Received: 27 December 2024 | Accepted: 22 April 2025

DOI: 10.1111/syen.12684

ORIGINAL ARTICLE

Systematic Entomology

Inconsistent performance of multi-type genomic data in phylogenomics of neuropteridan insects, with solutions toward conflicting results

Ruyue Zhang[1] | Liming Wang[1] | Shuo Tian[1] | Yang Liu[1] | Yunlan Jiang[1] | Xiaofan Zhou[2] | Ding Yang[3] | Xingyue Liu[4] | Yuyu Wang[1]

[1]College of Plant Protection, Hebei Agricultural University, Baoding, China

[2]College of Plant Protection, South China Agricultural University, Guangzhou, China

[3]Department of Entomology, China Agricultural University, Beijing, China

[4]Institute of Zoology, Chinese Academy of Sciences, Beijing, China

Correspondence

Ding Yang, Department of Entomology, China Agricultural University, Beijing 100193, China.
Email: dyangcau@126.com

Xingyue Liu, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China.
Email: xingyue_liu@yahoo.com

Yuyu Wang, College of Plant Protection, Hebei Agricultural University, Baoding 071001, China.
Email: wangyy_amy@126.com

Funding information

National Natural Science Foundation of China, Grant/Award Numbers: 32170451, 32130012, 32170448; Basic Research Project of Shijiazhuang for University in Hebei Province, Grant/Award Number: 241791137A; Science and Technology Planning Project of Baoding, Grant/Award Number: 2472P018; Basic Scientific Research Business Expenses of Hebei Provincial Universities, Grant/Award Number: KY2023061; Institute Go-Sailing Program (ShengKeChuang-QA); Earmarked Fund for CARS-27

Editor: Bonnie Blaimer

Abstract

Reconstructing the tree of life is facing challenges in inferring accurate and robust phylogeny based on large data in the genomic era. Currently, universal single-copy orthologs (USCOs), ultraconserved elements (UCEs) and mitochondrial genomes (mitogenomes) are widely used to reconstruct phylogeny. In this study, the higher-level phylogeny of lacewings and allied orders (Neuropterida) is reconstructed based on USCOs, UCEs and mitogenomes assembled from 42 newly sequenced low-coverage genomes (above 32.80X), representing all orders and all families except Rhachiberothidae, under various types of data filtering, model selection and strategies of tree reconstruction. Using relatively conservative criteria, we demonstrate that the topology based on amino acid matrices of the USCOs filtered by multifactorial strategies under the site heterogeneity model (LG + PMSF (C20)) is the most robust. The average bootstrap support (ABS) values, an important criterion in gene filtering, exhibit large variation among different repetitions. Applying fossil calibrations at deeper nodes close to the root of the phylogeny is demonstrated to facilitate more accurate estimation of evolutionary timescales by comparing three different calibration schemes (deeper nodes, shallower nodes and a combination of both). These results highlight the complexity of genomic data and offer an integrative solution to overcome systematic error in phylogenomic inference.

KEYWORDS

divergence time, low-coverage genome, mitochondrial genome, Neuropterida, phylogeny

INTRODUCTION

The rapid development of next-generation sequencing (NGS) technologies has facilitated a paradigm shift in phylogenetic and evolutionary studies in the inference of the tree of life (Behura, 2015; Boutte et al., 2022; Chan & Ragan, 2013; Mbanyana et al., 2024; Sinaiko et al., 2024; Wachi et al., 2017; Young & Gillung, 2019). Currently, universal single-copy orthologs (USCOs), ultraconserved elements (UCEs) and mitochondrial genomes (mitogenomes) are frequently used in comparative phylogenomics (Allen et al., 2017; Dai et al., 2023; Yu et al., 2022; Zhang et al., 2019; Zhang et al., 2022; Zhao et al., 2023). However, it is widely acknowledged within the phylogenomics

community that these types of genomic data may lead to conflicts in tree topology, especially at the deep nodes (Shen et al., 2017; Simion et al., 2017; Steenwyk et al., 2023; Whelan et al., 2017; Yu et al., 2022). Factors responsible for such conflicts involve sampling strategies, systematic error (e.g. compositional heterogeneity and rate heterogeneity) (Cai et al., 2020; Foster, 2004; Kapli et al., 2021; Kocot et al., 2017; Li et al., 2023; Philippe et al., 2011; Szantho et al., 2023) and biological factors (e.g. incomplete lineage sorting (ILS), horizontal gene transfer (HGT), etc.) (Edwards, 2009; Liu et al., 2015; Liu et al., 2024; Steenwyk et al., 2023; Young & Gillung, 2019). To address systematic error in phylogenetic inference, gene filtering (Fernandez et al., 2016; Moles et al., 2021; Parks et al., 2018; Shen, Salichos, & Rokas, 2016), more appropriate evolutionary models (Cai et al., 2023; Dong et al., 2022; Lartillot et al., 2007; Lartillot & Philippe, 2004; Rinke et al., 2021; Wang et al., 2018; Yu et al., 2022) and recoding strategies (Giacomelli et al., 2022; Hernandez & Ryan, 2021; Kosiol et al., 2004; Simion et al., 2017; Susko & Roger, 2007) are often applied. Moreover, the coalescent-based method accounts for the presence of ILS in individual gene trees, allowing for different evolutionary histories among genes, as well as phylogenetic conflicts between gene trees (Edwards, 2009; Liu et al., 2015; Steenwyk et al., 2023; Young & Gillung, 2019). However, various filtering methods are not always effective in reducing inconsistencies, and the best method, if any, is not clear.

Significant advancements have been made in methodologies for estimating evolutionary rates and divergence times, and how to improve the reliability of molecular clocks has been an important issue in fields such as evolutionary biology and biogeography (Duchene et al., 2014; Kovacs et al., 2024; Li et al., 2019; Luo et al., 2023; Marshall, 2008; Püschel et al., 2021). Node-dating approaches use fossil data indirectly to inform calibration prior densities on internal node ages (Drummond et al., 2006; Nguyen & Ho, 2020; Yang & Rannala, 2006). Numerous studies have been conducted on the effects of substitution models (Abadi et al., 2019; Tao et al., 2020), molecular clock models (Ho & Duchene, 2014; Lepage et al., 2007) and fossil sampling (Inoue et al., 2010; O'Reilly & Donoghue, 2020; Tamura et al., 2018). Regardless, there are few comparative analyses of divergence time estimations utilizing diverse molecular datasets under strictly controlled sampling conditions (Zhang et al., 2025). In addition, empirical evidence indicates that divergence time estimates are sensitive to the location and number of fossil calibrations (Duchene et al., 2014; Püschel et al., 2019; Villaverde et al., 2021). How exactly different datasets and various fossil choices affect these results deserves further exploration.

Neuropterida are one of the key groups of Holometabola, with moderate species diversity (ca. 6600 extant species worldwide) with an archaic evolutionary history (dating back to the Carboniferous) and remarkable morphological and ecological diversification. Among the orders of Neuropterida, that is, Raphidioptera, Megaloptera and Neuroptera, the sister group relationship between Megaloptera and Neuroptera has been widely accepted based on a series of phylogenetic studies using morphological and genomic data. However, the relationships among some families of Neuroptera remain controversial, particularly deeper relationships among groups of families (Beutel et al., 2010; Lu et al., 2021; Vasilikopoulos et al., 2020; Winterton et al., 2010; Winterton et al., 2017; Winterton et al., 2018). This issue hinders the reliable estimation of the time scale for the diversification of Neuropterida and further affects studies on the evolution of this group in various aspects.

Herein, we present a case study on the higher-level phylogeny of Neuropterida to explore conflict in phylogenetic inferences based on the USCOs, UCEs and mitogenomes, assembled from low-coverage genomes (mostly newly sequenced). This study aims to: (1) assess the performance of the three data types in reconstructing the higher phylogeny of Neuropterida, (2) systematically compare multiple analytical strategies for reconstruction and (3) investigate the performance of three different fossil calibration schemes by data type in divergence time estimation.

MATERIALS AND METHODS

Taxon sampling, DNA extraction and sequencing

We sampled 57 species in this study from all orders and most families of Neuropterida (except Rhachiberothidae). Specimens were preserved in absolute ethanol at −20°C in Hebei Agricultural University. The low-coverage genomes of 42 species are newly sequenced, while 11 published genomes of Neuropterida were downloaded from GenBank (http://www.ncbi.nlm.nih.gov) (Table S1). Four species of Coleoptera were selected as outgroups: *Agrilus cyanescens* (Ratzeburg) (Polyphaga; Buprestidae), *Agelastica alni* (Linnaeus) (Polyphaga: Chrysomelidae), *Anaspis maculata* Geoffroy (Polyphaga: Scraptiidae) and *Ophonus ardosiacus* (Lutshnik) (Adephaga: Carabidae).

Total DNA was extracted from the thoracic muscle tissues of specimens using TIANamp Genomic DNA Kit (Tiangen) according to the standard protocol. The purity, concentration and integrity of genomic DNA were detected by NanoDrop 2000, Qubit fluorometer and Agilent 4200 Bioanalyzer. An Illumina paired-end library with a single species was generated from the pooled genomic DNA with an average insert size of 300 bp by the Novogene Bioinformatics Technology Co., Ltd. (Beijing, China). The clean data was generated by filtering raw data with Fastp v0.23.4 (Chen et al., 2018) to remove adapters, reads containing poly-N and low-quality reads.

Genome assembly and marker acquisition

Genome size was assessed through k-mer frequency distribution analysis. Jellyfish v1.1.10 (Marcais & Kingsford, 2011) was used to tackle k-mer counting. The genome size was estimated using GenomeScope v2.0.1 (Vurture et al., 2017). Newly sequenced genomes were assembled using PLWS v1.0.5 (http://github.com/xtmtd/PLWS) (Zhang et al., 2019). Quality control and normalization were conducted using BBTools v37.93 (Bushnell, 2014). Sequencing errors were corrected using Lighter v1.1.1 (Song et al., 2014). Contig assemblies were

performed using Minia v3.00-alpha1 with multiple k-mer (21, 41, 61, 81, 101, 121) (Chikhi & Rizk, 2013). Redundant contigs were removed using Redundans v0.13c (Pryszcz & Gabaldón, 2016). Contig scaffolding using base N was performed using BESST v2.2.8 (Sahlin et al., 2014). Gap filling (i.e. base N replaced by A\T\G\C) was performed using GapCloser v1.12 (Xu et al., 2020). The completeness of the assemblies was assessed using BUSCO v3.0.2 (Waterhouse et al., 2018) against the endopterygota_odb10 ($n = 2124$) database following standard parameters.

The USCOs were extracted from genomes with BUSCO v3.0.2 (Waterhouse et al., 2018) using the endopterygota_odb10 as reference. The default standard deviations (σ) of the mean USCO length were modified to 2σ to allow more USCOs to be classified as 'complete', as incomplete 'fragmented' USCO sequences were not output by the BUSCO pipeline.

The UCE probes of Neuropterida have been designed in accordance with Script3 of PLWS v1.0.5 (Zhang et al., 2019). Nine representative species of Neuropterida and one representative species of Coleoptera (*A. maculata*) with high-quality genomes were selected for the design of UCE probes (Table S2). The genome assembly of *Chrysoperla carnea* (Stephens) was selected as the base genome (accession GCA_905475395.1). ART v2.5.1 (Huang et al., 2012) was used to resample error-free, paired-end reads of 100 bp terminal at a coverage of 2X. Stampy v1.0.32 (Lunter & Goodson, 2011) was used to align the simulated original sequence to the base genome with a substitution rate of 0.05. The unmapped reads were deleted using SAMtools v1.19.2 (Li et al., 2010). Bedtools v2.31.0 (Quinlan & Hall, 2010) was used to merge overlapping or nearly overlapping intervals. The probes were aligned to genomic sequences to ensure that these conserved sequences were located in each genome. PHYLUCE v1.5.0 (Faircloth, 2016) was used for UCE probe design. The final probe set was designed by selecting the conserved genes shared by the eight species and removing the duplicates. For UCE loci extraction, the initial input files were all 57 genome assemblies and the newly created probe set of Neuropterida ('Neuropterida-5Kv1', $n = 5341$). PHYLUCE v1.5.0 (Faircloth, 2016) was also used to obtain UCEs from assembled genomes.

Mitochondrial genome assembly and annotation

Approximately 5 Gb high-quality reads were extracted randomly using BBTools v37.93 (reformat.sh) (Bushnell, 2014) from the clean data for each species. The mitogenome was assembled using NOVOPlasty v4.3.1 (Dierckxsens et al., 2017) and annotated by MITOS Web Server (Bernt et al., 2013) under default settings and then checked by manual proofreading according to relative species. ClustalX v1.83 (Thompson et al., 2003) was used to compare protein-coding genes (PCGs) and rRNA genes with published Neuropterida mitogenome sequences to determine accurate boundaries. The control region was identified by the boundary of the *rrnS* and *tRNA*^{Ile}.

Matrix generation

Amino acid sequences of USCOs and nucleotide sequences of UCEs were used for subsequent analyses. MAFFT v7.182 (Rozewicki et al., 2019) was used to align all genes (L-INS-i). To keep phylogenetically informative sites and remove others, alignments were trimmed by ClipKIT v2.2.1 (Steenwyk et al., 2020) with the strict parameter. The matrix concatenation and statistics have been conducted using FASConCAT-g v1.04 (Kück & Meusemann, 2010). Subsequently, all USCOs and UCEs were concatenated into USCO_all and UCE_all matrices, respectively. In this study, there were five successive strategies for filtering loci (Figure 1). For Strategy I, genes were filtered based on the average bootstrap support (ABS) values of the gene trees. Individual gene trees of each gene were estimated with IQ-TREE v1.6.10 (Minh et al., 2020) with the automatically selected best model. Eight sub-datasets were established for comprising genes whose maximum likelihood (ML) trees had ABS values across all internodes greater than or equal to 70, 75, 80 and 85 (USCO_I_abs70, USCO_I_abs75, USCO_I_abs80, USCO_I_abs85, UCE_I_abs70, UCE_I_abs75, UCE_I_abs80 and UCE_I_abs85). For strategy II, ten matrices (USCO_II_70, USCO_II_75, USCO_II_80, USCO_II_85, USCO_II_90, UCE_II_70, UCE_II_75, UCE_II_80, UCE_II_85 and UCE_II_90) were constructed with 70%, 75%, 80%, 85% and 90% taxa completeness for downstream analyses to reduce the proportion of missing data. For Strategy III, eight matrices (USCO_III_70_abs70, USCO_III_75_abs75, USCO_III_80_abs80, USCO_III_85_abs85, UCE_III_70_abs70, UCE_III_75_abs75, UCE_III_80_abs80 and UCE_III_85_abs85) were generated under the combined filtering conditions of Strategy I and Strategy II. For Strategy IV, the number of parsimony-informative sites per gene was detected using Phykit v1.2.1 (Steenwyk et al., 2021). The loci with the number of parsimony-informative sites less than 100 were removed. BaCoCa v1.1 (Kück & Struck, 2014) was used to assess heterogeneity using relative composition frequency variability (RCFV) values. The loci with RCFV values greater than 0.3 (RCFV>0.3) were removed. Then, the loci with SRH (stationary, reversible and homogeneous) (Naser-Khdour et al., 2019) model violations were filtered using IQ-TREE v1.6.10 (Minh et al., 2020). Subsequently, the remaining USCOs and UCEs were concatenated into USCO_IV and UCE_IV matrices, respectively. For Strategy V, the filtering methods based on the gene tree were added to Strategy IV. Firstly, Treeshrink v1.3.7 (Mai & Mirarab, 2018) was used to remove genes with abnormally long branches using multiple α thresholds (0.05). The degree of violation of a molecular clock (DVMC) and treeness of genes was calculated using Phykit v1.2.1 (Steenwyk et al., 2021). The genes with DVMC>1 or treeness<0.3 were removed. Subsequently, matrices (USCO_V and UCE_V) were concatenated from the remaining USCOs and UCEs, respectively.

For mitogenomes, each PCG was aligned individually using MAFFT v7.182 (Rozewicki et al., 2019) based on codon-based multiple alignments with the L-INS-i strategy. Two rRNA genes were individually aligned using MAFFT v7.182 (Rozewicki et al., 2019) with the G-INS-i strategy. The unaligned sites were excluded using Gblocks v

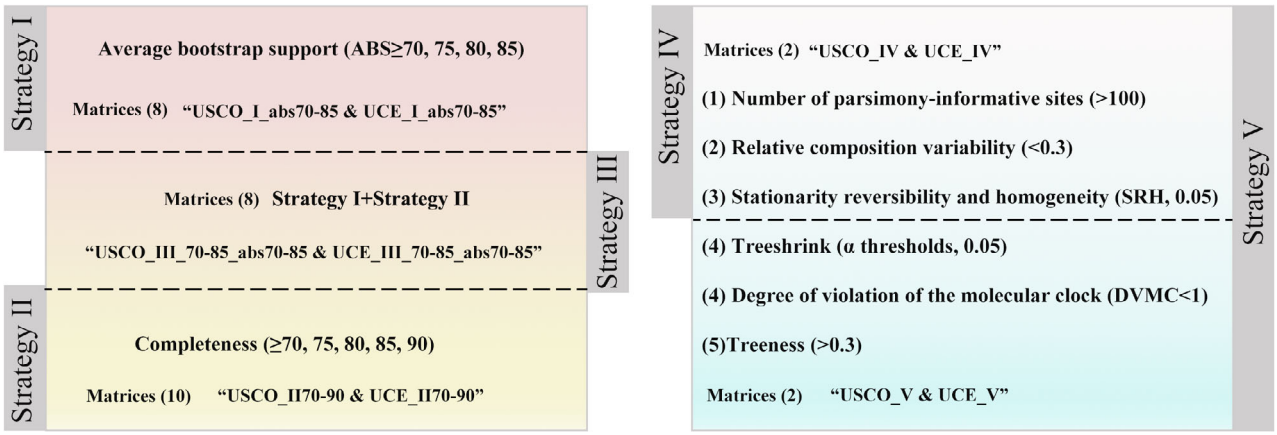

FIGURE 1 Five filtering strategies and corresponding matrices (Strategy I-V).

0.91b (Talavera & Castresana, 2007) with default parameters. Then alignments of individual mitogenome genes were concatenated to generate five matrices. The PCG matrix included all three codon positions of 13 PCGs. The PCGR matrix included all three codon positions of 13 PCGs and two rRNA genes. The PCG12 matrix included the first and second codon positions of 13 PCGs. The PCG12R matrix included the first and second codon positions of 13 PCGs and two rRNA genes. The AA matrix included amino acid sequences of 13 PCGs.

Phylogenetic reconstruction and topology testing

For the USCOs and UCEs matrices, phylogenetic trees were inferred based on concatenation and coalescent-based analyses. The concatenated ML analyses were inferred using IQ-TREE v1.6.10 (Minh et al., 2020) with the optimization model automatically selected for each gene (−m MFP) for all matrices. Dayhoff6 recoding (Hernandez & Ryan, 2021) corresponding to the six-dayhoff's amino acid families was first performed using Phylogears v2.2.0 (Tanabe, 2008), and then reconstructed using IQ-TREE v1.6.10 (Minh et al., 2020) with '-m GTR + R' based on all USCO matrices. The phylogenetic trees were reconstructed under posterior mean site frequency (PMSF) model with '-m LG + C20 + FO + R' using IQ-TREE v1.6.10 (Minh et al., 2020) for all USCO matrices. The corresponding concatenated partitioning ML tree was used as the initial guide tree. The resulting PMSF tree was used as a new guide tree, resulting in a new PMSF profile and a new tree. The PMSF tree inference process was repeated iteratively at least twice until no topological changes occurred. ML trees for individual genes were inferred using IQ-TREE v1.6.10 (Minh et al., 2020) with the optimization model automatically selected (−m MFP). The coalescent-based analyses were generated by individual gene trees as mentioned above using 100 replicates in ASTRAL-III v5.6.1 (Zhang et al., 2018).

For mitogenomes, phylogenies were estimated under the site-homogeneous model and the site-heterogeneous model. For the concatenated ML analyses of all mitogenome matrices, partitioning schemes and substitution models were selected automatically using IQ-TREE v1.6.10 (Minh et al., 2020) for 1000 UFBoot2 bootstraps (Hoang et al., 2018) and 1000 SH-aLRT replicates (Guindon et al., 2010). Bayesian inference (BI) was performed using MrBayes v3.1.2 (Ronquist & Huelsenbeck, 2003) with four Markov Chain Monte Carlo (MCMC) chains for 10 million generations, sampling every 1000 trees and discarding the first 25% of generations as burn-in for all mitogenome matrices. A phylogenetic tree based on the PCGR matrix under the site-heterogeneous model (CAT-GTR) was also reconstructed with PhyloBayes v3.3 (Lartillot et al., 2009). Two independent MCMC chains were run at the same time for separate analysis until the maxdiff was less than 0.1, and the consensus tree was obtained by discarding 25% of the samples as burn-in. Individual gene trees were obtained with IQ-TREE v1.6.10 (Minh et al., 2020) with '-m MFP'.

Then, topology testing was performed in IQ-TREE v1.6.10 (Minh et al., 2020) to select the most robust topology for USCOs and UCEs matrices, respectively. Approximately Unbiased (AU), Weighted Kishino-Hasegawa (WKH) and Weighted Shimodaira-Hazegawa (WSH) tests were calculated in IQ-TREE v1.6.10 (Minh et al., 2020). The tree estimated from previous PMSF inference on the USCO_V matrix was used as the guide tree, using the USCO_V matrix and 22 topologies as input files with the model 'LG + C20 + F + R' for the topology test for USCOs. The UCE_V matrix and three topologies were used as input file with the model 'GTR + I + G' for the topology test for UCEs.

Relative frequency analyses of the gene trees in the USCO, UCE and mitogenome datasets were performed in DiscoVista v1.0 (Sayyari et al., 2018) to visualize the frequencies of the three possible gene tree topologies around the major discordance, respectively. For every focal split, a bar graph shows the quartet frequencies around that branch.

Divergence time estimation

Estimation of divergence time of Neuropterida was performed using MCMCTree in PAML v4.9j (Yang, 2007). For the analysis with fossil calibrations, three different calibration schemes were used as calibration points (Table S3). The first scheme of fossil calibrations was set in

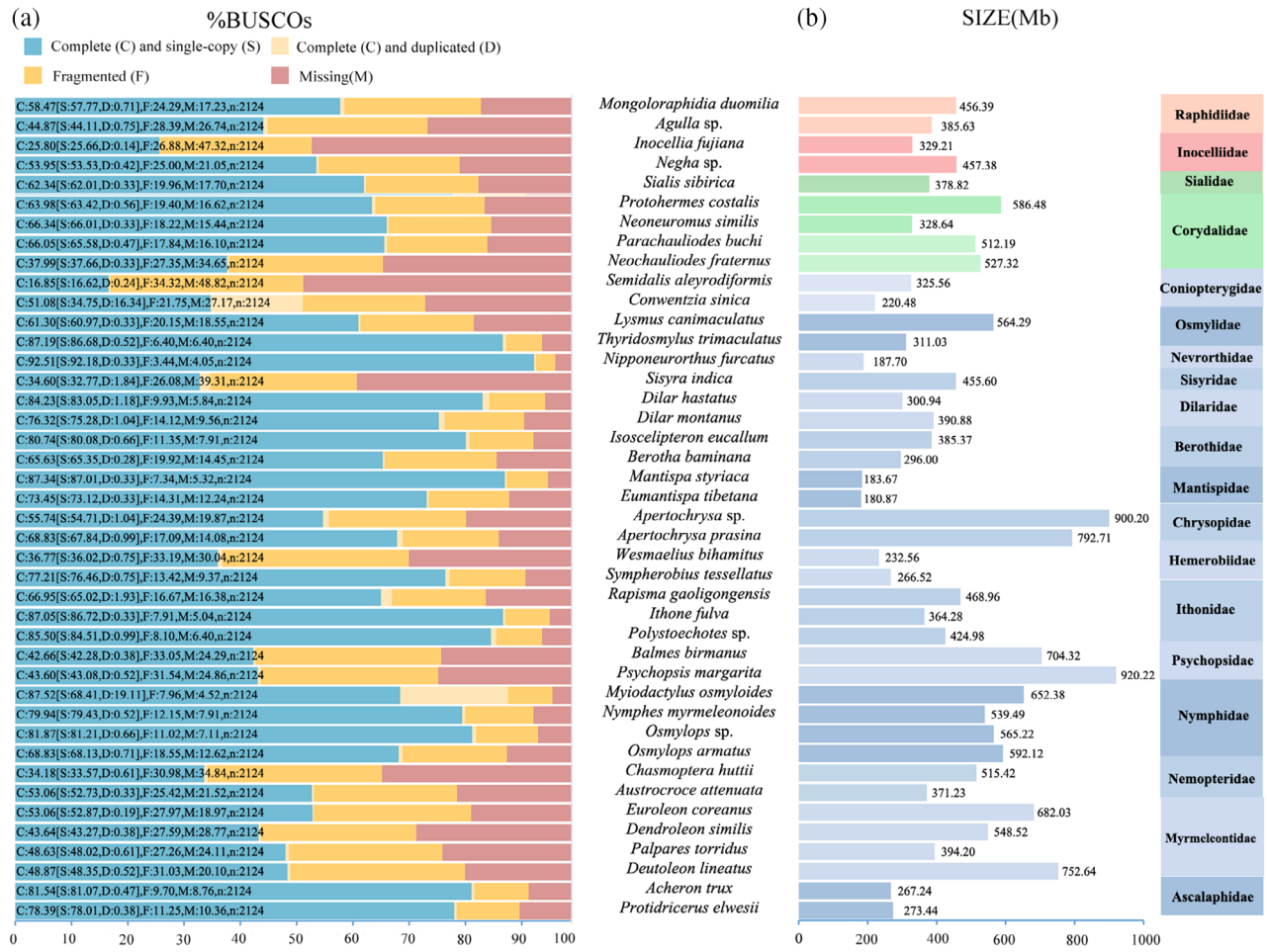

FIGURE 2 Genome completeness assessment against the specific endopterygota_odb10 set ($n = 2114$) and genome size of 42 Neuropterida species.

deeper nodes closer to the root, the second scheme of fossil calibrations was set in shallower nodes and the last scheme was a combination of the two previous schemes of fossil calibrations. The USCO_V, UCE_V and PCGR matrices and their corresponding robust topologies were selected as input files, respectively, because they demonstrated the strongest phylogenetic signal. The independent rates clock model (clock = 2) was used for further calculation. The parameters of the mean substitution rate and the rate drift were set as rgene_gamma = '2 20 1' and sigma2_gamma = '1 10 1', respectively. Two independent MCMC runs were performed for 100 million generations with the 25% as burn-in. The analysis was terminated until all the ESS exceeded 200 using Tracer v1.6 (Rambaut et al., 2018).

RESULTS

Genome assembly

Approximately 15.44–69.89 Gb clean data was generated for each sample, with coverage ranging from 32.80 X (*Balmes birmanus* McLachlan, Psychopsidae) to 202.70 X (*Mantispa styriaca* (Poda), Mantispidae) obtained (Figure S1 and Table S4). The genomes of 42 Neuropterida species were newly assembled in this study, with genome sizes ranging from 180.87 Mb (*Eumantispa tibetana* Yang, Mantispidae) to 920.22 Mb (*Psychopsis margarita* Tillyard, Psychopsidae) (Figure 2 and Table S5). The lengths of scaffold N_{50} were shortest in *Nipponeurorthus furcatus* Liu, H. Aspöck *et* U. Aspöck (3.78 kb, Nevrorthidae) and longest in *Wesmaelius bihamitus* (Yang) (116.90 kb, Hemerobiidae). The number of scaffolds ranged from 27,488 (*N. furcatus*, Nevrorthidae) to 398,979 (*P. margarita*, Psychopsidae). The GC content ranges from 25.37% for *Neoneuromus similis* Liu, Hayashi *et* Yang (Corydalidae) to 37.17% for *W. bihamitus* (Hemerobiidae). Additional statistics including average read coverage, the number of scaffolds, assembly size, longest scaffold length, scaffold N_{50} and BUSCO completeness are provided in Table S5.

Marker capture and matrix generation

For the extraction of USCOs, almost all species were extracted with more than 50% complete and single-copy USCOs, except *Semidalis aleyrodiformis* (Stephens) (Coniopterygidae), *Conwentzia sinica* Yang

(Coniopterygidae) and *Inocellia fujiana* Yang (Inocelliidae), which may be due to the age of the specimens or the small size of the samples (Table S6). There were 2124 USCOs obtained for the subsequent analyses. After removing loci containing fewer than three sequences, 2117 loci were retained. The USCO_all matrix contained 774,120 amino acid sites. After filtering according to Strategy I, there are four matrices (USCO_I_abs70–85) containing 443,858–756,620 amino acid sites. After filtering according to Strategy II, there are five matrices (USCO_II_70–90) including 246,421–627,733 amino acid sites. After filtering according to Strategy III, there are four matrices (USCO_III_70–85_abs70–85) including 227,899–613,265 amino acid sites. After filtering according to Strategy IV and Strategy V, the concatenated USCO_IV and USCO_V matrices contained 229,165 and 189,562 amino acid sites, respectively (Table S7).

For the UCE probe design, there were 3,590,180–15,155,374 reads from 10 genome assemblies modelled by ART v2.5.1 (Huang et al., 2012). There were 0.67%–10.16% simulated reads mapped to the base genome (*C. carnea*). The results of the PHYLUCE probe design showed that the provisional Neuropterida UCE probe set contained 16,269 baits and 8404 conserved loci. A set of 5341 target genes shared by eight species was selected for the final probe design. After deleting the duplicates, the final UCE probe set of Neuropterida was obtained, which contained 90,672 probe decoys and 5341 conserved loci. Then, there were a total of 3823 UCEs with 3,384,479 nucleotide sites extracted. The UCE_all matrix contained 3,384,479 nucleotide sites. After filtering according to Strategy I, there were four matrices (UCE_I_abs70–85) contained 939,334–3,155,238 nucleotide sites. After filtering according to Strategy II, there are five matrices (UCE_II_70–90) including 171,412–1,988,741 nucleotide sites. After filtering according to Strategy III, there are four matrices (UCE_III_70–85_abs70–85) including 210,992–1,957,185 nucleotide sites. After filtering according to Strategy IV and Strategy V, the UCE_IV and UCE_V matrices contained 958,077 and 725,499 nucleotide sites, respectively (Table S7).

A total of 21 complete mitogenomes were assembled, with lengths ranging from 16,307 bp in *Sisyra indica* Needham (Sisyridae) to 17,001 bp in *Dilar hastatus* Zhang, Liu, H. Aspöck *et* al U. Aspöck (Dilaridae) (Table S8). A total of 18 nearly complete mitogenomes were assembled, with only the control region being difficult to assemble due to its high AT content. The length of PCGs in the presently sampled species of Neuropterida ranges from 11,060 bp in *Mongoloraphidia duomilia* (Yang) (Raphidiidae) to 11,206 bp in *N. furcatus* (Nevrorthidae) (Table S8). The changes in the length of PCG, *rrnL* and *rrnS* were relatively small, and the changes in the length of the control region were relatively large. The PCG and PCGR matrix contained 11,055 and 12,887 bp, respectively. The PCG12 and PCG12R matrices contained 7370 and 9202 bp, respectively. The AA matrix contained 3559 amino acid sites.

Unstable ABS values

Based on the results of six repetitions of the phylogenetic reconstruction using all USCOs loci (2117 × 6) with the same command, ABS (the absolute value of the subtraction of the six repetitions) changed greatly, ranging from 0 to 18.5 (Figure 3a). The total number of changes of the ABS in six repetitions was 31,413 (98.92%) times. The number of changes of ABS in six repetitions equal to or greater than 1 was about 12,673 (39.91%) times, equal to or greater than 2 was 5736 (18.06%) times. The number of changes in ABS in three repetitions equal to or greater than 3 and 4 was 608 and 1627 times, respectively. The number of changes in ABS in three repetitions equal to or greater than 5 was 902 times. Notably, the astral trees obtained from the six repetitions exhibited inconsistent topology (Figure 3b). While the first five replications consistently supported Chrysopidae as the sister group to Myrmeleontoidea, the sixth replication demonstrated Hemerobiidae as the sister group to the Myrmeleontoidea.

Phylogeny of Neuropterida based on USCOs

Topologies of phylogenetic trees of Neuropterida based on USCOs under inferred models were consistent at most nodes (Figure S2). Raphidioptera were recovered as monophyletic and as the sister group to the rest of Neuropterida. Megaloptera were also recovered as monophyletic and the sister group to the monophyletic Neuroptera. Each family of Neuropterida (Raphidioptera, Megaloptera and Neuroptera) was strongly supported as reciprocally monophyletic except for Myrmeleontidae. Coniopterygidae were recovered as the sister group to the rest of the Neuroptera. Dilaridae were demonstrated as the sister group to all remaining Neuroptera except for Coniopterygidae and Osmyloidea. Nemopteridae were demonstrated as the sister group to the branch including Myrmeleontidae and Ascalaphidae. However, the phylogenetic positions of Hemerobiidae and Chrysopidae were inconsistent, with five different topologies being recovered across different analyses (Figure S2A). Under the first topology (T1), Hemerobiidae were recovered as the sister group to the rest of Neuroptera except for Coniopterygidae, Osmyloidea and Dilaridae, while Chrysopidae were recovered as the sister group to Mantispoidea. Under the second topology (T2), Chrysopidae were recovered as the sister group to the branch including Hemerobiidae + Myrmeleontoidea. Under the third topology (T3), the branch of Chrysopidae + Mantispoidea was recovered as the sister group to the branch including Hemerobiidae + Myrmeleontoidea. Under the fourth topology (T4), Hemerobiidae were recovered as the sister group to the branch including (Mantispoidea + (Chrysopidae + Myrmeleontoidea)). Under the fifth topology (T5), Mantispoidea were recovered as the sister group to the branch including (Hemerobiidae + (Chrysopidae + Myrmeleontoidea)). Most concatenated ML analyses with USCO matrices generated topology T1 except for USCO_I_abs85 and USCO_IV matrices. Most phylogenetic analyses using the Dayhoff6 recoding scheme based on USCO matrices supported topology T2 except for USCO_II_85, USCO_III_75_abs75, USCO_III_85_abs85 and USCO_IV matrices. Most phylogenetic analyses under the PMSF model based on USCO matrices supported topology T3 except for the USCO_I_abs85 matrix. The concatenated partitioning ML analyses with the USCO_IV matrix, as well as the ML analysis using the

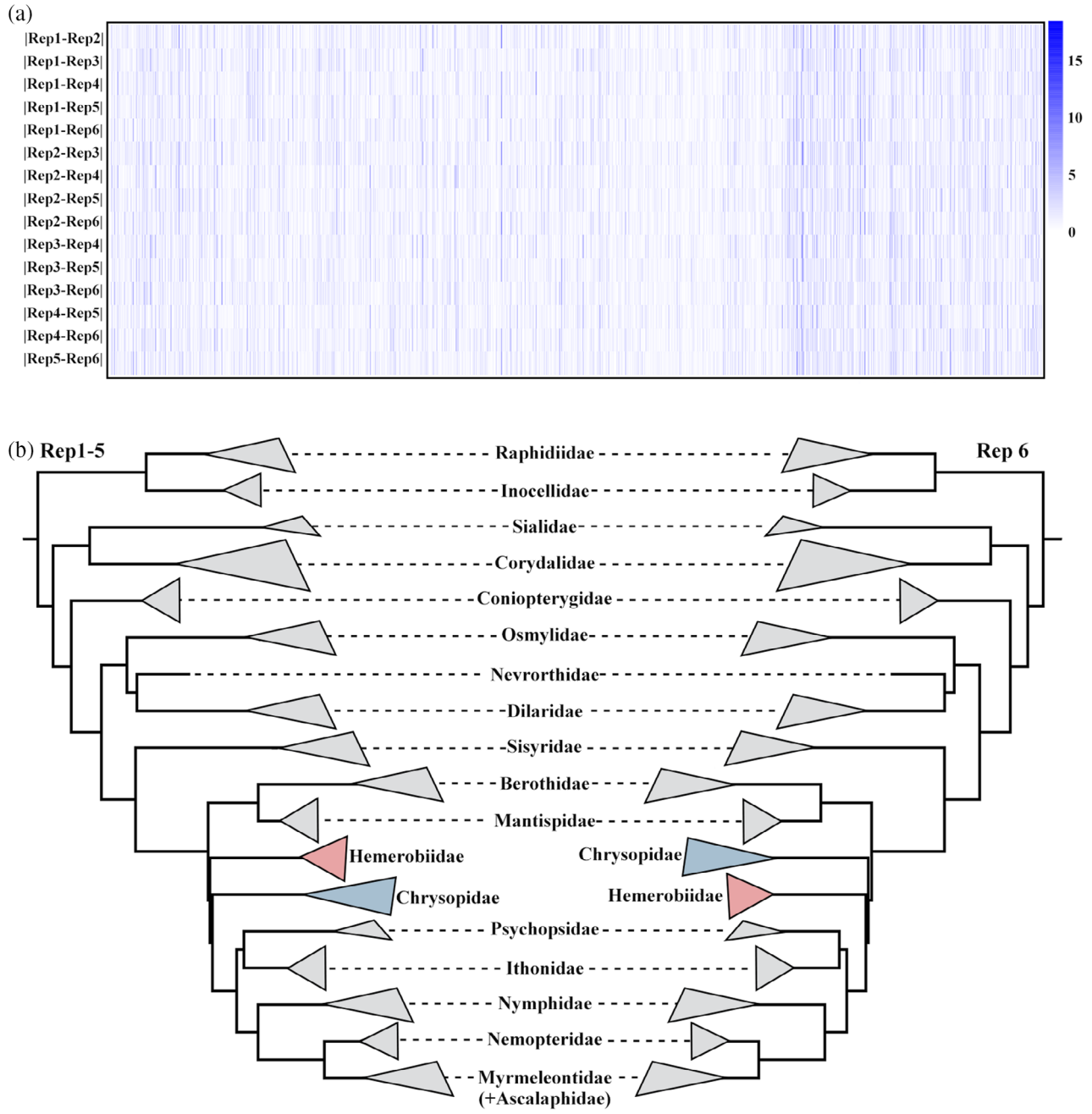

FIGURE 3 Comparative visualization of average bootstrap support variability across six replicate analyses. (a) Heatmap representation of absolute sum of the difference in average bootstrap support values among the six replicates of the gene trees reconstructed from all USCOs (2117). (b) Corresponding topological variations in six replicate ASTRAL trees. The phylogenetic positions of Chrysopidae (marked in blue) and Hemerobiidae (marked in pink) showed inconsistency among the six ASTRAL trees.

Dayhoff6 recoding scheme with USCO_III_75_abs75 and USCO_IV matrices, consistently supported topology T4. Most coalescent analyses of USCO matrices, conducted under ABS-based filtering (Strategy I), completeness-based filtering (Strategy II) and their combined approach (Strategy III), consistently supported topology T5.

Besides, there were also some minor controversies within Chrysopidae, Nymphidae and Myrmeleontidae (plus Ascalaphidae). Six topologies of Chrysopidae were inferred using all USCO matrices, with the majority of nodes consistent except for the position of *Nothochrysa sinica* Yang, *Italochrysa pardalina* Yang *et* Wang and *Apertochrysa* sp. (Figure S2B). Under the first (a) and second topologies (b), *Apertochrysa* sp. was recovered as the sister group to *N. sinica*, while *I. pardalina* was recovered as the sister group to the rest of Chrysopinae (a) or the clade comprising *Chrysopa* Leach and *Chrysoperla* Steinmann (b). Under the third topology (c), *Apertochrysa* sp. was recovered as the sister group to the clade comprising *Chrysopa* and *Chrysoperla*. Under the fourth topology (d), *Apertochrysa* sp. was recovered as the sister group to *Chrysoperla*. None of these four topologies (a, b, c and

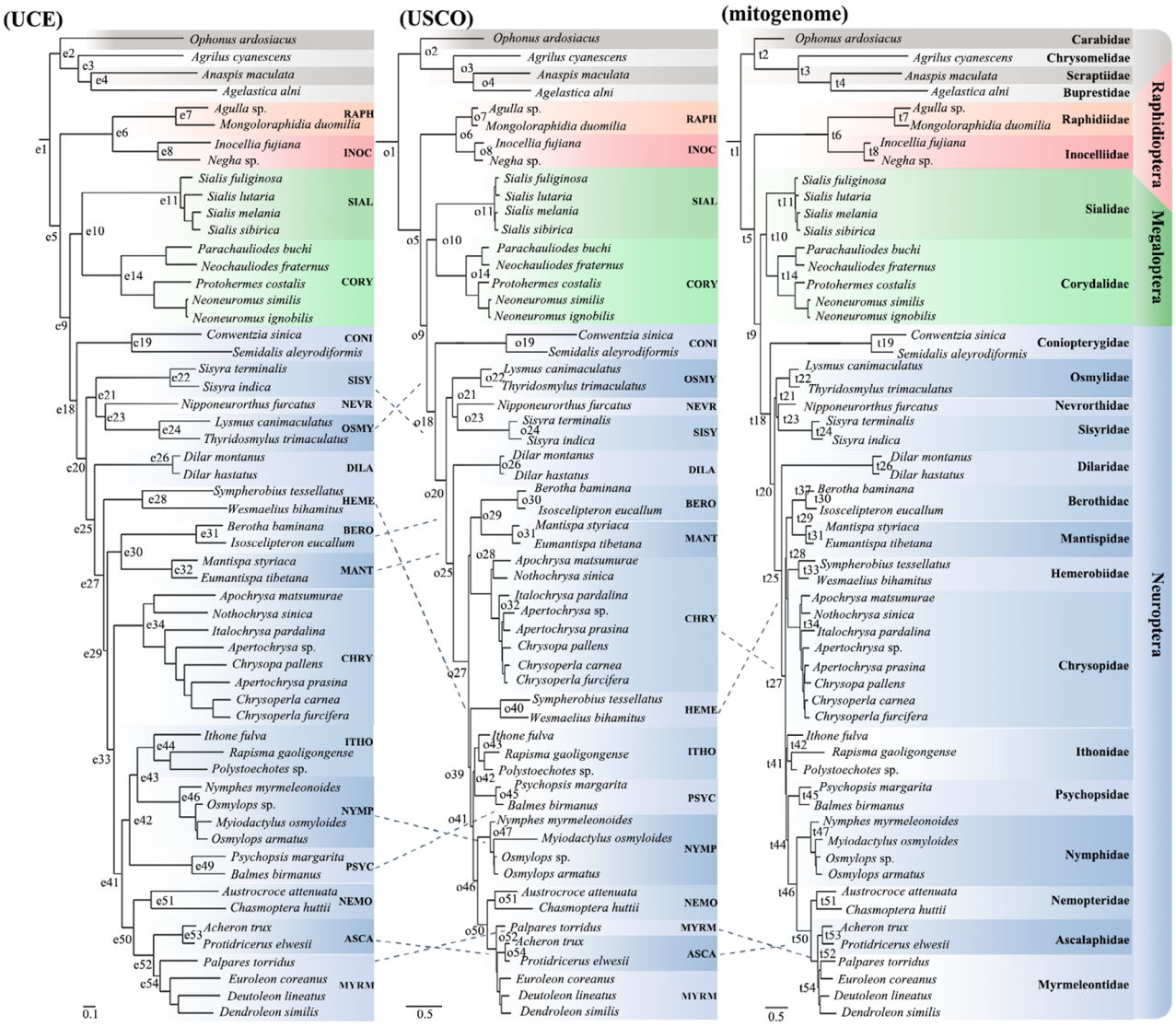

FIGURE 4 Comparison of preferred phylogenetic trees of Neuropterida based on USCO, UCE and mitogenome. Four-letter prefixes of exemplar names represent codes for respective family names. Conflicting tips between data types are indicated with dotted lines between the topologies. The node numbers are consistent with those in Figures 5 and 6, Tables S9–S11. USCO, Universal single-copy ortholog; UCE, ultraconserved element; mitogenome, mitochondrial genome.

d) supported the monophyly of *Apertochrysa* Tjeder. Under the fifth topology (e), Nothochrysinae (*N. sinica*) was recovered as the sister group to Chrysopinae. Under the sixth topology (f), Nothochrysinae (*N. sinica*) was recovered as the sister group to Apochrysinae. Using the concatenated ML analyses and the ML analyses with Dayhoff6 recoding scheme, USCO_IV and USCO_V matrices supported the fifth topology and the sixth topology (f), respectively. Both the concatenated partitioning ML analyses and the ML analyses with PMSF model based on USCO_IV and USCO_V matrices supported the sixth topology (f).

The phylogenetic relationships within the Nymphidae were inconsistent according to different matrices and models (Figure S2C). Almost all analyses of USCO matrices supported *Nymphes* Leach as the sister group to (*Myiodactylus* Brauer + *Osmylops* Banks) (labelled as 'β'), whereas ML analyses using Dayhoff6 recoding scheme (USCO_all, USCO_I_abs75, USCO_II_70–90 and USCO_III_75–85_abs75–85) supported *Myiodactylus* being the sister group to *Nymphes* + *Osmylops* (labelled as '&').

The monophyly of Myrmeleontidae was inconsistent, resulting in three topologies (Figure S2D). Under the first topology (labelled as '*'), Myrmeleontidae was recovered as the sister group to Ascalaphidae and *Euroleon* Esben-Petersen was recovered as the sister group to *Dendroleon* Brauer + *Deutoleon* Navás. Under the second topology (labelled as '#'), Myrmeleontidae was paraphyletic relative to Ascalaphidae, while *Euroleon* was recovered as the sister group to *Dendroleon* + *Deutoleon*. Under the third topology (labelled as '^'), Myrmeleontidae were recovered as the sister group to Ascalaphidae, while *Dendroleon* were recovered as the sister group to *Euroleon* + *Deutoleon*. ML analyses using Dayhoff6 recoding scheme and under PMSF model, as well as concatenated ML analyses with 10 matrices

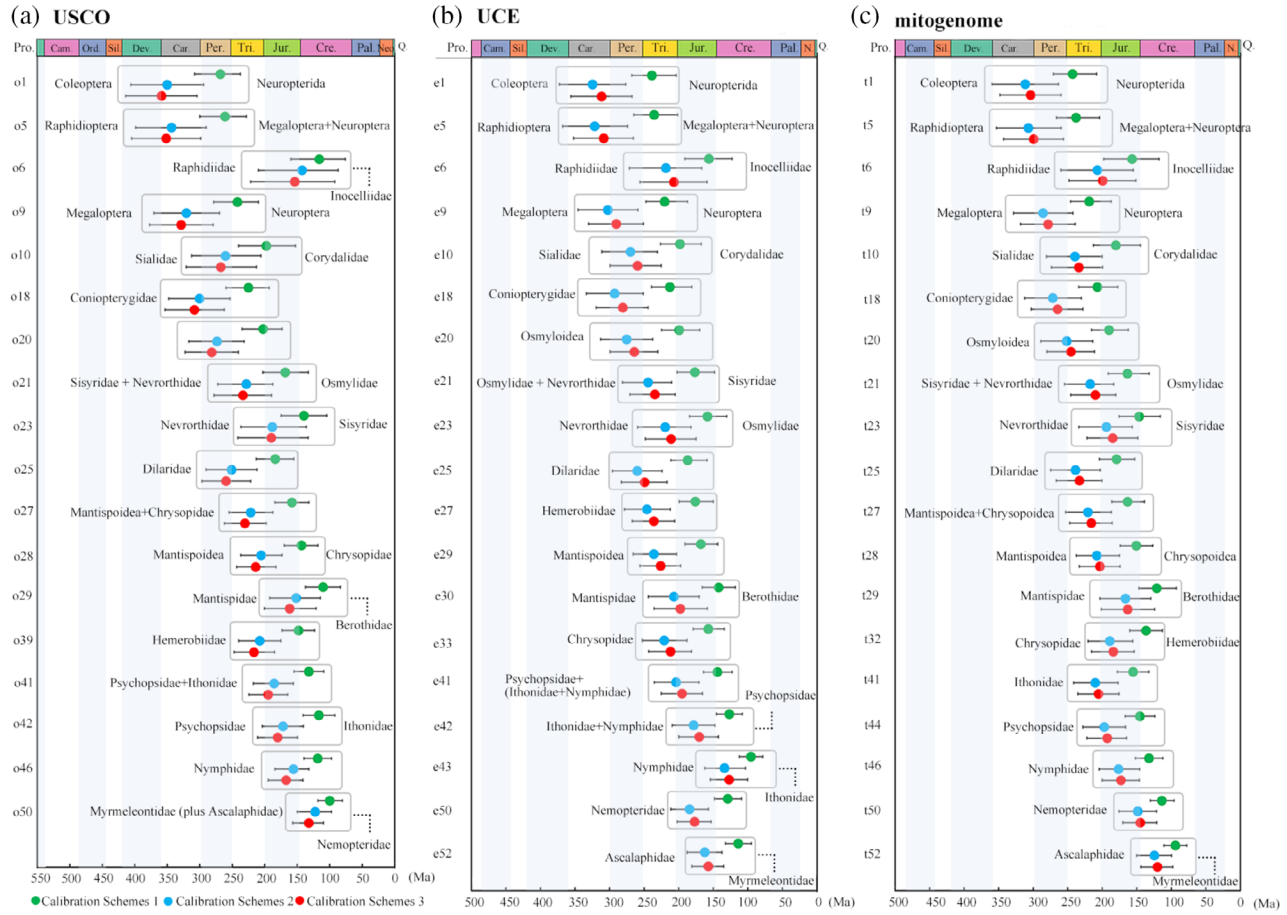

FIGURE 5 Comparison of divergence time estimates for major splitting events based on analysis of USCO_V (a), UCE_V (b) and PCGR (c) matrices using three different calibration schemes, respectively. Each box represents a splitting event with its two branching lineages (if applicable) shown left and right sides. The circle within boxes represents the mean of the posterior estimate, and the whiskers mark the upper and lower 95% highest posterior density of the age estimates.

(USCO_all, USCO_I_abs70, USCO_II_70–90 and USCO_III_70–80_abs70–80) also recovered the first topology (*). ML analyses using Dayhoff6 recoding scheme with USCO_V matrix, as well as ML partitioning analyses and ML analyses under PMSF model with USCO_II_90 matrix recovered the first topology (*). All ASTRAL analyses using USCO matrices supported the second topology (#). Some other analyses also support the second topology (#), including concatenated ML analyses with USCO_I_abs75, USCO_I_abs80, USCO_IV and USCO_V, ML analyses using Dayhoff6 recoding scheme with USCO_I_abs75, USCO_I_abs80 and USCO_IV matrices and ML analyses under PMSF model with USCO_I_abs75, USCO_I_abs80, USCO_IV and USCO_V matrices. ML analyses using Dayhoff6 recoding scheme and under PMSF model, as well as concatenated ML analyses based on USCO_I_abs85 and USCO_III_85_abs85 matrices supported the third topology (^).

The topology test based on USCOs indicated that T3- $f^{\#}_{\beta}$ was strongly supported by all hypothesis tests (AU, WKH and WSH) performed herein (Figure 4, Table 1). Results based on the USCO dataset recovered Chrysopidae as the sister group to Mantispoidea, but Hemerobiidae as the sister group to Myrmeleontoidea. Myrmeleontidae were demonstrated to be paraphyletic relative to Ascalaphidae.

Phylogeny of Neuropterida based on UCEs

Phylogenetic trees inferred from different UCE matrices were largely congruent, but the position of Hemerobiidae was controversial (Figure S3). Hemerobiidae were demonstrated to be the sister group to Mantispoidea + (Chrysopidae + Myrmeleontoidea) (Ti) or Chrysopidae + Myrmeleontoidea (Tii) (Figure S3a). Almost all concatenated ML analyses with UCE matrices (excluding UCE_II_85 and UCE_II_90 matrices) supported topology Ti. The coalescent analyses with UCE matrices (excluding UCE_IV and UCE_V matrices) supported topology Tii. Additionally, all analyses of UCE matrices did not support the monophyly of *Osmylops* (labelled as 'o'), except the concatenated ML analyses for UCE_II_90 matrix (labelled 'β' as for USCOs) (Figure S3b,c).

The topology test based on UCEs indicated that Ti_o was strongly supported by all hypothesis tests (AU, WKH and WSH) herein performed (Figure 4, Table 1). According to the two robust phylogenies

TABLE 1 Comparison of tree topology based on USCOs and UCEs, plus signs denote the 95% confidence sets, minus signs denote significant exclusion.

Loci type	Topology	lLogL	deltaL	Bp-RELL	p-WKH	p-WSH	p-AU	Method and matrix
USCO	T1-a^{*}_{β}	−7027740.478	14,056	0−	0−	0−	1.26e-29−	Partition: USCO_all, USCO_I_abs70−75, USCO_II_70−90, USCO_III_70_abs70, USCO_III_75_abs75
	T1-$a^{\#}_{\beta}$	−7027709.703	14,026	0−	0−	0−	4.07e-42−	Partition: USCO_I_abs75−80
	T1- $a^{\wedge}_{\beta}$	−7028175.128	14,491	0−	0−	0−	6.61e-71−	Partition: USCO_III_85_abs85
	T1-$f^{\#}_{\beta}$	−7013959.798	275.72	0−	0−	0−	6.26e-09−	Partition: USCO_V
	T2-$a^{\#}_{\beta}$	−7027715.107	14,031	0−	0−	0−	3.29e-41−	Dayhoff6: USCO_I_abs80
	T2-$a^{\#}_{\&}$	−7029415.22	15,731	0−	0−	0−	5.2e-69−	Dayhoff6: USCO_I_abs75
	T2-$b^{\wedge}_{\beta}$	−7029145.635	15,462	0−	0−	0−	1.83e-06−	Partition: USCO_I_abs85; Dayhoff6: USCO_I_abs85; ASTRAL: USCO_I_abs85
	T2-$a^{\wedge}_{\&}$	−7029872.314	16,188	0−	0−	0.0106−	3.09e-32−	Dayhoff6: USCO_II_90
	T2-a^{*}_{β}	−7027747.79	14,064	0−	0−	0−	1.11e-56−	Dayhoff6: USCO_I_abs70, USCO_II_70−80, USCO_III_70_abs70
	T2-$a^{*}_{\&}$	−7029872.314	16,188	0−	0−	0.0096−	3.09e-32−	Dayhoff6: USCO_all, USCO_III_80_abs80
	T2-$c^{\#}_{\beta}$	−7015395.727	1711.7	0−	0−	0−	3.92e-06−	ASTRAL: USCO_IV, USCO_V
	T2-f^{*}_{β}	−7013974.494	290.42	0−	0−	0.0001−	8.79e-05−	Dayhoff6: USCO_V
	T3-$a^{\#}_{\beta}$	−7027428.861	13,745	0−	0−	0−	6.8e-38−	PMSF: USCO_I_abs75, USCO_I_abs80
	T3-a^{*}_{β}	−7027460.067	13,776	0−	0−	0−	5.08e-07−	PMSF: USCO_all, USCO_I_abs70, USCO_II_70−90, USCO_III_70−80_abs70−80
	T3-$a^{\wedge}_{\beta}$	−7027894.874	14,211	0−	0−	0−	3.95e-40−	PMSF: USCO_III_85_abs85
	T3-$f^{\#}_{\beta}$	**−7013684.076**	**0**	**1+**	**1+**	**1+**	**1+**	**PMSF: USCO_IV, USCO_V**
	T4-$e^{\#}_{\beta}$	−7014559.5	875.42	0−	0−	0−	5.83e-06−	Partition: USCO_IV; Dayhoff6: USCO_IV
	T4-$a^{*}_{\&}$	−7029865.63	16,182	0−	0−	0−	9.33e-68−	Dayhoff6: USCO_III_75_abs75
	T5-$a^{*}_{\&}$	−7029845.606	16,162	0−	0−	0−	2.33e-72−	Dayhoff6: USCO_II_85
	T5-$a^{\wedge}_{\&}$	−7030271.108	16,587	0−	0−	0−	8.07e-76−	Dayhoff6: USCO_III_85_abs85
	T5-$c^{\#}_{\beta}$	−7015791.575	2107.5	0−	0−	0-	4.46e-06−	ASTRAL: USCO_all, USCO_I_abs70−85, USCO_II_70−85, USCO_III_70_abs70, USCO_III_75_abs75
	T5-$d^{\#}_{\beta}$	−7016048.576	2364.5	0−	0−	0-	5.95e-34−	ASTRAL: USCO_II_90
UCE	**Tio**	**−17910284.56**	**0**	**1+**	**1+**	**1+**	**1+**	Partition: UCE_all, UCE_I_abs70−85, UCE_II_70−80, USCO_III_70−85_abs70−85, UCE_IV, UCE_V; ASTRAL: UCE_IV, UCE_V
	Tiio	−17910695.06	410.5	0−	0−	0−	4.83e-35−	Partition: UCE_II_85; ASTRAL: UCE_all, UCE_I_abs70−85, UCE_II_70−75, UCE_III_70−85_abs70−85
	Tiiβ	−17910785.19	500.63	0−	0−	0−	1.04e-92−	Partition: UCE_II_90

Abbreviations: AU, approximately unbiased; RELL, resampled estimated log-likelihood; UCE, ultraconserved element. Best-fit hypotheses are given in boldface; USCO, universal single-copy ortholog; WKH, weighted Kishino−Hasegawa; WSH, weighted Shimodaira−Hasegawa.

supported by hypothesis tests using USCOs and UCEs, there were four main differences at the family level. First, the clade of (Sisyridae + (Osmylidae + Nevrorthidae)) was supported by the UCE matrices, whereas the clade of (Osmylidae + (Sisyridae + Nevrorthidae)) was supported by the USCO matrices. Second, Chrysopidae + Myrmeleontoidea was supported by the UCE matrices, whereas the sister group relationship between Chrysopidae and Mantispoidea was recovered using USCO matrices. Third, Hemerobiidae were recovered as the sister group to Mantispoidea + (Chrysopidae + Myrmeleontoidea) based on UCE matrices or as the sister group to Myrmeleontoidea based on USCO matrices. Lastly, the sister group relationship between Psychopsidae and Ithonidae was supported based on the USCO matrices, with Nymphidae as the sister group to the remaining Myrmeleontoidea, whereas the clade of (Psychopsidae + (Ithonidae + Nymphidae)) was supported by the UCE matrices. Additionally, the monophyly of Myrmeleontidae was supported based on all UCE matrices.

Phylogeny of Neuropterida based on mitochondrial genomes

The phylogeny inferred from the mitogenome data is largely consistent with that from USCOs, although there are still some topological

conflicts among the results from different matrices (Figure S4). Most results supported Osmylidae + (Sisyridae + Nevrorthidae) (H1, H2, H4 and H5), except the ML analyses of PCG12, PCGR and AA matrices (H3 and H6). Almost all results supported that Dilaridae were the sister group to all remaining Neuroptera except Coniopterygidae and Osmyloidea, while the ML analyses of PCG and PCGR matrices supported H5 and H6, respectively. Most results supported Nymphidae as the sister group to all remaining Myrmeleontoidea except Psychopsidae, while the ML analyses of PCG and PCGR matrices supported H5 and H6. Most results supported that Myrmeleontidae were paraphyletic with Ascalaphidae nested within it. However, the ML and BI trees of the PCGR matrix under the homogeneous model, as well as the BI tree of the PCGR matrix under the heterogeneous model (CAT-GTR) all supported the monophyly of Myrmeleontidae.

Notably, only the BI tree under the heterogeneous model recovered Coniopterygidae as the sister group to the rest of Neuroptera, while all other analyses under the homogeneous model assigned Coniopterygidae as the sister group to Raphidioptera (H1). Based on extensive morphological and molecular evidence supporting the monophyly of Neuroptera (Aspöck et al., 2001; Song et al., 2018; Song, Li, Zhai, et al., 2019; Vasilikopoulos et al., 2020; Wang et al., 2017; Winterton et al., 2010; Winterton et al., 2018; Zhao et al., 2014; Zhao et al., 2024), the sister group relationship between Coniopterygidae and Raphidioptera is implausible, most likely due to compositional heterogeneity and long-branch attraction (LBA) (Ban et al., 2022; Gao et al., 2022; Susko & Roger, 2021; Xiong et al., 2023; Zou et al., 2020). In many studies, the heterologous models (i.e. CAT-GTR) have been demonstrated to be effective in solving phylogenetic problems of compositional heterogeneity based on mitogenomes (Lartillot et al., 2007; Song, An, et al., 2016; Tian et al., 2023). Recent studies have demonstrated that the third codon positions of PCGs in the mitogenome of Neuropterida are significantly saturated (Song, Li, Zhai, et al., 2019; Tian et al., 2023). Moreover, the site-heterogeneous CAT-GTR model has been shown to be least sensitive to long-branch attraction (Cai, 2024; Cai et al., 2023; Song, Li, Yin, et al., 2019). Therefore, the BI (CAT-GTR) tree represents the most robust phylogeny based on mitogenome data (Figure 4).

However, in this BI tree, the topologies concerning the positions of Chrysopidae, Hemerobiidae and Ithonidae are not consistent with that inferred from the USCOs. Specifically, the mitogenome analysis supported Hemerobiidae as the sister group to Chrysopidae, whereas USCO analyses consistently recovered Chrysopidae as sister to Mantispoidea and Hemerobiidae as sister to Myrmeleontoidea. Furthermore, while mitogenome data placed Ithonidae as sister group to the remaining Myrmeleontoidea, USCO analyses consistently supported Ithonidae as the sister group to Psychopsidae.

Relative branching quartet frequency analyses

The results of the relative frequency of gene trees showed conflicting signals for several key nodes within the backbone phylogeny of Neuropterida (Figure S5). The clade of Nevrorthidae + Sisyridae was supported by most gene trees of the USCO dataset (44%), whereas Osmylidae + Nevrorthidae was supported by a larger proportion of gene trees in the UCE (43%) and mitogenome (51%) datasets. Chrysopidae + Mantispoidea was supported by most gene trees in the mitogenome datasets (42%), whereas this topology and the other possible topologies were supported by subequal proportions of gene trees in the USCO dataset (37%). Chrysopidae + Myrmeleontoidea was supported by most gene trees in the USCO (40%) and UCE (48%) datasets, whereas one of the other possible topologies was supported by most gene trees in the mitogenome datasets. Hemerobiidae + Myrmeleontoidea was supported by most gene trees in the USCO dataset (36%), whereas the other possible topologies were supported by most gene trees in the UCE and mitogenome datasets. Hemerobiidae as the sister group to the rest of Neuroptera except Coniopterygidae, Osmyloidea and Dilaridae was supported by most gene trees in the USCO, UCE and mitogenome datasets. This topology was also present in T1 and T4 of the USCO dataset.

The clade of Psychopsidae + Ithonidae as sister group to the rest of Myrmeleontoidea was supported by most gene trees in the USCO (42%) and UCE (45%) datasets, whereas the clade of Psychopsidae + Ithonidae and one of the other possible topologies was supported by equal proportions of gene trees in mitogenome datasets (40%). The clade of Psychopsidae + Ithonidae as sister group to the rest of Myrmeleontoidea, followed by Nymphidae, was supported by most gene trees in the USCO (55%) and mitogenome (78%) datasets, whereas this topology was not supported by most gene trees in the UCE dataset. The clade of Ithonidae + Nymphidae, as well as the clade of Psychopsidae + (Nymphidae + Ithonidae) were supported by most gene trees in the UCE dataset, whereas other possible topologies were recovered by most gene trees of USCOs and mitogenome datasets. Ithonidae as the sister group to the rest of Myrmeleontoidea was supported by most gene trees in the USCOs (60%) and UCE (67%) datasets, whereas most gene trees supported other possible topologies in mitogenome datasets. The clade of Psychopsidae as sister group to the rest of Myrmeleontoidea except Ithonidae was not supported by most gene trees in the USCO and UCE datasets, whereas this topology and one of the other possible topologies were supported by equal proportions of gene trees in mitogenome dataset (40%). Nymphidae as sister group to the rest of Myrmeleontoidea except Psychopsidae and Ithonidae was supported by most gene trees in the USCO (56%), UCE (41%) and mitogenome (78%) datasets. The monophyly of Myrmeleontidae, as well as the sister group relationship between Myrmeleontidae and Ascalaphidae, were supported by most gene trees in UCE and mitogenome datasets, whereas this topology was not supported by most gene trees in the USCO dataset.

Divergence time estimation

The divergence time estimation under three calibration schemes respectively using USCO_V, UCE_V and PCGR matrices (Tables S9–S11) was compared. Divergence time estimates obtained from calibration scheme 2 (primarily deep node calibrations) and

scheme 3 (combined deep and shallow node calibrations) demonstrated remarkable consistency across all data types, with no statistically significant discrepancies observed (Figure 5 and Tables S9–11). However, the estimation ages of major divergence events under calibration scheme 1 (most calibrations at young nodes) recovered much younger ages than those obtained from the other two calibration schemes, a pattern consistently observed across all three datasets (Figure 5). Besides, our results suggested that the mitogenomes and UCE tend to underestimate the divergence time compared to USCOs. For instance, Coleoptera and Neuropterida began to diverge during the Permian using the USCO_V matrix rather than during the Triassic using the PCGR and UCE_V matrices based on calibration schemes 1. Based on calibration schemes 2 and 3, Coleoptera and Neuropterida diverged at the start of the Carboniferous using the USCO_V matrix rather than at the end of the Carboniferous using the PCGR and UCE_V matrices. According to established studies, increasing the number of fossil calibrations improves the reliability of different time estimates (Duchene et al., 2014; Ho & Phillips, 2009; Hug & Roger, 2007; Linder et al., 2005; Lukoschek et al., 2012; Marshall, 2008; Villaverde et al., 2021). In general, calibrations at the root or at deeper nodes are preferred over those at shallower nodes (Duchene et al., 2014; Ho, 2007; Mello & Schrago, 2014; Püschel et al., 2019; Rannala & Yang, 2007; Sauquet et al., 2012; Yang & Rannala, 2006). Calibration scheme 3 was set to include more fossil markers than calibration scheme 2 and may be more accurate based on previous studies (Duchene et al., 2014; Ho & Phillips, 2009; Hug & Roger, 2007; Linder et al., 2005; Lukoschek et al., 2012; Marshall, 2008; Villaverde et al., 2021).

The chronogram presents the divergence times under the calibration scheme 3 using USCO_V matrix (Figure 6). The 95% highest posterior density (HPD) values for each node are given in Table S9. The phylogenetic split between Coleoptera and Neuropterida was estimated at the beginning of the Carboniferous ca. 358.61 Ma (HPD = 304.22–414.18 Ma), which was older than most of the other studies based on AHE data (~267 Ma) (Winterton et al., 2018), mitogenome data (~311.34 Ma) (Wang et al., 2017) and transcriptome data (~321.35 Ma) (Vasilikopoulos et al., 2020) (Figure S6). Neuroptera and Megaloptera diverged from Raphidioptera at ca. 351.51 Ma (HPD = 298.49–404.92 Ma). Neuroptera diverged from Megaloptera at ca. 328.44 Ma (HPD = 279.39–377.16 Ma) during the Mississippian Carboniferous, which was older than some other studies based on AHE data (~277.75 Ma) (Winterton et al., 2018), mitogenome data (~297.47 Ma) (Wang et al., 2017) and transcriptome data (~300.05 Ma) (Vasilikopoulos et al., 2020). Sialidae diverged from Corydalidae at ca. 267.51 Ma (HPD = 212.74–320.84 Ma) during the Triassic, while Raphidiidae and Inocelliidae split during the Jurassic. Within Neuroptera, Coniopterygidae were the earliest group to diverge at the Upper Pennsylvanian in the Carboniferous. Osmyloidea started to diversify during the Upper Triassic (mean = 233.45 Ma, HPD = 189.54–278.02 Ma). Both Sisyridae (mean = 65.64 Ma, HPD = 32.30–107.71 Ma) and Osmylidae (mean = 51.34 Ma, HPD = 29.28–81.02 Ma) started to diversify during the Paleogene. The divergence between Nevrorthidae and Sisyridae was estimated to have occurred during the Lower Jurassic (mean = 189.65 Ma, HPD = 133.06–241.34 Ma). Dilaridae diverged from its sister group, that is, all neuropteran families except Osmyloidea and Coniopterygidae, at ca. 259.93 Ma, which was consistent with an estimate based on mitogenome data (~258.71 Ma) (Wang et al., 2017), but older than estimates using transcriptome data (~227.90 Ma) (Vasilikopoulos et al., 2020) and AHE data (~239.73 Ma) (Winterton et al., 2018). Lastly, the crown groups of various neuropteran families (e.g., Chrysopidae, Hemerobiidae, Nemopteridae, Ithonidae and Myrmeleontidae) started to diversify during the Cretaceous, while those of the remaining extant families (i.e., Dilaridae, Berothidae, Mantispidae, Psychopsidae and Nymphidae) started to diversify during the Paleogene.

DISCUSSION

Selection of molecular data type in higher-level phylogenetic inference

Comparative analysis of the three genomic datasets revealed strong support for most phylogenetic relationships in Neuropterida. However, there remain conflicts regarding relationships among families of Osmyloidea, and the phylogenetic position of Hemerobiidae, Chrysopidae, Psychopsidae, Ithonidae and Nymphidae. According to relative branching quartet frequency analyses, gene tree conflicts caused by different types of molecular markers are prevalent (Prasanna et al., 2020; Romiguier et al., 2016; Wu, 2012; Xi et al., 2014; Zhao et al., 2023). More attention should be focused on gene tree conflicts caused by biological factors in the future.

The sister group relationship between Nevrorthidae and Sisyridae was supported based on USCO and mitogenome datasets, but not supported based on all UCE analyses (Figure 3; Figure S5). Larvae of Nevrorthidae and Sisyridae are exclusively aquatic in Neuroptera while only certain subfamilies of Osmylidae can be considered as semiaquatic at most (Aspöck et al., 2017; Winterton et al., 2018). This supports the sister group hypothesis between Nevrorthidae and Sisyridae, and that Osmylidae could represent either an intermediate, predominantly semiaquatic or secondarily terrestrial group as some of the more derived Osmylidae are indeed exclusively terrestrial (Winterton et al., 2018). The sister group relationship between Sisyridae and Nevrorthidae is also in agreement with selected morphological data, such as the secondarily multisegmented antennae of the larvae (Jandausch et al., 2019), and some fossil evidence such as the peculiar feature of female tergum 9 and the distal fusion of ScP and RA (Lu et al., 2018; Nakamine et al., 2023). The phylogenetic positions of Chrysopidae and Hemerobiidae inferred from the USCO dataset and the UCE dataset, respectively, are currently supported by a lack of morphological evidence. The clade of Chrysopidae + Hemerobiidae based on mitogenome datasets were determined based on the morphological characteristics such as campodeiform body shape of the larva, and similar larval head shape (Beutel et al., 2010; Garzón-Orduña et al., 2016; Randolf et al., 2014; Winterton et al., 2010).

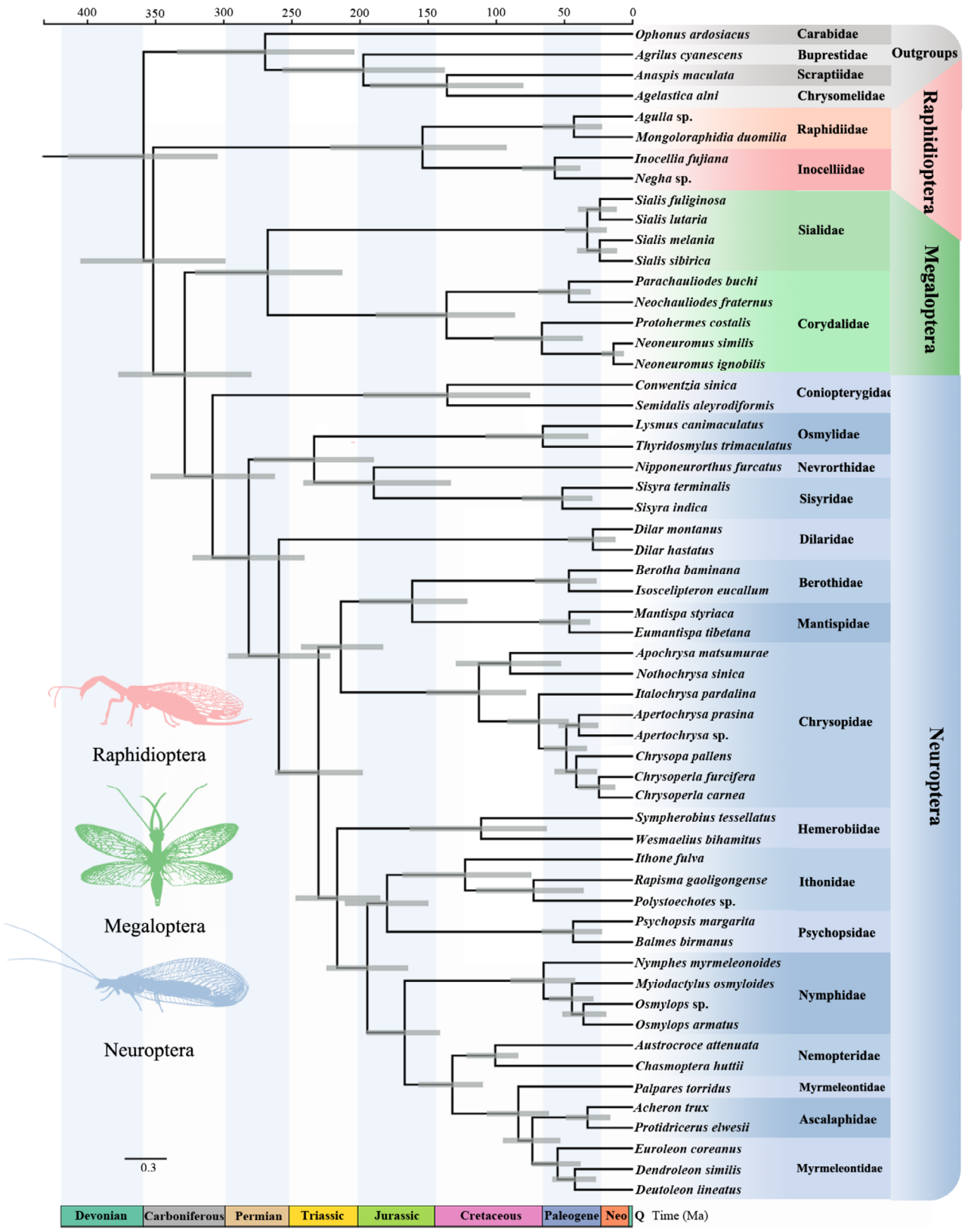

FIGURE 6 Divergence time estimation within Neuropterida using USCO_V matrix. Node numbers refer to values of mean ages, and ranges are provided in Supplementary Table S9.

However, this type of larvae is also found in families such as Psychopsidae, Dilaridae, Berothidae, Rhachiberothidae, Osmylidae and Nevrorthidae (Grebennikov, 2004), suggesting a plesiomorphic condition. This part of the neuropteran tree was not robustly resolved. Placement of Ithonidae as sister to Psychopsidae based on USCO datasets is reasonable, with heavily sclerotized head capsules and similar

'gula'-like sclerites (Engel et al., 2018), as well as the dorsal location of the occipital foramen in the head capsule (Winterton et al., 2018). Nymphidae as sister group to the rest of Myrmeleontoidea except Psychopsidae and Ithonidae based on USCO and mitogenome datasets was supported by most gene trees in the USCO, UCE and mitogenome datasets, which is congruent with some morphological characters, such as the presence of toothed jaws and dolichasterine setae (Badano et al., 2017; Badano et al., 2018; Jandausch et al., 2019).

After comparing the phylogenetic trees reconstructed based on the three types of genomic data, combining the evidence of morphology and biological habits, the phylogenetic tree constructed based on the amino acid sequences of the USCO dataset may be more credible. Inferring phylogenetic trees from nucleotide sequences can be misleading due to a number of factors including substitutional saturation (Duchene et al., 2022; Philippe et al., 2011), GC content heterogeneity (Regier et al., 2010; Rota-Stabelli et al., 2013), as well as codon-usage bias (Parvathy et al., 2022; Stenoien, 2005). Thus, amino acids may be preferred to nucleotides in phylogenomic analyses of ancient relationships based on the comparative analyses of both nucleotide and amino acid sequence data, such as beetles (McKenna et al., 2019; Vasilikopoulos et al., 2019) and hexapods (Schwentner et al., 2017). Besides, the mitogenome may provide limited information compared with the nuclear genome (Campbell et al., 2022; Roe & Sperling, 2007; Teske et al., 2018; Toews & Brelsford, 2012), with strong base compositional and mutational rate heterogeneity (Jiang et al., 2017; Song et al., 2018; Song, Li, et al., 2016; Tian et al., 2023; Wang et al., 2017). In fact, some studies suggested that the mitogenome is ineffective at some deep nodes (Song, Li, Zhai, et al., 2019; Talavera & Vila, 2011). The application of UCE and mitogenome data, particularly in the higher-level phylogenomic studies, requires careful consideration and methodological scrutiny. More studies incorporating combined morphological and molecular datasets to further validate and refine the phylogenetic relationships inferred here should be conducted in the future.

Impact of concatenation and multi-species coalescent approaches, model selection and data filtering on phylogenetic inference

In most cases, the results of the same analytical approach with different matrices were mostly consistent in this study. Topologies generated based on concatenated ML analyses may suffer from systematic errors due to potential problems such as stochastic sampling error, horizontal gene transfer, gene duplication, ILS, excessive heterogeneity and LBA (Kubatko & Degnan, 2007; Susko & Roger, 2021; Zhou et al., 2017). Most concatenated partitioning ML analyses with USCOs matrices generated topology T1 except for USCO_I_abs85 and USCO_IV matrices. Partitioning attempts take into account differences in the rate and pattern of evolution among loci by estimating independent molecular evolutionary models for a subset of loci that are thought to evolve in a similar way, but there are still some limitations resulting from among-site compositional heterogeneity (Cai et al., 2023; Kumar et al., 2012; Lanfear et al., 2017). Most ML analyses using the Dayhoff6 recoding scheme, except for USCO_II_85, USCO_III_75_abs75, USCO_III_85_abs85 and USCO_IV matrices, recovered topology T2. Phylogenetic analyses using the Dayhoff6 recoding scheme can further reduce compositional heterogeneity, but the method may also cause the reduction of phylogenetic signal (Feuda et al., 2017; Giacomelli et al., 2022; Hernandez & Ryan, 2021; Laumer et al., 2019; Phillips et al., 2004). For all the 16 USCO matrices in this study, the RCFV values of the original matrices and recoded matrices, as well as the ABS for reconstructing phylogenetic trees, were calculated. The recoded matrices displayed lower RCFV values than original matrices (16/16), while most of the ABS values were reduced (12/16) (Figure S7). Most ML analyses under the PMSF model, except for the USCO_II_85 matrix, recovered topology T3 that was significantly supported by the topology tests. The PMSF model has been used in other studies to reduce sequence heterogeneity and produced more plausible phylogenetic inferences (Dong et al., 2022; Rinke et al., 2021; Wang et al., 2018; Yu et al., 2022). Therefore, the selection of appropriate models is crucial in phylogenetic reconstruction and should be routinely employed.

For USCOs and UCEs, a multi-species coalescent approach was also used to reconstruct the phylogenetic tree, which is considered to be highly robust when the ILS level is high and the number of gene trees is sufficiently large (Mirarab et al., 2014; Zhang et al., 2018). However, the multi-species coalescent approach yielded family-level topologies (with relatively low support for some nodes) based on USCO matrices that are not consistent with concatenated ML analyses in this study (Figure S2). The discordance between ASTRAL (coalescent) and concatenation results, such as the paraphyly of Myrmeleontidae, highlights potential challenges in resolving contentious nodes. While this study did not explicitly explore whether these conflicts stem from ILS or model misspecification in gene tree estimation, such factors are known to contribute to topological inconsistencies in phylogenomic analyses (Degnan & Rosenberg, 2009; Gillung et al., 2018; Murillo-Ramos et al., 2023; Owen & Miller, 2022; Yu et al., 2024; Zhong et al., 2013). Related studies incorporating more sophisticated methods to disentangle these effects, such as explicit ILS modelling or improved gene tree estimation approaches, should be conducted to provide further insights into these conflicts in the future. Herein, the multi-species coalescent approach yielded family-level topologies that are not consistent with concatenated ML analyses using UCE matrices filtered by Strategy III and IV, both with shallow nodes poorly resolved (Figure S3). Furthermore, we performed six independent replicates of gene tree reconstruction, which yielded inconsistent ASTRAL tree topologies (Figure 3). More advanced models may be needed to infer the gene trees used to infer the ASTRAL trees. We recommend that concatenation and the multi-species coalescent approach always be used in combination.

Some studies have suggested that genes with high ABS values have stronger phylogenetic signals (Salichos & Rokas, 2013; Wang et al., 2019). However, the ABS values are unstable, with approximately 98.92% of the variation being greater than 0, indicating that

the results are not repeatable in this study. The reasons and impacts of this are not well understood. In phylogenetics, bootstrap was introduced to assess the confidence in estimated phylogenetic trees (Felsenstein, 1985). Subsequently, a series of studies have discussed the meaning of bootstrap support (also known as the bootstrap percentage), the nature of its bias and corrections for the bias (Sanderson & Wojciechowski, 2000; Simon, 2022). When data are analysed under the correct model, nonparametric bootstrapping is conservative (Erixon et al., 2003). However, the single filtering strategy using ABS did not work well because some individual genes had very large ABS variations that affected matrix generation. In addition, using the same genes, the ASTRAL trees derived from tree-building based on the ancestry method were inconsistent. More attention needs to be paid to the potential causes and effects of ABS variation, using more advanced models for tree building, or in combination with other filtering methods (e.g., compositional heterogeneity) in the future.

Impact of calibration strategy on estimates of divergence times

Extensive empirical studies have repeatedly shown that the mitogenome tends to overestimate divergence times compared with results from more slowly evolving nuclear genes in the absence of younger effective calibration points (Near et al., 2012; Shen, Liang, et al., 2016; Zheng et al., 2011). However, the mitogenomes and UCEs are shown to underestimate the divergence time compared with USCOs based on our results. This result is plausible because the topologies derived from mitogenomes and UCEs are inconsistent and there are evolutionary rate differences. Future studies should explicitly incorporate evolutionary rate heterogeneity when integrating nuclear and mitochondrial genomic data in phylogenetic analyses. The divergence times of most nodes close to the root estimated using the USCO datasets would be earlier than those estimated using the UCE datasets. Both are slightly earlier than those estimated using the mitogenome data, elucidating that data type also affects divergence time estimation. The divergence time estimations appear to be sensitive to the calibration position in the phylogenetic tree (Duchene et al., 2014; Ho, 2007; Mello & Schrago, 2014; Near & Sanderson, 2004; Püschel et al., 2019; Rannala & Yang, 2007; Sauquet et al., 2012; Villaverde et al., 2021; Yang & Rannala, 2006). In general, calibrations at the root or at deeper nodes were preferred over those at shallower nodes (Duchene et al., 2014; Ho, 2007; Mello & Schrago, 2014; Püschel et al., 2019; Rannala & Yang, 2007; Sauquet et al., 2012; Yang & Rannala, 2006). Previous studies confirmed that increasing the number of fossil calibrations improves the reliability of divergence time estimation (Duchene et al., 2014; Ho & Phillips, 2009; Hug & Roger, 2007; Linder et al., 2005; Lukoschek et al., 2012; Marshall, 2008; Villaverde et al., 2021). In this study, using fossil calibration nodes close to the root may improve the accuracy of divergence time estimation, while increasing the number of fossil calibrations for shallow nodes did not have this effect. Estimates of substitution rates are largely based on branches located between the calibration node and the tips, so deeper calibrations capture a greater proportion of the overall genetic variation (Duchene et al., 2014). Indeed, the selection and calibration of fossils has always been a difficult problem and is worth further exploration.

CONCLUSIONS

The higher-level phylogeny of Neuropterida was reconstructed using three molecular markers (USCOs, UCEs and mitogenomes) under strictly controlled sampling conditions. The topology recovered by amino acid matrices of USCOs under site-heterogeneous models based on a multi-criteria filtering strategy was believed to be the most robust and in line with morphological and biological evidence. Notably, the ABS values were not stable and were highly variable for many genes, suggesting that this alone is not a reliable filtering strategy for genes. It is necessary to improve the ABS-based filtering strategy in future studies, such as using more advanced models to reconstruct individual gene trees. Divergence time estimation indicated that Neuropterida originated at the start of the Carboniferous, which is older than most studies have previously estimated. The divergence time of key taxa in the evolution of Neuropterida inferred using the USCOs dataset is more ancient than that inferred using the UCE and mitogenome datasets. Calibration points at deeper nodes close to the phylogenetic root may help to accurately estimate evolutionary timescales. These findings underscore the importance of employing multiple analytical strategies and diverse data types in phylogenomic studies to optimize phylogenetic inference and mitigate potential systematic errors.

AUTHOR CONTRIBUTIONS

Ruyue Zhang: Writing – original draft; formal analysis; data curation; writing – review and editing. **Liming Wang:** Formal analysis; data curation; visualization; writing – review and editing. **Shuo Tian:** Formal analysis; data curation; visualization. **Yang Liu:** Formal analysis; data curation. **Yunlan Jiang:** Data curation; visualization. **Xiaofan Zhou:** Writing – review and editing; methodology. **Ding Yang:** Writing – review and editing; resources; supervision. **Xingyue Liu:** Writing – review and editing; supervision; resources. **Yuyu Wang:** Writing – review and editing; project administration; conceptualization; funding acquisition; investigation; resources.

ACKNOWLEDGEMENTS

We sincerely appreciate Dr. Shaun L. Winterton for his help in collecting and identifying specimens, as well as critically reading and editing the draft of this manuscript. We are also grateful to the Editor-in-chief Dr. Bonnie Blaimer and two anonymous reviewers for reviewing the manuscript and providing very helpful advice that significantly improved the manuscript.

FUNDING INFORMATION

This work was supported by the National Natural Science Foundation of China (32170451, 32130012, 32170448), the Basic Research

Project of Shijiazhuang for University in Hebei Province (241791137A), the Science and Technology Planning Project of Baoding (2472P018), the Project of Basic Scientific Research Business Expenses of Hebei Provincial Universities (KY2023061), the Institute Go-Sailing Program (ShengKeChuang-QA) and the Earmarked Fund for CARS-27.

CONFLICT OF INTEREST STATEMENT

No potential conflicts of interest were reported by the authors.

DATA AVAILABILITY STATEMENT

The genome data underlying this article are available in the GenBank Nucleotide Database (PRJNA1132765) and can be accessed with the accession number listed in Supplementary Table S1. The molecular matrices and the tree files in this study are available on Dryad (doi: 10.5061/dryad.6hdr7srbt).

ORCID

Yunlan Jiang https://orcid.org/0009-0007-1644-0915
Ding Yang https://orcid.org/0000-0002-7685-3478
Xingyue Liu https://orcid.org/0000-0002-9168-0659
Yuyu Wang https://orcid.org/0000-0002-8645-303X

REFERENCES

Abadi, S., Azouri, D., Pupko, T. & Mayrose, I. (2019) Model selection may not be a mandatory step for phylogeny reconstruction. *Nature Communications*, 10, 1–11.

Allen, J.M., Boyd, B., Nguyen, N.P., Vachaspati, P., Warnow, T., Huang, D.I. et al. (2017) Phylogenomics from whole genome sequences using aTRAM. *Systematic Biology*, 66, 786–798.

Aspöck, U., Aspöck, H. & Liu, X. (2017) The Nevrorthidae, mistaken at all times: phylogeny and review of present knowledge (Holometabola, Neuropterida, Neuroptera). *Deutsche Entomologische Zeitschrift*, 64, 77–110.

Aspöck, U., Plant, J.D. & Nemeschkal, H.L. (2001) Cladistic analysis of Neuroptera and their systematic position within Neuropterida (Insecta: Holometabola: Neuropterida: Neuroptera). *Systematic Entomology*, 26, 73–86.

Badano, D., Aspöck, U., Aspöck, H. & Cerretti, P. (2017) Phylogeny of Myrmeleontiformia based on larval morphology (Neuropterida: Neuroptera). *Systematic Entomology*, 42, 94–117.

Badano, D., Engel, M.S., Basso, A., Wang, B. & Cerretti, P. (2018) Diverse cretaceous larvae reveal the evolutionary and behavioural history of antlions and lacewings. *Nature Communications*, 9, 1–14.

Ban, X.C., Shao, Z.K., Wu, L.J., Sun, J.T. & Xue, X.F. (2022) Highly diversified mitochondrial genomes provide new evidence for interordinal relationships in the Arachnida. *Cladistics*, 38, 452–464.

Behura, S.K. (2015) Insect phylogenomics. *Insect Molecular Biology*, 24, 403–411.

Bernt, M., Donath, A., Juhling, F., Externbrink, F., Florentz, C., Fritzsch, G. et al. (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, 69, 313–319.

Beutel, R.G., Friedrich, F. & Aspöck, U. (2010) The larval head of Nevrorthidae and the phylogeny of Neuroptera (Insecta). *Zoological Journal of the Linnean Society*, 158, 533–562.

Boutte, J., Fishbein, M. & Straub, S.C.K. (2022) NGS-Indel coder v2.0: a streamlined pipeline to code indel characters in phylogenomic data. In: Pereira-Santana, A., Gamboa-Tuz, S.D. & Rodríguez-Zapata, L.C. (Eds.) *Plant comparative genomics*. New York, NY: Springer US, pp. 61–72.

Bushnell, B. (2014) BBTools software package. Available from: http://sourceforge.net/projects/bbmap [Accessed 20th June 2023].

Cai, C. (2024) Ant backbone phylogeny resolved by modelling compositional heterogeneity among sites in genomic data. *Communications Biology*, 7, 1–7.

Cai, C., Tihelka, E., Liu, X. & Engel, M.S. (2023) Improved modelling of compositional heterogeneity reconciles phylogenomic conflicts among lacewings. *Palaeoentomology*, 6, 49–57.

Cai, C., Tihelka, E., Pisani, D. & Donoghue, P.C.J. (2020) Data curation and modeling of compositional heterogeneity in insect phylogenomics: a case study of the phylogeny of Dytiscoidea (Coleoptera: Adephaga). *Molecular Phylogenetics and Evolution*, 147, 1–7.

Campbell, E.O., MacDonald, Z.G., Gage, E.V., Gage, R.V. & Sperling, F.A.H. (2022) Genomics and ecological modelling clarify species integrity in a confusing group of butterflies. *Molecular Ecology*, 31, 2400–2417.

Chan, C.X. & Ragan, M.A. (2013) Next-generation phylogenomics. *Biology Direct*, 8, 1–6.

Chen, S., Zhou, Y., Chen, Y. & Gu, J. (2018) Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34, i884–i890.

Chikhi, R. & Rizk, G. (2012) Space-efficient and exact de bruijn graph representation based on a bloom filter, proceedings of the 12th international conference on algorithms in bioinformatics. *Algorithms for Molecular Biology*, 8, 1–9.

Dai, M., He, S.L., Chen, B. & Li, T.J. (2023) Phylogeny of *Rhynchium* and its related genera (hymenoptera: Eumeninae) based on universal single-copy orthologs and ultraconserved elements. *Insects*, 14, 1–14.

Degnan, J.H. & Rosenberg, N.A. (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24, 332–340.

Dierckxsens, N., Mardulyn, P. & Smits, G. (2017) NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, 45, 1–9.

Dong, W., Li, E., Liu, Y., Xu, C., Wang, Y., Liu, K. et al. (2022) Phylogenomic approaches untangle early divergences and complex diversifications of the olive plant family. *BMC Biology*, 20, 1–25.

Drummond, A.J., Ho, S.Y., Phillips, M.J. & Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, 4, 699–710.

Duchene, D.A., Mather, N., van der Wal, C. & Ho, S.Y.W. (2022) Excluding loci with substitution saturation improves inferences from phylogenomic data. *Systematic Biology*, 71, 676–689.

Duchene, S., Lanfear, R. & Ho, S.Y. (2014) The impact of calibration and clock-model choice on molecular estimates of divergence times. *Molecular Phylogenetics and Evolution*, 78, 277–289.

Edwards, S.V. (2009) Is a new and general theory of molecular systematics emerging? *Evolution*, 63, 1–19.

Engel, M.S., Winterton, S.L. & Breitkreuz, L.C.V. (2018) Phylogeny and evolution of Neuropterida: where have wings of lace taken us? *Annual Review of Entomology*, 63, 531–551.

Erixon, P., Svennblad, B., Britton, T. & Oxelman, B. (2003) Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Systematic Biology*, 52, 665–673.

Faircloth, B.C. (2016) PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, 32, 786–788.

Felsenstein, J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39, 783–791.

Fernandez, R., Edgecombe, G.D. & Giribet, G. (2016) Exploring phylogenetic relationships within Myriapoda and the effects of matrix composition and occupancy on phylogenomic reconstruction. *Systematic Biology*, 65, 871–889.

Feuda, R., Dohrmann, M., Pett, W., Philippe, H., Rota-Stabelli, O., Lartillot, N. et al. (2017) Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. *Current Biology*, 27, 3864–3870.

Foster, P.G. (2004) Modeling compositional heterogeneity. *Systematic Biology*, 53, 485–495.
Gao, J.W., Yuan, X.P., Wu, H., Xiang, C.Y., Xie, M., Song, R. et al. (2022) Mitochondrial phylogenomics of Acanthocephala: nucleotide alignments produce long-branch attraction artefacts. *Parasites & Vectors*, 15, 1–12.
Garzón-Orduña, I.J., Menchaca-Armenta, I., Contreras-Ramos, A., Liu, X. & Winterton, S.L. (2016) The phylogeny of brown lacewings (Neuroptera: Hemerobiidae) reveals multiple reductions in wing venation. *BMC Evolutionary Biology*, 16, 1–19.
Giacomelli, M., Rossi, M.E., Lozano-Fernandez, J., Feuda, R. & Pisani, D. (2022) Resolving tricky nodes in the tree of life through amino acid recoding. *iScience*, 25, 1–23.
Gillung, J.P., Winterton, S.L., Bayless, K.M., Khouri, Z., Borowiec, M.L., Yeates, D. et al. (2018) Anchored phylogenomics unravels the evolution of spider flies (Diptera, Acroceridae) and reveals discordance between nucleotides and amino acids. *Molecular Phylogenetics and Evolution*, 128, 233–245.
Grebennikov, V.V. (2004) Grub-like larvae of Neuroptera (Insecta): a morphological review of the families Ithonidae and Polystoechotidae and a description of Oliarces clara. *European Journal of Entomology*, 101, 409–417.
Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59, 307–321.
Hernandez, A.M. & Ryan, J.F. (2021) Six-state amino acid recoding is not an effective strategy to offset compositional heterogeneity and saturation in phylogenetic analyses. *Systematic Biology*, 70, 1200–1212.
Ho, S.Y. & Duchene, S. (2014) Molecular-clock methods for estimating evolutionary rates and timescales. *Molecular Ecology*, 23, 5947–5965.
Ho, S.Y. & Phillips, M.J. (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, 58, 367–380.
Ho, S.Y.W. (2007) Calibrating molecular estimates of substitution rates and divergence times in birds. *Journal of Avian Biology*, 38, 409–414.
Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q. & Vinh, L.S. (2018) UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35, 518–522.
Huang, W., Li, L., Myers, J.R. & Marth, G.T. (2012) ART: a next-generation sequencing read simulator. *Bioinformatics*, 28, 593–594.
Hug, L.A. & Roger, A.J. (2007) The impact of fossils and taxon sampling on ancient molecular dating analyses. *Molecular Biology and Evolution*, 24, 1889–1897.
Inoue, J., Donoghue, P.C. & Yang, Z. (2010) The impact of the representation of fossil calibrations on Bayesian estimation of species divergence times. *Systematic Biology*, 59, 74–89.
Jandausch, K., Beutel, R.G. & Bellstedt, R. (2019) The larval morphology of the spongefly *Sisyra nigra* (Retzius, 1783) (Neuroptera: Sisyridae). *Journal of Morphology*, 280, 1742–1758.
Jiang, Y., Garzón-Orduña, I.J., Winterton, S.L., Yang, F. & Liu, X. (2017) Phylogenetic relationships among tribes of the green lacewing subfamily Chrysopinae recovered based on mitochondrial phylogenomics. *Scientific Reports*, 7, 1–9.
Kapli, P., Flouri, T. & Telford, M.J. (2021) Systematic errors in phylogenetic trees. *Current Biology*, 31, R51–R66.
Kocot, K.M., Struck, T.H., Merkel, J., Waits, D.S., Todt, C., Brannock, P.M. et al. (2017) Phylogenomics of Lophotrochozoa with consideration of systematic error. *Systematic Biology*, 66, 256–282.
Kosiol, C., Goldman, N. & Buttimore, N.H. (2004) A new criterion and method for amino acid classification. *Journal of Theoretical Biology*, 228, 97–106.
Kovacs, T.G.L., Walker, J., Hellemans, S., Bourguignon, T., Tatarnic, N.J., McRae, J.M. et al. (2024) Dating in the dark: elevated substitution rates in cave cockroaches (Blattodea: Nocticolidae) have negative impacts on molecular date estimates. *Systematic Biology*, 73, 532–545.
Kubatko, L.S. & Degnan, J.H. (2007) Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, 56, 17–24.
Kück, P. & Meusemann, K. (2010) FASconCAT: convenient handling of data matrices. *Molecular Phylogenetics and Evolution*, 56, 1115–1118.
Kück, P. & Struck, T.H. (2014) BaCoCa-a heuristic software tool for the parallel assessment of sequence biases in hundreds of gene and taxon partitions. *Molecular Phylogenetics and Evolution*, 70, 94–98.
Kumar, S., Filipski, A.J., Battistuzzi, F.U., Kosakovsky Pond, S.L. & Tamura, K. (2012) Statistics and truth in phylogenomics. *Molecular Biology and Evolution*, 29, 457–472.
Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2017) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34, 772–773.
Lartillot, N., Brinkmann, H. & Philippe, H. (2007) Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. *BMC Evolutionary Biology*, 7, 1–14.
Lartillot, N., Lepage, T. & Blanquart, S. (2009) PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*, 25, 2286–2288.
Lartillot, N. & Philippe, H. (2004) A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Molecular Biology and Evolution*, 21, 1095–1109.
Laumer, C.E., Fernandez, R., Lemer, S., Combosch, D., Kocot, K.M., Riesgo, A. et al. (2019) Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proceedings of the Royal Society B: Biological Sciences*, 286, 1–10.
Lepage, T., Bryant, D., Philippe, H. & Lartillot, N. (2007) A general comparison of relaxed molecular clock models. *Molecular Biology and Evolution*, 24, 2669–2680.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N. et al. (2010) The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25, 2078–2079.
Li, H.T., Yi, T.S., Gao, L.M., Ma, P.F., Zhang, T., Yang, J.B. et al. (2019) Origin of angiosperms and the puzzle of the Jurassic gap. *Nature Plants*, 5, 461–470.
Li, Y.D., Engel, M.S., Tihelka, E. & Cai, C. (2023) Phylogenomics of weevils revisited: data curation and modelling compositional heterogeneity. *Biology Letters*, 19, 1–8.
Linder, H.P., Hardy, C.R. & Rutschmann, F. (2005) Taxon sampling effects in molecular clock dating: an example from the African Restionaceae. *Molecular Phylogenetics and Evolution*, 35, 569–582.
Liu, G., Pan, Q., Dai, Y., Wang, X., Li, M., Zhu, P. et al. (2024) Phylogenomics of Afrotherian mammals and improved resolution of extant Paenungulata. *Molecular Phylogenetics and Evolution*, 195, 108047.
Liu, L., Wu, S. & Yu, L. (2015) Coalescent methods for estimating species trees from phylogenomic data. *Journal of Systematics and Evolution*, 53, 380–390.
Lu, X., Lin, A., Wang, D. & Liu, X. (2021) The first mitochondrial genome of spongillafly from Asia (Neuroptera: Sisyridae: *Sisyra aurorae* Navas, 1933) and phylogenetic implications of Osmyloidea. *Mitochondrial DNA Part B-Resources*, 6, 2369–2370.
Lu, X., Xia, F., Wang, B., Aspöck, U. & Liu, X. (2018) Taxonomic notes on *Cretarophalis patrickmuelleri* Wichard, 2017 (Insecta: Neuroptera: Nevrorthidae) from the mid-cretaceous of Myanmar, and its phylogenetic significance. *Zootaxa*, 4370, 591–600.
Lukoschek, V., Scott Keogh, J. & Avise, J.C. (2012) Evaluating fossil calibrations for dating phylogenies in light of rates of molecular evolution: a comparison of three approaches. *Systematic Biology*, 61, 22–43.
Lunter, G. & Goodson, M. (2011) Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Research*, 21, 936–939.

Luo, A., Zhang, C., Zhou, Q.S., Ho, S.Y.W. & Zhu, C.D. (2023) Impacts of taxon-sampling schemes on Bayesian tip dating under the fossilized birth-death process. *Systematic Biology*, 72, 781–801.

Mai, U. & Mirarab, S. (2018) TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics*, 19, 23–40.

Marcais, G. & Kingsford, C. (2011) A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, 27, 764–770.

Marshall, C.R. (2008) A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. *American Naturalist*, 171, 726–742.

Mbanyana, N., Blaimer, B.B., le Roux, J.J., van Noort, S., Brady, S.G. & Wossler, T.C. (2024) Out of the desert: Paleoclimatic changes drove the diversification of arid-adapted Ocymyrmex ants in southern Africa. *Molecular Phylogenetics and Evolution*, 191, 1–10.

McKenna, D.D., Shin, S., Ahrens, D., Balke, M., Beza-Beza, C., Clarke, D.J. et al. (2019) The evolution and genomic basis of beetle diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 24729–24737.

Mello, B. & Schrago, C.G. (2014) Assignment of calibration information to deeper phylogenetic nodes is more effective in obtaining precise and accurate divergence time estimates. *Evolutionary Bioinformatics*, 10, 79–85.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. et al. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37, 1530–1534.

Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S. & Warnow, T. (2014) ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30, i541–i548.

Moles, J., Cunha, T.J., Lemer, S., Combosch, D.J. & Giribet, G. (2021) Tightening the girdle: phylotranscriptomics of Polyplacophora. *Journal of Molluscan Studies*, 87(2), 1–5. Available from: https://doi.org/10.1093/mollus/eyab019

Murillo-Ramos, L., Twort, V., Wahlberg, N. & Sihvonen, P. (2023) A phylogenomic perspective on the relationships of subfamilies in the family Geometridae (Lepidoptera). *Systematic Entomology*, 48, 618–632.

Nakamine, H., Yamamoto, S., Takahashi, Y. & Liu, X. (2023) A remarkable new genus of Nevrorthidae (Neuroptera, Osmyloidea) from mid-cretaceous Kachin amber of northern Myanmar. *Deutsche Entomologische Zeitschrift*, 70, 113–120.

Naser-Khdour, S., Minh, B.Q., Zhang, W., Stone, E.A. & Lanfear, R. (2019) The prevalence and impact of model violations in phylogenetic analysis. *Genome Biology and Evolution*, 11, 3341–3352.

Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P. et al. (2012) Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 13698–13703.

Near, T.J. & Sanderson, M.J. (2004) Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 359, 1477–1483.

Nguyen, J.M.T. & Ho, S.Y.W. (2020) Calibrations from the fossil record. In: Ho, S.Y.W. (Ed.) *The molecular evolutionary clock: theory and practice*. Cham: Springer International Publishing, pp. 117–133.

O'Reilly, J.E. & Donoghue, P.C.J. (2020) The effect of fossil sampling on the estimation of divergence times with the fossilized birth-death process. *Systematic Biology*, 69, 124–138.

Owen, C.L. & Miller, G.L. (2022) Phylogenomics of the Aphididae: deep relationships between subfamilies clouded by gene tree discordance, introgression and the gene tree anomaly zone. *Systematic Entomology*, 47, 470–486.

Parks, M.B., Wickett, N.J. & Alverson, A.J. (2018) Signal, uncertainty, and conflict in phylogenomic data for a diverse lineage of microbial eukaryotes (diatoms, Bacillariophyta). *Molecular Biology and Evolution*, 35, 80–93.

Parvathy, S.T., Udayasuriyan, V. & Bhadana, V. (2022) Codon usage bias. *Molecular Biology Reports*, 49, 539–565.

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T., Manuel, M., Worheide, G. et al. (2011) Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biology*, 9, 1–10.

Phillips, M.J., Delsuc, F. & Penny, D. (2004) Genome-scale phylogeny and the detection of systematic biases. *Molecular Biology and Evolution*, 21, 1455–1458.

Prasanna, A.N., Gerber, D., Kijpornyongpan, T., Aime, M.C., Doyle, V.P. & Nagy, L.G. (2020) Model choice, missing data, and taxon sampling impact phylogenomic inference of deep basidiomycota relationships. *Systematic Biology*, 69, 17–37.

Pryszcz, L.P. & Gabaldón, T. (2016) Redundans: an assembly pipeline for highly heterozygous genomes. *Nucleic Acids Research*, 44, 1–10.

Püschel, H.P., Bertrand, O.C., O'Reilly, J.E., Bobe, R. & Püschel, T.A. (2021) Divergence-time estimates for hominins provide insight into encephalization and body mass trends in human evolution. *Nature Ecology & Evolution*, 5, 808–819.

Püschel, H.P., O'Reilly, J.E., Pisani, D., Donoghue, P.C.J. & Rahman, I. (2019) The impact of fossil stratigraphic ranges on tip-calibration, and the accuracy and precision of divergence time estimates. *Palaeontology*, 63, 67–83.

Quinlan, A.R. & Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26, 841–842.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Systematic Biology*, 67, 901–904.

Randolf, S., Zimmermann, D. & Aspoeck, U. (2014) Head anatomy of adult *Nevrorthus apatelios* and basal splitting events in Neuroptera (Neuroptera: Nevrorthidae). *Arthropod Systematics & Phylogeny*, 72, 111–136.

Rannala, B. & Yang, Z. (2007) Inferring speciation times under an episodic molecular clock. *Systematic Biology*, 56, 453–466.

Regier, J.C., Shultz, J.W., Zwick, A., Hussey, A., Ball, B., Wetzer, R. et al. (2010) Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature*, 463, 1079–1083.

Rinke, C., Chuvochina, M., Mussig, A.J., Chaumeil, P.A., Davín, A.A., Waite, D.W. et al. (2021) A standardized archaeal taxonomy for the genome taxonomy database. *Nature Microbiology*, 6, 946–959.

Roe, A.D. & Sperling, F.A. (2007) Patterns of evolution of mitochondrial cytochrome c oxidase I and II DNA and implications for DNA barcoding. *Molecular Phylogenetics and Evolution*, 44(1), 325–345. Available from: https://doi.org/10.1016/j.ympev.2006.12.005

Romiguier, J., Cameron, S.A., Woodard, S.H., Fischman, B.J., Keller, L. & Praz, C.J. (2016) Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. *Molecular Biology and Evolution*, 33, 670–678.

Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19, 1572–1574.

Rota-Stabelli, O., Lartillot, N., Philippe, H. & Pisani, D. (2013) Serine codon-usage bias in deep phylogenomics: pancrustacean relationships as a case study. *Systematic Biology*, 62, 121–133.

Rozewicki, J., Li, S., Amada, K.M., Standley, D.M. & Katoh, K. (2019) MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Research*, 47, W5–W10.

Sahlin, K., Vezzi, F., Nystedt, B., Lundeberg, J. & Arvestad, L. (2014) BESST- efficient scaffolding of large fragmented assemblies. *BMC Bioinformatics*, 15, 1–11.

Salichos, L. & Rokas, A. (2013) Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature*, 497, 327–331.

Sanderson, M.J. & Wojciechowski, M.F. (2000) Improved bootstrap confidence limits in large-scale phylogenies, with an example from neo-Astragalus (Leguminosae). *Systematic Biology*, 49, 671–685.

Sauquet, H., Ho, S.Y., Gandolfo, M.A., Jordan, G.J., Wilf, P., Cantrill, D.J. et al. (2012) Testing the impact of calibration on molecular divergence times using a fossil-rich group: the case of *Nothofagus* (Fagales). *Systematic Biology*, 61, 289–313.

Sayyari, E., Whitfield, J.B. & Mirarab, S. (2018) DiscoVista: interpretable visualizations of gene tree discordance. *Molecular Phylogenetics and Evolution*, 122, 110–115.

Schwentner, M., Combosch, D.J., Pakes Nelson, J. & Giribet, G. (2017) A phylogenomic solution to the origin of insects by resolving crustacean-hexapod relationships. *Current Biology*, 27, 1818–1824.

Shen, X.X., Hittinger, C.T. & Rokas, A. (2017) Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nature Ecology & Evolution*, 1, 1–23.

Shen, X.X., Liang, D., Chen, M.Y., Mao, R.L., Wake, D.B. & Zhang, P. (2016) Enlarged multilocus data set provides surprisingly younger time of origin for the Plethodontidae, the largest family of salamanders. *Systematic Biology*, 65, 66–81.

Shen, X.X., Salichos, L. & Rokas, A. (2016) A genome-scale investigation of how sequence, function, and tree-based gene properties influence phylogenetic inference. *Genome Biology and Evolution*, 8, 2565–2580.

Simion, P., Philippe, H., Baurain, D., Jager, M., Richter, D.J., Di Franco, A. et al. (2017) A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. *Current Biology*, 27, 958–967.

Simon, C. (2022) An evolving view of phylogenetic support. *Systematic Biology*, 71, 921–928.

Sinaiko, G., Cao, Y. & Dietrich, C.H. (2024) Phylogenomics of the leafhopper genus *Neoaliturus* distant, 1918 (Hemiptera: Cicadellidae: Deltocephalinae) reveals genetically divergent lineages in the invasive beet leafhopper. *Molecular Phylogenetics and Evolution*, 195, 1–11.

Song, F., Li, H., Jiang, P., Zhou, X., Liu, J., Sun, C. et al. (2016) Capturing the phylogeny of Holometabola with mitochondrial genome data and Bayesian site-heterogeneous mixture models. *Genome Biology and Evolution*, 8, 1411–1426.

Song, L., Florea, L. & Langmead, B. (2014) Lighter: fast and memory-efficient sequencing error correction without counting. *Genome Biology*, 15, 1–13.

Song, N., An, S.H., Yin, X.M., Zhao, T. & Wang, X.Y. (2016) Insufficient resolving power of mitogenome data in deciphering deep phylogeny of Holometabola. *Journal of Systematics and Evolution*, 54, 545–559.

Song, N., Li, X., Yin, X., Li, X., Yin, J. & Pan, P. (2019) The mitochondrial genomes of palaeopteran insects and insights into the early insect relationships. *Scientific Reports*, 9, 1–11.

Song, N., Li, X.X., Zhai, Q., Bozdogan, H. & Yin, X.M. (2019) The mitochondrial genomes of Neuropteridan insects and implications for the phylogeny of Neuroptera. *Genes*, 10, 1–19.

Song, N., Lin, A. & Zhao, X. (2018) Insight into higher-level phylogeny of Neuropterida: evidence from secondary structures of mitochondrial rRNA genes and mitogenomic data. *PLoS One*, 13, 1–21.

Steenwyk, J.L., Buida, T.J., 3rd, Li, Y., Shen, X.X. & Rokas, A. (2020) ClipKIT: a multiple sequence alignment trimming software for accurate phylogenomic inference. *PLoS Biology*, 18(12), 1–17. Available from: https://doi.org/10.1371/journal.pbio.3001007

Steenwyk, J.L., Buida, T.J., Labella, A.L., Li, Y., Shen, X.X. & Rokas, A. (2021) PhyKIT: a broadly applicable UNIX shell toolkit for processing and analyzing phylogenomic data. *Bioinformatics*, 37, 2325–2331.

Steenwyk, J.L., Li, Y., Zhou, X., Shen, X.X. & Rokas, A. (2023) Incongruence in the phylogenomics era. *Nature Reviews Genetics*, 24, 834–850.

Stenoien, H.K. (2005) Adaptive basis of codon usage in the haploid moss Physcomitrella patens. *Heredity*, 94, 87–93.

Susko, E. & Roger, A.J. (2007) On reduced amino acid alphabets for phylogenetic inference. *Molecular Biology and Evolution*, 24, 2139–2150.

Susko, E. & Roger, A.J. (2021) Long branch attraction biases in phylogenetics. *Systematic Biology*, 70, 838–843.

Szantho, L.L., Lartillot, N., Szollosi, G.J. & Schrempf, D. (2023) Compositionally constrained sites drive long-branch attraction. *Systematic Biology*, 72, 767–780.

Talavera, G. & Castresana, J. (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, 56, 564–577.

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 Evolutionary Biology*, 11, 1–15.

Tamura, K., Tao, Q. & Kumar, S. (2018) Theoretical foundation of the RelTime method for estimating divergence times from variable evolutionary rates. *Molecular Biology and Evolution*, 35, 1770–1782.

Tanabe, A.S. (2008) Phylogears, 2008. Available from: http://www.fifthdimension.jp/ [Accessed 20th October 2023].

Tao, Q., Barba-Montoya, J., Huuki, L.A., Durnan, M.K. & Kumar, S. (2020) Relative efficiencies of simple and complex substitution models in estimating divergence times in phylogenomics. *Molecular Biology and Evolution*, 37, 1819–1831.

Teske, P.R., Golla, T.R., Sandoval-Castillo, J., Emami-Khoyi, A., van der Lingen, C.D., von der Heyden, S. et al. (2018) Mitochondrial DNA is unsuitable to test for isolation by distance. *Scientific Reports*, 8, 1–9.

Thompson, J.D., Gibson, T.J. & Higgins, D.G. (2003) Multiple sequence alignment using ClustalW and ClustalX. *Current Protocols in Bioinformatics*, 2003, 2.3.1–2.3.22.

Tian, S., Jiang, Y., Lai, Y., Wang, S., Liu, X. & Wang, Y. (2023) New mitogenomes of the green lacewing tribe Ankylopterygini (Neuroptera: Chrysopidae: Chrysopinae) and phylogenetic implications of Chrysopidae. *Insects*, 14, 1–16.

Toews, D.P. & Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, 21, 3907–3930.

Vasilikopoulos, A., Balke, M., Beutel, R.G., Donath, A., Podsiadlowski, L., Pflug, J.M. et al. (2019) Phylogenomics of the superfamily Dytiscoidea (Coleoptera: Adephaga) with an evaluation of phylogenetic conflict and systematic error. *Molecular Phylogenetics and Evolution*, 135, 270–285.

Vasilikopoulos, A., Misof, B., Meusemann, K., Lieberz, D., Flouri, T., Beutel, R.G. et al. (2020) An integrative phylogenomic approach to elucidate the evolutionary history and divergence times of Neuropterida (Insecta: Holometabola). *BMC Evolutionary Biology*, 20, 1–24.

Villaverde, T., Maguilla, E., Luceño, M. & Hipp, A.L. (2021) Assessing the sensitivity of divergence time estimates to locus sampling, calibration points, and model priors in a RAD-seq phylogeny of *Carex* section *Schoenoxiphium*. *Journal of Systematics and Evolution*, 59, 687–697.

Vurture, G.W., Sedlazeck, F.J., Nattestad, M., Underwood, C.J., Fang, H., Gurtowski, J. et al. (2017) GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics*, 33, 2202–2204.

Wachi, N., Matsubayashi, K.W. & Maeto, K. (2017) Application of next-generation sequencing to the study of non-model insects. *Entomological Science*, 21, 3–11.

Wang, H.C., Minh, B.Q., Susko, E. & Roger, A.J. (2018) Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Systematic Biology*, 67, 216–235.

Wang, Y., Liu, X., Garzón-Orduña, I.J., Winterton, S.L., Yan, Y., Aspöck, U. et al. (2017) Mitochondrial phylogenomics illuminates the evolutionary history of Neuropterida. *Cladistics*, 33, 617–636.

Wang, Y., Zhou, X., Wang, L., Liu, X., Yang, D. & Rokas, A. (2019) Gene selection and evolutionary modeling affect phylogenomic inference of Neuropterida based on transcriptome data. *International Journal of Molecular Sciences*, 20, 1–17.

Waterhouse, R.M., Seppey, M., Simao, F.A., Manni, M., Ioannidis, P., Klioutchnikov, G. et al. (2018) BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution*, 35, 543–548.

Whelan, N.V., Kocot, K.M., Moroz, T.P., Mukherjee, K., Williams, P., Paulay, G. et al. (2017) Ctenophore relationships and their placement as the sister group to all other animals. *Nature Ecology & Evolution*, 1, 1737–1746.

Winterton, S.L., Hardy, N.B. & Wiegmann, B.M. (2010) On wings of lace: phylogeny and Bayesian divergence time estimates of Neuropterida (Insecta) based on morphological and molecular data. *Systematic Entomology*, 35, 349–378.

Winterton, S.L., Lemmon, A.R., Gillung, J.P., Garzon, I.J., Badano, D., Bakkes, D.K. et al. (2018) Evolution of lacewings and allied orders using anchored phylogenomics (Neuroptera, Megaloptera, Raphidioptera). *Systematic Entomology*, 43, 330–354.

Winterton, S.L., Zhao, J., Garzón-Orduña, I.J., Wang, Y. & Liu, Z. (2017) The phylogeny of lance lacewings (Neuroptera: Osmylidae). *Systematic Entomology*, 42, 555–574.

Wu, Y. (2012) Coalescent-based species tree inference from gene tree topologies under incomplete lineage sorting by maximum likelihood. *Evolution*, 66, 763–775.

Xi, Z., Liu, L., Rest, J.S. & Davis, C.C. (2014) Coalescent versus concatenation methods and the placement of Amborella as sister to water lilies. *Systematic Biology*, 63, 919–932.

Xiong, Z., He, D., Guang, X. & Li, Q. (2023) Novel tRNA gene rearrangements in the mitochondrial genomes of poneroid ants and phylogenetic implication of Paraponerinae (hymenoptera: Formicidae). *Life (Basel)*, 13, 1–17.

Xu, M., Guo, L., Gu, S., Wang, O., Zhang, R., Peters, B.A. et al. (2020) TGS-GapCloser: a fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *GigaScience*, 9(9), 1–11. Available from: https://doi.org/10.1093/gigascience/giaa094

Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24, 1586–1591.

Yang, Z. & Rannala, B. (2006) Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution*, 23, 212–226.

Young, A.D. & Gillung, J.P. (2019) Phylogenomics-principles, opportunities and pitfalls of big-data phylogenetics. *Systematic Entomology*, 45, 225–247. Available from: https://doi.org/10.1111/syen.12406

Yu, D., Ding, Y., Tihelka, E., Cai, C., Hu, F., Liu, M. et al. (2022) Phylogenomics of elongate-bodied springtails reveals independent transitions from aboveground to belowground habitats in deep time. *Systematic Biology*, 71, 1023–1031.

Yu, H.Y., Chu, K.H., Tsang, L.M. & Ma, K.Y. (2024) Incomplete lineage sorting and long-branch attraction confound phylogenomic inference of Pancrustacea. *Frontiers in Ecology and Evolution*, 12, 1–22.

Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. (2018) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, 19, 15–30.

Zhang, D., Niu, Z.Q., Luo, A.R., Orr, M.C., Ferrari, R.R., Jin, J.F. et al. (2022) Testing the systematic status of *Homalictus* and *Rostrohalictus* with weakened cross-vein groups within Halictini (Hymenoptera: Halictidae) using low-coverage whole-genome sequencing. *Insect Science*, 29, 1819–1833.

Zhang, F., Ding, Y., Zhu, C.D., Zhou, X., Orr, M.C., Scheu, S. et al. (2019) Phylogenomics from low-coverage whole-genome sequencing. *Methods in Ecology and Evolution*, 10, 507–517.

Zhang, J., Lin, L., Mu, Y., Brelsford, A. & Purcell, J. (2025) A comparison of phylogenomic inference pipelines for low-coverage whole-genome sequencing in Formica ants. *Systematic Entomology*, 50, 611–629.

Zhao, C., Huang, M., Yang, D. & Liu, X. (2024) Comparative morphology of the wing base structure illuminates higher-level phylogeny of Holometabola. *Insects*, 15, 1–23.

Zhao, C., Liu, X. & Yang, D. (2014) Wing base structural data support the sister relationship of Megaloptera and Neuroptera (Insecta: Neuropterida). *PLoS One*, 9, 1–23.

Zhao, M., Kurtis, S.M., White, N.D., Moncrieff, A.E., Leite, R.N., Brumfield, R.T. et al. (2023) Exploring conflicts in whole genome phylogenetics: a case study within Manakins (Aves: Pipridae). *Systematic Biology*, 72, 161–178.

Zheng, Y., Peng, R., Kuro-o, M. & Zeng, X. (2011) Exploring patterns and extent of bias in estimating divergence time from mitochondrial DNA sequence data in a particular lineage: a case study of salamanders (order Caudata). *Molecular Biology and Evolution*, 28, 2521–2535.

Zhong, B., Liu, L., Yan, Z. & Penny, D. (2013) Origin of land plants using the multispecies coalescent model. *Trends in Plant Science*, 18, 492–495.

Zhou, Y., Duvaux, L., Ren, G., Zhang, L., Savolainen, O. & Liu, J. (2017) Importance of incomplete lineage sorting and introgression in the origin of shared genetic variation between two closely related pines with overlapping distributions. *Heredity*, 118, 211–220.

Zou, H., Jakovlic, I., Zhang, D., Hua, C.J., Chen, R., Li, W.X. et al. (2020) Architectural instability, inverted skews and mitochondrial phylogenomics of isopoda: outgroup choice affects the long-branch attraction artefacts. *Royal Society Open Science*, 7, 1–13.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Characteristics of newly sequenced genomes in this study inferred from GenomeScope. len: genome haploid length; uniq: percentage of the genome unique length; het: heterozygosity; kcov: coverage of hybrid peak; err: read error rate; dup: duplicates; k: k-mer size.

Figure S2. Possible phylogenetic topologies of Neuropterida from different matrices and model based on USCOs. Controversies have centred on the phylogenetic relationships between the families of the Neuropterida (A, T1–T5), the phylogenetic relationships within Chrysopidae (B, a–f), the phylogenetic relationships within the Nymphidae (C, labelled as '&, o and β'), the paraphyly of the Myrmeleontidae relative to Ascalaphidae, and the relationships within Myrmeleontidae (D, labelled as '*, # and ^'). All possible topologies according to different matrices and models (E). Different background colour indicates different family-level relationships of Neuropterida, respectively (T1–T5). Topologies that are confusing and will not be considered are labelled '?'.

Figure S3. Possible phylogenetic topologies of Neuropterida from different matrices and model based on UCEs. Trees in boxes on right highlight the alternative topologies yielded in our study (A, Ti and Tii). Different background colour indicates different family-level relationships of Neuropterida, respectively (C, Ti and Tii). Topology that are confusing and will not be considered are labelled '?'.

Figure S4. Possible phylogenetic topologies of Neuropterida from different matrices and model based on mitochondrial genome data (H1–H6). Different background colour indicates different family-level relationships of Neuropterida, respectively (H1–H6). ML, Maximum likelihood, BI, Bayesian inference.

Figure S5. Relative frequency analysis using DiscoVista for the gene trees of all loci of USCOs, UCEs and mitochondrial genes. Fifteen branches (a–o) showed topological discordance among analyses. Four-letter prefixes of exemplar names represent codes for respective

family names. Red bar shows the relative frequency of the topology consistent in our species tree topology, whereas blue and teal bars show alternative topologies. The dotted lines at 0.33 indicate the 1/3 threshold for the frequency of gene trees supporting the two minority topologies given multi-species coalescent. USCO, Universal single-copy ortholog; UCE, ultraconserved element.

Figure S6. Comparison of divergence time estimates for the crown clade sharing across four studies.

Figure S7. Comparative analysis of phylogenetic matrices before and after Dayhoff6 recoding. (A) Differential measurement of relative compositional heterogeneity between original and recoded matrices (original minus recoded). (B) Comparative assessment of phylogenetic support, represented as average bootstrap support values (original minus recoded matrices) across all nodes.

Table S1. Taxonomic information, GenBank/SRA accession numbers and references for all analysed individuals. *: published study.

Table S2. Species taxonomic information used for design of Neuropterida UCE probe sets.

Table S3. Minimum age constraints used for the divergence times analysis.

Table S4. Based on K-mer assessment of genome size of Neuropterida ($k = 21$).

Table S5. Genome assembly information of Neuropterida genomic data. S: Complete and single-copy BUSCOs; D: Complete and duplicated BUSCOs; F: Fragmented BUSCOs; M: Missing.

Table S6. Universal single-copy ortholog extraction in this study. C: Complete; S: Complete and single-copy BUSCOs; D: Complete and duplicated BUSCOs; F: Fragmented BUSCOs; M: Missing.

Table S7. Summary of USCO amino acid and UCE nucleotide matrices for phylogenetic analyses.

Table S8. The size of PCGs, *rrnL*, *rrnS* and control region, respectively, among all Neuropterida mitochondrial genomes studied. –: incomplete mitochondrial genomes.

Table S9. Summary of divergence time estimates for each node in the maximum likelihood tree based on USCO_V matrix using three calibrations schemes.

Table S10. Summary of divergence time estimates for each node in the maximum likelihood tree based on UCE_V matrix using three calibrations schemes.

Table S11. Summary of divergence time estimates for each node based on PCGR matrix using three calibrations schemes.

How to cite this article: Zhang, R., Wang, L., Tian, S., Liu, Y., Jiang, Y., Zhou, X. et al. (2025) Inconsistent performance of multi-type genomic data in phylogenomics of neuropteridan insects, with solutions toward conflicting results. *Systematic Entomology*, 50(4), 855–875. Available from: https://doi.org/10.1111/syen.12684