

Research article

Open Access

The mitochondrial genome of the 'twisted-wing parasite' *Mengenilla australiensis* (Insecta, Strepsiptera): a comparative study

Dino P McMahon*, Alexander Hayward and Jeyaraney Kathirithamby

Address: Department of Zoology, University of Oxford, The Tinbergen Building, South Parks Road, Oxford, OX1 3PS, UK

Email: Dino P McMahon* - dino.mcmahon@zoo.ox.ac.uk; Alexander Hayward - alexander.hayward@zoo.ox.ac.uk; Jeyaraney Kathirithamby - jeyaraney.kathirithamby@zoo.ox.ac.uk

* Corresponding author

Published: 14 December 2009

Received: 15 July 2009

BMC Genomics 2009, 10:603 doi:10.1186/1471-2164-10-603

Accepted: 14 December 2009

This article is available from: <http://www.biomedcentral.com/1471-2164/10/603>

© 2009 McMahon 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: Strepsiptera are an unusual group of sexually dimorphic, entomophagous parasitoids whose evolutionary origins remain elusive. The lineage leading to *Mengenilla australiensis* (Family Mengenillidae) is the sister group to all remaining extant strepsipterans. It is unique in that members of this family have retained a less derived condition, where females are free-living from pupation onwards, and are structurally much less simplified. We sequenced almost the entire mitochondrial genome of *M. australiensis* as an important comparative data point to the already available genome of its distant relative *Xenos vesparum* (Family Xenidae). This study represents the first in-depth comparative mitochondrial genomic analysis of Strepsiptera.

Results: The partial genome of *M. australiensis* is presented as a 13421 bp fragment, across which all 13 protein-coding genes (PCGs), 2 ribosomal RNA (rRNA) genes and 18 transfer RNA (tRNA) sequences are identified. Two tRNA translocations disrupt an otherwise ancestral insect mitochondrial genome order. A+T content is measured at 84.3%, C-content is also very skewed. Compared with *M. australiensis*, codon bias in *X. vesparum* is more balanced. Interestingly, the size of the protein coding genome is truncated in both strepsipterans, especially in *X. vesparum* which, uniquely, has 4.3% fewer amino acids than the average holometabolan complement. A revised assessment of mitochondrial rRNA secondary structure based on comparative structural considerations is presented for *M. australiensis* and *X. vesparum*.

Conclusions: The mitochondrial genome of *X. vesparum* has undergone a series of alterations which are probably related to an extremely derived lifestyle. Although *M. australiensis* shares some of these attributes; it has retained greater signal from the hypothetical most recent common ancestor (MRCA) of Strepsiptera, inviting the possibility that a shift in the mitochondrial selective environment might be related to the specialization accompanying the evolution of a small, morphologically simplified completely host-dependent lifestyle. These results provide useful insights into the nature of the evolutionary transitions that accompanied the emergence of Strepsiptera, but we emphasize the need for adequate sampling across the order in future investigations concerning the extraordinary developmental and evolutionary origins of this group.

Background

Strepsiptera are an unusual group of obligate endoparasitoid insects [1]. They occur as a small (approx. 600 spp.) monophyletic insect order with uncertain evolutionary origins, inclusive of any clear understanding over the group's nearest extant relative. Strepsiptera parasitize 7 orders of insects, including silverfish (Thysanura); cockroaches (Blattaria); mantids (Mantodea); crickets and grasshoppers (Orthoptera); bugs (Hemiptera); wasps, ants and bees (Hymenoptera) and flies (Diptera). Understandably, the very uniqueness of Strepsiptera makes their placement within insects using morphological taxonomic methods a contentious task. Strepsiptera have been placed as the sister-group to myriad different insect groups, from beetles to true flies, even being placed outside of Holometabola: each hypothesis being founded on one or two 'key' characteristics that at one time or another have come under question [2-15]. Confident assertions of classification have especially been restricted because intermediate forms have largely gone extinct and are unrecorded in the fossil record (see [16] for a notable exception). Simplification of gross morphology during strepsipteran evolutionary specialization can also be viewed as a significant component of this problem: observable morphological variation is low and unevenly distributed between extremely dimorphic sexes.

Females are especially strongly simplified (lacking in most typical adult characters such as wings, legs or mouthparts), and in most strepsipteran species, they remain in the living host until the end of the reproductive cycle. Conversely, males metamorphose in a typical holometabolan fashion; developing wings and a usual suite of adult insect characters (such as antennae, mouthparts and compound eyes) only to leave the host immediately in search of a mate; usually in the form of a female-containing host [17-19]. Males deliver sperm through a brood canal opening in the cephalothorax (the modified "head") of the female, who as a reproductive adult is only partially exposed to the environment, as an extrusion between the tergites or sternites of living hosts. After fertilization, females are capable of producing many hundreds of thousands of active 1st instar larvae who emerge from the cephalothorax.

Members of the strepsipteran lineage, Mengenillidae, are the sister-group to Stylopodia, a clade that includes all other extant Strepsiptera [[20], unpublished data]. This family represents a transitional phase in the evolutionary specialization of Strepsiptera, whereby both sexes leave the host before pupation, and females do not reproduce or release progeny in the unusual fashion outlined above. The level of simplification in free living mengenillid females is much less extreme than in Stylopodia (e.g. *X. vesparum*); legs, mouthparts and compound eyes are all

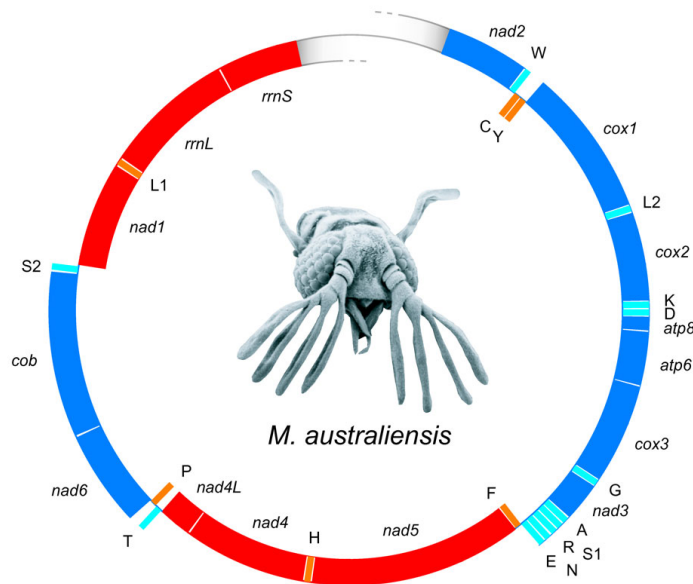
still present, although strongly reduced. The invention of a completely endoparasitoid female was probably one of the most important novelties leading to the radiation of this unique group [21]. Evolutionary studies attempting to unravel the developmental and ecological phenomena that make Strepsiptera so biologically interesting therefore require that species from before this important transition occupy a central role in research.

Insect mitochondrial genomes provide a useful medium to deepen and connect comprehension of microevolutionary forces of populations, like neutral drift and selective sweeps, to macroevolutionary events affecting species and/or deeper levels of divergence. The abundance of mitochondria in most metazoan cell and tissue types makes mtDNA an easily obtainable, universally plentiful marker, where a lack of introns or duplicate genes and non-coding variable spacer regions make amplification of mtDNA relatively uncomplicated [22] (although pseudogenes co-opted by the nucleus from the mitochondrial genome can create separate difficulties [23-26]). The near-complete sequence of the mt genome of *Xenos vesparum* is available, and as a member of the more specialized clade Stylopodia, it is important to place it into context. The mitochondrial genome of *Mengenilla australiensis* (Family Mengenillidae) was therefore chosen for sequencing. Genome arrangement, nucleotide content, codon usage and the secondary structure of ribosomal RNA genes are each comparatively assessed between two distant Strepsiptera, and more widely across Holometabola (a major division of insects defined by the process of complete metamorphosis, including the 4 very species-rich orders; Lepidoptera, Diptera, Hymenoptera, Coleoptera and 6 smaller orders).

Results and discussion

Genome composition

The near-entire mt genome fragment of *M. australiensis* is presented as a 13421 bp sequence, across which 13 protein-coding genes (PCGs), a large (16S) and a small (12S) subunit rRNA gene (*rnl* and *rns* respectively) and 18 tRNA sequences can be identified (Figure 1). PCGs are in expected positions, but two tRNA translocations disrupt an otherwise ancestral genome arrangement. Serine₁ (S1: AGN) moves to a position between Alanine (A) and Arginine (R), whereas Valine (V) is either lost, or transferred to a position in a region that spans the flanking region, potentially anywhere between tRNA-Ile and *nad2*. All start codons across the PCGs (*nad2* is only partial) begin with the typical M- or I- residue [27], and end in TAA, TAG, TA, or T. The control region itself, including up to 4 flanking tRNA genes could not be amplified. A variety of approaches were explored, none of which were successful in spanning the presumed A+T region. Similar difficulties were encountered during the amplification of the mt

**Figure 1**

Summary schematic of the mt genome of *M. australiensis*. Genes transcribed on the leading (α) strand are given in blue (PCGs) or light blue (tRNAs) and run clockwise. Genes of the opposite polarity are given in red (PCGs) or orange (tRNAs). Abbreviations are described in the text.

genome of *X. vesparum* [28]. In our study, genome-specific primers designed to span the A+T region also performed well when reverse complement versions were implemented. It is hypothesized that the region is unusually long and/or too problematic for even advanced *Taq* polymerases to amplify, as a result of extreme repetitiveness or secondary structural folding issues (or both).

Figure 2 compares the genome organization of *M. australiensis* and *X. vesparum* against the inferred ancestral arrangement [29,30]; the former having just two gene order rearrangements, compared to 4 in the latter. In the context of variation across a dataset containing 68 mitochondrial genomes taken from 6 major holometabolous groups compiled from Genbank (see methods), this would be consistent with a mostly ancestral arrangement being retained between the origins of the major holometabolous radiations, since every order contains representative ancestral (or near-ancestral as in *M. australiensis*) genome arrangements. Within certain orders, especially Hymenoptera (and certain hemimetabolous groups not discussed here [31]), lineages have undergone significant independent modifications (see [32] for a full review of hymenopteran mt genome evolution). Some synapomorphic alterations to this ancestral arrangement also exist in certain orders. These include a tRNA arrangement shift

from IQM to MIQ across Lepidoptera [33] (outside of the fragment shown in Figure 2), and a likely shift from WCY to CWY in Neuroptera - a modification that is not shared by the other neuropterid orders [34]. But largely, the 68 genomes corroborate a hypothesis positing that the holometabolous orders emerged during a (relatively) brief period of ancient rapid radiation. Or in the context of phylogenetic tree shape, as an initial phase of short internal nodes, during which there would be insufficient time for changes in mitochondrial genome organization to become fixed between orders, followed by a longer period of external node expansion [35] in which enough time would pass for modifications to emerge independently along multiple branches within and between orders. With the inclusion of *M. australiensis*, whose genome does not share any of the alterations of *X. vesparum*, and is more representative of an ancestral insect arrangement, this also appears consistent for Strepsiptera: although greater sampling is required for a general picture of mt genome structural evolution to emerge.

The lengths of PCGs are truncated in both strepsipteran species, and especially so in *X. vesparum*. Across the 68 holometabolous dataset, *M. australiensis* has a reduction in mean content of 6.2 amino acids per gene. In *X. vesparum*, this reaches 11.8 amino acid deletions per gene. *X.*

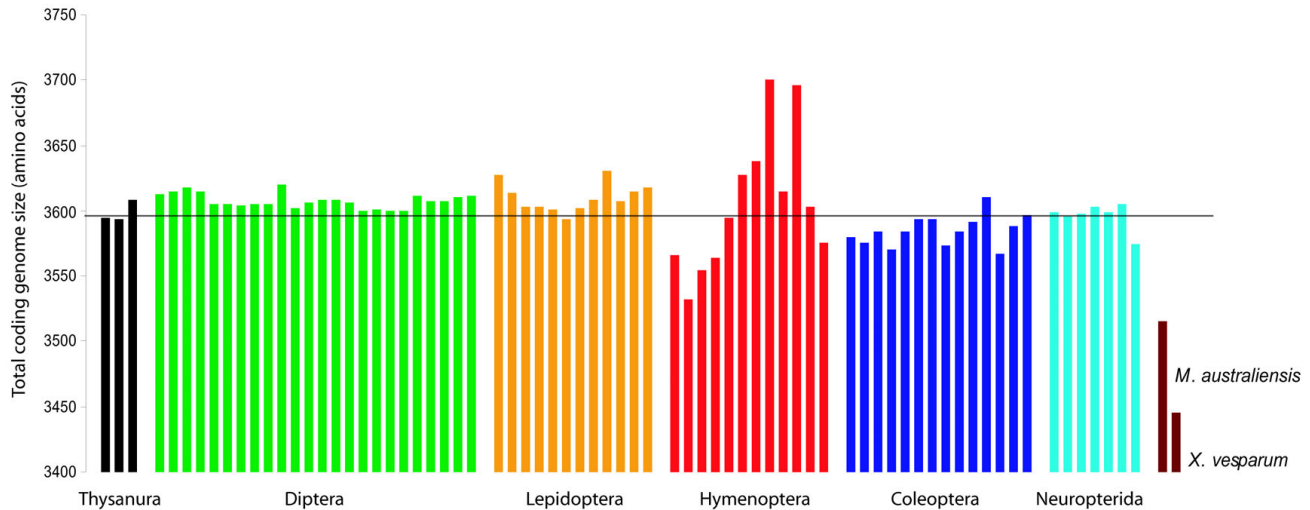


Figure 3

Comparison of the total mt coding genome size (in amino acids) across Holometabola. Black: Thysanura; Green: Diptera; Orange: Lepidoptera; Red: Hymenoptera; Blue: Coleoptera; Turquoise: Neuropterida; Burgundy: Strepsiptera. Each bar represents one species' genome. The superimposed line indicates the mean coding genome size across all 68 taxa.

Overall nucleotide composition is typically A+T rich in *M. australiensis* (84.3%; 6th highest amongst holometabolans insects), and like in *X. vesparum*, C-skew is high (+0.27). A scatter of the variation in skew and A+T% across the 68 holometabolans genomes (plus 3 thysanuran genomes) is shown in Figures 4A and 4B. Notably, in Figure 4B, of the most C-skewed and G+C% poor genomes (occupying the lower right quarter of this graph), 9 (of 12) data points are hymenopteran, 2 (of 2) are strepsipteran and 1 (of 12) is lepidopteran. A similar graph that including hemimetabolous insects [41] shows that only 1 other genome (*Schizaphis graminum*; Aphidoidea, 1 of 11 hemipteran genomes) is so C-skewed and A+T% rich. Conversely, the thysanuran genomes and a subset of Coleoptera are very A-skewed and distinctly more A+T% balanced (Figure 4A). This pattern is discussed in more detail in the following section.

Codon Usage

PCGs from Strepsiptera and the holometabolans dataset were imported into INCA v2.1 [42] to analyze codon usage. *M. australiensis* and *X. vesparum* are compared in Figure 5: codon bias is significantly relaxed in *X. vesparum* (2-tailed paired T-Test of ENC/MILC vectors (see methods); $P < 1 \times 10^{-5}$), and an apparent switch in threonine (T) preference from ACU to ACA has occurred. In Figure 6A, these levels are assessed in context, alongside the holometabolans dataset using the ENC and MILC measures of codon bias for individual PCGs (see methods; [43]). In *X. vesparum*, all 12 PCGs occupy a much more

relaxed, spread apart region of ENC/MILC space (the higher the value, the less codon bias), with average gene ENC and MILC values of 36.26 and 0.71 respectively. In *M. australiensis* these are 29.63 and 0.56, and in Holometabola, they average 32.77 (ENC) and 0.62 (MILC) over all genes. It follows that the values of codon bias among PCGs in *M. australiensis* should be found in the densest cluster of background holometabolans PCGs. The thysanuran genomes (empty black circles), alongside three coleopteran genomes (forming the isolated cluster of empty blue circles with MILC values > 1) have particularly balanced codon usage. Although beetles generally follow this rule; *Pyrophorus divergens*, *Tetraphalerus bruchi* and *Tribolium castaneum* in particular, have unusually high MILC values. Interestingly, unlike the thysanurans, these do not result in correspondingly elevated ENC values, possibly because of differences in the way MILC and ENC estimate bias [43].

These results are consistent with prevailing neutral mutational theories positing that genomic G+C content is the most significant factor in determining codon bias between organisms [44-47]. This may explain why *X. vesparum* has significantly relaxed ENC-MILC values, whose global G+C content is 5% higher than *M. australiensis*. For the thysanuran and beetle genomes, although G+C content in the 3rd codon position of individual PCGs is not necessarily elevated (Figure 6B), it is global G+C content that appears to matter: it has been documented that codon bias can be accurately predicted from inter-

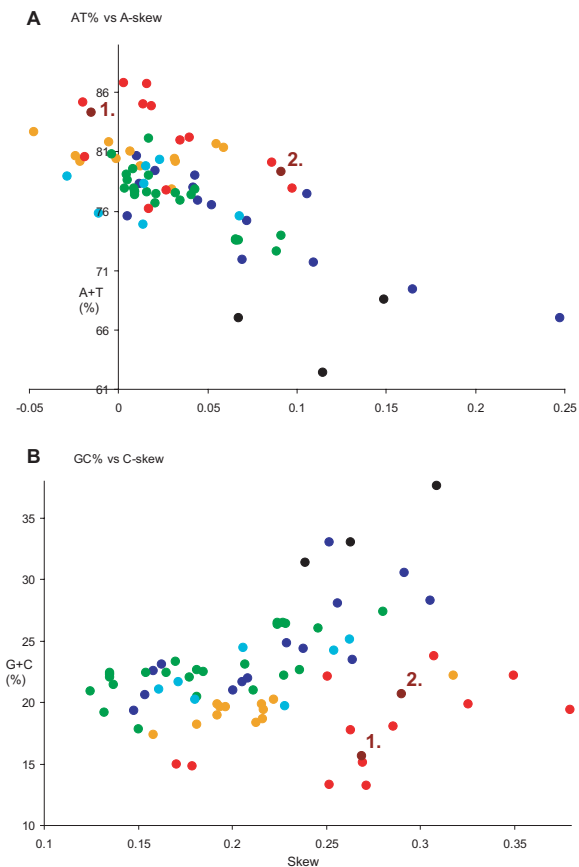


Figure 4
Graphical summary of nucleotide content across 68 holometabolan mt genomes. Measured in bp percentage (Y-axis) and level of nucleotide skew (X-axis) (described in methods, following [41]). Green: Diptera; Orange: Lepidoptera; Red: Hymenoptera; Blue: Coleoptera; Turquoise: Neuropterida; Burgundy: Strepsiptera. 1 and 2 refer to *M. australiensis* and *X. vesparum* respectively. A) A+T% vs A-skew B) G+C% vs C-skew. 3 genomes from Thysanura (in black), the host group for *M. australiensis* are also shown (see methods).

genic regions [44], and it is unsurprising that these genomes should also have the highest genome wide G+C content across the 68 Holometabola dataset (Additional file 2). It is noticeable that the coleopteran *T. bruchi* has a very A-skewed genome (Figure 4A), which could also partly explain some of the discrepancies appearing between the different methods.

Secondary structure of ribosomal and transfer RNAs

With the addition of *M. australiensis*, a comparative revision of strepsipteran mitochondrial rRNA secondary structure was possible. Figures 7 and 8 are complete models for the *rrnL* in *M. australiensis* and *X. vesparum* respec-

tively. The addition of new comparative evidence (from *M. australiensis* and *Apis mellifera* [48]) since Carapelli *et al.* (2006, [28]) enables some improvement over the structural model for *X. vesparum*. These modifications are highlighted in blue. Regions in which sequence variability is still too high for good structural covariation are highlighted in red. Overall, *M. australiensis* predictions are largely consistent with current insect consensus models but substantial portions of secondary structure remain problematic in *X. vesparum*. For example, the helix at the base of domain IV in *M. australiensis* is supported by comparative evidence from *Drosophila melanogaster* and *Apis mellifera*. In *X. vesparum*, most of this primary sequence in this helix is highly divergent and covarying base-pairing cannot be identified. Where this helix terminates in the single-strand bulge however, the base pairing AUG-UAU is supported by *M. australiensis* (AAG-UUU) which provides additional corroborative support for the bps in *D. melanogaster* and *A. mellifera*, which have, AAG-UUU and AGG-UCU respectively.

rrnS secondary structural models for *M. australiensis* and *X. vesparum* are presented in Figure 9. Absence of sequence across domain I precludes a complete analysis, but partial models for domains II and III are generally conserved across Strepsiptera. Regions highlighted in red are more labile, and adequate consensus models for these structures are still lacking. The structural predictions for 18 tRNA genes are given in Figure 10. There is less scope for variation within extremely length-constrained tRNA genes. Despite this, the T Ψ P stem is still too variable to be useful in a deep comparative framework: even between *M. australiensis* and *X. vesparum*, base-pairings demonstrate little covariation. The three remaining stem-loop structures are more conserved across Strepsiptera, and more useful in a comparative structural context. Within these regions, *X. vesparum* demonstrates more substantial modifications to typical insect tRNA structure than *M. australiensis* - consisting largely of single base or base-pair substitutions/compensations. For example, the proximal base-pairs of the acceptor stem of tRNA-Pro (P) in *X. vesparum* appears as UCAG-CUGA. In *M. australiensis*, Lepidoptera (*Ochrogaster lunifer* [41]) and Hymenoptera (*Vanhornia eucnemidarum* [39]) it is CAAA-UUUG. Similarly, the proximal DHU stem consists of a canonical A-U base-pairing in *X. vesparum*, where in *M. australiensis* and other insect groups it appears as a non-canonical G-U. Further, the base-pairing C-G, found across disparate insect lineages is substituted by A-U in the proximal stems of the tRNA-Lys (K) acceptor and the tRNA-Glu (E) DHU helices of *X. vesparum*.

Conclusions

The mitochondrial genome of *X. vesparum* displays a number of characteristics that are not shared by its distant

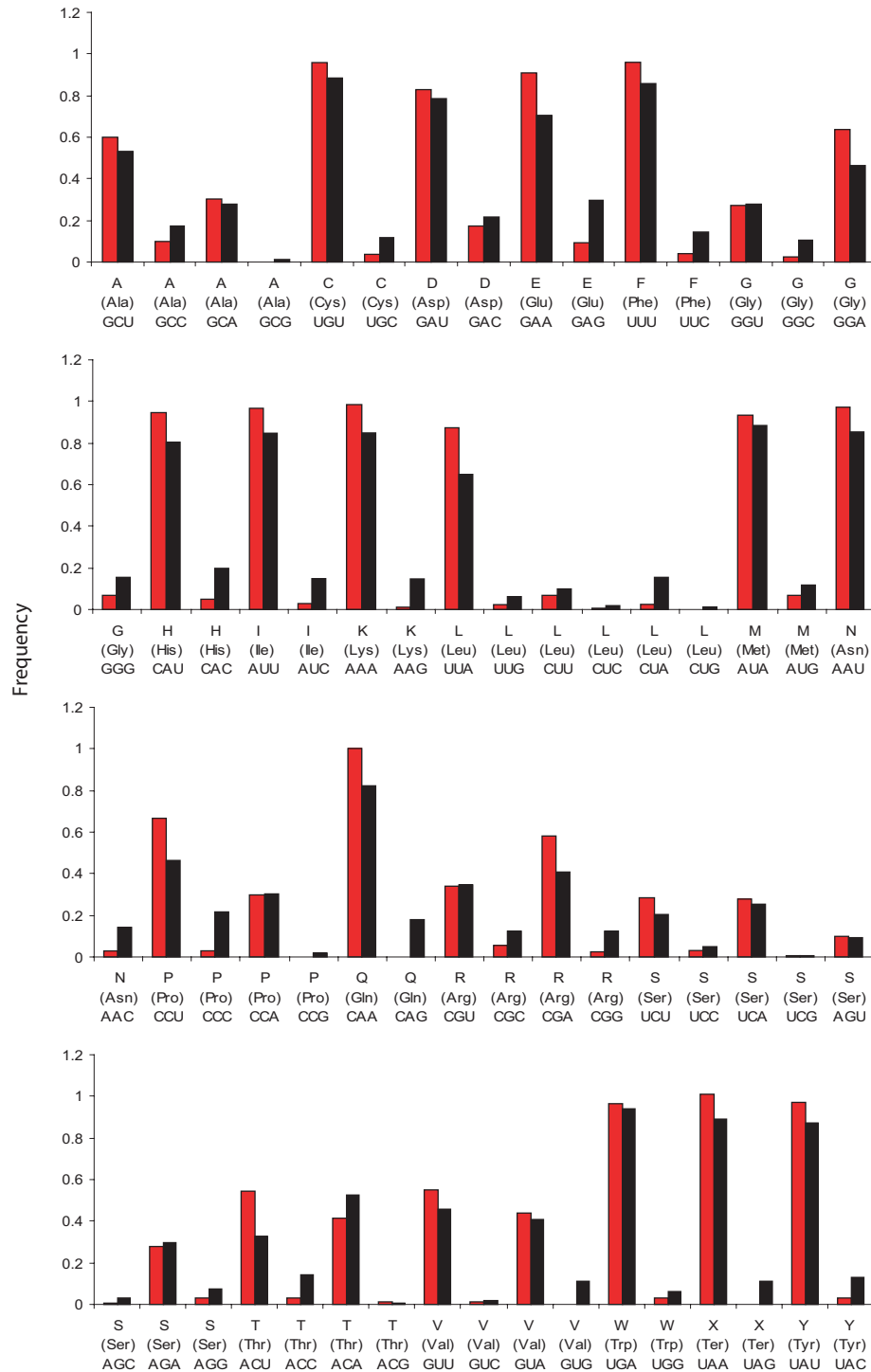


Figure 5
Codon usage in *M. australiensis* (red) and *X. vesparum* (black). Codon frequencies are given (out of 1) for individual amino acids.

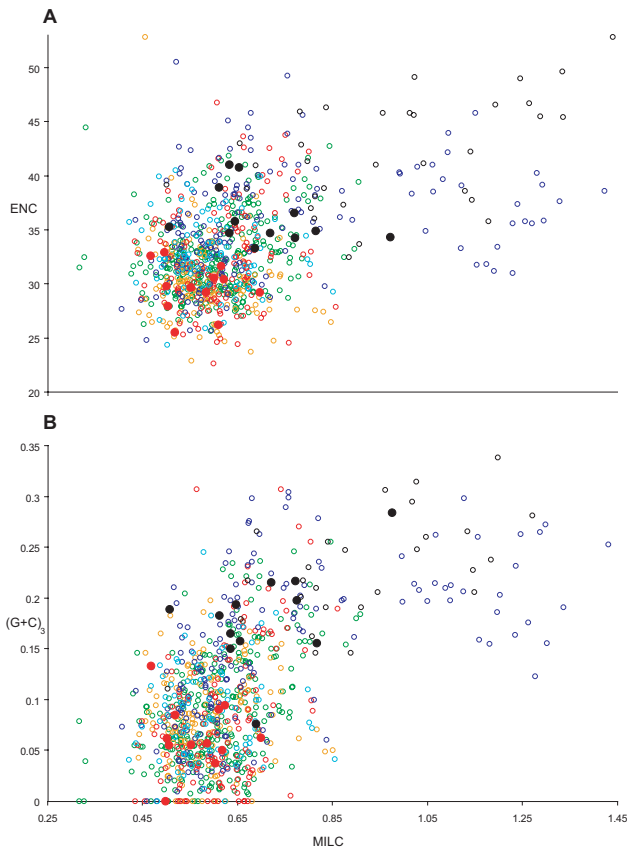


Figure 6
Evaluation of codon bias across Holometabola. A) Codon bias in *M. australiensis* (red filled circle) and *X. vesparum* (black filled circle) measured using ENC and MILC indices (described in methods), are compared against remaining Holometabola, whose data points are given as empty circles: the colour scheme follows taxonomic group as in Figure 3 and 4. Data points refer to individual PCGs. B) MILC values are compared against $(G+C)_3$; which is G+C content at the 3rd codon position of each PCG. This serves as an approximate measure of background nucleotide composition across individual PCGs.

relative *M. australiensis*. In the latter, only 2 tRNA translocations disrupt an otherwise ancestral insect genome organization, whereas in *X. vesparum*, 3 tRNA translocations and at least 2 duplications have occurred (Figure 2). Codon bias in *X. vesparum* is also significantly relaxed: *M. australiensis* occupies a much more typical region of ENC/MILC space (Figures 5 and 6). Further, rRNA secondary structural model predictions in *M. australiensis* are much more consistent with current insect consensus structures (Figures 7, 8, 9 and 10). In *X. vesparum*, increased variability at the primary sequence level precludes the assembly of well corroborated models in several parts of the rRNA

genes. Both strepsipteran genomes are unusually C-skewed and A+T% rich. Interestingly, PCGs are noticeably truncated in both strepsipteran taxa, but especially so in *X. vesparum* which has lost nearly twice as many codons (154) as *M. australiensis* (81), equating to 4.3% and 2.2% fewer codons than the average holometabolan genome respectively.

These results provide useful insight into the evolution of strepsipteran mitochondria spanning the shift from a largely free-living insect into an extremely unique and simplified endoparasitoid, since the emergence of Strepsiptera over 100 million years ago [49]. The modifications to the mitochondrial genomes of *X. vesparum* and *M. australiensis* appear consistent with organisms that have evolved extremely derived lifestyles. *M. australiensis* represents a transitional phase in the evolutionary specialization of Strepsiptera, and the composition, architecture and structure of its genome reflects this. These observations raise important questions about the changing selective environment in (mitochondrial) genomes belonging to small, and almost entirely host-dependent endoparasitoids. Future investigations into the evolutionary and developmental origins of this unique biological system must ensure that strepsipteran species from Mengenillidae are adequately sampled: wide sampling across this group is crucial if the underlying ancestral signal is to be maximized.

Methods

Specimen description

DNA was used from a small set of free-flying males collected in light traps, leaving little possibility for contamination through thysanuran host DNA. Nuclear and mitochondrial gene alignments from a widely sampled Strepsiptera taxon set (unpublished) confirm this. We also include 3 thysanuran genomes in Figure 8 to emphasize the genome composition differences between *M. australiensis* and its target host group. Specimens were collected on the 16th March 2006 in Australia, Queensland, Blackdown Tablelands NP, South Mimosa Creek, 50 m dstr road, 794 mao. Coordinates as follows: S23 47.687'E 149 04. 195'. Light trap loc 34 (Collectors were N. Jönsson, T. Malm & D. Williams). Complete species name: *Mengenilla australiensis* Kifune & Hirashima 1983.

Genome isolation and sequencing

M. australiensis specimens were extracted for PCR either by macerating tissue and digesting overnight at 50°C with Chelex 100 grade resin (Bio-Rad Cat. no 142-1253, Hemel Hempstead, UK) and enzyme Proteinase K (BIO-LINE, Cat. no BIO-37037, London, UK) or by employing a QIAGEN blood and tissue column extraction protocol. Amplification of the *M. australiensis* mitochondrion was carried out by amplifying three ~3 kb fragments, with

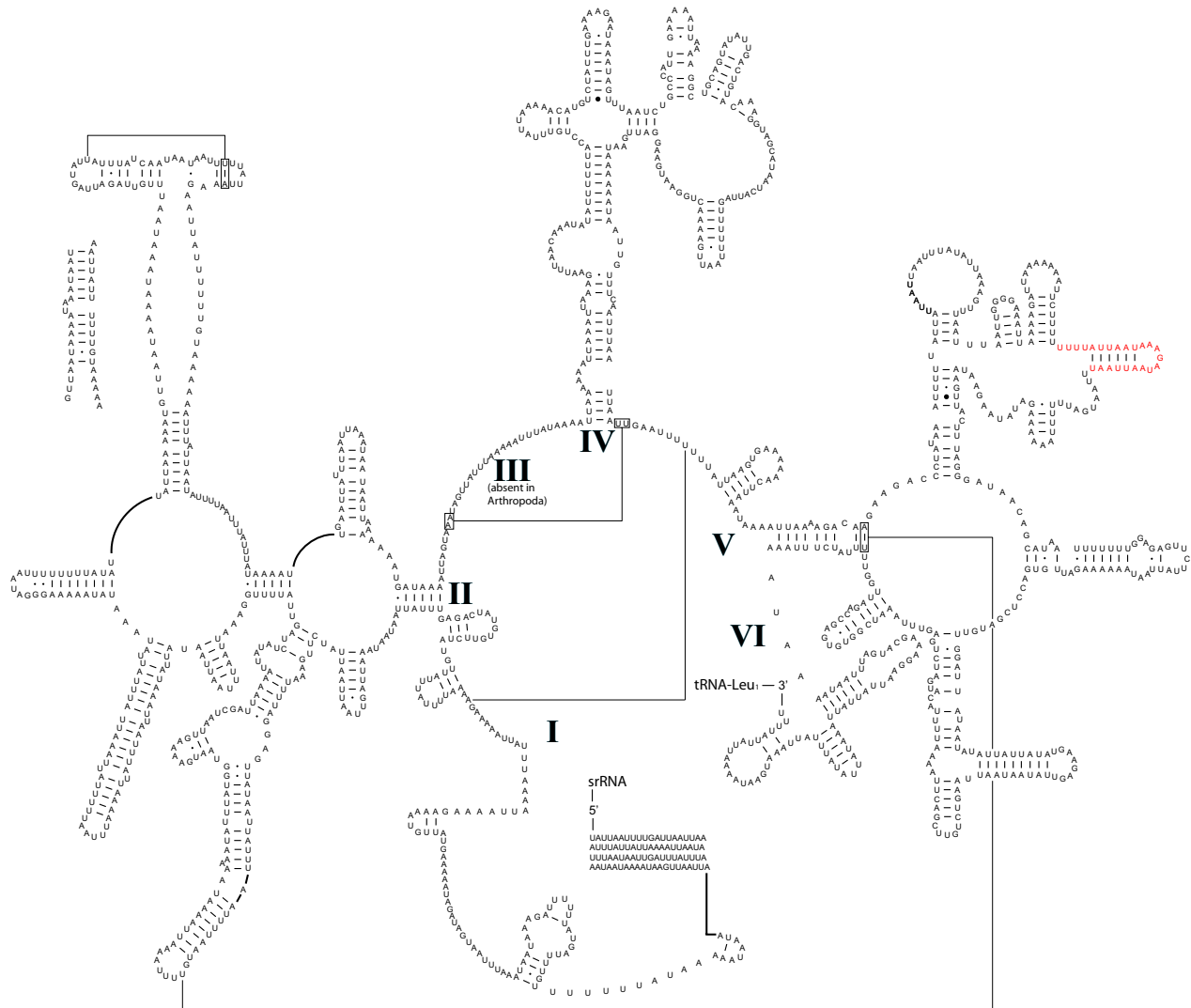


Figure 7
Predicted secondary structure of the *rrnL* in *M. australiensis*. Regions in red receive little comparative support and remain ambiguous. Canonical A-U and G-C base pairings are given as '-'. G-U, G-A, and other non-canonical bp combinations are denoted as, '•', 'm' and 'l' respectively. Predicted tertiary interactions are represented by a line. Missing data are coded as 'N'.

intervening sections that did not overlap being amplified by standard PCR. Long PCR mixes containing 1 µl BIO-X-ACT long DNA polymerase (BIOLINE, Cat. no. BIO-21049), 1 µl dNTPs, 25 µl 2× Polymate additive (BIOLINE, Cat. no. BIO-37041) for A+T rich sequences, 5 µl 10× opti-buffer, 4 µl MgCl₂ solution, 3 µl forward/reverse primers and 5-10 µl of genomic DNA were made up to 50 µl reactions with ultrapure (NANOpure Diamond™) water. Long PCR cycling parameters were as follows: 92 °C for 2 minutes; 10 cycles of: 92 °C for 10 seconds, 53 °C for 30 seconds and 68 °C for 13 minute; 28 cycles of: 92 °C for

15 seconds, 53 °C for 30 seconds and 68 °C for 14 minute; with an extension step of 7 minutes at 68 °C. Long PCR fragments were gel extracted (QIAGEN Cat no. 28704, Hemel Hempstead, UK) and used as templates in subsequent sequencing reactions. Standard PCR mixes contained 0.5 µl *Taq* polymerase, 1 µl dNTPs, 5 µl MgCl₂ free buffer, 4 µl MgCl₂ solution, 3 µl forward/reverse primers and 2 µl of genomic DNA, made up to 50 µl reactions with ultrapure (NANOpure Diamond™) water. Standard PCR cycling parameters were as follows: 94 °C for 2 minutes, then 35 cycles of 94 °C for 1 minute, T °C for 45 seconds

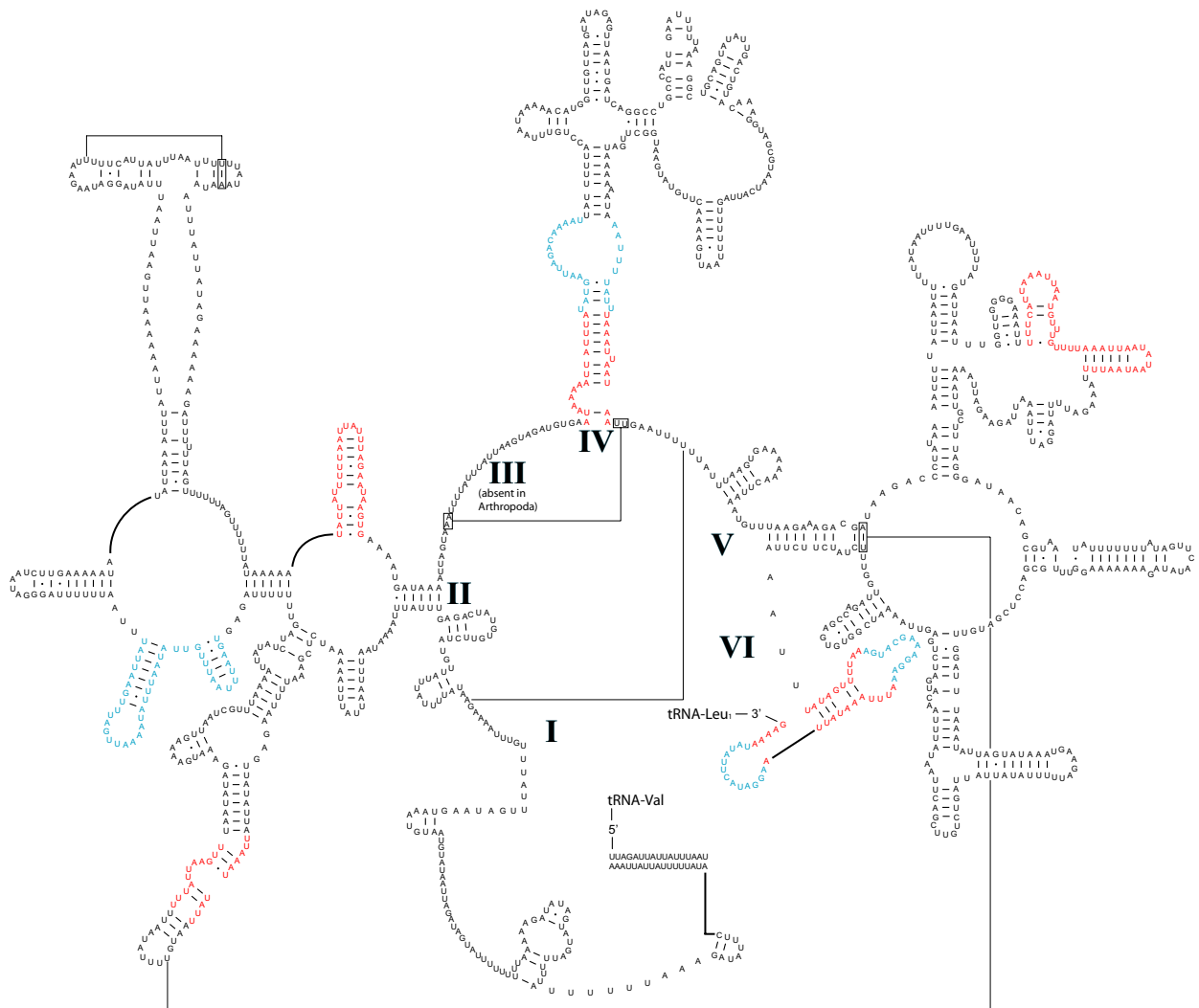


Figure 8
Revised secondary structural prediction of the *rrnL* in *X. vesparum*. Regions in red receive little comparative support and remain ambiguous; regions in blue are structural predictions which have improved with the inclusion of comparative data from *M. australiensis*. Structural annotations follow figure 7.

(*T* varied with primer pair) and 72°C for 1 minute, with an extension step of 5 minutes at 72°C. For difficult regions, Phusion high fidelity *Taq* was utilized, following the manufacturer's instructions. Standard PCR products were cleaned using 2 µl of Shrimp Alkaline Phosphatase (SAP) and 3 µl of Exonuclease I (in a 1:10 SAP buffer dilution) per 50 µl PCR reaction, or fragments were extracted from agarose gels and sequenced directly. Problematic fragments that did not sequence directly were inserted into pGEM-T Easy plasmid vectors (Promega, Southampton, UK), and multiple clones were sequenced for each fragment. Sequencing reactions were performed using the BigDye® Terminator v3.1 cycle sequencing kit with the fol-

lowing modifications: 1 µl of BigDye®, 1.5 µl 5× Buffer, 1 µl primer (3.2 pmol), 1-3 µl genomic DNA made up to 10 µl reactions with ultrapure (NANOpure Diamond™) water. Sequence reads were generated using an Applied Biosystems 3730xl DNA Analyzer. Sequences were imported into GridinSoft Notepad (Lite Edition; <http://notepad.gridinsoft.com>) and BioEdit (version 7.0.5.3; [50]) for analysis. Primer pairs employed are given in Additional file 3. The annotated genome fragment can be found in Genbank under the accession number [GU188852](http://www.ncbi.nlm.nih.gov/Genbank/CG1188852). Voucher specimens and extractions are deposited in the Hope Entomological Collections, Oxford University Museum of Natural History, Oxford, UK, and

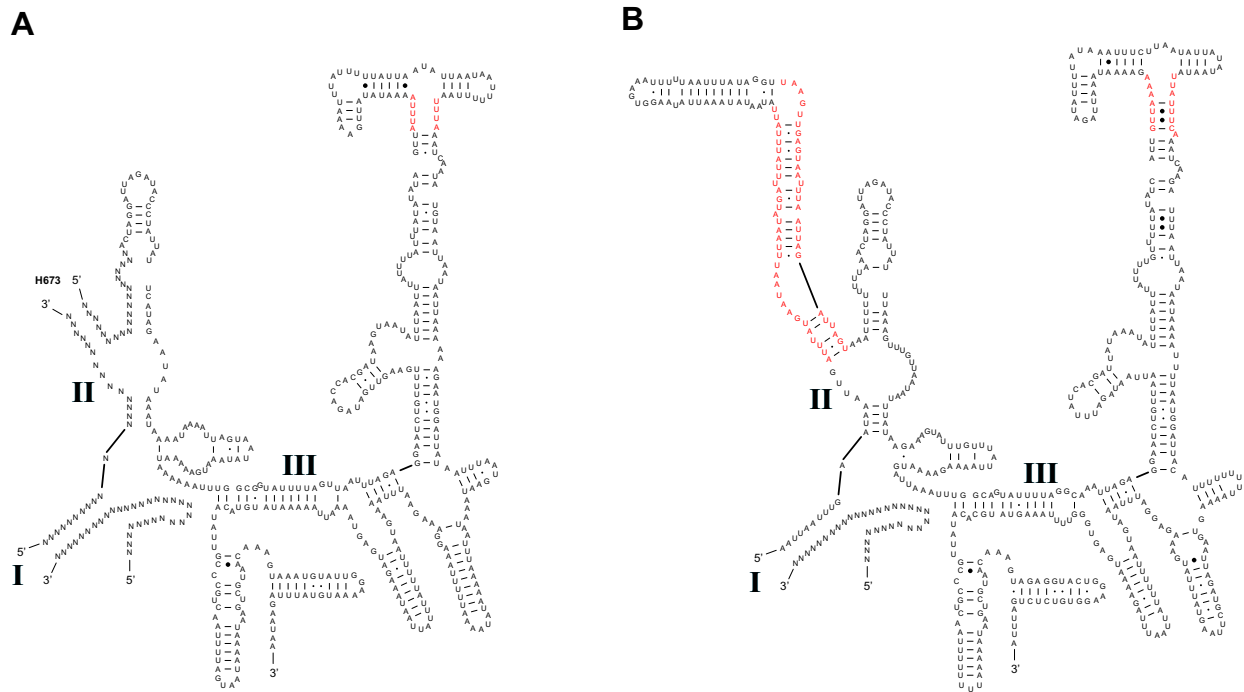


Figure 9

Secondary structural predictions for the *rrnS* in Strepsiptera. A) *M. australiensis*, B) *X. vesparum*. Regions in red receive little comparative support and remain ambiguous. Structural annotations follow figure 7.

the Swedish Museum of Natural History, Stockholm, Sweden.

Genome characterization

PCGs and their boundaries were diagnosed via sequence comparison across alignments derived from the 68 holometabolite genome dataset. Start codons found to be in-frame and not overlapping with upstream genes on the same strand were usually identified as 5' gene boundaries. Stop codons were typically identified as TAA, TAG, TA, or T. A complete annotation of the *M. australiensis* genome is given in Additional file 4. One non-coding fragment, 24 bp in length, was identified between the stop codon of *nad1* and the beginning of tRNA-Ser₂. Secondary structural predictions for ribosomal and transfer RNAs were ascertained by implementing a comparative structural framework [51-55] (for an online tutorial see the jRNA web site: <http://hymenoptera.tamu.edu/rna/>), in conjunction with a thermodynamic-based RNA folding algorithm (*mfold*; [56]). *mfold* gave initial raw estimates of structure: a tool that is especially useful for variable domains with highly divergent primary sequence. In this approach, structures with the lowest thermodynamic stability values were most highly considered (although structures with only marginal thermodynamic differences were also eval-

uated) and robust comparative evidence where available, took priority. The following ribosomal RNA structural templates from the Comparative RNA Web (CRW) site database [57] were used: *Escherichia coli*, *Drosophila melanogaster*, *Drosophila virilis*, *Chorthippus parallelus* and *Apis mellifera* [48]. For tRNA genes, the following models were used as comparative data points: *X. vesparum* (Strepsiptera: [28]), *O. lunifer* (Lepidoptera: [41]), *V. eucnemidarum* (Hymenoptera: [39]) and *Cryptopygus antarcticus* (Collembola: [58]), structures were verified for accuracy in tRNA-Scan [59]. Gene order was assessed across Holometabola using the inferred ancestral insect arrangement as a comparative standard [29,30]. As a measure of gene length fluctuation, total and individual gene length was counted across strepsipteran PCGs, and compared across the dataset of 68 holometabolite genomes. To include *nad2* gene, the 5' portion (approx. 100 amino acids) was excluded from analysis so that *M. australiensis* could be assessed alongside the entire dataset.

The 68 holometabolite mitochondrial genome dataset was compiled from previously published data across the 6 major holometabolite divisions: Diptera [60-74], Lepidoptera [33,75-81], Hymenoptera [39,82-87], Coleoptera [88-93], Neuropterida [34,94,95] and Strepsiptera [28].

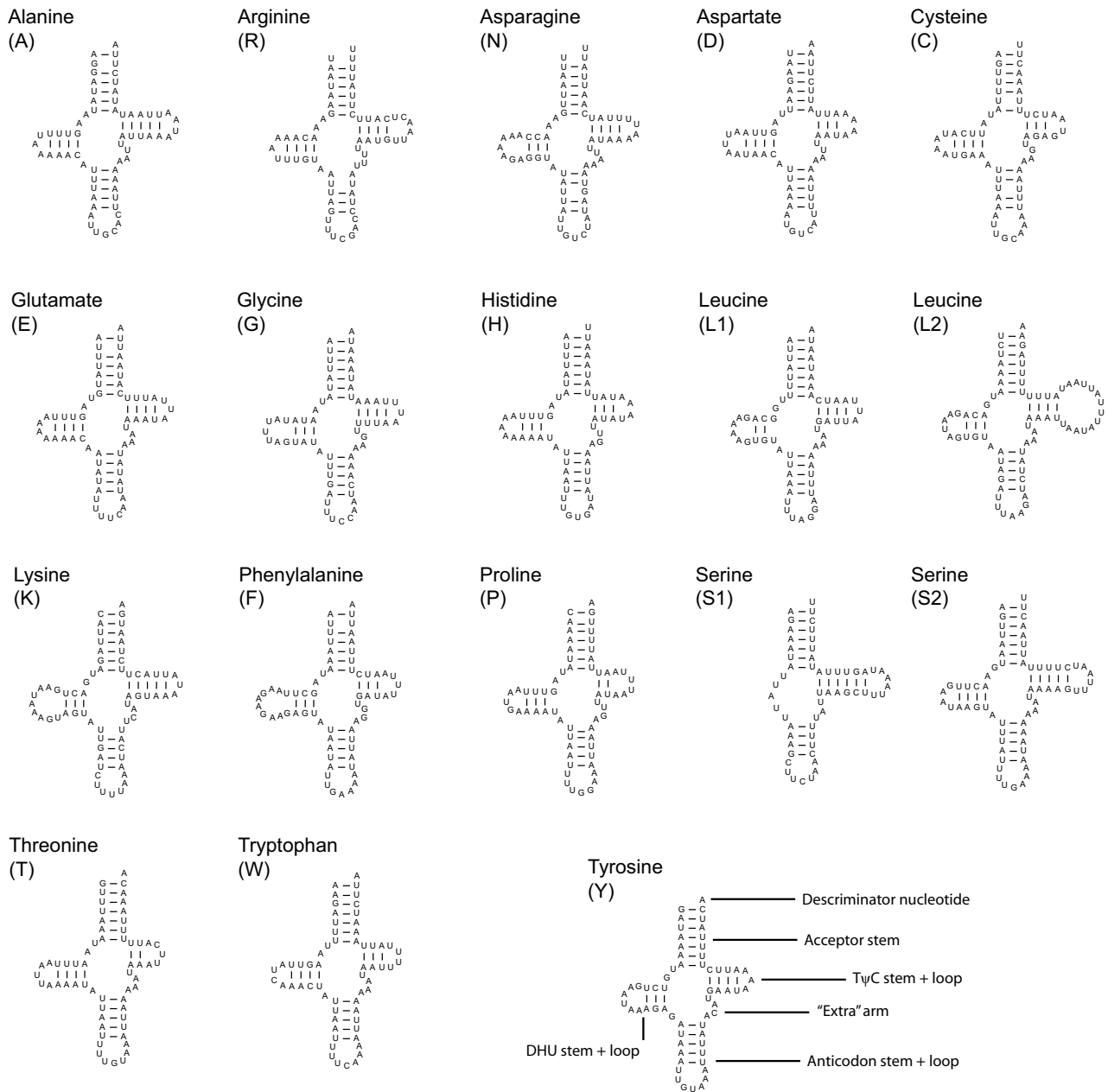


Figure 10
Secondary structural predictions for *M. australiensis* tRNA genes. Pertinent aspects relating to comparative structural differences between *X. vesparum* [28] and other insect lineages [39,41,58] are discussed in the main text.

Nucleotide content and Codon usage

A+T% and G+C% values were calculated for the α strand of the *M. australiensis* genome, which runs clockwise in the direction of transcription from *nad2* (Figure 1). Genomic A- and G-skew measures were calculated using total nucleotide % values in the following manner: $[A-T]/[A+T]$ and $[G-C]/[G+C]$ following Perna & Kocher [96]. A complete list of nucleotide content values for the 68

holometabolan dataset is given in Additional file 3. Codon usage was investigated by importing the complete cDNA datasets of 68 available holometabolan mitochondrial genomes, including *X. vesparum* and the sequenced genome of *M. australiensis* into INCA v2.1 [42]. Two independent measures of codon bias: ENC (Effective Number of Codons used) and MILC (Measure Independent of Length and Composition) are implemented, whose statis-

tical justification and behaviour under simulation differ markedly [43]. These two independent measures were combined as a single vector statistic as follows: $\sqrt{(\text{ENC}^2 + \text{MILC}^2)}$. The shortest gene, *atp8* was also excluded from analyses due to artificially increased ENC/MILC bias in genes shorter than around 100 amino acids. *nad4L* was of borderline length (approx. 90-105 amino acids) and was retained, although *nad4L* from *Drosophila sechellia* and *Melipona bicolor* produced arbitrary ENC 'cut-off' values of 61, because ENC behaves less stably at borderline gene lengths; these data points were removed. Readers are referred to [43] for a detailed explanation.

Abbreviations

mt: mitochondrial; PCG: protein-coding gene; A+T region: the putative control region; α strand: leading strand of transcription; ENC: Effective Number of Codons used; MILC: Measure Independent of Length and Composition; *rrnL* and *rrnS*: large (16S) and small (12S) subunit ribosomal RNA (rRNA) gene; tRNA genes are denoted as single letter amino acid IUPAC-IUB abbreviations; *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; MRCA: Most Recent Common Ancestor.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

DPM and AH devised and completed laboratory work. DPM conducted analyses and early manuscript draughts. All authors contributed to the final version of the manuscript.

Additional material

Additional file 1

Coding gene lengths. Spreadsheet containing coding gene lengths (in amino acids) across the Holometabola dataset.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-603-S1.XLS>]

Additional file 2

Nucleotide content. spreadsheet of genomic nucleotide content, A-skew and C-skew across the Holometabola dataset.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-603-S2.XLS>]

Additional file 3

Primers. List of primer pairs used in the study.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-603-S3.XLS>]

Additional file 4

Gene annotation. Detailed summary of genome fragment architecture.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-603-S4.XLS>]

Acknowledgements

We owe gratitude to Niklas Jönsson, Tobias Malm and Dawn Williams for collection of material, to Clive Cook Director of the Queensland Parks and Wildlife Service - Northern Region for the Collecting Permit (# WITK03408905), and to B. Viklund (Swedish Museum of Natural History, Stockholm, Sweden) for donation of material. Thanks must also be extended to Peter Holland for useful discussions, the Leverhulme Trust (grant no. F/08 502G) for funding to JK, and the Elizabeth Hannah Jenkinson fund (University of Oxford) for an award to DPM.

References

- Kathirithamby J: **Host-parasitoid associations in Strepsiptera.** *Annu Rev Ent* 2009, **54**:227-249.
- Crowson RA: **The phylogeny of Coleoptera.** *Annu Rev Ent* 1960, **5**:111-134.
- Crowson RA: **The Biology of the Coleoptera.** Academic Press, New York; 1981.
- Whiting MF, Wheeler WC: **Insect homeotic transformation.** *Nature* 1994, **368**:696.
- Kristensen NP: **Forty years' insect phylogenetics systematics.** *Zool Beitr* 1995, **36**:83-124.
- Whiting MF, Carpenter JC, Wheeler QD, Wheeler WC: **The Strepsiptera Problem: Phylogeny of the Holometabolous Insect Orders Inferred From 18S and 28S Ribosomal DNA sequences and Morphology.** *Syst Biol* 1997, **46**:1-68.
- Huelsenbeck JP: **Is the Felsenstein Zone a fly trap?** *Syst Biol* 1997, **46**:69-74.
- Huelsenbeck JP: **Systematic bias in phylogenetic analysis: is the Strepsiptera problem solved?** *Syst Biol* 1998, **47**:519-537.
- Hwang UW, Kim W, Tautz D, Friedrich M: **Molecular phylogenetics at the Felsenstein zone: approaching the Strepsiptera problem using 5.8S and 28S rDNA sequences.** *Mol Phyl Evol* 1998, **9**:470-480.
- Huelsenbeck JP: **A Bayesian perspective of the Strepsiptera problem.** *Tidschr Ent* 2001, **144**:165-178.
- Whiting MF: **Long branch distractions and the Strepsiptera.** *Syst Biol* 1998, **47**:134-138.
- Hünefeld F, Beutel RG: **The sperm pumps of Strepsiptera and Antliophora (Hexapoda).** *J Zool Syst Evol Res* 2005, **43**:297-306.
- Wiegmann BM, Trautwein MD, Kim JW, Cassel BK, Bertone MA, Winterton SL, Yeates DK: **Single-copy nuclear genes resolve the phylogeny of the holometabolous insects.** *BMC Biology* 2009, **7**:34.
- Beutel RG, Pohl H: **Endopterygote systematics - where do we stand and what is the goal (Hexapoda, Arthropoda)?** *Syst Ent* 2006, **31**:202-219.
- Kinzelbach R: **The systematic position of Strepsiptera (Insecta).** *Am Entomol* 1990, **35**:292-303.
- Pohl H, Beutel RG, Kinzelbach R: **Protoxenidae fam. nov. (Insecta, Strepsiptera) from Baltic amber - a 'missing link' in Strepsipteran phylogeny.** *Zool Scr* 2005, **34**:57-69.
- Kinzelbach RK: **Strepsiptera.** *Die Tierwelt Deutschlands* 1978, **65**:166.
- Kathirithamby J: **Review of the order Strepsiptera.** *Syst Ent* 1989, **14**:41-62.
- Kathirithamby J: **Strepsiptera.** In *The Insects of Australia: A textbook for students and research workers* 2nd edition. Edited by: Naumann ID, Carne PB, Lawrence JF, Nielson ES, Spradberry JP, Taylor RW, Whittten MJ, Littlejohn MJ. Melbourne: Melbourne University Press; 1993:684-695.
- Pohl H, Beutel RG: **The phylogeny of Strepsiptera.** *Cladistics* 2005, **21**:328-374.

21. Pohl H, Beutel RG: **The evolution of Strepsiptera.** *Zoology* 2007, **111**:318-338.
22. Harrison RG: **Animal mitochondrial DNA as a genetic marker in population and evolutionary biology.** *Trends Ecol Evol* 1989, **4**:6-11.
23. Zhang D-X, Hewitt GM: **Nuclear integrations: challenges for mitochondrial DNA markers.** *Trends Ecol Evol* 1996, **11**:247-251.
24. Perna NT, Kocher TD: **Mitochondrial DNA: molecular fossils in the nucleus.** *Curr Biol* 1996, **6**:128-129.
25. Pons J, Vogler AP: **Complex Pattern of Coalescence and Fast Evolution of a Mitochondrial rRNA Pseudogene in a Recent Radiation of Tiger Beetles.** *Mol Biol Evol* 2005, **22**:991-1000.
26. Song H, Buhay JE, Whiting MF, Crandall KA: **Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified.** *P Natl Acad Sci USA* 2008, **105**:13486-13491.
27. Wolstenholme DR: **Animal mitochondrial DNA: structure and evolution.** *Int Rev Cytol* 1992, **141**:173-216.
28. Carapelli A, Vannini L, Nardi F, Boore JL, Beani L, Dallai R, Frati F: **The mitochondrial genome of the entomophagous endoparasite *Xenos vesparum* (Insecta: Strepsiptera).** *Gene* 2006, **376**:248-259.
29. Boore JL, Collins TM, Stanton D, Daehler LL, Brown WM: **Deducing the pattern of arthropod phylogeny from mitochondrial DNA rearrangements.** *Nature* 1995, **376**:163-165.
30. Boore JL, Lavrov D, Brown WM: **Gene translocation links insects and crustaceans.** *Nature* 1998, **393**:667-668.
31. 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.
32. Downton M, Cameron SL, Dowavic JJ, Austin AD, Whiting MF: **Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral.** *Mol Biol Evol* 2009, **26**:1607-1617.
33. 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.
34. Cameron SL, Sullivan J, Song H, Miller KB, Whiting MF: **A mitochondrial genome phylogeny of the Neuropterida (lace-wings, alderflies and snakeflies) and their relationship to the other holometabolous insect orders.** *Zool Scr* 2009, **38**:575-590.
35. Rokas A, Carroll SB: **Bushes in the Tree of Life.** *PLOS Biol* 2006, **4**:1899-1904.
36. Maeta Y, Takahashi K, Shimada N: **Host body size as a factor determining the egg complement of Strepsiptera, an insect parasite.** *Int J Insect Morphol Embryol* 1998, **27**:27-37.
37. Castro LR, Austin AD, Downton M: **Contrasting rates of mitochondrial molecular evolution in parasitic Diptera and Hymenoptera.** *Mol Biol Evol* 2002, **19**:1100-1113.
38. Eggleton P, Belshaw R: **Insect parasitoids: an evolutionary overview.** *Philos Trans R Soc Lond* 1992, **337**:1-20.
39. Castro LR, Ruberu K, Downton M: **Mitochondrial genomes of *Vanhornia eucnemidarum* (Apocrita: vanhorniidae) and *Primeuchroes* spp. (Aculeata: chrysididae): evidence of rearranged mitochondrial genomes within the Apocrita (Insecta: Hymenoptera).** *Genome* 2006, **49**:752-766.
40. Downton M, Austin AD: **Increased genetic diversity in mitochondrial genes is correlated with the evolution of parasitism in the Hymenoptera.** *J Mol Evol* 1995, **41**:958-965.
41. Salvato P, Simonato M, Battisti A, Negrisolo E: **The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae).** *BMC Genomics* 2008, **9**:331.
42. Supek F, Vlahovicek K: **INCA: synonymous codon usage analysis and clustering by means of self-organizing map.** *Bioinformatics* 2004, **20**:2329-2330.
43. Supek F, Vlahovicek K: **Comparison of codon usage measures and their applicability in prediction of microbial gene expressivity.** *BMC Bioinformatics* 2005, **6**:182.
44. Chen SL, Lee W, Hottes AK, Shapiro L, McAdams HH: **Codon usage between genomes is constrained by genome-wide mutational processes.** *Proc Natl Acad Sci USA* 2004, **101**:3480-3485.
45. Kanaya S, Kinouchi M, Abe T, Kudo Y, Yamada Y, Nishi T, Mori H, Ikemura T: **Analysis of codon usage diversity of bacterial genes with a self-organizing map (SOM): characterization of horizontally transferred genes with emphasis on the *E. coli* O157 genome.** *Gene* 2001, **276**:89-99.
46. Knight RD, Freeland SJ, Landweber LF: **A simple model based on mutation and selection explains trends in codon and amino acid usage and GC composition within and across genomes.** *Genome Biol* 2001, **2**:RESEARCH0010.
47. Herschberg R, Petrov DA: **Selection on codon bias.** *Annu Rev Genet* 2008, **42**:287-299.
48. Gillespie JJ, Johnston JS, Cannone JJ, Gutell RR: **Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements.** *Insect Mol Biol* 2006, **15**:657-686.
49. Grimaldi D, Kathirithamby J, Schawaroch V: **Strepsiptera and tringula in Cretaceous amber.** *Insect Syst Evol* 2005, **36**:1-20.
50. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucl Acids Symp Ser* 1999, **41**:95-98.
51. Kjer KM, Baldrige GD, Fallon AM: **Mosquito large subunit ribosomal RNA: simultaneous alignment of primary and secondary structure.** *Biochim Biophys Acta* 1994, **1217**:147-155.
52. Kjer KM: **Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from the frogs.** *Mol Phylogenet Evol* 1995, **4**:314-330.
53. Gillespie JJ: **Characterizing regions of ambiguous alignment caused by the expansion and contraction of hairpin-stem loops in ribosomal RNA molecules.** *Mol Phylogenet Evol* 2004, **33**:936-943.
54. Gillespie JJ, McKenna CH, Yoder MJ, Gutell RR, Johnston JS, Kathirithamby J, Cognato AI: **Assessing the odd secondary structural properties of nuclear small subunit ribosomal RNA sequences (18S) of the twisted-wing parasites (Insecta: Strepsiptera).** *Insect Mol Biol* 2005, **14**:625-643.
55. Gillespie JJ, Yoder MJ, Wharton RA: **Predicted secondary structures for 28S and 18S rRNA from Ichneumonidae (Insecta: Hymenoptera: Apocrita): Impact on sequence alignment and phylogeny estimation.** *J Mol Evol* 2005, **61**:114-137.
56. Zuker M: **Mfold web server for nucleic acid folding and hybridization prediction.** *Nucleic Acids Res* 2003, **31**:3406-3415.
57. Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, Feng B, Lin N, Madabusi LV, Müller KM, Pandey N, Shang Z, Yu N, Gutell RR: **The Comparative RNA Web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs.** *BMC Bioinformatics* 2002, **3**:2. Correction: *BMC Bioinformatics*. **3**:15.
58. Carapelli A, Comandi S, Convey P, Nardi F, Frati F: **The complete mitochondrial genome of the Antarctic springtail *Cryptopygus antarcticus* (Hexapoda: Collembola).** *BMC Genomics* 2008, **9**:315.
59. 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.
60. Lessinger AC, Junqueira ACM, Lemos TA, Kemper EL, da Silva FR, Vettore AL, Arruda P, Azeredo-Espin AML: **The mitochondrial genome of the primary screwworm fly *Cochliomyia hominivorax* (Diptera: Calliphoridae).** *Insect Mol Biol* 2000, **9**:521-529.
61. Ballard JWO: **Comparative genomics of mitochondrial DNA in *Drosophila simulans*.** *J Mol Evol* 2000, **51**:64-75.
62. 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.
63. 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.
64. Oliveira MT, Barau JG, Junqueira ACM, Feijao PC, da Rosa AC, Abreu CF, Azeredo-Espin AML, Lessinger AC: **Structure and evolution of the mitochondrial genomes of *Haematobia irritans* and *Stomoxys calcitrans*: The Muscidae (Diptera: Calyptratae) perspective.** *Mol Phylogenet Evol* 2008, **48**:850-857.

65. Stevens JR, Westr H, Wall R: **Mitochondrial genomes of the sheep blowfly, *Lucilia sericata*, and the secondary blowfly, *Chrysomya megacephala*.** *Med Vet Entomol* 2008, **22**:89-91.
66. Yu DJ, Xu L, Nardi F, Li JG, Zhang RJ: **The complete nucleotide sequence of the mitochondrial genome of the oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae).** *Gene* 2007, **396**:66-74.
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. 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.
69. 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.
70. Ballard JWO: **Comparative genomics of mitochondrial DNA in members of the *Drosophila melanogaster* subgroup.** *J Mol Evol* 2000, **51**:48-63.
71. 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.
72. 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.
73. 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.
74. Matsumoto Y, Yanase T, Tsuda T, Noda H: **Species-specific mitochondrial gene rearrangements in biting midges and vector species identification.** *Med Vet Entomol* 2009, **23**:47-55.
75. 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.
76. 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 silkmoth, *Bombyx mandarina*, and its close relative, the domesticated silkmoth, *Bombyx mori*.** *Mol Biol Evol* 2002, **19**:1385-1389.
77. 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**:217-225.
78. Yang L, Wei ZJ, Hong GY, Jiang ST, Wen LP: **The complete nucleotide sequence of the mitochondrial genome of *Phthonandria atrilineata* (Lepidoptera: Geometridae).** *Mol Biol Rep* 2009, **36**:1441-1449.
79. 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 boisduvalii* (Lepidoptera: Saturniidae) and comparison with other lepidopteran insects.** *Gene* 2008, **413**:49-57.
80. Lee ES, Shin KS, Kim MS, Park H, Cho S, Kim CB: **The mitochondrial genome of the smaller tea tortrix *Adoxophyes honmai* (Lepidoptera: Tortricidae).** *Gene* 2006, **373**:52-57.
81. Liu Y, Li Y, Pan M, Dai F, Zhu X, Lu C, Xiang Z: **The complete mitochondrial genome of the Chinese oak silkmoth, *Antheraea pernyi* (Lepidoptera: Saturniidae).** *Acta Biochim Biophys Sin (Shanghai)* 2008, **40**:693-703.
82. Castro LR, Dowton M: **The position of the Hymenoptera within the Holometabola as inferred from the mitochondrial genome of *Perga condei* (Hymenoptera, Symphyta, Pergidae).** *Perga condei* 2005, **34**:469-479.
83. Cameron SL, Dowton M, Castro LR, Ruberu K, Whiting MF, Austin AD, Diement K, Stevens J: **Mitochondrial genome organization and phylogeny of two vespid wasps.** *Genome* 2008, **51**:800-808.
84. 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.
85. Silvestre D, Dowton M, Arias MC: **The mitochondrial genome of the stingless bee *Melipona bicolor* (Hymenoptera, Apidae, Meliponini): sequence, gene organization and a unique tRNA translocation event conserved across the tribe Meliponini.** *Genet Mol Biol* 2008, **31**:451-460.
86. Crozier RH, Crozier YC: **The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization.** *Genetics* 1993, **133**:97-117.
87. Wei S, Tang P, Zheng L, Shi M, Chen X: **The complete mitochondrial genome of *Evania appendigaster* (Hymenoptera: Evaniidae) has low A+T content and a long intergenic spacer between *atp8* and *atp6*.** *Mol Biol Rep* 2009 in press.
88. Arnoldi FGC, 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.
89. 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**:978-985.
90. Friedrich M, Muqim N: **Sequence and phylogenetic analysis of the complete mitochondrial genome of the flour beetle *Tribolium castaneum*.** *Mol Phylogenet Evol* 2003, **26**:502-512.
91. Li X, Ogoh K, Ohba N, Liang X, Ohmiya Y: **Mitochondrial genomes of two luminous beetles, *Rhagophthalmus lufengensis* and *R. ohbai* (Arthropoda, Insecta, Coleoptera).** *Gene* 2007, **392**:196-205.
92. 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**:513-526.
93. Sheffield NC, Song H, Cameron SL, Whiting MF: **A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles.** *Mol Biol Evol* 2008, **25**:2499-2509.
94. Beckenbach AT, Stewart JB: **Insect mitochondrial genomics 3: the complete mitochondrial genome sequences of representatives from two neuropteroid orders: a dobsonfly (order Megaloptera) and a giant lacewing and an owlfly (order Neuroptera).** *Genome* 2009, **52**:31-38.
95. Hua J, Li M, Dong P, Xie Q, Bu W: **The mitochondrial genome of *Protohermes concolorus* Yang et Yang 1988 (Insecta: Megaloptera: Corydalidae).** *Mol Biol Rep* 2009, **37**:1757-1765.
96. Perna NT, Kocher TD: **Patterns of Nucleotide Composition at Fourfold Degenerate Sites of Animal Mitochondrial Genomes.** *J Mol Evol* 1995, **41**:353-358.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

