détermination

des familles (Lepidoptera: Noctuoidea). *Alexanor* 25:131-151.

Nieukerken E, Kaila L, Kitching I, et al. 2011. Order Lepidoptera Linnaeus, 1758. *Zootaxa*

3148:212-221.

Raupach MJ, Astrin JJ, Hannig K, et al. 2010. Molecular species identification of Central

European ground beetles (Coleoptera: Carabidae) using nuclear rDNA expansion segments

and DNA barcodes. *Frontiers in Zoology* 7:1-15.

Regier JC, Mitter C, Mitter K, et al. 2016. Further progress on the phylogeny of Noctuoidea

(Insecta: Lepidoptera) using an expanded gene sample. *Systematic Entomology* 42:1-11.

Simon C, Frati F, Beckenbach A, et al. 1994. Evolution, Weighting, and Phylogenetic Utility of

Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain

Reaction Primers. *Annals of the Entomological Society of America* 87:651-701.

Speidel W, Fanger H. 2010. The phylogeny of the Noctuidae (Lepidoptera). *Systematic*

Entomology 21:219-252.

Sun YX, Wang L, Wei GQ, et al. 2016. Characterization of the Complete Mitochondrial Genome

of *Leucoma salicis* (Lepidoptera: Lymantriidae) and Comparison with Other Lepidopteran

Insects. *Scientific Reports* 6:39153.

Tamura K, Peterson D, Peterson N, et al. 2011. MEGA5: Molecular Evolutionary Genetics

Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony

Methods. *Mol. Biol. Evol.* 28:2731-2739.

- Thompson JD, Gibson TJ, Plewniak F. 1997. The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic. Acids. Res.* 25:4876-4882.
- Wilson AC, Cann RL, Carr SM, et al. 2010. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society* 26:375-400.
- Wolstenholme DR. 1992. Animal mitochondrial DNA: structure and evolution. *International Review of Cytology* 141: 173.
- Yang L, Wei ZJ, Hong GY, et al. 2009. The complete nucleotide sequence of the mitochondrial genome of *Phthonandria atrilineata* (Lepidoptera: Geometridae). *Molecular Biology Reports* 36:1441-1449.
- Yang X, Cameron SL, Lees DC, et al. 2015. A mitochondrial genome phylogeny of owlet moths (Lepidoptera: Noctuoidea), and examination of the utility of mitochondrial genomes for lepidopteran phylogenetics. *Molecular Phylogenetics and Evolution* 85:230-237.
- Yang XS, Xue DY, Han HX. 2013. The complete mitochondrial genome of *Biston panterinaria* (Lepidoptera: Geometridae), with phylogenetic utility of mitochondrial genome in the Lepidoptera. *Gene* 515:349-358.
- Yu S, Chen C, Jin G, et al. 2017. Comparative mitochondrial genome analysis of *Daphnis nerii* and other lepidopteran insects reveals conserved mitochondrial genome organization and phylogenetic relationships. *PLoS ONE* 12:e0178773.
- Zahiri R, Holloway JD, Kitching IJ, et al. 2012. Molecular phylogenetics of Erebidae (Lepidoptera, Noctuoidea). *Systematic Entomology* 37:102-124.

- 421 Zahiri R, Kitching IJ, Lafontaine JD, et al. 2011. A new molecular phylogeny offers hope for a
422 stable family level classification of the Noctuoidea (Lepidoptera). *Zoologica Scripta*
423 40:158-173.
- 424 Zhang DX, Hewitt GM. 1997. Insect mitochondrial control region: A review of its structure,
425 evolution and usefulness in evolutionary studies. *Biochemical Systematics and Ecology*
426 25:99-120.
- 427 Zhang HL, Zeng HH, Huang Y, et al. 2013. The complete mitochondrial genomes of three
428 grasshoppers, *Asiotmethis zacharjini*, *Filchnerella helanshanensis* and *Pseudotmethis*
429 *rubimarginis* (Orthoptera: Pamphagidae). *Gene* 517:89-98.
- 430 Zhang J, Zhou C, Gai Y, et al. 2008. The complete mitochondrial genome of *Parafronurus youi*
431 (Insecta: Ephemeroptera) and phylogenetic position of the Ephemeroptera. *Gene* 424:8-24.
- 432 Zhu BJ, Liu QN, Dai LS, et al. 2013 Characterization of the complete mitochondrial genome of
433 *Diaphania pyloalis* (Lepidoptera: Pyralididae). *Gene* 527:283-291.

Table 1 (on next page)

Details of the primers used to amplify the mitochondrial DNA of *D. stuposa*

1 **Table 1.** Details of the primers used to amplify the mitochondrial DNA of *D. stuposa*

Primer name	Orientation	Annealing position (bp)	Nucleotide sequence (5'-3')	PCR length
S1F	F	10170-10192	ACTTTAAAACTTCAAAGAAAAA	5668
S1R	R	93-116	ACTTAATTTATCCTATCAGAATAA	
S2F	F	15136-15153	CGCAACTGCTGGCACAAA	7360
S2R	R	6748-6774	GAAGAGAAGTTTATAGTGGATGAGGTT	
S3F	F	6360-6389	TAAGCTGCTAACTTAATTTTATAGT	4775
S3R	R	11112-11134	GTAATAAATTCCTCGTCCAATAT	

2

Table 2(on next page)

Composition and skew in different Noctuoidea mitogenomes.

1 **Table 2.** Composition and skew in different Noctuoidea mitogenomes.

Species	Size (bp)	A%	G%	T%	C%	A+T %	AT skewness	GC skewness
Whole genome								
<i>D. stuposa</i>	15721	39.98	7.5	40.38	12.14	80.36	-0.005	-0.236
<i>A. plana lacteata</i>	15416	40.08	7.49	40.26	12.16	80.34	-0.002	-0.238
<i>V.a virilis</i>	15417	40.18	7.56	40.22	12.05	80.4	0.000	-0.229
<i>G. minora</i>	15801	40.97	6.77	40.52	11.75	81.49	0.006	-0.269
<i>R. prominens</i>	15343	40.25	7.8	40.82	11.13	81.07	-0.007	-0.176
<i>O. lunifer</i>	15593	40.09	7.56	37.75	14.6	77.84	0.030	-0.318
<i>A. ipsilon</i>	15377	40.38	7.71	40.87	11.04	81.25	-0.006	-0.178
PCG								
<i>D. stuposa</i>	11334	33.86	10.91	44.63	10.61	78.49	-0.137	0.014
<i>A. plana lacteata</i>	11211	33.87	10.92	44.76	10.45	78.63	-0.138	0.022
<i>V. virilis</i>	11203	33.14	11.16	45.43	10.27	78.57	-0.156	0.042
<i>G. minora</i>	11237	34.72	10.11	44.98	10.2	79.7	-0.129	-0.004
<i>R. prominens</i>	11216	33.64	10.57	46	9.8	79.64	-0.155	0.038
<i>O. lunifer</i>	11266	32.47	12.08	43.26	12.19	75.73	-0.142	-0.005
<i>A. ipsilon</i>	11211	34.24	10.64	45.56	9.55	79.8	-0.142	0.054
AT RICH								
<i>D. stuposa</i>	406	43.6	2.46	48.77	5.17	92.37	-0.056	-0.355
<i>A. plana lacteata</i>	328	46.04	1.22	48.48	4.27	94.52	-0.026	-0.556
<i>V. virilis</i>	362	44.48	1.1	50.55	3.87	95.03	-0.064	-0.557
<i>G. minora</i>	449	43.21	2.67	49.44	4.68	92.65	-0.067	-0.273
<i>R. prominens</i>	342	44.15	2.34	49.42	4.09	93.57	-0.056	-0.272
<i>O. lunifer</i>	319	44.51	1.57	48.9	5.02	93.41	-0.047	-0.524
<i>A. ipsilon</i>	332	46.08	1.51	48.8	3.61	94.88	-0.029	-0.410

2

Table 3(on next page)

List of annotated mitochondrial genes of *D. stuposa*

1 **Table 3.** List of annotated mitochondrial genes of *D. stuposa*

Gene name	Start	Stop	Strand	Length	Anti-codon	Start codon	End codon	Intergenic nucleotides
<i>trnM</i>	1	68	J	68	CAT	/	/	2
<i>trnI</i>	71	138	J	68	GAT	/	/	8
<i>trnQ</i>	147	215	N	69	TTG	/	/	55
<i>nad2</i>	271	1284	J	1014	/	ATT	TAA	-2
<i>trnW</i>	1283	1350	J	68	TCA	/	/	-8
<i>trnC</i>	1343	1409	N	67	GCA	/	/	22
<i>trnY</i>	1432	1496	N	65	GTA	/	/	9
<i>coxI</i>	1506	3041	J	1536	/	CGA	TAA	-5
<i>trnL2</i>	3037	3103	J	67	TAA	/	/	0
<i>cox2</i>	3104	3820	J	717	/	ATA	TAA	-35
<i>trnK</i>	3786	3856	J	71	CTT	/	/	0
<i>trnD</i>	3857	3923	J	67	GTC	/	/	0
<i>atp8</i>	3924	4085	J	162	/	ATC	TAA	-7
<i>atp6</i>	4079	4756	J	678	/	ATG	TAA	27
<i>cox3</i>	4784	5572	J	789	/	ATG	TAA	2
<i>trnG</i>	5575	5640	J	66	TCC	/	/	0
<i>nad3</i>	5641	5994	J	354	/	ATT	TAA	34
<i>trnA</i>	6029	6085	J	57	TGC	/	/	105
<i>trnR</i>	6191	6256	J	66	TCG	/	/	10
<i>trnN</i>	6267	6332	J	66	GTT	/	/	8
<i>trnS1</i>	6341	6406	J	66	GCT	/	/	32
<i>trnE</i>	6439	6506	J	68	TTC	/	/	50
<i>trnF</i>	6557	6624	N	68	GAA	/	/	-17
<i>nad5</i>	6608	8368	N	1761	/	ATT	TAA	-3
<i>trnH</i>	8366	8433	N	68	GTG	/	/	-65
<i>nad4</i>	8369	9772	N	1404	/	ATG	TAA	42
<i>nad4l</i>	9815	10102	N	288	/	ATG	TAA	14
<i>trnT</i>	10117	10181	J	65	TGT	/	/	0
<i>trnP</i>	10182	10246	N	65	TGG	/	/	7
<i>nad6</i>	10254	10784	J	531	/	ATT	TAA	14
<i>cob</i>	10799	11959	J	1161	/	ATG	TAA	-2
<i>trnS2</i>	11958	12025	J	68	TGA	/	/	22
<i>nad1</i>	12048	12986	N	939	/	ATG	TAA	1
<i>trnL1</i>	12988	13055	N	68	TAG	/	/	65
<i>rrnL</i>	13121	14428	N	1308	/	/	/	37
<i>trnV</i>	14466	14533	N	68	TAC	/	/	0
<i>rrnS</i>	14534	15315	N	782	/	/	/	0
AT-rich	15316	15721	/	406	/	/	/	/

region

2

Figure 1

Map of the mitogenome of *Dysgonia stuposa*.

The tRNA genes are labeled according to the IUPAC-IUB three-letter amino acids; *cox1*, *cox2* and *cox3* refer to the cytochrome c oxidase subunits; *cob* refers to cytochrome b; *nad1-nad6* refer to NADH dehydrogenase components.

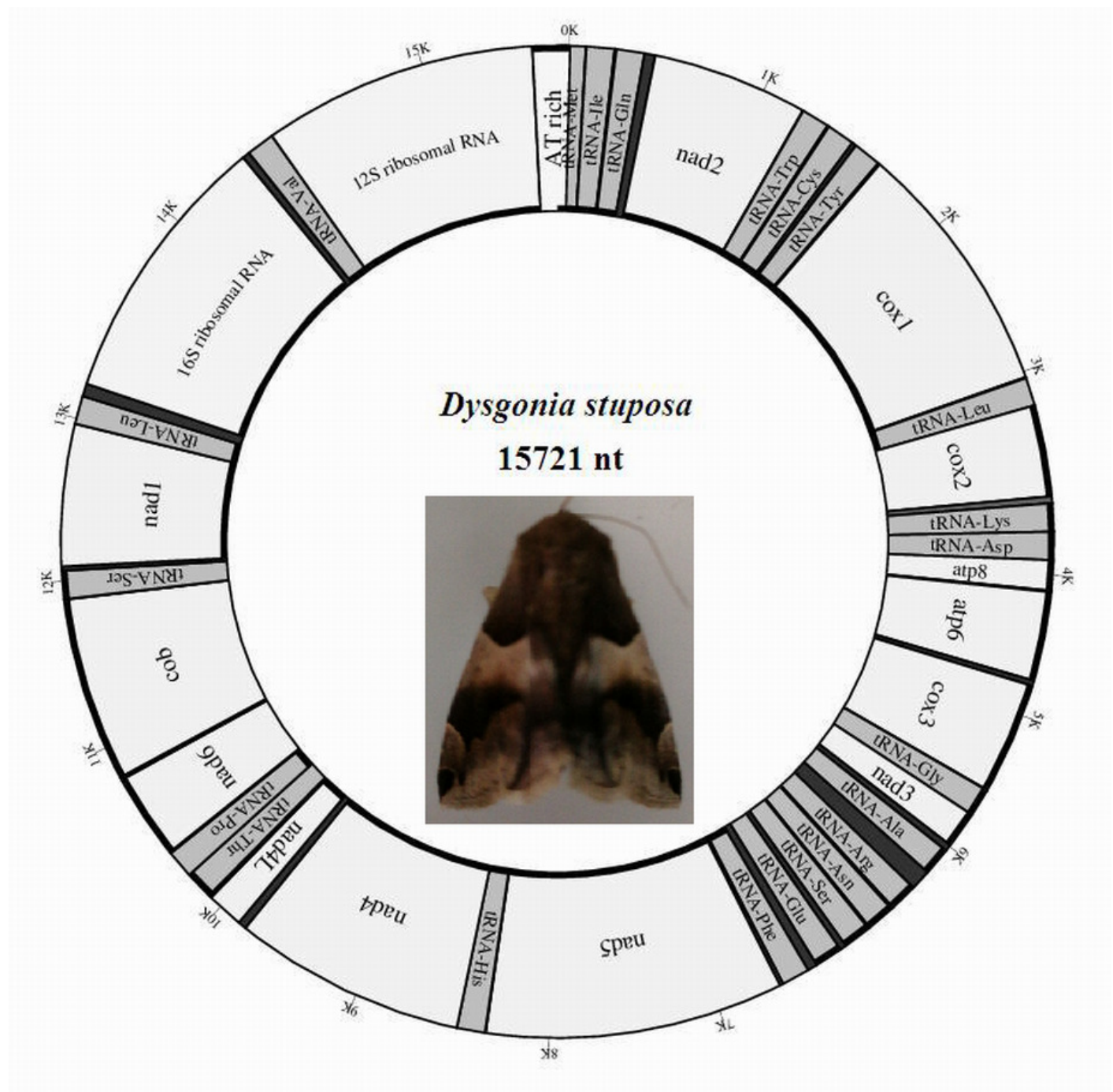


Figure 2

Comparison of codon usage within the mitochondrial genome in the Noctuoidea.

Lowercase letters (a, b, c and d) above the name of species represent the family (a: Erebiidae, b: Nolide, c: Notodontidae, d: Noctuidae).

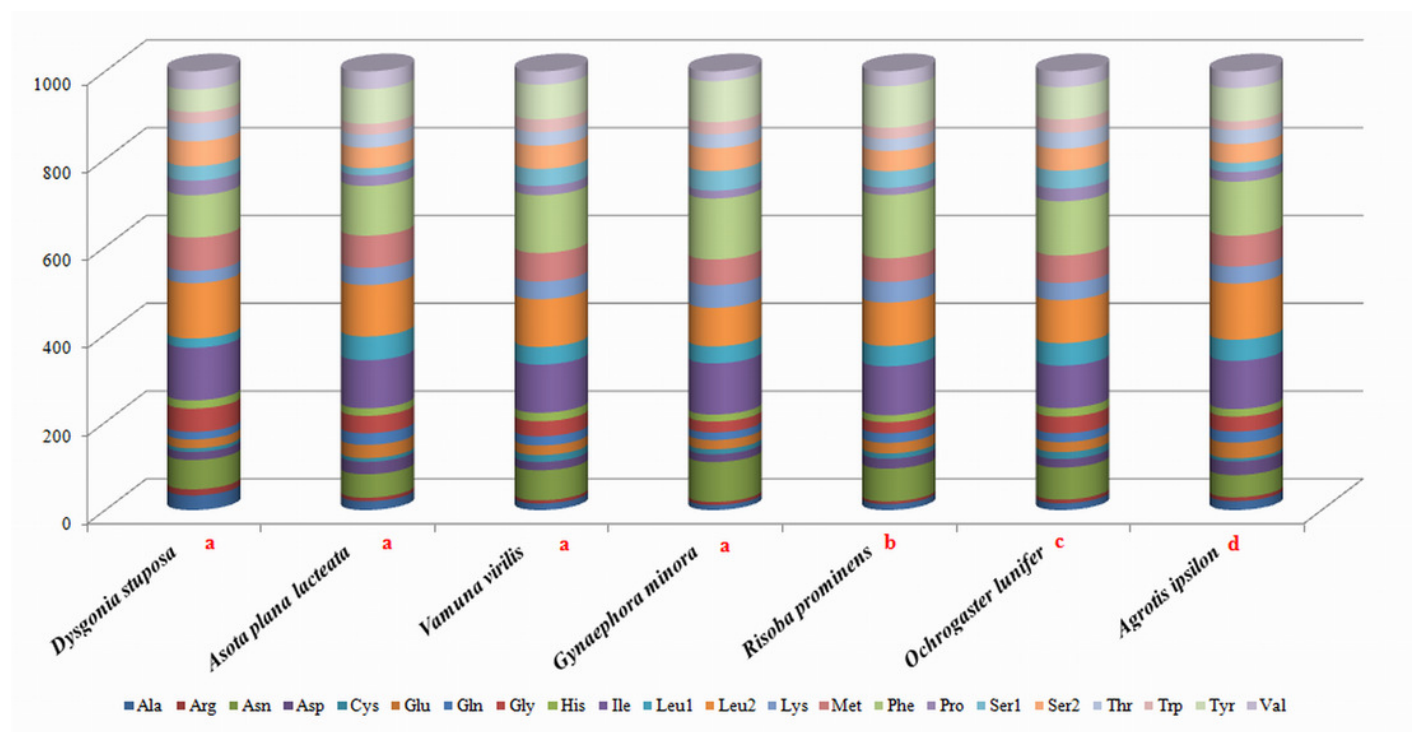


Figure 3

Codon distribution in members of the Noctuoidea. CDspT = codons per thousand codons.

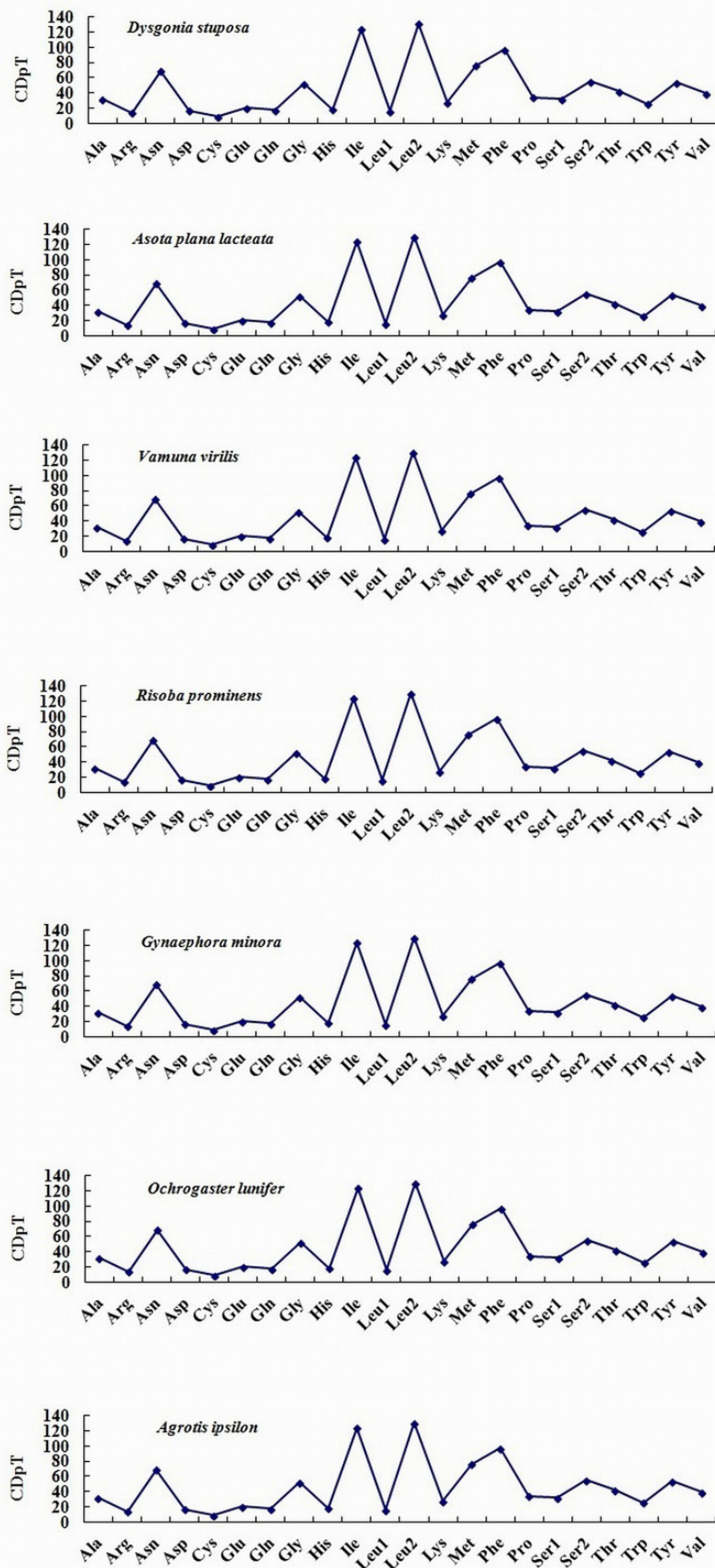


Figure 4

Relative Synonymous Codon Usage (RSCU) of the mitochondrial genome of four families in the Noctuoidea.

Codon families are plotted on the x-axis. Codons above the bars are absent in the mitogenome.

Figure 5

Predicted secondary structures of the 22 tRNA genes of the *D. stuposa* mitogenome.

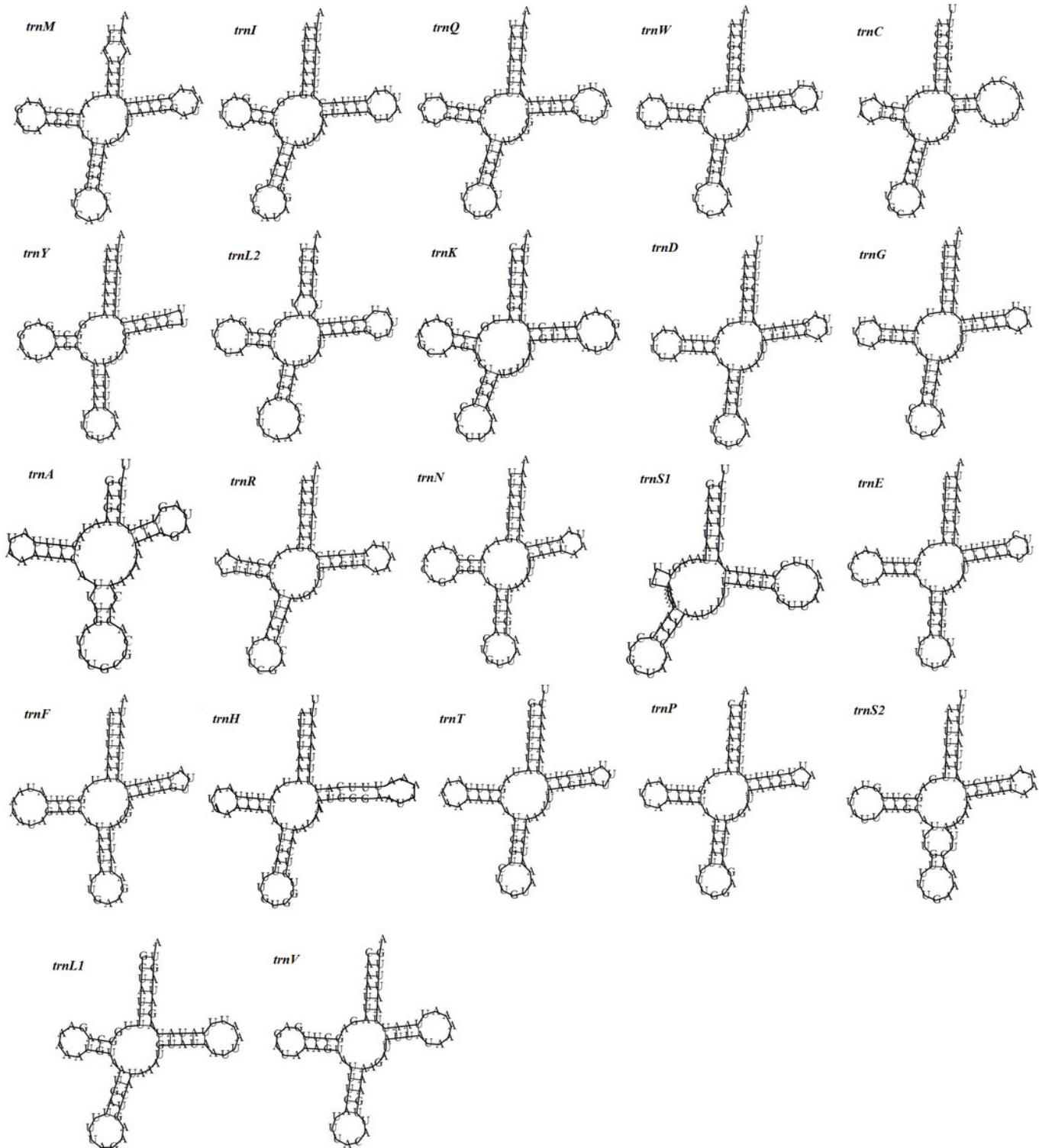


Figure 6

Alignment of the intergenic spacer region and the A+T-rich region of *Dysgonia stuposa*

(A) Alignment of the intergenic spacer region between *trnS2* and *nad1* of different Noctuoidea species. (B) Features present in the A+T-rich region of *Dysgonia stuposa*. The 'ATAGA' motif is shaded. The poly-T stretch is underlined and the poly-A stretch is double underlined. The single microsatellite 'AT' repeat sequence is indicated by dotted underlining.

A	<i>Dysgonia stuposa</i> (Erebidae)	ATACTAAAAATAATCAACAAAA
	<i>Asota plana lacteata</i> (Erebidae)	ATACTAAAAATAATCAATAA
	<i>Vamuna virilis</i> (Erebidae)	ATACTAAAAATAATTAATT
	<i>Gynaephora minora</i> (Erebidae)	TATACTAAAAAAAATTATACAATTA
	<i>Risoba prominens</i> (Nolide)	ATACTAAAAATAATTAA
	<i>Ochrogaster lunifer</i> (Notodontidae)	ATACTAAAAATAATTAA
	<i>Agrotis ipsilon</i> (Noctuidae)	ATACTAAAAAAAATTAAA

B

rrnS-15,315-TTTATATGCACAATTTCTCACATAGATTTTTTTTTT
TTTTTTTTTTTTATATTTAAATTTTATTATATAATATTATTTTAT
ATTAAAATATTTAATATAATTATTAAATATTAAATAATTTCTTT
TTCTTTTTTCTTCATACTATTCATATTGAAACCTAATTTGGAA
ATTAAACAATTACAATTCTTAAAAATTACAATATATTAATATAA
TTAATAATAATTTTTCTTAATAAGTTAATGAATTATAAATATTT
TAATTTATTTAAAAATTTAATATATATATATAAATATTAATTTTA
TAAAAATTTAATATATATATATATATATAATTTTAAAGAAAAT
TATTATTTAATTATGTATTTAAACCATTTTAAATAATAATGCAT
ATAAATAAAAAAAAATA-15,721-trnM

Figure 7

Phylogenetic relationships of Noctuoidea

(A) Tree showing the phylogenetic relationships among 29 species, constructed using Maximum Likelihood with 1000 bootstrap replicates. (B) Tree constructed using Bayesian Inference (BI) MCMC consensus tree, with posterior probabilities shown at nodes. *Bombyx mori* (AY048187) and *Antheraea pernyi* (AY242996) were used as outgroups.

