Research Paper

The First Mitochondrial Genomes of Antlion (Neuroptera: Myrmeleontidae) and Split-footed Lacewing (Neuroptera: Nymphidae), with Phylogenetic Implications of Myrmeleontiformia

Yan Yan¹, Yuyu Wang¹, Xingyue Liu¹*, Shaun L. Winterton², Ding Yang¹*

¹. Department of Entomology, China Agricultural University, Beijing 100193, China.
². California State Collection of Arthropods, California Department of Food & Agriculture, Sacramento, California 95832, United States of America.

*Corresponding authors: E-mail: xingyue_liu@yahoo.com, dyangcau@126.com.

© Ivyspring International Publisher. This is an open-access article distributed under the terms of the Creative Commons License (http://creativecommons.org/licenses/by-nc-nd/3.0/). Reproduction is permitted for personal, noncommercial use, provided that the article is in whole, unmodified, and properly cited.

Received: 2014.04.21; Accepted: 2014.07.21; Published: 2014.08.15

Abstract

In the holometabolous insect order Neuroptera (lacewings), the cosmopolitan Myrmeleontidae (antlions) are the most species-rich family, while the closely related Nymphidae (split-footed lacewings) are a small endemic family from the Australian-Malesian region. Both families belong to the suborder Myrmeleontiformia, within which controversial hypotheses on the interfamilial phylogenetic relationships exist. Herein, we describe the complete mitochondrial (mt) genomes of an antlion (Myrmeleon immanis Walker, 1853) and a split-footed lacewing (Nymphes myrmeleonoides Leach, 1814), representing the first mt genomes for both families. These mt genomes are relatively small (respectively composed of 15,799 and 15,713 bp) compared to other lacewing mt genomes, and comprise 37 genes (13 protein coding genes, 22 tRNA genes and two rRNA genes). The arrangement of these two mt genomes is the same as in most derived Neuroptera mt genomes previously sequenced, specifically with a translocation of trnC. The start codons of all PCGs are started by ATN, with an exception of cox1, which is ACG in the M. immanis mt genome and TCG in N. myrmeleonoides. All tRNA genes have a typical clover-leaf structure of mitochondrial tRNA, with the exception of trnS1(AGN). The secondary structures of rrnL and rrnS are similar with those proposed insects and the domain I contains nine helices rather than eight helices, which is common within Neuroptera. A phylogenetic analysis based on the mt genomic data for all Neuropterida sequenced thus far, supports the monophyly of Myrmeleontiformia and the sister relationship between Ascalaphidae and Myrmeleontidae.

Key words: mitochondrial genome, Neuroptera, Myrmeleontiformia, phylogeny.

Introduction

Neuroptera (lacewings) are a holometabolous insect order, belonging to the superorder Neuroptera, with ca. 6000 extant species sorted in 16 families. The order consists of elegant insects, and notable for their transparent wings laced with intricate venation, and is characterized by the association of the male ninth gonocoxites with the gonarcus and in the larvae by the peculiar composite sucking mouthparts, discontinuous gut and Malphigian tubules modified for silk production [1, 2]. The earliest definitive neuropteran is hitherto known from the Late Permian [2], indicating a long evolutionary history of this order. The ongoing discoveries of a species rich and taxonomically diverse extinct fossil neuropteran fauna around the world suggest that this order may be less diverse today than in the past (provide some recently
published references of fossil neuropteran described).

The phylogenetic relationships among the various neuropteran families are still unresolved with competing hypotheses concerning the number and composition of suborders as well as the higher-level relationships. The morphology-based three-suborder concept [3-5] divides Neuroptera into Nevrorthiiformia, Myrmeleontiformia and Hemerobiiformia with the latter two being sister groups [4]. On the other hand, several other studies based on morphological as well as molecular data resulted in phylogenies differing in multiple aspects of interfamilial relationships except for the almost universal support for a paraphyletic Hemerobiiformia and a monophyletic Myrmeleontiformia [3-8].

Myrmeleontiformia is always recovered as a well-supported monophyletic group in almost all studies. This suborder consists of five extant families, i.e., Ascalaphidae, Myrmeleontidae, Nemopteridae, Nymphidae and Psychopsidae [1, 4, 8, 9]. Ascalaphidae (owlflies) and Psychopsidae (ant lions) are cosmopolitan in distribution, and Ascalaphidae contains ca. 430 species whereas Myrmeleontidae represents the most species-rich lacewing family with ca. 1630 species [10]. Nemopteridae (spoon- and thread-winged lacewings) is a distinctive group characterized by highly specialized (e.g., elongated, thread-like or spoon-shaped) hind wings and distributed in the Afrotropical, Palaeartic, Australian and Neotropical regions [11]. Nymphidae (split-footed lacewings) is a small family endemic to the Australian-Malesian region with ca. 35 species [12]. Psychopsidae (silky lacewings) is also a spectacular group with butterfly-like wing shape and narrowly distributed in Australia, Asia and Africa with ca. 26 extant species [13].

Despite the undisputed monophyly of Myrmeleontiformia, the interfamilial relationships within this suborder are still controversial. MacLeod identified a series of characteristics that define the group based on the larval head capsule, all associated with increase size of the head and wider articulation of the jaws resulting in more generalized prey selection [9]. Relationships among the various families of Myrmeleontiformia vary among authors. Almost all recent studies agree that Psychopsidae are sister to remaining Myrmeleontiformia, that Ascalaphidae and Myrmeleontidae are sister families and that Nymphidae are an intermediate clade between the two [4, 5, 8, 9, 14-17]. Yet, the phylogenetic position of Nemopteridae remains problematic with some authors proposing, based solely on morphological data, that Nemopteridae are sister to Psychopsidae in a clade sister to all other Myrmeleontiformia [4, 7]; more synthetic analyses of combined molecular and morphological data find Nemopteridae as sister to Ascalaphidae+Myrmeleontidae [8].

The mitochondrial genome analysis is becoming increasingly common in phylogenetic studies, population genetics and phylogeography [18], providing more phylogenetic information and multiple genome-level characteristics than shorter sequences of individual genes [19-22]. More than 600 insect mt genomes of arthropods have been sequenced, most of them from the mega-diverse orders such as Hymenoptera, Hymenoptera, Coleoptera, Diptera, and Lepidoptera [23, 24]. However, there are only nine mt genomes of Neuroptera fully sequenced and published to date, representing five families, i.e. Ascalaphidae, Chrysopidae, Ithonidae, Mantispidae and Osmyridae [23, 25]. Within Myrmeleontiformia, the mt genomes of only two species of Ascalaphidae (Libelloides macaronius Scopoli, 1763 and Ascaloptynx appendiculatus (Fabricius, 1793)) have been sequenced [30, 31].

Herein, we present two complete mt genomes of Myrmeleon immanis Walker, 1853 and Nymphes myrmeleonoides Leach, 1814, as the first representatives of Myrmeleontidae and Nymphidae, respectively. We compare the genomic structure and composition, such as gene order, nucleotide composition, codon usage, secondary structure of tRNAs, compositional biases, and investigate the phylogenetic status of Myrmeleontidae and Nymphidae within Neuroptera based on a phylogenomic analysis.

Results and Discussion

The mt genome of Myrmeleon immanis is a typical circular DNA with 15,799 bp (GenBank accession number KJ461323), which comprises 37 genes with a major A+T-rich region of 971 bp (Figure 1, Supplementary Material: Table S1), while the mt genome of Nymphes myrmeleonoides is a similar circular DNA 15,713 bp in length (GenBank accession number KJ461322) with a noncoding region of 812 bp (Figure 1, Supplementary Material: Table S2). The mt genomes of the two species are slightly smaller than those of other sequenced Neuroptera, whose size range from 15,877 bp (A. appendiculatus) to 16,723 bp (Chrysopa pallens). Both mt genomes consist of 22 transfer RNA genes, 13 protein-coding genes and two ribosomal RNA genes. Twenty-three genes are located on the J strand, while the remaining fourteen genes are located on the N strand. In N. myrmeleonoides the gene overlaps are 19 bp distributed in seven gene junctions, while in M. immanis the gene overlaps are 29 bp located in 12 gene junctions.

Genome organization

Gene arrangement is identical in both M. immanis and N. myrmeleonoides mt genomes. Compared
with Drosophila yakuba (Burla) (Diptera: Drosophilidae), whose genome exhibits the typical ground pattern for insect mt genomes [26], the gene arrangement is relatively conserved except for one rearrangement in M. immanis and N. myrmeleonoides (Figure 1, Supplementary Material: Tables S1-S2). In most insect mt genomes, the assumed ancestral arrangement for these three tRNA genes is trnW-trnC-trnY, but in M. immanis and N. myrmeleonoides, the arrangement is trnC-trnW-trnY. Among the known mt genomes of Neuroptera, the arrangement of W-C-Y is clearly plesiomorphic and is found in Raphidioptera, Megaloptera and a single family of Neuroptera (Osmymidae: Thyridosmylus) [23]; all other Neuroptera have the rearrangement C-W-Y, which is a derived characteristic for the bulk of the family [23, 25, 27-31].

Protein-coding genes

Among the 13 PCGs in the M. immanis and N. myrmeleonoides mt genomes, nine PCGs are located in the J strand, while the other four are located in the N strand (Figure 1). The length of all PCGs for M. immanis is 11,189 bp and 11,182 bp in N. myrmeleonoides. The overall AT content of the 13 PCGs is 74.3% in the M. immanis mt genome, ranging from 67.5% (cox1) to 83.5% (nad6), while in the N. myrmeleonoides mt genome, the AT content of the 13 PCGs is 79.0%, ranging from 72.9% (cox1) to 86.4% (nad6).

Generally, the start codon is mostly ATT and ATG for the PCGs, but ATA represents an alternative start codon for a few PCGs in the mt genomes of M. immanis and N. myrmeleonoides. In the M. immanis mt genome, ATA is a start codon for nad1, nad3, and nad5, while in N. myrmeleonoides, ATA as the start codon is only found in nad1 (Tables S1-S2). In cox1 the start codon is ACG in the M. immanis and TCG in N. myrmeleonoides.

Most PCGs have a start codon as ATG in nad4, nad4l, cox2, cox3, atp6, and cyt b. However, the other PCGs in Neuroptera families contain some unusual start codons, like TTG, ACG, TCG, TTA, CTG. Comparing with the codon usage of the PCGs in the Neuroptera species, it is obvious that the start codon of cox1 is highly variable (Supplementary Material: Table S3). The start codons vary mostly in cox1 with many types of uncommon start codons like CTG in Mongoloraphidia harmandi, TTA in A. appendiculatus and TCG in several other species. The nad1 genes in Neuroptera also show the various start codons, such as TTG, ATG and ATA.

Stop codons are far less variable than start codons in the Neuroptera. Most stop codons of PCGs are complete or incomplete TAA, except in nad1 and nad3 which have TAG as the end codon in several species. The end codon of the 13 PCGs in M. immanis and N. myrmeleonoides are exactly the same. The end codon of nad1 is TAG. The cox1 and cox2 end with a thymine followed by a tRNA gene instead of an end codon. The other PCGs end with a conserved triplet of TAA. The triplet stop codon (TAA, TAG) always has a one or two bp overlap with the down-stream tRNA gene, and these overlaps could be the ‘backup’ to prevent translation read through if the transcripts are not cleaved correctly [32].

In the mt genomes of both M. immanis and N. myrmeleonoides, an ‘ATGATAA’ overlap is present between atp8 and atp6. This feature is shared by other species with known mt genomes in Neuroptera, in which the atp8 and atp6 are located on the J strand.
with an overlap of 7 bp. On the N strand, an overlap of ‘ATGTTAA’ between nad4 and nad4l was also observed in the *M. immanis* and *N. myrmeleonoides* mt genomes. However, within Myrmeleontiformia, the overlap of nad4 and nad4l is not conserved because the mt genomes of two owlfly species (*Ascaloptynx appendiculatus* [30] and *Libelloides macaronius* [31]) show a contiguous pattern between nad4 and nad4l.

**Transfer RNAs**

In the 22 tRNA genes, 14 genes are located on the J strand, while the others are encoded on the N strand. The length of a single tRNA ranges from 64 bp to 72 bp in both *M. immanis* and *N. myrmeleonoides*. The aminoacyl (AA) stem and the AC loop are conserved with 7 bp, but the AA arm of *trnI* in the *M. immanis* mt genome has just 6 bp. The DHU arm and TΨC arm are more variable, ranging from 3-5 bp. The AC arm is always 5 bp with some exceptions like 4 bp of *trnK*. Except for *trnS1(AGN)*, all the other tRNAs form into the typical cloverleaf structure, and the DHU arm of *trnS1(AGN)* in the two species is a large loop instead of the conserved stem-and-loop structure (Figures 2-3). Some base pairs are not the classic A-U and C-G based on the secondary structure. In tRNAs of *M. immanis* and *N. myrmeleonoides* mt genomes, there are some mismatched base pairs such as G-U and U-U (Supplementary Material: Tables S4-S5). In *Myrmeleon immanis*, there are 20 G-U mismatched base pairs, and these mismatched base pairs are located in the AA stem (3 bp), the AC arm (4), the TΨC arm (5), and the DHU arm (8). However, in *N. myrmeleonoides*, there are 15 G-U mismatched base pairs located in the AA arm (3), AC arm (2), TΨC arm (1), and DHU arm (9). As for the U-U mismatch, the *trnW* gene has this type of mismatch in both mt genomes, and was observed in the anticodon (AC) arm in *N. myrmeleonoides* and the TΨC arm in *M. immanis*.

**Ribosomal RNAs**

The length of *rrnL* and *rrnS* of the *M. immanis* mt genome is 1,304 bp and 777 bp, respectively, and the length of the *rrnL* and *rrnS* in *N. myrmeleonoides* is 1,320 bp and 786 bp. The overall structures of the two rRNAs largely resemble the previously published secondary structures for *L. macaronius* [31].

A proposed secondary structure model for *rrnS* of *M. immanis* and *N. myrmeleonoides* is shown in Figure 4. The secondary structure of *rrnS* contains three domains and 34 helices, which is similar with that of some other endopterygote orders (i.e. Coleoptera, Diptera, Hymenoptera, and Lepidoptera) [33-36]. There are a limited number of non-canonical pairings in the *rrnS* of the two mt genomes. The *rrnS* has 189 canonical pairings (A-U and G-C) and 42 non-canonical pairings (G-U: 31, U-U: 5, C-U: 2, G-A: 3 and C-A: 1).

A proposed secondary structure model for *rrnL* of *M. immanis* and *N. myrmeleonoides* is shown in Figure 5. The secondary structure of *rrnL* is also generally congruent with the secondary structure models of previously published endopterygote insects, and have 5 canonical domains (I–II, IV–VI) with an absence of domain III. The five domains and lack of domain III is also a typical trait in arthropods [37]. The secondary structure of *rrnL* has 50 helices. Domain I contains nine helices rather than eight helices which is typical for neuropteran *rrnL* [23]. The *rrnL* has 321 canonical pairings (A-U and G-C) and 65 non-canonical pairings (G-U: 50, U-U: 11, C-U: 1, G-A: 1 and C-A: 2). The distribution of the conserved nucleotides is uneven in the secondary structure of *rrnL*, with lowest level of invariable positions on domains I-II and highest level on domain IV.

**Non-coding region**

The major A+T rich-region, i.e. the control region, of the *M. immanis* and *N. myrmeleonoides* mt genomes is located between *rrnS* and *trnl*, and the A+T base composition of this region is 86.8% and 93.1%, respectively. The high rate of A+T base composition is caused by the repeated motifs containing almost exclusively A/T at different lengths. In the control region of the *M. immanis* mt genome, the most abundant TA motif occurs 194 times, AT occurs 187 times, TAA occurs 78 times, TTA occurs 27 times, AATT occurs 12 times, and TTATATAT occurs 3 times. In the control region of *N. myrmeleonoides*, the TA motif occurs 198 times, AT occurs 201 times, TAA occurs 80 times, AATT occurs 38 times, TTTA occurs 13 times, and ATATAA occurs 10 times. Additionally, some short tandem repeat sequences exist in the control region of *M. immanis*, e.g. (AT)₃, (AT)₄, (AT)₅, (TA)₃, (TA)₄, (TA)₅, (TA)₆, (TAT)₃, and (ATT). Whereas, in the control region of *N. myrmeleonoides*, the short tandem repeat sequences are (TA)₃, (TA)₄, (TA)₅, (TA)₆, (TA)₉, (TA)₉, (AT)₅, (AT)₆, (AT)₉, (TAA)₃, and (TTTA). These tandem repeat elements can be considered as microsatellites and are useful in the study of geographical population structure [38].

The secondary structure of control region is simpler in the mt genomes of *M. immanis* and *N. myrmeleonoides* than that in Hymenopteran [39], Hemipteran [40] or Lepidopteran [41] insects with less conserved blocks or long tandem repeats. There is a poly-T stretch with 12 bp in the control region of the *M. immanis* and *N. myrmeleonoides* mt genomes, and an area with 24 bp identical sequences upstream from poly-T is present *M. immanis* and *N. myrmeleonoides*, i.e. ATGGTTCAATAAAAATAATTCTCTTTTTTTT.
TTTT. This region is highly conserved within Neuroptera.

There is another long non-coding region in the mt genome of *M. immanis* between *trnI* and *trnQ* with 70 bp, and a similar non-coding region in the same position of the mt genome of *N. myrmeleonoides* is present with 58 bp. Additionally, a non-coding region is located between *trnS2(UCN)* and *nad1*, with 16 bp in the *N. myrmeleonoides* mt genome and 21 bp in the *M. immanis* mt genome. The other non-coding regions of the *M. immanis* mt genome are rather short, ranging from 1 to 9 bp, while *N. myrmeleonoides* has another long non-coding region of 29 bp between *trnQ* and *trnM*. The other short non-coding regions range from 1 to 14 bp in the *N. myrmeleonoides* mt genome.

Figure 2. Secondary structures of 22 tRNA genes of *Myrmeleon immanis*. 
Figure 3. Secondary structures of 22 tRNA genes of Nymphes myrmecoleonides.
Figure 4. Secondary structures of rrnS of *Myrmeleon immanis* and *Nymphes myrmeleonoides*. Red circles indicate the conserved nucleotides in the rrnS of *M. immanis* and *N. myrmeleonoides*. Gray circles indicate the variable nucleotides in the rrnS of *M. immanis*. Orange circles indicate the variable nucleotides in the rrnS of *N. myrmeleonoides*. Pink arrows indicate the replacement of different nucleotides in *N. myrmeleonoides*. Blue arrows indicate the insertion of additional nucleotides in *N. myrmeleonoides*. Each helix is numbered progressively from the 5' to the 3' end together with the first nucleotide belonging to the helix itself. Domains are labeled with Roman numerals. Tertiary structures are denoted by boxed bases joined by solid lines.
Codon usage

All the genetic codons in the Neuropterida mtDNAs use standard invertebrate mitochondrial genetic code. We exclude the stop codons since some of them are incomplete. This result (Figure 6) shows the relative synonymous codon usage (RSCU). It is obvious that the preferred codon usage is A/T in their third position rather than G/C. As for the four-fold degenerate codon usage, the third position prefers T than A in most codon families except Arg, Gly, Ser1, and Ser2 in several species. The result also shows the missing part of the codon in Neuropterida families, and these missing codons all have high content of GC with G or C in the third position. In the mt genome of *M. immanis*, the missing codons are Arg (CGC), Gly (GGC) and Ser1 (AGG), while in the mt genome of *N. myrmeloonoides*, the missing parts are Arg (CGC, CGG), Gln (CAG), Ser1 (AGG), Thr (ACG) and Val (GUC). In the Neuropterida mtDNAs, the missing codon mostly occurs in Arg and Ser1 with absence of CGC and AGG. The mt genomes of *M. immanis* and *N. myrmeloonoides* also show a significant bias to the A/T. The top six most frequently used triplets in these two mt genomes are the same, i.e. UUU (Phe), UUA (Leu), AUU (Ile), AUA (Met), UAU (Tyr) and AAU (Asn), and all these triplets consist of A and T, which may be due to the high AT composition in the mt genomes and PCGs of the *M. immanis* and *N. myrmeloonoides*.

Base composition

The mt genomes of *M. immanis* and *N. myrmeloonoides* show a strong bias toward A and T (*M. immanis*: A=38.7%, T=36.9%, C=14.4%, G=10.1%; *N. myrmeloonoides*: A=39.8%, T=40.7%, C=11.1%, G=8.4%), which is very typical of invertebrate mt genomes (Figure 7). The three parameters: AT-skew, GC-skew, and A+T content are usually used in the investigation of the nucleotide-compositional behavior of mt genomes [42, 43]. Results of the comparative analysis of the A+T%, AT- and GC-Skew within the sequenced Neuropterida mt genomes is summarized in a three-dimensional scatter-plot chart (Figure 7). Considering the whole mt genome, the A+T% in *N. myr-
meleonoides is higher than any other species of sequenced mt genomes in Neuropterida, while the A+T% in M. immanis is lower than the average AT content. The AT- and GC-Skew indicate the strand bias in the base composition [44-46]. The average AT-Skew of sequenced Neuropterida mt genomes is 0.0138, ranging from -0.0367 in Apochrysa matsumurae to 0.0711 in L. macaronius, whereas the average GC-Skew is -0.1863, spanning from -0.2590 in Corydalus cornutus to -0.1374 in Chrysoperla nipponensis (Supplementary Material: Table S6). The AT- and GC-Skew are 0.0238 and -0.1755 in the M. immanis mt genome, while they are -0.0112 and -0.1385 in N. myrmeleonoides.
Figure 6. Relative synonymous codon usage (RSCU) of the mt genomes of Neuroptera. Stop codon is not given.

The three parameters of the Neuroptera genomes indicate some phylogenetic relationships of certain families (Figure 7). Three species of Chrysopidae, Chrysoperla nipponensis, Chrysopa pallens and Apochrysa matsumurae (blue balls), show the similar content of A+T and AT/GC-skew. Two species of Ascalaphidae, Libelloides macaronius and Ascaloptynx appendiculatus, also show the similar content of A+T and AT-skew with a little difference in GC-skew. Moreover, the genome of Myrmeleon immanis shows the similar content of those three parameters with that of A. appendiculatus and L. macaronius, while it is different from that of N. myrmeleonoides with a discrepancy of A+T% and AT-skew. This result confirms the close relationship between Ascalaphidae and Myrmeleontidae relative to Nymphidae.

Phylogeny

Among the four datasets used in the phylogenetic analyses, there are 13,257 sites in the PCG123R matrix (containing all three codon positions of PCGs, plus two rRNA genes), 11,076 sites in the PCG123 matrix (containing all three codon positions of PCGs), 9,565 sites in the PCR12R matrix (containing the first and the second codon positions of PCGs), 7,384 sites in the PCG12 matrix (containing the first and the second codon positions of PCGs), plus two rRNA genes) and 7,384 sites in the PCG12 matrix.
(containing the first and the second codon positions of PCGs).

The phylogenetic trees inferred from both Bayesian and ML analyses have very similar topologies across the four datasets. The branch support values of the PCG123R matrix are higher than the other matrices. Therefore, we show the phylogenetic trees of the PCG123R matrix in Figure 8. All species of Myrmeleontiformia in our analysis formed a monophylum, supporting the monophyly of this suborder. Second, the undoubted sister relationship between Ascalaphidae and Myrmeleontidae is also corroborated. Lastly, placement of Nymphidae as sister to to Ascalaphidae and Myrmeleontidae is corroborated [1, 4]. However, Nymphidae was alternatively considered to be the sister to Psychopidae + Nemopteridae based on morphological data from the genital sclerites [7]. The recent comprehensive phylogenetic study on Neuropterida combining morphological and molecular data recovered that Nymphidae is the sister of the lineage comprising Nemopteridae + (Ascalaphidae + Myrmeleontidae) [8]. Our mt phylogenomic result is consistent with this hypothesis that Nymphidae together with Ascalaphidae and Myrmeleontidae form a monophyletic group. However, lack of Nemopteridae and Psychopidae samples precludes any assessment of the placement of these families relative to the rest of Myrmeleontiformia. Future mt genomic study should be focused on Psychopidae and Nemopteridae as well as some significant subfamilial taxa (e.g. Stilbopteryginae) in order to unravel the higher phylogeny within Myrmeleontiformia.
Material and methods

Sampling and DNA extraction

The specimen of *M. immanis* was collected on August 13, 2011 at Qianguo prairie, Songyuan, Jilin Province, China by Yuyu Wang using the light trap and *N. myrmeleonoides* was collected on November 13, 2007 by S. L. Winterton at the base of Mt. Coot-tha, Slaughter Falls, Brisbane, Queensland, Australia. After collection, specimens were preserved in 95% ethanol and deposited at -20°C in the Entomological Museum of China Agricultural University (CAU) for long-term preservation. Total DNA of *M. immanis* and *N. myrmeleonoides* was extracted by using the TIANGEN (TIANamp Genomic DNA Kit) with the muscle of the thorax. The quality of DNAs was assessed through electrophoresis in a 1% agarose gel and staining with Gold View (nucleic acid stain replacing EB).

PCR amplification and sequencing

We amplified the mitochondrial genomes of *M. immanis* and *N. myrmeleonoides* by using overlap fragments which about 300-1200 bp. These fragments were amplified by PCR (Polymerase Chain Reaction) with 30 s at 95°C, 40 cycles of 10 s at 95°C, 50 s at 40°C, 2 min at 65°C, and 10 min at 65°C, then finishing with 10°C indefinitely. All the PCR used the NEB Long Taq DNA polymerase (New England BioLabs, Ipswich, MA) and the PCR products were detected by agarose gel electrophoresis to ensure the correctness of the result and the approximate size of the fragments. The primers of the fragments are listed in Supplementary Material: Table S8 for *M. immanis* mt genome and Table S9 for *N. myrmeleonoides*.

All fragments were sequenced in both directions by using the BigDye Terminator Sequencing Kit (Applied Bio Systems) and the ABI 3730XL Genetic Analyzer (PE Applied Biosystems, San Francisco, CA, USA) with two vector-specific primers and internal primers for primer walking.

Bioinformatic analysis

All sequenced fragments of the *M. immanis* and *N. myrmeleonoides* mt genomes were spliced manually and the uncertain site, which is located at the overlap of two fragments, is determined by comparing the original alignments. The 22 tRNA genes were identified by using the tRNAscan-SE Search Server v.1.21 [47] with a Cove cutoff score of 1 and the prediction of the genetic code followed the invertebrate mitochondrial DNA, while some tRNA genes were identified by the multiple sequence alignment of the tRNAs in Neuroptera. The secondary structures of tRNAs were also estimated by tRNAscan-SE Search Server v.1.21.

The boundaries of 13 PCGs in the *M. immanis* and *N. myrmeleonoides* mt genomes were identified by comparing the boundaries of same PCGs in other Neuroptera mt genomes with MEGA 4.0 [48]. The rRNA genes and the control region were identified by the boundary of the tRNA genes and comparing with other insect mt genomes. The base composition and the codon usage were analyzed with MEGA 4.0 [48]. In addition, the AT-skew and GC-skew in the mt genomes were used for study on the feature of the base composition and calculated with the formulae: AT-skew = (A-T)/(A+T) and GC-skew = (G-C)/(G+C) [49]. We also used the OriginPro 9.0 [50] to generate the three-dimensional scatter plot of the mt genomes within the Neuroptera. The secondary structures of the *rrnL* and *rrnS* were inferred based on the model of *Drosophila yakuba* [26] and *Libelloides macaronius* [31]. Sequence motifs in the control region were identified using Spectral Repeat Finder program [51].

Phylogenetic analysis

To investigate the phylogenetic status of Myrmeleontidae and Nymphidae within Neuroptera, we selected all eight species of Neuroptera in which the complete mitochondrial genomes have been sequenced (see Supplementary Material: Table S7). *Mangoloraphidia harmandi* (Raphidioptera), *Neochauiodes punctatolosus* (Megaloptera), *Sialis hamata* (Megaloptera) and *Corydalus cornutus* (Megaloptera) were selected as outgroups because of the close relationship among Neuroptera, Megaloptera and Raphidioptera.

We used the Clustal X [52] for alignment of sequences of PCGs. Additionally, the two rRNAs alignments were conducted by the G-blocks Server (http://molevol.cmima.csic.es/castresana/Gblocks_server.html) based on more stringent selection. By using the MrBayes Version 3.1.2 [53] and a PHYML online web server [54, 55], we constructed the phylogenetic trees and the evolutionary model was estimated by Modeltest 3.7 [56], resulting the GTR+I+G model was optimal for analysis nucleotide alignments (PCGs and rRNAs). In the Bayesian inference, two simultaneous runs were conducted with 10,000,000 generations. Each set was sampled with a burnin of 25% of every 1000 generations. Trees inferred prior to stationarity were discarded as burnin, and the remaining trees were used to construct a 50% majority-rule consensus tree. In the ML analysis, the nodal supporting values were assessed by bootstrap re-sampling (BP) [57] calculated using 100 replicates. We used all the PCGs and delete any regions of ambiguous alignment, and subsequently concatenated all 13 PCGs (PCG123) (excluding the stop codon) to construct the phylogenetic tree initially, then included additionally the *rrnL* and *rrnS* (PCG123R) sequences.
to reconstruct the phylogenetic tree as a second result. We compared phylogenetic estimates to determine the most appropriate tree by considering the nodal supporting values.

**Supplementary Material**

Tables S1 – S9.

http://www.ijbs.com/v10p0895s1.pdf

**Abbreviations**

tmt: mitochondrial; PCG: protein-coding gene; atp6 and atp8: genes for the ATPase subunits 6 and 8; cox1-cox3: genes for cytochrome c oxidase subunits I-III; cytB: a gene for apocytochrome b; nad1-nad6 and nad4l: genes for NADH dehydrogenase subunits 1-6 and 4L; rRNA: large (16S) rRNA subunit (gene); rRNA: small (12S) rRNA subunit (gene); trnX (where X is replaced by one letter amino acid code of the corresponding amino acid), transfer RNA.

**Acknowledgements**

This research was supported by the National Natural Science Foundation of China (Nos. 31322501 and 31320103902), the National Key Basic Research Program of China (973 Program) (No. 2013CB127600), the Foundation for the Author of National Excellent Doctoral Dissertation of PR China (No. 2011178) and the National Science Foundation of the United States of America (DEB-1144119).

**Competing Interests**

The authors have declared that no competing interest exists.

**References**


32. 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.
38. Wei S, Shi M, Sharkey MJ, Van Achterberg C, Chen X. Comparative mitogenomic genomics of Braconidae (Insecta: Hymenoptera) and the phylogenetic utility of