

Research article

Open Access

# The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae)

Paola Salvato<sup>†1</sup>, Mauro Simonato<sup>†1</sup>, Andrea Battisti<sup>1</sup> and Enrico Negrisolò<sup>\*2</sup>

Address: <sup>1</sup>Department of Environmental Agronomy and Vegetal Productions-Entomology, University of Padova, Agripolis, Viale dell'Università 16, 35020 Legnaro, Italy and <sup>2</sup>Department of Public Health, Comparative Pathology and Veterinary Hygiene, University of Padova, Agripolis, Viale dell'Università 16, 35020 Legnaro, Italy

Email: Paola Salvato - [paola.salvato@unipd.it](mailto:paola.salvato@unipd.it); Mauro Simonato - [mauro.simonato@unipd.it](mailto:mauro.simonato@unipd.it); Andrea Battisti - [andrea.battisti@unipd.it](mailto:andrea.battisti@unipd.it); Enrico Negrisolò\* - [enrico.negrisolò@unipd.it](mailto:enrico.negrisolò@unipd.it)

\* Corresponding author †Equal contributors

Published: 15 July 2008

Received: 21 May 2008

BMC Genomics 2008, 9:331 doi:10.1186/1471-2164-9-331

Accepted: 15 July 2008

This article is available from: <http://www.biomedcentral.com/1471-2164/9/331>

© 2008 Salvato et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Abstract

**Background:** Knowledge of animal mitochondrial genomes is very important to understand their molecular evolution as well as for phylogenetic and population genetic studies. The Lepidoptera encompasses more than 160,000 described species and is one of the largest insect orders. To date only nine lepidopteran mitochondrial DNAs have been fully and two others partly sequenced. Furthermore the taxon sampling is very scant. Thus advance of lepidopteran mitogenomics deeply requires new genomes derived from a broad taxon sampling. In present work we describe the mitochondrial genome of the moth *Ochrogaster lunifer*.

**Results:** The mitochondrial genome of *O. lunifer* is a circular molecule 15593 bp long. It includes the entire set of 37 genes usually present in animal mitochondrial genomes. It contains also 7 intergenic spacers. The gene order of the newly sequenced genome is that typical for Lepidoptera and differs from the insect ancestral type for the placement of *trnM*. The 77.84% A+T content of its  $\alpha$  strand is the lowest among known lepidopteran genomes. The mitochondrial genome of *O. lunifer* exhibits one of the most marked C-skew among available insect Pterygota genomes. The protein-coding genes have typical mitochondrial start codons except for *cox1* that present an unusual CGA. The *O. lunifer* genome exhibits the less biased synonymous codon usage among lepidopterans. Comparative genomics analysis study identified *atp6*, *cox1*, *cox2* as *cox3*, *cob*, *nad1*, *nad2*, *nad4*, and *nad5* as potential markers for population genetics/phylogenetics studies. A peculiar feature of *O. lunifer* mitochondrial genome it that the intergenic spacers are mostly made by repetitive sequences.

**Conclusion:** The mitochondrial genome of *O. lunifer* is the first representative of superfamily Noctuoidea that account for about 40% of all described Lepidoptera. New genome shares many features with other known lepidopteran genomes. It differs however for its low A+T content and marked C-skew. Compared to other lepidopteran genomes it is less biased in synonymous codon usage. Comparative evolutionary analysis of lepidopteran mitochondrial genomes allowed the identification of previously neglected coding genes as potential phylogenetic markers. Presence of repetitive elements in intergenic spacers of *O. lunifer* genome supports the role of DNA slippage as possible mechanism to produce spacers during replication.

## Background

Animal mitochondrial genomes (mtDNAs) are usually circular molecules spanning 16–20 kbp that contain 13 protein-coding genes (PCGs), 2 ribosomal RNA and 22 transfer (tRNA) genes [1]. Non-coding control elements, that regulate the transcription and replication of the genome, are also present in mtDNAs [1,2]. Mitochondrial genomes are very important subject for different scientific disciplines including animal health, comparative and evolutionary genomics, molecular evolution, phylogenetics and population genetics. However, current knowledge on mtDNAs is very uneven as well exemplified by sequences available in GenBank that were obtained mostly from vertebrate taxa. Insects constitute the most species-rich class among animals with almost a million of taxa described to date [3]. Within the insects, the Lepidoptera (butterflies plus moths) order accounts for more than 160,000 species [4]. Despite this huge taxonomic diversity the existing information on lepidopteran mtDNA is very limited. Complete sequences have been determined for the two butterflies *Coreana raphaelis* and *Artogeia melete*, and for the seven moths *Adoxophyes honmai*, *Antheraea pernyi*, *Bombyx mori*, *Bombyx mandarina* and *Manduca sexta*, *Phthorandria atrilineata* and *Saturnia boisduvalii* [5-9] while near complete sequences exist for *Ostrinia furnacalis* and *Ostrinia nubilalis* [10] (Table 1). Current genomic knowledge of Lepidoptera is very scanty and the covered taxon-sampling is extremely poor and limited to six superfamilies among the 45–48 known, and to 9 families of the recognized 120 [4]. A better understanding of the lepidopteran mtDNA requires an expansion of taxon and genome samplings. We were able to fully sequence the mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer*. The newly determined mtDNA is the first complete sequence for the Superfamily Noctuoidea, a very large assemblage that accounts for about 40% of all described Lepidoptera. [4]. In the present paper the *Ochrogaster* genome is described and compared with mtDNAs of other lepidopterans as well as pterygote Insecta.

## Results and discussion

### Genome organization, structure and composition

The mtDNA genome of *O. lunifer* is a circular molecule 15593 bp long. It includes the entire set of 37 genes usually present in animal mtDNAs [1], i.e., 13 PCGs, 22 tRNA genes, and 2 ribosomal genes (Figure 1). The mtDNA genome of *O. lunifer* contains also 7 intergenic spacers (s1–s7), spanning at least 15 bp, described in a paragraph below. Genes on the same strand are overlapped (e.g. *trnM* vs. *trnI*; *atp8* vs. *atp6*), contiguous, separated by few nucleotides or by intergenic spacers (e.g. *nad3* vs. *trnA*; *trnC* vs. *trnY*). Genes on opposite strands exhibit a similar behavior (Figure 1).

The *O. lunifer* mtDNA has the typical lepidopteran gene order [8,9] that differs from the ancestral gene order of insects [1] for the placement of *trnM*. In the ancestral type (e.g. *Drosophila yakuba* mtDNA) the order in the  $\alpha$  strand is: A+T region, *trnI*, *trnQ*, *trnM*, *nad2*. In all lepidopteran mtDNAs, sequenced to date, the order is: A+T region, *trnM*, *trnI*, *trnQ*, *nad2* which implies the translocation of *trnM* [5-11]. This placement of *trnM* is a molecular feature exclusive to lepidopteran mtDNAs. Further genome sequencing is necessary to establish if this feature is a mitochondrial signature of the whole order Lepidoptera.

The composition of the  $\alpha$  strand of *O. lunifer* mtDNA is A = 6252 (40.09%), T = 5886 (37.75%), G = 1179 (7.56%) and C = 2276 (14.60%).

The A+T% and G+C% values for the  $\alpha$  strand as well as the A- and G-skews [12] were calculated for all available complete mtDNA genomes of Pterygota and are presented in the scatter plots of Figure 2.

The average A+T% value for the analyzed mtDNAs set is  $76.63 \pm 4.84$ . The highest A+T% values are shared by the mtDNAs of three bees (*Apis mellifera*, *Bombus ignitus* and *Melipona bicolor*) and two bugs (*Aleurodicus dugesii* and *Schizaphis graminum*). All lepidopteran mtDNAs but *O. lunifer* exhibit high A+T% values. The A+T content of *O. lunifer* mtDNA is 77.84% that represents the lowest value for lepidopteran complete mtDNAs [5-8,10]. The lowest A+T contents are found in the termite mtDNAs (*Reticulitermes* spp.). Extreme A+T values are also shared by species having highly re-arranged gene order [13]. However the possession of a re-arranged genome is not sufficient *per se* to have an A+T content drastically departing from the average (e.g. *Aleurochiton aceris* and *Bemisia tabaci*). The A+T values appear to be linked to taxonomic relatedness at low rank (i.e. genus, family) (e.g. species of *Drosophila*, species of *Bactrocera*, members of family *Apidae*). The relation is not true at higher ranks (i.e. superfamily; order) where patterns become inconsistent and the A+T content can be very different among species as exemplified by Hemiptera (*A. dugesii* vs. *Triatoma dimidiata*).

The average A-skew is  $0.04214 \pm 0.11350$  and most of pterygote mtDNAs are slightly to moderately A-skewed with values ranging from 0.00287 (*B. ignitus*) to 0.18247 (*Locusta migratoria*). The lepidopteran A-skews vary from -0.04748 (*C. raphaelis*) to 0.05872 (*B. mori*) with the *O. lunifer* mtDNA exhibiting a slight A-skew (0.03015). The *Reticulitermes* mtDNA genomes, having the lowest A+T% content, exhibit a very pronounced A-skew. Most marked T-skews are observed in the mtDNA genomes of *Campanulotes bidentatus* and *Trialeuroides vaporarium* that have low A+T% content and gene-orders different than insect ancestral gene order [1,14,15]. Gene order re-arrangement is

Table 1: List of taxa analyzed in present paper

Order	Family	Species	Acc. number	Reference	
<b>Orthoptera</b>	Acrididae	<i>Locusta migratoria</i>	<a href="#">NC 001712</a>	[42]	
	Acrididae	<i>Oxya chinensis</i>	<a href="#">NC 010219</a>	Hang and Zhang, unpublished	
	Tettigonidae	<i>Anabrus simplex</i>	<a href="#">NC 009967</a>	[43]	
	Tettigonidae	<i>Ruspolia dubia</i>	<a href="#">NC 009876</a>	[44]	
	Gryllotalpidae	<i>Gryllotalpa orientalis</i>	<a href="#">NC 006678</a>	[45]	
<b>Isoptera</b>	Rhinotermitidae	<i>Reticulitermes flavipes</i>	<a href="#">NC 009498</a>	[13]	
	Rhinotermitidae	<i>Reticulitermes hageni</i>	<a href="#">NC 009501</a>	[13]	
	Rhinotermitidae	<i>Reticulitermes virginicus</i>	<a href="#">NC 009500</a>	[13]	
	Rhinotermitidae	<i>Reticulitermes santonensis</i>	<a href="#">NC 009499</a>	[13]	
<b>Mantophasmatodea</b>	Mantophasmatidae	<i>Sclerophasma paresisense</i>	<a href="#">NC 007701</a>	[46]	
<b>Mantodea</b>	Mantidae	<i>Tamolanica tamolana</i>	<a href="#">NC 007702</a>	[46]	
<b>Blattaria</b>	Blattidae	<i>Periplaneta fuliginosa</i>	<a href="#">NC 006076</a>	[47]	
<b>Plecoptera</b>	Pteronarcyidae	<i>Pteronarcys princeps</i>	<a href="#">NC 006133</a>	[48]	
<b>Phthiraptera</b>	Boopidae	<i>Heterodoxus macropus</i>	<a href="#">NC 002651</a>	[16]	
	Philopteridae	<i>Bothriometopus macrocnemis</i>	<a href="#">NC 009983</a>	[49]	
	Philopteridae	<i>Campanulotes bidentatus</i>	<a href="#">NC 007884</a>	[15]	
<b>Hemiptera</b>	Aleyrodidae	<i>Aleurochiton aceris</i>	<a href="#">NC 006160</a>	[14]	
	Aleyrodidae	<i>Aleurodicus dugesii</i>	<a href="#">NC 005939</a>	[14]	
	Aleyrodidae	<i>Bemisia tabaci</i>	<a href="#">NC 006279</a>	[14]	
	Aleyrodidae	<i>Neomaskellia andropogonis</i>	<a href="#">NC 006159</a>	[14]	
	Aleyrodidae	<i>Tetraleurodes acaciae</i>	<a href="#">NC 006292</a>	[14]	
	Aleyrodidae	<i>Trialeurodes vaporariorum</i>	<a href="#">NC 006280</a>	[14]	
	Aphididae	<i>Schizaphis graminum</i>	<a href="#">NC 006158</a>	[14]	
	Cicadellidae	<i>Homalodisca coagulata</i>	<a href="#">NC 006899</a>	Baumann and Baumann, unpublished	
	Psyllidae	<i>Pachypsylla venusta</i>	<a href="#">NC 006157</a>	[14]	
	Aphrophoridae	<i>Philaenus spumarius</i>	<a href="#">NC 005944</a>	[50]	
	Reduviidae	<i>Triatoma dimidiata</i>	<a href="#">NC 002609</a>	[51]	
<b>Psocoptera</b>	Lepidopsocidae	<i>Lepidopsocid</i> sp. RS-2001	<a href="#">NC 004816</a>	[52]	
<b>Thysanoptera</b>	Thripidae	<i>Thrips imaginis</i>	<a href="#">NC 004371</a>	[53]	
<b>Coleoptera</b>	Cerambycidae	<i>Anoplophora glabripennis</i>	<a href="#">NC 008221</a>	An et al., unpublished	
	Chrysomelidae	<i>Crioceris duodecimpunctata</i>	<a href="#">NC 003372</a>	[54]	
	Elateridae	<i>Pyrophorus divergens</i>	<a href="#">NC 009964</a>	[55]	
	Lampyridae	<i>Pyrocoelia rufa</i>	<a href="#">NC 003970</a>	[56]	
	Tenebrionidae	<i>Tribolium castaneum</i>	<a href="#">NC 003081</a>	[57]	
<b>Lepidoptera</b>	Tortricoidea	Tortricidae	<i>Adoxophyes honmai</i>	<a href="#">NC 008141</a>	[7]
	Bombycoidea	Saturniidae	<i>Antheraea pernyi</i>	<a href="#">NC 004622</a>	Liu et al., unpublished
	Bombycoidea	Bombycidae	<i>Bombyx mandarina</i>	<a href="#">NC 003395</a>	[5]
	Bombycoidea	Bombycidae	<i>Bombyx mori</i>	<a href="#">NC 002355</a>	Lee et al., unpublished
	Bombycoidea	Saturniidae	<i>Saturnia boisduvalii</i>	<a href="#">NC 010613</a>	[9]
	Geometroidea	Geometridae	<i>Phthonandria atrilineata</i>	<a href="#">NC 010522</a>	Yang et al., unpublished
	Papilionoidea	Pieridae	<i>Artogeia melete</i>	<a href="#">NC 010568</a>	Hong et al., unpublished
	Papilionoidea	Lycaenidae	<i>Coreana raphaelis</i>	<a href="#">NC 007976</a>	[6]
	Sphingoidea	Sphingidae	<i>Manduca sexta</i>	<a href="#">EU286785</a>	[8]
	<b>Noctuoidea</b>	<b>Notodontidae</b>	<b><i>Ochrogaster lunifer</i></b>	<a href="#">AM946601</a>	<b>This paper</b>
	Pyraloidea	Crambidae	<i>Ostrinia furnacalis</i>	<a href="#">NC 003368</a>	[10]
	Pyraloidea	Crambidae	<i>Ostrinia nubilalis</i>	<a href="#">NC 003367</a>	[10]
	<b>Diptera</b>	Ceratopogonidae	<i>Culicoides arakawae</i>	<a href="#">NC 009809</a>	Matsumoto, unpublished
		Culicidae	<i>Aedes albopictus</i>	<a href="#">NC 006817</a>	Ho et al., unpublished
		Culicidae	<i>Aedes aegypti</i>	<a href="#">NC 010241</a>	Lobo et al., unpublished
Culicidae		<i>Anopheles gambiae</i>	<a href="#">NC 002084</a>	[58]	
Culicidae		<i>Anopheles quadrimaculatus</i> A	<a href="#">NC 000875</a>	[59]	
Calliphoridae		<i>Cochliomyia hominivorax</i>	<a href="#">NC 002660</a>	[60]	
Calliphoridae		<i>Lucilia sericata</i>	<a href="#">NC 009733</a>	Cibrario et al., unpublished	
Calliphoridae		<i>Chrysomya putoria</i>	<a href="#">NC 002697</a>	[61]	
Drosophilidae		<i>Drosophila melanogaster</i>	<a href="#">NC 001709</a>	[62]	
Drosophilidae		<i>Drosophila mauritiana</i>	<a href="#">NC 005779</a>	[63]	
Drosophilidae		<i>Drosophila sechellia</i>	<a href="#">NC 005780</a>	[63]	
Drosophilidae		<i>Drosophila simulans</i>	<a href="#">NC 005781</a>	[63]	

**Table 1: List of taxa analyzed in present paper (Continued)**

	Drosophilidae	<i>Drosophila yakuba</i>	<a href="#">NC_001322</a>	[64]
	Oestridae	<i>Dermatobia hominis</i>	<a href="#">NC_006378</a>	Azeredo-Espin et al., unpublished
	Muscidae	<i>Haematobia irritans</i>	<a href="#">NC_007102</a>	Lessinger et al., unpublished
	Nemestrinidae	<i>Trichophthalma punctata</i>	<a href="#">NC_008755</a>	[65]
	Syrphidae	<i>Simosyrphus grandicornis</i>	<a href="#">NC_008754</a>	[65]
	Tabanidae	<i>Cydistomyia duplonotata</i>	<a href="#">NC_008756</a>	[65]
	Tephritidae	<i>Ceratitis capitata</i>	<a href="#">NC_000857</a>	[66]
	Tephritidae	<i>Bactrocera carambolae</i>	<a href="#">NC_009772</a>	Ye et al., unpublished
	Tephritidae	<i>Bactrocera dorsalis</i>	<a href="#">NC_008748</a>	Yu et al., unpublished
	Tephritidae	<i>Bactrocera oleae</i>	<a href="#">NC_005333</a>	[67]
	Tephritidae	<i>Bactrocera papayae</i>	<a href="#">NC_009770</a>	Ye et al., unpublished
	Tephritidae	<i>Bactrocera philippinensis</i>	<a href="#">NC_009771</a>	Ye et al., unpublished
<b>Hymenoptera</b>	Apidae	<i>Apis mellifera ligustica</i>	<a href="#">NC_001566</a>	[68]
	Apidae	<i>Bombus ignitus</i>	<a href="#">DQ870926</a>	[69]
	Apidae	<i>Melipona bicolor</i>	<a href="#">NC_004529</a>	Silvestre and Arias, unpublished
	Vanhornidae	<i>Vanhornia eucnemidarum</i>	<a href="#">NC_008323</a>	[70]

not necessarily linked to strong A/T-skew as proved by the highly rearranged, but low skew, genome of *Heterodoxus macropus* [16].

The average G+C% content is  $23.37 \pm 4.84$ . The G+C% pattern among various species is obviously opposite to the A+T% thus it does not require further comments. More composite is the G/C-skew distribution. The average G-skew is  $-0.16006 \pm 0.138235$ . Most of pterygota mtDNAs are C-skewed with G-skew values ranging from  $-0.32827$  (*Vanhornia eucnemidarum*) to  $-0.01250$  (*Heterodoxus macropus*). The main exception is represented by the mtDNA of bugs, while the highest G-skewed genome is that of *C. bidentatus*. Most of lepidopteran mtDNAs share very similar G-skew values that are included within the bulk of mtDNAs. The notable exception is represented by the newly determined mtDNA of *O. lunifer* that exhibits the second most pronounced C-skew (G-skew =  $-0.31751$ ) among analyzed genomes.

G-skew can be markedly different even in species belonging to the same genus and having a very similar G+C content as well exemplified by *Reticulitermes santonensis* and *Reticulitermes virginicus* mtDNAs. The same reasoning applies at high taxonomic rank to the Hemiptera. The mtDNA of *C. bidentatus* exhibits very high A-skew and G-skew. However, this feature is not a general rule and extreme A-skew and G-skew are not necessarily reciprocally linked, as proved by species of genus *Reticulitermes* that exhibit very strong A-skews but not G-skews.

The list of currently available mtDNAs reveals that there is a strong bias in term of taxon sampling both at low and high taxonomic ranks within Pterygota. A direct consequence is that present knowledge of base composition and A/G skews reflects such biases and addition of a single taxon can change our view on these features. This point is well exemplified by the *O. lunifer* mtDNA that exhibits a A+T percentage different than other lepidopteran mtD-

NAs that share high A+T contents [8,9]. Thus a broad and more balanced taxon sampling appears to be a mandatory goal to investigate and identify general patterns for the parameters considered above.

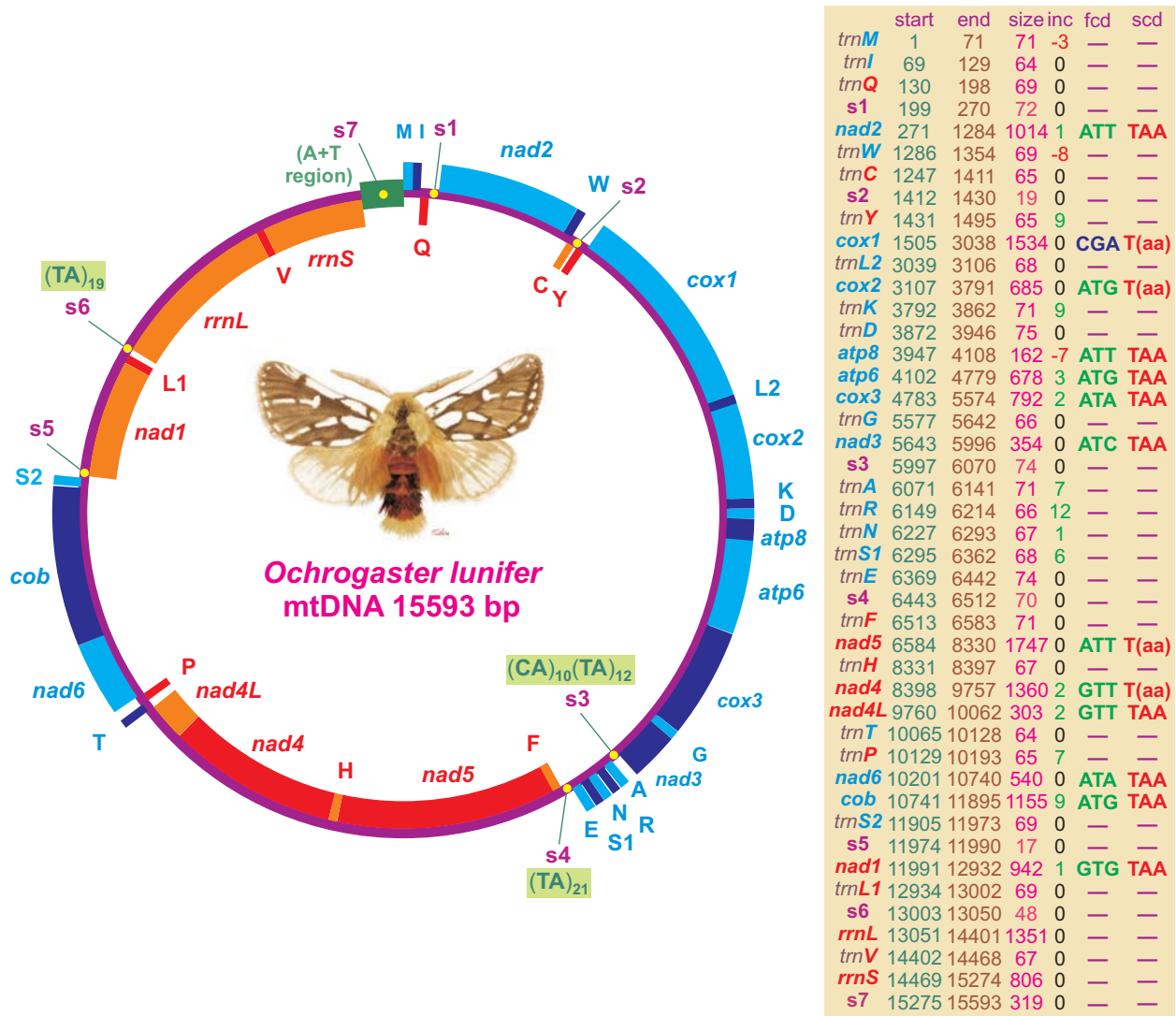
#### Protein-coding genes

The mtDNA of *O. lunifer* contains the full set of PCGs usually present in animal mtDNA. PCGs are arranged along the genome according to the standard order of Insects [1] (Figure 1). The putative start codons of PCGs are those previously known for animal mtDNA i.e. ATN, GTG, TTG, GTT [17] with the only exception represented by the CGA start codon of *cox1* gene. This non-canonical putative start codon is found also in the butterfly *A. melete* and in the moths *A. honmai*, *B. mori*, *B. mandarina*, *M. sexta* and *P. atrilineata* [5-8]. In the butterfly *C. raphaelis* the tetranucleotide TTAG is the putative start codon [6] and the six nucleotide TATTAG has been suggested as putative start codon for the moths *O. nubilalis* and *O. furnicalis* [10]. An unusual start codon for *cox1* gene is known in various arthropod mtDNA [e.g. [18]].

The *cox1*, *cox2*, *nad5*, and *nad4* genes of *O. lunifer* mtDNA have incomplete stop codons. The presence of incomplete stop codons is a feature shared with all lepidopteran mtDNAs sequenced to date [5-10] and more in general with many arthropod mtDNAs [1].

The *atp8* and a *atp6* of *O. lunifer* are the only PCGs having a seven nucleotides overlap (Figure 1). This feature is common to all lepidopteran mtDNA genomes known [5-10] and is found in many animal mtDNAs [1].

The abundance of codon families and Relative Synonymous Codon Usage (RSCU) [19] in PCGs were investigated for all available lepidopteran mtDNAs and the results are summarized in Figures 3 and 4. All first codons as well as stop codons, complete and incomplete, were

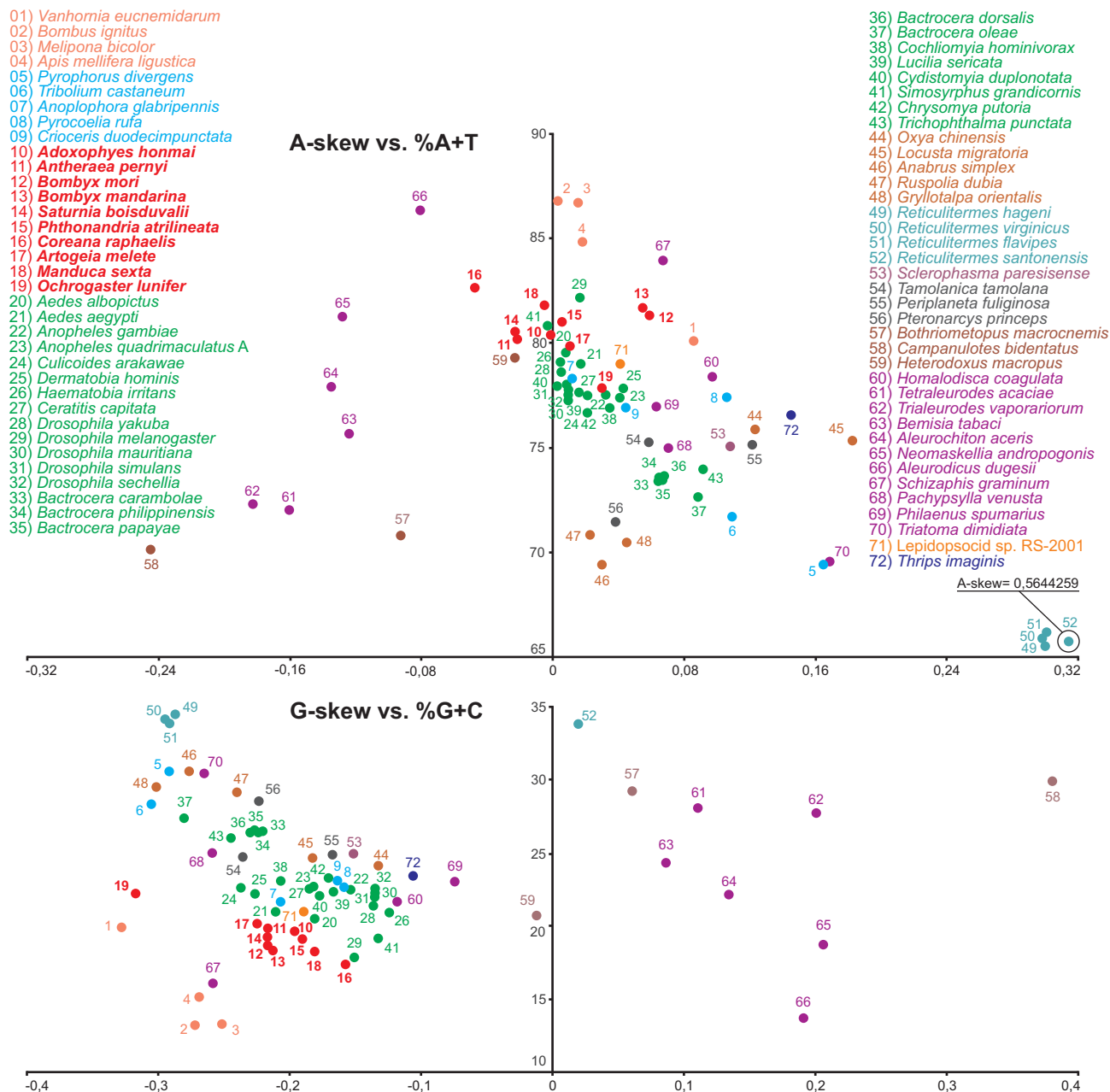


**Figure 1**  
**Map of the mitochondrial genome of *O. lunifer*.** Genes coded in the  $\alpha$  strand (clockwise orientation) are blue or cyan colored. Genes coded in the  $\beta$  strand (anti-clockwise orientation) are red or orange colored. Alternation of colors was applied for clarity. Start, first position along  $\alpha$  strand; end, last position along  $\alpha$  strand; size, size of the sequence; inc, intergenic nucleotides; fcd, first codon; scd, stop codon. Incomplete stop codons are presented with parentheses. Negative inc values refer to overlapping nucleotides for genes located in the same or different strands. Gene names are the standard abbreviations used in this paper; tRNA genes are indicated by the single letter IUPAC-IUB abbreviation for their corresponding amino acid in the draw. s1–s7, intergenic spacers.

excluded from the analysis to avoid biases due to unusual putative start codons and incomplete stop codons.

Total number of non-stop codons (CDs) used by the 12 analyzed mtDNAs is very similar ranging from 3695 of *C. raphaelis* to 3732 of *O. lunifer*. The codon families exhibit a very similar behavior among considered species. The eight codon families with at least 50 CDs per thousand CDs (Leu2, Ile, Phe, Met, Asn, Ser2, Gly, Tyr) encompass

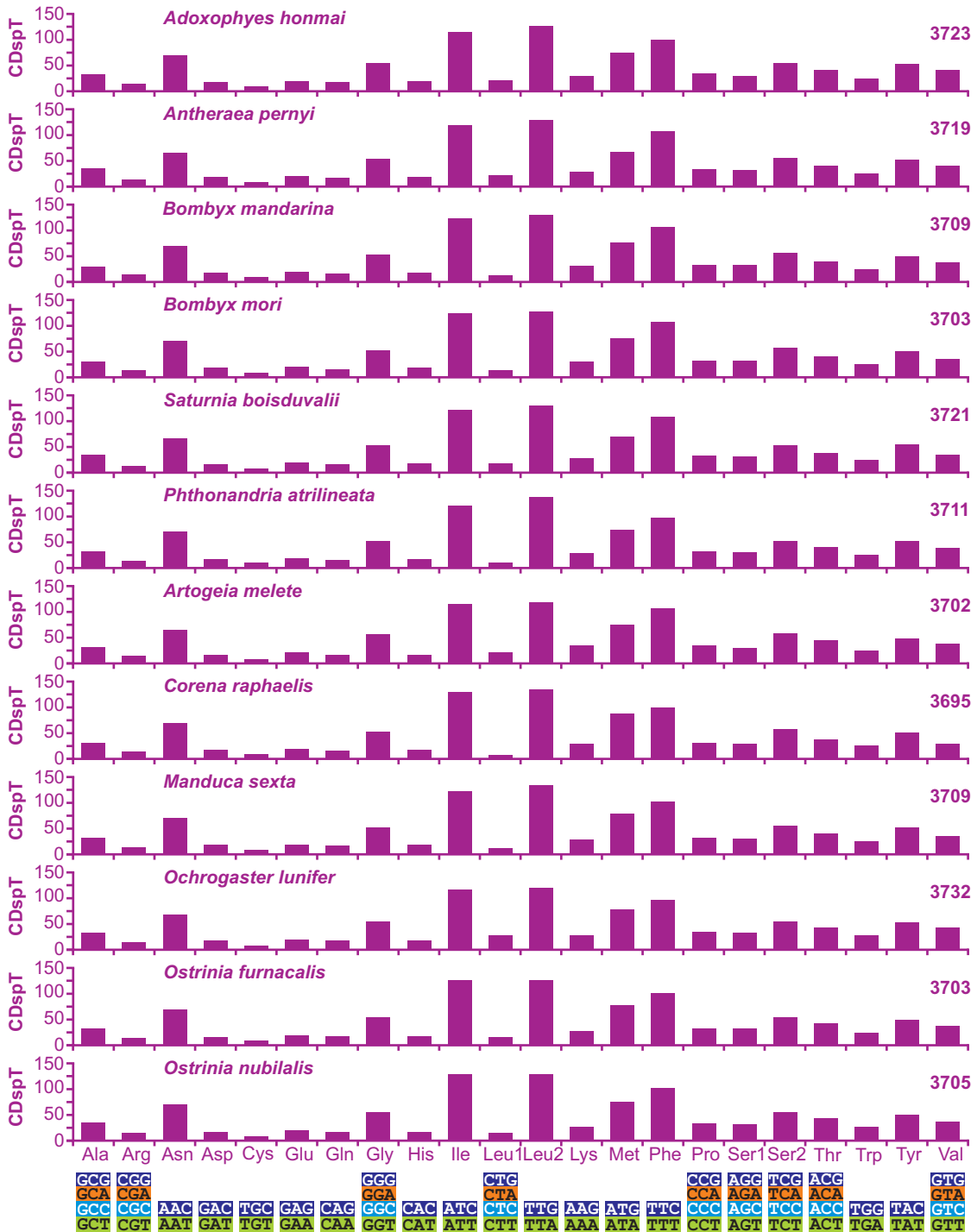
an average  $65.82\% \pm 1.20\%$  of all CDs. The three families with at least 100 CDs per thousand CDs (Leu2, Ile, Phe) account for an average  $35.36\% \pm 0.98\%$  of all CDs (Figure 3). The A+T rich CDs are favored over synonymous CDs with lower A+T content as proved by RSCU results (Figure 4). This point is well exemplified by the Leu2 family where the TTA codon accounts for the large majority of CDs in the family (see below). Invertebrate mitochondrial code includes 62 amino-acid encoding codons [1].



**Figure 2**  
**A-skew vs. A+T% and G-skew vs. G+C% in the Pterygota mtDNAs.** Values were calculated on  $\alpha$  strands for full length mtDNA genomes. The X axis provides the skews values, while the Y axis provides the A+T/G+C values. Named of species are colored according to their taxonomic placement at Order level (see Table I).

Among the 12 analyzed genomes the total number of used codons results to be directly linked to the A+T content. The *C. raphaelis* mtDNA, having the highest A+T% content (see Figure 2) uses 52 codons, and never utilized the 10 G+C rich codons listed in Figure 2. Conversely, *O.*

*lunifer* mtDNA, characterized by the lowest A+T% among considered lepidopteran genomes, uses all 62 codons. Differences in the number of used CDs are present between species of the same genus (e.g. *B. mandarina* vs. *B. mori*) even if the discrepancies appear circumscribed to

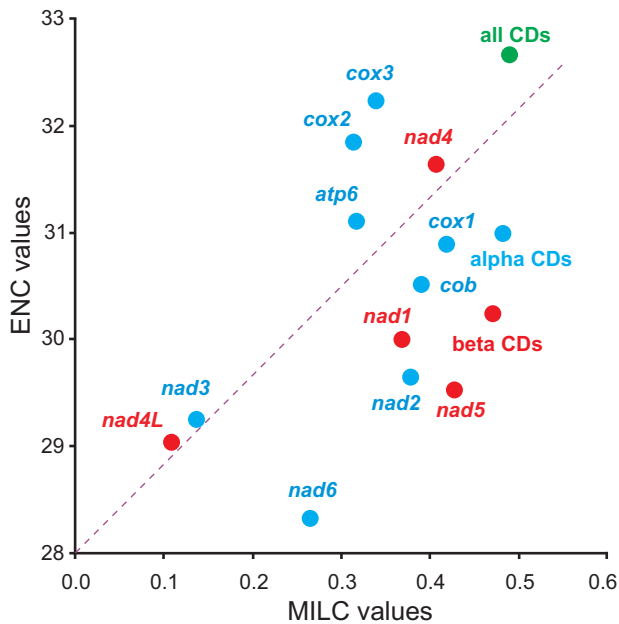


**Figure 3**  
**Codon distribution in lepidopteran mtDNAs.** Numbers to the left refer to the total number of codon. CDspT, codons per thousands codons. Codon Families are provided on the x axis.



**Figure 4**  
**Relative Synonymous Codon Usage (RSCU) in lepidopteran mtDNAs.** Codon Families are provided on the x axis. Red-colored codon, codon not present in the genome. Codon Families are provided on the x axis.





**Figure 5**  
**Scatter plot graphic of MILC vs. ENC calculated for PCGs of lepidopteran mtDNAs.** Dots correspond to average values calculated for different genes. PCGs on  $\alpha$  strand are blue-colored, PCGs on  $\beta$  strand are red colored. All pooled PCGs are presented as a green dot plot. Genes nomenclature as in main text.

G+C rich CDs with very limited use (e.g. GCG and CGC). The Leu1 (average =  $11.73 \pm 3.82\%$ ) and Leu2 (average =  $88.44 \pm 3.89\%$ ) codon families are very differently represented in lepidopteran PCGs while Ser1 (average =  $34.95 \pm 3.67\%$ ) and Ser2 (average =  $64.05 \pm 1.09\%$ ) exhibit a more balanced composition.

Four amino acid residues (Leu, Ile, Phe and Ser) account for more than 44.50% (average =  $45.68 \pm 0.58\%$ ) of all residues forming the 13 mitochondrial proteins. The Leu and Ile amino acids share hydrophobic lateral chains, Phe is also hydrophobic and Ser exhibits an aliphatic behavior [20] thus their massive presence is striking but not surprising for membrane proteins.

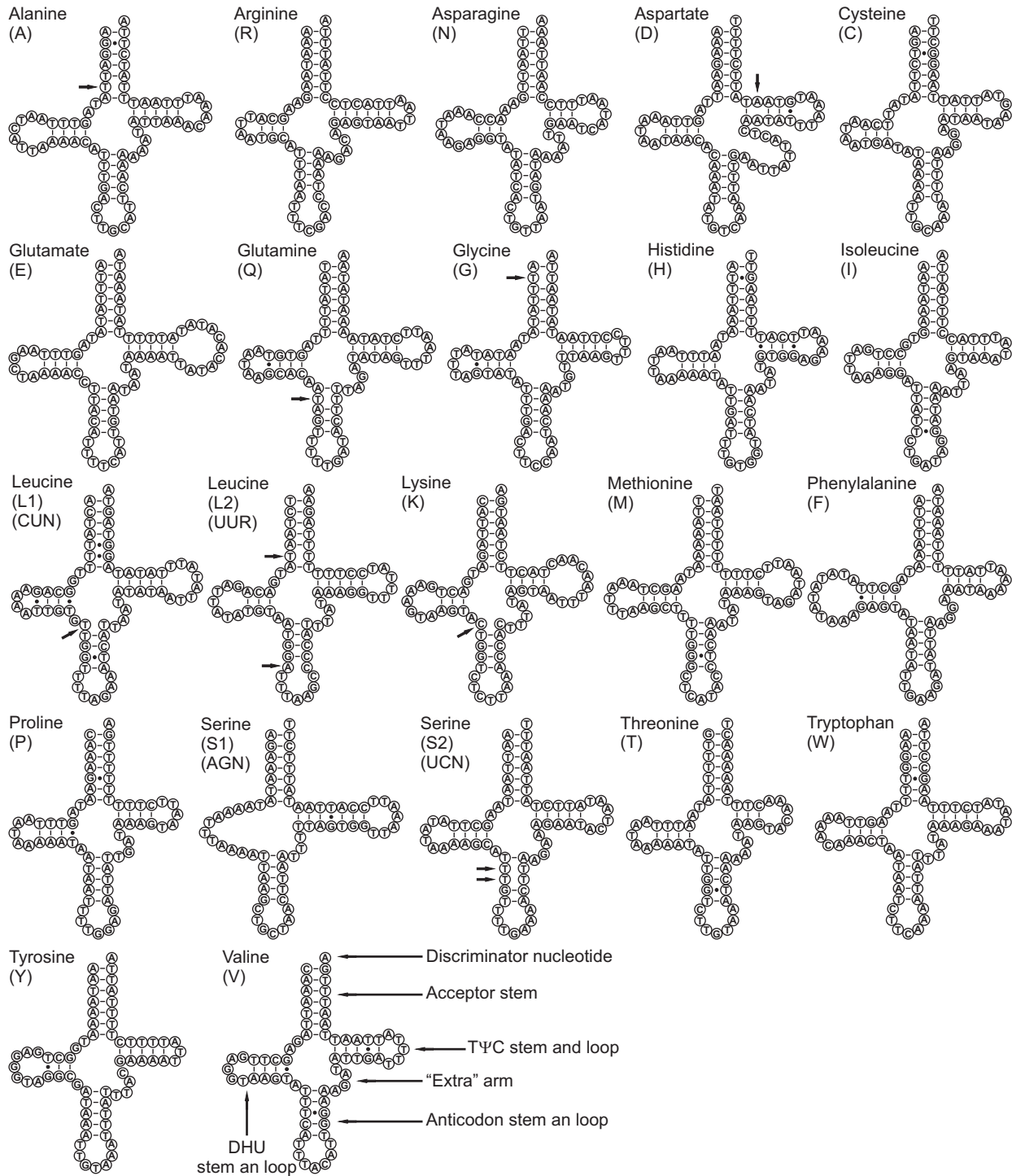
Codon usage by single PCGs was investigated by calculating the two indices ENC (Effective number of codon used) [21] and MILC (Measure Independent of Length and Composition) [22]. Both indices, based on different approaches [21,22] provide a measure of codon variability of PCGs. The ENC and MILC estimate the codon variability in a way that allows comparison among sequences having different lengths as is the case of various PCGs.

Genes exhibiting a higher diversity in codon usage have generally a higher number of variable sites, a prerequisite to be potential phylogenetic markers. Thus the use of ENC and MILC scores, according to the new approach presented in this paper, is a way to study PCG sequences variability on a codon perspective. The best scores of both indices should allow to identify the more diverse PCGs in a approach complementary to the usual method based on evolutionary distances among orthologous sequences (e.g. [8]). The assessment of genetic variability is an interesting point. Indeed some PCGs are standard marker for species recognition [23] or have been extensively used as phylogenetic markers in Lepidoptera while others have received so far limited or no attention. Understanding the genetic diversity of each PCG is a prerequisite to determine its phylogenetic usefulness. The ENC and MILC values were calculated for all PCGs but *atp8* that contains too a few codons to get reliable ENC/MILC estimations [22]. Calculations were extended also to all 13 PCGs pooled as well as to the pooled PCGs belonging to  $\alpha$  and  $\beta$  strands respectively. The scatter plot analysis is provided in Figure 5. As expected the greatest diversity in codon usage is found when all codons are considered. Good codon diversity is found also when all PCGs of  $\alpha$  or  $\beta$  strands are considered. More interesting is the behavior of single genes. In this latter case sequences well established as phylogenetic markers (i.e. *cox1*, *cob*, *nad5*, and *cox2*) are intermixed with PCGs poorly or not considered by researchers (e.g. *cox3*, *nad4*, *nad1*, *nad2*). Our results suggest that the neglected PCGs should be considered as potential markers thus extending the number of mtDNA PCGs sampled for population as well as phylogenetic markers. Findings, based on codon diversity, must be integrated with direct comparisons of sequences [8] that allow to better define the optimal task that each gene can perform i.e. to be used at low taxonomic level or at high taxonomic level.

#### Transfer and ribosomal RNA genes

*Ochrogaster* genome has the characteristic 22 tRNAs set (Figure 6) present in most of animal mtDNAs [1]. All tRNAs present the typical clover leaf secondary structure but *trnS1* lacks the DHU stem. This feature is shared with the *C. raphaelis* mtDNA [6] but is not a general feature of lepidopteran mtDNA as proved by *A. honmai* that has all tRNAs with a complete clover leaf structure [7]. In general, the lack of DHU arm in *trnS1* is a common condition in metazoan mtDNAs [24].

The *trnA*, *trnD*, *trnG*, *trnK*, *trnL1*, *trnL2*, *trnQ*, and *trnS2* of *Ochrogaster* mtDNA show mismatches in their stems. Mismatches are located mostly in the acceptor and anticodon stems with a single exception represented by *trnD* that exhibits the mismatch on the T $\Psi$ C stem. Mismatches on tRNA stems are known also for the *trnA*, *trnL1*, *trnL2*, and *trnQ*, of *C. raphaelis* [6]. Mismatches observed in tRNAs



**Figure 6**  
**Secondary structures of transfer tRNAs in *O. lunifer* mtDNA.**

are corrected through RNA-editing mechanisms that are well known for arthropod mtDNA [e.g. [24]].

Preliminary analysis performed on *rrnL* and *rrnS* of *O. lunifer* revealed that these genes are capable of folding into structures (data not shown) similar to those already produced for lepidopteran mitochondrial ribosomal subunits [8,25,26]. Further studies, that extend the taxon sampling, are currently in progress in our lab to better define *rrnL* and *rrnS* structures within the Thaumetopoeinae subfamily that includes also *O. lunifer*.

### Non coding regions

The mtDNA genome of *O. lunifer* contains 7 intergenic spacers (s1-s7) spanning at least 15 bp (Figures 1 and 7). The features of s1-s7 spacers are presented below with reference to the  $\alpha$  strand for orientation and sequence motifs description.

The s1 spacer, located between *trnQ* and *nad2*, appears to be the result of a duplicated segment (Figure 7). The s1 spacer is present in all 12 lepidopteran mtDNAs so far sequenced while it is absent in other insects [8]. While the genomic location is constant the sequence divergence is high among species [8]. Further investigation with a broad taxon sampling within the Lepidoptera is necessary to assess if the s1 spacer is a constant molecular signature of lepidopteran mtDNA.

The s2 spacer, placed between *trnC* and *trnY*, derives from the triplication of a six nucleotides motif with minor changes (Figure 7). An 11 bp spacer between *trnC* and *trnY* is found also in the mtDNA of *A. melete* and shares the ACAATT motif with the s2 spacer of *O. lunifer*. Because no other known lepidopteran mtDNA exhibits such a spacer its presence in *A. melete* and *O. lunifer* has to be interpreted as the result of independent events.

Spacer s3, located between *nad3* and *trnA*, exhibits a partial duplicated segment and a poly-T motif within the first 30 nt. The second half of s3 spacer is characterized by two microsatellite repeats (CA)<sub>10</sub>(TA)<sub>12</sub>. Spacers having the same genomic location, and containing TA microsatellites are found also in *B. mori* and *B. mandarina* mtDNA genomes.

Spacer s4, inserted between *trnE* and *trnF*, contains a 5' microsatellite (TA)<sub>23</sub>, while the 3' half seems to be the triplication of a 10 nucleotides motif with some changes (Figure 7). A spacer characterized by a different motif (TATTA)<sub>31</sub>, but having the same genomic placement, is found in the *A. honmai* mtDNA genome.

The spacer s5, located between *trnS2* and *nad1*, contains the ATACTAA motif which is conserved across the Lepi-

doptera order [8]. This motif is possibly fundamental to site recognition by the transcription termination peptide (mtTERM protein) [2]. Spacer s5 is present in most insect mtDNAs even if the nucleotide sequence can be quite divergent [8].

The s6 spacer is located between *trnS2* and *-rrnL* and exhibits a di-nucleotide microsatellite (TA)<sub>19</sub> directly in contact with the 3' end of *rrnL* gene. To date spacer s6 is known only for the mtDNA of *O. lunifer*.

The s7 spacer coincides with the A+T region. Several features common to the Lepidoptera A+T region [8] are present in the s7 spacer. The O<sub>R</sub> $\beta$  (origin of the  $\beta$  strand replication) is located 21 bp downstream from *rrnS* gene in *B. mori* [27]. It contains the motif ATAGA followed by an 18 bp poly-T stretch. A very similar pattern occurs in *O. lunifer* where the ATAGA motif is located 17 bp downstream from *rrnS* gene and is followed by a 20 bp poly-T stretch (Figure 7). A microsatellite-like (AT)<sub>7</sub>(TA)<sub>3</sub> element preceded by the ATTTA motif is present in the 3' third of *O. lunifer* s7 spacer. The presence of a microsatellite preceded by the ATTTA motif is also a feature found in the A+T regions of other Lepidoptera [8]. Finally a 10 bp poly-A is present immediately upstream *trnM*. This poly-T (in the  $\beta$  strand) element is still a common feature of the A+T region in Lepidoptera [8,28]. No large repeated segments were detected in the A+T region of *O. lunifer*. This arrangement is consistent with other lepidopteran A+T regions while markedly contrasts with patterns observed in other insect orders [8,29].

Intergenic spacers containing repeated elements are scattered all over the lepidopteran mtDNAs while repeated elements are restricted mostly to the A+T region in other insects [8]. Most parts of spacers of *O. lunifer* are made by repeated motifs. Predominance of repeated elements suggest that mtDNA expansion can be achieved through a miss-pairing duplication mechanism, i.e. DNA slippage, during genome replication. Several intergenic spacers are restricted to a single butterfly/moth species and have not counterparts even within Lepidoptera. Thus it is plausible to suggest that spacers production occurs independently and recursively within Lepidoptera. It remains unknown while this feature is so prominent in moths and butterflies and apparently limited, reduced or absent in other insect mtDNAs sequenced to date. This behavior requires further investigation provided that mtDNA intergenic spacers are found in non-insect Arthropoda as well as other animal phyla [e.g. [18,30]].

### Conclusion

The mitochondrial genome of *O. lunifer* is the first sequenced mtDNA for a representative of the Noctuoidea a superfamily that includes about 40% of all described



genome sequencing will establish if this feature characterizes the whole order Lepidoptera. The mtDNA of *O. lunifer* exhibits a peculiar low A+T content and marked C-skew. Compared to other lepidopteran genomes it is less biased in synonymous codon usage. Comparative analysis on codon usage among lepidopteran mitochondrial genomes identified *atp6*, *cox1*, *cox2*, *cox3*, *cob*, *nad1*, *nad2*, *nad4*, and *nad5* as potential markers for phylogenetic and population genetic studies. Most of the genes listed above have been previously neglected for the tasks suggested here. The massive presence of repetitive elements in intergenic spacers of *O. lunifer* genome lead us to suggest an important role of DNA slippage as possible mechanism to produce spacers during replication.

## Methods

### Sample origin and DNA extraction

An ethanol-preserved larva specimen of *Ochrogaster lunifer* collected in Australia (Suburb of Kenmore, Queensland, 25th February 2005) by Myron P. Zalucki (University of Queensland) was used as starting material for this study. Total DNA was extracted by applying a salting-out protocol [31]. Quality of DNA was assessed through electrophoresis in a 1% agarose gel and staining with ethidium bromide.

### PCR amplification and sequencing of *Ochrogaster lunifer* mtDNA

PCR amplification was performed using a mix of insect universal primers [32,33] and primers specifically designed on the *O. lunifer* sequences. For a full list of successful primers as well as PCR conditions see Additional file 1. The PCR products were visualized in electrophoresis in a 1% agarose gel and staining with ethidium bromide. Each PCR product represented by a single electrophoretic band was purified with the ExoSAP-IT kit (Amersham Biosciences) and directly sequenced. Sequencing of both strands was performed at the BMR Genomics service (Padova, Italy) on automated DNA sequencers mostly employing the primers used for PCR amplification.

### Sequence assembly and annotation

The mtDNA final consensus sequence was assembled using the SeqMan II program from the Lasergene software package (DNASTar, Madison, WI). Genes and strands nomenclature used in this paper follows Negrisolo et al. [18].

Sequence analysis was performed as follows. Initially the mtDNA sequence was translated into putative proteins using the Transeq program available at the EBI web site. The true identity of these polypeptides was established using the BLAST program [34,35] available at the NCBI web site. Gene boundaries were determined as follows. The 5' ends of PEGs were inferred to be at the first legiti-

mate in-frame start codon (ATN, GTG, TTG, GTT; [17]) in the open reading frame (ORF) that was not located within the upstream gene encoded on the same strand. The only exception was *atp6*, which has been previously demonstrated to overlap with its upstream gene *atp8* in many mtDNAs [17]. The PCG terminus was inferred to be at the first in-frame stop codon encountered. When the stop codon was located within the sequence of a downstream gene encoded on the same strand, a truncated stop codon (T or TA) adjacent to the beginning of the downstream gene was designated as the termination codon. This codon was thought to be completed by polyadenylation to a complete TAA stop codon after transcript processing. Finally pair-wise comparisons with orthologous proteins were performed with ClustalW program [36] to better define the limits of PCGs.

Irrespectively of the real initiation codon, a formyl-Met was assumed to be the starting amino acid for all the proteins as previously proved for other mitochondrial genomes [37,38].

The transfer RNA genes were identified using the tRNAscan-SE program [39] or recognized manually as sequences having the appropriate anticodon and capable of folding into the typical cloverleaf secondary structure [17].

The boundaries of the ribosomal *rnmL* gene were assumed to be delimited by the ends of the *trnV-s6* pair. The 3' end of *rnmS* gene was assumed to be delimited by the start of *trnV* while the 5' end was determined through comparison with orthologous genes of other Lepidoptera so far sequenced.

### Genomic analysis

Nucleotide composition was calculated with the EditSeq program included in the Lasergene software package. The GC-skew = (G-C)/(G+C) and AT-skew = (A-T)/(A+T) were used [12] to measure the base compositional difference between the different strands or between genes coded on the alternative strands. The Relative Synonymous Codon Usage (RSCU) values were calculated with MEGA 4 program [40].

The codon usage by analyzed genomes was investigated by calculating the two indices ENC (Effective Number of Codon used) [21] and MILC (Measure Independent of Length and Composition) [22]. ENC and MILC values were calculated with the INCA 2.1 program [41].

### Abbreviations

mtDNA: mitochondrial DNA; *atp6* and *atp8*: ATP synthase subunits 6 and 8; *cob*: apocytochrome b; *cox1-3*: cytochrome c oxidase subunits 1–3; *nad1-6* and *nad4L*: NADH dehydrogenase subunits 1–6 and 4L; *rnmS* and *rnmL*: small

and large subunit ribosomal RNA (rRNA) genes; *trnX*: transfer RNA (tRNA) genes, where X is the one-letter abbreviation of the corresponding amino acid; s1-s7: mitochondrial genomic spacers; A+T region: the putative control region; PCG: protein coding gene; RSCU: Relative Synonymous Codon Usage; ENC, MILC: Measure Independent of Length and Composition; aa: amino acids; nt: nucleotides; bp: base pairs.

### Authors' contributions

PS and MS carried out the molecular experiments. AB and EN designed and coordinated all experiments. EN performed the genomic analyses. All authors contributed to the manuscript and then read and approved the final version.

### Additional material

#### Additional file 1

Additional file 1. List of primers and PCR conditions used in the sequencing of *Ochogaster lunifer* mtDNA.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-331-S1.pdf>]

### Acknowledgements

We express our sincere thanks to Myron P. Zalucki (School of Integrative Biology, University of Queensland, Brisbane, Australia) who kindly provided the specimen of *Ochogaster lunifer* used in present study. We thank Filippo Calore (Albignasego, Padova, Italy) who painted the icon of *O. lunifer* included in Figure 1, using as template a picture publically available at the CSIRO web site. Finally we thank two anonymous referees that provided very useful suggestions that helped to improve the manuscript.

### References

- Boore JL: **Animal mitochondrial genomes.** *Nucleic Acids Res* 1999, **27**:1767-1780.
- Taanman JW: **The mitochondrial genome: structure, transcription, translation and replication.** *Biochim Biophys Acta* 1999, **1410**:103-123.
- Resh VH, Cardé RG: **Insecta, Overview.** In *Encyclopedia of Insects* Edited by: Resh VH, Cardé RG. Academic Press, Burlington MA, USA; 2003:564-566. 1266pp.
- Powell JA: **Lepidoptera (Moths, Butterflies).** In *Encyclopedia of Insects* Edited by: Resh VH, Cardé RG. Academic Press, Burlington MA, USA; 2003:631-663. 1266pp.
- Yukuhiro K, Sezutsu H, Itoh M, Shimizu K, Banno Y: **Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkworm, *Bombyx mandarina* and its close relative, the domesticated silkworm, *Bombyx mori*.** *Mol Biol Evol* 2002, **19**:1385-1389.
- Kim I, Lee EM, Seol KY, Yun EY, Lee YB, Hwang JS, Jin BR: **The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae).** *Insect Mol Biol* 2006, **15**(2):217-225.
- Lee E-S, Shin KS, Kim M-S, Park H, Cho S, Kim C-B: **The mitochondrial genome of the smaller tea tortrix *Adoxophyes honmai* (Lepidoptera: Tortricidae).** *Gene* 2006, **373**:52-57.
- Cameron SL, Whiting MF: **The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths.** *Gene* 2008, **408**:112-123.
- Hong MY, Lee EM, Jo YH, Park HC, Kim SR, Hwang JS, Jin BR, Kang PD, Kim KG, Han YS, Kim I: **Complete nucleotide sequence and organization of the mitogenome of the silk moth *Caligula boissiduvallii* (Lepidoptera: Saturniidae) and comparison with other lepidopteran insects.** *Gene* 2008, **413**:49-57.
- Coates BS, Sumerford DV, Hellmich RL, Lewis LC: **Partial mitochondrial genome sequences of *Ostrinia nubilalis* and *Ostrinia furnicalis*.** *Int J Biol Sci* 2005, **1**:13-18.
- Taylor MFJ, McKechnie SW, Pierce N, Kreitman M: **The lepidopteran mitochondrial control region: structure and evolution.** *Mol Biol Evol* 1993, **10**:1259-1272.
- Perna NT, Kocher TD: **Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes.** *J Mol Evol* 1995, **41**:353-358.
- Cameron SL, Whiting MF: **Mitochondrial genomic comparisons of the subterranean termites from the genus *Reticulitermes* (Insecta: Isoptera: Rhinotermitidae).** *Genome* 2007, **50**:188-202.
- Thao ML, Baumann L, Baumann P: **Organization of the mitochondrial genomes of whiteflies, aphids, and psyllids (Hemiptera, Sternorrhyncha).** *BMC Evolutionary Biology* 2004, **4**:25.
- Covacin C, Shao R, Cameron S, Barker SC: **Extraordinary number of gene rearrangements in the mitochondrial genomes of lice (Phthiraptera: Insecta).** *Insect Mol Biol* 2006, **15**:63-68.
- Shao R, Campbell NJ, Barker SC: **Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera).** *Mol Biol Evol* 2001, **18**:858-865.
- Wolstenholme DR: **Animal mitochondrial DNA: structure and evolution.** *Int Rev Cytol* 1992, **141**:173-216.
- Negrisol E, Minelli A, Valle G: **Extensive gene order rearrangement in the mitochondrial genome of the centipede *Scutigera coleoptrata*.** *J Mol Evol* 2004, **58**:413-423.
- Sharp PM, Tuohy TMF, Mosurski KR: **Codon usage in yeast: Cluster analysis clearly differentiates highly and lowly expressed genes.** *Nucleic Acids Research* 1986, **14**:5125-5143.
- Patthy L: **Protein evolution.** 2nd edition. Blackwell, London; 2008:374.
- Wright F: **The 'effective number of codons' used in a gene.** *Gene* 1990, **87**:23-29.
- Supek F, Vlahoviček K: **Comparison of codon usage measures and their applicability in prediction of microbial gene expressivity.** *BMC Bioinformatics* 2005, **6**:182.
- Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN: **DNA barcodes distinguish species of tropical Lepidoptera.** *Proc Natl Acad Sci USA* 2006, **103**:968-971.
- Lavrov DV, Brown WM, Boore JL: **A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*.** *Proc Natl Acad Sci USA* 2000, **97**:13738-13742.
- Niehuis O, Naumann CM, Misof B: **Identification of evolutionary conserved structural elements in the mt SSU Rrna of Zygaenoidea (Lepidoptera): a comparative sequence analysis.** *Org Divers Evol* 2006, **6**:17-32.
- Niehuis O, Yen S-H, Naumann CM, Misof B: **Higher phylogeny of zygaenid moths (Insecta: Lepidoptera) inferred from nuclear and mitochondrial sequence data and the evolution of larval cuticular cavities for chemical defence.** *Mol Phylogenet Evol* 2006, **39**:812-829.
- Saito S, Tamura K, Aotsuka T: **Replication origin of mitochondrial DNA in insects.** *Genetics* 2005, **171**:1695-1705.
- Vila M, Björklund M: **The utility of the neglected mitochondrial control region for evolutionary studies in Lepidoptera (Insecta).** *J Mol Evol* 2004, **58**:280-290.
- Zhang DX, Hweitt GM: **Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies.** *Biochem Syst Ecol* 1997, **25**:99-120.
- Boore JL: **The complete sequence of the mitochondrial genome of *Nautilus macromphalus* (Mollusca: Cephalopoda).** *BMC Genomics* 2006, **7**:182.
- Patwary MU, Kenchington EL, Bird CJ, Zouros E: **The use of random amplified polymorphic DNA markers in genetic studies of the sea scallop *Placopecten magellanicus* (Gmelin, 1791).** *J Shellfish Res* 1994, **13**:547-553.

32. Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P: **Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers.** *Ann Entomol Soc Am* 1994, **87**:651-704.
33. Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT: **Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA.** *Annual Review of Ecology, Evolution, and Systematics* 2006, **37**:545-579.
34. Altschul SF, Gish W, Miller W, Myers EV, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403-410.
35. Tatusova TA, Madden TL: **BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences.** *FEMS Microbiol Lett* 1999, **174**(2):247-250.
36. Thompson JD, Higgins DG, Gibson TJ: **Clustal-W – improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22**:4673-4680.
37. Smith AE, Marcker KA: **N-formylmethionyl transfer RNA in mitochondria from yeast and rat liver.** *J Mol Biol* 1968, **38**:241-243.
38. Fearnley IM, Walker JE: **Initiation codons in mammalian mitochondria: differences in genetic code in the organelle.** *Biochemistry* 1987, **26**:8247-8251.
39. Lowe TM, Eddy SR: **tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence.** *Nucleic Acids Res* 1997, **25**:955-964.
40. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0.** *Mol Biol Evol* 2007, **24**:1596-1599.
41. Supek F, Vlahoviček K: **INCA: synonymous codon usage analysis and clustering by means of self-organizing map.** *Bioinformatics* 2004, **20**:2329-2330.
42. Flook PK, Rowell CH, Gellissen G: **The sequence, organization, and evolution of the *Locusta migratoria* mitochondrial genome.** *J Mol Evol* 1995, **41**:928-941.
43. Fenn JD, Cameron SL, Whiting MF: **The complete mitochondrial genome sequence of the Mormon cricket (*Anabrus simplex*: Tettigoniidae: Orthoptera) and an analysis of control region variability.** *Insect Mol Biol* 2007, **16**:239-252.
44. Zhou Z, Huang Y, Shi F: **The mitochondrial genome of *Ruspolia dubia* (Orthoptera: Conocephalidae) contains a short A+T-rich region of 70 bp in length.** *Genome* 2007, **50**:855-866.
45. Kim I, Cha SY, Yoon MH, Hwang JS, Lee SM, Sohn HD, Jin BR: **The complete nucleotide sequence and gene organization of the mitochondrial genome of the oriental mole cricket, *Gryllotalpa orientalis* (Orthoptera: Gryllotalpidae).** *Gene* 2005, **353**:155-168.
46. Cameron SL, Barker SC, Whiting MF: **Mitochondrial genomics and the new insect order Mantophasmatodea.** *Mol Phylogenet Evol* 2006, **38**(1):274-279.
47. Yamauchi MM, Miya MU, Nishida M: **Use of a PCR-based approach for sequencing whole mitochondrial genomes of insects: two examples (cockroach and dragonfly) based on the method developed for decapod crustaceans.** *Insect Mol Biol* 2004, **13**:435-442.
48. Stewart JB, Beckenbach AT: **Insect mitochondrial genomics 2: The complete mitochondrial genome sequence of a giant stonefly, *Pteronarcys princeps*, asymmetric directional mutation bias, and conserved plecopteran A+T-region elements.** *Genome* 2006, **49**:815-824.
49. Cameron SL, Johnson KP, Whiting MF: **The mitochondrial genome of the screamer louse *Bothriometopus* (phthiraptera: ischnocera): effects of extensive gene rearrangements on the evolution of the genome.** *J Mol Evol* 2007, **65**:589-604.
50. Stewart JB, Beckenbach AT: **Insect mitochondrial genomics: the complete mitochondrial genome sequence of the meadow spittlebug *Philaenus spumarius* (Hemiptera: Auchenorrhyncha: Cercopoidae).** *Genome* 2005, **48**:46-54.
51. Dotson EM, Beard CB: **Sequence and organization of the mitochondrial genome of the Chagas disease vector, *Triatoma dimidiata*.** *Insect Mol Biol* 2001, **10**:205-215.
52. Shao R, Downton M, Murrell A, Barker SC: **Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of Insects.** *Mol Biol Evol* 2003, **20**:1612-1619.
53. Shao R, Barker SC: **The highly rearranged mitochondrial genome of the plague thrips, *Thrips imaginis* (Insecta: Thysanoptera): convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes.** *Mol Biol Evol* 2003, **20**:362-370.
54. Stewart JB, Beckenbach AT: **Phylogenetic and genomic analysis of the complete mitochondrial DNA sequence of the spotted asparagus beetle *Crioceris duodecimpunctata*.** *Mol Phylogenet Evol* 2003, **26**(3):513-526.
55. Arnoldi FG, Ogoh K, Ohmiya Y, Viviani VR: **Mitochondrial genome sequence of the Brazilian luminescent click beetle *Pyrophorus divergens* (Coleoptera: Elateridae): mitochondrial genes utility to investigate the evolutionary history of Coleoptera and its bioluminescence.** *Gene* 2007, **405**:1-9.
56. Bae JS, Kim I, Sohn HD, Jin BR: **The mitochondrial genome of the firefly, *Pyrocoelia rufa*: complete DNA sequence, genome organization, and phylogenetic analysis with other insects.** *Mol Phylogenet Evol* 2004, **32**(3):978-985.
57. Friedrich M, Muqim N: **Sequence and phylogenetic analysis of the complete mitochondrial genome of the flour beetle *Tribolium castaneum*.** *Mol Phylogenet Evol* 2003, **26**(3):502-512.
58. Beard CB, Hamm DM, Collins FH: **The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects.** *Insect Mol Biol* 1993, **2**:103-124.
59. Mitchell SE, Cockburn AF, Seawright JA: **The mitochondrial genome of *Anopheles quadrimaculatus* species A: complete nucleotide sequence and gene organization.** *Genome* 1993, **36**:1058-1073.
60. Lessinger AC, Martins Junqueira AC, Lemos TA, Kemper EL, da Silva FR, Vettore AL, Arruda P, Azeredo-Espin AM: **The mitochondrial genome of the primary screwworm fly *Cochliomyia hominivorax* (Diptera: Calliphoridae).** *Insect Mol Biol* 2000, **9**:521-529.
61. Junqueira AC, Lessinger AC, Torres TT, Da Silva FR, Vettore AL, Arruda P, Azeredo-Espin AM: **The mitochondrial genome of the blowfly *Chrysomya chloropyga* (Diptera: Calliphoridae).** *Gene* 2004, **339**:7-15.
62. Lewis DL, Farr CL, Kaguni LS: ***Drosophila melanogaster* mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons.** *Insect Mol Biol* 1995, **4**:263-278.
63. Ballard JW: **Comparative genomics of mitochondrial DNA in members of the *Drosophila melanogaster* subgroup.** *J Mol Evol* 2000, **51**:48-63.
64. Clary DO, Wolstenholme DR: **The mitochondrial DNA molecular of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code.** *J Mol Evol* 1985, **22**:252-271.
65. Cameron SL, Lambkin CL, Barker SC, Whiting MF: **A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high precision.** *Syst Entomol* 2007, **32**:40-59.
66. Spanos L, Koutroumbas G, Kotsyfakis M, Louis C: **The mitochondrial genome of the mediterranean fruit fly, *Ceratitis capitata*.** *Insect Mol Biol* 2000, **9**:139-144.
67. Nardi F, Carapelli A, Dallai R, Frati F: **The mitochondrial genome of the olive fly *Bactrocera oleae*: two haplotypes from distant geographical locations.** *Insect Mol Biol* 2003, **12**:605-611.
68. Crozier RH, Crozier YC: **The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization.** *Genetics* 1993, **133**:97-117.
69. Cha SY, Yoon HJ, Lee EM, Yoon MH, Hwang JS, Jin BR, Han YS, Kim I: **The complete nucleotide sequence and gene organization of the mitochondrial genome of the bumblebee, *Bombus ignitus* (Hymenoptera: Apidae).** *Gene* 2007, **392**:206-220.
70. Castro LR, Ruberu K, Downton M: **Mitochondrial genomes of *Vanhornia eucnemidarum* (Apocrita: Vanhorniidae) and *Primeuchroeus* spp. (Aculeata: Chrysididae): Evidence of rearranged mitochondrial genomes within the Apocrita (Insecta: Hymenoptera).** *Genome* 2006, **49**:752-766.