

## Compositional phylogenomic modelling resolves the ‘Zoraptera problem’: Zoraptera are sister to all other polyneopteran insects

Erik Tihelka<sup>1,2</sup>, Michael S. Engel<sup>3,4</sup>, Jesus Lozano-Fernandez<sup>5,6</sup>, Mattia Giacomelli<sup>5</sup>, Ziwei Yin<sup>7</sup>, Omar Rota-Stabelli<sup>8,9</sup>, Diying Huang<sup>1</sup>, Davide Pisani<sup>2,5</sup>, Philip C.J. Donoghue<sup>2</sup> & Chenyang Cai<sup>1,2\*</sup>

<sup>1</sup>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, People’s Republic of China

<sup>2</sup>School of Earth Sciences, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, United Kingdom

<sup>3</sup>Division of Entomology, Natural History Museum, University of Kansas, Lawrence, Kansas, USA.

<sup>4</sup>Department of Ecology & Evolutionary Biology, University of Kansas, Lawrence, Kansas, USA

<sup>5</sup>School of Biological Sciences, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, United Kingdom

<sup>6</sup>Institute of Evolutionary Biology (CSIC-UPF), Barcelona, Spain

<sup>7</sup>Laboratory of Systematic Entomology, College of Life Sciences, Shanghai Normal University, 100 Guilin Road, Shanghai n/a, China

<sup>8</sup>Research and Innovation Centre, Fondazione Edmund Mach, 38010 San Michele all Adige (TN), Italy

<sup>9</sup>Center Agriculture Food Environment, University of Trento, 38010 San Michele all Adige, Italy

\*e-mail: [cycal@nigpas.ac.cn](mailto:cycal@nigpas.ac.cn)

**Abstract:** The evolution of wings propelled insects to their present mega-diversity. However, interordinal relationships of early-diverging winged insects and the timescale of their evolution are difficult to resolve, in part due to uncertainties in the placement of the enigmatic and species-poor order Zoraptera. The ‘Zoraptera problem’ has remained a contentious issue in insect evolution since its discovery more than a century ago. This is a key issue because different placements of Zoraptera imply dramatically different scenarios of diversification and character evolution among polyneopteran. Here, we investigate the systematic placement of Zoraptera using the largest protein-coding gene dataset available to date, deploying methods to mitigate common sources of error in phylogenomic inference, and testing historically proposed hypotheses of zorapteran evolution. We recover Zoraptera as the earliest-diverging polyneopteran order, while earwigs (Dermaptera) and stoneflies (Plecoptera) form a monophyletic clade (Dermoplecoptera) sister to the remainder of Polyneoptera. The morphology and palaeobiology of stem-zorapterans are informed by Mesozoic fossils. The gut content and mouthparts of a male specimen of *Zorotypus nascimbenei* from Kachin amber (Cretaceous) reveal a fungivorous diet of Mesozoic zorapterans, akin to extant species. Based on a set of 42 justified fossil and stratigraphic calibrations, we recover a Devonian origin of winged insects and Polyneoptera, suggesting that these groups coincided with the rise of arborescence during the diversification of early terrestrial plants, fungi, and animals. Our results provide a robust framework for understanding the pattern and timescale of early winged insect diversification.

**Keywords:** Zoraptera problem, Pterygota, stem-group fossil, compositional heterogeneity, topological conflict, molecular clock

## Introduction

The ability to fly paved the way for the winged insects (Pterygota) to become the most diverse and abundant animal group, and today pterygote insects account for over 98% of hexapod species richness<sup>1</sup>. However, the interordinal relationships within Polyneoptera (an ancient and diverse clade of winged insects with a fossil record extending to the Carboniferous) remain unclear<sup>2-7</sup>. Polyneoptera includes the first great diversification of neopteran insects, those winged insects with the novel ability to fold their wings over the body at rest, thereby protecting wings when not in use and also allowing wings to take on many novel secondary specializations<sup>1</sup>. Accordingly, Polyneoptera constitutes the first major proliferation of insect lineages that remain significant components of the insect fauna to this day. The evolutionary roots of the eleven extant polyneopteran orders (earwigs, grasshoppers, mantises, roaches, stick insects, and allies) have proven difficult to resolve<sup>6,8-10</sup>.

Uncertainties in polyneopteran relationships have implications for our understanding of the early diversification of neopteran winged insects since reconstructing the morphology and biology of the ancestral polyneopteran is contingent on resolving the early divergences in the clade. The timescale of polyneopteran evolution has likewise remained difficult to constrain, with molecular clock analyses of different topologies yielding a wide range of dates spanning the Early Devonian to the Middle Triassic<sup>4,7,11-15</sup>.

Crucial to disentangling the early evolution of Polyneoptera is the phylogenetic position of the enigmatic order Zoraptera, sometimes referred to as ‘groundlice’ or ‘angel insects’. With less than 50 described extant species, zorapterans are the third smallest and arguably the least understood insect order<sup>16</sup>. Since their discovery in 1913<sup>17</sup>, the phylogenetic position of zorapterans has represented one of the most persistent problems in reconstructing the insect tree of life. They have been placed in a bewildering number of positions, including as sister to Acercaria (Hemiptera and allies), all Holometabola (insects that undergo full metamorphosis), Dictyoptera (Mantodea + Blattodea), Eukinolabia (Phasmatodea + Embioptera), earwigs (Dermaptera), and webspinners (Embiodea)<sup>18-25</sup>. Over the last decade, phylogenetic analyses of morphological and molecular data have unequivocally supported the placement of zorapterans within Polyneoptera, but systematic affinities with almost all polyneopteran clades have been suggested<sup>2,4,16,21,26-28</sup>, leading Beutel and Weide<sup>21</sup> to dub this the ‘Zoraptera problem’, reflecting its resistance to resolution. The Zoraptera problem results, at least in part, from the insects’ unusual biology and associated morphological specializations; the minute adults occur as either fully-sighted (i.e., well-developed compound eyes) alate individuals or blind and wingless morphs that live in colonies inside decaying logs or termite nests. Zorapterans are primarily distributed in the pantropics, although some species extend northwards in North America and Asia. Aside from their unusual morphology and high rates of molecular evolution<sup>18</sup>, incongruences in resolving the systematic placement of zorapterans are exacerbated by their scant fossil record, which discloses little about their affinities<sup>29</sup>.

To investigate the phylogenetic position of Zoraptera, we reanalysed the most comprehensive genome-scale dataset available<sup>2</sup> including representatives of all major clades of winged insects and all representatives of extant polyneopterans, and we used it to test alternative hypotheses of zorapteran evolution. To understand why past studies reached contrasting results for the relationships of the major orders, we investigated the effects of common sources of systematic error in phylogenomic inference through dataset curation and use of models that can account for site-specific compositional heterogeneity<sup>30</sup>. Complementing our transcriptome analyses, we discuss the morphology of fossil zorapterans from the mid-Cretaceous amber from northern Myanmar. An exquisitely preserved specimen of *Zorotypus nascimbenei* reveals that a fungivorous diet was already present among stem-

zorapterans. Finally, we integrate transcriptomic and palaeontological data to elucidate the timescale of polyneopteran evolution and propose a revised classification of Zoraptera.

## Results

**Phylogenomics confidently assesses the placement of Zoraptera.** Resolution of conflicting hypotheses on the early evolution of Polyneoptera has to account for common sources of error in phylogenetic analyses and demonstrate why past analyses arrived at divergent topologies<sup>31</sup>. Tree reconstruction artifacts are frequently recovered when models are used that make incorrect assumptions about the data<sup>30,32</sup>. To explore the effect of different substitution models on the recovered phylogeny of Polyneoptera, we tested models of increasing complexity<sup>33</sup>: (i) standard site-homogeneous amino acid substitution models LG+G and GTR+G, variations of which have been used in previous studies<sup>2,4</sup>, (ii) the multi-matrix mixture model LG4X, and (iii) the compositionally site-heterogeneous infinite mixture model CAT-GTR+G and the compositionally site-heterogeneous LG+C60+F+G model. Among-site heterogeneous mixture models such as CAT-GTR+G have been shown to offer improved resilience to long-branch attraction artefacts (LBA)<sup>34</sup> and invariably provides better fit to the data than standard site-homogeneous models<sup>31,35–39</sup>. In our cross-validation analysis, CAT-GTR+G fitted the dataset better than the site-homogeneous LG+G and GTR+G models (CAT-GTR+G > LG+G: CV score = 34,548.8 ± 344.907; CAT-GTR+G > GTR+G: CV score = 29,639.7 ± 293.483). We thus treat the CAT-GTR+G tree as the most trustworthy topology (Fig. 1).

Analyses conducted using the site-homogeneous models LG+G, GTR+G, and the LG4X model, recovered a monophyletic Polyneoptera with a Zoraptera + Dermaptera clade (bootstrap value = 63–98) as sister to the remaining orders (Figs S16–23). As expected, these results mirror those obtained by recent transcriptomic-based phylogenies using site-homogeneous models<sup>2,4</sup>, and highlight the importance of testing the effect of compositionally-heterogeneous models to ensure a complete exploration of the data.

Our CAT-GTR+G and LG+C60+F+G analyses recovered a monophyletic Polyneoptera with maximum support (Bayesian posterior probabilities [BPP] = 1; bootstrap values = 100), in congruence with recent phylotranscriptomic studies<sup>2,4,6</sup>. Our results however differ from recent phylogenomic analyses in the placement of the early-diverging orders. Zoraptera was recovered as sister to the remaining orders with full support (BPP = 1; bootstrap = 100). Dermaptera and Plecoptera formed a well-supported clade (BPP = 1; bootstrap 90–97), sister to the remainder of Polyneoptera, excluding Zoraptera. The relationships among the remaining polyneopteran orders were in line with previous transcriptomic phylogenies. Orthoptera was sister to the remainder of Polyneoptera (excluding Zoraptera, Dermaptera, and Plecoptera), Mantophasmatodea + Grylloblattodea (Notoptera) formed a clade sister to Phasmatodea + Embioptera (Eukinolabia) dubbed Mecynoptera, and Mantodea were sister to Blattodea (Dictyoptera). The position of Zoraptera remained stable across the four analysed datasets with different taxon sampling adjusted to exclude rogue taxa and to the effect of just using a single species per order. However, analyses of a 33-taxon dataset with CAT-GTR+G, sampling only one representative of each polyneopteran order, yielded an identical position of Zoraptera, but recovered Orthoptera as sister to Plecoptera, instead of as sister to the remaining polyneopterans as in analyses with denser taxon sampling. Analyses of the 33-taxon dataset with the LG+C60+F+G model recovered Zoraptera as sister to Dermaptera with low support (bootstrap = 85). Outside Polyneoptera, the monophyly of Palaeoptera and Acercaria, supported by traditional morphology-based classification<sup>20</sup> but questioned by some recent analyses<sup>4</sup>, was recovered with full support (although support for Acercaria was low in analyses of the 102-taxon matrix under both models and the 106-taxon dataset under the LG+C60+F+G model; Figs S5, 8, 9).

**Evaluating alternative hypotheses of polyneopteran evolution.** To compare support for alternative hypotheses of early polyneopteran evolution, we ran topology tests in IQ-TREE using the site-heterogeneous LG+C60+F+G model. Model selection suggested CAT-GTR+G to be the best fitting model. However, CAT-GTR+G is not implemented in IQ-TREE. Accordingly, to run our topology tests we used LG+C60+F+G, which despite not the best fitting model, can still account for compositional heterogeneity in a maximum likelihood framework. Congruently with the majority of recent morphological and morphological studies<sup>2,4,22,40</sup>, the Approximately Unbiased (AU) test rejected placements of Zoraptera outside of Polyneoptera. In addition, AU tests rejected the placement of Zoraptera outside of ‘basal Polyneoptera’, as well as sister to Plecoptera, sister to Polyneoptera excluding Dermaptera, or sister to a clade including Dermaptera and Plecoptera ( $P_{AU} = 0$ ; Fig. 2, Tab. S1). The Haplocercata first topology favoured by analyses with site-homogeneous models<sup>2,4</sup>, with Zoraptera as sister to Dermaptera, was also rejected ( $P_{AU} > 0.006$ ; Fig. 2B). Out of 20 tested topologies, only the ‘Zoraptera-first’ topology recovered in our CAT-GTR+G analyses was not rejected ( $P_{AU} > 0.995$ ; Fig. 2A).

**Cretaceous fossils shed light on the biology of stem-group zorapterans.** Fossils, in particular stem-zorapterans, may provide an independent line of evidence for evaluating competing hypotheses of polyneopteran relationships. However, all known zorapteran fossils from the Mesozoic only narrowly diverge from the body plan of extant zorapterans. To date, all described zorapterans were placed into the genus *Zorotypus*, with the sole exception of *Xenozorotypus burmiticus* from Cretaceous Burmese amber. *Xenozorotypus* possesses plesiomorphic characters such as the presence of the vein (M3+4) in the hind wing, which is otherwise absent from all extant zorapterans<sup>29</sup>. Other Cretaceous stem-zorapterans are included in the subgenus *Octozoros*.

Here we describe a single apterous male zorapteran *Zorotypus nascimbenei*<sup>29</sup> (full morphological description provided in the Supplementary Information). Along with other members of the subgenus *Octozoros*, *Z. nascimbenei* is set apart from all crown-group zorapterans by its reduced antennae with eight segments, and plesiomorphic presence of a strong and expanded empodium on the metatarsus and jugate setae<sup>41</sup>. The exquisite preservation of *Z. nascimbenei* reveals novel details of stem-zorapteran palaeobiology. Over a hundred dark ~6 µm biconcave discal fungal spores are preserved around the appendages and attached to the abdomen of the fossil (Fig. 3c, f: fs). Their occurrence in small clusters (Fig. 3c, f, e: fsc) suggests that they were vectored by the zorapteran prior to entombment and not dispersed independently, as similar preservation is also found in pollen grains in amber transported by pollinators<sup>42</sup>. The translucent abdomen reveals similar darkened specks inside (Fig. 2c: fs-abd). The spore’s complete entombment in the amber matrix and resolution limitations of light microscopy preclude us from assigning them to a formal palynotaxon. Fungivory in *Z. nascimbenei* is also supported by the densely setose galeal brush of the fossil (Fig. 3b: gb). Apically setose galeae are also found in crown-group zorapterans and are thought to function in moving small food particles to the oral cavity<sup>21</sup>. The grinding molae of the mandibles (Fig. 3b: md) may have been used to crush fungal matter prior to ingestion<sup>21</sup>. The association of *Z. nascimbenei* with fungal spores provides the earliest direct evidence of fungivory in Zoraptera and indicates that this feeding strategy had been exploited at least as early as the mid-Cretaceous. Modern zorapterans are opportunistic omnivores, feeding on fungal spores, hyphae, predaceous on springtails, mites and nematodes, occasionally cannibalistic, and scavenging dead arthropods<sup>43</sup>.

**Timescale of polyneopteran evolution.** Previous divergence time analyses estimated an Early Devonian to the Middle Triassic origin of Polyneoptera, depending on the topology and analytical methods used<sup>4,7,11–15</sup>. Some of these estimates are untenable based on their incongruence with the palaeontological record, since polyneopterans have a fossil record extending from the Carboniferous<sup>44</sup>. To reconcile the timescale of polyneopteran evolution with the fossil record and our resolution of the Zoraptera problem, we established 42 new calibration points, more than have been used in any previous molecular clock analysis of winged insects. Furthermore, all calibrations are fully justified with respect to their stratigraphic age and phylogenetic position, in line with best practice recommendations<sup>45</sup>. Contrary to some previous analyses, we specified objective maximum age constraints based on the absence of given lineages in well-explored fossil deposits preserving insects. The full justified list of calibrations is provided in the Supplementary Information. These serve as a basis for future analyses of insect evolution, as well as our own.

We recovered a middle Silurian – Early Devonian (427–399 Ma) origin of the winged insects (Pterygota; Fig. 1). Our molecular clock analyses recovered a late Early to Late Devonian (402–365 Ma) origin of crown-Polyneoptera. These dates correspond broadly with previous molecular clock estimates of insect diversification<sup>4,46,47</sup>. They fall within the notorious ‘hexapod gap’, a period between the Early Devonian and Late Carboniferous which preserves little terrestrial sedimentary rock exposures and where insect fossils are absent<sup>48</sup>.

Under the ‘Zoraptera-first’ scenario, the divergence of the lineage that gave rise to Zoraptera is backdated to the Devonian (Tab. 1). This significantly predates previous estimates that dated the origin of the lineage in the early Permian–Carboniferous<sup>4,16</sup>. The ancestral lineage of Dermaptera and Plecoptera diverged in a Late Devonian Pennsylvanian interval (368–314 Ma). This is in congruence with the fossil record; stem-Plecoptera and †Protelytroptera (a group of Palaeozoic dermapteran-like insects) co-occur in some Permian localities<sup>13,49,50</sup>, suggesting that their earliest common ancestor must have existed before this date and a putative protelytropteran is known from the late Carboniferous<sup>51</sup>.

The lineages giving rise to the crown-clades Eukinolabia and Dictyoptera originated in the Late Pennsylvanian – Early Triassic (300–249 Ma) and early Permian – Late Triassic (290–227 Ma), respectively. The origin of termites (clade Isoptera of Blattodea) was dated to the Early Jurassic – earliest Cretaceous (187–135 Ma).

**Phylogeny of Zoraptera.** With the position of Zoraptera in the insect tree of life now becoming increasingly clear, the last remaining aspect of the ‘Zoraptera problem’ pertains to the internal phylogeny and classification of the order itself. As one of the most morphologically homogeneous insect orders, finding reliable external morphological characters for zorapteran classification has proven notoriously difficult. Zorapterans have been traditionally placed into the single genus *Zorotypus* belonging to the single family Zorotypidae<sup>20</sup>. In the past, some morphological workers have proposed organising extant species into as many as eight separate genera on the basis of wing venation<sup>19</sup> and morphology of the mouthparts and tarsi<sup>52</sup>, but this alternative classification scheme has not been widely adopted as some of the proposed generic characters are continuous across taxa or highly variable within species<sup>24</sup>. Recent molecular analyses identified multiple clades within the order<sup>16</sup>, leading Kočárek et al.<sup>53</sup> to divide the order into two families and nine genera. However, the internal relationships within Zoraptera show incongruences between analysed datasets<sup>16,53</sup>, questioning the stability of this new classification. To re-evaluate the contentious intraordinal relationships of Zoraptera, we mined GenBank for publicly available zorapteran sequences. The resultant dataset included 5 molecular markers with a nucleotide alignment of approximately 1,700 bp in length (data occupancy: 65.0%). We also prepared a

morphological dataset sampling 10 extinct and extant zorapteran species, with a stonefly as the outgroup.

The molecular analyses run with four models converged on a broadly similar topology of Zoraptera, with the site-heterogeneous CAT-GTR+G and the best-fitting model GTR+F+R3 under a maximum likelihood (ML) framework supporting an identical tree (Fig. 4). In all analyses, Zoraptera formed two main clades with strong to moderate support (BPP = 0.79–1; bootstrap value = 63–100): (i) a clade more or less corresponding to ‘Clade 1’ of Matsumura et al.<sup>16</sup> (‘Zorotypidae’ of Kočárek et al.<sup>53</sup>) where males possess asymmetrical genitalia, and (ii) a second clade including males with symmetrical genitalia (‘Clade 2’ in<sup>16</sup>, ‘Spiralizeridae’ in<sup>53</sup>). However, relationships among some species and the recently proposed genera were incongruent with previous analyses, rendering Kočárek’s et al.<sup>53</sup> Zorotypinae, Spermozorinae, and *Spermozoros* paraphyletic. Unlike in Matsumura et al.<sup>16</sup>, we recovered *Z. barberi* as sister to Clade 1 with moderate support, not Clade 2. The position of *Z. novobritannicus* within Spiralizeridae was incongruent among the analyses with different models (Figs S27–30).

A morphological phylogenetic analysis recovered *Xenozorotypus* as sister to the remaining zorapterans. The subgenus *Octozoros* is the next diverging clade. Relationships within *Zorotypus* were poorly resolved, with *Z. asymmetricus* + *Z. barberi*, and *Z. neotropicus* + *Z. magnicaudelli* + *Z. weidneri* forming poorly supported clades (bootstrap value = 31–41).

Given that morphological and molecular studies conducted by Matsumura et al.<sup>16</sup>, Kočárek et al.<sup>53</sup> and in this study support a deep split in the genus *Zorotypus*, into taxa with symmetrical and asymmetrical genitalia, we update zorapteran classification accordingly. We consider extant zorapterans with asymmetrical genitalia as belonging to the subgenus *Zorotypus sensu stricto* (equivalent to ‘Zorotypidae’ of Kočárek et al.<sup>53</sup>), while taxa with symmetrical male genitalia are assigned to the subgenus *Centrozoros* stat. rev. (Tab. 2). This conservative arrangement accommodates uncertainties in *Zorotypus* phylogeny (discussed above) and can be expanded, pending more extensive taxon and gene sampling in future studies. A detailed classification of Zoraptera with a list of species is provided in the Supplementary Information.

## Discussion

### Phylogenomics, morphology and fossils: congruence in the placement of Zoraptera.

Zoraptera represents a typical fast-evolving group, connected by a long branch to the rest of the taxa as shown in many previous analyses<sup>54,55</sup>, suggesting that its placement might be prone to be affected by tree reconstruction artifacts. Zoraptera is recovered as sister to Dermaptera (Haplocercata first) only with less fitting models, both compositionally site-homogeneous models or compositionally site-heterogeneous models with an insufficient number of compositional categories. Analyses with a higher number of substitutional categories (e.g., LG+C60+F+G and CAT-GTR+G) consistently support Zoraptera as the earliest-diverging polyneopteran order, while Dermaptera and Plecoptera form the next branch. As such, our results lend support to the ‘Zoraptera first’ hypothesis, while we regard the Haplocercata hypothesis as a phylogenetic artefact caused by systematic error.

Morphological evidence, from both extant and fossil insects, provides complementary evidence for the placement of Zoraptera. Zorapterans possess numerous plesiomorphies with respect to the remainder of Polyneoptera, such as the morphology of the pro- and mesothoracic spinae, pleural arm, and membranous anapleural suture<sup>56</sup>. A close relationship between Zoraptera and the early-diverging Dermaptera and Plecoptera has been suggested by several morphological studies. Silvestri<sup>17</sup>, who described Zoraptera in 1913, discussed superficial similarities between the order and Dermaptera. Both notably share gregarious

habits<sup>57</sup> and maternal care in some species<sup>58</sup>. Zoraptera, Plecoptera, and Dermaptera share the absence of the muscle *Musculus hypopharyngomandibularis*, a plesiomorphy shared with Palaeoptera<sup>59</sup>. Zorapterans and dermapterans are the only polyneopteran groups with holocentric chromosomes, an apparently plesiomorphic condition otherwise found in Palaeoptera<sup>2,60</sup>. *Zorotypus hubbardi* shares the number of male diploid chromosomes ( $2n = 38$ ) with the dermapteran *Prolabia* (within Polyneoptera, this condition is otherwise present only in some Isoptera and Phasmatodea)<sup>61,62</sup>. Kuznetsova *et al.*<sup>62</sup> hypothesised an ancestral karyotype of  $2n = 40$  for Zoraptera, which is shared with the early-diverging dermapteran family Labiduridae, and within Polyneoptera is otherwise found only in some Isoptera<sup>61</sup>. Friedrich *et al.*<sup>56</sup> found similarities in the thoracic skeletal structures of Zoraptera and Plecoptera, such as in the morphology of the prothoracic basisternum which is partly reduced and spoon-shaped in both orders. Both orders likewise possess reduced ovipositors<sup>63</sup>. In the head, Zoraptera and Plecoptera share tormae without mesal extensions and 5 incisivi on the left and right mandibles<sup>59</sup>. Zorapteran winged morphs have a kidney-shaped frontal spot also present in some plecopterans<sup>64</sup>, although the homology of this structure is unclear<sup>21</sup>.

Kukalová-Peck and Peck<sup>19</sup> observed that the head shape in Zoraptera with antennal articulations close to the ventral cranial margin, represents a plesiomorphy within Insecta, similar to fossil pterygotes from the Palaeozoic. According to the interpretation of Haas and Kukalová-Peck<sup>65</sup>, Zoraptera, Protelytroptera, Dermaptera, and Plecoptera all share a mp-cua cross-vein (arcus). Protelytroptera, Dermaptera, and Plecoptera share a large anal fan on the hind wing in their ground plan.

The Dermaptera + Plecoptera clade is supported by wing base structure, namely the morphology of the ventral basisubcostale, and the articulation between the antemedian notal wing process and first axillary sclerite<sup>66</sup>. Both groups share the absence of male gonostyli and a functional ovipositor (both shared with Zoraptera), and paired male gonopore<sup>66,67</sup>. An early-diverging Dermaptera + Plecoptera clade (dubbed Dermoplectopterida) was recovered in a *18S* analysis<sup>68</sup> and some early phylotranscriptomic studies<sup>6,40</sup>, and is now supported by our own analyses accounting for common sources of systematic error. We herein formally propose the name Eupolyneoptera Engel, Tihelka, & Cai clade nov. nom. for all Polyneoptera excluding Zoraptera (see the Supplementary Information for a full systematic treatment and synapomorphies).

**The nature of the ancestral polyneopteran.** Hennig<sup>20</sup> regarded the last ancestor of “Paurometabola” (i.e., his concept of Polyneoptera, but excluding Zoraptera and Plecoptera) as probably a ground-dwelling insect with forewings modified into leathery tegmina that lived during the Carboniferous. Recently, molecular studies have attempted to reconstruct morphology, life history, and chromosomal system of the last common ancestor of Polyneoptera<sup>2,61,69</sup>. Reconstructing the root of Polyneoptera is heavily dependent on the resolving the inferred relationships of the early-diverging orders.

Our analyses suggest that the earliest common ancestor of crown-polyneopterans lived in the Devonian. This coincides with the diversification of vascular plants and animals on land<sup>70</sup>. The Devonian saw an abrupt spread of vegetation and increase in plant size<sup>71-73</sup>, which would have provided structured habitats rich in plant biomass such as leaf litter, dead wood, and spaces under bark that are occupied by early-diverging polyneopterans today<sup>2</sup>. These diversified spaces would have made the newly folded wings of these early neopterans beneficial for insects invading plant crevices and other spaces, allowing for access to microhabitats previously inaccessible by palaeopterous insects. It also freed wings, when not in use, to evolve secondary specializations, such as protective covers, etc. (e.g., Dermaptera). Since the three basalmost polyneopteran orders, Zoraptera, Dermaptera, and Plecoptera, all represent opportunistic predators, scavengers, phytophages, and detritivores<sup>43,74</sup>, this suggests

such a biology was plesiomorphic to the clade and that the increase in trophic complexity during the Devonian may have opened new niches for the group that gave rise to Polyneoptera. All groups that constitute the diet of extant zorapterans, namely terrestrial fungi, springtails, possible insects, mites, and nematodes make their first appearance in the fossil record in the Devonian<sup>75,76</sup>. Diverse terrestrial fungi first occur in the fossil record in the Devonian<sup>77</sup>; notably the specimen of the stem-zorapteran *Z. nascimbenei* herein possesses specializations for feeding on fungal spores. Since our understanding of insect diversity in the Devonian is severely hampered by the rarity of well-preserved terrestrial sediments of this age<sup>48</sup>, further analyses of the fossil record will be required to test if the basal radiation of insects coincided with the diversification of land plants and animals in the Devonian. However, because unequivocal fossil stem-polyneopterans remain to be identified, the age of total-group Polyneoptera may predate the Devonian.

The ancestral polyneopteran is unlikely to have resembled extant Zoraptera or stem-zorapterans, as numerous characteristics of the order such as the greatly reduced wing venation evidently represent specializations for highly derived, subcortical life in late-stage decaying wood and consequences of miniaturisation, as zorapterans don't exceed 3 mm in length<sup>6,20,21</sup>. With Zoraptera as sister to the remaining Polyneoptera, it is possible that the ancestral polyneopteran could have exhibited gregarious behaviour and biparental care for offspring<sup>78</sup>, although these are more likely derived specializations of Zoraptera, just as maternal care is indicative of crown-Dermaptera but was absent in stem-Dermaptera<sup>79</sup>. Resolution of the position of stoneflies (Plecoptera), which possess aquatic nymphs, as sister to Dermaptera within Dermopteroptera also suggests that the ancestral Polyneopteran was likely terrestrial, contrary to the widespread view that the earliest flying insects originated in aquatic habitats<sup>80,81</sup>.

## Conclusion

The relationships among the early-diverging orders of Polyneoptera, one of the three major clades of insects with folding wings (Neoptera), represent one of the most persistent problems in insect evolution. We reanalysed the most complete protein-coding gene dataset available for the group to interrogate common sources of tree reconstruction errors to infer the early evolution of Polyneoptera. We recover a novel topology, with the enigmatic Zoraptera as sister to the remaining polyneopteran orders, and Dermaptera + Plecoptera (Dermopteroptera), as the next branch, in congruence with some previous molecular and morphological studies. By experimenting with different models of molecular evolution, we show that previous topologies within Polyneoptera are only supported by analyses conducted within a narrow analytical window using poorly fitting substitution models. A newly discovered fossil of the Cretaceous stem zorapteran *Zorotypus nascimbenei* preserves fungal spores in its abdomen, providing insights into the dietary diversity of Mesozoic zorapterans. Together with its specialised mouthparts and the presence of a spore brush on the apex of galea, *Octozoros* provide the earliest direct evidence of feeding on fungal spores in Polyneoptera and further illuminates the ancestral biology of the group. Our results highlight the importance of adequate modelling of molecular evolution and seeking a consensus of molecular and morphological evidence in resolving contentious problems in insect phylogeny.

## Methods

**Dataset assembly and phylogenetic reconstruction.** For the phylogenomic analyses, we used the multiple sequence alignment generated by Wipfler et al.<sup>2</sup>, sampling all polyneopteran orders represented by 106 species. With 3,014 protein-coding single-copy genes, it represents the most comprehensive dataset for Polyneoptera compiled to date. To



mitigate errors due to poor alignment and incorrect identification of orthologs<sup>30</sup>, the decisive amino acid (AA) dataset was trimmed with BMGE using BLOSUM95 with -h set to 0.4<sup>82</sup>. Trimming reduced the original dataset by 84.7% (from 1,246,506 AA sites to 191,242) and increased data occupancy from 52.2% to 92.0%. The exclusion of highly variable/uncertain regions reduced the average pairwise distance among taxa by 50% (from 0.30 to 0.15). We used the trimmed alignment to generate four matrices with variable taxon sampling. We excluded two and six taxa that behaved as rogue taxa in preliminary analyses, giving datasets with 106, 102, and 100 taxa. The excluded taxa were *Liposcelis* (Psocoptera) and *Pediculus* (Phthiraptera) in the 104-taxon dataset, in addition to *Apachyus* (Dermaptera), *Brachyptera* (Plecoptera), *Leuctra*, and *Perla* (Plecoptera) in the 100-taxon dataset. Wipfler *et al.*<sup>2</sup> pointed out that the position of Zoraptera may be unstable because of unbalanced taxon sampling, since the order is often represented by only a single species in phylogenetic analyses. To investigate this possibility, we produced a fourth matrix which included only one representative of each order (33 taxa dataset).

The status of Zoraptera as a fast-evolving group prone to systematic bias has been long known<sup>54,55</sup>. To overcome this problem, Wipfler *et al.*<sup>2</sup> used ModelFinder in IQ-TREE to partition their decisive alignment and analyse individual partitions with best-fitting default site-homogeneous models and the LG4X mixture model. In a recent simulation study, Wang *et al.*<sup>83</sup> found that, on a simulated and empirical datasets, partitioning data and analysing partitions with site-homogeneous models is still prone to systematic errors and underperforms in comparison with analysing concatenated datasets with site-heterogeneous models. Hence, we tested the effect of using heterogeneous models of varying computational complexity on polyneopteran phylogeny. We tested the performance of the site-homogeneous models LG+G and GTR+G, variations of which have been used in past polyneopteran phylogenomic studies<sup>2,4</sup>, the LG4X mixture model, and the compositionally site-heterogeneous LG+C60+F+G and CAT-GTR+G, the latter of which has been shown theoretically and empirically to better describe substitution in metazoan genomes<sup>33-35,84</sup>. To compare the fit of these models to our polyneopteran dataset, we run a 10-fold Bayesian cross-validation analysis in PhyloBayes MPI 1.7<sup>36</sup> with 10 replicates.

Analyses with the site-homogeneous models LG+G and GTR+G use a single amino acid substitution matrix to describe the biochemical properties of amino acids. Variations of these models have been widely used in previous analyses of polyneopteran phylogeny. In our analyses, they were conducted in IQ-TREE 1.6.3<sup>85</sup>. Analyses employing the LG4X, that accounts for substitution rate heterogeneity across the alignment by estimating four different LG+G substitution matrices<sup>86</sup>, were likewise performed in IQ-TREE. The compositionally site-heterogeneous profile mixture model LG+C60+F+G implements 60 equilibrium frequency categories<sup>87</sup> and was run in IQ-TREE. Support values for all analyses were generated using 1,000 ultra-fast bootstraps.

Analyses with the site-heterogeneous infinite mixture model CAT-GTR+G were performed in PhyloBayes by running two independent Markov chain Monte Carlo (MCMC). The bpcomp program was used to evaluate convergence by generating output of the largest (maxdiff) and mean (meandiff) discrepancies observed across all bipartitions. Due to computational constraints, only the 100- and 33-taxon datasets were run to complete convergence (maxdiff<sub>100</sub> = 0; maxdiff<sub>25</sub> = 0.097). The remaining analyses did not converge (maxdiff<sub>102,106</sub> = 1), however the recovered topology was identical to the converged datasets only with a few nodes poorly supported.

**Topology tests.** To evaluate support for 20 alternative hypotheses of polyneopteran evolution, we conducted topology tests using the 100-taxon dataset in IQ-TREE using the site-heterogeneous LG+C60 model. This model was used because the best fitting CAT-

GTR+G is not implemented in IQ-Tree, but LG+C60+F+G (similarly to CAT-GTR+G) can still accommodate (despite it is not best fit), compositional heterogeneity and would fit the data better than the other models available in IQ-Tree. To perform the topology tests the Approximately Unbiased (AU) test was used, with 10,000 resamplings performed using the RELL method. We focused on testing historical and recent hypotheses of Zoraptera as well as Dermaptera and Plecoptera phylogeny, since the relationships of these three orders have produced the most incongruence in phylogenetic studies of Polyneoptera<sup>3</sup>.

**Molecular clock analyses.** To establish a timescale of polyneopteran evolution congruent we followed the best-practice recommendations for the selection of calibrations outlined by Parham et al.<sup>45</sup>. Each fossil-based node calibration is based on a museum-curated fossil specimen or group of specimens, phylogenetic and age justifications for each are provided in the Supplementary Information. In total, 42 calibrations spread equitably throughout the tree were used, the calibrated nodes are displayed in Fig. S1.

For example, the roached *Qilianiblattea namurensis* from the Tupo Formation in northwestern China<sup>44,88</sup> was used to calibrate the node representing stem-Dictyoptera. The preserved right forewing shares with extant Blattodea a deeply concave CuP vein<sup>89</sup>, but the RA with branches translocated to RP indicates that the fossil belongs to stem Dictyoptera<sup>44</sup>. Roachoids' such as *Q. namurensis* were roach-like insects abundant in the Paleozoic and some of the earliest winged insects in the fossil record. While some authors consider Paleozoic roachoids as close to extant Blattodea<sup>90</sup>, they are excluded from crown Blattodea and Mantodea by wing venation character and most notably by the presence of long external ovipositors in females<sup>91</sup>. Previous analyses have shown that excluding Carboniferous roachids from molecular clock analyses leads to underestimates of polyneopteran timescale of diversification<sup>11</sup>. The fossiliferous horizon yielding the specimen has been dated to the Namurian B/C or to the Bashkirian (latest Duckmantian), based on the presence of a characteristic ammonoid and conodont fauna<sup>88,92</sup>. Trümper et al.<sup>92</sup> proposed an upper bound on the age of the insect-bearing bed of ~315 Ma, which we used as the minimum constraint on the node. The maximum age, 428.9 Ma, is taken from the two oldest Lagerstätten preserving terrestrial animals, the Přídolian Ludford Lane in Shropshire<sup>93</sup>, England and the Pragian Rhynie chert from Aberdeenshire, Scotland

The geological timescale follows Ogg et al.<sup>94</sup>. Molecular clock analyses were run in MCMCTree implemented in PAML 4.7<sup>95,96</sup>. We obtained 200,000 trees with a sampling frequency of 50 and discarded 10,000 as burn-in. Default parameters were set as follows: 'cleandata=0', 'BDparas=1 1 0', 'kappa\_gamma=6 2', 'alpha\_gamma=1 1', 'rgene\_gamma=2 20', 'sigma2\_gamma=1 10', and 'finetune=1: 0.1 0.1 0.1 0.01 .5'. Convergence was assessed by plotting posterior mean times from the first run against the second run. Analyses were run using a uniform prior distribution between minimum and maximum age constraints. We used soft bounds with maximum and minimum tail probabilities of 2.5%. Analyses both with the autocorrelated rates clock (AC) model and independent rates (IR) clock model, and our results integrate the results from both. To ensure our priors were appropriate, we ran the MCMCTree analysis without sequence data to calculate the effective priors, compared to the specified priors.

**Palaeontology.** The studied amber inclusion originates from mines near the summit of the Noiye Bum hill in the Hukawng Valley, Kachin State, northern Myanmar. Radiometric dating of the amber deposit has provided an early Cenomanian age, ~99 Ma, while palaeontological evidence suggests that the amber is no older than late Albian<sup>97,98</sup>. The specimen has been cut with a razor blade and polished using diatomite powder. Photographs under normal visible light were taken with a Canon EOS 5D Mark III digital camera, equipped with a Canon MP-

E 65 mm macro lens (F2.8, 1–5×), and with an attached Canon MT-24EX twin flash. Epifluorescence photomicrographs were taken using a Zeiss Axio Imager 2 microscope under the eGFP mode (Zeiss Filter Set 10; excitation/emission: 450–490/515–565 nm). Extended depth of field images were digitally compiled in Zerene Stacker 1.04. The type specimen is deposited in the amber collection of the Nanjing Institute of Geology and Palaeontology (NIGP), Chinese Academy of Sciences under the accession number NIGP175112. The amber piece was purchased from a local amber miner in late 2016, complying with the laws of Myanmar and China, and is open to legitimate study<sup>99</sup>.

**Molecular and morphological phylogeny of Zoraptera.** To elucidate the intraordinal relationships within Zoraptera, morphological data from extant and fossil species as well as five gene markers were integrated in a total evidence phylogenetic reconstruction. The molecular markers used were the publicly available sequences of the nuclear *18S rRNA*, and *H3*, and the mitochondrial *12S rRNA*, *16S rRNA*, and *COI* mined from GenBank. In total, 18 zorapteran taxa identified to species level alongside five plecopteran and dermapteran outgroups were included. Accessions are provided in Supplementary Table 2. The sequences of *COI* and *H3* were unambiguously aligned owing to few gaps and their codon-based structure using the MUSCLE algorithm implemented in the MEGA X<sup>100</sup>. The ribosomal RNAs *12S*, *16S*, and *18S*, were aligned in MAFFT using the E-INS-I algorithm<sup>101</sup>. The third codon of *COI* was excluded to reduce data heterogeneity. We analysed our data with the Bayesian site-heterogeneous CAT-GTR+G model implemented in PhyloBayes. We also tested the fit of a range of the default maximum likelihood site-homogeneous models implemented in IQ-TREE. The results are displayed in Tab S3; GTR+F+R3 was the best-fitting model by the corrected Akaike Information Criterion (AICc = 20871.1) and TVM+F+I+G by the Bayesian Information Criterion (BIC = 21152.0), while JC had the poorest fit to the alignment overall (AICc = 23033.1; BIC = 23264.6). We thus used these three ML models to run further analyses of the nucleotide alignment.

For a morphological phylogenetic analysis, we scored 12 morphological characters for 10 ingroup taxa. The principal objective was to address the relationships of the extinct genus *Xenozorotypus*, the extinct subgenus *Octozoros*, and crown-group zorapterans (*Zorotypus s. s.*). Within the crown-group of *Zorotypus*, we sampled representatives of all key clades recovered by molecular analyses<sup>16</sup>. We included the unusual Miocene species *Z. goeleti* which possesses dimeric cerci and may thus represent the sister group to the remainder of *Zorotypus s. s.*<sup>24</sup>. A single plecopteran representative, *Perla cephalotes*, was used as the outgroup. The character matrix is available in Tabs S4 and S5. Maximum parsimony analyses were conducted in TNT v. 1.5<sup>102</sup> using implied weighting. The recommended concavity value (*K*) of 12 was used, as these have been shown to achieve higher accuracy against homoplastic characters<sup>103</sup>. Collapsing rules were set to ‘none’ and the analysis was run using default settings in ‘New Technology Search’. A majority-rule consensus tree of the resultant four most parsimonious trees was computed. To assess tree support, nonparametric bootstrap analysis was run with 1,000 replicates. Character states were mapped using ASADO v. 1.61<sup>104</sup>.

### Data availability

Analysed files and results have been uploaded to MendeleyData repository (doi: 10.17632/pk747fvxxp.1).

## References

1. Grimaldi, D. & Engel, M. S. *Evolution of the Insects*. (Cambridge University Press, 2005).
2. Wipfler, B., Letsch, H., Frandsen, P. B. & et al. Evolutionary history of Polyneoptera and its implications for our understanding of early winged insects. *Proc. Natl. Acad. Sci.* **116**, 3024–3029 (2019).
3. Beutel, R. G., Wipfler, B., Gottardo, M. & Dallai, R. Polyneoptera or “Lower Neoptera”-new light on old and difficult phylogenetic problems. *Atti Accad. Naz. Ital. Entomol.* **61**, 113–142 (2013).
4. Misof, B. *et al.* Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763–767 (2014).
5. Kjer, K. M., Simon, C., Yavorskaya, M. & Beutel, R. G. Progress, pitfalls and parallel universes: a history of insect phylogenetics. *J. R. Soc. Interface* **13**, 20160363 (2016).
6. Letsch, H. & Simon, S. Insect phylogenomics: New insights on the relationships of lower neopteran orders (Polyneoptera). *Syst. Entomol.* **38**, 783–793 (2013).
7. Evangelista, D. A. *et al.* An integrative phylogenomic approach illuminates the evolutionary history of cockroaches and termites (Blattodea). *Proc. Roy. Soc. B.* **286**, 20182076 (2019).
8. Chesters, D. The phylogeny of insects in the data-driven era. *Syst. Entomol.* **45**, 540–551 (2019).
9. Beutel, R. G., Yavorskaya, M. I., Mashimo, Y., Fukui, M. & Meusemann, K. The phylogeny of Hexapoda (Arthropoda) and the evolution of megadiversity. *Proc. Arthropod. Embryol. Soc. Jpn.* **51**, 1–15 (2017).
10. Trautwein, M. D., Wiegmann, B. M., Beutel, R., Kjer, K. M. & Yeates, D. K. Advances in insect phylogeny at the dawn of the postgenomic era. *Annu. Rev. Entomol.* **57**, 449–468 (2012).
11. Tong, K. J., Duchêne, S., Ho, S. Y. W. & Lo, N. Comment on “Phylogenomics resolves the timing and pattern of insect evolution”. *Science* **349**, 487–487 (2015).

12. Gaunt, M. W. & Miles, M. A. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. *Mol. Biol. Evol.* **19**, 748–761 (2002).
13. Davis, R. B., Baldauf, S. L. & Mayhew, P. J. Many hexapod groups originated earlier and withstood extinction events better than previously realized: inferences from supertrees. *Proc. Roy. Soc. B.* **277**, 1597–1606 (2010).
14. Wang, Y. *et al.* Fossil record of stem groups employed in evaluating the chronogram of insects (Arthropoda: Hexapoda). *Sci. Rep.* **6**, 38939 (2016).
15. Montagna, M. *et al.* Recalibration of the insect evolutionary time scale using Monte San Giorgio fossils suggests survival of key lineages through the End-Permian Extinction. *Proceedings of the Royal Society B: Biological Sciences* **286**, 20191854 (2019).
16. Matsumura, Y. *et al.* The evolution of Zoraptera. *Syst. Entomol.* **45**, 349–364 (2020).
17. Silvestri, F. Descrizione di un nuovo ordine di insetti. *Boll. Lab. Zool. Portici* **7**, 193–209 (1913).
18. Mashimo, Y. *et al.* 100 years Zoraptera—a phantom in insect evolution and the history of its investigation. *Insect Syst. Evol.* **45**, 371–393 (2014).
19. Kukalová-Peck, J. & Peck, S. B. Zoraptera wing structures: evidence for new genera and relationship with the blattoid orders (Insecta: Blattoneoptera). *Syst. Entomol.* **18**, 333–350 (1993).
20. Hennig, W. *Die Stammesgeschichte der Insekten*. (Waldemar Kramer, 1969).
21. Beutel, R. G. & Weide, D. Cephalic anatomy of *Zorotypus hubbardi* (Hexapoda: Zoraptera): New evidence for a relationship with Acercaria. *Zoomorphology* **124**, 121–136 (2005).
22. Wheeler, W. C., Whiting, M., Wheeler, Q. D. & Carpenter, J. M. The phylogeny of the extant hexapod orders. *Cladistics* **17**, 113–169 (2001).
23. Terry, M. D. & Whiting, M. F. Mantophasmatodea and phylogeny of the lower neopterous insects. *Cladistics* **21**, 240–257 (2005).

24. Engel, M. S. & Grimaldi, D. A. A winged *Zorotypus* in Miocene amber from the Dominican Republic (Zoraptera: Zorotypidae), with discussion on relationships of and within the order. *Acta Geol. Hispan.* **35**, 149–164 (2000).
25. Boudreaux, H. B. *Arthropod Phylogeny with Special Reference to Insects*. (Wiley-Interscience, 1979).
26. Wang, Y. *et al.* A unique box in 28S rRNA is shared by the enigmatic insect order Zoraptera and Dictyoptera. *PLoS One* **8**, (2013).
27. Yoshizawa, K. & Johnson, K. P. Aligned 18S for Zoraptera (Insecta): phylogenetic position and molecular evolution. *Mol. Phylogenet. Evol.* **37**, 572–580 (2005).
28. Lozano-Fernandez, J. *et al.* Pancrustacean evolution illuminated by taxon-rich genomic-scale data sets with an expanded remipede sampling. *Genome Biol. Evol.* **11**, 2055–2070 (2019).
29. Engel, M. S. & Grimaldi, D. A. The first Mesozoic Zoraptera (Insecta). *Am. Mus. Novit.* **2002**, 1–20 (2002).
30. Philippe, H. *et al.* Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biol.* **9**, e1000602 (2011).
31. Tihelka, E. *et al.* Fleas are parasitic scorpionflies. *Palaeontol.* **3**, 641–653 (2020).
32. Kapli, P., Flouri, T. & Telford, M. J. Systematic errors in phylogenetic trees. *Curr. Biol.* **31**, R59–R64 (2021).
33. Redmond, A. K., McLysaght, A. & McLysaght, A. Evidence for sponges as sister to all other animals from partitioned phylogenomics with mixture models and recoding. *Nat. Commun.* **12**, 1783 (2021).
34. Lartillot, N., Brinkmann, H. & Philippe, H. Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. *BMC Evol. Biol.* **7**, S4 (2007).
35. Feuda, R. *et al.* Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. *Curr. Biol.* **27**, 3864–3870.e4 (2017).

36. Lartillot, N., Rodrigue, N., Stubbs, D. & Richer, J. PhyloBayes MPI: Phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* **62**, 611–615 (2013).
37. Pisani, D. *et al.* Genomic data do not support comb jellies as the sister group to all other animals. *Proc. Natl. Acad. Sci.* **112**, 15402–15407 (2015).
38. Lozano-Fernandez, J. *et al.* Increasing species sampling in chelicerate genomic-scale datasets provides support for monophyly of Acari and Arachnida. *Nat. Commun.* **10**, 1–8 (2019).
39. Lartillot, N. & Philippe, H. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21**, 1095–1109 (2004).
40. Simon, S., Narechania, A., DeSalle, R. & Hadrys, H. Insect phylogenomics: exploring the source of incongruence using new transcriptomic data. *Genome Biol. Evol.* **4**, 1295–1309 (2012).
41. Engel, M. Phylogeny of the Zoraptera. *Entomol. Abh.* **61**, 147–148 (2003).
42. Cai, C. *et al.* Beetle pollination of cycads in the Mesozoic. *Curr. Biol.* **28**, 2806–2812.e1 (2018).
43. Choe, J. C. Zoraptera of Panama with a review of the morphology, 1 systematics, and 2 biology of the order. in *Insects of Panama and 3 Mesoamerica: Selected Studies* (eds. Quintero, D. & Aiello, A.) 249–256 (Oxford University Press, 1992).
44. Guo, Y., Béthoux, O., Gu, J. & Ren, D. Wing venation homologies in Pennsylvanian ‘cockroachoids’ (Insecta) clarified thanks to a remarkable specimen from the Pennsylvanian of Ningxia (China). *J. Syst. Palaeontol.* **11**, 41–46 (2013).
45. Parham, J. F. *et al.* Best practices for justifying fossil calibrations. *Syst. Biol.* **61**, 346–359 (2011).
46. Rota-Stabelli, O., Daley, A. C. & Pisani, D. Molecular timetrees reveal a Cambrian colonization of land and a new scenario for ecdysozoan evolution. *Curr. Biol.* **23**, 392–398 (2013).
47. Rehm, P. *et al.* Dating the arthropod tree based on large-scale transcriptome data. *Mol. Phylogenet. Evol.* **61**, 880–887 (2011).
48. Schachat, S. R., Labandeira, C. C., Saltzman, M. R., Cramer, B. D., Payne, J. L., & Boyce, C. K. Phanerozoic pO<sub>2</sub> and the early evolution of terrestrial animals. *Proc. R. Soc. Lond. B* **285**, 20172631 (2018).

49. Béthoux, O., Cui, Y., Kondratieff, B., Stark, B. & Ren, D. At last, a Pennsylvanian stem-stonefly (Plecoptera) discovered. *BMC Evol. Biol.* **11**, 248 (2011).
50. Schubnel, T., Perdu, L., Roques, P., Garrouste, R. & Nel, A. Two new stem-stoneflies discovered in the Pennsylvanian Avion locality, Pas-de-Calais, France (Insecta: 'Exopterygota'). *Alcheringa* **43**, 430–435 (2019).
51. Beutel, R. G., Yan, E. V. & Kukalová-Peck, J. Is †Skleropectera (†*Stephanastus*) an order in the stemgroup of Coleopterida (Insecta)? *Insect Syst. Evol.* **50**, 670–678 (2019).
52. Chao, R.-F. & Chen, C.-S. *Formosozoros newi*, a new genus and species of Zoraptera (Insecta) from Taiwan. *Pan-Pac. Entomol.* **76**, 24–27 (2000).
53. Kočárek, P., Horká, I. & Kundera, R. Molecular phylogeny and infraordinal classification of Zoraptera (Insecta). *Insects* **11**, 51 (2020).
54. Song, N. *et al.* Molecular phylogeny of Polyneoptera (Insecta) inferred from expanded mitogenomic data. *Sci. Rep.* **6**, 36175 (2016).
55. Ma, C., Wang, Y., Wu, C., Kang, L. & Liu, C. The compact mitochondrial genome of *Zorotypus medoensis* provides insights into phylogenetic position of Zoraptera. *BMC Genomics* **15**, 1156 (2014).
56. Friedrich, F. & Beutel, R. G. The thorax of *Zorotypus* (Hexapoda, Zoraptera) and a new nomenclature for the musculature of Neoptera. *Arthropod Struct. Devel.* **37**, 29–54 (2008).
57. Sauphanor, B. & Sureau, F. Aggregation behaviour and interspecific relationships in Dermaptera. *Oecologia* **96**, 360–364 (1993).
58. Choe, J. C. Sexual selection and mating system in *Zorotypus gurneyi* Choe (Insecta: Zoraptera): I. Dominance hierarchy and mating success. *Behav. Ecol. Sociobiol.* **34**, 87–93 (1994).
59. Matsumura, Y. *et al.* Cephalic anatomy of *Zorotypus weidneri* New, 1978: new evidence for a placement of Zoraptera. *Arthr. Syst. Phyl.* **73**, 85–105 (2015).
60. Melters, D. P., Paliulis, L. V., Korf, I. F. & Chan, S. W. L. Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. *Chromosome Res.* **20**, 579–593 (2012).



61. Sylvester, T. *et al.* Lineage-specific patterns of chromosome evolution are the rule not the exception in Polyneoptera insects. *Proc. Roy. Soc. B.* **287**, 20201388 (2020).
62. Kuznetsova, V. G., Nokkala, S. & Shcherbakov, D. E. Karyotype, reproductive organs, and pattern of gametogenesis in *Zorotypus hubbardi* Caudell (Insecta: Zoraptera, Zorotypidae), with discussion on relationships of the order. *Can. J. Zool.* (2002) doi:10.1139/z02-074.
63. Hünefeld, F. The genital morphology of *Zorotypus hubbardi* Caudell, 1918 (Insecta: Zoraptera: Zorotypidae). *Zoomorphol.* **126**, 135–151 (2007).
64. Weidner, H. Die Ordnung Zoraptera oder Bodenläuse. *Entomol. Z.* **79**, 29–51 (1969).
65. Haas, F. & Kukulová-Peck, J. Dermaptera hindwing structure and folding: new evidence for familial, ordinal and superordinal relationships within Neoptera (Insecta). *Eur. J. Entomol.* **98**, 445–509 (2001).
66. Yoshizawa, K. Monophyletic Polyneoptera recovered by wing base structure. *Syst. Entomol.* **36**, 377–394 (2011).
67. Kristensen, N. P. Phylogeny of extant hexapods. in *The Insects of Australia, 2nd ed.* 125–140 (Melbourne University Press, 1991).
68. Misof, B. *et al.* Towards an 18S phylogeny of hexapods: accounting for group-specific character covariance in optimized mixed nucleotide/doublet models. *Zool.* **110**, 409–429 (2007).
69. Shelomi, M., Wipfler, B., Zhou, X. & Pauchet, Y. Multifunctional cellulase enzymes are ancestral in Polyneoptera. *Insect Mol. Biol.* **29**, 124–135 (2020).
70. Dahl, T. W. *et al.* Devonian rise in atmospheric oxygen correlated to the radiations of terrestrial plants and large predatory fish. *Proc. Natl. Acad. Sci.* **107**, 17911–17915 (2010).
71. Stein, W. E., Berry, C. M., Hernick, L. V. & Mannolini, F. Surprisingly complex community discovered in the mid-Devonian fossil forest at Gilboa. *Nature* **483**, 78–81 (2012).
72. Strullu-Derrien, C. *et al.* The earliest wood and its hydraulic properties documented in c. 407-million-year-old fossils using synchrotron microtomography. *Bot. J. Linn. Soc.* **175**, 423–437 (2014).

73. Hetherington, A. J., Dolan, L. & Dolan, L. Stepwise and independent origins of roots among land plants. *Nature* **561**, 235–238 (2018).
74. Figueroa, J. M. T. de & López-Rodríguez, M. J. Trophic ecology of Plecoptera (Insecta): a review. *Eur. J. Zool.* **86**, 79–102 (2019).
75. Whalley, P. & Jarzembowski, E. A. A new assessment of *Rhyniella*, the earliest known insect, from the Devonian of Rhynie, Scotland. *Nature* **291**, 317–317 (1981).
76. Dunlop, J. A. & Garwood, R. J. Terrestrial invertebrates in the Rhynie chert ecosystem. *Phil. Trans. Roy. Soc. B.* **373**, 20160493 (2018).
77. Berbee, M. L. *et al.* Genomic and fossil windows into the secret lives of the most ancient fungi. *Nat. Rev. Microbiol.* **18**, 717–730 (2020).
78. Gilbert, J. D. J. & Manica, A. The evolution of parental care in insects: A test of current hypotheses. *Evolution* **69**, 1255–1270 (2015).
79. Nel, A., Aria, C., Garrouste, R. & Waller, A. Evolution and palaeosynecology of the Mesozoic earwigs (Insecta: Dermaptera). *Cretac. Res.* **33**, 189–195 (2012).
80. Kukalova-Peck, J. Origin and evolution of insect wings and their relation to metamorphosis, as documented by the fossil record. *J. Morphol.* **156**, 53–125 (1978).
81. Marden, J. H. & Kramer, M. G. Surface-skimming stoneflies: a possible intermediate stage in insect flight evolution. *Science* **266**, 427–430 (1994).
82. Criscuolo, A. & Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, 210 (2010).
83. Wang, H.-C., Susko, E. & Roger, A. J. The relative importance of modeling site pattern heterogeneity versus partition-wise heterotachy in phylogenomic inference. *Syst. Biol.* **68**, 1003–1019 (2019).
84. Blanquart, S. & Lartillot, N. A site- and time-heterogeneous model of amino acid replacement. *Mol. Biol. Evol.* **25**, 842–858 (2008).

85. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
86. Le, S. Q., Dang, C. C. & Gascuel, O. Modeling protein evolution with several amino acid replacement matrices depending on site rates. *Mol. Biol. Evol.* **29**, 2921–2936 (2012).
87. Si Quang, L., Gascuel, O. & Lartillot, N. Empirical profile mixture models for phylogenetic reconstruction. *Bioinformatics* **24**, 2317–2323 (2008).
88. Zhang, Z., Schneider, J. W. & Hong, Y. The most ancient roach (Blattodea): a new genus and species from the earliest Late Carboniferous (Namurian) of China, with a discussion of the phylomorphogeny of early blattids. *J. Syst. Palaeontol.* **11**, 27–40 (2013).
89. Prokop, J. *et al.* Late Palaeozoic Paoliida is the sister group of Dictyoptera (Insecta: Neoptera). *J. Syst. Palaeontol.* **12**, 601–622 (2014).
90. Vršanský, P., Liang, J. H. & Ren, D. Advanced morphology and behaviour of extinct earwig-like cockroaches (Blattida: Fuziidae fam. nov.). *Acta Geol. Carpat.* **60**, 449–462.
91. Correia, P., Schubnel, T. & Nel, A. What is the roachoid genus *Eneriblatta* (Dictyoptera: Phyloblattidae) from the Carboniferous of Portugal. *Hist. Biol.* **in press**, (2019).
92. Trümper, S. *et al.* Age and depositional environment of the Xiaheyan insect fauna, embedded in marine black shales (Early Pennsylvanian, China). *Palaeogeog. Palaeoclim. Palaeoecol.* **538**, 109444 (2020).
93. Jeram, A. J., Selden, P. A. & Edwards, D. Land animals in the Silurian: arachnids and myriapods from Shropshire, England. *Science* **250**, 658–661 (1990).
94. Ogg, J. G., Ogg, G. & Gradstein, F. M. *A Concise Geologic Time Scale: 2016*. (Elsevier, 2016).
95. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
96. dos Reis, M. & Yang, Z. Approximate likelihood calculation on a phylogeny for Bayesian estimation of divergence times. *Mol. Biol. Evol.* **28**, 2161–2172 (2011).

97. Shi, G. *et al.* Age constraint on Burmese amber based on U–Pb dating of zircons. *Cretaceous Res.* **37**, 155–163 (2012).
98. Mao, Y. Y. *et al.* Various amberground marine animals on Burmese amber with discussions on its age. *Palaeontomol.* **1**, 91–103 (2018).
99. Engel, M. S. Myanmar: palaeontologists must stop buying conflict amber. *Nature* **584**, 525–525 (2020).
100. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
101. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
102. Goloboff, P. A. & Catalano, S. A. TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics* **32**, 221–238 (2016).
103. Goloboff, P. A., Torres, A. & Arias, J. S. Weighted parsimony outperforms other methods of phylogenetic inference under models appropriate for morphology. *Cladistics* **34**, 407–437 (2018).
104. Nixon, K. C. *ASADO version 1.5 Beta. Program and Documentation Distributed by the Author.* (Self-published, 2004).

### **Authors' contributions**

C.C., E.T. and P.C.J.D. conceived and designed the study, C.C. and E.T. conducted the phylogenomic analyses, E.T. conducted topology tests, the molecular clock analyses, and analyses of zorapteran intraordinal phylogeny, Z.Y. and C.C. processed the photomicrograph data and studied *Z. nascimbenei* morphology. E.T., Z.Y. and C.C. drafted the manuscript, to which J.L.-F., M.S.E., D.H., M.G., O. R.-S., P.C.J.D. and D.P. contributed.

### **Competing interests**

The authors declare no competing interests.

**Correspondence** and requests for materials should be addressed to C.C.

**Funding.** This work has been supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB26000000), the National Natural Science Foundation of China (42072022, 41688103), the Second Tibetan Plateau Scientific Expedition and Research project (2019QZKK0706), the Biotechnology and Biological Sciences Research Council (BB/N000919/1; BB/T012773/1), the High Performance Computing facility at the University of Bristol and a Newton International Fellowship from the Royal Society awarded to C.C. and P.C.J.D.

**Figure 1.** Timescale of Polyneoptera evolution from analyses with the site-heterogeneous model CAT-GTR+G of the 100-taxon dataset. All nodes are fully supported (BPP = 1). Ages were estimated based on 42 calibrated nodes, integrating the results of analyses using independent rates (IR) and autocorrelated rates (AC) molecular clock models in MCMCTree. Abbreviations: Noto., Notoptera; Sil., Silurian; Q., Quaternary. Numbered nodes indicate the calibrations, see Supplementary Information for full list of calibrations and Fig. S1 for a full calibrated tree.

**Figure 2.** Competing hypotheses of Zorapteran placement, compared with Approximately Unbiased (AU) tests conducted with the LG+C60+F+G model in IQ-TREE. **a** Zoraptera first hypothesis supported by our analyses with the CAT-GTR+G and LG+C60+F+G models and transcriptomic analyses in Simon et al.<sup>40</sup> and Letsch and Simon<sup>6</sup>. **b** Hapllocercata first hypothesis supported by transcriptomic analyses of Misof et al.<sup>4</sup>, Wipfler et al.<sup>2</sup>. **c** Dermaptera first hypothesis, sister to remaining Polyneoptera recovered by Wipfler et al.<sup>2</sup> with the multispecies coalescent method. **d** Zoraptera + Plecoptera first hypothesis, supported by Letsch and Simon<sup>6</sup> analysis using unreduced matrix.  $P$ -value > 0.05: topology not rejected;  $P$ -value < 0.05: topology rejected significantly;  $P$ -value = 0: topology rejected with high significance. The 20 tested topologies are listed in full in Tab. S1.

**Figure 3.** Stem-group zorapteran *Zorotypus nascimbenei* in mid-Cretaceous amber from northern Myanmar (~99 Ma; NIGP175112). **a** Habitus in ventral view. **b** mouthparts in dorsal view. **c** Abdominal apex in dorsal view. **d** Head-prothorax junction in dorsal view. **e** Metafemur and metatibia. **f** Metatibial apex and metatarsi. Abbreviations: a1–8, antennomeres 1–8; abd, abdomen; ce, cercus; cr, cervix; cl, pretarsal claw; fs, fungal spore; fsc, cluster of fungal spores; fsc-abd; putative fungal spores preserved inside the abdomen; g, galea, gb, apical galeal setose brush; gu, gula; la, lacinia; lb, labial palpus; md, mandible; mp, maxillary palpus; mst, mesothorax; mtt, metathorax; mtt1–2, metatarsomeres 1–2; mttf, metafemur; mtti, metatibia; pm, postmentum; prm, prementum; pt, prothorax, sp1–6, metafemoral spines 1–6; spB, basal metafemoral spine. Scale bars: 200  $\mu\text{m}$  (**a**), 100  $\mu\text{m}$  (**c–f**), 50  $\mu\text{m}$  (**b**).



**Figure 4.** Intraordinal relationships within Zoraptera. **a** Five-gene nucleotide alignment analysed with the CAT-GTR+G model in PhyloBayes. **b** Morphological dataset sampling extinct species, under maximum parsimony in TNT (CI = 0.895; RI = 0.778); character states are mapped in Fig. S31. Support values presented as BPP (**a**) and bootstrap values with 1,000 replicates (**b**).

**Table 1.** Ninety-five per cent HPD age estimates of divergences between major winged insect (Pterygota) crown groups. Full dated trees are provided in Figs S24 and S25.

<b>Clade</b>	<b>95% HDP (Ma)</b>
Pterygota	427–399
Palaeoptera	411–367
Acercaria	387–342
Holometabola	344–390
Polyneoptera	402–365
Zoraptera	402–365
Dermopteroptera	368–315
Dermaptera	170–128
Plecoptera	315–223
Orthoptera	317–284
Orthoptera	317–284
Notoptera	307–180
Mantophasmatodea	96–7
Grylloblattodea	136–30
Eukinolabia	300–249
Embioptera	170–123
Phasmatodea	276–200
Dictyoptera	290–227
Blattodea	245–161
Isoptera	187–135
Mantodea	214–105

**Table 2.** Updated classification of Zoraptera proposed herein, highlighting adopted taxonomic changes. A full checklist of described species and their geographical distribution is provided in the Supplementary Information.

---

Family Zorotypidae Silvestri, 1913 **sensu nov.**

Genus †*Xenozorotypus* Engel & Grimaldi, 2002

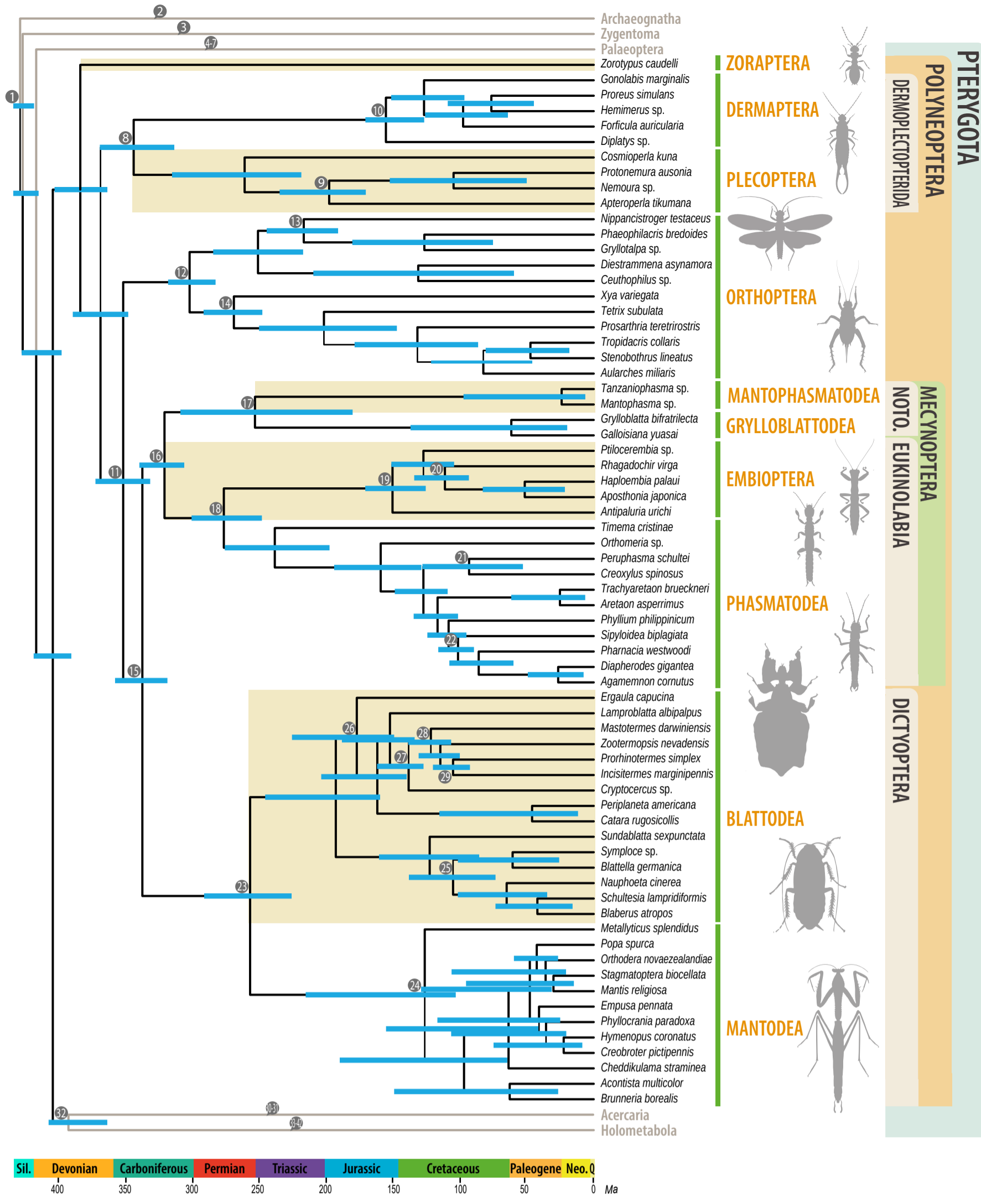
Genus *Zorotypus* Silvestri, 1913 **sensu nov.**

Subgenus †*Octozoros* Engel, 2003 **stat. nov.**

Subgenus *Zorotypus* Silvestri, 1913 **stat. rev.**

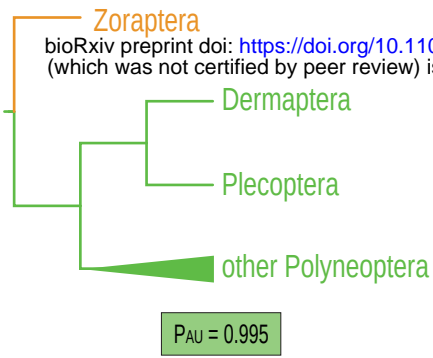
Subgenus *Centrozoros* Kukalová-Peck & Peck, 1993 **stat. rev.**

---

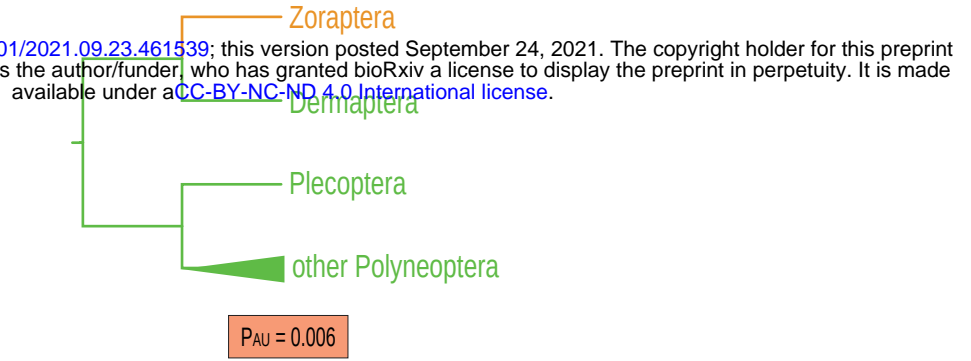


bioRxiv preprint doi: <https://doi.org/10.1101/2021.09.23.461539>; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

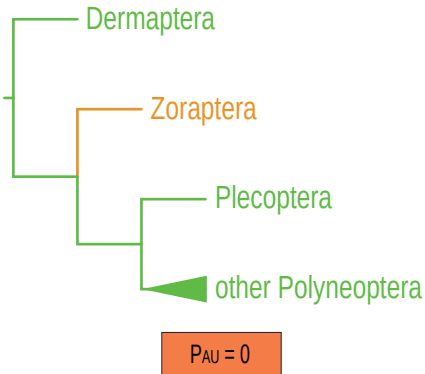
### a Zoraptera first



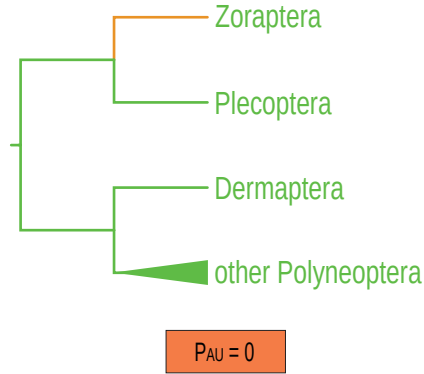
### b Haplocercata first



### c Dermaptera first

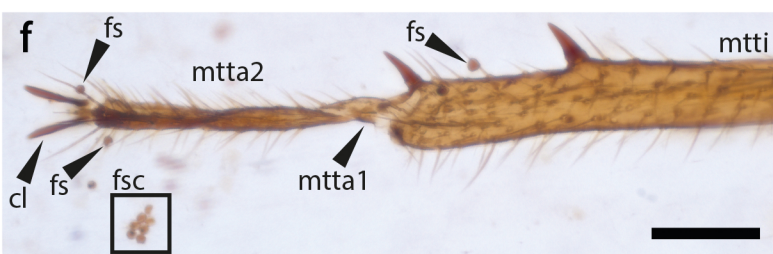
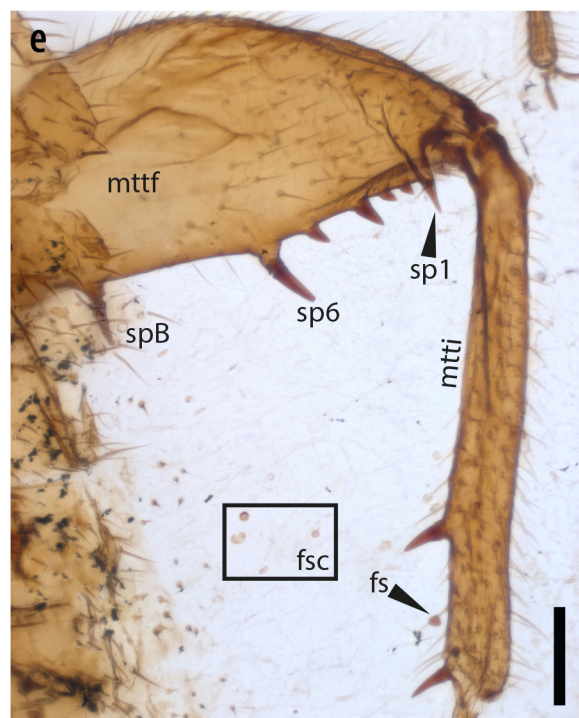
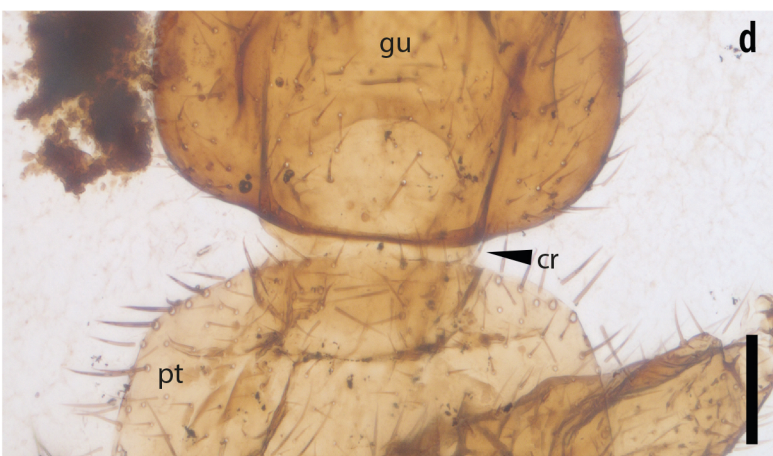
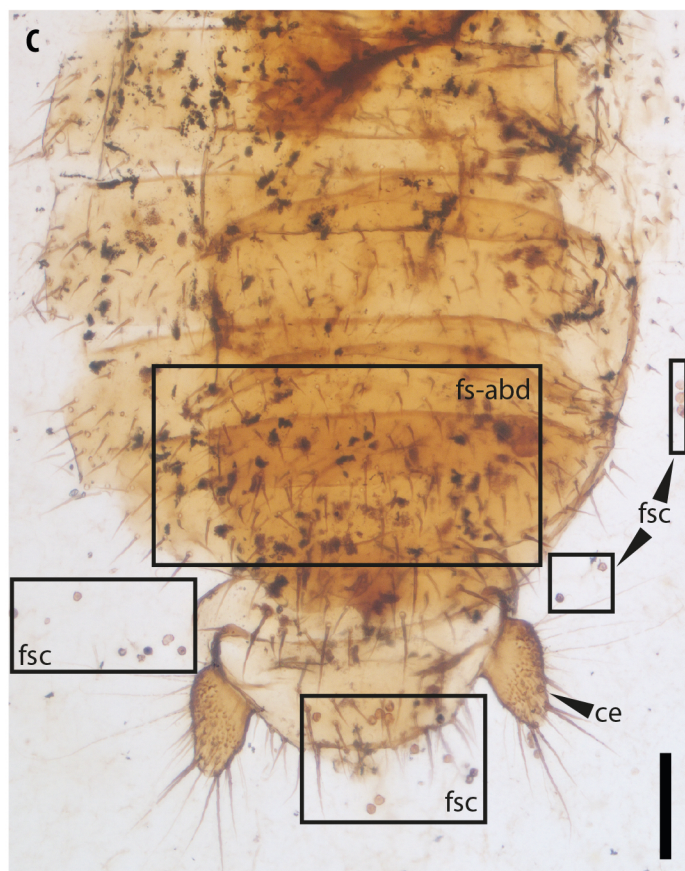
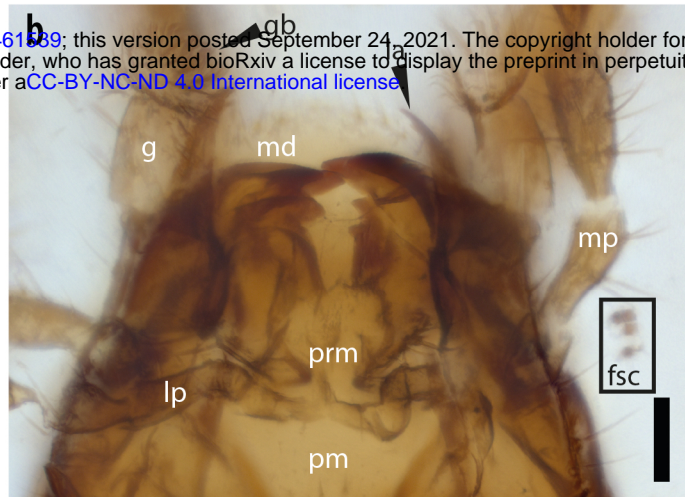
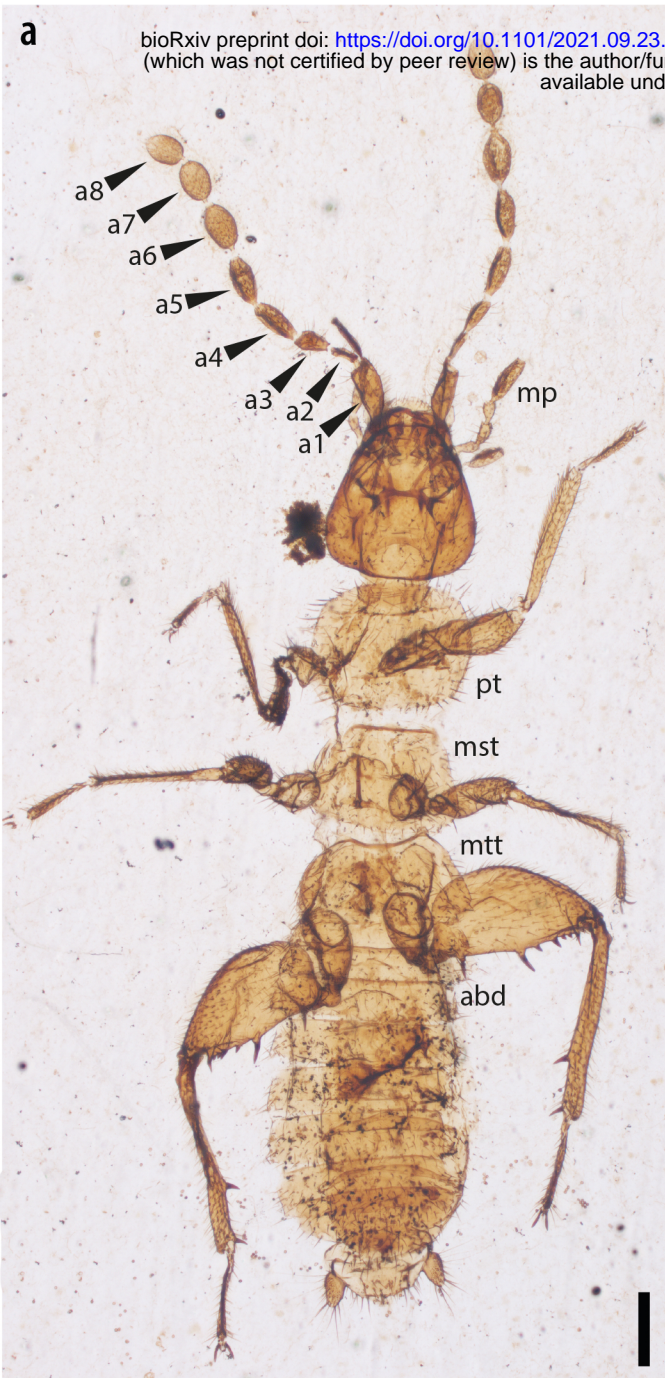


### d Zoraptera + Plecoptera first

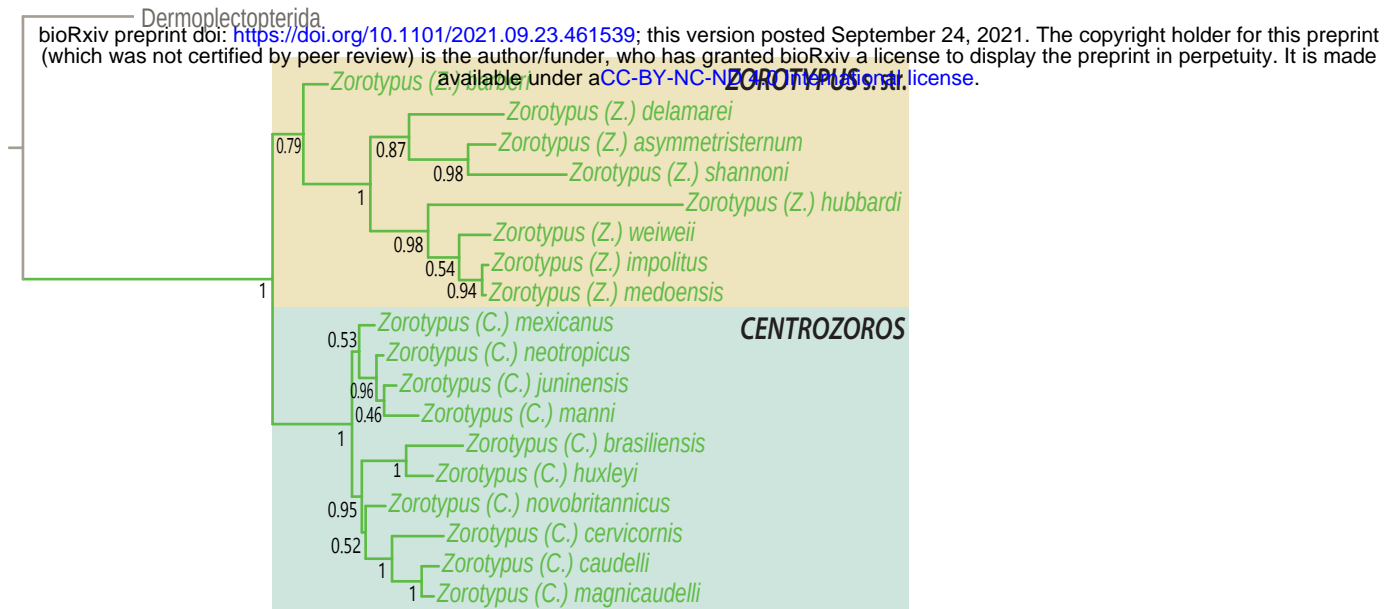


P-value





## a 5-gene dataset



## b Morphology

