

A Unique Box in 28S rRNA Is Shared by the Enigmatic Insect Order Zoraptera and Dictyoptera

Yanhui Wang¹✉, Michael S. Engel^{2,3}✉, Jose A. Rafael⁴✉, Kai Dang¹, Haoyang Wu¹, Ying Wang¹, Qiang Xie^{1*}, Wenjun Bu^{1*}

1 Institute of Entomology, College of Life Sciences, Nankai University, Tianjin, China, **2** Division of Entomology (Paleoentomology), Natural History Museum, London, England, **3** Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas, United States of America, **4** Instituto Nacional de Pesquisas da Amazônia, INPA, Manaus, Amazonas, Brazil

Abstract

The position of the Zoraptera remains one of the most challenging and uncertain concerns in ordinal-level phylogenies of the insects. Zoraptera have been viewed as having a close relationship with five different groups of Polyneoptera, or as being allied to the Paraneoptera or even Holometabola. Although rDNAs have been widely used in phylogenetic studies of insects, the application of the complete 28S rDNA are still scattered in only a few orders. In this study, a secondary structure model of the complete 28S rRNAs of insects was reconstructed based on all orders of Insecta. It was found that one length-variable region, D3-4, is particularly distinctive. The length and/or sequence of D3-4 is conservative within each order of Polyneoptera, but it can be divided into two types between the different orders of the supercohort, of which the enigmatic order Zoraptera and Dictyoptera share one type, while the remaining orders of Polyneoptera share the other. Additionally, independent evidence from phylogenetic results support the clade (Zoraptera+Dictyoptera) as well. Thus, the similarity of D3-4 between Zoraptera and Dictyoptera can serve as potentially valuable autapomorphy or synapomorphy in phylogeny reconstruction. The clades of (Plecoptera+Dermaptera) and ((Grylloblattodea+Mantophasmatodea)+(Embiodea+Phasmatodea)) were also recovered in the phylogenetic study. In addition, considering the other studies based on rDNAs, this study reached the highest congruence with previous phylogenetic studies of Holometabola based on nuclear protein coding genes or morphology characters. Future comparative studies of secondary structures across deep divergences and additional taxa are likely to reveal conserved patterns, structures and motifs that can provide support for major phylogenetic lineages.

Citation: Wang Y, Engel MS, Rafael JA, Dang K, Wu H, et al. (2013) A Unique Box in 28S rRNA Is Shared by the Enigmatic Insect Order Zoraptera and Dictyoptera. PLoS ONE 8(1): e53679. doi:10.1371/journal.pone.0053679

Editor: Ben J. Mans, Onderstepoort Veterinary Institute, South Africa

Received: July 5, 2012; **Accepted:** December 3, 2012; **Published:** January 3, 2013

Copyright: © 2013 Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the National Science Foundation of China (grant numbers J0930005, 30970350, 30725005). The participation of Michael S. Engel was supported by U.S. National Science Foundation grant DEB-0542909. Jose A. Rafael was financially supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), grant number 472237/2009-8 and Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) by the financial support to the Project PRONEX, Edital 016/2006, Proc. 1437/2007. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: qiangxie@nankai.edu.cn (QX); wenjunbu@nankai.edu.cn (WJB)

✉ These authors contributed equally to this work.

Introduction

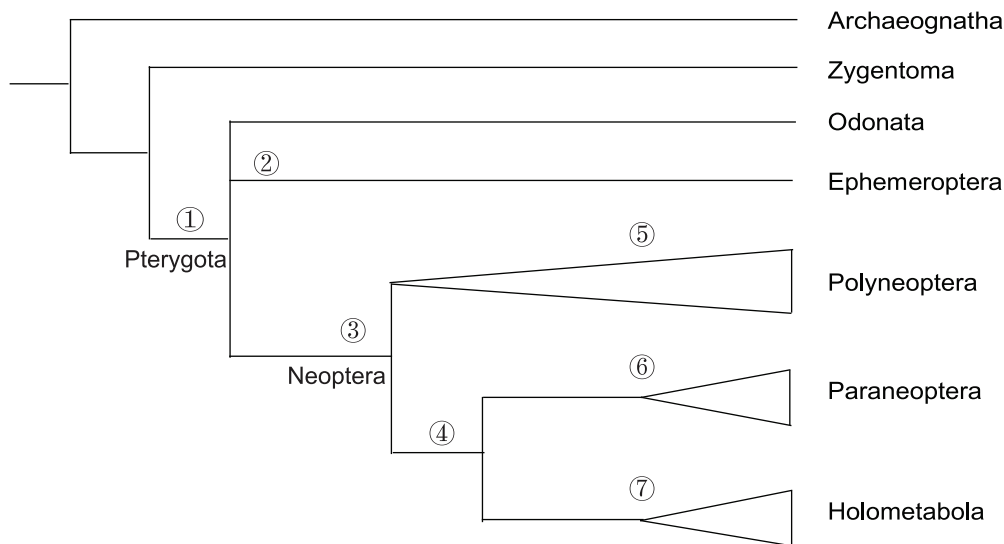
Insects are the most diverse group of living organisms. The Insecta are comprised of the primitively wingless orders Archaeognatha (bristletails) and Zygentoma (silverfish), and the winged lineages of Odonata (dragonflies and damselflies), Ephemeroptera (mayflies), and the hyperdiverse Neoptera (all other insect lineages). The Neoptera themselves are divided into three large groups or supercohorts, the Polyneoptera, Paraneoptera, and Holometabola [1–6]. The currently recognized relationships between the orders of insects are summarized in Figure 1. With respect to the Neoptera, the monophyly of the Neoptera, Paraneoptera, and Holometabola have been overwhelmingly supported from morphological, paleontological, molecular, as well as combined analytical studies (Figure 1).

Compared to the Paraneoptera and Holometabola, there are many more uncertainties regarding the phylogenetic interrelationships, and even the monophyly as a whole, of the Polyneoptera.

The Polyneoptera include the Dermaptera (earwigs), Plecoptera (stoneflies), Blattaria (roaches), Isoptera (termites), Mantodea (mantises), Zoraptera (angel insects), Notoptera [Grylloblattodea (ice crawlers) and Mantophasmatodea (rock crawlers)], Embiodea (webspinners), Orthoptera (crickets, grasshoppers, katydids, and wetas), and Phasmatodea (stick and leaf insects). Among these eleven orders, the monophyly of the Dictyoptera (Blattaria, Isoptera, and Mantodea) is the most widely accepted. In recent years, significant data from multiple sources has supported the unification of the Mantophasmatodea and Grylloblattodea as the order Notoptera [6,7–13]. Of more controversy are the several studies uniting the stoneflies and earwigs [11,14–16], although as noted by Yoshizawa [17] such a grouping is unstable and the Dermaptera remain challenging to place between the dictyopteran, plecopteroid, and orthopteroid orders.

Beyond the three clades mentioned above, there is virtually no consensus regarding the phylogeny of the Polyneoptera. The gaps between various taxonomic systems for the Polyneoptera are quite

a)



b)

Nodes	References & Dataset
①	Hennig 1969/1981 (morphology); Boudreaux 1979 (morphology); Kristensen 1991 (morphology); Kukalová-Peck 1991 (morphology); Kristensen 1995 (morphology); Wheeler et al. 2001 (18S+28S, morphology); Kjer et al. 2006 (molecular); Misof et al. 2007 (18S); von Reumont et al. 2009 (18S+28S);
②	Monophyletic Palaeoptera: Martynov 1923, 1925 (morphology); Hennig 1969/1981 (morphology); Kukalová-Peck 1991 (morphology); Kjer et al. 2006 (molecular)
	Metapterygota: Kristensen 1975, 1981, 1991, 1995 (morphology); Wheeler et al. 2001 (morphology); Grimaldi and Engel 2005 (paleontological);
③	Chiastomyaria: Boudreaux 1979 (morphology); Kjer 2004 (18S); Yoshizawa et al. 2005 (18S); Misof et al. 2007 (18S); von Reumont et al. 2009 (18S+28S); Simon et al. 2009 (EST); Lin et al. 2010 (Mitochondrial genome)
	Hennig 1969/1981 (morphology); Boudreaux 1979 (morphology); Kukalová-Peck 1991 (morphology); Wheeler et al. 2001 (morphology); Kjer 2004 (18S); Terry et al. 2005 (18S); Kjer et al. 2006 (molecular); von Reumont et al. 2009 (18S+28S); Ishiwata et al. 2010 (nuclear genes)
④	Boudreaux 1979 (morphology); Kukalová-Peck 1991 (morphology); Wheeler et al. 2001 (morphology)
⑤	Paraphyly: Hennig 1969/1981 (morphology); Kristensen 1995 (morphology); Beutel et al. 2001 (morphology); Kjer 2004 (18S); Terry et al. 2005 (18S); Yoshizawa et al. 2005 (18S); Beutel et al. 2006 (morphology); Kjer et al. 2006 (molecular); Misof et al. 2007 (18S); von Reumont et al. 2009 (18S+28S)
	Monophyly: Boudreaux 1979 (morphology); Kukalová-Peck 1991 (morphology); Wheeler et al. 2001; Ishiwata et al. 2010; Yoshizawa 2011
⑥	Boudreaux 1979 (morphology); Kukalová-Peck 1991 (morphology); Wheeler et al. 2001 (morphology); Beutel and Gorb 2001, 2006 (morphology); Kjer 2004(18S); Grimaldi and Engel 2005 (morphology, fossil); Kjer et al. 2006 (molecular); Misof et al. 2007 (18S)
⑦	Hennig 1969/1981 (morphology); Boudreaux 1979 (morphology); Kukalová-Peck 1991 (morphology); Kristensen 1995 (morphology); Whiting et al. 1997 (18S+28S, morphology); Wheeler et al. 2001 (morphology); Kjer 2004 (18S); Yoshizawa et al. 2005 (18S); Kjer et al. 2006 (molecular); Misof et al. 2007 (18S); Wiegmann et al. 2009 (nuclear genes); Ishiwata et al. 2010 (nuclear genes); Mckenna et al. 2010 (9-genes)

Figure 1. Currently recognized phylogeny of the Insecta. a) Summary cladogram, b) list of morphological and molecular studies. doi:10.1371/journal.pone.0053679.g001

large (Table 1). The most considerable change in position remains that of the Zoraptera. This order is thought to be closely related either to Paraneoptera [4,5,18,19], Holometabola [20], within

Dictyoptera [15,21,22], with Dictyoptera+Dermaptera [23], Embiidea [6,17,24–27], Dermaptera [9,12], or alongside Plecoptera+Dermaptera [28].

Table 1. Some traditional classifications of living orders in the Supercohort Polyneoptera.

Hennig (1953, 1969, 1981)	Blattopteroidea	Problematic "Orthopteroidea"	Remaining orders	
	Blattaria, Isoptera, Mantodea	Dermaptera, Grylloblattodea, Orthoptera, Phasmatodea	Plecoptera, Embiodea	
Sharov (1968)	Orthopteroidea			Remaining orders
	Dictyoptera, Dermaptera, Grylloblattodea, Orthoptera, Phasmatodea			Plecoptera, Embiodea, Zoraptera
Boudreaux (1979)	Orthopterodida			Remaining orders
	Dictyoptera, Dermaptera, Grylloblattodea, Orthoptera, Phasmatodea, Zoraptera			Plecoptera, Embiodea
Kristensen (1991, 1995)	Dictyoptera	Remaining orders		
	Blattaria, Isoptera, Mantodea	Dermaptera, Grylloblattodea, Orthoptera, Phasmatodea, Embiodea, Plecoptera		
Kukalová-Peck (1991)	Blattoneoptera	Orhthoneoptera		Pleconeoptera
	Dictyoptera, Dermaptera, Grylloblattodea, Zoraptera	Orthoptera, Phasmatodea, Embiodea		Plecoptera
Grimaldi & Engel (2005), Arillo & Engel (2006)	Blattodea	Orthopterida	Dermapterida	Plecopterida
	Blattaria, Isoptera, Mantodea	Notoptera, Orthoptera, Phasmatodea	Dermaptera	Plecoptera, Embiodea, Zoraptera

doi:10.1371/journal.pone.0053679.t001

Hennig [4,5] and Kristensen [18,29] placed the Zoraptera within the Paraneoptera based on their reduced numbers of tarsomeres, Malpighian tubes, and ganglia. However, the reduction of tarsomeres occurs independently in numerous insect lineages (e.g., Plecoptera, Dermaptera, Isoptera) and cannot be considered definitive evidence of relationship in isolation from other character sources. Subsequently, Minet and Bourgoin [24] allied the Zoraptera with the Embiodea based on a unique metatibial musculature, and Engel and Grimaldi [25,26] expanded this suite of characters to include the loss of gonostyli, reduced cerci, enlarged metafemora, narrow and paddle-shaped forewings, presence of apterous morphs, dehiscent wings, and gregarious behavior. Yoshizawa [17,30] later added numerous wing base characters to the list of synapomorphies supporting Zoraptera+Embiodea (= Mystroptera). Others have argued that Zoraptera share some synapomorphies with the Dictyoptera, such as a highly derived flight motor, reduced phragmata, and greatly reduced indirect flight musculature [20,21], but the latter two reductions occur homoplastically numerous times across many orders and cannot be considered strongly indicative of relationship. Simultaneously, molecular-based phylogenetic placements of the Zoraptera also have been uncertain [9,12,15,28,31]. The main reasons for these controversies may be due to the analytical methods employed [14,32–34] as well as the inclusion of dubious sequences [34,35].

Ribosomal DNA sequences have been playing a major role in molecular phylogenetic studies in insects for the past two decades [36]. Analysis of 18S rDNA (also known as small subunit ribosomal DNA, SSU rDNA) has been used extensively in previous studies of insects at the ordinal level [9,11,12,14,15,28,31,37–40]. However, analysis of 28S rDNA (also known as large subunit ribosomal rDNA, LSU rDNA) has not been employed in previous studies of insect phylogeny as generally as 18S has been. And in the cases of including 28S rDNA as one of the molecular markers, only partial segments, which vary from approximately 350 to 2000 nt, have been examined. It has been suggested that 28S rDNA contains significant phylogenetic signal for studying wide-ranging relationships [41–44]. Additionally, 28S rDNA shares many features with 18S rDNA, such as dramatic length variations, but is approxi-

mately two times the length of 18S rDNA and includes more variable regions, therefore representing a great suite of available data. The amplification of 28S rDNA is more likely interfered by the hairpin structures or tandem replicates of single nucleotides or oligonucleotides. As a result of these challenges, the application of complete 28S rDNA sequences in systematics has been hampered.

Among the aforementioned studies, only a few have employed information regarding the secondary structure of rRNAs [11,12,14,15]. In fact, phylogenetic studies based on rDNAs can benefit considerably from information regarding rRNA secondary structure. First, the secondary structure of rRNA can be used to improve alignments and thus, improve the accuracy of tree construction [45–48]. The length variation in these sequences leads to ambiguous alignments, i.e., alternative arrangements of gaps. In addition, the hyper-length variation of some local regions may even result in incorrectly determined positional homology at a large scale [12]. For example, an length variable region (LVR) in the V4 region of 18S rDNA that does not exist in some species, such as *Pandinus imperator* (Arachnida, Scorpiones) [Genbank:AY210831] ranges up to a length of 1,349 nt in *Cubaris murina* (Crustacea, Isopoda) [Genbank:AJ287064] [49]. As length variation increases, it becomes increasingly difficult or even impossible to infer optimal alignment from multiple sequences using computational methods and manual procedures. Therefore, the alignment of sequences that include hyper-length-variable regions can be more problematic. Second, some LVRs can serve as synapomorphies for certain monophyletic groups. The members in a clade may share the same length or the same tendency for elongation of LVRs [12,50].

A complete comparative study on the secondary structure of 18S rRNAs among insect orders was previously carried out by Gillespie et al. [51], Misof et al. [48], and Xie et al. [12]. For 28S rRNAs, there are two numbering systems for LVRs, which were alternatively referred to as D (divergent) domains [52,53] or expansion segments [54,55]. Twenty-two variable regions have been recognized for the major eukaryotic lineages [56], and some of these regions are hyper-length-variable regions. Among the insects, the complete or nearly complete secondary structures of 28S rRNAs have thus far been published for only a few species: *Drosophila melanogaster* (Diptera, Drosophilidae) [55,56], *Aedes*

albopictus (Diptera, Culicidae) [56,57], *Acyrtosiphon pisum* (Hemiptera, Aphididae) [58], *Tenebrio* sp. (Coleoptera, Tenebrionidae) [59], *Apis mellifera* (Hymenoptera, Apidae) [60] and *Synthemis eustalacta* (Odonata, Synthemistidae) [61], and a comparative study of the full secondary structures of 28S rRNAs among all orders of insects is still lacking.

Herein, we examine the controversial position of Zoraptera utilizing complete 18S and 28S rDNA sequences in independent studies on secondary structure and phylogeny. In this study, we positioned all of the variable regions of the 28S rDNA sequences of insects and refined the boundaries of the conserved motifs based on the principles of co-variation [62–64] and compensatory or semi-compensatory substitution [65]. The regions of the 28S rDNA and 18S rDNA sequences with conserved lengths were used to reconstruct a phylogeny for Insecta with particular emphasis on the zorapterans, attempting to identify putative autapomorphies or synapomorphies for certain lineages.

Materials and Methods

Taxon Sampling

A dataset comprised of 28S and 18S rDNA sequences from 67 species was compiled representing all orders within the Insecta except for the Strepsiptera (Table S1), due to its only half sequenced length of 28S rDNA. The complete 28S rDNA sequences of 20 species and 18S rDNA sequences of 10 species were sequenced for the first time in this study. The newly acquired 28S rDNA sequences included two sequences for each of the orders Zoraptera, Archaeognatha, Thysanoptera, Neuroptera, and Psocoptera; one sequence from each of the orders Phthiraptera, Embiodea, Zygentoma, Megaloptera, and Trichoptera; and five sequences from Hemiptera.

Molecular Experiments

Depending on the individual size of the sampled species, genomic DNA was extracted from either thorax tissue or the whole body except for the abdomen of ETOH-preserved insect specimens. Total genomic DNA was isolated using the CTAB-based method [66]. The primer sets used for amplification as well as sequencing were listed in Table S2. The functions of these primers were annotated in File S1. The PCR protocol for 28S rDNA included an initial denaturation at 94°C for 1 minute, followed by 30 cycles of 30 seconds at 94°C, 30 seconds-1 minute at 48–55°C and 1–2 minutes at 72°C, ending with a final extension at 72°C for 8–10 minutes. The thermal cycling program for 18S rDNA followed Johnson and Clayton [67]. All fragments were sequenced in both directions with the HiSeq 2000 sequencing system. A more detailed description of the molecular experiments was provided in the supplementary material File S1.

Alignment and Phylogenetic Analysis

Sequence assembly was carried out using BioEdit 7.0 [68], and MEGA 5.01 [69], DAMBE 4.5.32 [70,71], and Mesquite 2.75 [72] were used to align, connect and transform the format, respectively. Weblogo 3.0 was used to consent the sequences of expansion segment D3-4 [73,74]. Reconstruction of secondary structure was realized by thermodynamic folding using RNAs-structure 5.3 [75] and comparative methods [12,45,48,54,62–65,76]. A more detailed description of the reconstruction of the secondary structure model was provided in the supplementary material File S1. Inkscape 0.48.2 was used for drawing the secondary structure (<http://inkscape.org/>). The secondary structure model of insect 18S rRNA followed published data [12]. The numbering system for LVRs of 28S rRNA followed the D system,

which roughly includes thirteen D domains [52,53]. All sequences were initially aligned using CLUSTAL X 2.0 software [77] and were then checked and corrected manually referring to the secondary structure models for 18S and 28S rRNAs. Nucleotides positions within which positional homology cannot be unambiguously aligned were eliminated during the process of phylogenetic reconstruction. The data matrix is attached as Dataset S1.

MrBayes 3.1.2 [78,79] was used for Bayesian analysis. jModeltest 0.1.1 [80] was used to choose an appropriate model of substitution and GTR+G+I was selected as the best model for the data matrix. We used the parallelized version of MrBayes run on a graphics processing unit (GPU) [81] to speed up the calculation, achieving an approximately thirty times greater efficiency according to Nvidia GTX 580. The number of generations was 5,000,000, and the sampling frequency was 100. The average standard deviation of split frequencies fell below 0.01 after 1,901,000 generations, and the generations before generation 1,901,000 were burned-in. ML analysis was performed using Treefinder version 2011 [82]. The model GTR [Optimum, Empirical]:G [Optimum]:5 was determined by the program to be the best one. The number of bootstrap replicates was 1000. The other parameters were used with their default values.

Results and Discussion

Complete 28S rDNA sequences of Zoraptera, Embiodea, Thysanoptera, Psocoptera, Phthiraptera, Neuroptera, and Megaloptera were provided for the first time in this study. These new data make each order of Insecta has at least one complete 28S rDNA sequence, except for Strepsiptera. These newly obtained 28S rDNA sequences make it possible to reconstruct the phylogeny of insects based on complete 18S and 28S rDNA sequences together. Besides, these complete sequences of 28S rDNAs also provide a background for further comparative studies of secondary structures of 28S rRNAs within each order of Insecta.

Both the monophyly of the Polyneoptera and the interrelationships within the Polyneoptera have been debated for quite some time. Results may vary based on the same type of data, such as morphological [4,5,19,31,83,84] or rDNA data [12,14,15,28] as well as between different types of evidence, such as mitochondrial and nuclear protein-coding gene (PCG) analyses [13,85]. In recent years, EST [86], nuclear PCG [13,87,88], and mitochondrial genome sequences [85,89] have been used to examine the order-level phylogeny of insects. Nevertheless, due to incomplete taxon sampling, incomplete data, and other hindrances, the positions of most orders within the Polyneoptera remain uncertain. In this study, the phylogenetic results based on the length-stable regions of complete rDNAs are summarized in Figure 2. The consensus tree simultaneously recovered many clades indicated separately in various studies, such as Notoptera (which include the Mantophasmatodea and Grylloblattodea) [8–10], Plecoptera and Dermaptera [11,14,15,17], Embiodea and Phasmatodea [9,13,29,90], and Condylognatha (which include the Hemiptera and Thysanoptera) [4,5,13,15,21,28,91].

Position of the Zoraptera

The clade (Zoraptera+Dictyoptera) received a posterior probability of 87% in the BI analysis and a bootstrap value of 52% in the ML analysis. This bootstrap value, and support in general, is rather low. However, the same topology exists in both the BI and ML results. The low support value may be due to the unique evolutionary patterns among the Zoraptera, such as an accelerated substitution rate and unique insertions and deletions (indels), as have been noted in studies using 18S rDNA as a molecular marker

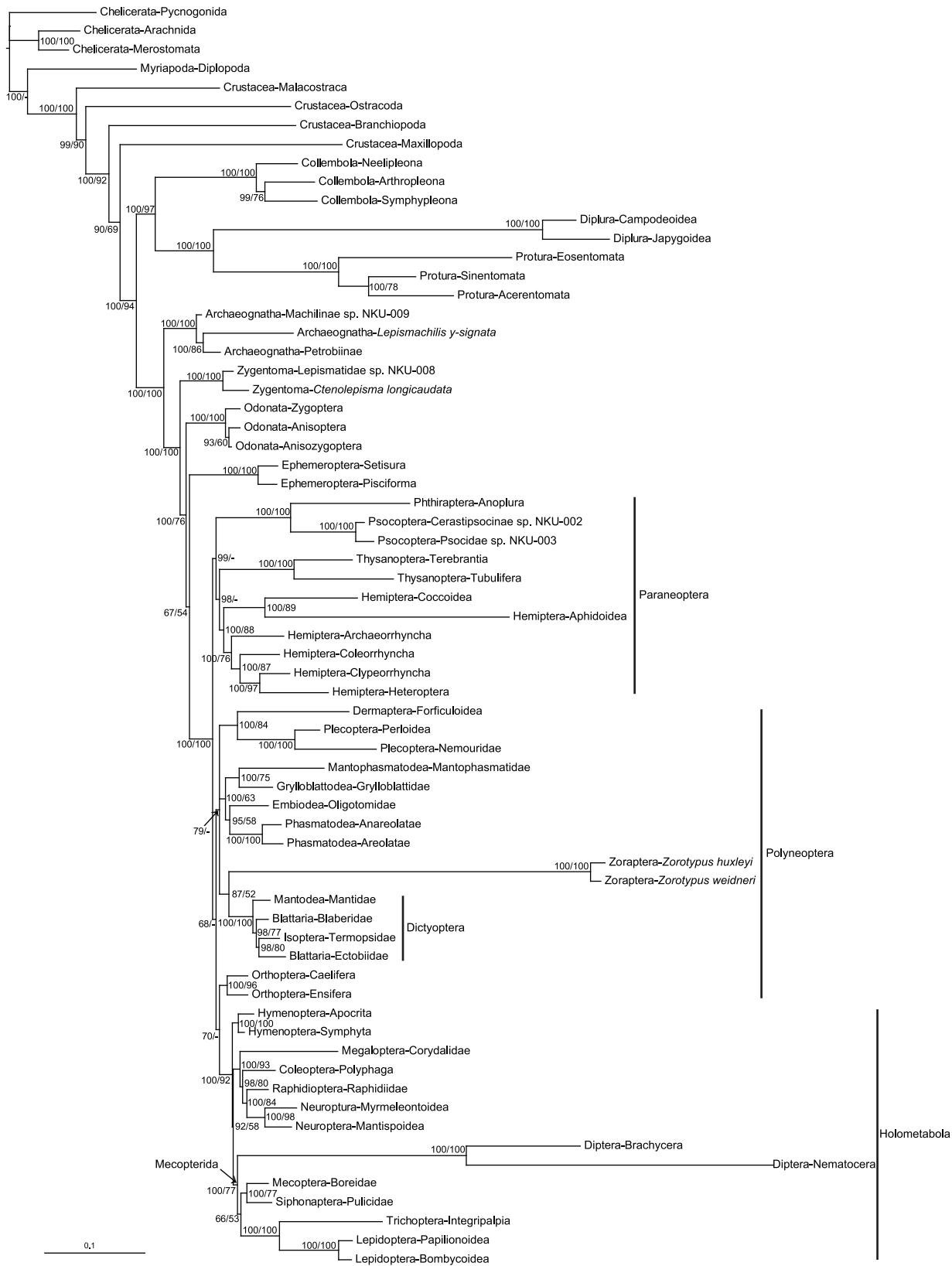


Figure 2. Phylogenetic tree inferred from the regions of the complete 18S and 28S rDNA sequences with conserved lengths. The numbers associated with the nodes are posterior probability values (first number) and bootstrap values (second number) obtained by Bayesian/ML analysis respectively. The lengths of the branches follow the phylogram of the Bayesian tree. doi:10.1371/journal.pone.0053679.g002

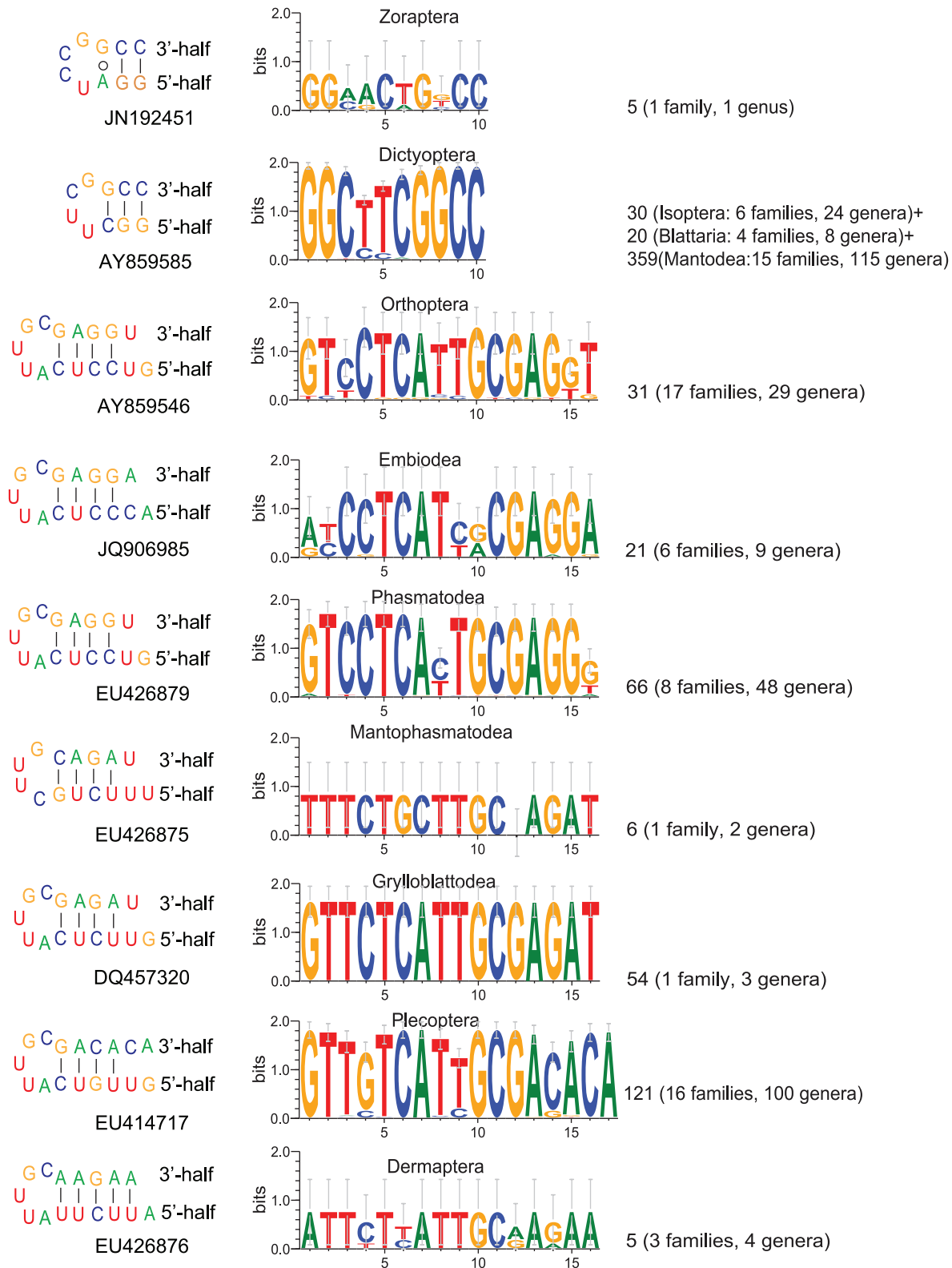


Figure 3. Consensus of the sequences of expansion segment D3-4 based on the homologues in GenBank. The left column is the secondary structure of the D3-4 'marker-box' of each polyneopteran order. The accession number in the bottom of each regional secondary structure stands for the corresponding sequence which is the same to the consensus sequence of each polyneopteran order or superorder. The middle column is the consensus result of the homologues in each polyneopteran order or superorder. The abscissa stand for the number of the bases, while the ordinate stand for the proportion of information content provided by each base in the same position. The right column is the number of sequences based on.

doi:10.1371/journal.pone.0053679.g003

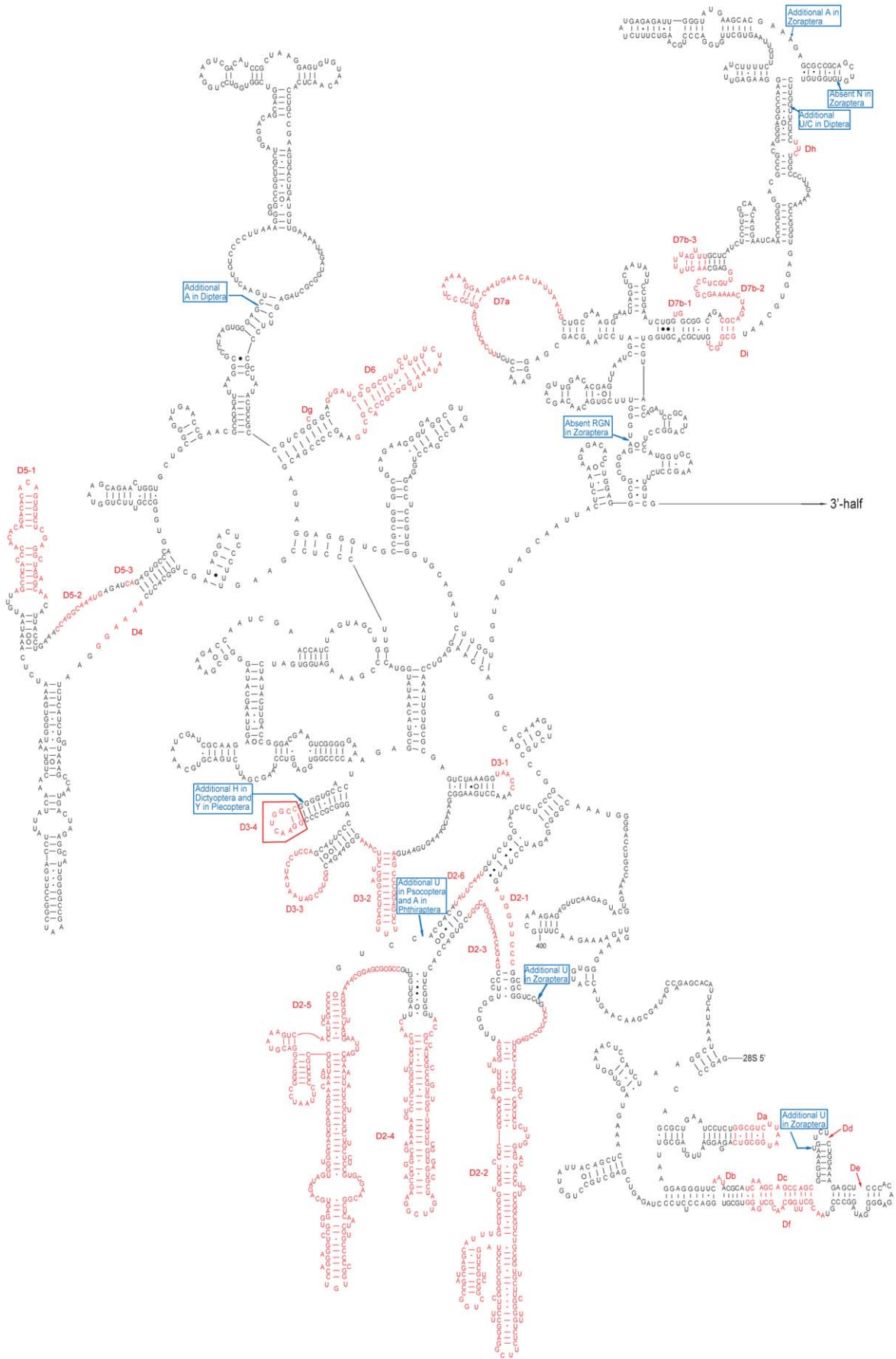


Figure 4. Secondary structure model of the 28S rRNA 5'-half of Zoraptera. This sequence is from *Zorotypus huxleyi* [Genbank:JN192451]. The length-variable regions are indicated in red. And the unique indels are marked with green color. The D3-4 box was highlighted with thick red lines. The Da–Dj numbering system for LVRs, which has not been taken into account previously, is a supplementary system to the D1–D12 coding system. Base pairing is indicated as follows: standard canonical pairs by lines (C–G, G–C, A–U, U–A); wobble G–U pairs by dots (G–U); A–G and A–C pairs by open circles (A–G, A–C); other non-canonical pairs by filled circles (e.g., U●U). doi:10.1371/journal.pone.0053679.g004

[11,15]. The special attributes of the Zoraptera can be observed in the BI results (Figure 2), in which the Zoraptera lineage is a rather long branch. The distinctive quality of Zoraptera rDNAs may be one of the reasons for the disputed status of this group in previous studies.

In the present study, independent secondary structure evidence also supports a sister relationship between the Dictyoptera and Zoraptera (Figure 3). Among all of the detected LVRs, we found that D3-4 is particularly special. The length and/or sequence of D3-4 is conservative in each order of Polyneoptera, but it can be divided into two types among the different orders of the supercohort (Figure 3). The Zoraptera and Dictyoptera shared one type of 10 nucleotides length, and the Plecoptera, Dermaptera, Orthoptera, Phasmatodea, Embiidea, Grylloblattodea, and Mantophasmatodea shared the other type of 16 ± 1 nucleotides length. There is a unique insertion special for Plecoptera around 3'-end, and there is a unique deletion of G special for Mantophasmatodea near the 3'-end. The length differences between the 10 nucleotides and the 15–17 nucleotides are extremely significant (File S1). This attribute makes D3-4 a good marker to indicate relationships within Polyneoptera. The clade (Zoraptera+Dictyoptera) shares a unique 10 nt long box in this expansion segment, and the D3-4 sequences of these two groups are also similar, i.e., GGYYYMKGCC in Dictyoptera and GGMRCWGBCC in Zoraptera. In Figure 4, if considering there is only one base pair between the D3-4 and the single insertion in some groups, these two parts can be alternatively viewed as a whole. However, even if this possibility is taken into account, the length and sequence of D3-4 is still most similar to those of Dictyoptera. Therefore, this alternative consideration on the boundary of D3-4 has no bias according to the competitive hypotheses of the phylogenetic position of Zoraptera.

Length-stable and length-variable rDNA regions are under different evolutionary constraints [46]. Most LVRs are distributed on the surface of the tertiary structure of rRNA, far from the functional site [49]. Although weak constraints of negative selection might also lead to homoplastic patterns in non-related taxa, the extremely significant differences between the lengths and similarity of the sequences make the probability quite low that the same pattern of D3-4 has originated independently in Zoraptera and Dictyoptera. Accordingly, a Zoraptera+Dictyoptera clade may have some credence if corroborated by future analyses and other forms of data. Besides the significance in understanding the evolutionary history of hexapods, resolving the phylogenetic position of Zoraptera is also very important because the resolving influences findings regarding shifts in diversification of hexapods [92].

The Dictyoptera represent the most universally accepted supra-ordinal grouping within the Polyneoptera. The three orders of Dictyoptera share several synapomorphies, including an extremely reduced ovipositor, mostly internal valvulae, perforated tentorium, and asymmetrical male genitalia [16,21,93]. At the molecular level, the length of one LVR in the 18S rRNA secondary structure was discovered to represent a synapomorphy of the Dictyoptera [12]. In the comparative analysis of 28S rRNA, a unique length of expansion segment D4 of 14 nt was found to be shared by all three orders of Dictyoptera and, thus, constitutes a further molecular

autapomorphy for Dictyoptera. Additionally, the three orders of Dictyoptera exhibit the same lengths for another 18 out of 40 LVRs.

The reduction of the ovipositor is a highly homoplastic character which occurs innumerable and independently across Polyneoptera, within orders, within families, sometimes even within a single genus, and, indeed, across all Insecta. Similarly, “mostly internal valvulae” also occurs many times independently. Additionally, Zoraptera do not have a perforated tentorium, and male genitalia are symmetrical. At present there is no compelling morphological evidence for a Zoraptera+Dictyoptera clade, and it will require future testing by expanded data sets.

Position of the Embiidea

Among the Polyneoptera, Embiidea is another order for which the position remains controversial. From a morphological perspective, Embiidea was hypothesized to be either a sister group of Plecoptera [21,31], Phasmatodea [18,29,90,93–96], Dermaptera [84], or Zoraptera [6,17,24–27]. However, in the past several years, most results based on molecular data support the Embiidea and Phasmatodea hypothesis [9,11,13,97]. The phylogenetic results of this study based on complete 18S and 28S rDNA sequences also support these two orders as closely related (Figure 2). Although the bootstrap value for Embiidea+Phasmatodea is low in this analysis, it is congruent with the Bayesian result of this study and the results of other molecular phylogenetic studies. In addition, these analyses supported the sister relationships between Grylloblattodea and Mantophasmatodea. These two monophyletic groups further formed a clade, which is congruent with Kjer et al. [11].

The clade ((Grylloblattodea+Mantophasmatodea)+(Embiidea+Phasmatodea)) was supported with high posterior probability values (100%). A clade including Grylloblattodea, Phasmatodea and Embiidea has been suggested based on 18S rDNA sequences [15], but without sampling Mantophasmatodea. Similarly, a clade including Grylloblattodea, Mantophasmatodea, and Phasmatodea has recently been indicated based on mitochondrial genomes [85], but without sampling Embiidea. This superordinal grouping, here called the “Mecynoptera hypothesis”, is novel to our study and deserves critical investigation. It is unclear to what extent this grouping may or may not be supported by existing paleontological data.

Based on morphological studies addressing fossils or living groups, Orthoptera had been viewed to have a close relationship with Phasmatodea [19,21,31,98–101]. However, in molecular phylogenetic studies, the position of Orthoptera has often been indeterminate [9,11,13–15,28,85]. In our study, the position of Orthoptera is shown as unresolved in the results of both the BI and ML analyses. However, Orthoptera share the same length and similar sequences of D3-4 with the Mecynoptera clade. Thus, these five orders may constitute a potential group, with Orthoptera as basal, but the phylogenetic signal in the available rDNA sequences is not sufficiently strong to make a definitive conclusion.

Taxonomy of Polyneoptera

Handlirsch [102] first suggested the existence of two subclasses of Polyneoptera: Orthopteroidea and Blattaeformia. Subsequently,

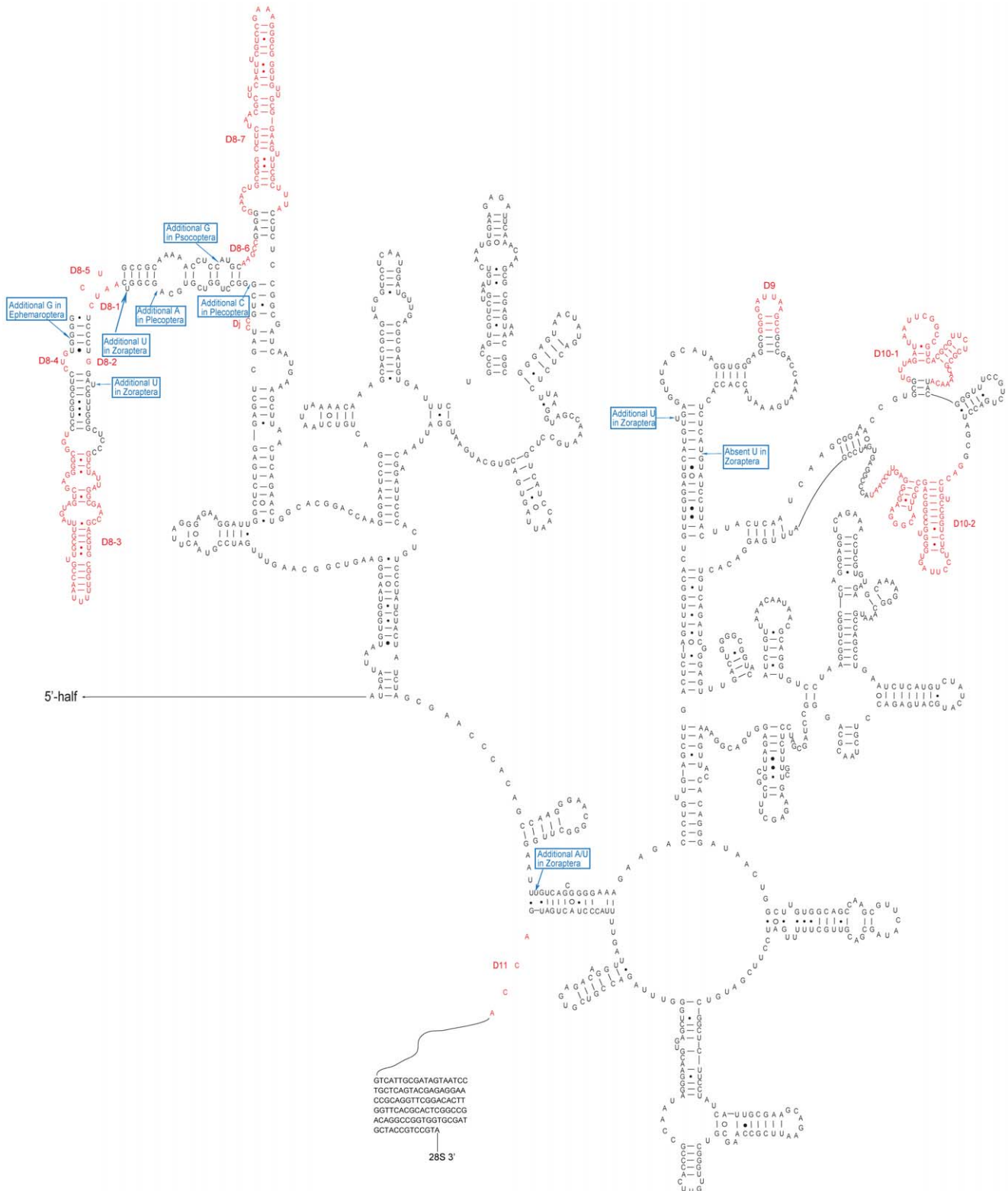


Figure 5. Secondary structure model of the 28S rRNA 3'-half of Zoraptera. This sequence is from *Zorotypus huxleyi* [Genbank:JN192451]. The length-variable regions are indicated in red. And the unique indels are marked with green color. The Da–Dj numbering system for LVRs, which has not been taken into account previously, is a supplementary system to the D1–D12 coding system. Base pairing is indicated as follows: standard canonical pairs by lines (C–G, G–C, A–U, U–A); wobble G–U pairs by dots (G–U); A–G and A–C pairs by open circles (A G, A C); other non-canonical pairs by filled circles (e.g., U●U).

doi:10.1371/journal.pone.0053679.g005

due to changes in hierarchical systems of subclasses, infraclasses, or superorders, the name Orthopteroidea has come to have different meanings for different researchers (Table 1). Orthopteroidea can indicate a group as small as consisting of only the Orthoptera, Phasmatodea, and Embiodea [96], or as large as including all of the Polyneoptera [9,31,103]. According to the results of this study, we tentatively suggest that there might be recognized a revised Dermoplecopterida for (Plecoptera+Dermaptera), the Blattopteroidea as (Dictyoptera+Zoraptera), and the Mecynoptera equal to the ((Embiodea+Phasmatodea)+(Grylloblattodea+Mantophasmatodea)).

The relationships between these four putative lineages, i.e., Dermoplecopterida, Blattopteroidea, Orthoptera, and Mecynoptera, are not effectively resolved in this study. In the future, combining rDNA results with analysis of nuclear PCGs may contribute to completely resolving the phylogeny of the Polyneoptera. It is also possible that the unresolved nodes within the Polyneoptera may be due to ancient rapid radiation [104–106], as rapid diversification would result in particularly short interdivergence times within which characters could accumulate. In fact, even phylogenomic studies based on EST may include unresolved or weakly supported nodes [105–114].

Phylogeny of the Eumetabola

The Paraneoptera are comprised of the Psocoptera (book lice and bark lice), Phthiraptera (lice), Hemiptera (true bugs), and Thysanoptera (thrips). Psocoptera and Phthiraptera are together referred to as the superorder Psocodea [115], and the monophyly of the Psocodea has been supported by numerous studies [11,13,14,28,116]. According to the phylogenetic relationships within the Paraneoptera, most morphological studies consistently view Thysanoptera as the sister group to Hemiptera, and these two orders are referred to as the superorder Condylognatha [4,5,21,29,31,91,116]. Compared to results based on morphological data, results based on molecular analyses may differ from each other. Among results from 18S rDNA analyses, the position of Thysanoptera has been shown to be close to either Psocodea [31] or Hemiptera [28]. In studies based on multiple nuclear genes, Thysanoptera was shown to be close to Hemiptera, supported by a high posterior probability in BI [13,15]. In this study, the monophyly of the Condylognatha was confirmed with a high probability in the Bayesian inference. Thus, the monophyly of the Condylognatha is now supported by evidence from morphological, nuclear PCG, and rDNA analyses.

Within the Holometabola, the phylogenetic results strongly supported a basal position for the Hymenoptera. The other orders of Holometabola are further segregated into two principal clades: Mecopterida (= Diptera+(Mecoptera+Siphonaptera)+(Trichoptera+Lepidoptera)); and Coleoptera+Neuropterida (= Megaloptera+Raphidioptera+Neuroptera) (Figure 2). This may be the first time that evidence from rDNAs has been consistent with that from nuclear PCGs [13,87,88,117]. In fact, if the partially sequenced rDNAs of Strepsiptera, as mentioned in the part Taxon Sampling of Material and Methods, are included in the taxon sampling of this study, its phylogenetic position is the sister group to Coleoptera in the Bayesian tree (Figure S1). This would make the phylogenetic result of Holometabola part more congruent with the result based on morphology [118] or other molecular markers [13,87,88,117,119,120].

The Impact of LVRs on Alignment and Phylogeny

A general secondary structure model for insect 28S rRNA was reconstructed (Figures 4 and 5, Figures S2 and S3), and there were a total of 40 LVRs detected. Most of the LVRs consisted of single

strands located around lateral or terminal bulges (Figures 4 and 5), while the others were internal bulges or multi-branched loops. The length variation of each LVR was summarized in Tables S3 (18S) and S4 (28S). According to the extent of length variability for each LVR of 28S rDNA, D2, D3, D5, D7, D8, and D10 were the most extensive LVRs or hyper-variable regions (Table S4). Variations in these six expansion segments accounted for approximately 87.5% of the total variability among all LVRs. Among the expansion segments, D8-3 was the most variable, ranging from 2 nt in Diptera [Genbank:L78065] to 524 nt in Neuroptera [Genbank:JQ259053]. Among the investigated groups, Phthiraptera [Genbank:JQ309932] and Strepsiptera [Genbank:HM156704] exhibited the most extensive LVRs of 28S rDNA.

The accuracy and quality of rDNA alignments are critical factors in molecular phylogenetic studies [14,12,45,47,50,56,121–123]. Nucleotide positions for which positional homology cannot be unambiguously or correctly aligned should be eliminated during the process of phylogenetic reconstruction [14,45,50]. In this work, the impact of LVRs on the alignment and, thus, on the phylogeny was considered. The tree obtained from the automatic alignment results for the combined 18S and 28S rDNA sequences (Dataset S2, Figures S4 and S5) yielded results presenting many contradictions compared to widely accepted opinions. For example, Dermaptera is imbedded within Holometabola; Archaeognatha is the sister group of Odonata; the Holometabola are paraphyletic; and Coleoptera and Ephemeroptera are sister groups. Therefore, in this case study, the comparative phylogenetic results reinforce the opinion that due to the improved alignment, the performance of rDNA regions with conserved lengths can be superior to that of the original sequences (Figure 2, Figures S4 and S5). However, it is imperative that the delimitation of conserved and variable regions be ascertained. Otherwise, more informative sites will be lost during the abandonment of ambiguous regions.

In this study, the secondary structure model of 28S rRNA reconstructed for eukaryotes [56,76] was specifically refined for insects. For a specific taxon, the hyper-variable regions summarized for eukaryotes can be divided into several small regions. After the comparative analysis of insect 28S rRNAs, six hyper-variable regions in the secondary structure model for eukaryotes were split into a number of sub-regions. For example, the hyper-variable region D8 was divided into eight sub-regions in this study (Figure 5). With respect to the phylogeny of the lower categories, these sub-regions can be further subdivided [59,121]. Group-specific analysis will be helpful in exploring possible evidence of common origins based on the secondary structure of rRNA. The expansion segment D3-4 of 28S rRNA is such a case.

In this study, the substitution models of base pairs were not applied. The use of specific mixed RNA/DNA substitution models in insect rRNA phylogenetics might not lead to more reasonable results, most likely due to substitutional saturation in unpaired regions [123,124]. In addition, compared to the biological background of structural studies of macromolecules, the substitution models of base pairs provided by current phylogenetic programs are not complete. Generally, current substitution models treat base pairs as only three different types, the canonical base pairs, GU-UG pairs, and all of the other modes of base pairs. The AC/CA and AG/GA base pairs, which exist subjectively in three dimensional structures of rRNAs [125], are not viewed as regular base pairs by current phylogenetic programs. The improvement of substitution models of base pairs in the future deserves being tested further.

Conclusions

The results of this work provided novel evidence to support the close relationship between Zoraptera and Dictyoptera from the views of secondary structure and phylogeny independently. Besides, the present analysis first provided the direct evidence to support the monophyly of the clade ((Embiodea+Phasmatodea) + (Grylloblattodea+Mantophasmatodea)). The results of this work also reached the highest congruence with the results of previous molecular phylogenetic studies of insects based on nuclear PCGs, especially those of Holometabola. Accordingly, these results for understanding the higher-level relationships and diversification of insects are of critical importance.

This study can also serve as one more case to support that, the LVRs can remarkably affect the result of alignment, and thereby the result of phylogeny. Based on the secondary structure model of the 28S rRNA reconstructed in this study, all of the LVRs were removed *a priori* and the complete rDNAs were aligned unambiguously. Due to the improved alignment, the performance of rDNA regions with conserved lengths can be superior to that of the original sequences.

Supporting Information

Figure S1 Bayesian tree inferred from analysis of the complete 18S and 28S rDNA sequences with conserved lengths. The sequences of Strepsiptera were included in this phylogenetic analysis. The number of generations was 10,000,000, the sampling frequency was 100, and the first 7,060,000 generations was discarded as “burnin”. This is a majority rule consensus tree, and the Bayesian posterior probability is given above each corresponding node.
(JPG)

Figure S2 Secondary structure model of the 28S rRNA 5'-half of Insecta. This sequence is from *Drosophila melanogaster* [GenBank:M21017]. The length-variable regions are indicated in red. And the unique indels are marked with green color. The D3-4 box was highlighted with thick red lines. Base pairing is indicated as follows: standard canonical pairs by lines (C-G, G-C, A-U, U-A); wobble G·U pairs by dots (G·U); A·G and A·C pairs by open circles (A G, A C); other non-canonical pairs by filled circles (e.g., U·U).
(JPG)

Figure S3 Secondary structure model of the 28S rRNA 3'-half of Insecta. This sequence is from *Drosophila melanogaster* [GenBank:M21017]. The length-variable regions are indicated in red. And the unique indels are marked with green color. Base pairing is indicated as follows: standard canonical pairs by lines (C-G, G-C, A-U, U-A); wobble G·U pairs by dots (G·U); A·G and A·C pairs by open circles (A G, A C); other non-canonical pairs by filled circles (e.g., U·U).
(JPG)

Figure S4 Tree obtained by Bayesian analysis of the complete 18S+28S rDNAs. The rDNA sequences were aligned by Clustal X, not adjusted by manual according to the secondary structures of the rDNAs. The number of generations was 5,000,000, the sampling frequency was 100, and the first 364,000 generations was discarded as “burnin”. This is a majority

rule consensus tree, and the Bayesian posterior probability is given above the node.

(JPG)

Figure S5 ML tree based on the automated alignment result of complete 18S+28S rDNAs. Numerals above the nodes are bootstrap values.
(JPG)

Table S1 Taxa sampling of 18S and 28S rDNAs. Accession numbers marked with an asterisk are newly sequenced for 28S rDNAs (and for 18S rDNAs when needed) in the present study. The sequences of Strepsiptera were not included in the phylogenetic analyses, but in the study of secondary structure rRNAs.
(XLS)

Table S2 Primer sets used for amplification and sequencing. The primer sets of DF1-FD1 and EE-GG were newly designed in this study and can be used universally for insects. The rest were specifically designed for some groups.
(XLS)

Table S3 The detailed information of the variable regions of the 18S rDNA for each taxon.
(XLS)

Table S4 The detailed information of the variable regions of the 28S rDNA for each taxon.
(XLS)

Dataset S1 The data matrix of the conservative parts of the 18S and 28S rDNA. This matrix only included the conservative region of the combined 18S and 28S rDNA. The variable regions were excluded in advance referred to the secondary structure of the 18S and 28S rRNA respectively.
(FAS)

Dataset S2 The data matrix of the combined 18S and 28S rDNA. This matrix was generated through alignment programs, which included the complete 18S and 28S rDNAs.
(FAS)

File S1 The detailed annotation of methods.
(DOC)

Acknowledgments

We thank Prof. Karl Kjer (Rutgers University, USA) for reviewing the manuscript and providing very helpful advices. We thank Dr. Jakob Damgaard (University of Copenhagen, Denmark), Prof. Ge-Xia Qiao (Institute of Zoology, Chinese Academy of Science, China), Prof. San-An Wu (Beijing Forestry University, China), and Ms. Zi-chen Wang (Guangzhou University of Chinese Medicine, China) for providing specimens. We thank Ms. Shuang Zhou, Mr. Geng-Yan Bai and Mr. Yu-Chen Zhang (Nankai University, China) for participating in reconstructing and illustrating the secondary structure. We thank for the Division of Entomology, University of Kansas Natural History Museum.

Author Contributions

Conceived and designed the experiments: QX. Performed the experiments: YHW YW. Analyzed the data: YHW QX HYW. Contributed reagents/materials/analysis tools: MSE JAR KD. Wrote the paper: YHW QX MSE JAR WJB.

References

- Martynov AV (1925) Ueber zwei Grundtypen der Flügel bei den Insekten und ihre Evolution. *Z Morphol Okol Tiere* 4: 465–501.
- Crampton GC (1938) The interrelationship and lines of descent of living insects. *Psyche* 45: 165–181.
- Hennig W (1953) Kritische Bemerkungen zum phylogenetischen System der Insekten. *Beitr Ent* 3: 1–85.
- Hennig W (1969) *Die Stammesgeschichte der Insekten*. Frankfurt am Main: Waldemar Kramer.

5. Hennig W (1981) *Insect Phylogeny*. New York: Wiley J & Sons.
6. Grimaldi DA, Engel MS (2005) *Evolution of the Insects*. Cambridge: Cambridge University Press.
7. Klass KD, Zompro O, Kristensen NP, Adis J (2002) Mantophasmatodea: a new insect order with extant members in the Afrotropics. *Science* 296: 1456–1459.
8. Engel MS, Grimaldi DA (2004) A new rock crawler in Baltic amber, with comments on the order (Mantophasmatodea: Mantophasmatidae). *Am Mus Novit* 3431: 1–11.
9. Terry MD, Whiting MF (2005) Mantophasmatodea and phylogeny of the lower neopterous insects. *Cladistics* 21: 240–257.
10. Arillo A, Engel MS (2006) Rock crawlers in Baltic amber (Notoptera: Mantophasmatodea). *Am Mus Novit* 3539: 1–10.
11. Kjer KM, Carle FL, Litman J, Ware J (2006) A molecular phylogeny of Hexapoda. *Arthropod Syst Phylogeny* 64: 35–44.
12. Xie Q, Tian XX, Qin Y, Bu WJ (2009) Phylogenetic comparison of local length plasticity of the small subunit of nuclear rDNAs among all Hexapoda orders and the impact of hyper-length-variation on alignment. *Mol Phylogenet Evol* 50: 310–316.
13. Ishiwata K, Sasaki G, Ogawa J, Miyata T, Su ZH (2010) Phylogenetic relationships among insect orders based on three nuclear protein-coding gene sequences. *Mol Phylogenet Evol* 58: 169–180.
14. Kjer KM (2004) Aligned 18S and insect phylogeny. *Syst Biol* 53: 506–514.
15. Yoshizawa K, Johnson KP (2005) Aligned 18S for Zoraptera (Insecta): Phylogenetic position and molecular evolution. *Mol Phylogenet Evol* 37: 572–580.
16. Kristensen NP (1991) Phylogeny of extant hexapods. In: Naumann ID (chi fed.). *The Insects of Australia. A Textbook for Students and Research Workers*. Melbourne: Melbourne University Press, 125–140.
17. Yoshizawa K (2011) Monophyletic Polyneoptera recovered by wing base structure. *Syst Entomol* 36: 377–394.
18. Kristensen NP (1975) The phylogeny of hexapod “orders.” A critical review of recent accounts. *J Zool Syst Evol Res* 13: 1–44.
19. Beutel RG, Gorb SN (2006) A revised interpretation of the evolution of attachment structures in Hexapoda (Arthropoda), with special emphasis on Mantophasmatodea. *Arthropod Syst Phylogeny* 64: 3–25.
20. Rasnitsyn AP (1998) On the taxonomic position of the insect order Zorotypida = Zoraptera. *Zool Anz* 237: 185–194.
21. Boudreaux HB (1979) *Arthropod phylogeny, with special reference to insects*. New York: John Wiley and Sons.
22. Smithers CN (1991) Zoraptera. In: Naumann I.D. (chi fed.): *The Insects of Australia. A Textbook for Students and Research Workers*. Melbourne: Melbourne University Press, 410–411.
23. Kukalová-Peck J, Peck SB (1993) Zoraptera wing structures: evidence for new genera and relationship with the blattoid orders (Insecta: Blattellodea). *Syst Entomol* 18: 333–350.
24. Minet J, Bourgoin T (1986) Phylogenie et classification des Hexapodes (Arthropoda). *Cah Liaison* 63: 23–28.
25. Engel MS, Grimaldi DA (2000) A winged *Zorotypus* in Miocene amber from the Dominican Republic (Zoraptera: Zorotypidae), with discussion on relationships of and within the order. *Acta Geol Hispanica* 35: 149–164.
26. Engel MS, Grimaldi DA (2002) The first Mesozoic Zoraptera (Insecta). *Am Mus Novit* 3362: 1–20.
27. Rafael JA, Godoi FSP, Engel MS (2008) A new species of *Zorotypus* from eastern Amazonia, Brazil (Zoraptera: Zorotypidae). *Trans Kansas Acad Sci* 111: 193–202.
28. Misof B, Niehuis O, Bischoff I, Rickert A, Erpenbeck D, et al. (2007) Towards an 18S phylogeny of hexapods: Accounting for group-specific character covariance in optimized mixed nucleotide/doublet models. *Zoology* 110: 409–429.
29. Kristensen NP (1981) Phylogeny of insect orders. *Annu Rev Entomol* 26: 135–157.
30. Yoshizawa K (2007) The Zoraptera problem: Evidence for Zoraptera+Embiodea from the wing base. *Syst Entomol* 32: 197–204.
31. Wheeler WC, Whiting M, Wheeler QD, Carpenter JM (2001) The phylogeny of the extant hexapod orders. *Cladistics* 17: 113–169.
32. Simmons MP (2004) Independence of alignment and tree search. *Mol Phylogenet Evol* 31: 874–879.
33. Morgan MJ, Kelchner SA (2010) Inference of molecular homology and sequence alignment by direct optimization. *Mol Phylogenet Evol* 56: 305–311.
34. Yoshizawa K (2010) Direct optimization overly optimizes data. *Syst Entomol* 35: 199–206.
35. Yoshizawa K (2011) Problems in direct optimization. *Taxa, Proc Jpn Soc Syst Zool* 30: 17–29.
36. Caterino MS, Cho S, Sperling FA (2000) The current state of insect molecular systematics: a thriving Tower of Babel. *Annu Rev Entomol* 45: 1–54.
37. Carmean D, Kimsey LS, Berbee ML (1992) 18S rDNA sequences and the holometabolous insects. *Mol Phylogenet Evol* 1: 270–278.
38. Pashley DP, McPherson BA, Zimmer EA (1993) Systematics of holometabolous insect orders based on 18S ribosomal RNA. *Mol Phylogenet Evol* 2: 132–142.
39. Chalwatzis N, Hauf J, van de Peer Y, Kinzelbach R, Zimmermann FK (1996) 18S ribosomal RNA genes of insects: primary structure of the genes and molecular phylogeny of the Holometabola. *Ann Entomol Soc Am* 89: 788–803.
40. Whiting MF, Carpenter JC, Wheeler QD, Wheeler WC (1997) The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Syst Biol* 46: 1–68.
41. von Reumont BM, Meusemann K, Szucsich NU, Dell’Ampio E, Gowri-Shankar V, et al. (2009) Can comprehensive background knowledge be incorporated into substitution models to improve phylogenetic analyses? A case study on major arthropod relationships. *BMC Evol Biol* 9: 119.
42. De Rijk P, Vandepuer Y, Vandenbroeck I, Dewachter R (1995) Evolution According to Large Ribosomal-Subunit RNA. *J Mol Evol* 41: 366–375.
43. Mallatt JM, Garey JR, Shultz JW (2004) Ecdysozoan phylogeny and Bayesian inference: first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin. *Mol Phylogenet Evol* 31: 178–191.
44. Mallatt JM, Giribet G (2006) Further use of nearly complete 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. *Mol Phylogenet Evol* 40: 772–794.
45. Kjer KM (1995) 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* 4: 314–330.
46. Buckley T, Simon C, Flook P, Misof B (2000) Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. *Insect Mol Biol* 9: 565–580.
47. Hickson RE, Simon C, Perrey SW (2000) The performance of several multiple-sequence alignment programs in relation to secondary-structure features for an rRNA sequence. *Mol Biol Evol* 17: 530–539.
48. Misof B, Niehuis O, Bischoff I, Rickert A, Erpenbeck D, et al. (2006) A hexapod nuclear SSU rRNA secondary-structure model and catalog of taxon-specific structural variation. *J Exp Zool Part B* 306B: 70–88.
49. Xie Q, Lin J, Qin Y, Zhou J, Bu WJ (2011) Structural diversity of eukaryotic 18S rRNA and its impact on alignment and phylogenetic reconstruction. *Protein & Cell* 2: 161–170.
50. Xie Q, Tian Y, Zheng LY, Bu WJ (2008) 18S rRNA hyper-elongation and the phylogeny of Ephemeroptera (Insecta: Hemiptera). *Mol Phylogenet Evol* 47: 463–471.
51. Gillespie JJ, McKenna CH, Yoder MJ, Gutell RR, Johnston JS, et al. (2005): Assessing the odd secondary structural properties of nuclear small subunit ribosomal RNA sequences (18S) of the twisted-wing parasites (Insecta: Strepsiptera). *Insect Mol Biol* 14: 625–634.
52. Hassouna N, Mithot B, Bachelier JP (1984) The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. *Nucleic Acids Res* 12: 3563–3583.
53. Ellis R, Sulston J, Coulson A (1986) The rDNA of *C. elegans*: sequence and structure. *Nucleic Acids Res* 14: 2345–2364.
54. Clark CG, Tague BW, Ware VC, Gerbi SA (1984) *Xenopus laevis* 28S ribosomal RNA: a secondary structure model and its evolutionary and functional implications. *Nucleic Acids Res* 12: 6197–6220.
55. Hancock JM, Tautz D, Dover GA (1988) Evolution of the secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster*. *Mol Biol Evol* 5: 393–414.
56. Schnare MN, Damberger SH, Gray MW, Gutell RR (1996) Comprehensive Comparison of Structural Characteristics in Eukaryotic Cytoplasmic Large Subunit (23 S-like) Ribosomal RNA. *J Mol Evol* 256: 701–719.
57. Kjer KM, Baldrige GD, Fallon AM (1994) Mosquito large subunit ribosomal RNA: simultaneous alignment of primary and secondary structure. *BBA-Genet Struct Expr* 1217: 147–155.
58. Amako D, Kwon OY, Ishikawa H (1996) Nucleotide sequence and presumed secondary structure of the 28S rRNA of pea aphid: Implication for diversification of insect rRNA. *J Mol Evol* 43: 469–475.
59. Gillespie J, Cannone J, Gutell R, Cognato A (2004) A secondary structural model of the 28S rRNA expansion segments D2 and D3 from rootworms and related leaf beetles (Coleoptera: Chrysomelidae; Galerucinae). *Insect Mol Biol* 13: 495–518.
60. Gillespie J, Johnston J, Cannone J, Gutell R (2006) Characteristics of the nuclear (18S, 5.8 S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements. *Insect Mol Biol* 15: 657–686.
61. Letsch HO, Greve C, Kück P, Fleck G, Stocsits RR, et al. (2009) Simultaneous alignment and folding of 28S rRNA sequences uncovers phylogenetic signal in structure variation. *Mol Phylogenet Evol* 53: 758–771.
62. Gutell RR, Weiser B, Woese CR, Noller HF (1985) Comparative anatomy of 16S-like ribosomal RNA. *Prog Nucleic Acid Res Mol Biol* 32: 155–216.
63. Gutell RR, Cannone JJ, Konings D, Gautheret D (2000) Predicting U-turns in ribosomal RNA with comparative sequence analysis. *J Mol Biol* 300: 791–803.
64. Gutell RR, Lee JC, Cannone JJ (2002) The accuracy of ribosomal RNA comparative structure models. *Curr Opin Struct Biol* 12: 301–310.
65. Dixon MT, Hillis DM (1993) Ribosomal RNA Secondary Structure: Compensatory Mutations and Implications for Phylogenetic Analysis. *Mol Biol Evol* 10: 256–267.
66. Reinecke A, Karlovsky P, Zebitz CPW (1998) Preparation and purification of DNA from insects for AFLP analysis. *Insect Mol Biol* 7: 95–99.
67. Johnson KP, Clayton DH (2000) Nuclear and mitochondrial genes contain similar phylogenetic signal for pigeons and doves (Aves: Columbiformes). *Mol Phylogenet Evol* 14: 141–151.

68. Hall T: BioEdit version 7.1.7. [website: <http://www.mbio.ncsu.edu/bioedit/bioedit.html>. Accessed: 2012 Nov 25.].
69. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* 28: 2731–2739.
70. Xia X (2000) Data analysis in molecular biology and evolution. Boston: Kluwer Academic publishers, 276.
71. Xia X, Xie Z (2001) DAMBE: software package for data analysis in molecular biology and evolution. *J Hered* 92: 371.
72. Maddison WP, Maddison DR: Mesquite: a modular system for evolutionary analysis. Version 2.75. [http://mesquiteproject.org/Mesquite_Folder/docs/mesquite/manual.html].
73. Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* 18: 6097–6100.
74. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome research* 14: 1188–1190.
75. Reuter J, Mathews D (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* 11: 129.
76. Taylor DJ, Devkota B, Huang AD, Topf M, Narayanan E, et al. (2009) Comprehensive molecular structure of the eukaryotic ribosome. *Structure* 17: 1591–1604.
77. Larkin M, Blackshields G, Brown N, Chenna R, McGettigan P, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.
78. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
79. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
80. Posada D (2008) jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol* 25: 1253–1256.
81. Zhou J, Liu X, Stones DS, Xie Q, Wang G (2011) MrBayes on a Graphics Processing Unit. *Bioinformatics* 27: 1255.
82. Jobb G (2011) *TREEFINDER* version of March 2011. Munich, Germany. Distributed by the author at www.treefinder.de.
83. Kristensen NP (1995) Forty years' insect phylogenetic systematics. *Zool Beitr N F* 36: 83–124.
84. Beutel RG, Gorb SN (2001) Ultrastructure of attachment specializations of hexapods, (Arthropoda): evolutionary patterns inferred from a revised ordinal phylogeny. *J Zool Syst Evol Res* 39: 177–207.
85. Piazzi F, Ricci A, Passamonti M (2011) The mitochondrial genome of *Bacillus stick* insects (Phasmatodea) and the phylogeny of orthopteroid insects. *Mol Phylogenet Evol* 58: 304–316.
86. Simon S, Strauss S, von Haeseler A, Hadrys H (2009) A Phylogenomic Approach to Resolve the Basal Pterygote Divergence. *Mol Biol Evol* 26: 2719–2730.
87. Wiegmann B, Trautwein M, Kim JW, Cassel B, Bertone M, et al. (2009) Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. *BMC biology* 7: 34.
88. McKenna DD, Farrell BD (2010) 9-Genes Reinforce the Phylogeny of Holometabola and Yield Alternate Views on the Phylogenetic Placement of Strepsiptera. *PLoS one* 5: e11887.
89. Talavera G, Vila R (2011) What is the phylogenetic signal limit from mitogenomes? The reconciliation between mitochondrial and nuclear data in the Insecta Class phylogeny. *BMC Evol Biol* 11: 315.
90. Whiting MF, Bradler S, Maxwell T (2003) Loss and recovery of wings in stick insects. *Nature* 421: 264–267.
91. Hamilton KGA (1981) Morphology and evolution of the Rhynchotan head (Insecta: Hemiptera, Homoptera). *Can Ent* 113: 953–974.
92. Davis RB, Baldauf SL, Mayhew PJ (2010). Many hexapod groups originated earlier and withstood extinction events better than previously realized: inferences from supertrees. *P Roy Soc B-Biol Sci* 277: 1597–1606.
93. Snodgrass RE (1935) *Principles of Insecta Morphology*. Ithaca: Cornell University Press.
94. Rahle W (1970) Untersuchungen an Kopf und Prothorax von *Embia ramburi* Rimsky-Korsakow 1906 (Embioptera, Embiidac). *Zool Jb Anat* 87: 248–330.
95. Thomas MA, Walsh KA, Wolf MR, McPherson BA, Marden JH (2000) Molecular phylogenetic analysis of evolutionary trends in stonefly wing structure and locomotor behavior. *Proc Natl Acad Sci USA* 97: 13178–13183.
96. Willmann R (2004) Phylogenetic relationships and evolution of insects. In: Cracraft J, Donoghue MJ. (Eds.), *Assembling the Tree of Life*. Oxford: Oxford University Press, 330–344.
97. Letsch HO, Meusemann K, Wipfler B, Schütte K, Beutel R, et al. (2012) Insect phylogenomics: results, problems and the impact of matrix composition. *P Roy Soc B-Biol Sci* 279: 3282–3290.
98. Kukalová-Peck J (1991) Fossil history and the evolution of hexapod structures. In: Naumann ID (chi fed.). *The Insects of Australia. A Textbook for Students and Research Workers*. Melbourne: Melbourne University Press, 141–179.
99. Haas F, Kukalová-Peck J (2001) Dermaptera hindwing structure and folding: new evidence for familial, ordinal and super ordinal relationships within Neoptera (Insecta). *Eur J Entomol* 98: 445–510.
100. Béthoux O, Nel A (2002) Venation pattern and revision of Orthoptera sensu nov. and sister groups. Phylogeny of Palaeozoic and Mesozoic Orthoptera sensu nov. *Zootaxa* 96: 1–88.
101. Sharov AG (1968) *Filogeniya Orthopteroidnykh Nasekomykh*. Tr Paleontol Inst Akad Nauk SSSR 118: 1–217.
102. Handlirsch A (1908) *Die Fossilien Insekten und die Phylogenie der Rezenten Formen: Ein Handbuch für Paläontologen und Zoologen*. Leipzig: Wilhelm Engelmann.
103. Richards OW, Davies RG (1977) *IMMS' General Textbook of Entomology*. Great Britain: St Edmundsbury Press.
104. Rokas A, Carroll SB (2006) Bushes in the tree of life. *PLoS Biol* 4: e325.
105. Whitfield JB, Lockhart PJ (2007) Deciphering ancient rapid radiations. *Trends Ecol Evol* 22: 258–265.
106. Whitfield JB, Kjer KM (2008) Ancient rapid radiations of insects: Challenges for phylogenetic analysis. *Annu Rev Entomol* 53: 449–472.
107. Burki F, Shalchian-Tabrizi K, Minge M, Skjæveland Å, Nikolaev SI, et al. (2007) Phylogenomics reshuffles the eukaryotic supergroups. *PLoS one* 2: e790.
108. Roeding F, Hagner-Holler S, Ruhberg H, Ebersberger I, Von Haeseler A, et al. (2007) EST sequencing of Onychophora and phylogenomic analysis of Metazoa. *Mol Phylogenet Evol* 45: 942–951.
109. Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, et al. (2008) Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* 452: 745–749.
110. Witek A, Herlyn H, Meyer A, Boell L, Bucher G, et al. (2008) EST based phylogenomics of Syndermata questions monophyly of Eurotatoria. *BMC Evol Biol* 8: 345.
111. Meusemann K, von Reumont BM, Simon S, Roeding F, Strauss S, et al. (2010) A phylogenomic approach to resolve the arthropod tree of life. *Mol Biol Evol* 27: 2451–2464.
112. Andrew DR (2011) A new view of insect-crustacean relationships II. Inferences from expressed sequence tags and comparisons with neural cladistics. *Arthropod Struct Dev* 40: 289–302.
113. Philippe H, Brinkmann H, Copley RR, Moroz LL, Nakano H, et al. (2011) Acoelomorph flatworms are deuterostomes related to Xenoturbella. *Nature* 470: 255–258.
114. von Reumont BM, Jenner RA, Wills MA, Dell'ampio E, Pass G, et al. (2011) Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of Hexapoda. *Mol Biol Evol* 29: 1031–1045.
115. Lyal CHC (1985) Phylogeny and classification of the Psocodea, with particular reference to the lice (Psocodea: Phthiraptera). *Syst Biol* 10: 145–165.
116. Yoshizawa K, Saigusa T (2001) Phylogenetic analysis of paraneopteran orders (Insecta: Neoptera) based on forewing base structure, with comments on monophyly of Auchenorrhyncha (Hemiptera). *Syst Biol* 26: 1–13.
117. Bonneton F, Brunet F, Kathirithamby J, Laudet V (2006) The rapid divergence of the ecdysone receptor is a synapomorphy for Mecoptera that clarifies the Strepsiptera problem. *Insect Mol Biol* 15: 351–362.
118. Beutel RG, Friedrich F, Hörschemeyer T, Pohl H, Hünfeldt F, et al. (2011): Morphological and molecular evidence converge upon a robust phylogeny of the megadiverse Holometabola. *Cladistics* 27: 341–355.
119. Longhorn SJ, Pohl HW, Vogler AP (2010) Ribosomal protein genes of holometabolous insects reject the Halteria, instead revealing a close affinity of Strepsiptera with Coleoptera. *Mol Phylogenet Evol* 55: 846–859.
120. Niehuis O, Hartig G, Grath S, Pohl H, Lehmann J, et al. (2012) Genomic and morphological evidence converge to resolve the enigma of Strepsiptera. *Curr Biol* 22: 1309–1313.
121. Gillespie JJ, Yoder MJ, Wharton RA (2005) Predicted Secondary Structure for 28S and 18S rRNA from Ichneumonidea (Insecta: Hymenoptera: Apocrita): Impact on Sequence Alignment and Phylogeny Estimation. *J Mol Evol* 61: 114–137.
122. Xie Q, Bu WJ, Zheng LY (2005) The Bayesian phylogenetic analysis of the 18S rRNA sequences from the main lineages of Trichophora (Insecta: Heteroptera: Pentatomomorpha). *Mol Phylogenet Evol* 34: 448–451.
123. Letsch HO, Kuck P, Stocsits RR, Misof B (2010) The Impact of rRNA Secondary Structure Consideration in Alignment and Tree Reconstruction: Simulated Data and a Case Study on the Phylogeny of Hexapods. *Mol Biol Evol* 27: 2507–2521.
124. Letsch HO, Kjer KM (2011) Potential pitfalls of modelling ribosomal RNA data in phylogenetic tree reconstruction: Evidence from case studies in the Metazoa. *BMC Evol Biol* 11: 146.
125. Nagaswamy U, Voss N, Zhang Z, Fox GE (2000) Database of noncanonical base pairs found in known RNA structures. *Nucl Acids Res* 28: 356–357.