International Journal of Biological Macromolecules 147 (2020) 1193–1203

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

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

The mitochondrial genomes of ladybird beetles and implications for evolution and phylogeny

Nan Song *, Xinxin Li, Xinming Yin, Xinghao Li, Yuqiang Xi

College of Plant Protection, Henan Agricultural University, Zhengzhou, China

ARTICLE INFO

Article history:
Received 3 July 2019
Received in revised form 8 October 2019
Accepted 8 October 2019
Available online 18 November 2019

Keywords:
Coccinellidae
Phylogeny
Food preference
Mitochondrial genome
Next-generation sequencing

ABSTRACT

Ladybirds formed the most familiar beetle group, namely the family Coccinellidae, whose internal relationships remain unclear. In particular, the subfamily relationships could not be well resolved in previous studies based on the conventional nuclear and/or mitochondrial gene fragments. In this study, we used next-generation sequencing to obtain new mitochondrial genomes (mitogenomes) from 13 species representing four coccinellid subfamilies (i.e., Coccinellinae, Epilachninae, Coccidulinae and Chilocorinae). Together with 24 existing mitogenome sequences of Cucujoidea, we conducted phylogenetic analyses to investigate the deep phylogenetic relationships in Coccinellidae, under maximum likelihood and Bayesian inference criteria. The analyses from nucleotide datasets resulted in a largely identical tree topology, where Epilachninae and Coccinellinae were monophyletic groups. The Scymninae and Coccidulinae were recovered as non-monophyletic. Amino acids differed from nucleotides in that the Epilachninae was retrieved as paraphyletic, with respect to *Epilachna admirabilis*. Ancestral state reconstruction suggested that the plant eating ladybird beetles arose within an aphidophagous/coccidophagous clade. In addition, three independent shifts toward coccidophagy and one shift toward mycophagy occurred in Coccinellidae.

1. Introduction

Ladybird beetles are well known to most people due to their beautifully colorful and spotty wing patterns. These beetle insects constitute the family Coccinellidae, which comprises approximately 6000 species assigned to nearly 360 genera [1]. An important aspect of ladybirds is that the predatory species of Coccinellidae are usually recognized as biological control agents. Most coccinellids are predaceous specialists on the insect pests of the hemipteran suborder Sternorrhyncha [2,3]. In the world, the first successful biological control example is the introduction of the Australian ladybird *Rodolia cardinalis* to California to control *Icerya purchasi* (cottony cushion scale) in citrus groves during the late 19th century.

Despite the familiarity and the great economic and ecological significance, there are few molecular phylogenetic studies focused on the Coccinellidae. The different classification systems at the subfamily level have been proposed for Coccinellidae depending on the morphological taxonomists: the two subfamilies scheme (the phytophagous group and the aphidophagous group) [4], three (Epilachninae, Coccinellinae and Lithophilinae) [5], six (Sticholotidinae, Coccidulinae, Scymninae, Chilocorinae, Coccinellinae and Epilachninae) [6–8], and seven (the additional subfamily of Ortaliinae into the scheme of six subfamilies) [9]. Yu (1994) investigated the higher-level relationships within Coccinellidae using a cladistic analysis of adult and larval morphological characters [10]. However, only the relationships of Coccinellini and Sukunahikonini were congruently supported by two kinds of characters [10]. The morphological characters alone seemed to be insufficient to solve the subfamily and tribe-level classification [9,11,12], challenges for resolving the higher-level relationships of Coccinellidae remain.

With the increasing availability of multiple genetic markers from various sequencing projects, molecular phylogenetic analyses have been widely used in the Coleoptera systematics. Recently, three studies have attempted to address the issues of the phylogenetic relationships among subfamilies and the evolution of food relationships of Coccinellidae by using DNA sequence data [11,13,14]. However, inconsistent results emerged from different gene fragments examined and taxon sampling included. All subfamilies was not supported as monophyletic, except for Coccinellinae [11,13,14]. Several phylogenetic studies on the higher-level relationships within Coleoptera also involved the family Coccinellidae [15–19]. The monophyly of the entire family Coccinellidae is

* Corresponding author.
E-mail address: songnan@henau.edu.cn (N. Song).

https://doi.org/10.1016/j.ijbiomac.2019.10.089
0141-8130/

often confirmed. The authors tend to recognize Endomychidae as one of the closest relatives of Coccinellidae [15,16,18].

Ladybird beetles have a wide range of food types in their diets. The subject of coccinellid food preferences has been comprehensively reviewed by Giorgi et al. [13] and Sloggett and Majerus [20]. According to the literature published [13], the subfamily Epilachninae and the genus *Bulaea* are the exclusive leaf-eaters, while several small groups within Coccinellinae, the Halyziini and Tythaspidini are the fungus feeders. Most of the remaining coccinellid species are predators of different insects. Moreover, many species are known to utilize alternative food resources when there no adequate preferred ones. The food items of predaceous Coccinellidae include aphids, psylloids, whiteflies, coccids, eggs and larvae of leaf beetles, bugs, and ants, as well as honeydew, pollen, plant sap, nectar, and various fungi [13,20–24]. The feeding habits of ladybird beetles could be associated with their fitness to the host population dynamics. For example, the availability of food may be one of the factors involved in egg production and oviposition of the aphidophagous ladybird beetles [25].

Mitochondrial genome (mitogenome) sequences have been shown to be useful for investigating beetle relationships at various taxonomic scales [26–30]. The advent of next-generation sequencing (NGS) technologies has made it easy to simultaneously acquire whole mitogenomes for numerous exemplars. As of January 2019, there are only 22 mitogenomes available for ladybird beetles in GenBank, and few published studies utilize mitogenomes to reconstruct the phylogeny of Coccinellidae.

In the present study, we used a method of next-generation sequencing of multiple pooled genomic DNA to obtain mitogenome sequences for 13 ladybird species, which represented the subfamilies: Coccinellinae (seven species), Epilachninae (four species), Coccidulinae (one species) and Chilocorinae (one species). The newly generated data were combined with the publicly available Cucujoidea mitogenomes to infer the higher-level phylogeny of Coccinellidae, with an attempt to provide insight into the evolution of feeding patterns.

2. Materials and methods

2.1. Taxon sampling

Specimens were collected from Zhengzhou and Xinyang, Henan province, China. No specific permits were required for the insects sampled for this study. Specimen identification was conducted by checking adult morphological characters (with reference to Ren et al. 2009; Yu 2010) [31,32], and blasting mitochondrial *cox1* gene fragment in public databases (i.e., BOLD, NCBI). Combined with published mitogenomes of Coccinellidae in GenBank, thirty-one species belonging to eight tribes of five subfamilies of Coccinellidae were used to construct ingroup taxa. Six outgroup species were included to represent six other families in Cucujoidea, of which the *Loberonotha olivascens* (Erotylidae) was used to root the tree. The detailed classification information, GenBank accession numbers, and the voucher ID for new sequences are listed in Table S1.

The mitogenome sequence of Coccinellidae sp. 1 EF-2015 (GenBank Accession: KT780638) downloaded from GenBank was identified as belonging to the genus *Scymnus* (Scymninae) by using the mitochondrial *cox1* gene to conduct the molecular identification in BOLD systems (sequence identity = 94.98% for the genus *Scymnus*) and in NCBI Standard Nucleotide BLAST (sequence identity = 99% for *Scymnus* sp. CO449, GenBank Accession: KP829557). Thus, the species name of this mitogenome sequence may be corrected to be *Scymnus* sp. in this study (see details in Table S1).

2.2. DNA extraction

Genomic DNA extraction was conducted from the thoracic muscle tissue of the single 100% ethanol preserved specimen, with the TIANamp Micro DNA Kit (TIANGEN BIOTECH CO., LTD) following the manufacturer's protocol. The DNA concentration was measured for each sample by using the nucleic acid protein analyzer (QUAWELL TECHNOLOGY INC.).

2.3. Mitogenome reconstruction

The assembly strategy of complete mitogenome is largely identical to that of Song et al. [33]. The uniform quantities of genomic DNA from each of samples were pooled to improve the efficiency of genome sequencing. Three mixed DNA pools were prepared, each of which contained other insect species having the distantly phylogenetic relationships to Coccinellidae. The sequencing libraries were constructed using Illumina TruSeqTM DNA Sample Prep Kit (Illumina, San Diego, CA, USA), with an insert size of 350 bp. The subsequent *de novo* genome sequencing were conducted on an Illumina HiSeq2500 platform at Shanghai OE Biotech CO., LTD, with a strategy of 150 paired-end sequencing. This produced 144,285,457, 152,492,351 and 170,937,533 paired-end reads for each library.

We used the NGS QC toolkit [34] to filter the data for quality control, which resulted in 134,957,025, 144,413,185 and 163,224,652 high-quality reads (cutoff read length for HQ = 70%, cutoff quality score = 20). The high-quality reads were used in *de novo* assembly using IDBA-UD v. 1.1.1 [35]. The assemblies were constructed using 200 for the setting of minimum size of contig, and an initial k-mer size of 40, an iteration size of 10, and a maximum k-mer size of 90.

Two mitochondrial gene fragments (i.e., *cox1*-5′ and *cob*-3′) were amplified with the primers in Song et al. [33] and pre-sequenced through Sanger sequencing. These sequences were used as baits to identify the mitochondrial contigs in the NGS data, with local-blasting implemented in BioEdit [36]. The mitochondrial sequences obtained were annotated with MITOS [37], using default settings and the invertebrate genetic code for mitochondria. Gene boundaries were further checked and refined by alignment against other published mitogenome sequences of Coccinellidae (the detailed species names are provided in Table S1). Mappings to the mitochondrial contigs were performed using BWA v. 0.7.5 [38]. Alignments produced in SAM format were converted to sorted BAM format by SAMtools v. 0.1.19 [39]. Statistics for nucleotide coverage were generated with Qualimap v.2.2.1 [40].

Table 1
The sequence length and statistics for the sequencing of each mitochondrial contig in the newly determined species.

Species	Length	Mapped bases	Mean coverage
Afissula kambaitana	14,407	42,32,389	294
Afissula sp.	17,420	52,38,232	301
Aiolocaria hexaspilota	17,396	1,01,48,959	583
Calvia muiri	17,130	50,24,079	293
Chilocorus bipustulatus	12,229	2,85,01,225	2331
Coccinella septempunctata	12,684	72,11,411	569
Coelophora saucia	11,776	9,34,158	79
Epilachna admirabilis	18,064	82,14,150	458
Harmonia axyridis	13,549	1,13,15,865	835
Henosepilachna vigintioctomaculata	16,818	24,11,499	143
Illeis cincta	15,756	32,16,062	204
Propylaea japonica	17,471	60,76,535	348
Rodolia quadrimaculata	13,871	1,08,01,050	773

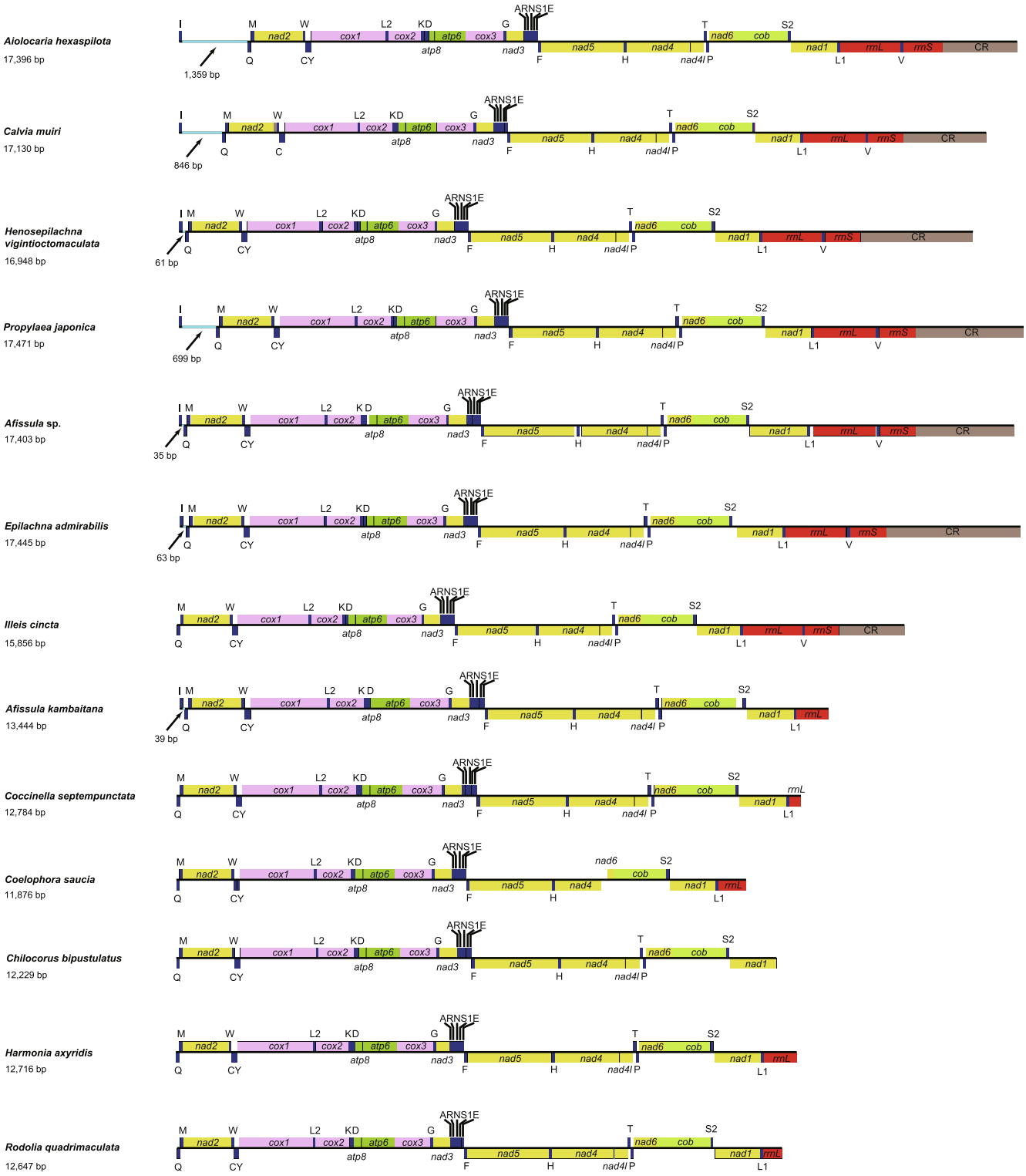

Fig. 1. Organizational maps of the 13 new mitogenomes sequenced in this study. Genes labelled above the line are transcribed in the same direction from left to right, while genes labelled below the line are transcribed in the same direction from right to left. The genes and intergenic spacers are scaled to their length in the mitogenome. Abbreviations: I, transfer RNA specifying Isoleucine; Q, transfer RNA specifying Glutamine; M, transfer RNA specifying Methionine; W, transfer RNA specifying Tryptophan; C, transfer RNA specifying Cysteine; Y, transfer RNA specifying Tyrosine; K, transfer RNA specifying Lysine; D, transfer RNA specifying Aspartic acid; L2, transfer RNA specifying Leucine, codon recognized by UUR; G, transfer RNA specifying Glycine; A, transfer RNA specifying Alanine; R, transfer RNA specifying Arginine; N, transfer RNA specifying Asparagine; S1, transfer RNA specifying Serine, codon recognized by AGN; E, transfer RNA specifying Glutamic acid; F, transfer RNA specifying Phenylalanine; H, transfer RNA specifying Histidine; T, transfer RNA specifying Threonine; P, transfer RNA specifying Proline; S2, transfer RNA specifying Serine, codon recognized by UCN; L1, transfer RNA specifying Leucine, codon recognized by CUN; V, transfer RNA specifying Valine; *cox1*, *cox2*, *cox3*, cytochrome oxidase subunits I, II, III; *cob*, cytochrome *b* apoenzyme; *nad 1–6*, *4L*, NADH dehydrogenase subunits 1–6, 4L; *atp6*, *atp8*, ATP synthase subunits 6, 8; *rrnL*, large ribosomal subunit; *rrnS*, small ribosomal subunit; CR, the putative control region.

Table 2
Saturation test for protein-coding genes and RNA genes.

Gene fragment	NumOTU	*Iss*	*Iss.cSym*	Psym	*Iss.cAsym*	Pasym
PCG#1	32	0.479	0.808	0.0000	0.553	0.0000
PCG#2	32	0.349	0.808	0.0000	0.553	0.0000
PCG#3	32	0.873	0.808	0.0000	0.553	0.0000
PCG#123	32	0.530	0.818	0.0000	0.572	0.0000
tRNAs	32	0.541	0.770	0.0000	0.483	0.0001
rRNAs	32	0.844	0.787	0.0000	0.545	0.0000
tRNAs + rRNAs	32	0.817	0.807	0.4540	0.550	0.0000

Note: *Iss*, index of substitution saturation; *Iss.cSym*, index of substitution saturation assuming a symmetrical true tree; Psym, probability of significant difference between *Iss* and *Iss.cSym* (two-tailed test); *Iss.cAsym*, index of substitution saturation assuming an asymmetrical true tree; Pasym, probability of significant difference between *Iss* and *Iss.cAsym* (two-tailed test).

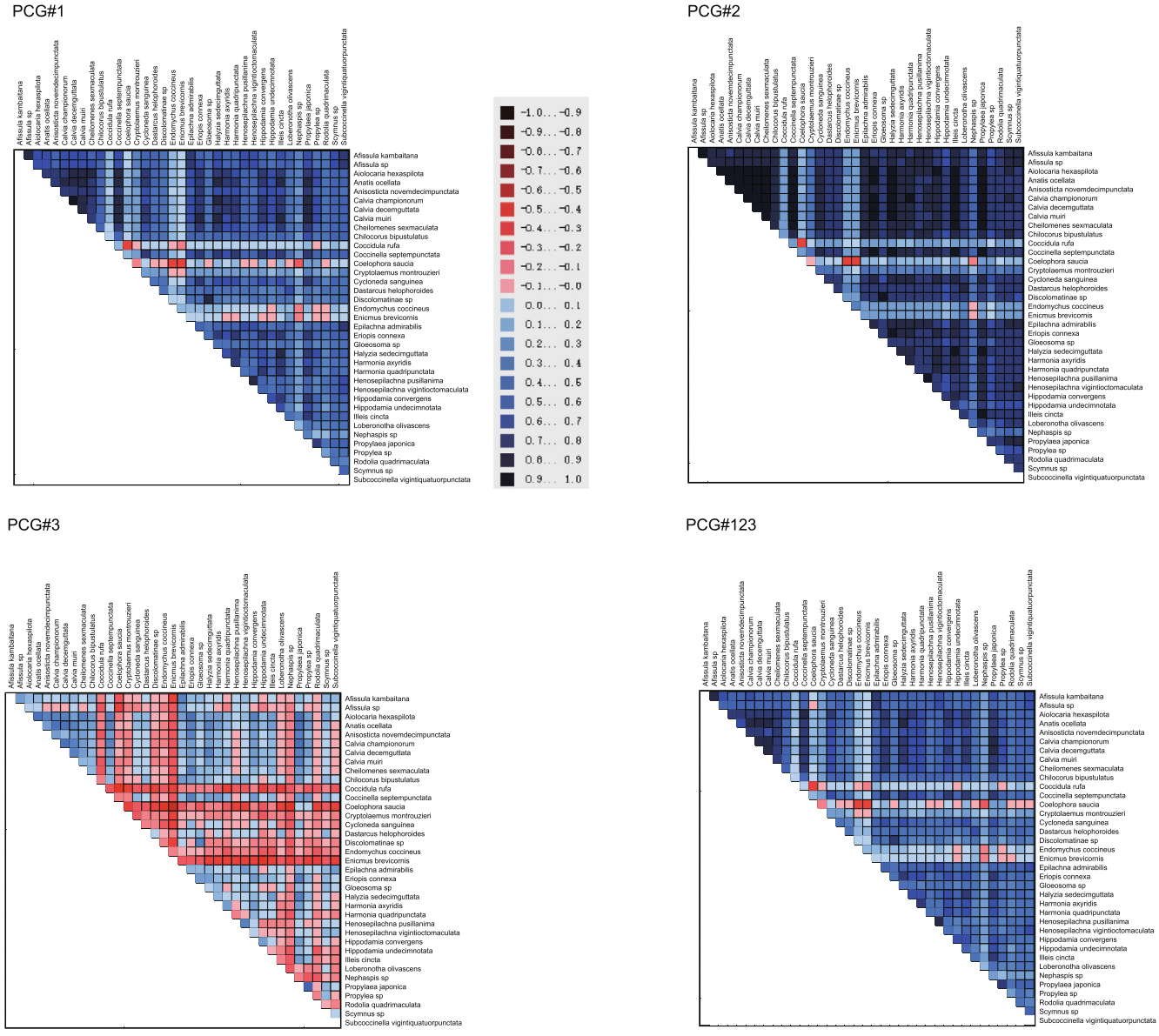

Fig. 2. Heterogeneous sequence divergence within datasets including only the first, the second, the third codon positions or all codon positions. The obtained mean similarity score between sequences was represented by a colored square. The scores were ranging from −1, indicating full random similarity, to +1, non-random similarity. The darker red indicated the higher randomized similarities between pairwise sequence comparisons. Blue indicated the opposite. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.4. Sequence alignment

We built alignments of protein-coding genes using TranslatorX [41], based on the invertebrate mitochondrial genetic code with MAFFT algorithm [42]. Each of tRNA and rRNA gene alignments was conducted using the online version of MAFFT [42], applying the E-INS-i strategy. The resulting alignments were checked in MEGA 6 [43]. Highly variable regions that were hard to align were trimmed using GBlocks [44], with all options for a less stringent selection. The gene losses present in the mitogenomes were coded as missing characters in the subsequent phylogenetic analysis. Finally, all alignments were concatenated to create two classes of datasets using FASconCAT_v1.0 [45], namely those with or without RNA genes: PCGRNA and PCG.

To reduce the effect of nucleotide compositional heterogeneity on phylogenetic estimate, the following three datasets were compiled for tree building: (1) PCGRNA: the 13 protein-coding genes including all codon positions combined with the 24 RNA genes; (2) PCG12RNA: the 13 protein-coding genes excluding the third-codon positions combined with the 24 RNA genes; (3) PCG_AA: the 13 protein-coding genes being translated into amino acids.

Potential saturation in different types of gene partitions were assessed using the index of substitution saturation (*Iss*) as implemented in DAMBE 5 [46]. The mean *ka* (nonsynonymous substitution rate) and *ks* (synonymous substitution rate) values were calculated by using DnaSP version 5 [47]. The sequence heterogeneity within datasets were analyzed by using AliGROOVE [48], with the default sliding window size.

Table 3
The substitution rate analyses conducted on the protein-coding genes using DnaSP.

Species	*ks*	*ka*	*ka*/*ks*
Illeis cincta	0.7984	0.2052	0.2570
Calvia championorum	0.7145	0.1863	0.2608
Harmonia quadripunctata	0.8255	0.2190	0.2653
Calvia muiri	0.7297	0.1953	0.2677
Afissula sp.	0.7672	0.2108	0.2748
Calvia decemguttata	0.6745	0.1899	0.2815
Halyzia sedecimguttata	0.6851	0.1937	0.2827
Coccinella septempunctata	0.6911	0.1981	0.2866
Propylea sp.	0.6499	0.1900	0.2923
Cycloneda sanguinea	0.7831	0.2290	0.2924
Anatis ocellata	0.6289	0.1872	0.2977
Coelophora saucia	0.6430	0.1924	0.2992
Hippodamia undecimnotata	0.6944	0.2113	0.3043
Propylaea japonica	0.6225	0.1896	0.3046
Hippodamia convergens	0.6821	0.2092	0.3067
Eriopis connexa	0.6422	0.1983	0.3087
Cheilomenes sexmaculata	0.6596	0.2048	0.3104
Subcoccinella vigintiquatuorpunctata	0.6971	0.2198	0.3153
Afissula kambaitana	0.6723	0.2130	0.3169
Enicmus brevicornis	0.8027	0.2548	0.3175
Coccidula rufa	0.6791	0.2289	0.3370
Anisosticta novemdecimpunctata	0.6334	0.2138	0.3375
Harmonia axyridis	0.6556	0.2215	0.3378
Aiolocaria hexaspilota	0.5917	0.2000	0.3380
Epilachna admirabilis	0.6099	0.2098	0.3439
Henosepilachna pusillanima	0.7152	0.2494	0.3487
Loberonotha olivascens	0.7446	0.2599	0.3490
Scymnus sp.	0.6256	0.2272	0.3632
Henosepilachna vigintioctomaculata	0.6700	0.2448	0.3653
Cryptolaemus montrouzieri	0.6371	0.2387	0.3746
Nephaspis sp.	0.6982	0.2635	0.3774
Discolomatinae sp.	0.7344	0.2797	0.3809
Chilocorus bipustulatus	0.6062	0.2327	0.3840
Dastarcus helophoroides	0.6422	0.2598	0.4045
Rodolia quadrimaculata	0.7054	0.2972	0.4213
Endomychus coccineus	0.5689	0.2606	0.4581
Gloeosoma sp.	0.5652	0.2770	0.4901

Note: *ks*, synonymous substitution rate; *ka*, nonsynonymous substitution rate.

2.5. Phylogenetic analyses

Tree searches were conducted under maximum likelihood (ML) and Bayesian inference (BI) criteria. The partitioned ML analyses were performed using IQ-TREE [49]. The prior data blocks were defined by gene types. The best-fitting substitution models for each partition were chosen using ModelFinder [50]. The partition schemes and models used are provided in Table S2. Branch support analysis was conducted using 1000 ultrafast bootstrap replicates. Bayesian analyses were conducted using PhyloBayes [51,52] as implemented in the CIPRES (Test) [53]. The site-heterogeneous CAT-GTR model was used for nucleotide analyses, while the CAT model was applied to amino acids. Two chains were run in parallel, and started from a random topology. The bpcomp program was used to calculate the largest (maxdiff) and mean (meandiff) discrepancy observed across all bipartitions. The program tracecomp was used to summarize the discrepancies and the effective sizes estimated for each column of the trace file. Analyses were terminated early if after a burn-in of 1000 cycles, the maxdiff value was lower than 0.1 and minimum effective size was higher than 300. A consensus tree was calculated from the saved trees by bpcomp program when the stationarity was reached.

We used Mesquite version 3.31 [54] to estimate ancestral states of food preferences, based on the framework of PCGRNA-ML tree. Information on the studied characters relevant to feeding patterns for each terminal taxon was retrieved from the literature. These characters were treated as discrete and unordered. Analysis of "Trace Character History" was conducted by applying 'Parsimony Ancestral States' as the method of ancestral state reconstruction.

3. Results

3.1. Mitogenome assembly

Thirteen new mitogenomes of ladybird beetles were identified from the individual contigs by bait sequences. All bait sequences can match to the large mitochondrial contigs with certainty (Identities $\geq$ 99% and E-Value = 0.0). The sequence length and statistics for the sequencing of each mitochondrial contig are presented in Table 1. Coverage reached over 200-fold in most cases.

3.2. Genome organization and structure

The complete mitogenomes were recovered for four species: *Aiolocaria hexaspilota*, *Calvia muiri*, *Henosepilachna vigintioctomaculata* and *Propylaea japonica*, which included the full 37 mitochondrial genes (13 protein-coding genes, 22 tRNA genes and two rRNA genes) and the entire control region. Three species of *Afissula* sp., *Epilachna admirabilis* and *Illeis cincta* had the nearly complete mitogenomes, which respectively contained 36–37 mitochondrial genes and a partial control region. The *Afissula* sp. and *E. admirabilis* respectively harbored the typical 37 mitochondrial genes, but had an incomplete control region. The *trnI* was not detected in *I. cincta.* The arrangement and orientation of the mitochondrial genes in the complete or nearly complete mitogenomes were identical to the presumed ancestral insect mitogenome [55]. The remaining six species newly sequenced had the partial mitogenomes, in which 31–35 mitochondrial genes were identified, respectively. For the partial mitogenomes, the missing regions were mainly located adjacent to the putative control region. That is, we failed to reconstruct the *trnV*, *rrnS* and the entire control region in the six species. The organizational maps of the new mitogenomes are shown in Fig. 1.

The organization of the ladybird beetle mitogenomes is generally compact, with no more than 100 nucleotides scattered in

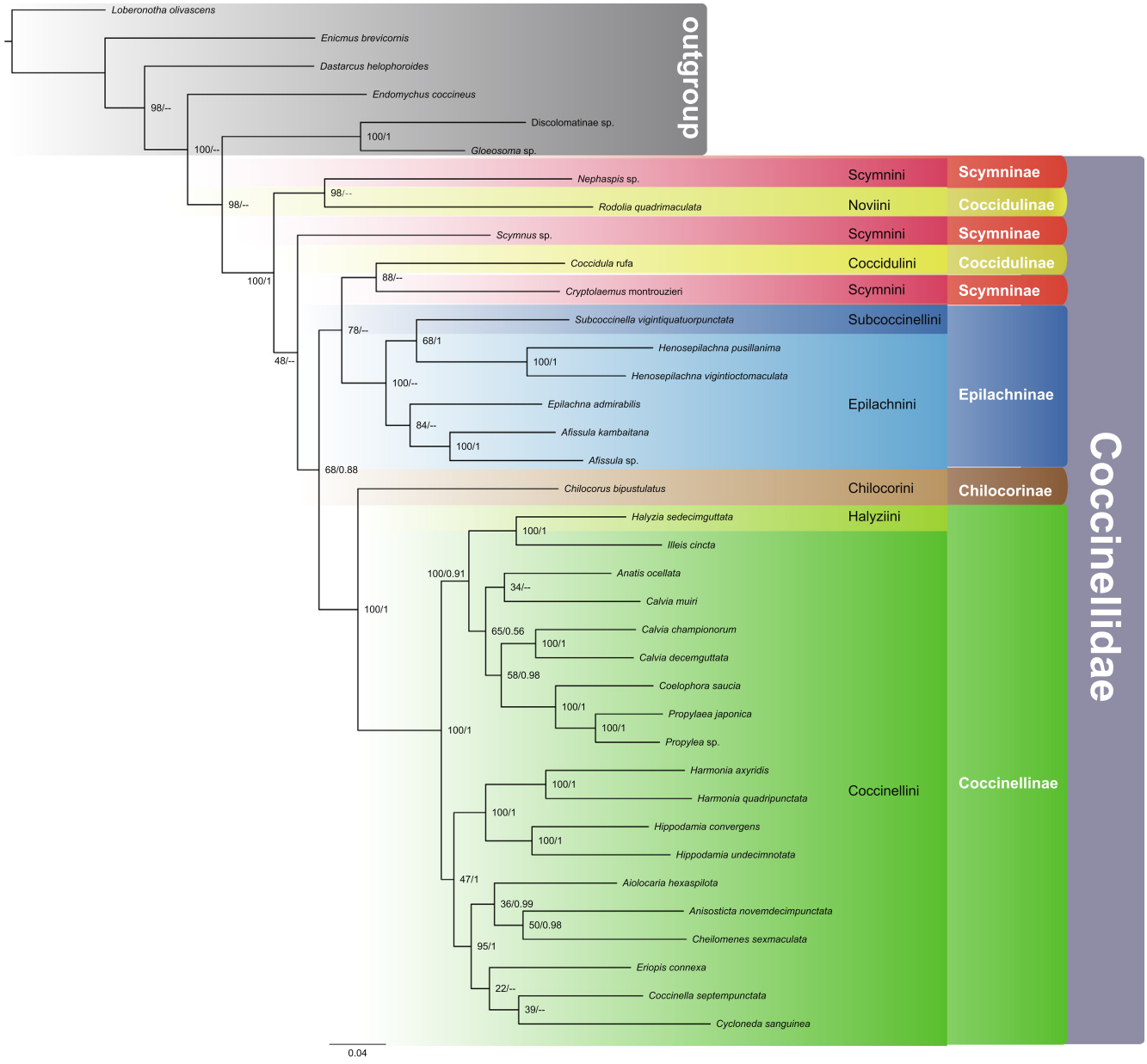

Fig. 3. Maximum likelihood tree inferred from the dataset of PCGRNA using the predefined partition schemes under the best-fitting models selected by ModelFinder. Node numbers show bootstrap support values (left) and Bayesian posterior probability support values (right). Scale bar represents substitutions/site. "–" indicates that the node is not recovered by BI analysis.

several intergenic spacers. Each intergenic spacer ranged in size between 1 and 20 bp. However, a large intergenic spacer between *trnI* and *trnQ* was detected in *A. hexaspilota*, *C. muiri* and *P. japonica*, all of which had a sequence length of more than 699 bp (Fig. 1). Compared with the sequenced ladybird beetle mitogenomes, we found that *Eriopis connexa*, *Harmonia quadripunctata*, *Anatis ocellata* and *Cheilomenes sexmaculata* also had a large intergenic spacer between *trnI* and *trnQ* (Fig. S1). All seven ladybird beetle species harboring the large intergenic spacer had a mitogenome size of more than 17 kb.

The secondary structures of mitochondrial tRNA genes were predicted for the newly sequenced species. The anticodons of tRNA genes were identical to those reported for *Drosophila yakuba* [56]. All tRNA genes can be folded in the form of a typical clover-leaf structure except for *trnS1* and *trnP*. The proposed secondary structures for the twenty-two tRNA genes of *A. hexaspilota* are shown in Fig. S2. As for *trnS1*, the dihydrouridine (DHU) arm was replaced with a simple loop due to unmatched base pairs. In five species (*A. hexaspilota*, *Calvia muiri*, *Coccinella septempunctata*, *Illeis cincta* and *P. japonica*), *trnP* displayed an unusual T ψ C arm, with a large T-loop motif (Figs. S2 and S3).

3.3. Data matrices

The saturation tests implied that there was no significant saturation in the subset including all codon positions and that including the first codon positions or the second codon positions or the tRNA genes, under the assumption of a symmetrical tree topology (Table 2). However, the third codon positions, the rRNA genes, and the combination of tRNA and rRNA genes exhibited significant saturation. AliGROOVE analyses of different codon positions of protein-coding gene alignments revealed the heterogeneity of sequence divergence (Fig. 2). Specifically, the alignment with only the third codon positions displayed the obviously randomized

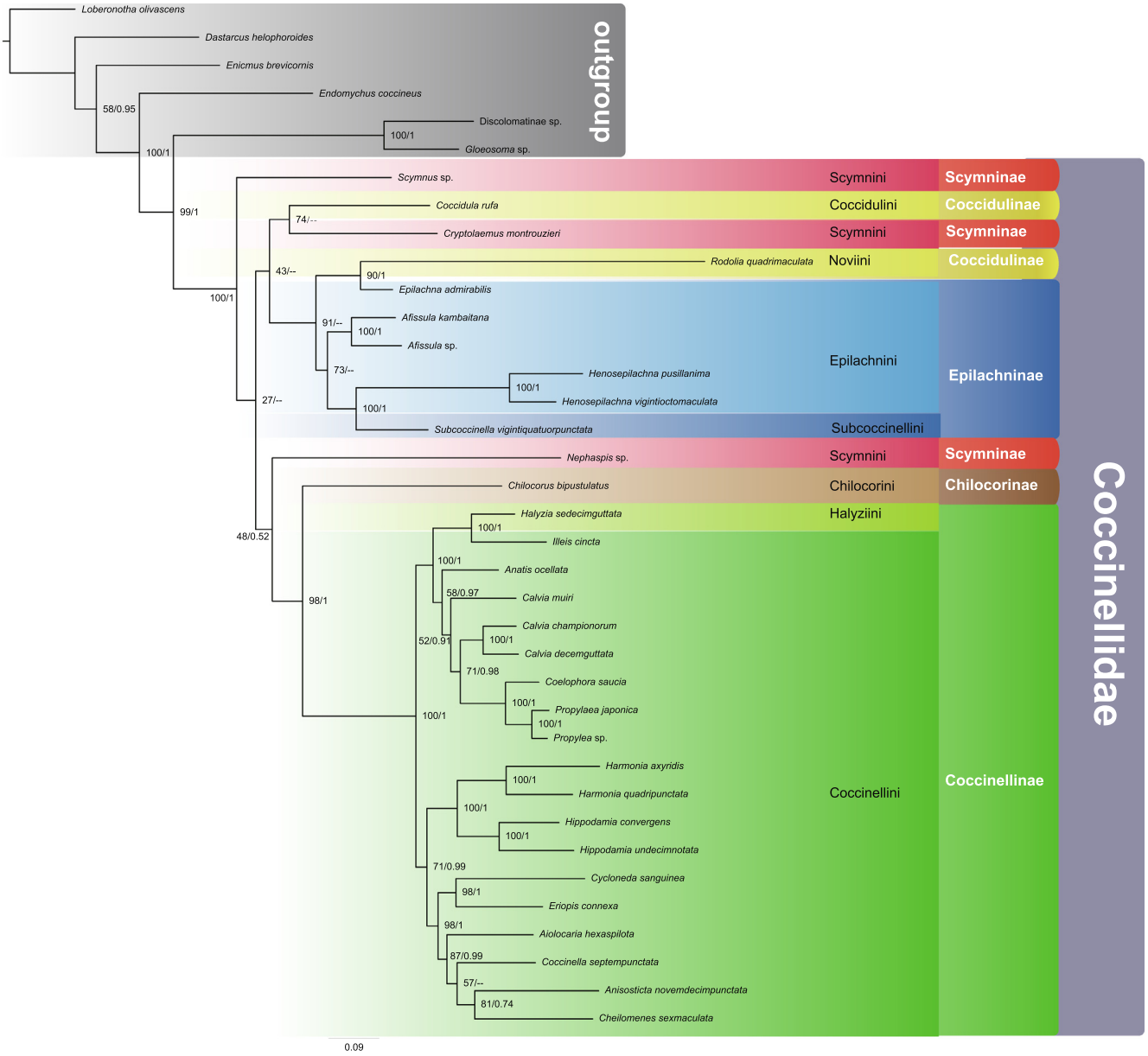

Fig. 4. Maximum likelihood tree inferred from the dataset of PCG_AA using the predefined partition schemes under the best-fitting models selected by ModelFinder. Node numbers show bootstrap support values (left) and Bayesian posterior probability support values (right). Scale bar represents substitutions/site. "–" indicates that the node is not recovered by BI analysis.

sequence similarities, with most pairwise comparisons having negative similarity scores. AliGROOVE analyses of various concatenated datasets showed low heterogeneity of sequence composition overall (Fig. S4). When compared with the nucleotide datasets, inferred amino acids from protein-coding genes reduced the effect of random sequence similarity and alignment ambiguity. Analyses of synonymous substitution (*ks*) and nonsynonymous substitution (*ka*) for protein-coding genes revealed that two outgroup species of *Endomychus coccineus* and *Gloeosoma* sp. had the highest sequence variation (*ka*/*ks*). The ingroup taxa shared the lower *ks* and *ka* values, especially in *A. ocellata*, *Calvia championorum*, *C. septempunctata* and *E. connexa* (Table 3).

3.4. Phylogenetic inference

Here, we presented results based on various datasets analyzed using a diversity of optimality criteria. Two inference methods (ML and BI) presented different reconstructions at some deep nodes. ML analyses placed a clade *Nephaspis* sp. + *Rodolia quadrimaculata* (ML-PCGRNA and ML-PCG12RNA) or the *Scymnus* sp. alone (ML-PCG_AA) as the most-basal lineage in Coccinellidae, whereas BI analyses consistently supported the *Scymnus* sp. as the earliest-diverging clade. In the Bayesian trees, relationships between subfamilies were poorly resolved due to the extremely short internal branch lengths (PP < 0.9), except for the sister group Chilocorinae + Coccinellinae. In contrast, ML analyses presented a clear hierarchical relationship among the five subfamilies (Figs. 3- 4 and Fig. S5).

All analyses, irrespective of inference method, strongly supported a monophyletic Coccinellidae (BP = 100, PP = 1.0). In the ML analyses on the nucleotides, removal of the third codon positions had no significant impact on the tree topology reconstruction (Fig. S5). Both nucleotide analyses under ML recovered a monophyletic Epilachninae. However, amino acid

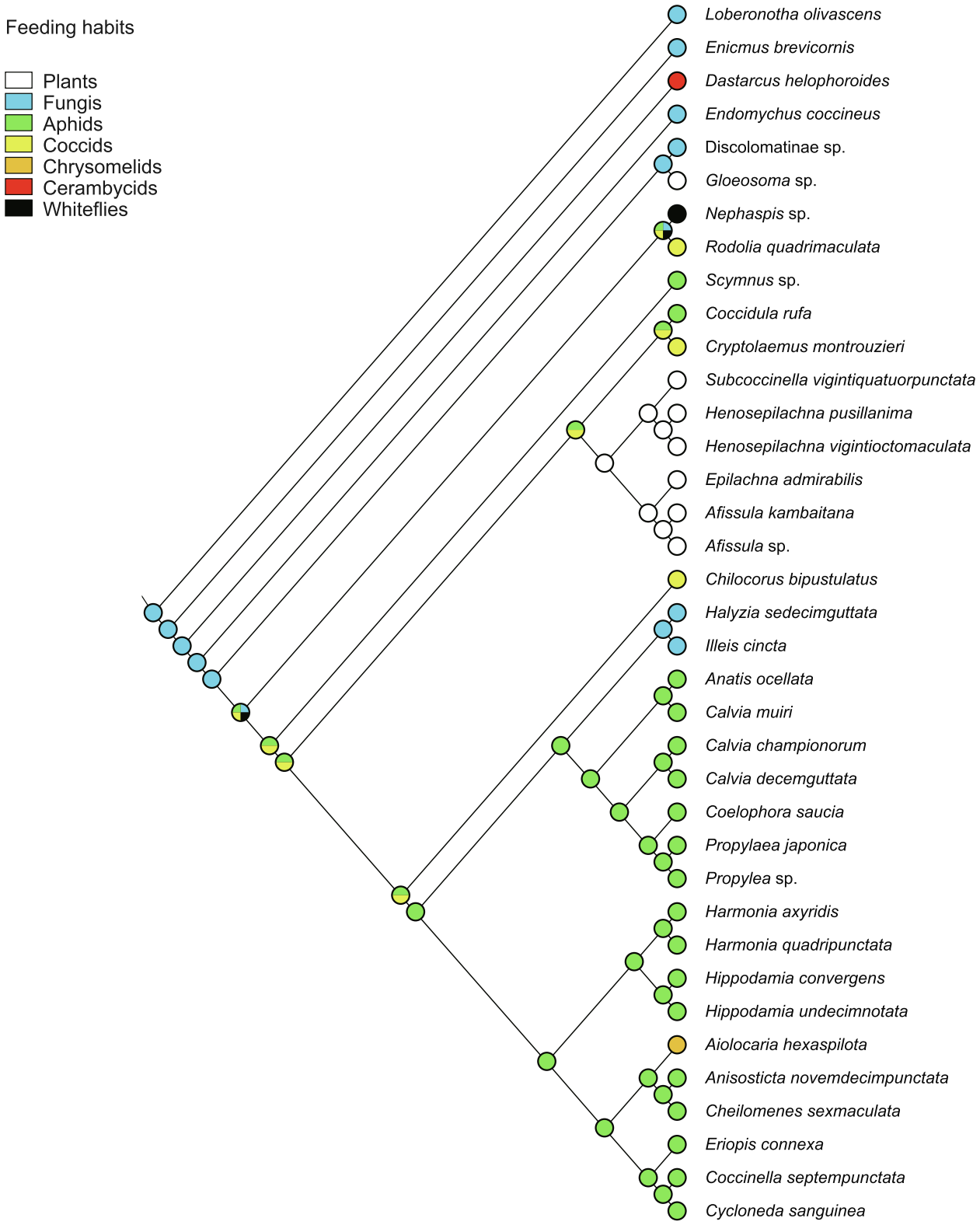

Fig. 5. Ancestral state reconstructions of food preferences based on the PCGRNA-ML tree performed under Mesquite using parsimony method. Probabilities of character states are presented at nodes with pie diagrams.

analysis supported the paraphyly of Epilachninae with respect to *R. quadrimaculata*. Other differences between nucleotide and amino acid analyses resided in the intrarelationships of the Epilachninae and of the Coccinellinae. In other respects, analyses of nucleotide datasets produced trees similar to those of the amino acid analyses. The monophyly of Coccinellinae was strongly supported (BP $\geq$ 95, PP = 1.0). The remaining subfamilies with more than two exemplars included were respectively split into two or more separated clades: Scymninae in three clades, and Coccidulinae in two clades. At the tribe level,

the Scymnini, Epilachnini and Coccinellini were resolved as non-monophyletic assemblages.

3.5. Ancestral state reconstructions of food preferences

Ancestral state reconstruction suggested that the ancestral feeding pattern for the whole Coccinellidae was ambiguous, which could either be predators of fungi, aphids, coccids or whiteflies (Fig. 5). Despite this, analysis indicated that the ladybird beetles with the predominantly predatory habit derived from within the fungus-feeding Cerylonid Series. In addition, the plant feeding ladybird beetles arose within an aphidophagous/coccidophagous clade. There have been at least three independent shifts toward coccidophagy in Coccinellidae. Fungi specialization occurred once independently.

4. Discussion

4.1. Characteristics of mitogenomes

Next generation sequencing allows effective reconstruction of full or partial mitogenomes from bulk insect samples [27,29,30,57,58]. However, the use of mitogenomes alone in the phylogenetic studies remains contentious [59–61]. This is mainly due to the limitations associated with characteristics of mitogenomes present in some specific lineages, for example, the aberrant nucleotide composition and the variable substitution rates. In this analysis, we found that several species in Coccinellini had a large intergenic spacer between *trnI* and *trnQ*. This intergenic spacer makes the length of the entire mitogenome to exceed 17 kb. The underlining causes of the large intergenic spacers in these mitogenomes are unknown. The substitution rate analyses showed that these species had a lower synonymous (*ks*) and nonsynonymous (*ka*) substitution rates (Table 3). This may lead to the poorly supported nodes comprising these species in the resulting trees (e.g., *C. muiri* and *E. connexa* in Fig. 3). Further researches are required to corroborate the potential link of the occurrence of large intergenic spacer and the relatively low substitution rate in the mitogenomes.

4.2. Phylogeny of Coccinellidae

Although the monophyly of the Coccinellidae is well supported by molecular [13–18] or morphological [7–9] data, the internal relationships remain equivocal. The monophyly of most of subfamilies within Coccinellidae could not be confirmed by previous studies [14]. Only one out of eight subfamilies has been definitely supported as monophyletic in recent molecular and/or morphological analyses [11,13,14], namely Coccinellinae. The subfamilies Sticholotidinae, Chilocorinae, Scymninae, Coccidulinae and Ortaliinae are often recognized as paraphyletic or polyphyletic [13,14]. In this analysis, the Coccinellinae was strongly supported as a monophyletic group, while the Scymninae and Coccidulinae were congruently recovered as non-monophyletic groups. These relationships are in line with the recent studies [11,13,14]. As for the monophyly of Epilachninae, previous analyses presented the conflict hypotheses. The studies of Giorgi et al. [13] and Seago et al. [11] supported Epilachninae as a monophyletic group, whereas Magro et al. [14] recovered Epilachninae as a paraphyletic group. Our nucleotide datasets returned the monophyly of Epilachninae, while amino acid analyses supported Epilachninae as paraphyletic. Thus, the issue on the monophyly of Epilachninae remains to be addressed.

This study provided the first mitogenome of Chilocorinae. A sister group relationship between Chilocorinae and Coccinellinae was consistently recovered by the current mitogenomic data, with strong nodal support. This result was also supported in the molecular study of Magro et al. [14]. Based on the analysis of morphological and multi-locus molecular data, Seago et al. [11] recovered a sister group relationship between Coccinellinae and a clade composed primarily of Chilocorinae. However, the morphology-based studies supported Epilachninae as a sister group of Coccinellinae [6–9]. Only one species representing Chilocorinae was included in this study, expanded taxonomic sampling will give more insights on the phylogenetic placement of this group.

4.3. Evolution of food relationships

Previous phylogenetic researches proposed that the Coccinellidae might arise from the mycophagous lineages of the Cerylonid Series [12,16–18]. Giorgi et al. [13] indicated a food preference transition from mycophagy to predation for the ancestors of extant Coccinellidae, and considered coccidophagy to be the ancestral condition of the whole Coccinellidae. However, the study of Magro et al. [14] did not produce a clear pattern of food preference evolution in Coccinellidae. This analysis implied that the ancestral state of food preference was uncertain, which may be predatory or mycophagous. The Coccoidea and Aphidoidea had already diversified in the Permian [62], which was often considered to be earlier than Coccinellidae. The latter might take place in the Cretaceous [15]. Therefore, the coccidophagy and aphidophagy are two alternative ancestral feeding conditions for Coccinellidae. The phytophagy observed in Epilachninae was nested within the aphidophagous and/or coccidophagous clades. This result is consistent with Giorgi et al. [13].

5. Conclusion

In this paper, we demonstrated that mitogenome sequences can provide insights into the higher-level phylogeny of Coccinellidae. The results supported Coccinellinae as a monophyletic group, and confirmed a sister group relationship between Chilocorinae and Coccinellinae. The Scymninae and Coccidulinae were resolved of non-monophyletic origin. The monophyly of Epilachninae was equivocal based on the current data, because nucleotide and amino acid analyses yielded conflict results. Phylogenetic reconstruction from mitogenomic data served as a fundamental step toward interpreting the evolution of feeding patterns. In future study, increasing taxonomic sample will contribute to determining the phylogenetic relationships and the evolutionary pattern of food preference in Coccinellidae.

Acknowledgements

This research is supported by grants from the Earmarked Fund for China Agriculture Research System (NO. CARS-27), the Science and Technology Innovation Fund of Henan Agricultural University (KJCX2019A10), Key Scientific Research Projects of Henan Province (Grant Nos. 14B210036, 16A210029, 18B210006), Science and Technology Innovation Fund of Henan Agricultural University (KJCX2019A10), Study on Technology of Chemical Fertilizer Reduction and Synergism in Tea Garden (2016YFD0200901), and Fruit Industry System of Henan Province (S2014-11-G03).

Declaration of Competing Interest

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2019.10.089.

References

[1] R.H. Arnett, J.H. Frank, M.C. Thomas, P.E. Skelley, American Beetles, Volume II: Polyphaga: Scarabaeoidea through Curculionoidea, CRC Press, NY, 2002.
[2] I. Hodek, A. Honěk, Scale insects, mealybugs, whiteflies and psyllids (Hemiptera, Sternorrhyncha) as prey of ladybirds, Biol. Control. 51 (2009) 232–243.
[3] J.J. Obrycki, J.D. Harwood, T.J. Kring, R.J. O'Neil, Aphidophagy by Coccinellidae: application of in agroecosystems, Biol. Control. 51 (2009) 244–254.
[4] L. Redtenbacher, Tetamen dispositionis generum et specierum Coleopterorum Pseudotrimeorum, Archiducatus Austriae (1844).
[5] R. Korschefsky, Coleopterum Catalogus, pars 118 Coccinellidae I, Schenklink, W Junk, 1931, p. 124.
[6] H. Sasaji, Phylogeny of the family Coccinellidae (Coleoptera), Occasional Publ. Biol. Lab. 35 (1968).
[7] H. Sasaji, Fauna Japonica, Coccinellidae (Insecta, Coleoptera), Academic Press of Japan Keigaku Publishing, Tokyo, 1971.
[8] H. Sasaji, Phylogenetic positions of some remarkable genera of the Coccinellidae (Coleoptera), with an attempt of the numerical methods, Mem. Faculty Educ. Fukui Univ. Ser. II (Nat. Sci.) 21 (1971) 55–73.
[9] I. Kovář, Phylogeny, in: Ecology of Coccinellidae, Kluwer Academic Publishers, Dordrecht, 1996, pp. 19–31.
[10] G. Yu, Cladistic analyses of the Coccinellidae (Coleoptera), Insect Sci. 1 (1994) 17–30.
[11] A.E. Seago, J.A. Giorgi, J. Li, A. Ślipiński, Phylogeny, classification and evolution of ladybird beetles (Coleoptera: Coccinellidae) based on simultaneous analysis of molecular and morphological data, Mol. Phylogenet. Evol. 60 (2011) 137–151.
[12] H.E. Escalona, A. Ślipiński, Generic revision and phylogeny of Microweiseinae (Coleoptera: Coccinellidae), Syst. Entomol. 37 (2012) 125–171.
[13] J.A. Giorgi, N.J. Vandenberg, J.V. McHugh, J.A. Forrester, S.A. Ślipiński, K.B. Miller, L.R. Shapiro, M.F. Whiting, The evolution of food preferences in Coccinellidae, Biol. Control. 51 (2009) 215–231.
[14] A. Magro, E. Lecompte, F. Magné, J.L. Hemptinne, B. Crouau-Roy, Phylogeny of ladybirds (Coleoptera: Coccinellidae): are the subfamilies monophyletic, Mol. Phylogenet. Evol. 54 (2010) 833–848.
[15] T. Hunt, J. Bergsten, Z. Levkanicova, A. Papadopoulou, O.St. John, R. Wild, P.M. Hammond, D. Ahrens, M. Balke, M.S. Caterino, J. Gómez-Zurita, I. Ribera, T.G. Barraclough, M. Bocakova, L. Bocak, A.P. Vogler, A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation, Science 318 (2007) 1913–1916.
[16] J.A. Robertson, M.F. Whiting, J.V. McHugh, Searching for natural lineages within the Cerylonid Series (Coleoptera: Cucujoidea), Mol. Phylogenet. Evol. 46 (2008) 193–205.
[17] L. Bocak, C. Barton, A.L.E.X. Crampton-Platt, D. Chesters, D. Ahrens, A.P. Vogler, Building the Coleoptera tree of life for > 8000 species: composition of public DNA data and fit with Linnaean classification, Syst. Entomol. 39 (2014) 97–110.
[18] D.D. Mckenna, A.L. Wild, K. Kanda, C.L. Bellamy, R.G. Beutel, M.S. Caterino, C.W. Farnum, D.C. Hawks, M.A. Ivie, M. Liz Jameson, R.A.B. Leschen, A.E. Marvaldi, J. V. Mchugh, A.F. Newton, J.A. Robertson, M.K. Thayer, M.F. Whiting, J.F. Lawrenve, A. Ślipiński, D.R. Maddison, B.D. Farrell, The beetle tree of life reveals that Coleoptera survived end-Permian mass extinction to diversify during the Cretaceous terrestrial revolution, Syst. Entomol. 40 (2015) 835–880.
[19] J.A. Robertson, A. Ślipiński, M. Moulton, F.W. Shockley, A. Giorgi, N.P. Lord, D.D. Mckenna, W. Tomaszewska, J. Forrester, K.B. Miller, M.F. Whiting, J.V. Mchugh, Phylogeny and classification of Cucujoidea and the recognition of a new superfamily Coccinelloidea (Coleoptera: Cucujiformia), Syst. Entomol. 40 (2015) 745–778.
[20] J.J. Sloggett, M. Majerus, Habitat preferences and diet in the predatory Coccinellidae (Coleoptera): an evolutionary perspective, Biol. J. Linn. Soc. 70 (2000) 63–88.
[21] I. Hodek, A. Honěk, Ecology of Coccinellidae, Kluwer, Dordrecht, Springer, Netherlands 10 (1996) 978–94.
[22] A. Ślipiński, Australian ladybird beetles (Coleoptera: Coccinellidae): their biology and classification, Australian Biological Resources Study Canberra (2007).
[23] B.Y. Seo, J. Park, W. Kwon, J. Park, The complete mitochondrial genome of Aiolocaria hexaspilota (Hope, 1831) (Coleoptera: Coccinellidae), Mitochondrial DNA B 4 (2019) 1472–1474.
[24] I. Hodek, Food relations, in: I. Hodek, A. Honěk (Eds.), Ecology of Coccinellidae, Kluwer Academic Publishers, Dordrecht, 1996, pp. 143–238.
[25] J.L. Hemptinne, A.F.G. Dixon, J. Coffin, Attack strategy of ladybird beetles (Coccinellidae): factors shaping their numerical response, Oecologia 90 (1992) 238–245.
[26] B.S. Coates, Assembly and annotation of full mitochondrial genomes for the corn rootworm species, *Diabrotica virgifera virgifera* and *Diabrotica barberi* (Insecta: Coleoptera: Chrysomelidae), using Next Generation Sequence data, Gene 542 (2014) 190–197.
[27] A. Crampton-Platt, M.J. Timmermans, M.L. Gimmel, S.N. Kutty, T.D. Cockerill, C. V. Khen, A.P. Vogler, Soup to tree: the phylogeny of beetles inferred by mitochondrial metagenomics of a Bornean rainforest sample, Mol. Biol. Evol. 32 (2015) 2302–2316.
[28] C. Gómez-Rodríguez, A. Crampton-Platt, M.J. Timmermans, A. Baselga, A.P. Vogler, Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages, Methods Ecol. Evol. 6 (2015) 883–894.
[29] M.J.T.N. Timmermans, C. Barton, J. Haran, D. Ahrens, C. Lorna Culverwell, A. Ollikainen, S. Dodsworth, P.G. Foster, L. Bocak, A.P. Vogler, Family-level sampling of mitochondrial genomes in Coleoptera: compositional heterogeneity and phylogenetics, Genome Biol. Evol. 8 (2016) 161–175.
[30] R.E. Nie, T. Breeschoten, M.J.T.N. Timmermans, K. Nadein, H.J. Xue, M. Bai, Y. Huang, X.K. Yang, A.P. Vogler, The phylogeny of Galerucinae (Coleoptera: Chrysomelidae) and the performance of mitochondrial genomes in phylogenetic inference compared to nuclear rRNA genes, Cladistics 34 (2018) 113–130.
[31] S.X. Ren, X.M. Wang, H. Pang, Z.Q. Peng, T. Zeng, Colored Pictorial Handbook of Ladybird Beetles in China, Science Press, Beijing, 2009.
[32] G. Yu, Chinese Lady Beetles (The Subfamily Coccinellinae), Published by Mechanical Industry Press, 2010.
[33] N. Song, H. Li, F. Song, W. Cai, Molecular phylogeny of Polyneoptera (Insecta) inferred from expanded mitogenomic data, Sci. Rep. 6 (2016) 36175.
[34] R.K. Patel, M. Jain, NGS QC Toolkit: a toolkit for quality control of next generation sequencing data, PloS ONE 7 (2012) e30619.
[35] Y. Peng, H.C.M. Leung, S.M. Yiu, F.Y.L. Chin, IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth, Bioinformatics 28 (2012) 1420–1428.
[36] T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucleic Acids Symp. Ser. 41 (1999) 95–98.
[37] M. Bernt, A. Donath, F. Juhling, F. Externbrink, C. Florentz, G. Fritzsch, J. Putz, M. Middendorf, P.F. Stadler, MITOS: Improved de novo metazoan mitochondrial genome annotation, Mol. Phylogenet. Evol. 69 (2013) 313–319.
[38] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics 25 (2009) 1754–1760.
[39] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools, Bioinformatics 25 (2009) 2078–2079.
[40] K. Okonechnikov, A. Conesa, F. García-Alcalde, Qualimap 2: advanced multisample quality control for high-throughput sequencing data, Bioinformatics 32 (2016) 292–294.
[41] F. Abascal, R. Zardoya, M.J. Telford, TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations, Nucleic Acids Res. 38 (2010) W7–W13.
[42] K. Katoh, D.M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability, Mol. Biol. Evol. 30 (2013) 772–780.
[43] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0, Mol. Biol. Evol. 30 (2013) 2725–2729.
[44] J. Castresana, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis, Mol. Biol. Evol. 17 (2000) 540–552.
[45] P. Kuck, K. Meusemann, FASconCAT: convenient handling of data matrices, Mol. Phylogenet. Evol. 56 (2010) 1115–1118.
[46] X.H. Xia, DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution, Mol. Biol. Evol. 30 (2013) 1720–1728.
[47] P. Librado, J. Rozas, DnaSP v5: a software for comprehensive analysis of DNA polymorphism data, Bioinformatics 25 (2009) 1451–1452.
[48] P. Kuck, S.A. Meid, C. Groß, J.W. Wägele, B. Misof, AliGROOVE - visualization of heterogeneous sequence divergence within multiple sequence alignments and detection of inflated branch support, BMC Bioinformatics 15 (2014).
[49] L.T. Nguyen, H.A. Schmidt, A. von Haeseler, B.Q. Minh, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies, Mol. Biol. Evol. 32 (2015) 268–274.
[50] S. Kalyaanamoorthy, B.Q. Minh, T.K.F. Wong, A. von Haeseler, L.S. Jermiin, ModelFinder: fast model selection for accurate phylogenetic estimates, Nat. Methods 14 (2017) 587–589.
[51] N. Lartillot, H. Philippe, A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process, Mol. Biol. Evol. 21 (2004) 1095–1109.
[52] N. Lartillot, T. Lepage, S. Blanquart, PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating, Bioinformatics 25 (2009) 2286–2288.
[53] M. Miller, W. Pfeiffer, T. Schwartz, Creating the CIPRES science gateway for inference of large phylogenetic trees, gateway computing environments workshop (GCE) 14 (2010) 1–8.
[54] W.P. Maddison, D.R. Maddison, Mesquite: a modular system for evolutionary analysis. Version 3.31 (2017).
[55] S.L. Cameron, Insect mitochondrial genomics: implications for evolution and phylogeny, Annu. Rev. Entomol. 59 (2014) 95–117.
[56] D.O. Clary, D.R. Wolstenholme, The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code, J. Mol. Evol. 22 (1985) 252–271.

[57] C.P. Gillett, A. Crampton-Platt, M.J. Timmermans, B.H. Jordal, B.C. Emerson, A.P. Vogler, Bulk de novo mitogenome assembly from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionoidea), Mol. Biol. Evol. 31 (2014) 2223–2237.
[58] M. Tang, M. Tan, G. Meng, S. Yang, X. Su, S. Liu, W. Song, Y. Li, Q. Wu, A. Zhang, X. Zhou, Multiplex sequencing of pooled mitochondrial genomes a crucial step toward biodiversity analysis using mito-metagenomics, Nucleic. Acids. Res. 42 (2014) e166.
[59] G. Talavera, R. Vila, What is the phylogenetic signal limit from mitogenomes? The reconciliation between mitochondrial and nuclear data in the Insecta class phylogeny, BMC Evol. Biol. 11 (2011) 315.
[60] M. Bernt, C. Bleidorn, A. Braband, J. Dambach, A. Donath, G. Fritzsch, A. Golombek, H. Hadrys, F. Jühling, K. Meusemann, M. Middendorf, B. Misof, M. Perseke, L. Podsiadlowski, B. von Reumont, B. Schierwater, M. Schlegel, M. Schrödl, S. Simon, P.F. Stadler, I. Stögerk, T.H. Struck, A comprehensive analysis of bilaterian mitochondrial genomes and phylogeny, Mol. Phylogenet. Evol. 69 (2013) 352–364.
[61] S. Simon, H. Hadrys, A comparative analysis of complete mitochondrial genomes among Hexapoda, Mol. Phylogenet. Evol. 69 (2013) 393–403.
[62] C.D.V. Dohlen, N.A. Moran, Molecular phylogeny of the Homoptera: a paraphyletic taxon, J. Mol. Evol. 41 (1995) 211–223.