# Article

# Improved modelling of compositional heterogeneity reconciles phylogenomic conflicts among lacewings

CHEN-YANG CAI[1, 2, *], ERIK TIHELKA[1, 2], XING-YUE LIU[3] & MICHAEL S. ENGEL[4]

[1]*State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, and Center for Excellence in Life and Paleoenvironment, Chinese Academy of Sciences, Nanjing 210008, China*
[2]*School of Earth Sciences, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, United Kingdom*
[3]*Department of Entomology, China Agricultural University, Beijing 100193, China*
[4]*Division of Entomology, Natural History Museum, and Department of Ecology & Evolutionary Biology, University of Kansas, Lawrence, KS, U.S.A.*

✉ *cycai@nigpas.ac.cn;* ⬤ *https://orcid.org/0000-0002-9283-8323*
✉ *wn20250@bristol.ac.uk;* ⬤ *https://orcid.org/0000-0002-5048-5355*
✉ *xingyue_liu@yahoo.com;* ⬤ *https://orcid.org/0000-0002-9168-0659*
✉ *msengel@ku.edu;* ⬤ *https://orcid.org/0000-0003-3067-077X*
*\*Corresponding author*

## Abstract

Exponential growth of large-scale data for Neuropterida, an iconic group of insects used in behavioural, ecological, and evolutionary studies, has greatly changed our understanding of the origin and evolution of lacewings and their allies. Recent phylogenomic studies of Neuropterida based on mitogenomes, anchored hybrid enrichment (AHE) data, and transcriptomes have yielded a well-resolved and largely congruent phylogeny. Some interfamilial relationships of lacewings, however, remain inconsistent among different phylogenomic studies. Here we re-analysed the genome-scale AHE and transcriptomic data for Neuropterida under the better fitting site-heterogeneous CAT-GTR+G4 model and recovered a strongly supported and congruent tree for the deeper phylogeny of Neuroptera. Integrating the smaller but more broadly sampled AHE and the larger but less-sampled transcriptomic data, we present a holistic phylogeny of Neuropterida from which to explore patterns of evolution across the clade. Our re-analyses of the largest available datasets of Neuropterida highlight the significance of modelling across-site compositional heterogeneity and model comparison in large-scale phylogenomic studies of insects.

**Keywords:** Neuropterida, phylogenomics, evolution, compositional heterogeneity, systematic error

## Introduction

In the recent decade the amount of molecular data generated for the exploration of insect evolution is staggering (Misof *et al*., 2014; Tihelka *et al*., 2021).

This is no less true for the study of the Neuropterida, an iconic group of insects used in behavioural, ecological, and evolutionary studies, as well as utilised in sustainable pest management strategies throughout the world. The Neuropterida are a well-established clade comprising three extant orders: Megaloptera (dobson-, fish- and alderflies), Raphidioptera (snakeflies), and the comparatively species-rich Neuroptera (lacewings, owlflies, antlions, and their relatives) (Wang *et al*., 2017; Engel *et al*., 2018; Winterton *et al*., 2018; Vasilikopoulos *et al*., 2020). Recently, comprehensive phylogenomic studies of Neuropterida based on mitogenomes (Wang *et al*., 2017), anchored hybrid enrichment (AHE) data (Winterton *et al*., 2018), and transcriptomes (Vasilikopoulos *et al*., 2020) have yielded a well-resolved and largely congruent phylogeny of Neuropterida. The interrelationships of the order are consistently recovered with high support, with Megaloptera as sister to Neuroptera and Raphidioptera diverging earlier, and as found in other studies based on smaller sampling of genes and taxa (Misof *et al*., 2014; Wang *et al*., 2019; Tihelka *et al*., 2021). Some key nodes to the backbone of Neuroptera are also seemingly well resolved. Within the diverse Neuroptera, Coniopterygidae are sister to all other lacewings, although this relationship is dependent on gene and model selections when the taxon sampling is limited (Wang *et al*., 2019). Additionally, clades are repeatedly recovered consisting of all Neuroptera excluding Dilaridae, Osmyloidea, and Coniopterygidae as well as those families of the Myrmeleontiformia and the Mantispoidea. Nonetheless, some interfamilial relationships of Neuroptera remain incongruent among different phylogenetic studies.

The monophyly of the superfamily Osmyloidea has been recently well supported by analysing AHE and transcriptomic data (Winterton *et al.*, 2018; Vasilikopoulos *et al.*, 2020), but not by mitogenomes (Wang *et al.*, 2017). Furthermore, the interrelationships of the three osmyloid families (Osmylidae, Nevrorthidae, and Sisyridae) vary among analyses. Similarly, the phylogenetic positions of Chrysopidae and Hemerobiidae and the relationships among myrmeleontoid families are incongruent among recent phylogenomic studies (Wang *et al.*, 2017; Winterton *et al.*, 2018; Vasilikopoulos *et al.*, 2020).

Reconstructing the insect tree of life in the phylogenomic era is a challenge, considering the data richness, compositional heterogeneity of molecular data, modelling of molecular evolution, and computational burden (Kapli *et al.*, 2020; Tihelka *et al.*, 2021; Lozano-Fernandez, 2022). Among the confounding factors for inferring a robust and consistent tree, modelling is one of the most critical components, especially for inferring deeper relationships and rapid radiations (Tihelka *et al.*, 2021; Kapli *et al.*, 2020).

Compositional and rate heterogeneity are among the most common sources of phylogenetic incongruence (Philippe *et al.*, 2011; Cai *et al.*, 2020; Tihelka *et al.*, 2021). Models often used in the most comprehensive phylogenomic analyses of Neuropterida based on AHE and transcriptomic data (Winterton *et al.*, 2018; Vasilikopoulos *et al.*, 2020) can not account for among-site compositional heterogeneity. Recent phylogenomic studies on beetles (Cai *et al.*, 2020; Cai *et al.*, 2022), fleas (Meusemann *et al.*, 2020; Tihelka *et al.*, 2020) and Neuropterida (Wang *et al.*, 2019) have clearly demonstrated that reducing site compositional heterogeneity in datasets combined with the utilisation of evolutionary models accounting also for compositional heterogeneity (*e.g.*, CAT-GTR+G4 model (Lartillot *et al.*, 2013; Lartillot, 2020)) has been shown to improve the fit of the model to the data and reduce long-branch attraction artefacts. Consequently, site-heterogeneous models have been widely used to resolve deep and rapid radiations and reconcile conflicts among analyses. Here we explore the phylogeny of Neuropterida by re-analysing the recently published AHE and transcriptomic data by using a better-fitting evolutionary model based on formal model testing, and our results clearly show that improved modelling of across-site compositional heterogeneity can reconcile phylogenomic conflicts of among lacewings.

**Materials and methods**

*Dataset collation*
The most complete transcriptomic data set (Supermatrix A) of Vasilikopoulos *et al.* (2020) was downloaded from the Dryad digital repository, https://doi.org/10.5061/dryad.1jwstqjrs. Since the interrelationships of Neuropterida (Neuroptera, Megaloptera, and Raphidioptera) have been confidently resolved in recent phylogenetic studies based on mitogenomes, transcriptomes, and anchored hybrid enrichment data (Misof *et al.*, 2014; Wang *et al.*, 2017; Winterton *et al.*, 2018; Wang *et al.*, 2019; Vasilikopoulos *et al.*, 2020), we generated a taxon-reduced matrix focusing on the order Neuroptera, and specifically those few areas of contention among analyses, by subsampling Supermatrix A (the most complete dataset, with 1,550,004 amino acid [AA] sites) from Vasilikopoulos *et al.* (2020) and filtering it using BMGE v.1.1 (Criscuolo & Gribaldo, 2010) (BLOSUM62 -h 0.1:0.4) to remove the constant AA sites and select evolutionarily conservative regions. The resultant dataset (386,322 AA sites) is represented by 49 taxa (three outgroups from Megaloptera and 46 ingroup taxa including all sampled genera) and 386,322 AA sites.

The original AA and NT supermatrices of the anchored hybrid enrichment (AHE) data, provided by Winterton *et al.* (2018) as Supporting Information, were downloaded from https://doi.org/10.1111/syen.12278. The information-free sites of the AA dataset were deleted manually. Since the three independent runs (chains) of our analysis of the complete 137-taxa nucleotide dataset under the CAT-GTR+G4 model did not converge, we further analysed another 133-taxa dataset with four rogue species excluded. The rogue taxa were determined by comparing the two chains of runs of the 137-taxa dataset.

*Phylogenetic analyses*
Phylogenomic analyses of both AHE and transcriptomic amino acid datasets were conducted using the LG4X+R and the site-heterogeneous (LG+C60+F+G) models with IQ-Tree v.1.6.3 and 1,000 ultrafast bootstraps (Nguyen *et al.*, 2015; Hoang *et al.*, 2018). For the LG+C60+F+G model applied to the AHE data, the posterior mean site frequency (PMSF) model (Wang *et al.*, 2018) was applied using the LG4X+R tree as the guide tree.

Phylogenetic analyses of the transcriptomic (amino acid) and AHE (amino acid and nucleotide) datasets were also performed under the compositionally site-heterogeneous infinite mixture model CAT-GTR+G4 in PhyloBayes MPI 1.7a (Lartillot *et al.*, 2013). For the PhyloBayes runs, two independent Markov chain Monte Carlo chains were run until convergence (maxdiff < 0.3) and the bpcomp program was used to generate output of the largest (maxdiff) and mean (meandiff) discrepancies observed across all bipartitions. As achieving convergence with the CAT-GTR+G4 model in PhyloBayes becomes challenging beyond datasets of around 20,000 positions even when using parallelization, our runs were regarded

as acceptable as long as the discrepancies between chains did not affect the key nodes of interest (Lartillot, 2020).

*Testing model adequacy*
For the comparatively small AHE amino acid data set, we used the comparatively efficient and reliable approaches, *i.e.*, the leave-one-out cross-validation (LOO-CV) and the widely applicable information criterion (wAIC), to estimate the relative fit of alternative models (CAT-GTR, LG and LG+C60) in the latest PhyloBayes MPI 1.9 (Lartillot, 2022). The LOO-CV and wAIC scores were compared to determine and select the best-fitting model.

# Results

*Reanalyses of the transcriptomic datasets of Vasilikopoulos et al. (2020)*
Our phylogenomic results based on the filtered transcriptomic dataset under both the LG4X+R (Le *et al.*, 2012) and the site-heterogeneous CAT-GTR+G4 models yielded an almost identical tree except for the internal relationships of Hemerobiidae (*Sympherobius elegans* and *Megalomus tortricoides*). The interfamilial relationships are maximally supported (MLB ≥ 95; BPP = 1, except Ascalaphidae) as in the focal tree inferred from analyses of the concatenated amino-acid sequence data of Vasilikopoulos *et al.* (2020). The phylogenetic relationships within the monophyletic Osmyloidea were confidently resolved (maximally supported in both analyses), with Osmylidae being sister to Nevrorthidae + Sisyridae, as found in all the concatenated amino-acid analyses in the original study (Vasilikopoulos *et al.*, 2020). Hemerobiidae were recovered with maximal support as sister to Ithonidae + Myrmeleontiformia, although it's notable that the myrmeleontiform family Psychopsidae was not sampled in the transcriptomic datasets.

*Reanalyses of the anchored hybrid enrichment data of Winterton et al. (2018)*
Our phylogenetic analyses based on the 137- and 133-taxa nucleotide data sets under the CAT-GTR+G4 model resulted in a congruent tree as shown in the Supplemental Data, although a few shallower nodes within Sisyridae, Hemerobiidae and Myrmeleontidae were not maximally supported. Our CAT-GTR+G4 trees are largely consistent with the tree under the site-homogeneous GTR+G model used in the original study (Winterton *et al.*, 2018), but differ significantly in the systematic positions of Hemerobiidae and families of Osmyloidea and Geoneuroptera.

Our analyses based on amino acid data (137 taxa) under the simpler LG4X+R and the site-heterogeneous

LG+C60 models resulted in an almost identical tree topology (except for some nodes within Hemerobiidae) with the results of Winterton *et al.* (2018). Similarities between our result and the original analysis under site-homogeneous models (Winterton *et al.*, 2018) include: 1) Nevrorthidae were sister to Osmylidae (maximum likelihood bootstrap [MLB] = 99); 2) Hemerobiidae were sister to all extant Neuroptera excluding Coniopterygidae, Dilaridae and Osmyloidea; and 3) Ascalaphidae were nested within Myrmeleontidae (low support, MLB = 75), as sister to Palparinae + Maulini.

By stark contrast, our focal analysis based on the amino acid data set (137 taxa) under the CAT-GTR+G4 model resulted in a tree (Fig. 1) largely consistent with the amino acid Bayesian tree (ExaBayes, under the partitioned model) in Winterton *et al.* (2018). However, under the better-fitting model our tree differs significantly in the relationships of some key families that have broader implications for understanding character evolution and natural classification: 1) Nevrorthidae were well supported as sister to Sisyridae, not Osmylidae (as in our CAT-GTR+G4 trees based on nucleotide data sets); 2) Ithonidae were sister to Psychopsidae + Myrmeleontoidea, rather than Myrmeleontoidea alone; and 3) the antlion subfamily Palparinae was supported as sister to Myrmeleontinae (BPP = 0.87), rather than Ascalaphidae (owlflies). Based on the comparative analyses of both nucleotide and amino acid sequence data of beetles (McKenna *et al.*, 2019; Vasilikopoulos *et al.*, 2019; Cai *et al.*, 2020; 2022), fleas (Meusemann *et al.*, 2020; Tihelka *et al.*, 2020), hexapods (Schwentner *et al.*, 2017), and pancrustacean (Rota-Stabelli *et al.*, 2013), it is clear that less saturated amino acids should be preferred to nucleotides in phylogenomic analyses of ancient relationships (Inagaki & Roger, 2006).

*Model comparison*
The LOO-CV (leave-one-out cross validation) scores and the wAIC (widely applicable information criterion) obtained here were quite close to each other, indicating that wAIC is a good approximation of LOO-CV for the large amino acid dataset.

Based on the model fitness scores given in Table 1, the CAT-GTR model was clearly a better fit than the site-heterogeneous LG+C60 model and the site-homogeneous LG model on the AHE amino acid dataset from Winterton *et al.* (Winterton *et al.*, 2018), both according to LOO-CV (Δcv1 = −19.85 + 22.18 = 2.33; Δcv2 = −19.85 + 23.45 = 3.60) and according to wAIC (ΔwAIC1 = −19.83 + 22.18 = 2.35; ΔwAIC2 = −19.83 + 23.44 = 3.61). As such, we selected the tree topology under the best fitting CAT-GTR model as our preferred tree for explaining patterns in the systematics and evolution of Neuroptera.
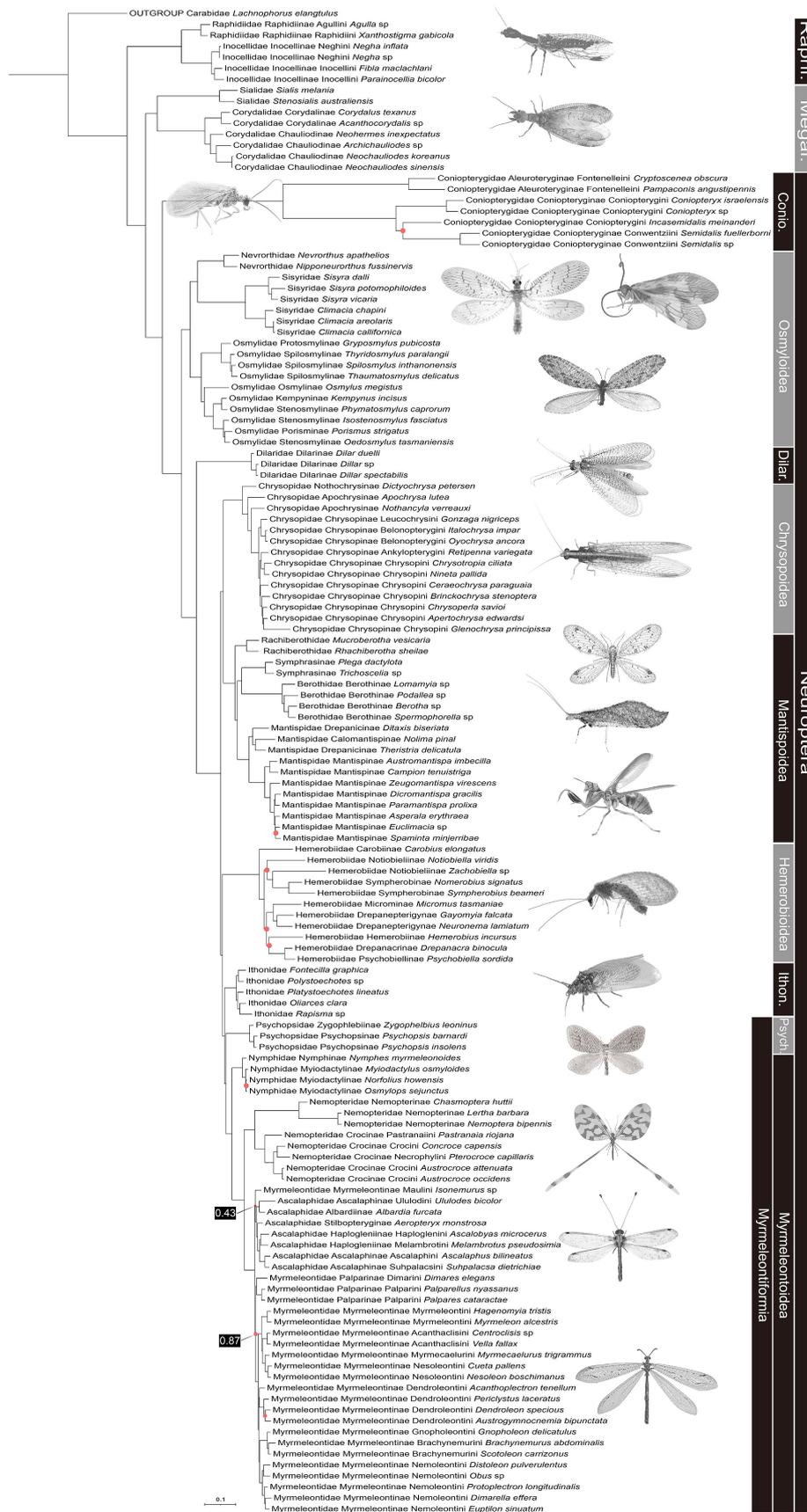
Raphi. | Megal. | Conio. | Osmyloidea | Dilar. | Chrysopoidea | Neuroptera | Mantispoidea | Hemerobioidea | Ithon. | Psych | Myrmeleontoidea | Myrmeleontiformia

OUTGROUP Carabidae *Lachnophorus elangtulus*
Raphidiidae Raphidiinae Agullini *Agulla* sp
Raphidiidae Raphidiinae Raphidiini *Xanthostigma gabicola*
Inocellidae Inocelliinae Neghini *Negha* sp
Inocellidae Inocelliinae Inocellini *Fibla maclachlani*
Inocellidae Inocelliinae Inocellini *Parainocellia bicolor*
Sialidae *Sialis melania*
Sialidae *Stenosialis australiensis*
Corydalidae Corydalinae *Corydalus texanus*
Corydalidae Corydalinae *Acanthocorydalis* sp
Corydalidae Chauliodinae *Neohermes inexpectatus*
Corydalidae Chauliodinae *Archichauliodes* sp
Corydalidae Chauliodinae *Neochauliodes koreanus*
Corydalidae Chauliodinae *Neochauliodes sinensis*
Coniopterygidae Aleuropteryginae Fontenelleini *Cryptoscenea obscura*
Coniopterygidae Aleuropteryginae Fontenelleini *Pampaconis angustipennis*
Coniopterygidae Coniopteryginae Coniopterygini *Coniopteryx israelensis*
Coniopterygidae Coniopteryginae Coniopterygini *Coniopteryx* sp
Coniopterygidae Coniopteryginae Coniopterygini *Incasemidalis meinanderi*
Coniopterygidae Coniopteryginae Conwentziini *Semidalis fuellerborni*
Coniopterygidae Coniopteryginae Conwentziini *Semidalis* sp
Nevrorthidae *Nevrorthus apathelios*
Nevrorthidae *Nipponeurorthus fussinervis*
Sisyridae *Sisyra dalli*
Sisyridae *Sisyra potomophiloides*
Sisyridae *Sisyra vicaria*
Sisyridae *Climacia chapini*
Sisyridae *Climacia areolaris*
Sisyridae *Climacia californica*
Osmylidae Protosmylinae *Gryposmylus pubicosta*
Osmylidae Spilosmylinae *Thyridosmylus paralangii*
Osmylidae Spilosmylinae *Spilosmylus inthanonensis*
Osmylidae Spilosmylinae *Thaumatosmylus delicatus*
Osmylidae Osmylinae *Osmylus megistus*
Osmylidae Kempyninae *Kempynus incisus*
Osmylidae Stenosmylinae *Phymatosmylus caprorum*
Osmylidae Stenosmylinae *Isostenosmylus fasciatus*
Osmylidae Porisminae *Porismus strigatus*
Osmylidae Stenosmylinae *Oedosmylus tasmaniensis*
Dilaridae Dilarinae *Dilar duelli*
Dilaridae Dilarinae *Dilar* sp
Dilaridae Dilarinae *Dilar spectabilis*
Chrysopidae Nothochrysinae *Dictyochrysa petersen*
Chrysopidae Apochrysinae *Apochrysa lutea*
Chrysopidae Apochrysinae *Nothancyla verreauxi*
Chrysopidae Chrysopinae Leucochrysini *Gonzaga nigriceps*
Chrysopidae Chrysopinae Belonopterygini *Italochrysa impar*
Chrysopidae Chrysopinae Belonopterygini *Oyochrysa ancora*
Chrysopidae Chrysopinae Ankylopterygini *Retipenna variegata*
Chrysopidae Chrysopinae Chrysopini *Chrysotropia ciliata*
Chrysopidae Chrysopinae Chrysopini *Nineta pallida*
Chrysopidae Chrysopinae Chrysopini *Ceraeochrysa paraguaia*
Chrysopidae Chrysopinae Chrysopini *Brinckochrysa stenoptera*
Chrysopidae Chrysopinae Chrysopini *Chrysoperla savioi*
Chrysopidae Chrysopinae Chrysopini *Apertochrysa edwardsi*
Chrysopidae Chrysopinae Chrysopini *Glenochrysa principissa*
Rachiberothidae *Mucroberotha vesicaria*
Rachiberothidae *Rhachiberotha sheilae*
Symphrasinae *Plega dactylota*
Symphrasinae *Trichoscelia* sp
Berothidae Berothinae *Lomamyia* sp
Berothidae Berothinae *Podallea* sp
Berothidae Berothinae *Berotha* sp
Berothidae Berothinae *Spermophorella* sp
Mantispidae Drepanicinae *Ditaxis biseriata*
Mantispidae Calomantispinae *Nolima pinal*
Mantispidae Drepanicinae *Theristria delicatula*
Mantispidae Mantispinae *Austromantispa imbecilla*
Mantispidae Mantispinae *Campion tenuistriga*
Mantispidae Mantispinae *Zeugomantispa virescens*
Mantispidae Mantispinae *Dicromantispa gracilis*
Mantispidae Mantispinae *Paramantispa prolixa*
Mantispidae Mantispinae *Asperala erythraea*
Mantispidae Mantispinae *Euclimacia* sp
Mantispidae Mantispinae *Spaminta minjerribae*
Hemerobiidae Carobiinae *Carobius elongatus*
Hemerobiidae Notiobiellinae *Notiobiella viridis*
Hemerobiidae Notiobiellinae *Zachobiella* sp
Hemerobiidae Sympherobiinae *Nomerobius signatus*
Hemerobiidae Sympherobiinae *Sympherobius beameri*
Hemerobiidae Microminae *Micromus tasmaniae*
Hemerobiidae Drepanepterygynae *Gayomyia falcata*
Hemerobiidae Drepanepterigynae *Neuronema lamiatum*
Hemerobiidae Hemerobiinae *Hemerobius incursus*
Hemerobiidae Drepanacrinae *Drepanacra binocula*
Hemerobiidae Psychobiellinae *Psychobiella sordida*
Ithonidae *Fontecilla graphica*
Ithonidae *Polystoechotes* sp
Ithonidae *Platystoechotes lineatus*
Ithonidae *Oliarces clara*
Ithonidae *Rapisma* sp
Psychopsidae Zygophlebiinae *Zygophlebius leoninus*
Psychopsidae Psychopsinae *Psychopsis barnardi*
Psychopsidae Psychopsinae *Psychopsis insolens*
Nymphidae Nymphinae *Nymphes myrmeleonoides*
Nymphidae Myiodactylinae *Myiodactylus osmyloides*
Nymphidae Myiodactylinae *Norfolius howensis*
Nymphidae Myiodactylinae *Osmylops sejunctus*
Nemopteridae Nemopterinae *Chasmoptera huttii*
Nemopteridae Nemopterinae *Lertha barbara*
Nemopteridae Nemopterinae *Nemoptera bipennis*
Nemopteridae Crocinae Pastranaiini *Pastrania riojana*
Nemopteridae Crocinae Crocini *Concroce capensis*
Nemopteridae Crocinae Necrophylini *Pterocroce capillaris*
Nemopteridae Crocinae Crocini *Austrocroce attenuata*
Nemopteridae Crocinae Crocini *Austrocroce occidens*
Myrmeleontidae Myrmeleontinae Maulini *Isonemurus* sp
Ascalaphidae Ascalaphinae Ululodini *Ululodes bicolor*
Ascalaphidae Albardiinae *Albardia furcata*
Ascalaphidae Stilbopteryginae *Aeropteryx monstrosa*
Ascalaphidae Haploglenine Haploglenini *Ascalobyas microcerus*
Ascalaphidae Haploglenine Melambrotini *Melambrotus pseudosimia*
Ascalaphidae Ascalaphinae Ascalaphini *Ascalaphus bilineatus*
Ascalaphidae Ascalaphinae Suhpalacsini *Suhpalacsa dietrichiae*
Myrmeleontidae Palparinae Dimarini *Dimares elegans*
Myrmeleontidae Palparinae Palparini *Palparellus nyassanus*
Myrmeleontidae Palparinae Palparini *Palpares cataractae*
Myrmeleontidae Myrmeleontinae Myrmeleontini *Hagenomyia tristis*
Myrmeleontidae Myrmeleontinae Myrmeleontini *Myrmeleon alcestris*
Myrmeleontidae Myrmeleontinae Acanthaclisini *Centroclisis* sp
Myrmeleontidae Myrmeleontinae Acanthaclisini *Vella fallax*
Myrmeleontidae Myrmeleontinae Myrmecaelurini *Myrmecaelurus trigrammus*
Myrmeleontidae Myrmeleontinae Nesoleontini *Cueta pallens*
Myrmeleontidae Myrmeleontinae Nesoleontini *Nesoleon boschimanus*
Myrmeleontidae Myrmeleontinae Dendroleontini *Acanthoplectron tenellum*
Myrmeleontidae Myrmeleontinae Dendroleontini *Periclystus laceratus*
Myrmeleontidae Myrmeleontinae Dendroleontini *Dendroleon specious*
Myrmeleontidae Myrmeleontinae Dendroleontini *Austrogymnocnemia bipunctata*
Myrmeleontidae Myrmeleontinae Gnopholeontini *Gnopholeon delicatulus*
Myrmeleontidae Myrmeleontinae Brachynemurini *Brachynemurus abdominalis*
Myrmeleontidae Myrmeleontinae Brachynemurini *Scotoleon carrizonus*
Myrmeleontidae Myrmeleontinae Nemoleontini *Distoleon pulverulentus*
Myrmeleontidae Myrmeleontinae Nemoleontini *Obus* sp
Myrmeleontidae Myrmeleontinae Nemoleontini *Protoplectron longitudinalis*
Myrmeleontidae Myrmeleontinae Nemoleontini *Dimarella effera*
Myrmeleontidae Myrmeleontinae Nemoleontini *Euptilon sinuatum*

0.43
0.87
0.1

**FIGURE 1.** Phylogram of Neuropterida relationships based on the CAT-GTR+G4 analysis of anchored hybrid enrichment (AHE) amino acid data. All branches have a support value of ≥ 0.93 Bayesian posterior probability (BPP) except those indicated by red dots (BPP < 0.9). Abbreviations: Conio., Coniopterygoidea; Dilar., Dilaroidea; Ithon., Ithonoidea; Megal., Megaloptera; Psych., Psychopsoidea; Raphi., Raphidioptera. Superfamilial classification is adapted from Engel *et al.* (2018).

**TABLE 1.** Comparing model fitness. Model comparison using PhyloBayes MPI-v.1.9 based on the AHE amino acid supermatrix. LOO-CV, leave-one-out cross validation; wAIC, widely applicable information criterion.

| Model type | Model | Criteria | |
|---|---|---|---|
| | | LOO-CV | wAIC |
| Site-heterogenous | CAT-GTR | -19.8451 | -19.8257 |
| | LG+C60 | -22.1823 | -22.1811 |
| Site-homogeneous | LG | -23.4451 | -23.4449 |

## Discussion

Morphological data, particularly of immatures, provide good but considerably limited data toward resolving higher-level relationships among families of Neuropterida, as the phylogenetic signal is considerably obscured by the overall generalised morphology of adults (although characters like genital sclerites are phylogenetically informative) across otherwise anciently diverging groups, combined with disparate larval morphologies associated with specialized life histories (Wang *et al*., 2017; Engel *et al*., 2018; Winterton *et al*., 2018). The vast molecular data available (especially transcriptomes and AHE data), however, have the potential to reconstruct the evolutionary tree of Neuropterida accurately and precisely, dependent on proper modelling of molecular evolution.

*Integrated phylogeny of Neuropterida*
We recovered a congruent and robust suite of interfamilial relationships for Neuropterida based on transcriptomic and AHE amino acid data sets under the site-heterogeneous CAT-GTR+G4 model (Fig. 2). Integrating the larger but less-sampled transcriptomic and the smaller but more broadly sampled AHE data, we are able to reconstruct a more holistic phylogeny of Neuropterida from which to explore patterns of evolution across the clade.

1) As supported in previous studies, Coniopterygidae (superfamily Coniopterygoidea) are sister to the remaining Neuroptera (*i.e.*, Euneuroptera). The monophyly of the superfamily Osmyloidea is corroborated, with Nevrothidae and Sisyridae recovered as sister groups and supported by their strictly aquatic larvae (Engel *et al*., 2018; Winterton *et al*., 2010).

2) Chrysopidae (Chrysopoidea), or green lacewings, are recovered as sister to the superfamily Mantispoidea (Berothidae, Mantispidae, Rhachiberothidae, and Symphrasinae), rejecting a close relationship between Chrysopidae and Hemerobiidae as revealed by mitogenomic data (Wang *et al*., 2017; Engel *et al*., 2018).

3) The mantis lacewing subfamily Symphrasinae was strongly supported (BPP = 1) as sister to the family Berothidae (beaded lacewings), but not to or within Mantispidae. We suggest that the small and archaic subfamily Symphrasinae may eventually deserve a familial rank, although broader sampling of Mantispoidea is welcomed in the future to better understand the systematic position of Symphrasinae in the tree of lacewings.

4) The phylogenetic position of the brown lacewings, or Hemerobiidae, is well supported as sister to Geoneuroptera (Ithonoidea, Myrmeleontoidea and Psychopsoidea; sensu Engel *et al*. (Engel *et al*., 2018).

5) As in Wang *et al*. (2017), Ithonidae (Ithonoidea) are supported as sister to Myrmeleontiformia, a clade encompassing Myrmeleontoidea and Psychopsoidea. The monophyly of Myrmeleontiformia is corroborated, and the interrelationships of Myrmeleontiformia are fully consistent with the prescient study of neuropteran phylogeny based on larval data (Withycombe, 1925), which was later confirmed by Sanger-sequencing data (Winterton *et al*., 2010) and a more formal phylogeny based on larval morphological, fossil, and behavioural characters (Badano *et al*., 2017; Badano *et al*., 2018).

6) Within Myrmeleontoidea, Nymphidae are sister to the clade Nemopteridae + (Ascalaphidae + Myrmeleontidae). Although Myrmeleontidae (antlions) are robustly supported as a monophylum by transcriptomic data, their monophyly is only weakly supported by the AHE nucleotide data under the CAT-GTR+G4 model. We suggest that a comprehensive phylogenomic study focusing on a denser sampling of Ascalaphidae and Myrmeleontidae and modelling compositional heterogeneity is required to fully understand such a rapid radiation of Neuroptera. The enigmatic subfamily Stilbopteryginae (restricted to Australia), placed in Myrmeleontidae in recent history or as its own family (Jones, 2019), is strongly supported as a lineage nested within Ascalaphidae and is therefore transferred to Ascalaphidae, in line with its more traditional position as near Albardiinae.

*The owlfly obstacle*
In our analyses of the AHE data, the subfamily Stilbopteryginae (represented by *Aeropteryx*) was always strongly supported as sister to Haplogleniinae + Ascalaphinae, rejecting its basal-most position within Ascalaphidae based on larval characters (Badano *et al*., 2017). Thus, we place the Australian endemic Stilbopteryginae as a subfamily of Ascalaphidae, as they
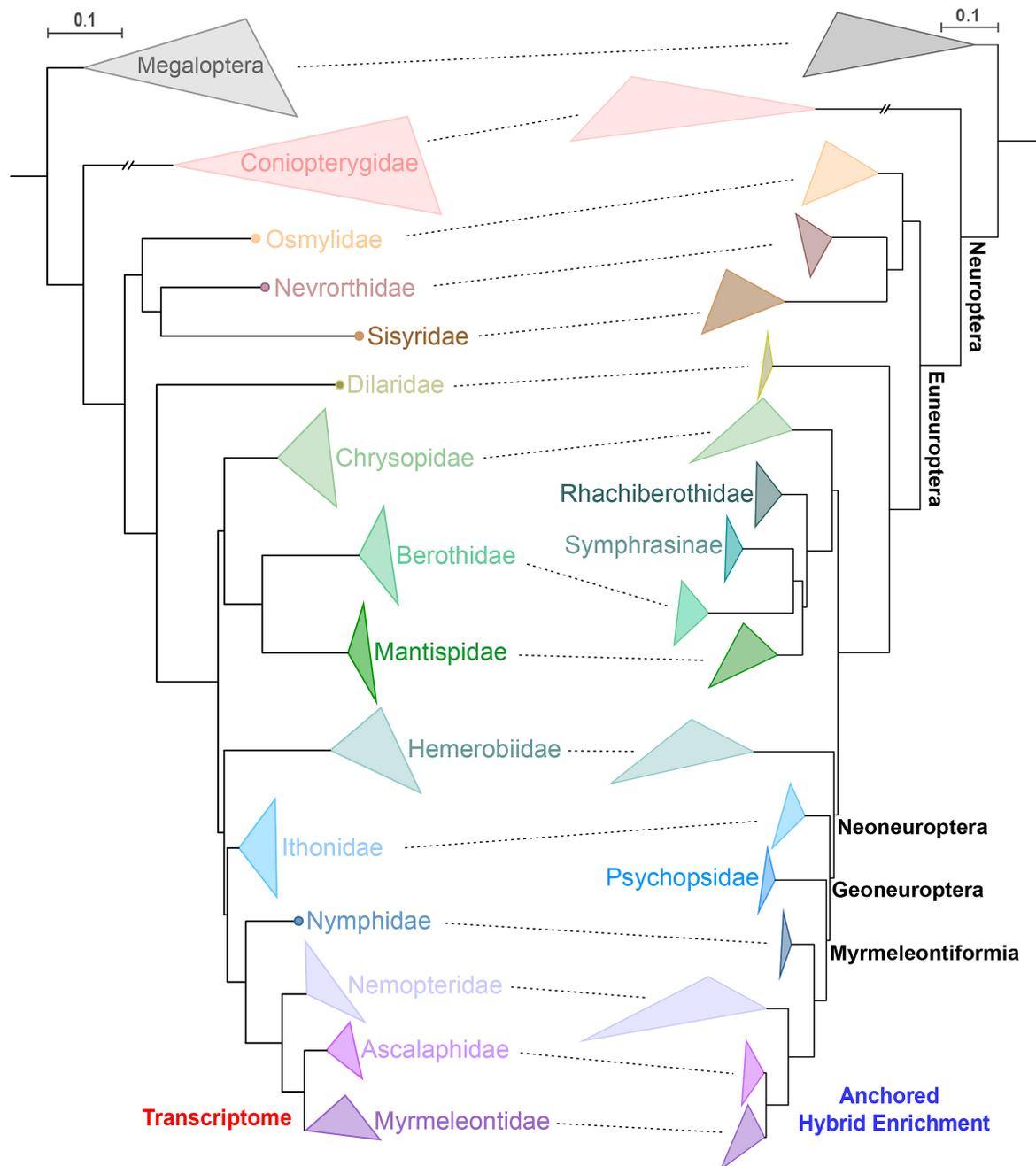
**FIGURE 2.** Congruent interfamilial relationships of Neuropterida inferred from transcriptomic (left) and AHE (right) amino acid data respectively under the site-heterogeneous CAT-GTR+G4 model. All branches have a strong support (BPP > 0.93), except for the monophyly of Myrmeleontidae (only weakly supported in AHE nucleotide data under the CAT-GTR+G4 model). Note that the transcriptomic dataset has fewer and sparser familial sampling than the AHE data.

also share notable characteristics with the latter, including the clubbed antennae, short hypostigmatic cells, and similar flight behaviour (Machado *et al*., 2019). Considering the fact that Stilbopteryginae are a putatively 'primitive' group of owlflies, albeit not the earliest-diverging lineage in our topology, Ascalaphidae probably represents a sister group to Myrmeleontidae, rather than stemmed from the latter (Machado *et al*., 2019). Our reanalyses of the nucleotide data of Winterton *et al*. (2018) support a sister-group relationship between Ascalaphidae and

Myrmeleontidae. We recommend caution, however, given that the internode branch lengths of the two lineages are extremely short, indicating a rapid radiation during their early evolution, presumably during the Cretaceous. We argue that the grouping of Palparinae and Ascalaphidae recovered under the worse fitting site-homogeneous models was probably a systematic error. Other data forms such as ultra-conserved elements (UCE) and whole genomes are promising for solving the perniciously persistent predicament of owlfly relationships. Based on

the phylogenomic evidence available, we suggest that it is more plausible to maintain historical precedence and regard owlflies as an independent family, Ascalaphidae.

Our reanalyses of the genome-scale datasets of Winterton *et al*. (2018) and Vasilikopoulos *et al*. (2020) under the site-heterogeneous CAT-GTR+G4 model recovered a well-resolved and strongly supported tree for the higher phylogeny of Neuropterida, resolving the long-lasting controversies in the phylogenetic positions of many key neuropteran families (Aspöck *et al*., 2012). Our re-analyses of the largest available datasets of Neuropterida highlight the significance of modelling across-site compositional heterogeneity and model comparison in large-scale phylogenomic studies of insects.

# Acknowledgements

# References

Aspöck, U., Haring, E. & Aspöck, H. (2012) The phylogeny of the Neuropterida: long lasting and current controversies and challenges (Insecta: Endopterygota). *Arthropod Systematics & Phylogeny*, 70, 119–129.

Badano, D., Aspöck, U., Aspöck, H. & Cerretti, P. (2017) Phylogeny of Myrmeleontiformia based on larval morphology (Neuropterida: Neuroptera): Phylogeny of Myrmeleonti-formia. *Systematic Entomology*, 42, 94–117. https://doi.org/10.1111/syen.12200

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, 3257. https://doi.org/10.1038/s41467-018-05484-y

Cai, C.Y., Tihelka, E., Giacomelli, M., Lawrence, J.F., Ślipiński, A., Kundrata, R., Yamamoto, S., Thayer, M.K., Newton, A.F., Leschen, R.A.B., Gimmel, M.L., Lü, L., Engel, M.S., Bouchard, P., Huang, D.Y., Pisani, D. & Donoghue, P.C.J. (2022) Integrated phylogenomics and fossil data illuminate the evolution of beetles. *Royal Society Open Science*, 9, 211771. https://doi.org/10.1098/rsos.211771

Cai, C.Y., 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, 106782. https://doi.org/10.1016/j.ympev.2020.106782

Criscuolo, A. & Gribaldo, S. (2010) BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evolutionary Biology*, 10, 210. https://doi.org/10.1186/1471-2148-10-210

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. https://doi.org/10.1146/annurev-ento-020117-043127

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. https://doi.org/10.1093/molbev/msx281

Inagaki, Y. & Roger, A.J. (2006) Phylogenetic estimation under codon models can be biased by codon usage heterogeneity. *Molecular Phylogenetics and Evolution*, 40, 428–434. https://doi.org/10.1016/j.ympev.2006.03.020

Jones, J.R. (2019) Total-evidence phylogeny of the owlflies (Neuroptera, Ascalaphidae) supports a new higher-level classification. *Zoologica Scripta*, 48, 761–782. https://doi.org/10.1111/zsc.12382

Kapli, P., Yang, Z.H. & Telford, M.J. (2020) Phylogenetic tree building in the genomic age. *Nature Reviews Genetics*, 21, 428–444. https://doi.org/10.1038/s41576-020-0233-0

Lartillot, N. (2020) PhyloBayes: Bayesian phylogenetics using site-heterogeneous models. *In*: Scornavacca, C., Delsuc, F. & Galtier, N. (Eds), *Phylogenetics in the Genomic Era*. No commercial publisher, Authors open access book, pp.1.5:1–1.5:16. Available at: https://hal.archives-ouvertes.fr/hal-02535342 (Accessed 10 Feb 2023)

Lartillot, N. (2022) Identifying the best approximating model in Bayesian phylogenetics: Bayes factors, cross-validation or wAIC? *bioRxiv*, 2022.04.22.489153. https://doi.org/10.1101/2022.04.22.489153

Lartillot, N., Rodrigue, N., Stubbs, D. & Richer, J. (2013) PhyloBayes MPI: Phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Systematic Biology*, 62, 611–615. https://doi.org/10.1093/sysbio/syt022

Le, S.Q., Dang, C.C. & Gascuel, O. (2012) Modeling protein evolution with several amino acid replacement matrices depending on site rates. *Molecular Biology and Evolution*, 29, 2921–2936. https://doi.org/10.1093/molbev/mss112

Lozano-Fernandez, J. (2022) A practical guide to design and assess a phylogenomic study. *Genome Biology and Evolution*, 14

(9), evac129.
https://doi.org/10.1093/gbe/evac129

Machado, R.J.P., Gillung, J.P., Winterton, S.L., Garzón-Orduña, I.J., Lemmon, A.R., Lemmon, E.M. & Oswald, J.D. (2019) Owlflies are derived antlions: anchored phylogenomics supports a new phylogeny and classification of Myrmeleontidae (Neuroptera). *Systematic Entomology*, 44, 418–450.
https://doi.org/10.1111/syen.12334

McKenna, D.D., Shin, S., Ahrens, D., Balke, M., Beza-Beza, C., Clarke, D.J., Donath, A., Escalona, H.E., Friedrich, F., Letsch, H., Liu, S.L., Maddison, D., Mayer, C., Misof, B., Murin, P.J., Niehuis, O., Peters, R.S., Podsiadlowski, L., Pohl, H., Scully, E.D., Yan, E.V., Zhou, X., Ślipiński, A. & Beutel, R.G. (2019) The evolution and genomic basis of beetle diversity. *Proceedings of the National Academy of Sciences*, 116, 24729–24737.
https://doi.org/10.1073/pnas.1909655116

Meusemann, K., Trautwein, M., Friedrich, F., Beutel, R.G., Wiegmann, B.M., Donath, A., Podsiadlowski, L., Petersen, M., Niehuis, O., Mayer, C., Bayless, K.M., Shin, S., Liu, S.L., Hlinka, O., Minh, B.Q., Kozlov, A., Morel, B., Peters, R.S., Bartel, D., Grove, S., Zhou, X., Misof, B. & Yeates, D.K. (2020) Are fleas highly modified Mecoptera? Phylogenomic resolution of Antliophora (Insecta: Holometabola). *bioRxiv*, 2020.11.19.390666.
https://doi.org/10.1101/2020.11.19.390666

Misof, B., Liu, S.L., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., Aberer, A.J., Aspöck, U., Aspöck, H., Bartel, D., Blanke, A., Berger, S., Böhm, A., Buckley, T.R., Calcott, B., Chen, J.Q., Friedrich, F., Fukui, M., Fujita, M., Greve, C., Grobe, P., Gu, S.C., Huang, Y., Jermiin, L.S., Kawahara, A.Y., Krogmann, L., Kubiak, M., Lanfear, R., Letsch, H., Li, Y.Y., Li, Z.Y., Li, J.G., Lu, H.R., Machida, R., Mashimo, Y., Kapli, P., McKenna, D.D., Meng, G.L., Nakagaki, Y., Navarrete-Heredia, J.L., Ott, M., Ou, Y.X., Pass, G., Podsiadlowski, L., Pohl, H., von Reumont, B.M., Schütte, K., Sekiya, K., Shimizu, S., Slipinski, A., Stamatakis, A., Song, W.H., Su, X., Szucsich, N.U., Tan, M.H., Tan, X.M., Tang, M., Tang, J.B., Timelthaler, G., Tomizuka, S., Trautwein, M., Tong, X.L., Uchifune, T., Walzl, M.G., Wiegmann, B.M., Wilbrandt, J., Wipfler, B., Wong, T.K.F., Wu, Q., Wu, G.X., Xie, Y.L., Yang, S.Z,, Yang, Q., Yeates, D.K., Yoshizawa, K., Zhang, Q., Zhang, R., Zhang, W.W., Zhang, Y.H., Zhao, J., Zhou, C.R., Zhou, L.L., Ziesmann, T., Zou, S.J., Li, Y.R., Xu, X., Zhang, Y., Yang, H.M., Wang, J., Wang, J., Kjer, K.M. & Zhou, X. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science,* 346, 763–767.
https://doi.org/10.1126/science.1257570

Nguyen, L.T., Schmidt, H.A., von Haeseler, A. & Minh, B.Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32, 268–274.
https://doi.org/10.1093/molbev/msu300

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G. & Baurain, D. (2011) Resolving difficult phylogenetic questions: Why more sequences are not enough. *PLoS Biology*, 9, e1000602.
https://doi.org/10.1371/journal.pbio.1000602

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.
https://doi.org/10.1093/sysbio/sys077

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.
https://doi.org/10.1016/j.cub.2017.05.040

Tihelka, E., Cai, C.Y., Giacomelli, M., Lozano-Fernandez, J., Rota-Stabelli, O., Huang, D.Y., Engel, M.S., Donoghue, P.C.J. & Pisani, D. (2021) The evolution of insect biodiversity. *Current Biology*, 31, R1299–R1311.
https://doi.org/10.1016/j.cub.2021.08.057

Tihelka, E., Giacomelli, M., Huang, D.Y., Pisani, D., Donoghue, P.C.J. & Cai, C.Y. (2020) Fleas are parasitic scorpionflies. *Palaeoentomology*, 3 (6), 641–653.
https://doi.org/10.11646/palaeoentomology.3.6.16

Vasilikopoulos, A., Balke, M., Beutel, R.G., Donath, A., Podsiadlowski, L., Pflug, J.M., Waterhouse, R.M., Meusemann, K., Peters, R.S., Escalona, H.E., Mayer, C., Liu, S.L., Hendrich, L., Alarie, Y., Bilton, D.T., Jia, F., Zhou, X., Maddison, D.R., Niehuis, O. & Misof, B. (2019) Phylogenomics of the superfamily Dytiscoidea (Coleoptera: Adephaga) with an evaluation of phylogenetic conflict and systematic error. *Molecular Phylogenetics and Evolution*, 135, 270–285.
https://doi.org/10.1016/j.ympev.2019.02.022

Vasilikopoulos, A., Misof, B., Meusemann, K., Lieberz, D., Flouri, T., Beutel, R.G., Niehuis, O., Wappler, T., Rust, J., Peters, R.S., Donath, A., Podsiadlowski, L., Mayer, C., Bartel, D., Böhm, A., Liu, S.L., Kapli, P., Greve, C., Jepson, J.E., Liu, X.Y., Zhou, X., Aspöck, H. & Aspöck, U. (2020) An integrative phylogenomic approach to elucidate the evolutionary history and divergence times of Neuropterida (Insecta: Holometabola). *BMC Evolutionary Biology*, 20, 64.
https://doi.org/10.1186/s12862-020-01631-6

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.
https://doi.org/10.1093/sysbio/syx068

Wang, Y.Y., Liu, X.Y., Garzón-Orduña, I.J., Winterton, S.L., Yan, Y., Aspöck, U., Aspöck, H. & Yang, D.Y. (2017) Mitochondrial phylogenomics illuminates the evolutionary history of Neuropterida. *Cladistics*, 33, 617–636.
https://doi.org/10.1111/cla.12186

Wang, Y.Y., Zhou, X.F., Wang, L.M., Liu, X.Y., Yang, D. &

Rokas, A. (2019) Gene selection and evolutionary modelling affect phylogenomic inference of Neuropterida based on transcriptome data. *International Journal of Molecular Sciences*, 20, 1072.
https://doi.org/10.3390/ijms20051072

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.
https://doi.org/10.1111/j.1365-3113.2010.00521.x

Winterton, S.L., Lemmon, A., Gillung, J.P., Garzon, I.J., Badano, D., Bakkes, D.K., Breitkreuz, L.C.V., Engel, M.S., Lemmon, E.M., Liu, X.Y., Machado, R.J.P., Skevington, J.H. & Oswald, J.D. (2018) Evolution of lacewings and allied orders using anchored phylogenomics (Neuroptera, Megaloptera, Raphidioptera). *Systematic Entomology*, 43, 330–354.
https://doi.org/10.1111/syen.12278

Withycombe, C.L. (1925) XV. Some aspects of the biology and morphology of the Neuroptera. With special reference to the immature stages and their possible phylogenetic significance. *Transactions of the Royal Entomological Society of London*, 72, 303–411.
https://doi.org/10.1111/j.1365-2311.1925.tb03362.x