Two mitochondrial genomes from the families Bethylidae and Mutillidae: Independent rearrangement of protein-coding genes and higher-level phylogeny of the Hymenoptera

Shu-Jun Wei, Qian Li, Kees van Achterberg, Xue-Xin Chen

Molecular Phylogenetics and Evolution 77 (2014) 1–10

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Two mitochondrial genomes from the families Bethylidae and Mutillidae: Independent rearrangement of protein-coding genes and higher-level phylogeny of the Hymenoptera

Shu-Jun Wei [a,b,*], Qian Li [a], Kees van Achterberg [c], Xue-Xin Chen [a,*]

[a] State Key Laboratory of Rice Biology and Ministry of Agriculture Key Laboratory of Agricultural Entomology, Institute of Insect Sciences, Zhejiang University, Hangzhou 310058, China
[b] Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
[c] Terrestrial Zoology, Naturalis Biodiversity Center, 2300 RA Leiden, Netherlands

ARTICLE INFO

Article history:
Received 5 September 2013
Revised 27 February 2014
Accepted 24 March 2014
Available online 1 April 2014

Keywords:
Mitochondrial genome
Protein-coding gene rearrangement
Hymenoptera
Phylogeny

ABSTRACT

In animal mitochondrial genomes, gene arrangements are usually conserved across major lineages but might be rearranged within derived groups, and might provide valuable phylogenetic characters. Here, we sequenced the mitochondrial genomes of *Cephalonomia gallicola* (Chrysidoidea: Bethylidae) and *Wallacidia oculata* (Vespoidea: Mutillidae). In *Cephalonomia* at least 11 tRNA and 2 protein-coding genes were rearranged, which is the first report of protein-coding gene rearrangements in the Aculeata. In the Hymenoptera, three types of protein-coding gene rearrangement events occur, i.e. reversal, transposition and reverse transposition. *Venturia* (Ichneumonidae) had the greatest number of common intervals with the ancestral gene arrangement pattern, whereas *Philotrypesis* (Agaonidae) had the fewest. The most similar rearrangement patterns are shared between *Nasonia* (Pteromalidae) and *Philotrypesis*, whereas the most differentiated rearrangements occur between *Cotesia* (Braconidae) and *Philotrypesis*. It is clear that protein-coding gene rearrangements in the Hymenoptera are evolutionarily independent across the major lineages but are conserved within groups such as the Chalcidoidea. Phylogenetic analyses supported the sister-group relationship of Orrussoidea and Apocrita, Ichneumonoidea and Aculeata, Vespidae and Apoidea, and the paraphyly of Vespoidea. The Evaniomorpha and phylogenetic relationships within Aculeata remain controversial, with discrepancy between analyses using protein-coding and RNA genes.

1. Introduction

Animal mitochondrial genomes are widely used in population genetics, species identification, phylogenetic and evolutionary studies because of their unique features of conserved gene content, lack of extensive recombination, maternal inheritance, relatively high evolutionary rate and abundant marker types (Boore et al., 1998; Curole and Kocher, 1999). Most animal mitochondrial genomes are approximately 16 Kb and contain 37 genes and an A + T-rich region (Zhang and Hewitt, 1997; Boore, 1999). Gene arrangements are usually conserved across major lineages but might be rearranged within groups, and may provide valuable phylogenetic characters (Boore et al., 1998; Dowton et al., 2002). Genome rearrangements can be characterized into minor rearrangements (tRNAs only) and major rearrangements (protein-coding and rRNA genes) according to the types of involved genes (Cameron et al., 2007a).

Previous studies have found gene rearrangements in many orders of Hexapoda; however, accelerated gene rearrangement events are restricted to the Hymenoptera (Dowton and Austin, 1999; Dowton et al., 2003), the hemipteroids (Shao et al., 2001a,b; Shao and Barker, 2003; Covacin et al., 2006; Cameron et al., 2007a, 2011) and the Protura (Chen et al., 2011). In the Phthiraptera, there is no variation in gene arrangement among species within a family, but there is much variation among the three suborders (Covacin et al., 2006). Protura have reverted to the *cox1–cox2* arrangement found in the ancestral bilaterian from the arrangement *cox1–trnL2–cox2* found in hexapod and crustacean groups (Boore et al., 1998; Chen et al., 2011). In the Hymenoptera, one of the most species-rich orders of insects, the

* Corresponding authors. Address: Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, 9 Shuguanghuayuan Middle Road, Haidian District, Beijing 100097, China (S.J. Wei). Address: Institute of Insect Sciences, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, China (X.X. Chen).
E-mail addresses: shujun268@163.com (S.J. Wei), xxchen@zju.edu.cn (X.X. Chen).

http://dx.doi.org/10.1016/j.ympev.2014.03.023
1055-7903/

gene rearrangement pattern is more complicated. Gene arrangement is conserved in the suborder "Symphyta" but accelerated in the Apocrita. Two hot spots of tRNA gene rearrangement within mitochondrial genomes have been studied across an extensive taxonomic sample (Dowton and Austin, 1999; Dowton et al., 2003), and the tRNA gene position was shown to be selectively neutral (Dowton et al., 2009b). However, the rearrangement of protein-coding genes is rare compared to the number of tRNA rearrangements in the Hymenoptera. Among the currently sequenced hymenopteran mitochondrial genomes, the rearrangement of protein-coding genes are found in three species of *Nasonia* (Chalcidoidea: Pteromalidae) (Oliveira et al., 2008), two species from the Agaonidae (Chalcidoidea) (Xiao et al., 2011), one species from the Aulacidae (Evanioidea) (Wei et al., 2013b), one species from the Ichneumonidae (Dowton et al., 2009b) and one species from the Microgastrinae (Ichneumonoidea: Braconidae) (Wei et al., 2010a).

Here, we sequenced two representative mitochondrial genomes from two families not previously studied, the Mutillidae (Aculeata: Vespoidea) and the Bethylidae (Aculeata: Chrysidoidea), and report a novel protein-coding gene rearrangement. The rearrangement of the protein-coding genes is compared and analyzed across the Hymenoptera. Finally, the phylogenetic relationships among major lineages of the Hymenoptera are reconstructed using mitochondrial genome sequences.

2. Materials and methods

2.1. DNA extraction, sequencing and genome annotation

The specimens of *Wallacidia oculata* (Fabricius, 1804) collected in Hainan (China) in May 2004 were identified with Lelej (2004) and Lelej and Brothers (2008), while the specimens of *Cephalonomia gallicola* (Ashmead, 1887) collected in Hangzhou (Zhejiang, China) in July 2008 was identified with Xu and He (2006). Both specimens were stored at −80 °C in 100% ethanol prior to DNA extraction in Zhejiang University. Total genomic DNA was extracted from individual specimens using the DNeasy tissue kit (Qiagen, Hilden, Germany) following the manufacturer's protocol.

A range of universal insect mitochondrial primers and hymenopteran mitochondrial primers modified from universal insect mitochondrial primers were used (Tables S1 and S2) (Simon et al., 1994, 2006). When necessary, species-specific primers were designed based on sequenced fragments and combined in various ways to bridge gaps. PCR and sequencing of both strands was conducted following the methods of Wei et al. (2009). Genome annotation, including the determination of the protein-coding genes, tRNA genes and rRNA genes, and the prediction of the secondary structures of tRNA genes, following the methods in Wei et al. (2009).

2.2. Comparison of gene rearrangement

Gene rearrangements were analyzed using the CREx software package (Bernt et al., 2007). The arrangement of the protein-coding, tRNA and rRNA genes in the two newly sequenced mitochondrial genomes was compared with the ancestral arrangement of the insect mitochondrial genome. The arrangement of protein-coding gene among all currently sequenced hymenopteran mitochondrial genomes was also analyzed. A distance matrix of the six patterns observed in Hymenoptera was calculated based on common intervals (A subset of genes that appear consecutively in two or more input gene orders). The rearrangement steps necessary to transform the ancestral arrangement into the derived patterns were analyzed.

2.3. Phylogenetic inference

2.3.1. Test of the substitution saturation

In the absence of saturation in the nucleotide substitution process, the uncorrected p-distances coincide with GTR + G + I distances; conversely, GTR + G + I distances are larger than p-distances when the level of saturation increases (Negrisolo et al., 2004). The level of saturation in different codon positions and genes was analyzed using scatter plot graphics, comparing the uncorrected p-distances with the distances calculated by applying the best-fit evolutionary model GTR + G selected by the jModeltest2 (Posada, 2009). All distances were calculated with PAUP*4.0 b10 (Swofford, 2002). The slope value, correlation coefficient and the average GTR distance value were considered in the evaluation of substitution saturation.

2.3.2. Taxa selection

There were 39 hymenopteran mitochondrial genomes available in GenBank in October 2013 (Table 1). In several families, genera and species, more than one mitochondrial genome has been sequenced. To avoid redundant sampling and the potential effect of a complex lower-level phylogeny on higher-level relationships, at most two representative species from each genus were included in the analyses. Since the genes *nad2*, *nad3* and *nad5* have not been sequenced from *Primeuchroeus* spp. (Chrysididae), this incomplete genome was not included in our analysis. The mitochondrial genome data from *Evania appendigaster* (Evaniidae) was initially included in our analyses, however, this taxa was always grouped within Aculeata with a long branch both in previous reports (Wei et al., 2010a; Mao et al., 2012) and in our preliminary analyses (results are not shown). Thus, we used one representative of the family Aulacidae from the Evanioidea instead of both Aulacidae and Evaniidae. As a result 25 species representing 16 families of the Hymenoptera were used for an analysis of phylogenetic relationships among major groups of Hymenoptera. Three superfamilies of "Symphyta" were included in the analyses, among which, Tenthredinoidea is well accepted to be basal to Cephoidea and Orussoidea (Sharkey, 2007; Dowton et al., 2009a; Sharkey et al., 2012; Klopfstein et al., 2013). Thus, the *Monocellicampa pruni* and *Perga condei* from Tenthredinoidea was set as outgroups in phylogenetic inferences.

2.3.3. Sequence alignment and gene selection

The amino acid sequences of the protein-coding genes and RNA genes were aligned independently using MAFFT v7.122 (Katoh and Standley, 2013). The alignment of the nucleotide sequences of the protein-coding genes was inferred from the amino acid sequence alignment using RevTrans v1.4 (Wernersson and Pedersen, 2003). In the Hymenoptera, many of the published mitochondrial genomes are incomplete, leading to the missing genes in some taxa used in phylogenetic analyses. First, we tested the exclusion of the protein-coding gene *nad2* and *nad6* in the analyses, which are both missing data from some taxa and display substitution saturation in codon positions. Secondly, the exclusion of *atp8* and *nad4l* were tested due to their demonstrated substitution saturation.

2.3.4. Data matrix, data partitioning and substitution model selection

The data matrix used in the phylogenetic analyses were generated in five combinations: first and second codon positions only (P12), all codon positions (P123), and amino acid sequences (AA) of the protein-coding genes, 22 tRNA and 2 rRNA genes (RNA), and the P123 and RNA genes (P123RNA). In order to standardize the partitioning strategy as recommended for phylogenetic analyses using mitochondrial genome (Cameron, 2014), the PartitionFinder v1.1.1 (Lanfear et al., 2012) was used to simultaneously

Table 1
The mitochondrial genomes currently sequenced in the Hymenoptera.

Species	Superfamily	Family	Accession number	References
Perga condei	Tenthredinoidea	Pergidae	AY787816	Castro and Dowton (2005)
Monocellicampa pruni	Tenthredinoidea	Tenthredinidae	JX566509	Wei et al. (2013a)
Cephus cinctus	Cephoidea	Cephidae	FJ478173	Dowton et al. (2009b)
Orussus occidentalis	Orussoidea	Orussidae	FJ478174	Dowton et al. (2009b)
Schlettererius cinctipes	Stephanoidea	Stephanidae	FJ478175	Dowton et al. (2009b)
Evania appendigaster	Evanioidea	Evanidae	FJ593187	Wei et al. (2010b)
Trissolcus basalis	Platygastroidea	Scelionidae	JN903532	Mao et al. (2012)
Pristaulacus compressus	Evanioidea	Aulacidae	KF500406	Wei et al. (2013b)
Vanhornia eucnemidarum	Proctotrupoidea	Vanhorniidae	NC008323	Castro et al. (2006)
Nasonia giraulti	Chalcidoidea	Pteromalidae	EU746609, EU746611, EU746614	Oliveira et al. (2008)
Nasonia longicornis	Chalcidoidea	Pteromalidae	EU746612, EU746616	Oliveira et al. (2008)
Nasonia vitripennis	Chalcidoidea	Pteromalidae	EU746615, EU746613, EU746610	Oliveira et al. (2008)
Philotrypesis pilosa	Chalcidoidea	Agaonidae	JF808722	Xiao et al. (2011)
Philotrypesis sp.	Chalcidoidea	Agaonidae	JF808723	Xiao et al. (2011)
Diadegma semiclausum	Ichneumonoidea	Ichneumonidae	EU871947	Wei et al. (2009)
Enicospilus sp.	Ichneumonoidea	Ichneumonidae	FJ478177	Dowton et al. (2009b)
Venturia canescens	Ichneumonoidea	Ichneumonidae	FJ478176	Dowton et al. (2009b)
Aphidius gifuensis	Ichneumonoidea	Braconidae	GU097658	Wei et al. (2010a)
Diachasmimorpha longicaudata	Ichneumonoidea	Braconidae	GU097655	Wei et al. (2010a)
Macrocentrus camphoraphilus	Ichneumonoidea	Braconidae	GU097656	Wei et al. (2010a)
Meteorus pulchricornis	Ichneumonoidea	Braconidae	GU097657	Wei et al. (2010a)
Phanerotoma flava	Ichneumonoidea	Braconidae	GU097654	Wei et al. (2010a)
Cotesia vestalis	Ichneumonoidea	Braconidae	FJ154897	Wei et al. (2010a)
Spathius agrili	Ichneumonoidea	Braconidae	FJ387020	Wei et al. (2010a)
Cephalonomia gallicola	Chrysidoidea	Bethylidae	FJ823227	This study
Wallacidia oculata	Vespoidea	Mutillidae	FJ611801	This study
Solenopsis geminata	Vespoidea	Formicidae	HQ215537	Shoemaker et al. (2006)
Solenopsis invicta	Vespoidea	Formicidae	HQ215538	Shoemaker et al. (2006)
Solenopsis richteri	Vespoidea	Formicidae	HQ215539	Shoemaker et al. (2006)
Solenopsis invicta	Vespoidea	Formicidae	HQ215540	Shoemaker et al. (2006)
Pristomyrmex punctatus	Vespoidea	Formicidae	AB556946	Hasegawa et al. (2011)
Pristomyrmex punctatus	Vespoidea	Formicidae	AB556947	Hasegawa et al. (2011)
Polistes humilis	Vespoidea	Vespidae	EU024653	Cameron et al. (2008)
Abispa ephippium	Vespoidea	Vespidae	NC011520	Cameron et al. (2008)
Philanthus triangulum	Apoidea	Crabronidae	JN871914	Kaltenpoth et al. (2012)
Apis mellifera ligustica	Apoidea	Apidae	NC001566	Crozier and Crozier (1993)
Bombus hypocrita sapporoensis	Apoidea	Apidae	NC011923	Hong et al. (2008)
Bombus ignitus	Apoidea	Apidae	NC010967	Cha et al. (2007)
Melipona bicolor	Apoidea	Apidae	NC004529	Silvestre and Arias (2006)

choose partitioning schemes and substitution models for each matrix. The maximum possible partition schemes inputted into the PartitionFinder software was defined by gene and by codon position for nucleotide sequences of protein-coding genes (39 partitions) and by gene for both amino acid sequences of protein-coding genes (13 partitions) and RNA genes (24 partitions). The "greedy" algorithm was used with branch lengths estimated as "unlinked" to search for the best-fit scheme.

2.3.5. Phylogenetic inference

Phylogenetic analyses were performed using Bayesian inference (BI) with the parallel version of MrBayes v 3.2.2 (Ronquist et al., 2012) and maximum likelihood (ML) with RAxML v7.9.6 (Stamatakis, 2006), both of which allows different models to be applied to different partitions. The GTR + G model was selected for each nucleotide partition while the MtArt + G + F model was selected for each amino acids partition. All Bayesian analyses were conducted with four independent Markov chains run for 10 million Metropolis-coupled (MCMC) generations, with tree sampling occurring every 1000 generations and a burn-in of 25% trees. After 10 million generations, all runs reached in stationarity examined by the program Tracer v1.4 (Rambaut and Drummond, 2007). For ML analyses, we used the GTRGAMMA model for each nucleotide partition and MtArtF model for each amino acid partition. For all analyses, 200 runs were conducted to find the best likelihood tree followed by 1000 bootstrap replicates. Phylogenetic trees were viewed and edited in Figtree v 1.4.0 (Rambaut, 2012).

3. Results and discussion

3.1. Sequencing results and genome size

The complete mitochondrial genome of *Wallacidia* (Genbank Accession number: FJ611801) and the nearly complete mitochondrial genome of *Cephalonomia* (Genbank Accession number: FJ823227) were determined to have lengths of 18,442 and 16,720 + bp, respectively (Fig. 1). The length of the mitochondrial genome of *Wallacidia* is longer than the typical insect mitochondrial genome, which is approximately 16,000 bp (Wolstenholme, 1992). However, within the Hymenoptera a longer mitochondrial genome was found in *Diadegma semiclausum* (18,728 bp), caused by a large insertion between the *cox1* and *cox2* genes and a large A + T-rich region (Wei et al., 2009). In the mitochondrial genome of *Cephalonomia*, a region between *trnP* and *cob* could not be amplified, which might be caused by high A + T content and the presence of long repeat sequences. According to the annotation of the sequenced portion of this genome, we inferred that the protein-coding gene *nad6* and the tRNA genes *trnC* and *trnS1* are located in the un-amplified portion of the genome. High A + T content is common in mitochondrial genomes of the suborder Apocrita (Crozier and Crozier, 1993; Cameron et al., 2008).

3.2. A + T-rich regions of the newly sequenced mitochondrial genomes

Both species have two long A + T-rich regions (Fig. 1). In the *Wallacidia* mitochondrial genome, the two A + T-rich regions are

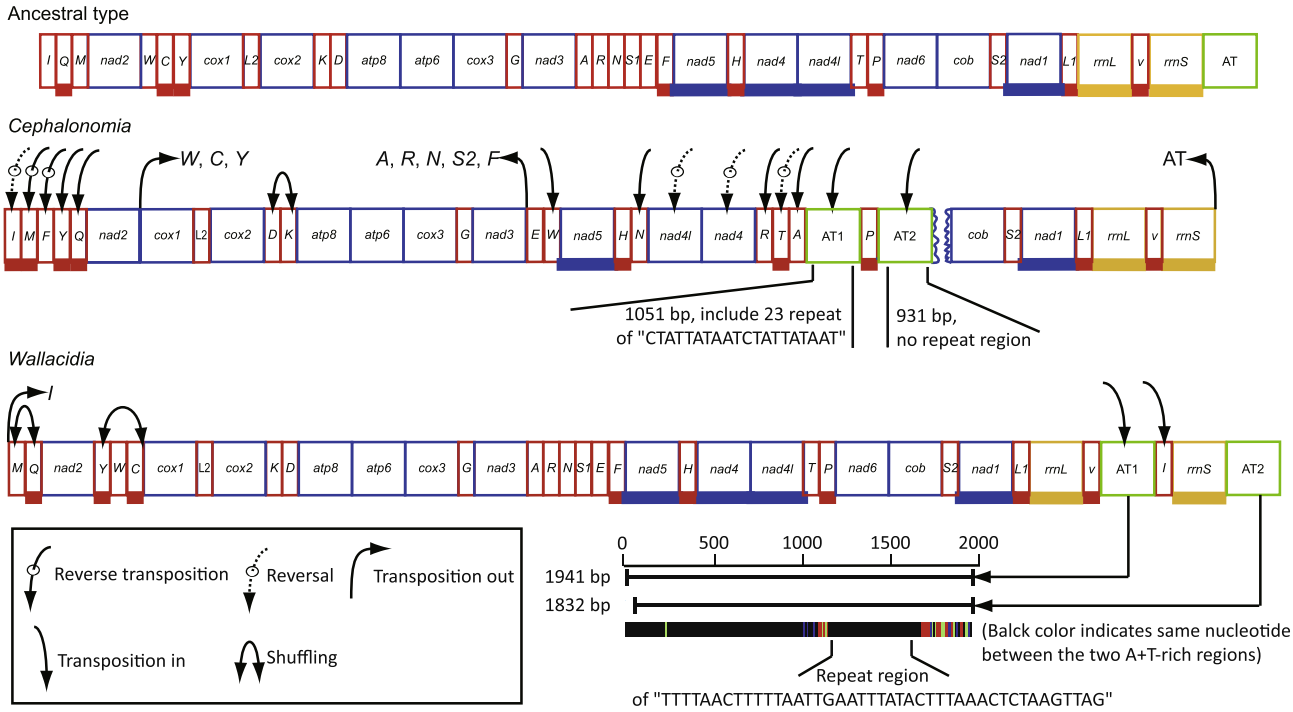

Fig. 1. Mitochondrial genome organization of *Cephalonomia gallicola* (Chrysidoidea: Bethylidae) and *Wallacidia oculata* (Vespoidea: Mutillidae). Boxes with underscores indicate that the gene is encoded on the minority strand. Rearranged protein-coding genes are shown in yellow-green. Gene abbreviations are as follows: *cox1*, *cox2*, and *cox3* refer to the cytochrome oxidase subunits, *cob* refers to cytochrome b, *nad1–nad6* refers to the NADH dehydrogenase components, and *rrnL* and *rrnS* refer to ribosomal RNAs. Transfer RNA genes are denoted by a one-letter symbol according to the IPUC-IUB single-letter amino acid codes. L1, L2, S1 and S2 denote tRNALeu (CUN), tRNALeu (UUR), tRNASer (AGN) and tRNASer (UCN), respectively. AT1 and AT2 indicate the A + T-rich regions. Region between wiggly lines indicates the unsequenced portion. Genes with squiggle line means the breakpoint caused by failure of sequencing. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

1941 bp and 1832 bp long, with an A + T content of 83.62% and 83.95%, respectively. These two regions, separated by the *trnI* and *rrnS* genes, are nearly identical except for the boundary regions, indicating that a concerted evolution might have occurred in this genome. Concerted evolution is common in insect A + T-rich regions most obviously in species with repeat units in their A + T-rich regions like termites (Cameron and Whiting, 2007), but also in species with non-tandemly repeated A + T-rich regions like thrips (Shao and Barker, 2003). The 3′ region is composed of 12 (in A + T-rich region 1) and 13 (in A + T-rich region 2) 43-bp repeat units. In the *Cephalonomia* mitochondrial genome, the two A + T-rich regions, separated by the *trnP* gene, are 1051 bp and 931 bp long, with an A + T content of 83.63% and 91.08%, respectively. In the 5′ region of A + T-rich region 1, there are 23 20-bp repeat units, whereas in the other A + T-rich region of this genome, there are no obvious repeats. The five conserved A + T-rich region elements proposed by Zhang and Hewitt (1997) were not found in either mitochondrial genome.

3.3. Gene arrangement in the mitochondrial genomes of Hymenoptera

In the *Wallacidia* mitochondrial genome, three tRNA genes, *trnI*, *trnQ*/*trnM* and *trnY*, are rearranged (Fig. 1), with 652 common intervals shared between *Wallacidia* and the ancestral insect mitochondrial gene arrangement as calculated by the CREx software package (Bernt et al., 2007). In the *Cephalonomia* mitochondrial genome, gene rearrangement is more extensive (Fig. 1), with at least 11 of 22 tRNA genes and two of 13 protein-coding genes rearranged and only 84 common intervals shared with the ancestral insect mitochondrial genome arrangement (Fig. 2). In the Hymenoptera, most of the rearranged genes are tRNAs (Dowton et al., 2009b). The rearrangement of protein-coding genes has been found in seven species from four families of Hymenoptera (Oliveira et al., 2008; Dowton et al., 2009b; Wei et al., 2010a; Xiao et al., 2011). The *Cephalonomia* mitochondrial genome is the first from the Aculeata that has a protein-coding gene rearrangement.

According to the definitions resulting from the CREx software and Dowton et al. (2003), we classified the animal mitochondrial gene rearrangement events into transposition (translocation), shuffling (transposition of adjacent genes), reversal (inversion) and reverse transposition (remote inversion). In the *Wallacidia* mitochondrial genome, the *trnI* gene was transposed upstream of the *rrnS* gene from the original region of *rrnS*–A + T-rich region–*trnI*–*trnQ*–*trnM*, forming a gene arrangement pattern of A + T-rich region 1–*trnI*–*rrnS*–A + T-rich region 2–*trnM*–*trnQ*. Tandem duplication followed by random loss (TDRL) of supernumerary genes is considered to be the dominant rearrangement mechanism in vertebrate mitochondrial genomes (San Mauro et al., 2006). TDRL has also been considered to be a likely mechanism for gene shuffling in insect mitochondrial genomes (Wei et al., 2009). Evidence for the TDRL mechanism is indicated by the pattern of gene order, the presence of pseudogenes or duplicated genes, and the position of intergenic spacers. In the *Wallacidia* mitochondrial genome two A + T-rich regions are present with high similarity, as discussed below, indicating that the transposition of the *trnI* gene might be caused by a tandem duplication of *rrnS*–A + T-rich region–*trnI* region, followed by the loss of the upstream *rrnS* gene and the downstream *trnI* gene. The other two rearrangement events, shuffling of the *trnQ* and *trnM* genes and shuffling of the *trnY*, *trnW* and *trnC* genes, could be explained by TDRL because there is a 29 bp intergenic region after the *trnQ* gene and a 246 bp intergenic region after the *trnC* gene.

In the *Cephalonomia* mitochondrial genome, the tRNA gene rearrangement events could be classified into transposition (*trnQ*,

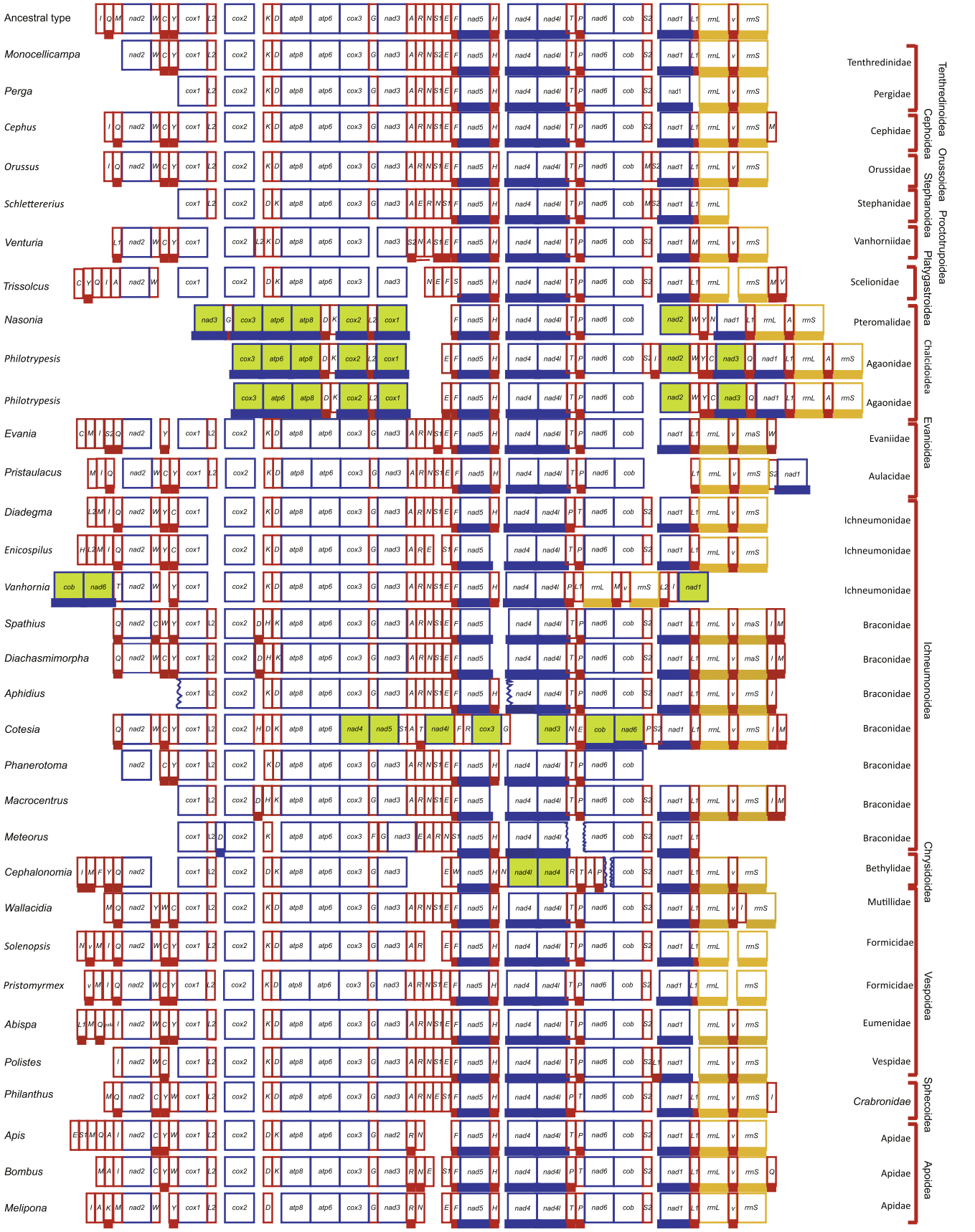

Fig. 2. Gene arrangements in the currently sequenced hymenopteran mitochondrial genomes. Symbols are as in Fig. 1.

trnW, *trnY*, *trnA*, *trnR* and *trnN*), shuffling (*trnD*/*trnK*), reversal (*trnI* and *trnT*) and reverse transposition (*trnM* and *trnF*), whereas the protein-coding gene rearrangement of the *nad4* and *nad4l* genes is a reversal. Downstream of the six translocated tRNA genes, intergenic nucleotides of at least 6 bp are present, indicating that a recombination or TDRL event could be involved in these genes. Shuffling of the *trnD* and *trnK* genes is common and has occurred independently in the multiple hymenopteran families (Dowton and Austin, 1999). The ancestral arrangement of *trnI–trnQ–trnM* is thoroughly disrupted by a local inversion of the *trnI* gene, the reverse transposition of the *trnM* gene and a subsequent insertion of the *trnF* and *trnY* genes. This region usually tends to have more tRNA genes in some members of the Hymenoptera, such as *Apis mellifera* and *Abispa ephippium* (Crozier and Crozier, 1993; Cameron et al., 2008).

In total, there are six arrangement patterns of the protein-coding genes in the currently sequenced hymenopteran mitochondrial genomes (Table 2). The CREx analysis indicated that *Venturia canescens* (Ichneumonidae) had the greatest number of common intervals with the ancestral gene arrangement pattern, whereas *Philotrypesis* (Agaonidae) had the fewest.

Among the 13 protein-coding genes in the mitochondrial genomes currently sequenced in Hymenoptera, 12 are rearranged in at least one species. All rearrangement types, except for shuffling, are found in protein-coding genes. Reversal is the most common rearrangement type, being found in all species with gene rearrangements except for *Pristaulacus compressus* (Evanioidea: Aulacidae). Transposition is less common and is found in *Pristaulacus* and the four heavily rearranged mitochondrial genomes of *Nasonia*, *Philotrypesis*, *Venturia* and *Cotesia*. Reverse transposition is found only once, in the mitochondrial genome of *Cotesia*.

Our analyses indicate that protein-coding gene rearrangements in Hymenoptera are randomly distributed across lineages but may be conserved within groups, such as the Chalcidoidea. Extensive sequencing of the mitochondrial genomes from the relatives of species with protein-coding genes rearrangements is necessary for meaningful investigation of their phylogenetic information.

3.4. Mitochondrial phylogenomics of major lineages in the Hymenoptera

3.4.1. Factors affecting phylogenetic reconstruction

We performed 22 phylogenetic reconstruction analyses with combinations of different data matrix and inference methods to test the influence of genes, codon position and inference methods on tree topology and nodal support (Table 3).

Table 2
Distance matrix based on common intervals among different protein-coding gene arrangements in the Hymenoptera, calculated with the CREx software.

	Ancestral	*Nasonia*	*Philotrypesis*	*Venturia*	*Cotesia*	*Cephalonomia*
Ancestral	150	68	46	110	60	106
Nasonia spp.	68	150	128	54	34	52
Philotrypesis spp.	46	128	150	36	28	34
Venturia canescens	110	54	36	150	48	88
Cotesia vestalis	60	34	28	48	150	60
Cephalonomia gallicola	106	52	34	88	60	126

The number of common intervals is a similarity measure. High numbers of common intervals mean similar gene orders. The *nad6* gene is missing in *Cephalonomia gallicola*. Although the *nad1* gene was rearranged in the *Pristaulacus compressus* (Evanioidea: Aulacidae) mitochondrial genome, the relative arrangement among the the 13 protein-coding genes was not changed. Thus, this species was not included in the CREx analysis.

Table 3
Summary of the conflict phylogenetic relationships inferred from different data matrix and methods.

Data matrix	Inference method	"Symphyta"	Parasitica	Proctotrupomorpha	Stephanoidea position	Ichneomonoidea position	Within Aculeata
13 AA	BI	(O + C 81)+Apoc*	Paraphyly	64	(S + E)+Proc 62	I + Acul 71	(B + M)+(F+(V + A 96))
13 AA	ML	(O + C 55)+Apoc*	Paraphyly	–	S + other Apoc 80	I + Plat 13	(B + M)+(F+(V + A 69))
13 P12	BI	C+(O + Apoc*)*	Paraphyly	94	S + other Apoc*	(E + I 82)+(Plat + Acul)	(B + M)+(F+(V + A*))
13 P12	ML	C+(O + Apoc*)*	Paraphyly	36	S + other Apoc 81	(E + I 26)+Acul	(B + M)+(F+(V + A 87))
13 P123	BI	C+(O + Apoc*)*	Paraphyly	100	S + other Apoc*	(E + I*)+(Plat + Acul)	(B + M)+(F+(V + A*))
13 P123	ML	C+(O + Apoc*)*	Paraphyly	47	S + other Apoc 86	(E + I 44)+Acul	F+((B + M)+(V + A 79))
11 AA	BI	O+(C + Apoc 57)*	Paraphyly	–	S + Platy 33	I + Acul*	(V + F)+((B + M)+A)
11 AA	ML	(O + C 55)+Apoc*	Paraphyly	–	S + Platy*	I + Acul 25	(B + M)+(F+(V + A 65))
11 P12	BI	C+(O + Apoc*)*	Paraphyly	83	(S + E 99)+Proc 68	I + Acul *	B+((M + F)+(V + A*))*
11 P12	ML	C+(O + Apoc*)*	Paraphyly	39	(S + E 27)+Proc 11	I + Acul 57	B+((M + F)+(V + A 85)) 74
11 P123	BI	C+(O + Apoc*)*	Paraphyly	84	(S + E 95)+Proc 72	I + Acul*	B+((M + F)+(V + A*))*
11 P123	ML	C+(O + Apoc*)*	Paraphyly	51	(S + E 18)+Proc 8	I + Acul 60	B+((M + F)+(V + A 73)) 84
9 AA	BI	O+(C + Apoc 82)*	Paraphyly	–	S + Platy 99	I + Acul*	(V + F)+((B + M)+A)
9 AA	ML	(O + C 48)+Apoc*	Paraphyly	48	S + Platy 29	I + Acul 25	(B + M)+(F+(V + A 43))
9 P12	BI	C+(O + Apoc*)*	Paraphyly	100	(S + E 91)+Proc*	I + Acul*	(B + M)+(F+(V + A*))
9 P12	ML	C+(O + Apoc*)*	Paraphyly	49	(S + E 27)+Proc 36	I + Acul 40	F+((B + M)+(V + A 78))
9 P123	BI	C+(O + Apoc*)*	Paraphyly	100	(S + E 84)+Proc 87	I + Acul 91	F+((B + M)+(V + A*))
9 P123	ML	C+(O + Apoc*)*	Paraphyly	57	(S + E 20)+Proc 24	I + Acul 38	F+((B + M)+(V + A 89))
24 RNA	BI	C+(O + Apoc*)*	Monophyly 62	–	S + Chal 61	E + I	M+(((V+(B + F))+A)
24 RNA	ML	C+(O + Apoc*)*	Monophyly 48	–	S + Chal 63	E + I	(B+(M + F))+(V + A 11)
11 P123 + RNA	BI	C+(O + Apoc*)*	Monophyly 89	100	S + other Apoc 89	E + I 81	B+((M + F)+(V + A*))*
11 P123 + RNA	ML	C+(O + Apoc*)*	Monophyly 53	65	S + Proc 28	E + I 55	B+((M + F)+(V + A 78))*

AA, amino acid sequences of protein-coding genes; P12, first and second codon positions of protein-coding genes, P123, all codon positions of protein-coding genes; RNA, nucleotide sequences of *lrna*, *srna* and 22 tRNA genes; 13 in the column of data matrix indicates full 13 protein-coding genes; 11 in the column of data matrix indicates the *nad2* and *nad6* were exclude from analyses; 9 in the column of data matrix indicates the *nad2*, *nad6* , *atp8* and *nad4l* were exclude from analyses; C, Cephoidea; O, Orussoidea; Apoc, Apocrita; Acul, Aculeata; S, Stephanoidea; E, Evanioidea; Plat, Platygastroidea; Proc, Proctotrupomorpha; I, Ichneumonoidea; B, Bethylidae; M, Mutillidae, F, Formicidae; V, Vespidae; A, Apoidea; the numbers after the group indicate bootstrap values for ML tree and 100 times of posterior probability for BI tree.
* Indicates the node is 1.0/100% supported.

Pairwise distance analyses indicated that multiple-substitutions were elevated in *atp8*, *nad2*, *nad4l* and *nad6* at first and second codon positions and for all protein-coding genes at third codon positions (Fig. 3, Table 4). The following removal of protein-coding genes and exclusion of codon position in phylogenetic analyses were conducted according to these results.

3.4.1.1. Removal of protein-coding genes. When all protein-coding genes were included, the Stephanidae was always placed at a sister-group position to all other Apocrita and the two long-branches of Bethylidae and Mutillidae grouped together. Additionally, tree topologies are sensitive to inference methods and exclusion of third codon position. Although adding characters with missing data can improve accuracy in phylogenetic analyses, there is a risk of long-branch attraction in some cases (Wiens and Moen, 2008; Thomson and Shaffer, 2010). In the subsequent analyses, the exclusion of *nad2* and *nad6*, which showed both missing data in some taxa and substitution saturation in first and second codon positions, improved the nodal support and congruence of tree topology between inference methods. Exclusion of the saturated *atp8* and *nad4l* did not alter the tree topology except for the relationships amongst the basal lineages of Aculeata (Table 4, Fig. 4). The negative effect of *nad2* and *nad6* in our analyses might be caused by joint effect of substitution saturation and missing data (Lemmon et al., 2009; Wiens and Morrill, 2011).

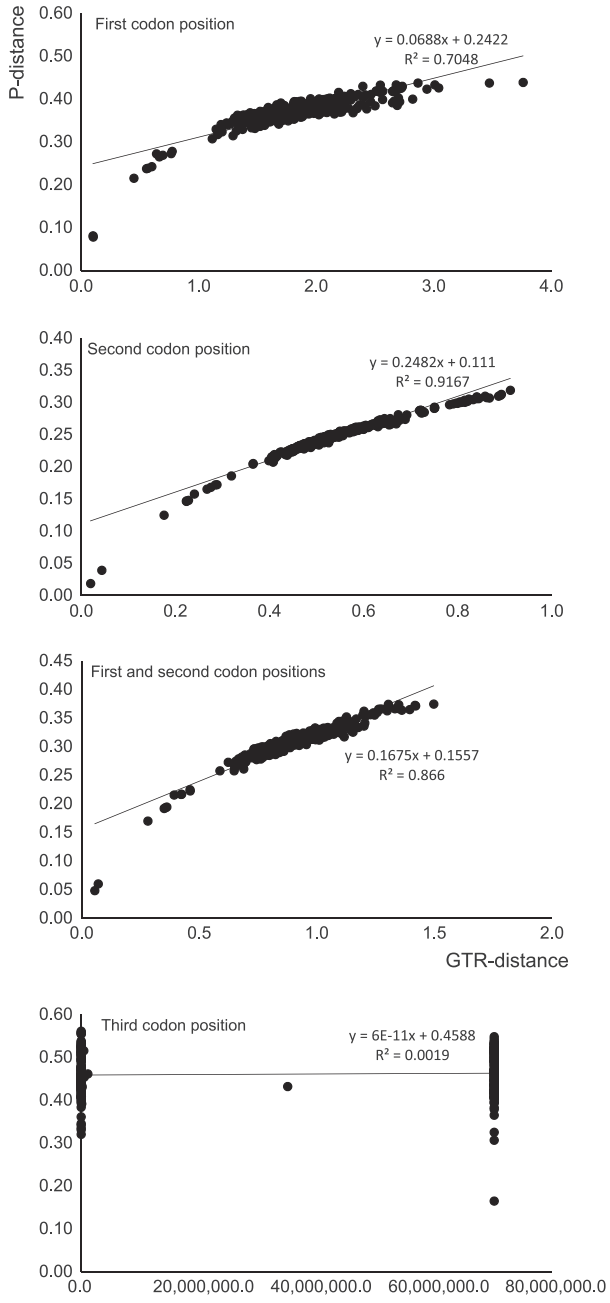

Fig. 3. Scatter plot graphics performed to test the saturation of the nucleotide substitution process in the mitochondrial protein-coding genes in Hymenoptera. Pairwise distances were calculated for the first, second, first and second, and third codon positions of the 13 mitochondrial protein-coding genes.

Table 4
Regression of the pairwise distances of the first and second codon positions of protein-coding genes and all positions of RNA genes.

Gene	Regression	*R*	Average GTR distance
atp6	$y = 0.0806x + 0.2307$	0.8797	1.0629
atp8	$y = 2E{-}05x + 0.4022$	0.1606	51.5384
cox1	$y = 0.4387x + 0.0498$	0.9834	0.2979
cox2	$y = 0.2223x + 0.1260$	0.9471	0.6454
cox3	$y = 0.0942x + 0.2129$	0.7645	0.8229
cytb	$y = 0.2737x + 0.1017$	0.9645	0.4941
nad1	$y = 0.1058x + 0.1999$	0.3335	1.0591
nad2	$y = 0.0032x + 0.4210$	0.8784	37.5498
nad3	$y = 0.03x + 0.3009$	0.9808	1.9455
nad4	$y = 0.0673x + 0.2419$	0.8713	1.4974
nad4 l	$y = 5E{-}07x + 0.3897$	0.4045	26069.8762
nad5	$y = 0.0872x + 0.2171$	0.8698	1.3574
nad6	$y = 8E{-}06x + 0.4346$	0.9137	6854.2320
rrnl	$y = 0.0979x + 0.1967$	0.8897	1.2587
rrns	$y = 0.058x + 0.2408$	0.8514	1.7998
tRNAs	$y = 0.1522x + 0.1511$	0.9279	0.7504

3.4.1.2. RNA genes. Date sets that include RNA genes support the monophyly of Parasitica with two major lineages: Evanioidea + Ichneumonoidea and Stephanidae + Proctotrupomorpha, or a sister-group relationship between Stephanidae and other Parasitica (Fig. 5). It is in conflict with results from protein-coding genes, where the paraphyly of Parasitica was supported. Additionally, the analyses of RNA sequences alone did not recover Proctotrupomorpha within Parasitica (Table 3).

3.4.1.3. Codon position. Many studies exclude the third codon positions due to substitution saturation when analyzing protein-coding genes (Wei et al., 2010a; Klopfstein et al., 2013). Our analyses showed that the exclusion of the third codon position did not alter the tree topology but reduced the support values in some nodes, suggesting that this position does not negatively effect phylogenetic reconstruction with complete mitochondrial genomes in Hymenoptera, as has been previously shown in Diptera (Cameron et al., 2007b) and Orthoptera (Fenn et al., 2008). The majority of previous studies have used relatively intuitive partitioning schemes such as by gene type, by gene and by codon position (Cameron, 2014). In our analyses, we chose the partitioning schemes and substitution models simultaneously implemented in the software PartitionFinder. The best-fit scheme was simultaneously divided by codon positions of protein-coding genes and gene groups (Table S3), rather than *a priori* partitioning schemes frequently used in mitochondrial phylogenomics. The partitioning strategy might optimize the information from the genes and codon positions (Leavitt et al., 2013).

3.4.1.4. Coding of protein-coding genes. After exclusion of the problematic genes of *nad2* and *nad6*, the nucleotide sequences of protein-coding genes resulted in congruent tree topology independent of inference method and the inclusion/exclusion of third

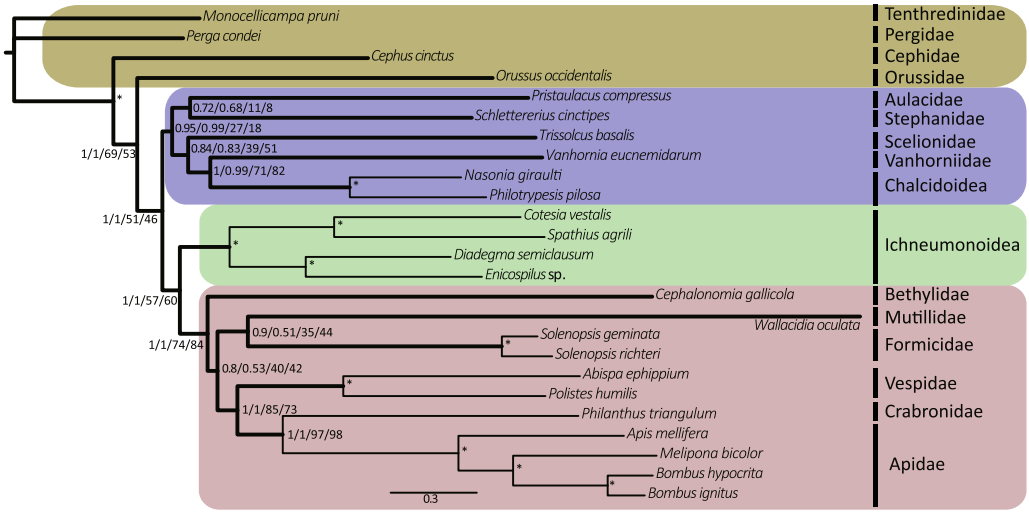

Fig. 4. Hymenoptera phylogeny based on nucleotide sequences of the mitochondrial protein-coding genes. Bayesian posterior probabilities and bootstrap values (1000 replicates) based on all codon positions and first and second codon positions of mitochondrial protein-coding genes (*nad2* and *nad6* are not included) are shown sequentially, separated by "/" near respective nodes. An asterisk indicates that the nodes were 1.0/100% supported by both inference methods.

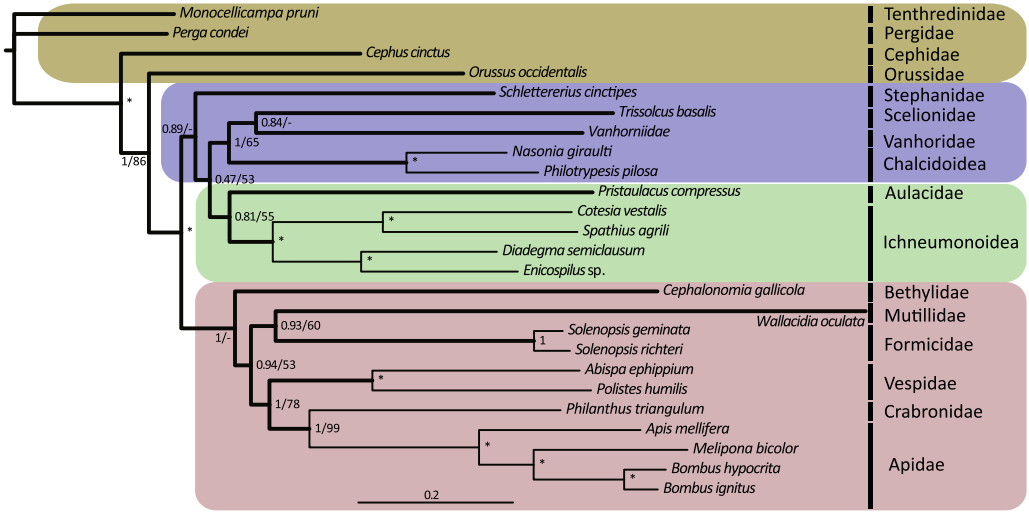

Fig. 5. Hymenoptera phylogeny based on nucleotide sequences of the the mitochondrial protein-coding and RNA genes. Bayesian posterior probabilities and bootstrap values (1000 replicates) based on all codon positions and of the protein-coding genes (*nad2* and *nad6* are not included), 2 rRNA and 22 tRNA genes are shown sequentially, separated by "/" near respective nodes. An asterisk indicates that the nodes were 1.0/100% supported by both inference methods.

codon position, although the nodal supports are different (Fig. 4, Table 4). The amino acid sequences of protein-coding genes has been claimed to be the best data set to be analyzed in higher-level phylogeny of in insects with mitochondrial genomes (Talavera and Vila, 2011). However, we found that the amino acid sequences could not avoid long branch attraction between Bethylidae and Mutillidae, or recover the well accepted sister-group relationship between Orussidae and Apocirta (Sharkey, 2007; Dowton et al., 2009a; Sharkey et al., 2012; Klopfstein et al., 2013). Nevertheless, the Evaniomorpha (with Stephanidae and Aulacidae included here) proposed by Rasnitsyn (1988), Proctotrupomorpha and its sister-group relationship to Evaniomorpha, the Ichneumonoidea + Aculeata were supported as recovered by nucleotide sequences of reduced set of 11 protein-coding genes.

3.4.1.5. Inference methods. The Bayesian and ML methods produced different topologies when we use the same data matrix except for that of nucleotide sequences of reduced set of 11 protein-coding genes. The ML searches using three amino acid matrices always grouped Orussidae and Cephidae together, rather than a well accepted sister-group relationship between Orussidae and Apocirta. After removal of the problematic *nad2* and *nad6* genes, both Bayesian and ML resulted in congruent tree topology independent of inference method and exclusion of third codon position (Fig. 4). However, the support values of the shared nodes are always higher in the Bayesian tree than those in ML tree generated from the same data matrix (Table 3). Our analyses support that the Bayesian inference in mitochondrial phylogenomics performs more reliably than the others, as revealed by previous studies (Dowton et al., 2009a; Gotzek et al., 2010; Talavera and Vila, 2011).

3.4.2. The phylogenetic relationships among major lineages in Hymenoptera

Our analyses based on different matrix recovered some well supported groups as well as conflicting relationships among major lineages within Hymenoptera. This is very similar to the previous

studies, both congruent and conflicting conclusions obtained from different data set and inference methods. The complicated and challenging phylogenetic relationships within Hymenoptera stimulated continuous researches for more than 100 years (Ashmead, 1896; Dowton and Austin, 1994; Ronquist et al., 1999; Dowton et al., 2009a; Klopfstein et al., 2013). After comparisons of the factors that might affect the phylogenetic inference, our analyses suggest that gene types (protein-coding vs RNA genes), selection of the used protein-coding genes and their coding method (nucleotide vs amino acid sequences) have strong influence on the resulting phylogenetic inference in our dataset. It should be noticed that the extraction of phylogenetic signal from hymenopteran mitochondrial genomic data is not simple (Dowton et al., 2009a; Gotzek et al., 2010). In our analyses, the nucleotide sequences of the reduced set of 11 protein-coding genes generated more reliable topology using both Bayesian and ML inference, although the nodal supports are lower in ML than those in Bayesian trees (Fig. 4).

Our analysis supported the basal position of "Symphyta" in Hymenoptera and its paraphyly, which is congruent with traditional views, with the Orussoidea sister to the suborder Apocrita (Rasnitsyn, 1988; Sharkey and Roy, 2002; Castro and Dowton, 2005; Dowton et al., 2009a).

Within the Apocrita, the superfamilies Ichneumonoidea and Chalcidoidea were fully supported in all analyses. The Proctotrupomorpha was supported. The Proctotrupomorpha has been well recovered many times by both morphological characters (Ronquist et al., 1999), single gene sequences (Dowton and Austin, 1994; Dowton et al., 1997), multiple gene analyses (Klopfstein et al., 2013) and concatenated genetic and morphological analyses (Dowton and Austin, 2001; Sharkey et al., 2012). The relationship (within Proctotrupomorpha) Platygastroidea + (Proctotrupoidea + Chalcidoidea) is congruent with two recent studies (Sharkey et al., 2012; Klopfstein et al., 2013). The Evaniomorpha (with Stephanidae + Aulacidae included here) was recovered (Fig. 4). When the Evaniomorpha was supported, it usually formed a sister-group relationship to Proctotrupomorpha, with the Ichneumonoidea sister to Aculeata (Dowton and Austin, 1994; Dowton et al., 1997; Rasnitsyn et al., 2004). Conflicting results came from analyses that incorporated RNA sequences, in which, the monophyly of Parasitica was supported (Fig. 5). The position of groups proposed within the Evaniomorpha is one of the most controversial issues in Hymenoptera (Sharkey, 2007; Klopfstein et al., 2013). Present phylogenetic analyses did not include representatives from Ceraphronoidea, Megalyroidea and Trigonalyoidea. Increasing taxonomic coverage in the sequencing of mitochondrial genomes will contribute to better understanding of the phylogenetic relationships in these groups (Cameron, 2014).

The monophyly of Aculeata was recovered in all analyses. Within this group, the superfamily of Apoidea (Crabronidae + Apoidae), and the sister-group relationship between Vespidae and Apoidea were well-supported. Vespoidea was recovered as a paraphyletic group, as in other studies (Pilgrim et al., 2008; Klopfstein et al., 2013). Most previous studies have placed Chrysidoidea (with the Bethylidae included) as a sister group to other Aculeata (Brothers, 1999), although the monophyly of Chrysidoidea is still controversial (Sharkey et al., 2012). Our analyses using nucleotide sequences of protein-coding genes recovered the relationship of Bethylidae + ((Mutillidae + Formicidae) + (Vespidae + Apoidea)) with strong nodal supports, while those using amino acid and RNA sequences resulted in an artifactual grouping of Bethylidae + Mutillidae, likely due to long branch attraction, inserted in the Aculeata that of variable position. The phylogeny of Aculeata remains unclear (Klopfstein et al., 2013). Further studies with denser sampling and more powerful molecular markers, as well as the suitable inference methods are necessary to reveal the phylogenetic relationships within Hymenoptera (Klopfstein et al., 2013).

Acknowledgments

We thank Jing-Xian Liu and Yue-Qian Lu for collecting the specimens and Prof. Jun-Hua He (Zhejiang University) and Prof. Zai-Fu Xu (South China Agricultural University) for identification of the species. Funding for this study was provided jointly by the State Key Program of National Natural Science Foundation of China (31230068), the National Natural Science Foundation of China (31101661, 30970384), the NSFC Innovative Research Groups (31021003), 973 Program (2013CB127600), the Beijing New Star Program on Science and Technology (2010B027) and the Beijing Excellent Talents Program (2010D002020000010).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2014.03.023.

References

Ashmead, W.H., 1896. The phylogeny of the Hymenoptera. Proc. Entomol. Soc. Wash., 323–336.
Bernt, M., Merkle, D., Ramsch, K., Fritzsch, G., Perseke, M., Bernhard, D., Schlegel, M., Stadler, P.F., Middendorf, M., 2007. CREx: Inferring genomic rearrangements based on common intervals. Bioinformatics 23, 2957–2958.
Boore, J.L., 1999. Animal mitochondrial genomes. Nucl. Acids Res. 27, 1767–1780.
Boore, J.L., Lavrov, D.V., Brown, W.M., 1998. Gene translocation links insects and crustaceans. Nature 392, 667–668.
Brothers, D.J., 1999. Phylogeny and evolution of wasps, ants and bees (Hymenoptera, Chrysidoidea, Vespoidea and Apoidea). Zool. Scr. 28, 233–250.
Cameron, S.L., 2014. Insect mitochondrial genomics: Implications for evolution and phylogeny. Annu. Rev. Entomol. 59, 95–117.
Cameron, S.L., Whiting, M.F., 2007. Mitochondrial genomic comparisons of the subterranean termites from the Genus *Reticulitermes* (Insecta: Isoptera: Rhinotermitidae). Genome 50, 188–202.
Cameron, S.L., Johnson, K.P., Whiting, M.F., 2007a. The mitochondrial genome of the screamer louse *Bothriometopus* (Phthiraptera: Ischnocera): Effects of extensive gene rearrangements on the evolution of the genome. J. Mol. Evol. 65, 589–604.
Cameron, S.L., Lambkin, C.L., Barker, S.C., Whiting, M.F., 2007b. A mitochondrial genome phylogeny of Diptera: Whole genome sequence data accurately resolve relationships over broad timescales with high precision. Syst. Entomol. 32, 40–59.
Cameron, S.L., Dowton, M., Castro, L.R., Ruberu, K., Whiting, M.F., Austin, A.D., Diement, K., Stevens, J., 2008. Mitochondrial genome organization and phylogeny of two vespid wasps. Genome 51, 800–808.
Cameron, S.L., Yoshizawa, K., Mizukoshi, A., Whiting, M.F., Johnson, K.P., 2011. Mitochondrial genome deletions and minicircles are common in lice (Insecta: Phthiraptera). BMC Genom. 12, 394.
Castro, L., Dowton, M., 2005. The position of the Hymenoptera within the Holometabola as inferred from the mitochondrial genome of *Perga condei* (Hymenoptera: Symphyta: Pergidae). Mol. Phylogenet. Evol. 34, 469–479.
Castro, L.R., Ruberu, K., Dowton, M., 2006. Mitochondrial genomes of *Vanhornia eucnemidarum* (Apocrita: Vanhorniidae) and *Primeuchroeus* spp. (Aculeata: Chrysididae): Evidence of rearranged mitochondrial genomes within the Apocrita (Insecta: Hymenoptera). Genome 49, 752–766.
Cha, S.Y., Yoon, H.J., Lee, E.M., Yoon, M.H., Hwang, J.S., Jin, B.R., Han, Y.S., Kim, I., 2007. The complete nucleotide sequence and gene organization of the mitochondrial genome of the bumblebee, *Bombus ignitus* (Hymenoptera: Apidae). Gene 392, 206–220.
Chen, W.J., Bu, Y., Carapelli, A., Dallai, R., Li, S., Yin, W.Y., Luan, Y.X., 2011. The mitochondrial genome of *Sinentomon erythranum* (Arthropoda: Hexapoda: Protura): An example of highly divergent evolution. BMC Evol. Biol. 11, 246.
Covacin, C., Shao, R., Cameron, S., Barker, S.C., 2006. Extraordinary number of gene rearrangements in the mitochondrial genomes of lice (Phthiraptera: Insecta). Insect Mol. Biol. 15, 63–68.
Crozier, R.H., Crozier, Y.C., 1993. The mitochondrial genome of the honeybee *Apis mellifera*: Complete sequence and genome organization. Genetics 133, 97–117.
Curole, J.P., Kocher, T.D., 1999. Mitogenomics: Digging deeper with complete mitochondrial genomes. Trends Ecol. Evol. 14, 394–398.
Dowton, M., Austin, A.D., 1994. Molecular phylogeny of the insect order Hymenoptera: Apocritan relationships. Proc. Natl. Acad. Sci. USA 91, 9911–9915.
Dowton, M., Austin, A.D., 1999. Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. Mol. Biol. Evol. 16, 298–309.
Dowton, M., Austin, A.D., 2001. Simultaneous analysis of 16S, 28S, COI and morphology in the Hymenoptera: Apocrita-evolutionary transitions among parasitic wasps. Biol. J. Linn. Soc. 74, 87–111.

Dowton, M., Austin, A., Dillon, N., Bartowsky, E., 1997. Molecular phylogeny of the apocritan wasps: The Proctotrupomorpha and Evaniomorpha. Syst. Entomol. 22, 245–255.
Dowton, M., Castro, L.R., Austin, A.D., 2002. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: The examination of genome 'morphology'. Invertebr. Syst. 16, 345–356.
Dowton, M., Castro, L.R., Campbell, S.L., Bargon, S.D., Austin, A.D., 2003. Frequent mitochondrial gene rearrangements at the Hymenopteran *nad3–nad5* junction. J. Mol. Evol. 56, 517–526.
Dowton, M., Cameron, S.L., Austin, A.D., Whiting, M.F., 2009a. Phylogenetic approaches for the analysis of mitochondrial genome sequence data in the Hymenoptera – A lineage with both rapidly and slowly evolving mitochondrial genomes. Mol. Phylogenet. Evol. 52, 512–519.
Dowton, M., Cameron, S.L., Dowavic, J.I., Austin, A.D., Whiting, M.F., 2009b. Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral. Mol. Biol. Evol. 26, 1607–1617.
Fenn, J.D., Song, H., Cameron, S.L., Whiting, M.F., 2008. A preliminary mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data. Mol. Phylogenet. Evol. 49, 59–68.
Gotzek, D., Clarke, J., Shoemaker, D., 2010. Mitochondrial genome evolution in fire ants (Hymenoptera: Formicidae). BMC Evol. Biol. 10, 300.
Hasegawa, E., Kobayashi, K., Yagi, N., Tsuji, K., 2011. Complete mitochondrial genomes of normal and cheater morphs in the parthenogenetic ant *Pristomyrmex punctatus* (Hymenoptera: Formicidae). Myrmecol News 15, 85–90.
Hong, M., Cha, S., Kim, D., Yoon, H., Kim, S., Hwang, J., Kim, K., Han, Y., Kim, I., 2008. Presence of several tRNA-like sequences in the mitochondrial genome of the bumblebee, *Bombus hypocrita sapporoensis* (Hymenoptera: Apidae). Genes Genom. 30, 307–318.
Kaltenpoth, M., Showers Corneli, P., Dunn, D.M., Weiss, R.B., Strohm, E., Seger, J., 2012. Accelerated evolution of mitochondrial but not nuclear genomes of hymenoptera: New evidence from crabronid wasps. PloS ONE 7, e32826.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Klopfstein, S., Vilhelmsen, L., Heraty, J.M., Sharkey, M., Ronquist, F., 2013. The hymenopteran tree of life: Evidence from protein-coding genes and objectively aligned ribosomal data. PLoS ONE 8, e69344.
Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701.
Leavitt, J.R., Hiatt, K.D., Whiting, M.F., Song, H.J., 2013. Searching for the optimal data partitioning strategy in mitochondrial phylogenomics: A phylogeny of Acridoidea (Insecta: Orthoptera: Caelifera) as a case study. Mol. Phylogenet. Evol. 67, 494–508.
Lelej, A.S., 2004. Catalogue of the Mutillidae (Hymenoptera) of the Oriental Region. Dalnauka, Vladivostok.
Lelej, A.S., Brothers, D.J., 2008. The genus-group names of Mutillidae (Hymenoptera) and their type species, with a new genus, new name, new synonymies, new combinations and lectotypifications. Magnolia Press, Auckland, New Zealand.
Lemmon, A.R., Brown, J.M., Stanger-Hall, K., Lemmon, E.M., 2009. The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. Syst. Biol. 58, 130–145.
Mao, M., Valerio, A., Austin, A.D., Dowton, M., Johnson, N.F., 2012. The first mitochondrial genome for the wasp superfamily Platygastroidea: The egg parasitoid *Trissolcus basalis*. Genome 55, 194–204.
Negrisolo, E., Minelli, A., Valle, G., 2004. The mitochondrial genome of the house centipede *Scutigera* and the monophyly versus paraphyly of myriapods. Mol. Biol. Evol. 21, 770–780.
Oliveira, D.C.S.G., Raychoudhury, R., Lavrov, D.V., Werren, J.H., 2008. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). Mol. Biol. Evol. 25, 2167–2180.
Pilgrim, E.M., von Dohlen, C.D., Pitts, J.P., 2008. Molecular phylogenetics of Vespoidea indicate paraphyly of the superfamily and novel relationships of its component families and subfamilies. Zool. Scr. 37, 539–560.
Posada, D., 2009. Selection of models of DNA evolution with jModelTest. Methods Mol. Biol. 537, 93–112.
Rambaut, A., 2012. FigTree, a graphical viewer of phylogenetic trees. Available from <http://tree.bio.ed.ac.uk/software/figtree>.
Rambaut, A., Drummond, A., 2007. Tracer v1.4. Available from <http://beast.bio.ed.ac.uk/Trace>.
Rasnitsyn, A.P., 1988. An outline of evolution of the hymenopterous insects. Orient. Insects 22, 115–145.
Rasnitsyn, A.P., Basibuyuk, H.H., Quicke, D.L.J., 2004. A basal chalcidoid (Insecta: Hymenoptera) from the earliest Cretaceous or latest Jurassic of Mongolia. Insect. Syst. Evol. 35, 123–135.
Ronquist, F., Rasnitsyn, A.P., Roy, A., Eriksson, K., Lindgren, M., 1999. Phylogeny of the Hymenoptera: A cladistic reanalysis of Rasnitsyn's (1988) data. Zool. Scr. 28, 13–50.
Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
San Mauro, D., Gower, D.J., Zardoya, R., Wilkinson, M., 2006. A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome. Mol. Biol. Evol. 23, 227–234.
Shao, R., Barker, S., 2003. The highly rearranged mitochondrial genome of the plague thrips, *Thrips imaginis* (Insecta: Thysanoptera): Convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes. Mol. Biol. Evol. 20, 362–370.
Shao, R., Campbell, N.J.H., Barker, S.C., 2001a. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). Mol. Biol. Evol. 18, 858–865.
Shao, R.F., Campbell, N.J.H., Schmidt, E.R., Barker, S.C., 2001b. Increased rate of gene rearrangement in the mitochondrial genomes of three orders of hemipteroid insects. Mol. Biol. Evol. 18, 1828–1832.
Sharkey, M.J., 2007. Phylogeny and classification of Hymenoptera. Zootaxa 1668, 521–548.
Sharkey, M.J., Roy, A., et al., 2002. Phylogeny of the Hymenoptera: A reanalysis of the Ronquist et al. (1999) reanalysis, emphasizing wing venation and apocritan relationships. Zool. Scr. 31, 57–66.
Sharkey, M.J., Carpenter, J.M., Vilhelmsen, L., Heraty, J., Liljeblad, J., Dowling, A.P.G., Schulmeister, S., Murray, D., Deans, A.R., Ronquist, F., 2012. Phylogenetic relationships among superfamilies of Hymenoptera. Cladistics 28, 80–112.
Shoemaker, D.D., Ahrens, M.E., Ross, K.G., 2006. Molecular phylogeny of fire ants of the *Solenopsis saevissima* species-group based on mtDNA sequences. Mol. Phylogenet. Evol. 38, 200–215.
Silvestre, D., Arias, M.C., 2006. Mitochondrial tRNA gene translocations in highly eusocial bees. Genet. Mol. Biol. 29, 572–575.
Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.
Simon, C., Buckley, T.R., Frati, F., Stewart, J.B., Beckenbach, A.T., 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. Annu. Rev. Ecol. Syst. 37, 545–579.
Stamatakis, A., 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
Swofford, D.L., 2002. PAUP*: Phylogenetic analysis using parsimony, version 4.0 b10. Sinauer Associates, Sunderland, Massachusetts.
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.
Thomson, R.C., Shaffer, H.B., 2010. Sparse supermatrices for phylogenetic inference: Taxonomy, alignment, rogue taxa, and the phylogeny of living turtles. Syst. Biol. 59, 42–58.
Wei, S.J., Shi, M., He, J.H., Sharkey, M.J., Chen, X.X., 2009. The complete mitochondrial genome of *Diadegma semiclausum* (Hymenoptera: Ichneumonidae) indicates extensive independent evolutionary events. Genome 52, 308–319.
Wei, S.J., Shi, M., Sharkey, M.J., van Achterberg, C., Chen, X.X., 2010a. Comparative mitogenomics of Braconidae (Insecta: Hymenoptera) and the phylogenetic utility of mitochondrial genomes with special reference to holometabolous insects. BMC Genom. 11, 371.
Wei, S.J., Tang, P., Zheng, L.H., Shi, M., Chen, X.X., 2010b. The complete mitochondrial genome of *Evania appendigaster* (Hymenoptera: Evaniidae) has low A+ T content and a long intergenic spacer between *atp8* and *atp6*. Mol. Biol. Rep. 37, 1931–1942.
Wei, S.J., Wu, Q.L., Liu, W., 2013a. Sequencing and characterization of the *Monocellicampa pruni* (Hymenoptera: Tenthredinidae) mitochondrial genome. Mitochondr. DNA, doi:10.3109/19401736.19402013.19819501.
Wei, S.J., Wu, Q.L., van Achterberg, K., Chen, X.X., 2013b. Rearrangement of the *nad1*gene in *Pristaulacus compressus* (Spinola) (Hymenoptera: Evanioidea: Aulacidae) mitochondrial genome. Mitochondr. DNA, doi:10.3109/19401736.19402013.19834436.
Wernersson, R., Pedersen, A.G., 2003. RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences. Nucl. Acids Res. 31, 3537.
Wiens, J.J., Moen, D.S., 2008. Missing data and the accuracy of Bayesian phylogenetics. J. Syst. Evol. 46, 307–314.
Wiens, J.J., Morrill, M.C., 2011. Missing data in phylogenetic analysis: Reconciling results from simulations and empirical data. Syst. Biol. 60, 719–731.
Wolstenholme, D.R., 1992. Animal mitochondrial DNA: Structure and evolution. Int. Rev. Cytol. 141, 173–216.
Xiao, J.H., Jia, J.G., Murphy, R.W., Huang, D.W., 2011. Rapid evolution of the mitochondrial genome in chalcidoid wasps (Hymenoptera: Chalcidoidea) driven by parasitic lifestyles. PLoS ONE 6, e26645.
Xu, Z.F., He, J.H., 2006. First record of *Cephalonomia gallicola* (Ashmead) (Hymenoptera: Bethylidae) on tobacco beetle from China. Entomotaxonomia 28, 309–310.
Zhang, D.X., Hewitt, G.M., 1997. Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. Biochem. Syst. Ecol. 25, 99–120.