



OPEN

Characterization of four mitochondrial genomes from superfamilies Noctuoidea and Hyblaeoidea with their phylogenetic implications

Rauf Ahmad Shah¹, Muzafar Riyaz¹, Savarimuthu Ignacimuthu² & Kuppusamy Sivasankaran¹

In the present study, the newly sequenced mitogenomes of three Noctuid and one Hyblaeoid (Insecta: Lepidoptera) species were annotated based on next-generation sequence data. The complete mitogenome lengths of *Oraesia emarginata*, *Actinotia polyodon*, *Odontodes seranensis*, and *Hyblaea puera* were 16,668 bp, 15,347 bp, 15,419 bp, and 15,350 bp, respectively. These mitogenomes were found to encode 37 typical mitochondrial genes (13 protein-coding, 22 transfer RNA, 2 ribosomal RNA) and a control region, similar to most Lepidoptera species. Maximum likelihood (ML) methods and Bayesian inference (BI) were used to reconstruct the phylogenetic relationships of the moths. This study showed the relationships of Noctuid families as follows: (Notodontidae + (Erebidae + (Nolidae + (Euteliidae + Noctuidae))). Furthermore, the species *H. puera* was separately clustered from the Noctuoidea member groups. Till now, the species from the superfamily Hyblaeoidea have not been discussed for their phylogenetic relationships. In this study, the complete mitochondrial genome of one species from the superfamily Hyblaeoidea was analysed.

Complete mitochondrial genomes (mitogenomes) have been widely used to infer phylogenetic relationships specifically for wide-ranging groups of organisms such as insects. Mitogenomes have been considerably used as molecular tools for phylogenetic investigations, and comparative and evolutionary relationship studies^{1–3}. The insect mitogenomes are relatively small in size, show rapid evolution rates, have low level of recombination, and possess maternal inheritance^{4–7}. Therefore, the utilization of mitogenome is expected to provide novel information concerning the classification of insects and assessments of their evolutionary features.

The insect mitochondrial genome is typically a double-stranded, circular molecule that is 14–19 kb in length and is composed of 13 protein-coding genes (PCGs): two ATPase genes (*atp6* and *atp8*), three cytochrome c oxidase genes 1–3 (*cox1-cox3*), one cytochrome B (*cob*), seven NADH dehydrogenase genes (*nad1-6* and *nad4L*), 22 transfer RNA (tRNA), two ribosomal RNA (*rrnL* and *rrnS*) genes and non-coding A + T-rich region^{2,5,8–10}.

There are about 1,57,424 described moth species worldwide, from 45 superfamilies, and belonging to 139 families¹¹. Analysis of the phylogenetic relationships of lepidopteran moths using mitogenomes has increased rapidly during the last decade¹². However, the primary focus has been on 11 moth superfamilies namely Bombycoidea, Cossioidea, Gelechioidea, Geometroidea, Hepialoidea, Noctuoidea, Pyraloidea, Tineoidea, Tortricoidae, Yponomeutoidea, and Zygaenoidea.

Noctuoidea is the largest superfamily of the order Lepidoptera, comprising 42,407 species¹¹. The monophyly of Noctuoidea is supported by the existence of a gained apomorphic character, metathoracic tympanal organ¹³. Phylogenetic studies of Noctuoidea were primarily analyzed using molecular methods based on one or two genes and with limited taxon sampling^{14–17}. The molecular phylogenetic relationship of Noctuoidea has been analyzed based on single mitochondrial (*cox1*) and seven nuclear genes (EF-1 α , wingless, RpS5, IDH, MDH, GAPDH, and CAD) from 152 species¹⁸. Zahiri et al.¹⁸ have proposed a novel perception, separating the traditional group of quadrifid noctuids, and re-establishing Erebidae and Nolidae as families. This result contrasted meaningfully

¹Division of Taxonomy and Biodiversity, Entomology Research Institute, Loyola Collège, Chennai, Tamil Nadu 600034, India. ²Xavier Research Foundation, St. Xavier's College, Palayamkottai, Tamil Nadu 627002, India. ✉email: ganesh_swamy2005@yahoo.com; sivasankaran@loyolacollege.edu

with previous investigations of both morphological and molecular studies. Nevertheless, this analysis failed to resolve the phylogenetic relationships between Erebidae subfamilies¹⁹. The phylogenetic relationships of the family Erebidae were analyzed using the mitogenomes^{20–34}. Several mitogenome sequencing studies were carried out in the family Noctuidae for their phylogenetic utilization^[35–42] and single mitogenomic analysis was done in the family Euteliidae¹².

Hyblaeoidea is among the smallest superfamilies in the order Lepidoptera and consists of only one family Hyblaeidae comprising only two genera *Hyblaea* and *Erythrochrus*. The family Hyblaeidae contains only 20 species distributed all over the new and old-world tropics and subtropics. The genus *Hyblaea* is known as a serious forest pest⁴³. Twort et al.⁴⁴ have done the whole genome sequence and analyzed the phylogenetic relationships of *Hyblaea puera* and *Hyblaea madagascariensis*. They used the dataset of 162 taxa for analyses, which showed the stable placement of *Hyblaea* as sister to the Pyraloidea member group with strongly supported values.

At present, 84 complete mitogenomes of Noctuoidea from eight families have been deposited in GenBank^{12,20,21,24–26,28,30,31,36,37,39,41,45–50}. In the present study, we sequenced the complete mitochondrial genomes of four species representing superfamilies Noctuoidea (*Oraesia emarginata*, *Actinotia polyodon*, *Odontodes seranensis*) and Hyblaeoidea (*Hyblaea puera*) for the first time. The mitogenomes of these species were annotated and the general characteristics of the mitogenome sequences were analyzed and compared. We analyzed phylogenetic relationships of mitogenomes from 90 lepidopteran taxa. In addition, the phylogenetic tree was reconstructed using the maximum likelihood method and Bayesian inference to evaluate the relationships among the lepidopteran moths.

Results and discussion

Genome organization and base composition. In this study, we sequenced and characterized the complete mitogenomes of three Noctuoidea species; *Oraesia emarginata* (GenBank Accession no. MW648382), *Actinotia polyodon* (GenBank Accession no. MW697903), *Odontodes seranensis* (GenBank Accession no. MW719565) and one Hyblaeoidea species *Hyblaea puera* (GenBank Accession no. MW885970). The sequences were deposited in GenBank (Table 1). The total lengths of the mitogenomes of *O. emarginata*, *A. polyodon*, *O. seranensis*, and *H. puera* were 16,668 bp, 15,347 bp, 15,419 bp, and 15,350 bp, respectively. The sizes of mitogenome sequenced so far in the superfamily Noctuoidea ranged from 15,229 bp in *Helicoverpa gelatopoeon* to 16,346 bp in *Spodoptera frugiperda*. The mitogenome sequence lengths of *A. polyodon*, *O. seranensis*, and *H. puera* fell within the range of mitogenome of other sequenced Noctuid moths. However, the mitogenome of *Oraesia emarginata* was larger than that of *Gynaephora jiuzeiensis*. The organization of newly sequenced mitogenomes of four species are presented in Fig. 1.

All four species comprised a distinctive metazoan mitogenome composition of 13 protein-coding genes PCGs viz. ATPase subunits 6 and 8 (*atp6* and *atp8*), cytochrome c oxidase subunits 1–3 (*cox1*, *cox2* and *cox3*), NADH dehydrogenase subunits 1–6 (*nad1*, *nad2*, *nad3*, *nad4*, *nad5* and *nad6*), subunit 4L *nad* (*nad4l*) and cytochrome B (*cob*), 22 transfer RNA (*tRNA*) genes, two ribosomal genes (*rrnL* and *rrnS*) and a control region (A + T-rich region). Four of the thirteen PCGs (*nad1*, *nad4*, *nad4l*, and *nad5*) and eight tRNAs (*trnQ*, *trnC*, *trnY*, *trnF*, *trnH*, *trnP*, *trnL1* and *trnV*) and two rRNAs were encoded on the N-strand, whereas the other 23 genes (9 PCGs and 14 tRNAs) and the control-region (A + T-rich) were encoded on the J-strand (Table 2). All genes were organized in the same way without the rearrangement phenomenon.

The nucleotide compositions of the four moth mitogenomes had a high A + T bias: 79.72% in *O. emarginata*, 81.69% in *A. polyodon*, 81.09% in *O. seranensis*, and 81.21% in *H. puera*. Among the 88 Noctuid species for which mtDNA data was available, the lowest A + T content was 77.83% in *O. lunifer*, while the highest A + T content was 81.69% in *Gabala argentata*. The mitogenome of *A. polyodon* was also highest among the known Noctuid mitogenomes. All four mitogenomes showed a negative AT-skew on the majority strand and negative GC-skew as it occurred mostly among other Noctuid mitogenomes. The AT and GC skew values on the majority strand of the four moths' species are *O. emarginata* (–0.002 and –0.200), *A. polyodon* (–0.008 and –0.167), *O. seranensis* (–0.023 and –0.196), and *H. puera* (–0.000 and –0.178) (Supplementary Table 1). Similar patterns of nucleotide negative skew have also been found in the mitogenomes of other Noctuid taxa^{36,45,46,49}.

Protein-coding genes and codon usage. The total lengths of the 13 PCGs of *O. emarginata*, *A. polyodon*, *O. seranensis* and *H. puera* were 11,182 bp, 11,213 bp, 11,208 bp, and 11,195 bp accounting for 67.08%, 73.06%, 72.68%, and 72.93% of the mitogenomes respectively. The locations and orientations of the 13 PCGs within the four mitogenomes were identical to those of most Noctuid species. The nucleotide PCGs translated into 3716–3725 amino acid-coding codons, excluding the stop codons. Similar to the PCGs *nad5* and *atp8* were observed to be the largest (1727–1746 bp) and smallest (162–165 bp) genes, respectively. The majority of PCGs stringently started with an ATN (ATG/ATT/ATA) start codon, except *cox1* gene which started with CGA in *O. emarginata*, *A. polyodon* and *H. puera* and with TTG in *O. seranensis* (Table 1). The majority of PCGs terminated with a complete and canonical stop codon (TAA/TAG) except in *O. emarginata* where gene *nad5* terminated with TTA. The genes *cox2*, *nad4* (*O. emarginata*, *A. polyodon*), *cox1*, *cox2*, *nad4* (*O. seranensis*), *cox1*, *cox2*, and *nad5* (*H. puera*) were found to have a truncated termination codon (T) and it might be altered by post-transcriptional polyadenylation. The presence of an incomplete stop codon was also a common phenomenon in metazoan mitochondrial genes (Sheffield et al. 2010). The average A + T contents of the 13 PCGs within the four mitogenomes ranged from 78.85 to 80.36% (Supplementary Table 1).

Relative synonymous codon usage (RSCU) was calculated in the mitogenomes of four lepidopterans (Fig. 2 and Supplementary Table 2). The most frequently utilized codons were almost similar within the four Noctuid species. UUU (Phe), UUA (Leu), AUU (Ile), AAU (Asn) and AAA (Lys) were the most consistently used codons (> 232) within the PCGs of the four mitogenomes; however, GUG (Val), ACG (Thr), CCG (Pro), GCC (Ala), and

Superfamily	Family	Subfamily	Species	Accession		
Noctuoidea	Erebidae	Calpinae	<i>Oraesia emarginata</i>	MW648382*		
			<i>Eudocima salamina</i>	MW683337		
			<i>Eudocima phalonia</i>	KY196412		
		Catocalinae	<i>Catocala deuteronympha</i>	KJ432280		
			<i>Grammodes geometrica</i>	KY888135		
			<i>Parallelia stuposa</i>	MK262707		
			Aganainae	<i>Asota plana lacteata</i>	KJ173908	
				Herminiinae	<i>Hydrillodes lentalis</i>	MH013484
			Hypeninae		<i>Paragabara curvicornuta</i>	KT362742
		Arctiinae		<i>Hyphantria cunea</i>	GU592049	
			<i>Vamuna virilis</i>	KJ364659		
			<i>Amata formosae</i>	KC513737		
			<i>Callimorpha dominula</i>	KP973953		
			<i>Nyctemera arctata albofasciata</i>	KM244681		
			<i>Paraona staudingeri</i>	KY827330		
			<i>Spilosoma lubricipeda</i>	MT591568		
			<i>Arctia plantaginis</i>	MW394229		
			<i>Lemyra melli</i>	KP307017		
			<i>Spilarctia subcarnea</i>	KT258909		
			<i>Cyana</i> sp. MT-2014	KM244679		
			<i>Eilema ussuricum</i>	MN696172		
		Lymantriinae	<i>Aglaomorpha histrio</i>	KY800518		
			<i>Lymantria dispar</i>	FJ617240		
			<i>Lymantria dispar asiatica</i>	KY923067		
			<i>Lymantria dispar japonica</i>	KY923060		
			<i>Lymantria</i> sp. AN-2017	KY923068		
			<i>Lymantria umbrosa</i>	KY923066		
			<i>Lymantria sugii</i>	MT265380		
			<i>Gynaephora minora</i>	KY688086		
			<i>Gynaephora menyuanensis</i>	KC185412		
			<i>Gynaephora jiuzeiensis</i>	KY688085		
			<i>Gynaephora ruoergensis</i>	KY688083		
			<i>Gynaephora qumalaiensis</i>	KJ507134		
			<i>Gynaephora qinghaiensis</i>	KJ507133		
			<i>Gynaephora aureata</i>	KY688084		
			<i>Euproctis pseudoconspersa</i>	KJ716847		
		Notodontidae	Phalerinae	<i>Euproctis similis</i>	KT258910	
				<i>Euproctis cryptosticta</i>	KY996558	
				<i>Euproctis seitzii</i>	MN916588	
				<i>Somena scintillans</i>	MH051839	
				<i>Lachana alpherakii</i>	KJ957168	
				<i>Orgyia postica</i>	MW355619	
				<i>Laelia suffusa</i>	MN908152	
				Thaumetopoeinae	<i>Phalera flavescens</i>	JF440342
					<i>Ochrogaster lunifer</i>	AM946601
					<i>Clostera anachoreta</i>	KX108766
					<i>Clostera anastomosis</i>	MH286069
Nolidae	Chloephorinae			<i>Thaumetopoea pityocampa</i>	MH286070	
		<i>Gabala argentata</i>	KJ410747			
		<i>Sinna extrema</i>	MG872330			
Euteliidae	Risobinae	<i>Risoba prominens</i>	KJ396197			
	Euteliinae	<i>Eutelia adulatricoides</i>	KJ185131			
Noctuidae	Stictopterinae	<i>Odontodes seranensis</i>	MW719565*			
		Amphipyriinae	<i>Actinotia polyodon</i>	MW697903*		
			<i>Sesamia inferens</i>	JN039362		
		<i>Spodoptera exigua</i>	JX316220			

Continued

Superfamily	Family	Subfamily	Species	Accession
			<i>Spodoptera litura</i>	KF543065
			<i>Spodoptera frugiperda</i>	KU877172
			<i>Spodoptera littoralis</i>	MT816470
			<i>Spodoptera exempta</i>	MT906792
		Noctuinae	<i>Agrotis ipsilon</i>	KF163965
			<i>Agrotis segetum</i>	KC894725
			<i>Striacosta albicosta</i>	KM488268
		Xyleninae	<i>Athetis lepigone</i>	MF152842
			<i>Athetis pallidipennis</i>	MT040606
		Heliolithinae	<i>Helicoverpa zea</i>	KJ930516
			<i>Helicoverpa punctigera</i>	KF977797
			<i>Helicoverpa armigera</i>	GU188273
			<i>Helicoverpa assulta</i>	MG437198
			<i>Helicoverpa gelotopoeon</i>	MG437199
			<i>Heliothis subflexa</i>	KT598688
			<i>Noctua pronuba</i>	KJ508057
		Acronictinae	<i>Acronicta psi</i>	KJ508060
		Hadeninae	<i>Mythimna separata</i>	KM099034
			<i>Mythimna pallidicosta</i>	MH027985
			<i>Mythimna loreyi</i>	MT506351
			<i>Leiomotopon simyrides</i>	MW255962
			<i>Protegera songi</i>	KY379907
			<i>Mamestra configurata</i>	CM017530
			<i>Anarta trifolii</i>	MN715147
		Plusiinae	<i>Ctenoplusia agnata</i>	KC414791
			<i>Ctenoplusia limbirena</i>	KM244665
			<i>Ctenoplusia albostrata</i>	MN495624
			<i>Trichoplusia ni</i>	MK714850
			<i>Diachrysis nadeja</i>	MT916722
Hyblaeoidea	Hyblaeidae		<i>Hyblaea puera</i>	MW885970*
Pyraloidea	Crambidae	Crambinae	<i>Chilo suppressalis</i>	MK207057
			<i>Diatraea saccharalis</i>	FJ240227
		Pyraustinae	<i>Ostrinia furnacalis</i>	MN793323
			<i>Ostrinia nubilalis</i>	AF442957
Geometroidea	Geometridae	Ennominae	<i>Biston panterinaria</i>	JX406146
			<i>Phthonandria atrilineata</i>	EU569764
Bombycoidea	Sphingidae	Sphinginae	<i>Manduca sexta</i>	AY616435
	Saturniidae	Saturniinae	<i>Actias selene</i>	JX186589
			<i>Antheraea pernyi</i>	MT890592
	Bombycidae	Bombycinae	<i>Bombyx mori</i>	AF149768
			<i>Bombyx mandarina</i>	MK251840
Lasiocampoidea	Lasiocampidae		<i>Kunugia undans</i>	KX822016
		Pinarinae	<i>Trabala vishnou guttata</i>	KU884483
Outgroups			<i>Papilio polytes</i>	KM014701
			<i>Trogonoptera brookiana</i>	LT999986

Table 1. List of the complete mitogenome of the superfamily Noctuoidea reported so far. * indicates the species that are used in this study.

UGC (Cys) were the smallest used codons (<10). We found average relative synonymous codons of 3737 (*A. polyodon*), 3736 (*O. seranensis*), 3727 (*O. emarginata*) and 3731 (*H. puera*), not including stop codons, that were predicted for codon usage of all the four mitogenomes.

Overlapping and intergenic spacer regions. The intergenic sequences and overlapping regions were analyzed. The small intergenic spacers (IGS) were seen ranging in size 1–63 bp, and totalling 174 bp in *A. polyodon*, 256 bp in *O. seranensis*, 527 bp in *O. emarginata* and 262 bp in *H. puera* (Table 2). The longest intergenic

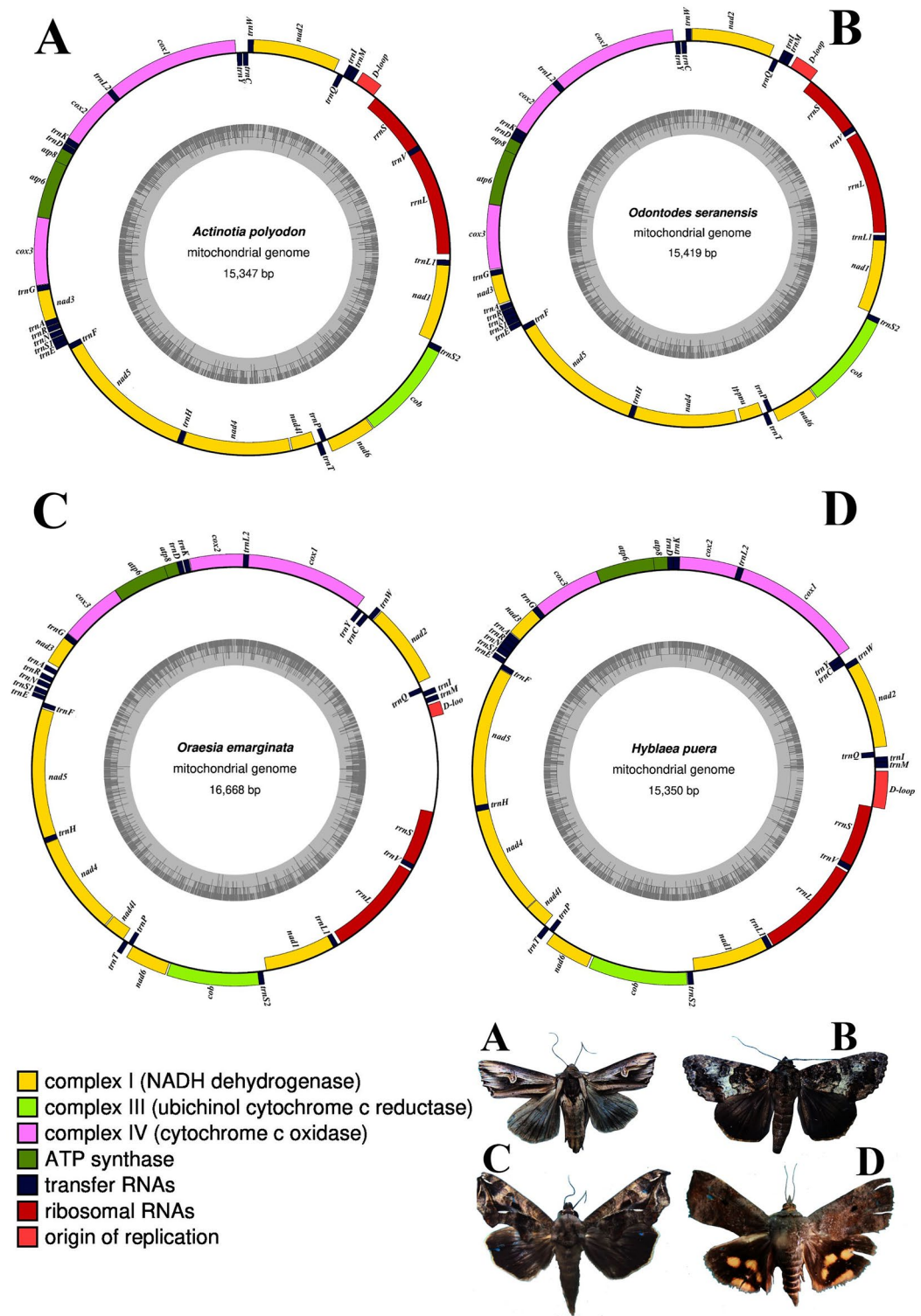


Figure 1. Circular maps of the newly sequenced complete mitochondrial genomes of (A) *Actinotia polyodon*, (B) *Odontodes seranensis*, (C) *Oraesia emarginata*, (D) *Hyblaea puera*.

spacer was located between *tRNA^{Gln}* and *nad2* with the length of 48 bp in *H. puera*, 56 bp in *A. polyodon*, 57 bp in *O. seranensis*, and 63 bp in *O. emarginata*.

In addition, the overlapping regions were also analyzed. The numbers of overlapping regions in four Noctuid moths were mostly inconstant from 1 to 45 bp. However, the longest overlapping region was located between *rrnL* and *rRNA^{Val}*, with a length of 45 bp in *A. polyodon* (Table 2).

Gene/strand	Position from	To	Anticodon	Start codon	Stop codon	Intergenic nucleotides
trnM/J	888/2512/2593/35	954/2580/2660/102	CAT			24/0/0/0
trnI/J	979/2581/2661/103	1046/2645/2727/167	GAT			-3/3/-3/-3
trnQ/N	1044/2643/2725/165	1112/2711/2793/233	TTG			63/56/57/48
nad2/J	1176/2768/2851/282	2189/3781/3864/1295		ATT/ATT/ATT/ATT	TAA/TAA/TAA/TAA	11/0/4
trnW/J	2201/3780/3869/1295	2270/3847/3938/1360	TCA			-8/-7/-8/-8
trnC/N	2263/3840/3931/1353	2334/3904/3997/1417	GCA			34/11/13/0
trnY/N	2369/3916/4011/1418	2433/3981/4077/1484	GTA			27/8/15/6
cox1/J	2461/3990/4093/1491	3996/5528/5626/3022		CGA/CGA/TTG/CGA	TAA/TAA/T/T	-5/-4/0/-1
trnL2/J	3992/5524/5627/3022	4058/5590/5693/3088	TAA			0/0/0/0
cox2/J	4059/5591/5694/3089	4734/6272/6375/3773		ATG/ATT/ATG/ATG	T/T/T/T	6/0/0/-2
trnK/J	4741/6273/6376/3771	4811/6343/6446/3841	CTT			12/3/-1/-1
trnD/J	4824/6347/6446/3841	4890/6414/6515/3907	GTC			0/0/0/0
atp8/J	4891/6415/6516/3908	5055/6576/6677/4072		ATA/ATT/ATT/ATT	TAA/TAA/TAA/TAA	-7/-7/-7/-7
atp6/J	5049/6570/6671/4066	5726/7247/7348/4740		ATG/ATG/ATG/ATG	TAA/TAA/TAA/TAA	-1/0/-1/-1
cox3/J	5726/7247/7348/4740	6514/8035/8136/5528		ATG/ATG/ATG/ATG	TAA/TAA/TAA/TAA	3/2/10/2
rnG/J	6518/8038/8147/5531	6584/8012/8212/5596	TCC			0/0/0/0
nad3/J	6585/8103/8213/5597	6938/8456/8566/5950		ATT/ATT/ATT/ATT	TAA/TAA/TAA/TAA	41/0/16/6
trnA/J	6980/8456/8583/5957	7045/8522/8649/6022	TGC			44/0/0/-1
trnR/J	7090/8526/8650/6022	7156/8591/8714/6087	TCG			21/11/-1/0
trnN/J	7178/8603/8714/6088	7243/8668/8779/6154	GTT			30/5/1/0
trnS1/J	7274/8674/8781/6155	7339/8739/8846/6220	GCT			9/0/3/19
trnE/J	7349/8740/8850/6240	7414/8806/8915/6306	TTC			15/2/4/-1
trnF/N	7430/8809/8920/6305	7495/8876/8986/6371	GAA			18/0/1/5
nad5/N	7514/8877/8988/6377	9244/10,622/10,733/8103		ATA/ATG/ATA/ATT	TTA/TAA/TAA/T	0/0/0/0
trnH/N	9245/10,623/10,734/8104	9311/10,689/10,799/8170	GTG			3/3/3/0
nad4/N	9315/10,693/10,803/8171	10,650/12,028/12,138/9511		ATG/ATG/ATG/ATG	T/T/T/TAA	15/20/51/6
nad4l/N	10,666/12,049/12,190/9518	10,953/12,339/12,477/9811		ATG/ATG/ATG/ATG	TAA/TAA/TAA/TAA	16/11/2/9
trnT/J	10,970/12,351/12,480/9821	11,035/12,416/12,544/9885	TGT			0/0/0/0
trnP/N	11,036/12,417/12,545/9886	11,100/12,481/12,610/9950	TGG			7/7/5/11
nad6/J	11,108/12,489/12,616/9962	11,631/13,022/13,149/10,495		ATA/ATG/ATT/ATG	T/TAA/TAA/TAA	18/11/13/24
cob/J	111,650/13,034/13,163/10,520	12,801/14,182/14,314/11,668		ATG/ATG/ATG/ATG	TAA/TAA/TAA/TAA	-1/2/-1/6
trnS2/J	12,801/14,185/14,314/11,675	12,868/14,250/14,380/11,742	TGA			44/18/21/16
nad1/N	12,913/14,269/14,402/11,759	13,851/15,207/15,340/12,694		ATT/ATG/ATG/ATG	TAA/ATT/TAA/TAG	0/1/0/1
trnL1/N	13,852/15,209/15,341/12,696	13,919/15,276/15,408/12,763	TAG			44/0/0/17
rrnL/N	13,964/1/20/12,781	15,280/1371/1356/14,056				32/-45/34/26
trnV/N	15,313/1327/1391/14,083	15,377/1393/1455/14,151	TAC			-1/0/0/0/-1
rrnS/N	15,377/1394/1456/14,152	16,114/2170/2236/14,913				0/0/0/-1
control region	601/2270/2249/14,912	862/2465/2592/15,350				

Table 2. Details on gene organization of four newly determined Lepidoptera mitogenomes. The order of the four species in the table is as follows: *Oraesia emarginata*/*Actinotia polyodon*/*Odontodes seranensis*/*Hyblaia puera*. J and N refer to the majority and minority strand respectively.

Transfer RNA genes (tRNA). The mitogenomes of *H. puera*, *A. polyodon*, *O. seranensis*, *O. emarginata* had 22 tRNA genes (Table 2). The total lengths of the 22 tRNA genes were 1479 bp (*O. emarginata*), 1472 bp (*A. polyodon*), 1476 bp (*O. seranensis*) and 1471 bp (*H. puera*); however individual tRNA genes typically ranged from 65 to 71 bp, among which, eight tRNAs were encoded on the N-strand and the remaining 14 were encoded on the J-strand. The putative secondary structures of tRNA genes recognized in these Noctuid mitogenomes are given in Supplementary Figs. 1–4. All the predicted tRNAs revealed the typical putative secondary structure except for *trnS1* (AGN) where dihydrouridine (DHU) arm lacked and formed a simple loop which has been found in many lepidopterans^{22,30,39,40}. The lack of dihydrouridine (DHU) arm in *trnS1* (AGN) was observed in all species, while TΨC arm disappeared only in *trnE* of *O. emarginata*. In addition, TΨC loop was seen lacking in *trnY* and *trnF* of *O. emarginata*, and *trnI*, *trnT*, *trnF* of *H. puera*. Several mismatching base pairs occurred in tRNA clover-leaf secondary structures in all four lepidopteran mitogenomes. A total of 16 mismatches of 9 U-G and 7 G-U wobble pairs were noticed in 13 tRNA genes of *A. polyodon*; 18 mismatches (1 U-U and 9 U-G) and 8 G-U wobble pairs were detected in 15 tRNA genes of *O. seranensis*; 20 mismatches (1 A-A, 3 U-U and 10 U-G) and 6 G-U wobble pairs were observed in 13 tRNA genes of *H. puera*; 18 mismatches (1 U-U and 10 U-G) and 7 G-U wobble pairs were observed in 14 tRNA genes of *O. emarginata*.

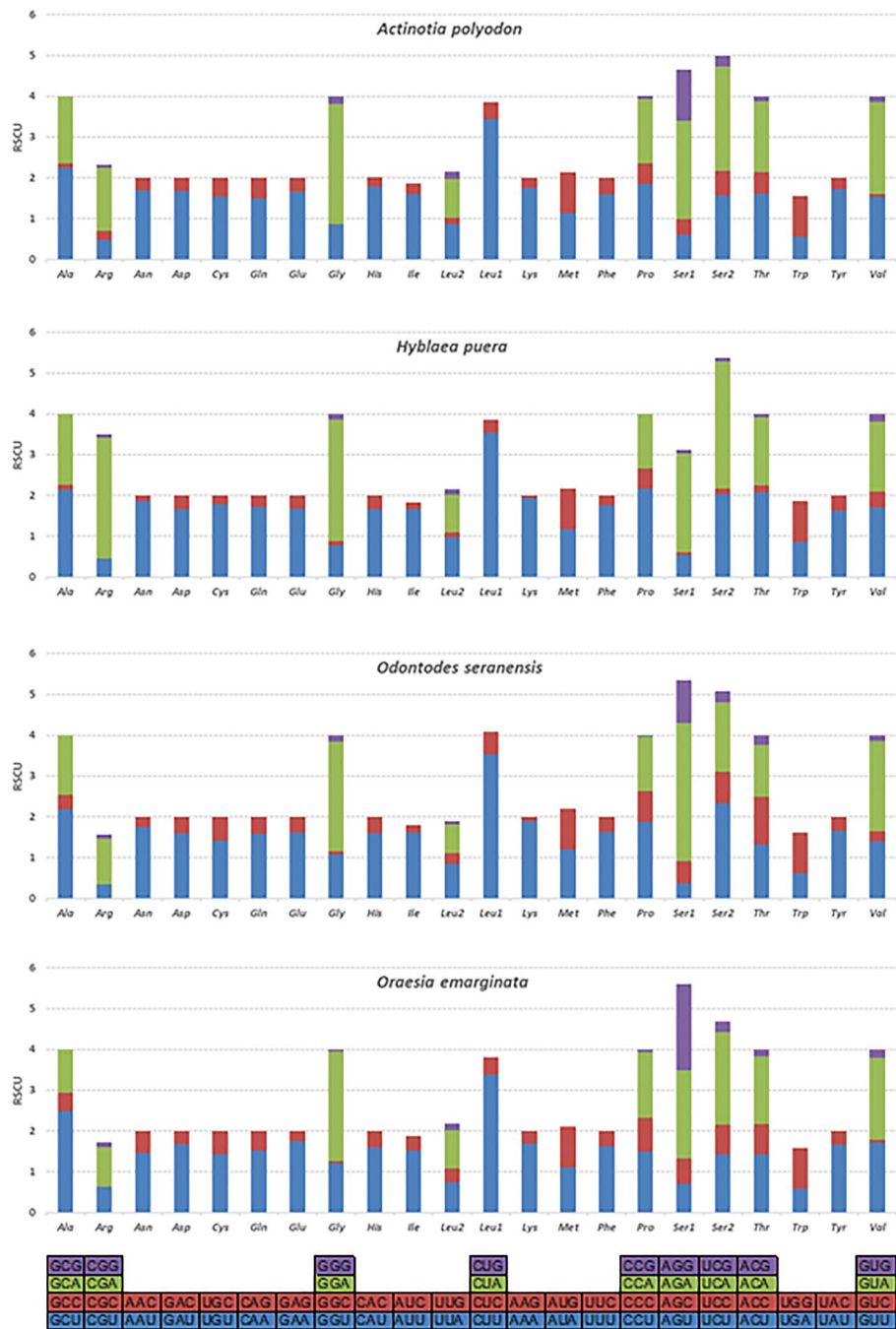


Figure 2. Relative Synonymous Codon Usage (RSCU) in Noctuoidea PCGs. The species name represents the superfamilies Noctuoidea & Hyblaeoidea. RSCU of *Actinotia polyodon*, RSCU of *Odontodes seranensis*, RSCU of *Oraesia emarginata* and RSCU of *Hyblaea puera*.

Ribosomal RNA genes. Two rRNA genes (*rrnL* and *rrnS*) are extremely conserved in Noctuid mitogenomes, and each of the four mitogenomes contained these two rRNA genes. *rrnL* gene lengths were 1370 bp for *A. polyodon*, 1336 bp for *O. seranensis*, 1276 bp for *H. puera* and 1317 bp for *O. emarginata*; whereas *rrnS* were 777, 781, 762, and 738 bp (Supplementary Table 1). The rRNA genes of currently sequenced mitogenomes displayed a negative AT skew (−0.007 to −0.036) and GC skew (−0.356 to −0.483). *rrnL* gene was located between *trnL1* and *trnV*, and *rrnS* was located between *trnV* and the control region (Table 2).

The A+T-rich region. The A+T-rich regions of *O. emarginata*, *A. polyodon*, *O. seranensis*, and *H. puera* were 287, 259, 343 and 439 bp in size respectively, all positioned between the *rrnS* and *tRNA^{Met}* (Table 2). The A+T content of these regions was 93.38%, 94.72%, 95.35% and 97.25%, respectively (Supplementary Table 1). The A+T-rich regions exhibited negative AT and GC-skew values. The conserved structure that connected the

<i>Odontodes seranensis</i> *	TTTATAAA	ATAGA	TTTTTTTTTTTTTTTTTATATT
<i>Hyblaea puera</i> *	TATTATAAAAAA	GTTACATAGA	TTTTTTTTTTTTTAAATTA AAAA
<i>Actinotia polyodon</i> *	ATACAAATTA	TTTTGAATAGA	TTTTTTTTTTTTTTTTTTA
<i>Spodoptera litura</i>	ATATGTAAATTTT	TAAATAGA	TTTTTTTTTTTTTTTTTATAT TAA
<i>Grammodes geometrica</i>	TATGTACAAATTTT	TCATATAGA	TTTTTTTTTTTTTTTTTATATT
<i>Helicoverpa armigera</i>	ATTTTTTAA	ATAGA	TTTTTTTTTTTTTTTTTATATT
<i>Catocala sp.</i>	TTTTACACA	TAGA	TTTTTTTTTTTTTTTTTATATT
<i>Eutelia adulatricoides</i>	TTTTTAA	ATAGA	TTTTTTTTTTTTTTTTTATAT TAA
<i>Amata formosae</i>	CTATTTTTTA	TAGA	TTTTTTTTTTTTTTTTTATAT TT
<i>Hydrillodes lentalis</i>	TAAATATAA	TTTTTAATAGA	TTTTTTTTTTTTTTTTTATAT TA
<i>Mythimna separata</i>	TATGTAAATTTT	TAAATAGA	TTTTTTTTTTTTTTTTTATAT TAA
<i>Asoata plana</i>	ATATACAAAATTTT	TAAATAGA	TTTTTTTTTTTTTTTTTATAT TA
<i>Protegira songi</i>	TATATGCAAATTA	TTTTGAATAGA	TTTTTTTTTTTTTTTTTATAT TAA
<i>Gabala argentata</i>	TATATATATTA	TTTTTCATAGA	TTTTTTTTTTTTTTTTTATAT TAAA
<i>Euproctis pseudoconsersa</i>	TTTTTTAA	ATAGA	TTTTTTTTTTTTTTTTTATAT TTA
<i>Vamuna virilis</i>	ATTTTTCA	TAGA	TTTTTTTTTTTTTTTTTATAT T TAA

Figure 3. Alignment of initiation site for A + T-rich region of 16 species completely sequenced lepidopteran mitogenomes. The boxed nucleotides indicate the conserved motif ATAGA and the shaded nucleotides indicate poly-T stretch. * newly sequenced mitogenomes presented in this study.

motif “ATAGA + poly-T stretch” was located downstream of the *rrnS* gene in the A + T-rich region of *H. puera*, *A. polyodon* and *O. seranensis*, which was not observed in *O. emarginata* (Fig. 3). We found that the motif ‘ATAGA’, might be the origin of light-strand replication⁵¹, directly connecting to the poly-T stretch in *A. polyodon* instead of (A)_n which connected to the poly-T structure in *H. puera* and *O. seranensis*. Multiple tandem repeats are naturally existing in the A + T-rich region of most lepidopteran insects. We detected the presence of tandem repeats in the mitochondrial A + T-rich region in *A. polyodon*, *O. seranensis* and *H. puera*, but not in *O. emarginata*. The A + T-rich region of *H. puera* consisted of three tandem repeats each of size 128 bp, 113 bp and 108 bp. Only one tandem repeat (103 bp) was found in the A + T-rich region of *A. polyodon*. In *O. seranensis*, the A + T-rich region consisted of two tandem repeats (57 bp and 47 bp). However, the tandem repeat was not observed in the A + T-rich region of *O. emarginata*; similarly, the tandem was also not present in *Dysgonia stuposa*³².

Additionally, two dinucleotide microsatellites and three motifs were detected in *H. puera*, referred to as (TA)₁₀, (TA)₃ in repeat 5, motif (ATAGA)₂, (ATTTA)₁₆, and TAATAATA. In *A. polyodon*, one dinucleotide microsatellite (TA)₇ and three motifs (ATTTA)₅, (ATAGA)₁, and TAATAATA were also observed. One dinucleotide (TA)₃, one trinucleotide (TAATAATA)₂ microsatellites, and three motifs (ATAGA)₂, (ATTTA)₄ and (ATATTA)₁₀ were found in *O. seranensis*. Similarly, one dinucleotide microsatellite (TA)₁₀ and two motifs (ATTTA)₄ and (ATATTA)₃ were found in *O. emarginata*. Furthermore, the ‘ATCTAA’ block in *H. puera* upstream of the origin of light-strand replication was different from the ‘ATACAA’ block in *A. polyodon* (Fig. 4).

Phylogenetic relationships. We performed the phylogenetic study on the mitogenomes of 99 lepidopteran species representing five Noctuid families (Erebidae, Euteliidae, Noctuidae, Nolidae, Notodontidae), one Hyblaecid (Hyblaecidae), one Pyraloid (Crambidae), one Geometroid (Geometridae), three Bombycoïd (Sphingidae, Saturniidae, Bombycidae), and one Lasiocampoid (Lasiocampidae) with two outgroup species (*Papilio polytes* and *Trogonoptera brookiana*) using the Maximum likelihood (ML) method and Bayesian inference (BI). The analyses were conducted on the dataset 13 PCGs + two rRNAs of the mitochondrial genomes which acquired similar tree topology (Figs. 5 and 6). We obtained the concatenated amino acid sequences to reconstruct the phylogenetic relationships (Figs. 5 and 6). The topology of the families based on mitogenomes in this study was consistent with the previous morphological and molecular studies^{11,12,32,52–54}.

The phylogenetic trees consisted of 6 clades corresponding to 12 major lepidopteran families (Figs. 5 and 6). The family Erebidae formed a major clade including 43 species with high bootstrap proportion and posterior probability (BP ≥ 100; PP: 1). This clade is further comprised of two subclades; subclade I with strongly supported values (BP ≥ 100; PP: 1) involving five subfamilies; Arctiinae Erebinae, Calpinae, Herminiinae, and Aganainae. The type species *Oraesia emarginata* belonging to the tribe Calpini was closely related to the *Catocala sp.*, *Grammodes geometrica*, and *Parallelia stuposa* with well-supported values (BP ≥ 100; PP: 1); subclade II comprised of 22 species belonging to the subfamilies Lymantriinae and Hypeninae with high nodal support (BP ≥ 96; PP: 1), out of which a single species *Paragabara curvicornuta* belonging to the subfamily Hypeninae was clustered separately in the same clade. The subfamily Arctiinae is closely related to the subfamily Erebinae rather than Lymantriinae. The Erebidae clade showed the relationship as; ((Aganainae + Calpinae + Herminiinae + Erebinae + (Arctiinae + (Hypeninae + Lymantriinae))).

The newly sequenced species *O. seranensis* and *Eutelia adulatricoides* clustered into a single clade with high bootstrap proportion (BP ≥ 100; PP: 1). These species belonged to the family Euteliidae and strongly supported a monophyletic group. This branch consisted of two subfamilies, Stictopterinae and Euteliinae. Formerly the Euteliinae and Stictopterinae were treated as separate subfamilies of Noctuidae^{55–57}. Later these two were placed

The A+T-rich regions

Odontodes seranensis

TTTTATAA **ATAGA**A **TTTTTTTTTTTTTTTTTT**ATATTA AAT **ATTTA** ATATTAATATTAATTTTAA
 The origin of minority strand replation (poly-T) Tandem repeats (47 bp) 43-89
 TTAATTTTA **ATATTA**ATTATTAACAATTAATAATTTCTTTTCTTCTTTATA **ATATTA**ATAT
 AA **ATAGA**AAA **ATTTA**AGAATTAAC AATAAAAATTCTTAAAAATAACTAT **ATATTAATATTA**A
 TA **ATATTA**ATTTTTCAATA **ATTTA**ATATATTAATATAAATTAATAAAATTAAT **ATTTA**ATATA
 Tandem repeats (57 bp) 215-271
 ATAATAATAATATAATATAATATTTTCTAA **ATATTA**TATTTTAGACCATTTTAAATAATTTTTCAT
 TAAATAAAAAAAAAAAAA

Actinotia polyodon

AAA **ATTTA**TTTAT **ATACAA**ATTATTTGA **ATAGA****TTTTTTTTTTTTTTTTTT**ATATTA AAAT **AT**
 The origin of minority strand replation (poly-T)
TTAATATAAATTATTAATTTTAAATAATTTCTTTTCTTTCTTTTATAATATTCTTATTA AAA
 ACTAAATTGGAAATTAACAATTA **ATATTA**ATAAAAATT **ATATTA**TATTAATATAATTAATTTTA
 Tandem repeats (103 bp) 170-272
ATTTATTGATTAATTAATG **TATATATATAT** **ATTTA**ATATATTA AATTTTAAATATATATATAAAA
 A micro satellite(TA)⁷
 TAATTAATTAATTATTTT **ATTTA**A **ATTTA**TTCTTAAACCGTTTTTAAATAAATTCATATAAAT
 AAATAAAAAAAAAAAAA

Hyblaea puera

TAAAAA **ATTTA**TTT **ATATTA**TAAAAAAGTTAC **ATAGA**A **TTTTTTTTTTTTTT**AAATTA AAAAAT
 The origin of minority strand replation (poly-T)
 AAAATA **ATATTA**TATAATACATAATAAATTAATA **ATTTA**TTAT **ATATTA**TATTATAAATTAATAA
 Tandem repeats (108 bp) 74-201
TTTATTTATTTA **ATATTA**ATAATAAATTAATA **ATTTA**TTGT **ATATTA**TATTA **ATTTA**AAATA **AT**
 Tandem repeats (113 bp) 89-201 Tandem repeats (128 bp) 100-207
TATTGATT **ATTTA**ATATAAATTATTAATTTTAAATAATCTTTCTTTTCTTTTTCATAAT **ATTT**
ATATAAAAATA **ATTTA**ATTATAATCA **ATTT** **ATATTA** **ATTTA**AAATAATA **ATTTA**TAAATATAATATA
 TAATTA **ATTTA**AAATAAATTAATAAATTTAT **ATTTA** **ATTTA**TATAA **ATTTA**AAAT **TATATATATAT**
 A micro satellite(TA)¹⁰
 ATATATATA **ATCTAAA** **ATAGAT**TATTAATTAATAAT **ATTTA**TTT

Oraesia emarginata

TTAAAAT **ATTTA**ATATAAATTATTAATTTTAAATAATTTCTTATTTTTTTTTTCATA **ATATTA**
TATTAAAAACCTAAATTGGAAATTAACAATATAAATTCGTA AAAAATAATAAAGAATTA **ATATT**
AATAATATTAATTTTCTTAATAAGTTAATAAATTTAATTTTAAATAAT **ATTTA**AAA **ATTTA**
TATATATATATATATATAATCTATTCAATTAATAATTTTCT **ATTTA**TAAAAATTTTACCGTTTTT
 A micro satellite(TA)¹⁰
 AATAAATTTTATTATAATAATAAAAA

Figure 4. Motifs and microsatellites found in the A + T-rich region of *Odontodes seranensis*, *Actinotia polyodon*, *Hyblaea puera* and *Oraesia emarginata*. These are indicated by specific colours and highlights. Motifs (ATAGA) are shown in dark blue high lights. Poly-T stretch are shown in darker gold accent highlights. Microsatellites (ATATTA) are shown in pink highlights. Microsatellites (ATTTA) are shown in green highlight. All tandem repeats are underlined. (ATACAA) block is shown in light blue highlight. Microsatellite (TA)¹⁰ and (TA)⁷ are shown in yellow colour.

as subfamilies of Erebidae⁵³. Afterwards, the position of subfamily Euteliinae was raised to the family level and Stictopterinae was placed into the subfamily of Euteliidae based on molecular study¹⁸. The present observation was well supported by the molecular study of Zahiri et al.¹⁸. The recently reconstituted family Nolidae with species *Gabala argentata*, *Sinna extrema* and *Risoba prominens* clustered in a single clade with high bootstrap proportion (BP ≥ 100; PP: 1) and observed more closely related to the clade (Euteliidae + Noctuidae), instead of Erebidae as proposed by Zahiri et al.⁵⁸.

The target species *A. polyodon* and thirty-one species belonging to the family Noctuidae were clustered into single branches with high nodal support values (BP ≥ 100; PP: 1). The species *A. polyodon* is clustered with Hadeninae clade with a high support value (BP ≥ 100; PP: 1). This phylogenetic analysis showed the main topology: ((Plusiinae + (Heliiothinae + (Amphipyrrinae + (Acronictinae + (Xyleninae + ((Hadeninae + Noctuidae)))))).

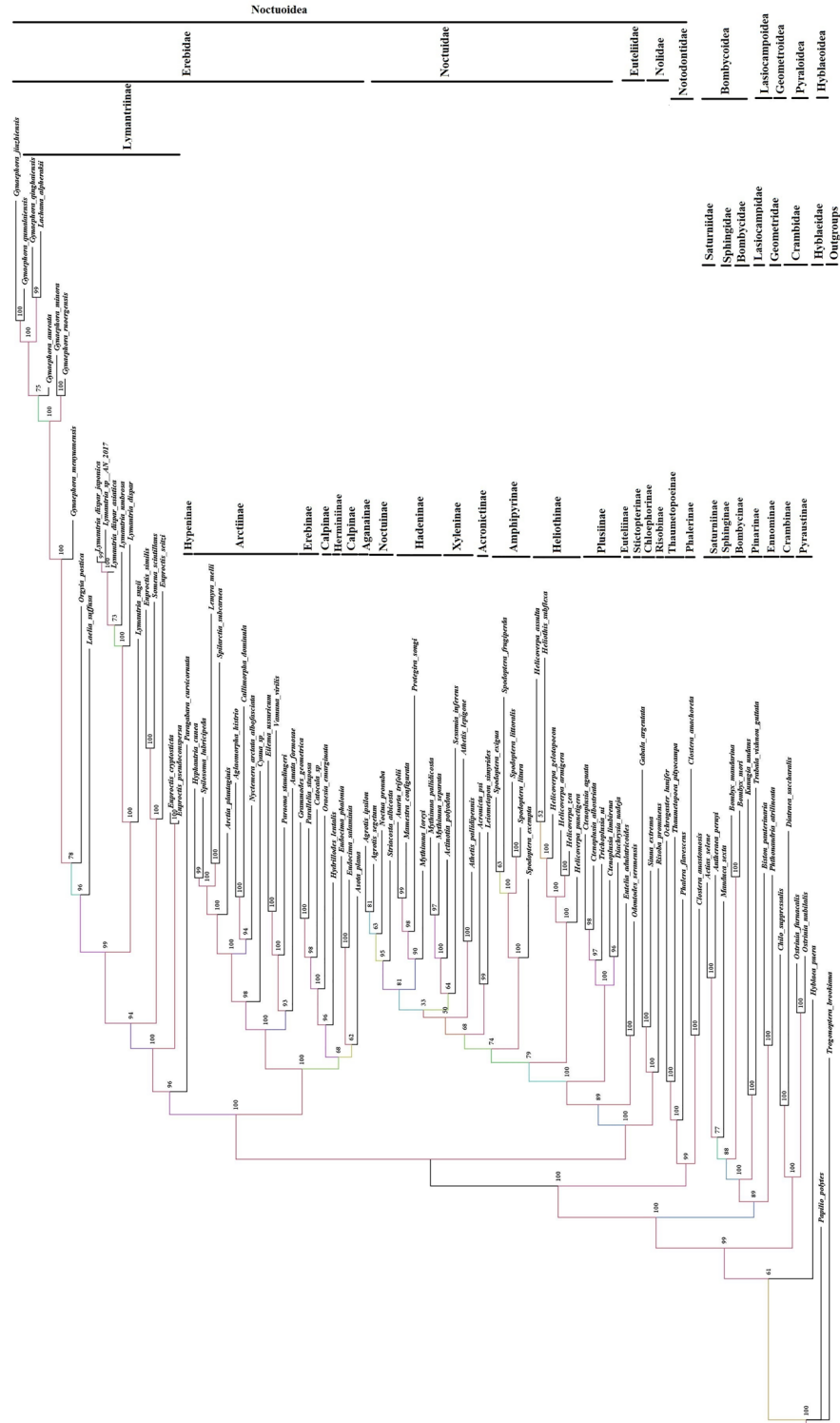


Figure 5. Phylogenetic tree of superfamily Noctuoidea moths using IQ-TREE. The phylogeny was reconstructed using 13 PCGs and two rRNA of the 90 species with maximum likelihood (ML) method (1000 replications). The species *Papilio polytes* and *Trigonoptera brookiana* mitogenomes were used as outgroups.

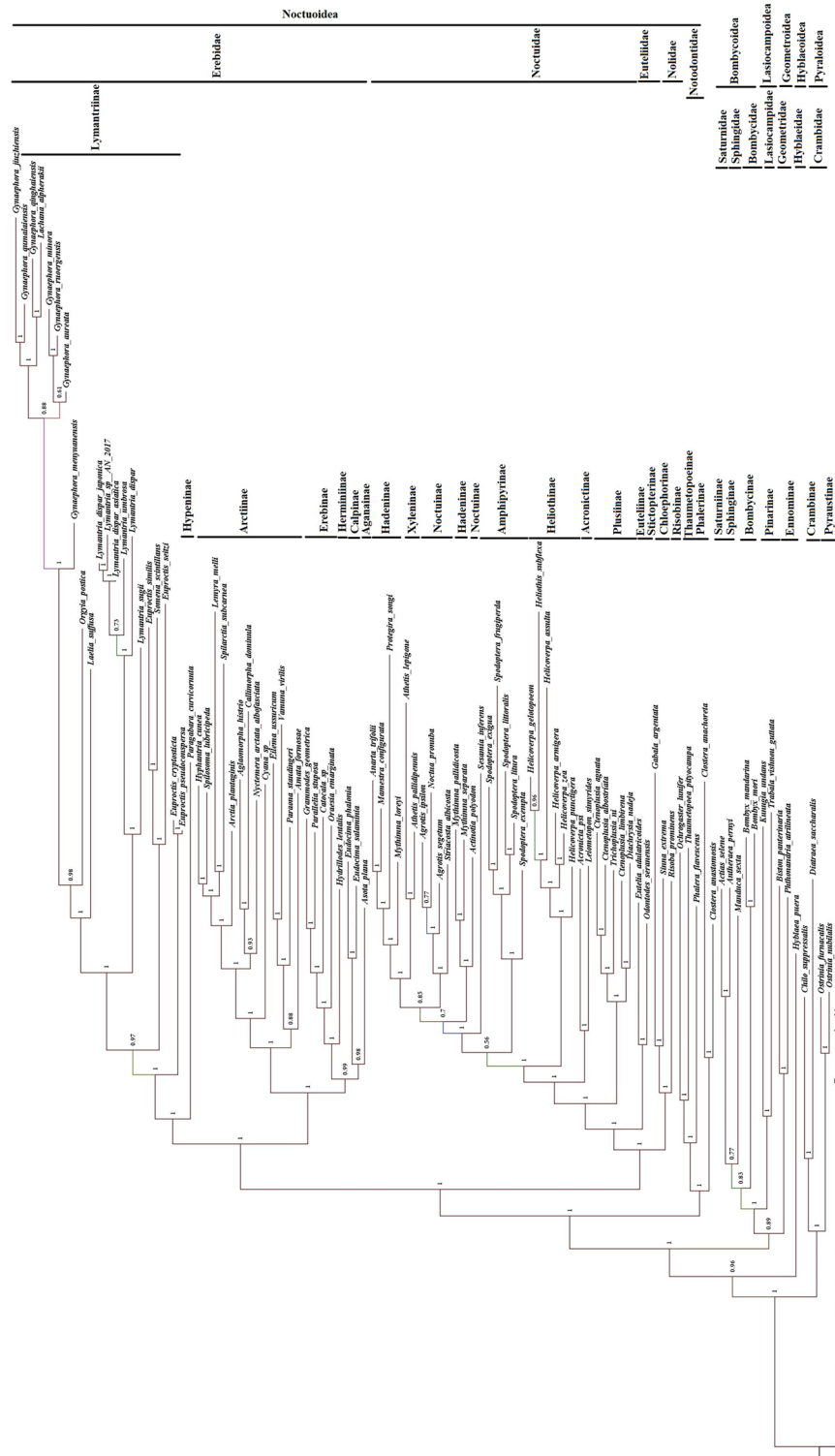


Figure 6. Phylogenetic tree of superfamily Noctuoidea moths using MrBayes. The phylogeny was reconstructed using 13 PCGs and two rRNA of the 90 species with Bayesian Inference. Posterior probability values lower than 50 were not shown.

Notodontiidae was strongly supported as a monophyletic group (BP \geq 99; PP: 1). The clade consisted of two subfamilies Phalerinae (*Phalera flavescens*), and Thaumetopoeinae (*Ochrogaster lunifer*, *Clostera anachoreta*, *Clostera anastomosis*, and *Thaumetopoea pityocampa*).

In the past decade, a number of studies have explored the molecular phylogenetic relationships among the Noctuoidea species. Zahiri et al.¹⁸ proposed the following among these families: (Notodontidae + (Euteliidae + (Noctuidae + Erebiidae + Nolidae)))). In comparison with this, Yang et al.¹² published different study in which the following assemblage was proposed: (Notodontidae + (Erebiidae + Nolidae + Euteliidae + Noctuidae))). All analyses clearly supported the monophyletic relationships of the 16 subfamilies within Noctuoidea (Figs. 5 and 6). The reformulated family Noctuidae clustered with the newly erected family Euteliidae. Our findings indicated that the branch of Noctuidae and Euteliidae was sister to the newly constituted family Nolidae. The family Erebiidae was sister to the clade of (Nolidae + (Euteliidae + Noctuidae)). Family Notodontidae members formed as a single clade consisting of subfamilies, Phalerinae and Thaumetopoeinae. Notodontidae was the sister group to the other Noctuid families. Our analysis revealed a topology within Noctuoidea as follows: (Notodontidae + (Erebiidae + (Nolidae + (Euteliidae + Noctuidae))))). The superfamily Noctuoidea relationships further confirmed that Noctuoidea was a monophyletic group, which was also supported by many previous mitogenome phylogenies^{12,32–34}.

In the present analyses, a total of 9 species were included belonging to the superfamilies Bombycoidea, Lasiocampoidea, and Geometroidea. The phylogenetic tree analyses showed that Saturniidae (*Actias selene* and *Antheraea pernyi*), Sphingidae (*Manduca sexta*), Bombycidae (*Bombyx mandarina* and *Bombyx mori*), Lasiocampidae (*Kunugia undans* and *Trabala vishnou guttata*) and Geometridae (*Biston panterinaria* and *Phtho-nandria atrilineata*) formed a clade with high nodal support values (BP \geq 89; PP: 0.89), this is consistent with earlier molecular study⁵⁹. The tree topologies indicate that the relationships are ((Geometridae + (Lasiocampidae + (Bombycidae + ((Sphingidae + Saturniidae)))))). The phylogenetic analyses also revealed the relationships in the superfamilies Bombycoidea, Geometroidea, Lasiocampoidea, and Noctuoidea with strongly supported values (BP \geq 100; PP: 1). This relationship is the resemblance to the novel Lepidoptera classification revised by van Nieukerken et al.¹¹ and the superfamilies are designated as the Macroheterocera clade.

Four species belonging to the family Crambidae formed a separate clade (BP \geq 99; PP: 1) which was placed sister to the family Hyblaeidae and both families belong to the clade Obtectomera¹¹. The present analysis is analogous to the molecular analysis by Twort et al.⁴⁴ which also showed that *Hyblaea* is sister to Pyraloidea. The newly sequenced species *H. puera* is separately clustered with moderate support in ML analysis and high support in BI analysis (BP \geq 61; PP: 0.96). This species was earlier classified under the family Noctuidae (Hampson, 1894). During the same year, the family Hyblaeidae was placed under the superfamily Pyraloidea based on the morphology characters⁶⁰. Afterward, it got its own superfamily rank Hyblaeoidea and was placed under the Obtectomera clade¹¹. In the present study, the species *H. puera* deviated from the Noctuoidea member groups. This mitogenome study is well supported by morphological¹¹ and molecular studies¹⁸. Unfortunately, the presence of only one mitogenome of Hyblaeidae restricted the discussion of its relationships; more species need to be added for a meaningful inference.

Conclusion

The complete mitochondrial genome sequences of *O. emarginata*, *O. seranensis*, *A. polyodon* and *H. puera* were successfully determined. The mitogenomes of these four moth species were all double-stranded single-circular molecules with similar gene arrangements (Fig. 1). The overall genomic characteristics (gene order, gene size, base composition, PCG codon usage, and tRNA cloverleaf structure) of the lepidopteran mitogenomes were typically constant with those of reported Lepidoptera mitogenomes. The longest intergenic spacer was present between *trnQ* and *nad2*; this was a unique feature in all sequenced species. Based on the phylogenetic analyses, the amino acid datasets supported the monophyly of Noctuoidea and its relationships (Notodontidae + (Erebiidae + (Nolidae + (Euteliidae + Noctuidae))). However, more mitochondrial genome samples need to be used to further resolve the relationships among the Noctuoidea.

Materials and methods

Sample collection and genomic DNA extraction. The samples of the four species, *O. emarginata* (11° 41' 181" N 76° 72' 07" E), *A. polyodon*, *O. seranensis* (10° 23' 5367" N 77° 49' 2933" E) and *H. puera* (10° 27' 045" N 77° 53' 3633" E) were collected from the Tamil Nadu part of Western Ghats. K. Sivasankaran identified all the species, which were preserved in absolute ethanol and stored at -80°C until DNA isolation. The genomic DNA was extracted from thorax tissue of moths using *Quick-DNA Tissue/Insect Microprep Kit* (Cat No-D6016-HSN CODE-38220090, Zymo Research, USA) with the manufacturer's protocol. The DNA samples and quality were checked using Nanodrop 1000 and confirmed with 1% agarose gel.

Mitogenome sequencing. The quality-check passed samples were subjected further for the library preparation. In brief, 100 ng of DNA was subjected to prepare indexed library using Truseq Nano library preparation kit (Illumina #20,015,964). Final libraries were quantified using Qubit 4.0 fluorometer (ThermoFisher #Q33238) using DNA HS assay kit (ThermoFisher #Q32851) following manufacturer's protocol. To identify the insert size of the library, we queried it on TapeStation 4150 (Agilent) utilizing highly sensitive D1000 screen tapes (Agilent # 5067–5582) following manufacturers' protocol. The next-generation sequencing was performed by Molsys Scientific Pvt. Ltd (Bangalore, India). Finally, NOVASEQ 6000 platform (Illumina, San Diego, California USA) was used to sequence 151 bp read lengths about 4 GB in size.

Sequence assembly and annotation. The raw sequences were assembled using the NOVOPLASTY Ver 4.2 (<https://github.com/ndierckx/NOVOPlasty>)⁶¹. The sequences' annotations were executed using MITOS2 (<http://mitos2.bioinf.uni-leipzig.de/index.py>)⁶² using the genetic code for invertebrate mitogenomes. The sequences were also annotated and verified for accurate lengths of the 13 protein-coding genes using CHLO-ROBOX-GeSeq-Annotation of Organellar Genomes (<https://chlorobox.mpimp-golm.mpg.de/geseq.html>)⁶³. The composition skewness was calculated using the formula: $AT\ skew = [A - T] / [A + T]$; similarly, $GCskew = [G - C] / [G + C]$ (<https://en.vectorbuilder.com/tool/gc-content-calculator>). The tRNA genes and their cloverleaf structures were predicted with MITOS2 software and analyzed by comparison with the nucleotide sequence of other lepidopteran tRNA sequences. Tandem repeats at the A + T-rich region were identified using the online Tandem Repeats Finder tool (<http://tandem.bu.edu/trf/trf.html>). Relative Synonymous Codon Usage (RSCU) of PCGs was determined using MEGA X⁶⁴. The circular maps of the four complete mitogenomes were drawn using the OGDRAW-Draw Organelle Genome Maps (<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>)⁶⁵.

Phylogenetic analyses. A total of 90 species (4 newly sequenced in this study, 86 available from GenBank) representing 7 families of Lepidoptera¹¹ were used to reconstruct the phylogenetic relationships among them. The ingroup consisted of 43 species of Erebiidae, 2 species of Euteliidae, 32 species of Noctuidae, 3 species of Nolidae, 5 species of Notodontidae, 1 species of Hyblaeidae, and 4 species of Crambidae. Species *Papilio polytes* and *Trogonoptera brookiana* mitogenomes were selected as outgroups (Table 1).

The amino acid sequences of 13 protein-coding genes and two rRNA genes were used in phylogenetic analysis. We used MAFFT to align and concatenate each of the 13 PCGs and rRNAs genes. Further the concatenated amino acid sequences from the 13 PCGs and rRNA genes were used for reconstructing the phylogenetic tree, which was performed using the Model-based Maximum Likelihood method using the IQ-TREE in PhyloSuite V1.2.2 program <https://github.com/dongzhang0725/PhyloSuite>⁶⁶. The appropriate model General Reversible mitochondrial (mtREV) Gamma distributed with invariant sites (G + I) was used to infer the phylogenetic relationships based on 5000 bootstraps of ultrafast replicates.

The analysis of Bayesian inference (BI) was conducted for the dataset. The BI analysis was performed through the MrBayes 3.2.6 in PhyloSuite V1.2.2⁶⁶ using the GTR + I + R model. Invgamma (+I + G proportion invariable, remaining gamma rate variations across sites were presented and performed. The convergence of Markov Chain Monte Carlo (MCMC), which was observed by the average standard deviation of split frequencies, reached below 0.01. Four chains (three hot and one cold) were run with a dataset for one million generations with the tree being sampled every 1000 generations with a burn-in of 2500. FigTree v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>) was practised to visualise the phylogenetic tree.

Received: 29 December 2021; Accepted: 28 September 2022

Published online: 07 November 2022

References

- Boore, J. L. The use of genome-level characters for phylogenetic reconstruction. *Trends Ecol. Evol.* **21**(8), 439–446. <https://doi.org/10.1016/j.tree.2006.05.009> (2006).
- Cameron, S. L. Insect mitochondrial genomics: Implications for evolution and phylogeny. *Annu. Rev. Entomol.* **59**(1), 95–117. <https://doi.org/10.1146/annurev-ento-011613-162007> (2014).
- Ma, Y. *et al.* The complete mitochondrial genomes of three bristletails (Insecta: Archaeognatha): The paraphyly of Machilidae and insights into Archaeognathan phylogeny. *PLoS ONE* **10**(1), e0117669. <https://doi.org/10.1371/journal.pone.0117669> (2015).
- Brown, W. M., George, M. & Wilson, A. C. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* **76**(4), 1967–1971. <https://doi.org/10.1073/pnas.76.4.1967> (1979).
- Boore, J. L. Animal mitochondrial genomes. *Nucleic Acids Res.* **27**(8), 1767–1780. <https://doi.org/10.1093/nar/27.8.1767> (1999).
- Curole, J. P. & Kocher, T. D. Mitogenomics: Digging deeper with complete mitochondrial genomes. *Trends Ecol. Evol.* **14**(10), 394–398. [https://doi.org/10.1016/S0169-5347\(99\)01660-2](https://doi.org/10.1016/S0169-5347(99)01660-2) (1999).
- Nardi, F., Carapelli, A., Dallai, R. & Frati, F. The mitochondrial genome of the olive fly Bactroceraoleae: Two haplotypes from distant geographical locations. *Insect Mol. Biol.* **12**(6), 605–611. <https://doi.org/10.1046/j.1365-2583.2003.00445.x> (2003).
- Shadel, G. S. & Clayton, D. A. Mitochondrial transcription initiation Variation and conservation. *J. Biol. Chem.* **268**(22), 16083–16086. [https://doi.org/10.1016/S0021-9258\(19\)85387-5](https://doi.org/10.1016/S0021-9258(19)85387-5) (1993).
- Jiang, S. T. *et al.* Characterization of the complete mitochondrial genome of the giant silkworm moth, *Eriogyna pyretorum* (Lepidoptera: Saturniidae). *Int. J. Biol. Sci.* **5**(4), 351–365. <https://doi.org/10.7150/ijbs.5.351> (2009).
- Wolstenholme, D. R. Animal mitochondrial DNA: Structure and evolution. *Int. Rev. Cytol.* **141**, 173–216. [https://doi.org/10.1016/S0074-7696\(08\)62066-5](https://doi.org/10.1016/S0074-7696(08)62066-5) (1992).
- Van Nieuwerkerken, E. J., Kaila, L., Kitching, I. J., Kristensen, N. P., Lees, D. C. & Minet, J. *et al.* Order Lepidoptera Linnaeus, 1758. In: *Zootaxa* Z.-Q. Zhang, Z.-Q. *et al.* (Ed.) *Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness*. *Zootaxa*, **3148**(1), 212–221. <https://doi.org/10.11646/zootaxa.3148.1.41> (2011).
- Yang, X., Cameron, S. L., Lees, D. C., Xue, D. & Han, H. A mitochondrial genome phylogeny of owlet moths (Lepidoptera: Noctuoidea), and examination of the utility of mitochondrial genomes for lepidopteran phylogenetics. *Mol. Phylogenet. Evol.* **85**, 230–237. <https://doi.org/10.1016/j.ympev.2015.02.005> (2015).
- Miller, J. S. Cladistics and classification of the Notodontidae (Lepidoptera, Noctuoidea) based on larval and adult morphology. *Bulletin of the AMNH*; no. 204. (1991)
- Weller, S. J., Pashley, D. P., Martin, J. A. & Constable, J. L. Phylogeny of Noctuid moths and the utility of combining independent nuclear and mitochondrial genes. *Syst. Biol.* **43**(2), 194–211. <https://doi.org/10.1093/sysbio/43.2.194> (1994).
- Mitchell, A. *et al.* Phylogenetic utility of elongation factor-1 alpha in Noctuoidea (Insecta: Lepidoptera): The limits of synonymous substitution. *Mol. Biol. Evol.* **14**(4), 38–390. <https://doi.org/10.1093/oxfordjournals.molbev.a025774> (1997).
- Fang, Q. Q. *et al.* Phylogenetic utility of the nuclear gene dopa decarboxylase in noctuid moths (Insecta: Lepidoptera: Noctuoidea). *Mol. Phylogenet. Evol.* **15**(3), 473–486. <https://doi.org/10.1006/mpev.1999.0758> (2000).

17. Mitchell, A., Mitter, C. & Regier, J. C. More taxa or more characters revisited: Combining data from nuclear protein-encoding genes for phylogenetic analyses of Noctuoidea (Insecta: Lepidoptera). *Syst. Biol.* **49**(2), 202–224. <https://doi.org/10.1093/sysbio/49.2.202> (2000).
18. Zahiri, R. *et al.* A new molecular phylogeny offers hope for a stable family level classification of the Noctuoidea (Lepidoptera). *Zool. Scr.* **40**(2), 158–173. <https://doi.org/10.1111/j.1463-6409.2010.00459.x> (2011).
19. Zahiri, R. *et al.* Molecular phylogenetics of Erebiidae (Lepidoptera, Noctuoidea). *Syst. Entomol.* **37**(1), 102–124 (2012).
20. Liao, F. *et al.* The complete mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). *Int. J. Biol. Sci.* **6**(2), 172–186. <https://doi.org/10.7150/ijbs.6.172> (2010).
21. Lu, H. F., Su, T. J., Luo, A. R., Zhu, C. D. & Wu, C. S. Characterization of the complete mitochondrial genome of diurnal moth *Amata emma* (Butler) (Lepidoptera: Erebiidae) and its phylogenetic implications. *PLoS ONE* **8**(9), e72410. <https://doi.org/10.1371/journal.pone.0072410> (2013).
22. Sun, Y. X. *et al.* Characterization of the complete mitochondrial genome of *Leucoma salicis* (Lepidoptera: Lymantriidae) and comparison with other lepidopteran insects. *Sci. Rep.* **6**, 39153. <https://doi.org/10.1038/srep39153> (2016).
23. Sun, Y. *et al.* Characterization of the complete mitochondrial genome of *Spilarctia robusta* (Lepidoptera: Noctuoidea: Erebiidae) and its phylogenetic implications. *Eur. J. Entomol.* **113**, 558–570 (2016).
24. Djoumad, A. *et al.* Comparative analysis of mitochondrial genomes of geographic variants of the gypsy moth, *Lymantria dispar*, reveals a previously undescribed genotypic entity. *Sci. Rep.* **7**(1), 14245. <https://doi.org/10.1038/s41598-017-14530-6> (2017).
25. Liu, Y. *et al.* The complete mitochondrial genome of *Euproctis similis* (Lepidoptera: Noctuoidea: Erebiidae) and phylogenetic analysis. *Int. J. Biol. Macromol.* **105**(1), 219–227. <https://doi.org/10.1016/j.ijbiomac.2017.07.033> (2017).
26. Sivasankaran, K. *et al.* Complete mitochondrial genome sequence of fruit-piercing moth *Eudocima phalonia* (Linnaeus, 1763) (Lepidoptera: Noctuoidea). *Genom. Data.* **14**, 66–81. <https://doi.org/10.1016/j.gdata.2017.09.004> (2017).
27. Chen, Z. S., Xu, Z., Ma, L. & Wang, X. The complete mitochondrial genome of *Paragabara curvicornuta* (Lepidoptera: Noctuoidea: Erebiidae). *Mitochondrial DNA B Resour.* **3**(1), 78–80. <https://doi.org/10.1080/23802359.2017.1422406> (2018).
28. Xin, Z. Z. *et al.* Comparative mitochondrial genome analysis of *Spilarctia subcarnea* and other noctuid insects. *Int. J. Biol. Macromol.* **107**(A), 121–128. <https://doi.org/10.1016/j.ijbiomac.2017.08.153> (2018).
29. Yang, M. *et al.* The complete mitochondrial genome of a medicinal insect, *Hydrillodes repugnalis* (Lepidoptera: Noctuoidea: Erebiidae), and related phylogenetic analysis. *Int. J. Biol. Macromol.* **123**, 485–493. <https://doi.org/10.1016/j.ijbiomac.2018.10.149> (2019).
30. Huang, Y. *et al.* Comparative mitochondrial genome analysis of *Grammodes geometrica* and other noctuid insects reveals conserved mitochondrial genome organization and phylogeny. *Int. J. Biol. Macromol.* **125**, 1257–1265. <https://doi.org/10.1016/j.ijbiomac.2018.09.104> (2019).
31. Li, J. *et al.* Mitochondrial genome characteristics of *Somena scintillans* (Lepidoptera: Erebiidae) and comparison with other Noctuoidea insects. *Genomics* **111**(6), 1239–1248. <https://doi.org/10.1016/j.ygeno.2018.08.003> (2019).
32. Sun, Y. *et al.* The complete mitochondrial genome of *Dysgonia stuposa* (Lepidoptera: Erebiidae) and phylogenetic relationships within Noctuoidea. *PeerJ* **8**, e8780. <https://doi.org/10.7717/peerj.8780> (2020).
33. Riyaz, M., Shah, R. A., Savarimuthu, I. & Kuppusamy, S. Comparative mitochondrial genome analysis of *Eudocima salamina* (Cramer, 1777) (Lepidoptera: Noctuoidea), novel gene rearrangement and phylogenetic relationship within the superfamily Noctuoidea. *Mol. Biol. Rep.* **48**(5), 4449–4463. <https://doi.org/10.1007/s11033-021-06465-z> (2021).
34. Sun, Y. *et al.* Organization and phylogenetic relationships of the mitochondrial genomes of *Speiredonia retorta* and other lepidopteran insects. *Sci. Rep.* **11**(1), 2957. <https://doi.org/10.1038/s41598-021-82561-1> (2021).
35. Chai, H. N. & Du, Y. Z. The complete mitochondrial genome of the pink stem borer, *Sesamia inferens*, in comparison with four other Noctuid moths. *Int. J. Mol. Sci.* **13**(8), 10236–56. <https://doi.org/10.3390/ijms130810236> (2012).
36. Wu, Q. L., Gong, Y. J., Gu, Y. & Wei, S. J. The complete mitochondrial genome of the beet armyworm *Spodoptera exigua* (Hübner) (Lepidoptera: Noctuidae). *Mitochondrial DNA* **24**(1), 31–33. <https://doi.org/10.3109/19401736.2012.716052> (2013).
37. Yin, J., Hong, G. Y., Wang, A. M., Cao, Y. Z. & Wei, Z. J. Mitochondrial genome of the cotton bollworm *Helicoverpa armigera* (Lepidoptera: Noctuidae) and comparison with other Lepidopterans. *Mitochondrial DNA* **21**(5), 160–169. <https://doi.org/10.3109/19401736.2010.503242> (2010).
38. Cao, S. S. & Du, Y. Z. Characterization of the complete mitochondrial genome of *Chilo auricilius* and comparison with three other rice stem borers. *Gene* **548**(2), 270–276. <https://doi.org/10.1016/j.gene.2014.07.045> (2014).
39. Liu, Q. N. *et al.* Characterization of the complete mitochondrial genome of the Oriental armyworm, *Mythimna separata* (Lepidoptera: Noctuidae). *Eur. J. Entomol.* **112**(3), 399–408 (2015).
40. Liu, Q. N. *et al.* The complete mitochondrial genome of fall armyworm *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *Genes Genomics* **38**(2), 205–216. <https://doi.org/10.1007/s13258-015-0346-6> (2016).
41. Liu, Q. N. *et al.* The complete mitochondrial genome of the common cutworm, *Spodoptera litura* (Lepidoptera: Noctuidade). *Mitochondrial DNA Part A DNA Mapping, Seq. Anal.* **27**(1), 122–123 (2016).
42. Perera, O. P., Walsh, T. K. & Luttrell, R. G. Complete Mitochondrial genome of *Helicoverpa zea* (Lepidoptera: Noctuidae) and expression profiles of mitochondrial-encoded genes in early and late embryos. *J. Insect Sci.* **16**(1), 40. <https://doi.org/10.1093/jisesa/iew023> (2016).
43. Hampson, G. F. *Fauna of British India including Ceylon and Burma (Moths)* Vol. 2 (Taylor & Francis, United Kingdom, 1894).
44. Twort, V., Minet, J., Wheat, C. & Wahlberg, N. Museomics of a rare taxon: Placing Whalleyanidae in the Lepidoptera tree of life. *Syst. Entomol.* **46**, 926–937. <https://doi.org/10.1101/2020.08.18.255182> (2021).
45. Salvato, P., Simonato, M., Battisti, A. & Negrisol, E. The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae). *BMC Genomics* **9**, 331. <https://doi.org/10.1186/1471-2164-9-331> (2008).
46. Sun, Q. Q. *et al.* Complete sequence of the mitochondrial genome of the Japanese buff-tip moth, *Phalera flavescens* (Lepidoptera: Notodontidae). *Genet. Mol. Res.* **11**(4), 4213–4225. <https://doi.org/10.4238/2012.September.10.2> (2012).
47. Wu, Q. L., Cui, W. X., Du, B. Z., Gu, Y. & Wei, S. J. The complete mitogenome of the turnip moth *Agrotis segetum* (Lepidoptera: Noctuidae). *Mitochondrial DNA* **25**(5), 345–347. <https://doi.org/10.3109/19401736.2013.800496> (2014).
48. Wu, Q. L., Cui, W. X. & Wei, S. J. Characterization of the complete mitochondrial genome of the black cutworm *Agrotis ipsilon* (Lepidoptera: Noctuidae). *Mitochondrial DNA* **26**(1), 139–140. <https://doi.org/10.3109/19401736.2013.815175> (2015).
49. Zhu, X. Y. *et al.* The complete mitochondrial genome of *Clostera anastomosis* (Lepidoptera: Notodontidae) and implication for the phylogenetic relationships of Noctuoidea species. *Int. J. Biol. Macromol.* **118**(1), 1574–1583 (2018).
50. Yang, Z. H. *et al.* The complete mitochondrial genome of *Sinna extrema* (Lepidoptera: Nolidae) and its implications for the phylogenetic relationships of Noctuoidea species. *Int. J. Biol. Macromol.* **137**, 317–326. <https://doi.org/10.1016/j.ijbiomac.2019.06.238> (2019).
51. Zhang, D. X. & Hewitt, G. M. Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. *Biochem. Syst. Ecol.* **25**(2), 99–120. [https://doi.org/10.1016/S0305-1978\(96\)00042-7](https://doi.org/10.1016/S0305-1978(96)00042-7) (1997).
52. Fibiger, M. & Lafontaine, J. D. A review of the higher classification of the Noctuoidea (Lepidoptera) with special reference to the Holarctic fauna. *Esperiana* **11**, 7–92 (2005).
53. Lafontaine, J. D. & Fibiger, M. Revised higher classification of the Noctuoidea (Lepidoptera). *Can. Entomol.* **138**(5), 610–635. <https://doi.org/10.4039/n06-012> (2006).

54. Mitchell, A., Mitter, C. & Regier, J. C. Systematics and evolution of the cutworm moths (Lepidoptera: Noctuidae): Evidence from two protein-coding nuclear genes. *Syst. Entomol.* **31**(1), 21–46. <https://doi.org/10.1111/j.1365-3113.2005.00306.x> (2006).
55. Holloway, J. D. *The moths of Borneo with special reference to Mont Kinabalu* (1–264). Malayan Nature Society and Sabah Foundation. (1976).
56. Kononenko, V. S. and A. Pinratana, A. An illustrated Catalogue of the Noctuidae (Insecta, Lepidoptera) in Thailand. *Moths of Thailand*, Vol. 3., Part. 1 (Subfamilies Herminiinae, Rivulinae, Hypeninae, Catocalinae, Aganainae, Euteliinae, Stictopterinae, Plusiinae, Pantheinae, Acronictinae and Agaristinae). *Moths of Thailand*, Vol. 3, Bangkok, 261. (2005).
57. Kononenko, V. S., Ahn, S. B. & Ronkay, L. Illustrated catalogue of Noctuidae in Korea (Lepidoptera). In: K. T. Park (eEd.), *Insects of Korea* (3), 1–509. (1998).
58. Zahiri, R. *et al.* Relationships among the basal lineages of Noctuidae (Lepidoptera, Noctuoidea) based on eight gene regions. *Zool. Scr.* **42**, 488–507 (2013).
59. Yang, X. S., Xue, D. Y. & Han, H. X. The complete mitochondrial genome of *Biston panterinaria* (Lepidoptera: Geometridae), with phylogenetic utility of mitochondrial genome in the Lepidoptera. *Gene* **515**, 349–358. <https://doi.org/10.1016/j.gene.2012.11.031> (2013).
60. Fletcher, D. S., & Nye, I. W. B. Generic Names of Moths of the World, vol. *Natural History*. Trustees of the British Museum, 5, xv. (1894).
61. Dierckxens, N., Mardulyn, P. & Smits, G. NOVOPlasty: *Denovo* assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **45**(4), e18. <https://doi.org/10.1093/nar/gkw955> (2017).
62. Bernt, M. *et al.* MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* **69**(2), 313–319. <https://doi.org/10.1016/j.ympev.2012.08.023> (2013).
63. Tillich, M. *et al.* GeSeq—versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **45**(1), 6–11 (2017).
64. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**(6), 1547–1549. <https://doi.org/10.1093/molbev/msy096> (2018).
65. Greiner, S., Lehwark, P. & Bock, R. Organellar genome DRAW (OGDRAW). Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research* version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* **47**(1), 59–64 (2019).
66. Zhang, D. *et al.* PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **20**(1), 348–355. <https://doi.org/10.1111/1755-0998.13096> (2020).

Acknowledgements

This study was funded by the Department of Science and Technology (DST)-Science and Engineering Research Board (SERB) [Grant Number: EMR/2017/000566]. The authors are grateful to the Director, Entomology Research Institute, Loyola College, Chennai, India for providing the facilities.

Author contributions

R.S. Experiment, data curation, and draft preparation. M.R. Experiment and data curation. S.I. Supervised, reviewed, and edited the manuscript. S.K. Conceptualization; Data curation, Investigation Software, Validation, Wrote the main manuscript text and prepared the figures.

Funding

The funding agency had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-21502-y>.

Correspondence and requests for materials should be addressed to K.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022