Cladistics (2020) 1–22

Cladistics

10.1111/cla.12436

Phylogenomics reveals accelerated late Cretaceous diversification of bee flies (Diptera: Bombyliidae)

Xuankun Li*[a,b], Luisa C. Teasdale[a], Keith M. Bayless[a], Allan G. Ellis[c], Brian M. Wiegmann[d], Carlos José E. Lamas[e], Christine L. Lambkin[f], Neal L. Evenhuis[g], James A. Nicholls[a], Diana Hartley[a], Seunggwan Shin[h,i], Michelle Trautwein[j], Andreas Zwick[a], Bryan D. Lessard[a] and David K. Yeates*[a]

[a]Australian National Insect Collection, CSIRO National Research Collections Australia, Canberra, ACT, 2601, Australia; [b]Research School of Biology, Australian National University, Canberra, ACT, 2601, Australia; [c]Botany and Zoology Department, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa; [d]Department of Entomology & Plant Pathology, North Carolina State University, Raleigh, NC, 27695, USA; [e]Museu de Zoologia da Universidade de São Paulo. Avenida Nazaré, 481 Ipiranga 04263-000, São Paulo, SP, Brazil; [f]Queensland Museum, PO Box 3300, South Brisbane, Qld, 4101, Australia; [g]J. Linsley Gressitt Center for Research in Entomology, Bishop Museum, 1525 Bernice Street, Honolulu, HI, 96817, USA; [h]Department of Biological Sciences, University of Memphis, Memphis, TN, 38152, USA; [i]School of Biological Sciences, Seoul National University, Seoul, 08826, Korea; [j]Entomology Department, Institute for Biodiversity Science and Sustainability, California Academy of Sciences, San Francisco, CA, 94118, USA

Received 23 December 2019; Revised 7 July 2020; Accepted 18 August 2020

Abstract

Bombyliidae is a very species-rich and widespread family of parasitoid flies with more than 250 genera classified into 17 extant subfamilies. However, little is known about their evolutionary history or how their present-day diversity was shaped. Transcriptomes of 15 species and anchored hybrid enrichment (AHE) sequence captures of 86 species, representing 94 bee fly species and 14 subfamilies, were used to reconstruct the phylogeny of Bombyliidae. We integrated data from transcriptomes across each of the main lineages in our AHE tree to build a data set with more genes (550 loci versus 216 loci) and higher support levels. Our overall results show strong congruence with the current classification of the family, with 11 out of 14 included subfamilies recovered as monophyletic. Heterotropinae and Mythicomyiinae are successive sister groups to the remainder of the family. We examined the evolution of key morphological characters through our phylogenetic hypotheses and show that neither the "sand chamber subfamilies" nor the "Tomophthalmae" are monophyletic in our phylogenomic analyses. Based on our results, we reinstate two tribes at the subfamily level (Phthiriinae stat. rev. and Ecliminae stat. rev.) and we include the genus *Sericosoma* Macquart (previously *incertae sedis*) in the subfamily Oniromyiinae, bringing the total number of bee fly subfamilies to 19. Our dating analyses indicate a Jurassic origin of the family (165–194 Ma), with the sand chamber evolving early in bee fly evolution, in the late Jurassic or mid-Cretaceous (100–165 Ma). We hypothesize that the angiosperm radiation and the hothouse climate established during the late Cretaceous accelerated the diversification of bee flies, by providing an expanded range of resources for the parasitoid larvae and nectarivorous adults.

Key words: Diptera, divergence times reconstruction, lower Brachycera, maximum likelihood

---

*Corresponding author:
E-mail address: xuankun.li@floridamuseum.ufl.edu (XL);david.yeates@csiro.au (DKY)

Introduction

Bombyliidae (bee flies) is one of the most species-rich families in lower Brachycera (Diptera), comprising 17 subfamilies, over 250 genera and over 5000 described species (Yeates, 1994; Evenhuis and Greathead, 1999; Yamaguchi et al., 2017). Although bee flies are divided into many subfamilies, over half of the generic and specific diversity is contained within two subfamilies: Anthracinae and Bombyliinae. Species of Bombyliidae display wide morphological diversity, ranging from the tiny, delicate, humpbacked black species of *Geron* Meigen, to bright, hairy, stout species of *Sisyromyia* White, and to elegant, elongate, yellow and black species of *Systropus* Wiedemann which are convincing sphecid wasp mimics. They range in size from approximately 2.5 mm wingspan in the genus *Glabellula* Bezzi, up to 65 mm wingspans in the genera *Palirika* Lambkin and Yeates and *Comptosia* Macquart (Yeates, 1994; Lambkin et al., 2003). Adult bee flies feed on nectar and pollen and are important pollinators (Armstrong, 1979; Heard et al., 1990). They play an important role in ecological communities and may have contributed to the diversification of some modern angiosperms (De Jager and Ellis, 2017). However, many questions about their evolutionary history remain unresolved, and the drivers that shaped their extant diversity are not settled.

Tachinidae and Bombyliidae are distantly related families in the Diptera Tree of Life (Wiegmann et al., 2011), but also the two largest parasitoid fly families. Their larvae are almost exclusively parasitoids, feeding on the immatures of a wide range of almost exclusively insect hosts including Orthoptera, Neuroptera, Coleoptera, Diptera, Lepidoptera and Hymenoptera (Yeates and Greathead, 1997; Stireman et al., 2006). All known Tachinidae are exclusively endoparasitoids with high host specificity driven by female host selection. Tachinidae is one of the largest families in Diptera and has approximately 10 000 described species (Pape et al., 2011), with speciation potentially driven by host specialization (Smith et al., 2007; Stireman et al., 2019). Bee flies, however, are mostly ectoparasitic and have broad host ranges, attacking multiple species at the level of host tribe or subfamily. While most adult females scatter eggs in suitable habitats, their planidia (first-instar larvae) make direct contact with the host. Because of this more indiscriminate behaviour of the adult females in some taxa, bombyliid larvae may be less host specific than tachinid larvae, resulting in lower species diversity than in Tachinidae (Yeates and Greathead, 1997). However, cryptic diversity has been found in some bee fly genera, where it may be driven by mate recognition systems (Yeates and Lambkin, 1998), suggesting that the current species diversity of Bombyliidae is underestimated. While significant progress has recently been made in understanding the relationships of Tachinidae (Cerretti et al., 2014; Winkler et al., 2015; Stireman et al., 2019), our knowledge of relationships in Bombyliidae has lagged behind. A study of the relationships of bee flies using two molecular markers (Trautwein et al., 2011) showed little concordance with the morphological phylogeny of Bombyliidae (Yeates, 1994).

Although bombyliids have long been considered a member of the superfamily Asiloidea, and sister to all the remaining asiloid families (Woodley, 1989; Yeates, 2002; Wiegmann et al., 2011), an alternate hypothesis suggests Bombyliidae is sister to the remaining Asiloidea + Eremoneura (Woodley et al., 2009; Trautwein et al., 2010; Sinclair et al., 2013). Some authors have treated the subfamily Mythicomyiinae as a separate family (e.g. Evenhuis and Greathead, 1999; Evenhuis, 2002). However, Mythicomyiinae falls within Bombyliidae in the most recent molecular studies, with a strongly supported monophyletic Bombyliidae that includes Mythicomyiinae, i.e. [Heterotropinae + (Mythicomyiinae + remaining Bombyliidae)] (Trautwein et al., 2010), and [(Heterotropinae + Mythicomyiinae) + remaining Bombyliidae] (Shin et al., 2018). Furthermore, the position of Mythicomyiinae is also supported as sister to the remaining Bombyliidae in morphological analyses (Yeates, 1994).

The great morphological diversity of Bombyliidae has always proved a challenge for phylogenetic reconstructions of the family based on morphology, and the relationships between bee fly subfamilies have not been resolved robustly or in a consistent manner across studies. In pioneering research, Becker (1913) indicated that Anthracinae was the earliest branch of the family, Usiinae and Aphoebantinae (now Aphoebantini of Anthracinae) were the most recently diverged subfamilies, and his tree also suggested that eight subfamilies were paraphyletic. Because of Becker's mistaken use of labile wing venation characters to divide the family into two large groups, this scheme was rejected by two modern phylogenetic studies (e.g. Mühlenberg, 1971; Yeates, 1994). Bezzi (1924) uprooted Becker's classification based on two different characters: the form of the hind eye margin in combination with the form of the postcranium. He proposed "Homeophthalmae" for the subfamilies with a simple hind eye margin and flat postcranium, and placed those subfamilies with an indentation on the hind margin of the eye and with a concave postcranium in "Tomophthalmae". He considered both groups to have originated from flies with simple hind eye margin and flat postcranium, and that Bombyliinae ("Homeophthalmae") diverged earlier than Anthracinae ("Tomophthalmae"). These two names ("Homeophthalmae" and "Tomophthalmae") proposed by Bezzi are not family-group names because they do not include a type genus (see Sabrosky, 1999).

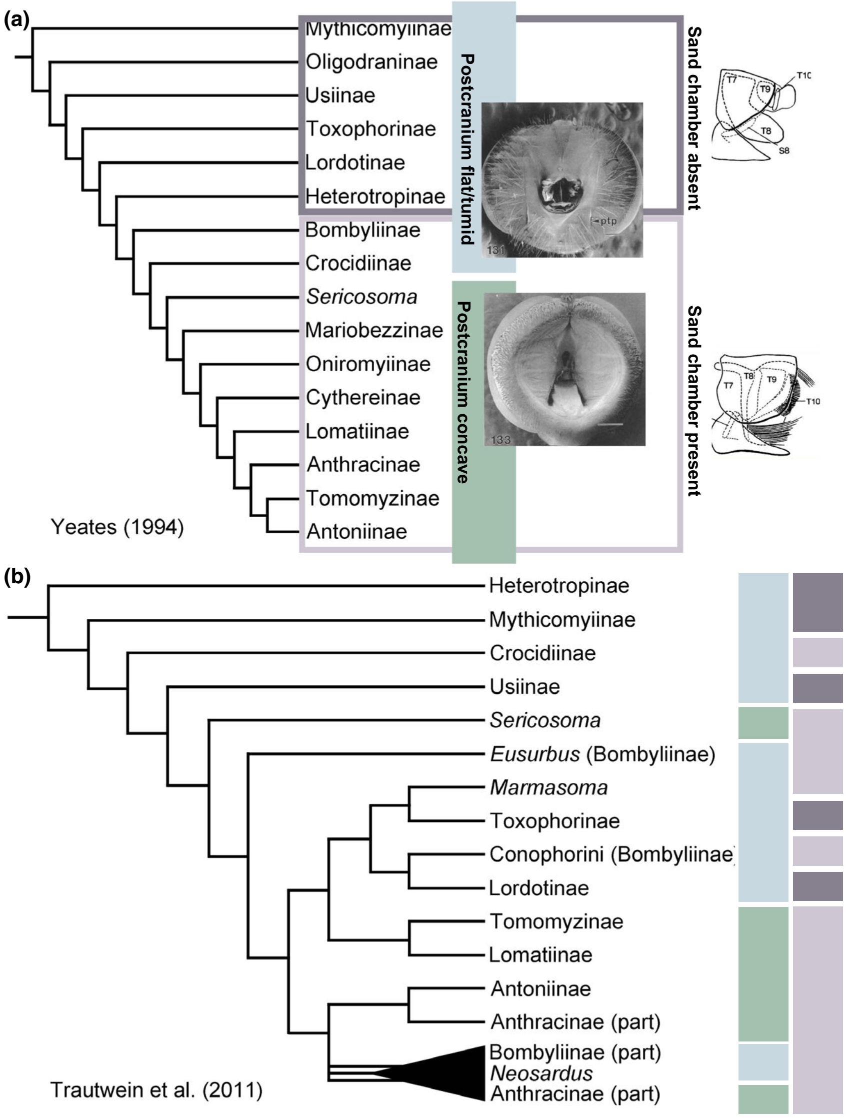


Fig. 1. Previous bombyliid phylogenies. (a) Morphological phylogeny by Yeates (1994); (b) Molecular phylogeny by Trautwein et al. (2011).

However, these names are commonly encountered when reviewing the history of Bombyliidae systematics over the past century. We therefore introduce them as we discuss previous classifications here, but exclude them from our figures to avoid confusion.

From the early 1970s to early 1990s, the subfamily classification of Bombyliidae was debated (e.g. Mühlenberg, 1971; Hull, 1973; Bowden, 1974; Hall, 1976; Theodor, 1983; Zaitsev, 1992), and disagreements persisted. Mühlenberg (1971) provided the first cladistic analysis of the Bombyliidae with 48 characters from female genitalia and eight characters from the head and wings. Mühlenberg's data matrix was reanalysed by Yeates (1994), showing that it did not support the monophyly of the "Tomophthalmae" or "Homeophthalmae". The most comprehensive taxonomic work on world Bombyliidae (Hull, 1973) introduced tribes into the classification to reduce the number of subfamilies, however, these were not widely accepted. The diagram of relationships proposed by Hull (1973) is similar to the classification by Becker (1913), and provided little additional insight. Zaitsev (1992) attempted to divide the Bombyliidae into families (Mythicomyiidae, Systropodidae, Phthiriidae, Usiidae and remaining Bombyliidae), revisiting the previous idea of Rohdendorf (1974). His work did not use a formal cladistic analytical method and was considered to be preliminary at that time. Zaitzev's hypothesis divided "Homeophthalmae" into small families without an assessment of their relationships.

Recent overall phylogenetic studies were based on morphological characters (Yeates, 1994) and molecular data (Trautwein et al., 2011), respectively. Yeates (1994) analyzed 154 characters from 63 Bombyliidae and 24 outgroup taxa, which: (i) reconstructed the phylogenetic relationships with 15 recognized subfamilies; (ii) recovered a monophyletic "Tomophthalmae" (with the apomorphic state of two occipital foramina) and a paraphyletic Homeophthalmae (with the plesiomorphic state of one occipital foramen); and (iii) supported as monophyletic the "sand chamber subfamilies" ("Psammophoridae" in Mühlenberg (1971)) (Fig. 1A). Trautwein et al. (2011) analyzed two nuclear gene fragments, protein-coding CAD (carbamoyl phosphate synthetase-aspartate transcarbamoylase-dihydroorotase) and ribosomal 28S, for 49 Bombyliidae and four outgroup taxa. Support was recovered for the monophyly of eight subfamilies and evidence was lacking to support the monophyly of "Tomophthalmae" or the "sand chamber subfamilies". Many backbone nodes were not strongly supported in their molecular phylogenetic hypotheses, demonstrating the complexity of bombyliid phylogenetics (Fig. 1B).

The sand chamber is arguably the most significant morphological character in bee fly evolution, and involves complex and integrated modifications of the tergites and sternites of female terminal abdominal segments eight, nine and ten. Morphological phylogenies have proposed that the sand chamber evolved once, but has been lost multiple times (Mühlenberg, 1970; Yeates, 1994). Female bee flies that possess a sand chamber (for example Bombyliinae, Anthracinae) are commonly seen dabbing sandy soil with the apex of their abdomen while they manipulate sand into their sand chamber with strong hair brushes on the posterior margin of tergite 8, and the acanthophorite spines on tergite 10; the sand is later used for coating their eggs. Females then flick the eggs into oviposition sites (e.g. bee nesting holes, burrows of tiger beetle larvae) while hovering. A recent study showed that the sand chamber can be lost in lineages that oviposit on the floor of a closed forest, where sand is not readily available (Li and Yeates, 2019). The adaptive significance of the sand coated eggs is unclear; it may be to reduce evaporation from the egg, or it may be to add weight to the egg so that it can be projected more firmly and accurately by the female.

To date, 85 species of fossil Bombyliidae belonging to 56 genera have been described, representing 10 subfamilies (Evenhuis, 2014; Myskowiak et al., 2016; Evenhuis, 2017; Ye et al., 2019). Most described bombyliid fossils are from the Cenozoic, especially the Miocene, Oligocene and Eocene, although more Mesozoic fossils have been described recently from mid-Cretaceous Burmese amber (Grimaldi, 2016; Ye et al., 2019). The oldest known bee fly fossil is from the Middle Jurassic (Mythicomyiinae: *Palaeoplatypygus zaitzevi* Kovalev; Kalugina and Kovalev, 1985), the earliest record of a developed sand chamber is from the mid-Cretaceous (Ye et al., 2019), and the oldest bee flies with a concave postcranium are described from the Eocene (Cythereinae: *Paleolomatia menatensis* Nel, 2008; and Lomatiinae: *Comptosia pria* Wedmann and Yeates, 2008). The origin of the Bombyliidae was estimated at around 170 Ma in the context of a phylogeny of all Diptera (Wiegmann et al., 2011), and the non-mythicomyiine lineage diverged 140–190 Ma (mean = 165 Ma) (Lamas and Nihei, 2007).

Here, we use phylogenomic data from hundreds of nuclear and mitochondrial genes to gain robust insights into the relationships of the Bombyliidae. We sequenced 18 transcriptomes and generated anchored hybrid enrichment (AHE) exome capture data for a further 95 terminals, including 94 bee fly species and 12 representatives of outgroup families. We generated matrices of 216 and 550 loci from these aligned sequence data (33 056 amino acid residues: 216AA; 99 174 nucleotides: 216NT; 131 672 amino acid residues: 550AA; 395 031 nucleotides: 550NT). Our sampled taxa represent 14 of 17 bombyliid subfamilies and 76% of the tribes. Three small subfamilies are missing from our analysis. Oligodraninae contains a single Palaearctic genus with three described species (Evenhuis, 1991). Xenoprosopinae is represented by South African *Xenonoprosopa paradoxa* Hesse, known from a single female specimen (Hesse, 1956; Hull, 1973; Yeates, 1994). Phthiraxiinae contains one species from Western Australia known from two specimens, and was recently described and subsequently raised to subfamily status (Li and Yeates, 2017; Yamaguchi et al., 2017). Our aim is to reconstruct the phylogeny of Bombyliidae, to test the monophyly of the included subfamilies, to reconstruct the relationships between the subfamilies, estimate the divergence times of major bee flies lineages, and trace the contributions of key innovations on their patterns of diversification.

Material and methods

Taxon sampling

Transcriptomes of 17 species and partial exomes based on Anchored Hybrid Enrichment (AHE) from 94 species were generated for the present study. Additionally, data for one outgroup taxon, *Hilarimorpha* sp.1 (Hilarimorphidae; Gillung et al., 2018) and one ingroup taxon, *Bombylius major* Linnaeus (Bombyliidae; Misof et al., 2014) were obtained from NCBI (Appendix S1).

Data included for ingroup Bombyliidae comprised 15 transcriptomes and AHE data for 86 species with the aim of representing as many subfamilies and tribes as possible. In total, 94 species of bee flies belonging to 76 genera were sequenced, representing 14 out of 17 subfamilies, and 29 out of 37 tribes. We sequenced the transcriptome and AHE data for two separate individuals of seven bee fly species (*Geron flavocciput* Evenhuis, *Comptosia australensis* (Schiner), *Thevenetimyia longipalpus* (Hardy), *Bombylius major*, *Mandella* sp.1, *Meomyia vetusta* (Walker), and *Anthrax maculatus* Macquart) to test for potential bias in the data collection and bioinformatics analyses strategy applied to these two different data sources. The classification we used was mainly based on the morphological phylogeny of Yeates (1994), with some new tribes established by more recent studies (Li and Yeates, 2017; Li and Yeates, 2019). The classification of Mythicomyiinae was based on the world catalog (Evenhuis, 2002) with subfamilies herein treated as tribes.

Transcriptomes of three species and AHE data from nine species in the lower Brachycera representing 12 families were used as outgroup taxa: Rhagionidae, Xylophagidae, Stratiomyidae, Tabanidae, Nemestrinidae, Hilarimorphidae, Asilidae, Mydidae, Apioceridae, Evocoidae, Therevidae and Scenopinidae.

After collection, specimens were initially preserved in 95% ethanol or RNAlater in the field, and then transferred to −80 °C for long-term storage (Appendix S1) at the Australian National Insect Collection (Canberra, Australia). Specimen collection and identification information are presented in Appendix S1.

RNA extraction, library preparation, and Illumina sequencing

New transcriptome data were generated at the Beijing Genomics Institute (BGI) in cooperation with the Australian National Insect Collection (TransANIC) and the 1KITE (https://1kite.org/) projects. For data from TransANIC, specimens were either preserved in RNAlater or shock-frozen in liquid nitrogen at the Australian National Insect Collection (ANIC). Depending on size, whole specimens or thoracic muscle tissue were sent to the Beijing Genomic Institute (BGI, China) for processing. Total RNA was extracted with TRIzol reagents (Invitrogen, Grand Island, NY, USA), and mRNA selected with Dynabeads mRNA Purification Kit (Invitrogen). The mRNA was sheared with RNA fragmentation reagents (Invitrogen), before cDNA libraries were built with Illumina TruSeq kits. The cDNA libraries of 11 distantly related species were equimolarly pooled for 150-bp paired-end sequencing on a HiSeq 4000. These pools included other flies, beetles and moths that are not part of this study.

For samples sequenced via the 1KITE (1000 Insect Transcriptome Evolution) project, RNA extraction, paired-end library building, and sequencing were performed as described in Misof et al. (2014), Peters et al. (2017) and Pauli et al. (2018). In short, samples were preserved in RNAlater and vouchers or partial vouchers were preserved where possible. RNA was extracted using TRIzol reagents (Invitrogen), concentrations were recorded with an Agilent 2100 Bioanalyzer (Santa Clara, CA, USA). RNA was separated with Dynabeads mRNA Purification Kit (Invitrogen) and sheared at 72 °C with the RNA fragmentation reagent (Ambion, Austin, Texas, USA). SuperScript™II Reverse Transcriptase (Invitrogen) and random N6 primer (IDT) were used to transcribe RNA fragments into first-strand cDNA, and RNase H (Invitrogen) and DNA polymerase I (New England BioLabs, Ipswich, MA, USA) were used to synthesize the second-strand cDNA. End-repair, adapter ligation, and size selection were then performed on the double-stranded cDNA. Fragment size and concentration of the cDNA was recorded again with an Agilent 2100 Bioanalyzer and an ABI StepOnePlus Real-Time PCR machine, then sequenced using Illumina HiSeq 2000 technology (San Diego, CA, USA).

DNA extraction, library preparation, and Illumina sequencing

Total DNA was extracted from thoracic muscle tissue (or whole bodies of specimens smaller than 3 mm long) using the Qiagen DNeasy Blood & Tissue Kit (Valencia, CA, USA), following the manufacturer's protocol for tissue, eluting with 50 μL of nuclease-free water. The remaining legs, head and wings were preserved in 95% ethanol as vouchers.

The DNA was fragmented by sonication with a Bioruptor Pico (Diagenode, Denville, NJ, USA), using 50 μL of the DNA extractions in 0.1 mL tubes (Cat. No. C30010015), targeting a modal fragment size of 350 base pairs. Libraries were built using the NEBNext Ultra II DNA Library Prep Kit (NEB #E7645L) with NEBNext Multiplex Oligos for Illumina (96 Index Primers) (NEB #E6609L), with two-sided size selection around a mode of 480 base pairs. Target enrichment through hybridization followed the Agilent SureSelectXT Target Enrichment Protocol (Version C2), with a modification at the end whereby the clean, target-enriched DNA suspension was incubated at 95 °C for 5 min, then immediately put onto the magnetic separator to take off the enriched supernatant. The probe set used was the published AHE Diptera Probe set (Young et al., 2016: Supplementary Information) and targeted 559 loci.

Enriched libraries were amplified using Phusion Hot Start II High-Fidelity DNA Polymerase (F549L) (Thermo Fisher Scientific, Waltham, MA, USA). PCR cycling consisted of an initial denaturing step at 98 °C for 30 s, followed by eight cycles of denaturing at 98 °C for 10 s, annealing at 60 °C for 30 s, elongation at 72 °C for 30 s, and a final elongation step at 72 °C for 10 min.

The 96 enriched and multiplexed libraries were sequenced using 150-bp paired-end reads on a single mid-output Nextseq 500 flow cell at the Biomolecular Resource Facility of the John Curtin School of Medical Research at the Australian National University (Canberra, ACT).

Sequence assembly and orthology prediction

For transANIC data, raw reads were pre-processed with SOAPnuke (v1.5.6; https://doi.org/10.1093/gigascience/gix120) and assembled with SOAPdenovo-Trans (v1.02; https://doi.org/10.1093/bioinformatics/btu077). Prior to orthology prediction, cross-sample contaminants due to index misspecification, for instance, were identified with an all-by-all reciprocal BLAST search with blastn, evalue 0.0000000001, and a custom decision-making script (Teasdale et al., in preparation) that flags potential sample cross-contaminants based on p-ident over 98.5% and relative coverage, then removed.

For transcriptome data provided by 1KITE, raw reads were quality checked, assembled with SOAPdenovo-Trans-31kmer v1.01 (Xie et al., 2014) and cleaned from potential contaminants as described by Peters et al. (2017).

Trinity v2.2 (Grabherr et al., 2011) was used for *de novo* assembly of raw transcriptome sequence data, and the Tadpole assembler (Bushnell et al., 2017) in BBmap v35.85 (Bushnell, 2014) was used for assembly of AHE data.

Orthograph v0.5.8 (Petersen et al., 2017) was used to infer the orthology of assembled sequence contigs. The ortholog reference set of (Pauli et al., 2018) 3145 orthologous sequences of single-copy protein-coding genes were generated from OrthoDB7 (Waterhouse et al., 2013) for five species: *Aedes aegypti* (Diptera, Culicidae), *Drosophila melanogaster* (Diptera, Drosophilidae), *Glossina morsitans* (Diptera, Glossinidae), *Bombyx mori* (Lepidoptera, Bombycidae) and *Danaus plexippus* (Lepidoptera, Nymphalidae).

For AHE data, 559 loci were targeted by the AHE Diptera probe set. However, 228 of these loci could not be detected from all five OrthoDB reference species. These sequences were identified using an all-by-all tblastx search (Teasdale et al., 2016).

After selection of potential orthologous sequences using Blast, custom bash scripts were used to map reads using BBmap v35.85 (Bushnell, 2014), call variants using GATK v4.1.1.0 (McKenna et al., 2010) and produce a final consensus sequence for each taxon

(Teasdale et al., 2016). All sequences with > 3% ambiguous sites or < 20% length of the alignment were removed from the final data set (Teasdale et al., 2016).

After orthology prediction, cross-sample contaminant check was performed with an all-by-all reciprocal BLAST search with blastn, evalue 0.0000000001, and a custom decision-making script (Teasdale et al., in preparation) that flags potential sample cross-contaminants based on p-ident over 98.5% and relative coverage. However, if potential contaminants were between species from the same tribe, we did not remove the sequences from either taxon. Contigs not matching Diptera or other insects were identified by an external contaminant check using NCBI BLAST, and were removed. The number of loci obtained for each terminal following the cleaning steps described above is presented in Appendix S1.

Data set construction

Nucleotide sequences were initially aligned using MAFFT v7.215 (Katoh and Standley, 2013) with the *L-INS-i* algorithm. Paralogs were then removed by a visual examination of the alignment, highlighting discrepancies to the majority rule consensus sequence using Aliview v1.22 (Larsson, 2014). The cleaned nucleotide sequences were re-aligned with MACSE v2 (Ranwez et al., 2018) to generate nucleotide alignments in reading frame and corresponding amino acid alignments. We used Aliscore v2.2 (Misof and Misof, 2009; Kück et al., 2010) to identify random similarity within alignments in the amino acid alignments, and Alicut to score amino acid alignments and remove problematic alignment sections from both amino acid and nucleotide alignments. Degen1 v1.4 (Regier et al., 2010; Zwick et al., 2012) was used to eliminate synonymous change in nucleotide alignments while retaining as much sequence information as possible. Only the fully degenerated nucleotide alignments were used for the subsequent steps, because the original nucleotide sequences have many fast-evolving sites that tend to introduce compositional biases, which interfere with phylogenetic signals in deep-level phylogenetics (Regier et al., 2010; Zwick et al., 2012).

Individual loci were concatenated using AMAS (Borowiec, 2016). MARE v0.1.2-rc (Meyer et al., 2011) was used with default weighting of information content (default = 3) to exclude genes with low information content. Two reduced data sets containing 368 loci (368AA and 368NT) were generated by selection of genes with more than 50% taxon coverage. A preliminary tree was generated in IQ-TREE with 1000 ultrafast bootstrap (UFBS) replicates (Minh et al., 2013; Hoang et al., 2018). Nine subgroups were defined based on the topology.

Two more inclusive data sets containing 916 loci (916AA and 916NT) were generated by including the 368 loci with more than 50% taxon coverage and adding 548 additional loci present in all nine subgroups, but with overall coverage less than 50%.

Several phylogenomic studies indicate that selecting loci with moderate evolutionary rates may provide better phylogenetic resolution for deep nodes (e.g. Klopfstein et al., 2017; Kuang et al., 2018). We therefore calculated uncorrected distances (p-distance) as a proxy measure of substitution rate variation across the 916AA data set constructed above (Teasdale et al., 2016; python script available at: https://github.com/lteasdale/p-distance_script.py). Twenty percent of the fastest and 20% of the slowest genes were excluded, and two data sets were constructed with 550 loci (131 672 amino acid residues: 550AA; 395 031 nucleotides: 550NT) (Appendix S2). Loci with more than 50% taxon coverage were reselected, and two 216 loci data sets were generated (33 056 amino acid residues: 216AA; 99 174 nucleotides: 216NT).

For both NT data sets (216NT and 550NT), synonymous changes were eliminated with Degen1 v1.4 (Regier et al., 2010; Zwick et al., 2012).

AliStat v1.3 (Misof et al., 2014) (available from: http://doi.org/10.4225/08/59309da8368e1) was used to inspect the data sets and generate heat maps of the distribution of missing data for data sets with different numbers of loci (Appendix S3). We used SymTest version 2.0.47 (https://github.com/ottmi/symtest) (Jermiin et al., 2004; Ababneh et al., 2006) to perform pairwise sequence comparisons with Bowker's matched-pairs test of symmetry (Bowker, 1948). Heat maps based on the inferred *P*-values, using default window and step sizes were generated for the four data sets (Appendix S3).

Phylogenetic analysis

Phylogenomic analyses were performed using the maximum likelihood (ML) method, and maximum parsimony (MP) method. ML analyses were carried out with IQ-TREE v1.7.0b12 (Nguyen et al., 2015) with a concatenation approach used for tree estimation on all four data sets.

A supermatrix approach was used for the two amino acid data sets (216AA and 550AA). We used a single LG model for amino acids (Le and Gascuel, 2008) including four matrices, each corresponding to one discrete gamma rate category (+Γ4 option; Le et al., 2012), and empirical amino acid frequencies estimated from the data (+F option). To compare the impact of partitioning, data partitioning and model testing were carried out in ModelFinder (Kalyaanamoorthy et al., 2017) implemented in IQ-TREE, using the rcluster algorithm (rcluster-percent = 10) to find the best partition scheme by possibly merging partitions (MF + MERGE), and the Bayesian information criterion (BIC). Two nucleotide data sets (216NT and 550NT) were partitioned by genes respectively, and a best-fitting substitution model for each partition was identified using ModelFinder implemented in IQ-TREE (MFP). For all ML analyses, to identify an optimal tree, 1000 separate ML searches were performed, and node support was calculated with 1000 non-parametric bootstrap (BS) replicates in IQ-TREE. After selection of the best topology (highest likelihood score), all BS values were mapped onto the best topology. BS values from different analyses were also mapped onto the selected topology to compare the supports from different analyses.

All AA and NT data sets were analysed using MP, carried out in TNT v.1.5 (Goloboff and Catalano, 2016) (new technology search, level 10, hits 20, gaps coded as missing data). Node support was assessed by jackknife (JK) resampling (1000 replicates at 36% deletion).

Exploration of node support

As concatenation analyses can return well-supported trees even when the level of gene incongruence is high (e.g. Jeffroy et al., 2006; Kumar et al., 2011), Quartet Sampling (Pease et al., 2018) was used to test the topology and provide more comprehensive and specific information on branch support.

Quartet sampling of internal node scores included a set of three scores: Quartet Concordance (QC: the more concordant the closer to 1, discordance increases as the score approaches −1), Quartet Differential (QD: the more equal the frequencies of discordant topologies the closer to 1, 0 indicates only one other discordant topology was found), and Quartet Informativeness (QI: 1 for all replicates informative, 0 for no replicates informative) (Pease et al., 2018).

Divergence time estimation

BEAST v2.4.5 (Bouckaert et al., 2014) was used through the CIPRES science gateway v3.3 (Miller et al., 2010) to estimate node ages. We included *Hilarimorpha* sp.1 (Hilarimorphidae) as the only outgroup, and removed AHE data for the seven species with both AHE and transcriptome sequences. In order to reduce the nucleotide alignment to a computationally feasible matrix, MARE was performed on the 216AA data set using the default settings and maintaining all the selected taxa. The overall information content was increased from 0.417 to 0.520, with 85 genes retained for 95 taxa.

The data set was partitioned by gene, with tree and clock models linked across partitions. GTR + G was selected as the substitution model. We applied an uncorrelated relaxed molecular clock model (Drummond et al., 2006) and a lognormal prior. The node dating approach with a birth-death tree prior (Kendall, 1948) was used, with the oldest fossil within a specific clade set as the minimum age via monophyly constraints (Heath et al., 2014). The Jurassic species *Palaeoplatypygus zaitzevi* was treated as stem Mythicomyiinae, while the two Cretaceous species of *Microburmyia* Grimaldi and Cumming were included within crown Bombyliidae exclusive of Heterotropinae and Mythicomyiinae. A total of 11 nodes were calibrated using fossils, more detailed information is provided in Appendix S6. Ages of fossils were obtained from the Fossilworks database (http://www.fossilworks.org/).

Four independent analyses with 200 million generations each were performed in BEAST. We evaluated the convergence and mixing of the MCMC chains in Tracer v1.6, to ensure the multiple runs converged on the same distribution and ascertained that effective sample sizes (ESS) exceeded 150. Generated tree files were combined and resampled with a frequency of 100 000 in LogCombiner (BEAST package) and a burn-in of 30%. The subsampled trees were used to summarize the maximum clade credibility tree by TreeAnnotator (BEAST package), with median heights as node heights.

Ancestral states reconstruction

We used Bayesian Binary MCMC (BBM) method implemented in RASP (Yu et al., 2015) to reconstruct the possible ancestral states of three characters in Bombyliidae on the phylogenetic tree generated by 550 AA data set. We used characters from Yeates (1994) as follows: For the sand chamber, taxa were coded as sand chamber absent (state 0) or present (state 1). For postcranium, taxa were coded as flat or slightly tumid (state 0), moderately produced (state 1), or cup-shaped (state 2). For eye hind margin, taxa were coded as complete (state 0) or with an indentation (state 1). The maximum number of areas was set to 1, and fixed JC + G (Jukes-Cantor + Gamma) was used in the analyses. The state was sampled every 100 generations with the total MCMC chains run simultaneously for 5 000 000 generations.

Diversification analyses

The time tree was trimmed to one individual per genus for the diversification analyses. Lineage through time (LTT) plots were used from the R package ape v3.4 (Paradis, 2012) to visualize the temporal pattern of lineage diversification. The mean age of the subfamily and higher-level divergence events were mapped as a histogram.

BAMM v2.5.0 (Rabosky et al., 2014) was used to analyze diversification rates through time and estimate rate shifts on the tree. The function "setBAMMpriors" in the R package BAMMtools v2.1.6 (Rabosky et al., 2014) was used to generate the priors, and the relevant sampling fractions of different clades were incorporated with the "sampleProbsFilename" argument (Appendix S7). The total number of species of each clade was based on the catalogue of Bombyliidae (Evenhuis and Greathead, 1999) and the catalogue of Mythicomyiidae (Evenhuis, 2002). We performed four independent Markov chains for 100 000 000 generations, and sampled every 10 000 generations. The coda v0.19-3 package (Plummer et al., 2013) was used to plot the log-likelihood trace and assess the MCMC convergence. All parameters had effective sample sizes over 8000. The function "plotRateThroughTime" was used to plot the diversification rate. The distinct shift configurations and their posterior probabilities were estimated, and the 95% probability shift scheme was summarized with the function "credibleShiftSet".

Data availability

Newly generated AHE and transcriptome data included herein is available from the NCBI SRA repository (Bioprojects PRJNA596828 for AHE data, PRJNA597068 for TransANIC transcriptomes). Accession numbers for the published data used here are provided in Appendix S1.

Results

Data sets construction and information content

Four data sets were generated for phylogenetic estimation. Initially a reduced nucleotide matrix including 368 loci with more than 50% coverage among taxa was selected to generate the preliminary tree. The preliminary tree was generated by IQ-TREE with 1000 ultrafast bootstrap (UFBS) replicates (Minh et al., 2013; Hoang et al., 2018) to check the topology, and nine subgroups appearing on the tree were subsequently defined for selection of further loci. We considered all the outgroup taxa as one subgroup, used strongly supported monophyletic families as natural subgroups, and aimed to have at least one transcriptome per subgroup whenever possible. Clades having only a few terminals were included in their most closely related larger subgroup. To expand this data set, we added loci that were present in at least one member of all nine subgroups, 916 loci were eventually selected.

We calculated uncorrected distances (p-distance) as a proxy measure of substitution rate variation across the 916AA data sets constructed above (ranging from 0.0036 to 0.4797, average of 0.1704; Appendix S2). Twenty percent of the fastest and 20% of the slowest genes were excluded, and the final two data sets were constructed with 550 loci (131 672 amino acid residues: 550AA; 395 031 nucleotides: 550NT). Loci with more than 50% taxa coverage were selected, and two 216 loci data sets were generated (33 056 amino acid residues: 216AA; 99 174 nucleotides: 216NT). In the final data sets, the proportion of total missing data in the 216 loci data sets was 39.55%, and in the 550 loci data sets was 72.60%.

Partitioning analyses on 216AA_MF + MERGE yielded 13 partitions, details of the partitioning and model selection results of 216AA_MF + MERGE, 550NT, and 216NT data sets are listed in Appendix S8.

Phylogenetic trees

Maximum likelihood (ML) trees generated from the four data sets yielded similar topologies, with 58.2%

of the nodes receiving significant support (BS ≥ 90) across all approaches, while 18.2% of the nodes received high support (95> BS ≥ 70) (Appendix S4). The topology resulting from analyses of the 550AA data set was selected for Figs 2–4. The topologies generated by the two 550 loci data sets (550AA and 550NT) were consistent except for the position of two ingroup terminal taxa (*Paracosmus* sp.1 and *Pantostomus* sp.1). The 550AA data set has higher support values on a larger number of nodes than other data sets, including higher QC values for most of the nodes.

The topology generated by 216AA_MF + MERGE was the same as 216AA_LG except for the position of two pairs of ingroup terminal taxa (*Pieza minuta* (Greene) and *Mythicomyia* sp.1, and *Paracosmus* sp.1 and *Pantostomus* sp.1, respectively). The 216AA_MF + MERGE topology did not show increased BS support, which may be because short genes were not informative enough to satisfy the model search using the traditional partitioning method. BS values from all other analyses were mapped on the selected topology (Fig. 2).

The monophyly of Bombyliidae was strongly supported by all five ML trees (ML BS = 100(550AA)/98 (550NT)/100(216AA_LG)/98(216NT)/100 (216AA_MF + MERGE); Quartet Sampling scores (QS)) = 0.56(Quartet Concordance, QC)/0(Quartet Differential, QD)/0.93(Quartet Informativeness, QI)), and 11 out of 17 recognized subfamilies were recovered as monophyletic (Fig. 2, Table 1). Three subfamilies (Heterotropinae, Oniromyiinae, and Crocidiinae) were only represented by a single taxon and, therefore, their monophyly was not tested here. The former two subfamilies each contain a single genus, however, Crocidiinae contains two tribes (Crocidiini and Desmatomyiini) with six and two genera, respectively (Lamas and Couri, 2004). Two tribes (Phthiriini and Eclimini) that had previously been included in established subfamilies (Usiinae and Bombyliinae, respectively), did not form monophyletic groups with the remaining members of their subfamilies. Therefore, we reinstated these two tribes as subfamilies: Phthiriinae stat. rev. and Ecliminae stat. rev. The *incertae sedis* genus *Sericosoma* Macquart was highly supported to be the sister group to *Oniromyia* Bezzi, representing the monogeneric subfamily Oniromyiinae, in all the analyses (BS = 97/95/89/94/88; QS = 0.55/0.57/0.9), so we move *Sericosoma* to the subfamily Oniromyiinae. Cythereinae and Bombyliinae are polyphyletic in our analyses, both forming three separate lineages.

Sections I and III. Sections I and III of the tree (see Fig. 2) were generally highly supported in all five tree reconstructions. Section I included the early branching lineages of Heterotropinae, Mythicomyiinae, and ((Usiinae + Crocidiinae) + Ecliminae stat. rev.), paraphyletic with respect to the remaining Bombyliidae, and section III was a monophyletic lineage of Lordotinae + (Bombyliinae + ((*Neosardus* + Antoniinae) + Anthracinae)).

Section II. The remaining parts of the tree have been marked as section II on Fig. 2. Within this section, four clades were recovered in all analyses: (i) a strongly supported monophyletic Toxophorinae (BS = 85/90/87/91/88, QS = 0.97/0.8/0.99) (labelled clade A); (ii) a well-supported clade of (*Enica* Macquart + Mariobezziinae) + Phthiriinae stat. rev. (BS = 84/90/84/92/99; QS = 0.42/0.5/0.69) (labelled clade B); (iii) a clade of Oniromyiinae + (*Marmasoma* White + (*Zentamyia* Li and Yeates + *Robertsmyia* Li and Yeates)) (BS = 33/68/63/75/66; QS = 0.01/0.96/ 0.84) (labelled clade C); (iv) a clade of (*Pantarbes* Osten Sacken + *Conophorus* Meigen) + (Tomomyzinae + Lomatiinae) (BS = 60/63/63/65/63; QS = 0.01/ 0.84/0.78) (labelled clade D). Although the monophyly of the four clades was supported by all of the analyses, the relationships between them and their relationship with section III were not resolved with strong support (Table 1).

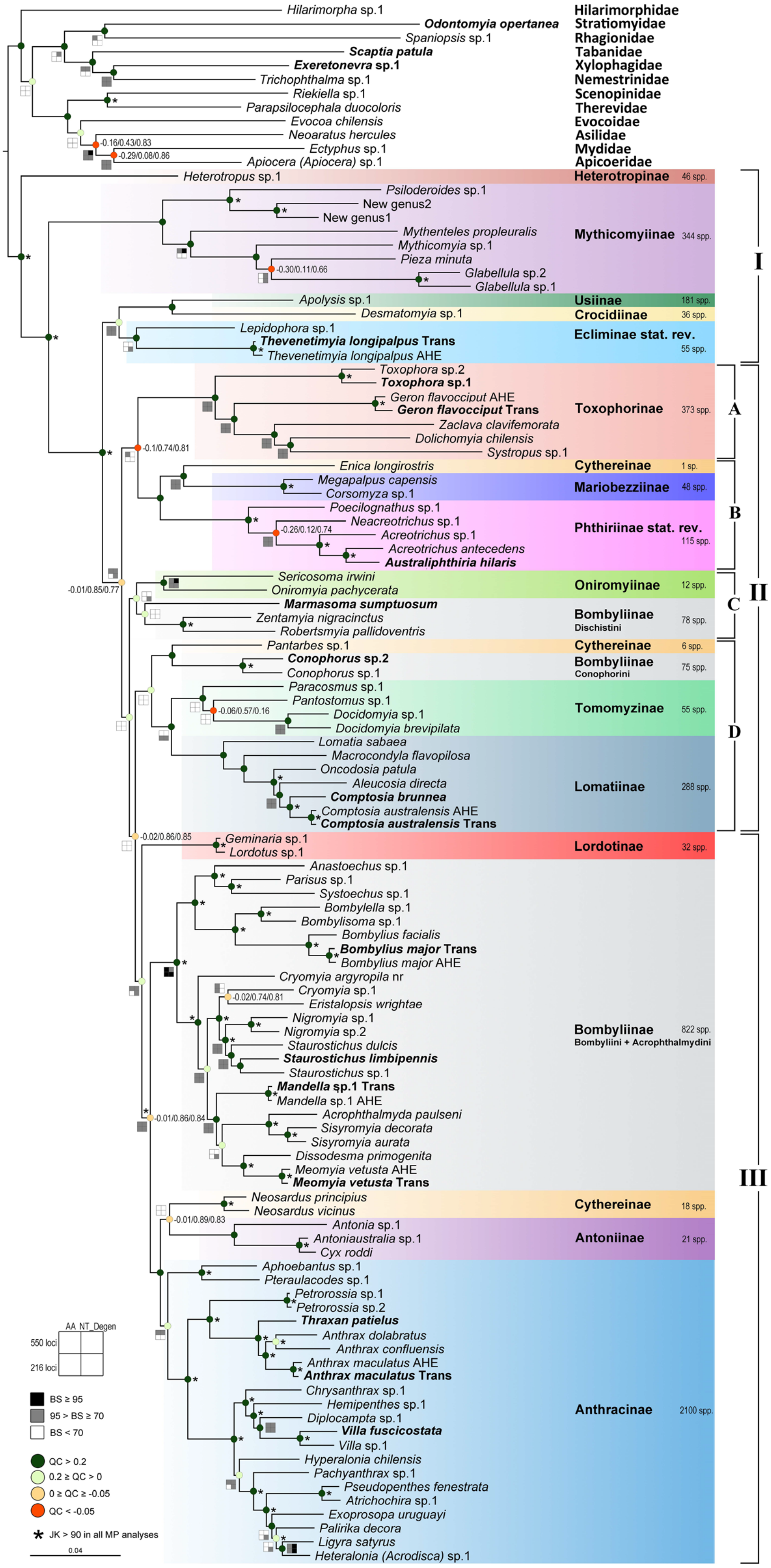
Both 550 loci data sets (550AA and 550NT) recovered the relationships within Sections II and III as (clade A + clade B) + (clade C + (clade D + III)) (Fig. 2), however, the two 216 loci data sets (216AA and 216NT) generated contradictory relationships with week support (Appendix S3).

Ancestral character state reconstruction

Ancestral character state reconstruction using Bayesian Binary MCMC (BBM) analyses traced the evolution of the sand chamber, the shape of the postcranium and the shape of hind eye margin. Ancestral Bombyliidae were estimated to lack the sand

---

Fig. 2. Phylogenetic tree of Bombyliidae. Topology generated based on the 550AA data set with LG + Γ4 + F model. Squares around the nodes are ML bootstrap (BS) values from different analyses, upper left: 550AA, upper right: 550NT_Degen, lower left: 216AA, lower right 216NT_Degen; nodes without squares indicate BS = 100 in all four analyses; black indicates BS ≥ 95, grey indicates 95> BS ≥ 70, and white indicates BS < 70. Circles on nodes are Quartet Concordance (QC) scores for internal branches: dark green indicates QC > 0.2, light green indicates 0.2 ≥ QC > 0, light orange indicates 0 ≥ QC ≥ – 0.05, and dark orange indicates QC < – 0.05. Quartet Sampling Internal Node Scores presented when QC ≤ 0, three numbers are: Quartet Concordance (QC), Quartet Differential (QD), and Quartet Informativeness (QI). Internal nodes supported by MP analysis of four data sets, jackknife values (> 90) in all four data sets are marked with an asterisk (*). Taxa name in bold indicates transcriptome data.


Hilarimorpha sp.1
Odontomyia opertanea
Spaniopsis sp.1
Scaptia patula
Exeretonevra sp.1
Trichophthalma sp.1
Riekiella sp.1
Parapsilocephala duocoloris
Evocoa chilensis
Neoaratus hercules
Ectyphus sp.1
Apiocera (Apiocera) sp.1
-0.16/0.43/0.83
-0.29/0.08/0.86
Hilarimorphidae
Stratiomyidae
Rhagionidae
Tabanidae
Xylophagidae
Nemestrinidae
Scenopinidae
Therevidae
Evocoidae
Asilidae
Mydidae
Apicoeridae
Heterotropus sp.1
Heterotropinae 46 spp.
Psiloderoides sp.1
New genus2
New genus1
Mythenteles propleuralis
Mythicomyia sp.1
Pieza minuta
Glabellula sp.2
Glabellula sp.1
-0.30/0.11/0.66
Mythicomyiinae 344 spp.
Apolysis sp.1
Usiinae 181 spp.
Desmatomyia sp.1
Crocidiinae 36 spp.
Lepidophora sp.1
Thevenetimyia longipalpus Trans
Thevenetimyia longipalpus AHE
Ecliminae stat. rev. 55 spp.
I
Toxophora sp.2
Toxophora sp.1
Geron flavocciput AHE
Geron flavocciput Trans
Zaclava clavifemorata
Dolichomyia chilensis
Systropus sp.1
-0.1/0.74/0.81
Toxophorinae 373 spp.
A
Enica longirostris
Cythereinae 1 sp.
Megapalpus capensis
Corsomyza sp.1
Mariobezziinae 48 spp.
Poecilognathus sp.1
Neacreotrichus sp.1
Acreotrichus sp.1
Acreotrichus antecedens
Australiphthiria hilaris
-0.26/0.12/0.74
Phthiriinae stat. rev. 115 spp.
B
-0.01/0.85/0.77
Sericosoma irwini
Oniromyia pachycerata
Oniromyiinae 12 spp.
Marmasoma sumptuosum
Zentamyia nigracinctus
Robertsmyia pallidoventris
Bombyliinae Dischistini 78 spp.
C
II
Pantarbes sp.1
Cythereinae 6 spp.
Conophorus sp.2
Conophorus sp.1
Bombyliinae Conophorini 75 spp.
Paracosmus sp.1
Pantostomus sp.1
Docidomyia sp.1
Docidomyia brevipilata
-0.06/0.57/0.16
Tomomyzinae 55 spp.
Lomatia sabaea
Macrocondyla flavopilosa
Oncodosia patula
Aleucosia directa
Comptosia brunnea
Comptosia australensis AHE
Comptosia australensis Trans
Lomatiinae 288 spp.
D
-0.02/0.86/0.85
Geminaria sp.1
Lordotus sp.1
Lordotinae 32 spp.
Anastoechus sp.1
Parisus sp.1
Systoechus sp.1
Bombylella sp.1
Bombylisoma sp.1
Bombylius facialis
Bombylius major Trans
Bombylius major AHE
Cryomyia argyropila nr
Cryomyia sp.1
Eristalopsis wrightae
-0.02/0.74/0.81
Nigromyia sp.1
Nigromyia sp.2
Staurostichus dulcis
Staurostichus limbipennis
Staurostichus sp.1
Mandella sp.1 Trans
Mandella sp.1 AHE
Acrophthalmyda paulseni
Sisyromyia decorata
Sisyromyia aurata
Dissodesma primogenita
Meomyia vetusta AHE
Meomyia vetusta Trans
-0.01/0.86/0.84
Bombyliinae 822 spp.
Bombyliini + Acrophthalmydini
Neosardus principius
Neosardus vicinus
Cythereinae 18 spp.
-0.01/0.89/0.83
Antonia sp.1
Antoniaustralia sp.1
Cyx roddi
Antoniinae 21 spp.
Aphoebantus sp.1
Pteraulacodes sp.1
Petrorossia sp.1
Petrorossia sp.2
Thraxan patielus
Anthrax dolabratus
Anthrax confluensis
Anthrax maculatus AHE
Anthrax maculatus Trans
Chrysanthrax sp.1
Hemipenthes sp.1
Diplocampta sp.1
Villa fuscicostata
Villa sp.1
Hyperalonia chilensis
Pachyanthrax sp.1
Pseudopenthes fenestrata
Atrichochira sp.1
Exoprosopa uruguayi
Palirika decora
Ligyra satyrus
Heteralonia (Acrodisca) sp.1
Anthracinae 2100 spp.
III
AA
NT_Degen
550 loci
216 loci
■ BS ≥ 95
▨ 95 > BS ≥ 70
□ BS < 70
● QC > 0.2
○ 0.2 ≥ QC > 0
○ 0 ≥ QC ≥ -0.05
● QC < -0.05
* JK > 90 in all MP analyses
0.04

chamber, and have the postcranium flat or slightly tumid, and a complete hind margin of the eye as found in their nearest relatives (Fig. 3).

Analysis of the sand chamber suggests a complicated evolutionary history. There are two possible optimizations of the evolution of the sand chamber in the early branching lineages of Bombyliidae: (i) a chamber originated after the split of Mythicomyiinae with a posterior probability (pp) of 0.5752, and was lost in Usiinae and the common ancestor of clade A + clade B (0.9050), but was re-gained in the clade of *Enica* + Mariobezziinae within clade B (pp = 0.9981); or (ii) the chamber evolved once after the split of Mythicomyiinae (pp = 0.5752), and was lost separately in Usiinae, Toxophorinae, and Phthiriinae stat. rev. In both scenarios, the sand chamber was present in the common ancestor of clade C + (clade D + section III), and was lost three times independently in *Zentamyia*, Lordotinae, and *Eristalopsis* Evenhuis (Fig. 3). If we consider the cost of evolving the sand chamber as much higher than its loss, then the sand chamber is more likely to have evolved once after the divergence of Mythicomyiinae, and was then lost six times in the present tree, in Usiinae, Toxophorinae, Phthiriinae stat. rev., *Zentamyia*, Lordotinae, and *Eristalopsis* (Fig. 3). Therefore, both “sand chamber subfamilies” (“Psammophoridae” *sensu* Mühlenberg, 1971) and “sand chamber-less subfamilies” are not monophyletic.

The cup-shaped postcranium (state 2) evolves twice separately in the clade of Tomomyzinae + Lomatiinae (pp = 0.9995) and the clade of (*Neosardus* + Antoniinae) + Anthracinae (pp = 0.9813) (Fig. 3). However, a moderately produced postcranium (state 1) evolves four times in our analysis, appearing independently in *Enica* + Mariobezziinae (pp = 0.9978), Oniromyiinae (pp = 0.9978), *Pantarbes*, and *Neosardus* (pp = 0.9995) (Fig. 3). Our results indicate that the “Homeophthalmae” (postcranium flat or slightly tumid), found to be paraphyletic by Yeates (1994) is polyphyletic, and “Tomophthalmae” (postcranium produced), found by Yeates (1994) to be monophyletic, is also polyphyletic.

Analysis of the indentation that occurs on the hind eye margin of some bee flies suggests that the indentation evolves separately in Lomatiinae (pp = 0.9995) and in the clade of (*Neosardus* + Antoniinae) + Anthracinae (Fig. 3). There are two possible, equally parsimonious, reconstructions for the indentation on the hind eye margin in the latter clade: (i) it evolved once in a common ancestor (pp = 0.4586), and was lost in *Neosardus* (pp = 0.9999); or (ii) the common ancestor possessed a complete hind eye margin (pp = 0.5414), and the indentation evolved separately in Antoniinae (pp = 0.9999) and Anthracinae (pp = 1). If we consider the cost of evolving the indentation on the hind eye margin is higher than its loss, then the indentation is more likely to have evolved twice in Lomatiinae and the clade of (*Neosardus* + Antoniinae) + Anthracinae, and was lost once in *Neosardus*.

Divergence time estimation

A reduced data set of 95 taxa and 85 genes was used to estimate the divergence times of bee fly subfamilies (Fig. 4). Our results indicated the origin of Bombyliidae in the Jurassic at 174 Ma [95% highest posterior density (HPD) = 194–165 Ma], with Mythicomyiinae diverging at approximately 166 Ma (95% HPD = 165–168 Ma). The crown clade of remaining Bombyliidae originated approximately 104 Ma (95% HPD = 110–100 Ma). The crown clades of the three most species-rich subfamilies, Anthracinae, Bombyliinae and Lomatiinae, all originated approximately 69–66 Ma, around the Cretaceous–Paleogene (K-Pg) boundary.

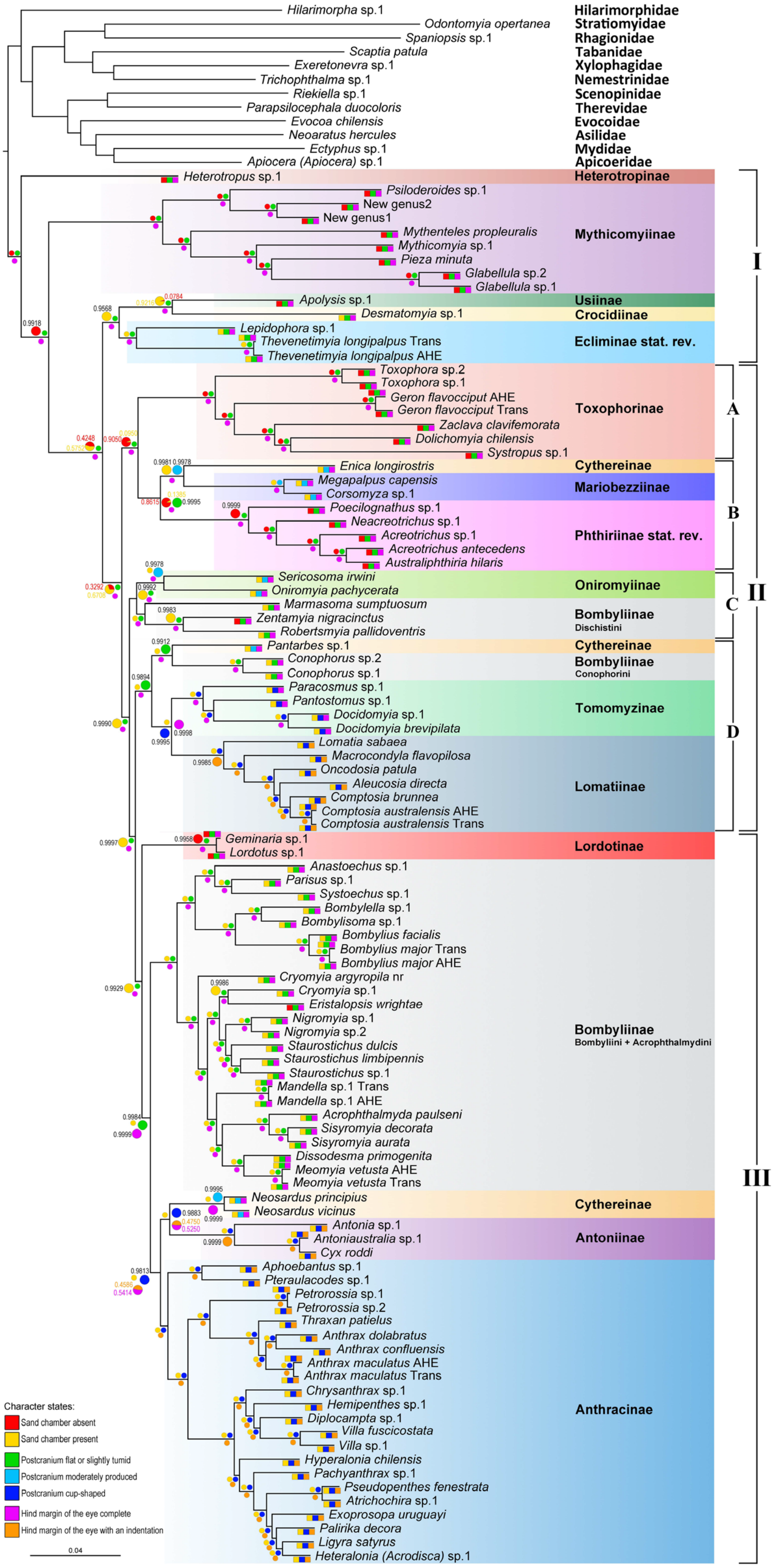
The 95% HPD for each node is given in Appendix S6, and the numbered nodes of the topology are presented in Appendix S5. All nodes of the tree were very well supported with posterior probabilities (PP) equal to 1 (Appendix S6).

Diversification dynamics

Based on the samples included in the BEAST analyses, we observed an increase in diversification during the late Cretaceous (100–66 Ma), a time when most of the subfamilies originated and the majority of higher-level diversification occurred in the Bombyliidae (Fig. 5A).

The BAMM analyses with non-random species sampling fractions incorporated for different clades indicates a dramatic increase of species diversification rate starting in the late Cretaceous (Fig. 5B). Diversification rate-shift analyses proposed seven configurations with highest possibilities, suggesting the diversification rate increased in the clade of Bombyliidae + ((*Nerosardus* + Antoniinae) + Anthracinae), approximately 81 Ma (95% HPD = 92–71 Ma) (Fig. 6).

Fig. 3. Ancestral character reconstruction of Bombyliidae. Topology generated by the 550AA data set with LG + Γ4+F model. Squares around distal branches are three pairs of character states: red indicates sand chamber absent, yellow indicates sand chamber present; green indicates postcranium flat or slightly tumid, blue indicates postcranium produced. Circles around the nodes are pie charts of ancestral character states obtained by Bayesian Binary MCMC (BBM) analysis. Bigger circles mark the nodes with possible ancestral states lower than 1, and with the posterior probability (pp) around them. Three characters (sand chamber, postcranium, and eye margin) around distal branches are listed from left to right; around nodes are listed upper left, upper right, and lower.


Hilarimorpha sp.1
Odontomyia opertanea
Spaniopsis sp.1
Scaptia patula
Exeretonevra sp.1
Trichophthalma sp.1
Riekiella sp.1
Parapsilocephala duocoloris
Evocoa chilensis
Neoaratus hercules
Ectyphus sp.1
Apiocera (Apiocera) sp.1
Hilarimorphidae
Stratiomyidae
Rhagionidae
Tabanidae
Xylophagidae
Nemestrinidae
Scenopinidae
Therevidae
Evocoidae
Asilidae
Mydidae
Apicoeridae
Heterotropus sp.1
Heterotropinae
Psiloderoides sp.1
New genus2
New genus1
Mythenteles propleuralis
Mythicomyia sp.1
Pieza minuta
Glabellula sp.2
Glabellula sp.1
Mythicomyiinae
I
Apolysis sp.1
Usiinae
Desmatomyia sp.1
Crocidiinae
Lepidophora sp.1
Thevenetimyia longipalpus Trans
Thevenetimyia longipalpus AHE
Ecliminae stat. rev.
Toxophora sp.2
Toxophora sp.1
Geron flavocciput AHE
Geron flavocciput Trans
Zaclava clavifemorata
Dolichomyia chilensis
Systropus sp.1
Toxophorinae
A
Enica longirostris
Cythereinae
Megapalpus capensis
Corsomyza sp.1
Mariobezziinae
Poecilognathus sp.1
Neacreotrichus sp.1
Acreotrichus sp.1
Acreotrichus antecedens
Australiphthiria hilaris
Phthiriinae stat. rev.
B
Sericosoma irwini
Oniromyia pachycerata
Oniromyiinae
Marmasoma sumptuosum
Zentamyia nigracinctus
Robertsmyia pallidoventris
Bombyliinae
Dischistini
C
II
Pantarbes sp.1
Cythereinae
Conophorus sp.2
Conophorus sp.1
Bombyliinae
Conophorini
Paracosmus sp.1
Pantostomus sp.1
Docidomyia sp.1
Docidomyia brevipilata
Tomomyzinae
Lomatia sabaea
Macrocondyla flavopilosa
Oncodosia patula
Aleucosia directa
Comptosia brunnea
Comptosia australensis AHE
Comptosia australensis Trans
Lomatiinae
D
Geminaria sp.1
Lordotus sp.1
Lordotinae
Anastoechus sp.1
Parisus sp.1
Systoechus sp.1
Bombylella sp.1
Bombylisoma sp.1
Bombylius facialis
Bombylius major Trans
Bombylius major AHE
Cryomyia argyropila nr
Cryomyia sp.1
Eristalopsis wrightae
Nigromyia sp.1
Nigromyia sp.2
Staurostichus dulcis
Staurostichus limbipennis
Staurostichus sp.1
Mandella sp.1 Trans
Mandella sp.1 AHE
Acrophthalmyda paulseni
Sisyromyia decorata
Sisyromyia aurata
Dissodesma primogenita
Meomyia vetusta AHE
Meomyia vetusta Trans
Bombyliinae
Bombyliini + Acrophthalmydini
Neosardus principius
Neosardus vicinus
Cythereinae
Antonia sp.1
Antoniaustralia sp.1
Cyx roddi
Antoniinae
Aphoebantus sp.1
Pteraulacodes sp.1
Petrorossia sp.1
Petrorossia sp.2
Thraxan patielus
Anthrax dolabratus
Anthrax confluensis
Anthrax maculatus AHE
Anthrax maculatus Trans
Chrysanthrax sp.1
Hemipenthes sp.1
Diplocampta sp.1
Villa fuscicostata
Villa sp.1
Hyperalonia chilensis
Pachyanthrax sp.1
Pseudopenthes fenestrata
Atrichochira sp.1
Exoprosopa uruguayi
Palirika decora
Ligyra satyrus
Heteralonia (Acrodisca) sp.1
Anthracinae
III
Character states:
🟥 Sand chamber absent
🟨 Sand chamber present
🟩 Postcranium flat or slightly tumid
🟦 Postcranium moderately produced
🟦 Postcranium cup-shaped
🟪 Hind margin of the eye complete
🟧 Hind margin of the eye with an indentation
0.04

Discussion

Impacts on the classification of Bombyliidae

Bee fly phylogeny has long been challenging to elucidate due to their ancient origins, their species richness and their extreme morphological disparity. Our phylogenomic data has largely resolved the backbone of the bee fly tree of life, as well as its timing and diversification. We consider the phylogenetic implications and re-evaluate the classification and status of several taxonomic groups as follows.

Ecliminae stat. rev.. The subfamily Ecliminae was first established by Hall (1969) and was moved to Bombyliinae (as tribe Eclimini) by Hull (1973). However, Bowden (1985), Greathead (1988), and Evenhuis (1991) all considered Ecliminae as a separate subfamily, until Yeates (1994) supported Hull (1973) by including Eclimini as a tribe of Bombyliinae based on phylogenetic analysis of morphological characters. While eclimine genera *Thevenetimyia* Bigot and *Marmasoma* were sampled in an earlier molecular phylogeny (Trautwein et al., 2011), they did not form a monophyletic clade. In the recent morphological phylogeny of Bombyliinae, Li and Yeates (2019) moved *Marmasoma* from Eclimini to its own tribe Marmasomini.

Our phylogenomic analyses recovered a polyphyletic Bombyliinae, with eclimine genera *Lepidophora* Westwood and *Thevenetimyia* forming a monophyletic group distant from the other four lineages containing bombyliine genera. The eclimine lineage was strongly supported as sister group to Usiinae + Crocidiinae in all our analyses. Therefore, we reinstate the Ecliminae (excluding Marmasomini) to its former taxonomic rank as a subfamily.

Phthiriinae stat. rev.. Usiinae and Phthiriinae are two subfamilies of small flies, many species of which are similar in appearance (Hull, 1973; Evenhuis, 1990). Yeates (1994) united them into one subfamily, Usiinae, based on the results of his morphological phylogeny. However, that classification was not followed by the world catalogue (Evenhuis and Greathead, 1999).

A weak sister-group relationship between Usiini and Phthiriini was proposed by Trautwein et al. (2011), although only one *Phthiria* Meigen was included to represent Phthiriini. Moreover, *Phthiria* was identified as a rogue taxon and was removed from the final tree. The monophyly of Usiini + Phthiriini was further supported by morphological evidence by Li and Yeates (2017) with the description of a new tribe, Phthiraxini, for the Usiinae.

The sister-group relationship between Usiini and Phthiriini was not supported in our study. The monophyly of Phthiriini was strongly supported by all the analyses, forming a strongly supported monophyletic clade with *Enica* + Mariobezziinae. Therefore, we reinstate the Phthiriinae to its former taxonomic rank as a subfamily.

A polyphyletic Bombyliinae. Hull (1973) expanded the concept of the subfamily Bombyliinae, including such divergent subfamilies as Cythereinae, Crocidiinae, Ecliminae, Corsomyzinae and Mariobezziinae of previous authors as tribes. Subsequent workers struggled to reconcile this broad concept with character data (Bowden, 1975, 1985; Evenhuis, 1991; Yeates, 1994). Even though half of Hull's accreted groups had already been excluded, 20 years later Yeates (1994) still described Bombyliinae as "a heterogeneous assemblage". Li and Yeates (2019) reconstructed the phylogeny of the Bombyliinae globally using morphological data, and divided the subfamily into nine tribes based on morphological characters. The subfamily Bombyliinae appears as five separate lineages on our phylogenomic trees. One we have discussed above, Ecliminae stat. rev.

The remaining Bombyliinae form three separate clades, however, these are largely consistent with the tribal classification: *Marmasoma* (Marmasomini) + (*Zentamyia* + *Robertsmyia*) (Dischistini) forming one clade, *Conophorus* (Conophorini) another, and the remaining Bombyliinae (Bombyliini + Acrophthalmydini) a third monophyletic lineage. While *Zentamyia* was placed within Bombyliini in the world phylogeny (Li and Yeates, 2019), the close relationship between *Zentamyia* and some Dischistini genera was also detected in the Australian generic phylogeny (Li and Yeates, 2018). Acrophthalmydini is nested within the Australian clade of Bombyliini in our analyses, rather than sister to it, as in Li and Yeates (2019). Divergence time results also suggest that the divergence between the South American genus *Acrophthalmyda* Bigot and the Australian genera may have occurred in response to the separation of southern South America, Antarctica and Australia (52–35 Ma).

Fig. 4. Estimated divergence times among lineages of Bombyliidae under the fossilized birth-death process, in BEAST 2. Topology generated by 85AA genes retained for 95 taxa, and 11 fossil calibrations. Scale is in Ma. Bars depict the 95% highest posterior probability density of each estimate. Pale blue node diamonds on chronogram represent minimum age constraints for those lineages. Mean ages and ranges are provided in Supplementary 6 and refer to nodes indicated in Appendix S5. Red line shows global temperature after Scotese (2016).

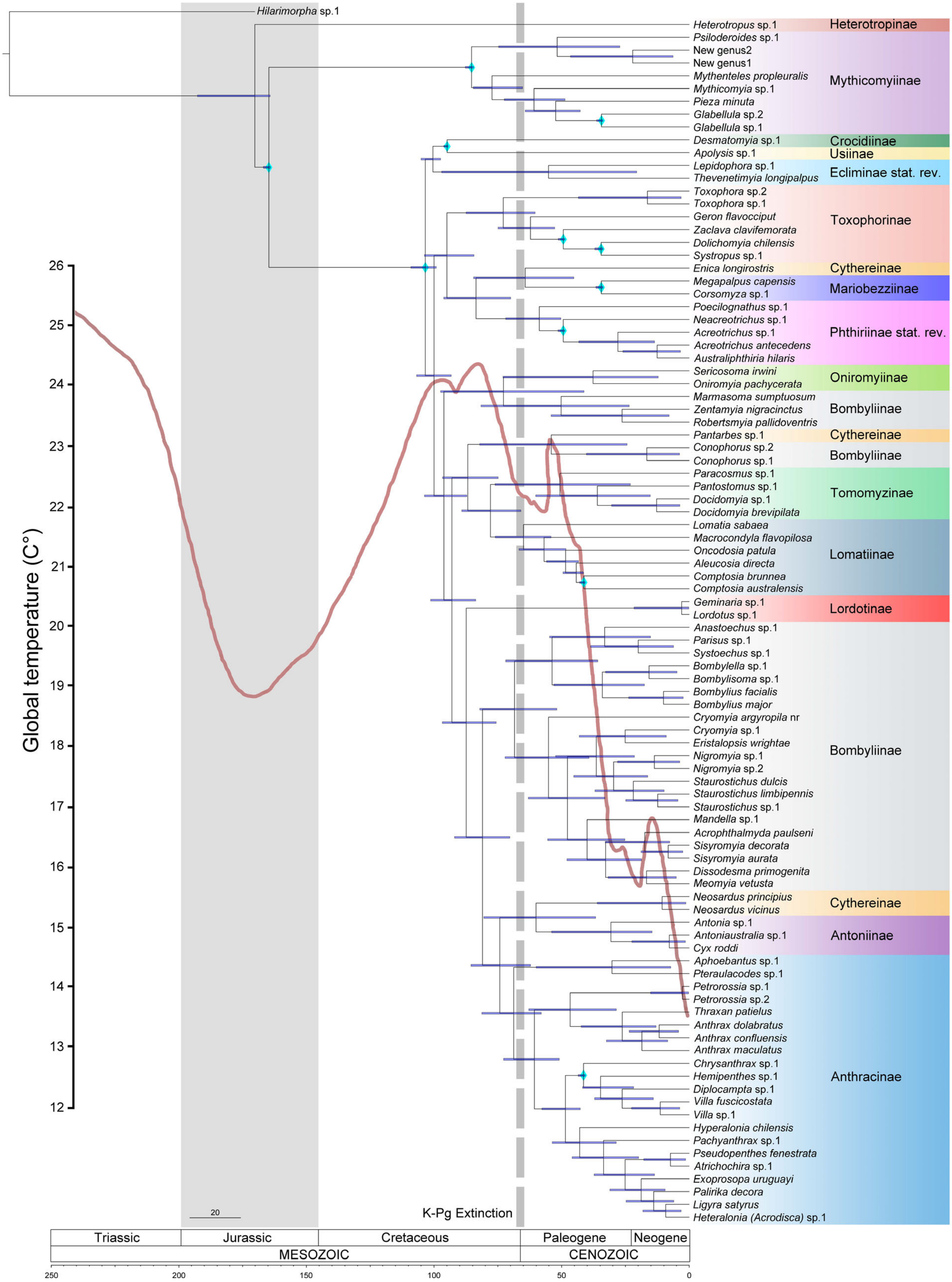

Hilarimorpha sp.1
Heterotropus sp.1
Heterotropinae
Psiloderoides sp.1
New genus2
New genus1
Mythenteles propleuralis
Mythicomyia sp.1
Pieza minuta
Glabellula sp.2
Glabellula sp.1
Mythicomyiinae
Desmatomyia sp.1
Crocidiinae
Apolysis sp.1
Usiinae
Lepidophora sp.1
Thevenetimyia longipalpus
Ecliminae stat. rev.
Toxophora sp.2
Toxophora sp.1
Geron flavocciput
Zaclava clavifemorata
Dolichomyia chilensis
Systropus sp.1
Toxophorinae
Enica longirostris
Cythereinae
Megapalpus capensis
Corsomyza sp.1
Mariobezziinae
Poecilognathus sp.1
Neacreotrichus sp.1
Acreotrichus sp.1
Acreotrichus antecedens
Australiphthiria hilaris
Phthiriinae stat. rev.
Sericosoma irwini
Oniromyia pachycerata
Oniromyiinae
Marmasoma sumptuosum
Zentamyia nigracinctus
Robertsmyia pallidoventris
Bombyliinae
Pantarbes sp.1
Cythereinae
Conophorus sp.2
Conophorus sp.1
Bombyliinae
Paracosmus sp.1
Pantostomus sp.1
Docidomyia sp.1
Docidomyia brevipilata
Tomomyzinae
Lomatia sabaea
Macrocondyla flavopilosa
Oncodosia patula
Aleucosia directa
Comptosia brunnea
Comptosia australensis
Lomatiinae
Geminaria sp.1
Lordotus sp.1
Lordotinae
Anastoechus sp.1
Parisus sp.1
Systoechus sp.1
Bombylella sp.1
Bombylisoma sp.1
Bombylius facialis
Bombylius major
Cryomyia argyropila nr
Cryomyia sp.1
Eristalopsis wrightae
Nigromyia sp.1
Nigromyia sp.2
Staurostichus dulcis
Staurostichus limbipennis
Staurostichus sp.1
Mandella sp.1
Acrophthalmyda paulseni
Sisyromyia decorata
Sisyromyia aurata
Dissodesma primogenita
Meomyia vetusta
Bombyliinae
Neosardus principius
Neosardus vicinus
Cythereinae
Antonia sp.1
Antoniaustralia sp.1
Cyx roddi
Antoniinae
Aphoebantus sp.1
Pteraulacodes sp.1
Petrorossia sp.1
Petrorossia sp.2
Thraxan patielus
Anthrax dolabratus
Anthrax confluensis
Anthrax maculatus
Chrysanthrax sp.1
Hemipenthes sp.1
Diplocampta sp.1
Villa fuscicostata
Villa sp.1
Hyperalonia chilensis
Pachyanthrax sp.1
Pseudopenthes fenestrata
Atrichochira sp.1
Exoprosopa uruguayi
Palirika decora
Ligyra satyrus
Heteralonia (Acrodisca) sp.1
Anthracinae
Global temperature (C°)
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
20
K-Pg Extinction
Triassic
Jurassic
Cretaceous
Paleogene
Neogene
MESOZOIC
CENOZOIC
250
200
150
100
50
0

Table 1
QS scores and BS values of the backbone nodes

	QS scores			BS values			
Relationships	qc	qd	qi	550AA	550NT	216AA	216NT
Bombyliidae	0.56	0.00	0.93	100	98	100	98
Mythicomyiinae	0.97	0.00	0.99	100	100	100	100
Ecliminae	0.76	0.00	0.79	62	68	63	76
Toxophorinae	0.97	0.80	0.99	85	90	87	91
Mariobezziinae	1.00	NA	1.00	100	100	100	100
Phthiriinae	0.86	0.13	0.97	99	100	99	100
Oniromyiinae	0.55	0.57	0.90	87	95	89	94
Bombyliinae (Dischistini)	0.12	0.47	0.84	58	56	47	64
Bombyliinae (Conophorini)	1.00	NA	1.00	100	100	100	100
Tomomyzinae	0.74	0.67	0.64	58	52	58	57
Lomatiinae	0.90	0.00	0.80	99	96	98	96
Lordotinae	1.00	NA	1.00	100	100	100	100
Bombyliinae (Bombyliini + Acrophthalmydini)	0.71	0.94	0.93	98	94	98	95
Neosardus	1.00	NA	1.00	100	100	100	100
Antoniinae	0.96	0.00	1.00	100	100	100	100
Anthracinae	0.06	0.61	0.84	80	75	66	54
Mythicomyiinae + remaining Bombyliidae	0.66	0.61	0.91	100	98	99	97
remaining Bombyliidae	0.82	0.25	0.89	99	97	99	96
Usiinae + Crocidiinae	0.94	0.67	0.92	97	99	97	97
Ecliminae + (Usiinae + Crocidiinae)	0.01	0.82	0.73	88	77	74	82
section II + section III	-0.01	0.85	0.77	90	76	65	77
calde A + clade B	-0.10	0.74	0.81	70	57	42	22
Enica + Mariobezziinae	0.30	0.22	0.91	85	94	77	93
clade B	0.73	0.62	0.95	98	95	98	95
calde C	0.01	0.96	0.84	33	68	63	75
clade D	0.01	0.84	0.78	60	63	63	65
Pantarbes + Conophorini	0.79	0.12	0.90	100	100	99	100
Tomomyzinae + Lomatiinae	0.46	0.82	0.76	67	64	70	70
clade C + (clade D + section III)	0.16	0.96	0.85	37	45	42	21
clade D + section III	-0.02	0.86	0.85	15	23	4	5
Section III	0.16	0.97	0.84	76	91	61	89
Bombyliinae + ((Neosardus + Antoniinae) + Anthracinae)	-0.01	0.86	0.84	92	93	92	94
(Neosardus + Antoniinae) + Anthracinae	0.33	0.83	0.87	95	99	95	99
Neosardus + Antoniinae	-0.01	0.89	0.83	68	53	60	38

Our molecular topology implies a monophyletic "Bombyliinae *sensu stricto*" including only Bombyliini and Acrophthalmydini. However, three tribes of Bombyliinae were not sampled in the present study, and three of the tribes included herein were only represented by a single genus each. Therefore, it is premature to revise the status of bombyliine tribes other than Ecliminae, or to redefine the concept of the world Bombyliinae, based on current evidence.

Tribal level relationships of Toxophorinae. The subfamily Toxophorinae is a well-supported monophyletic group including three tribes: Toxophorini, Gerontini, and Systropodini (Bowden, 1980; Yeates, 1994). Yeates (1994) proposed a sister relationship between Gerontini and Systropodini, which was supported by their rounded head and reduced sperm pump. However, Li and Yeates (2019) suggested Toxophorini and Systropodini were sister to each other based on the absence of occipital apodemes, narrowed squama, slender and elongate abdomen, and the long projection on the posterolateral corner of the male epandrium. The tribal level phylogeny was not resolved by Trautwein et al. (2011). Our phylogeny infers the relationship of Toxophorini + (Gerontini + Systropodini), supporting the hypothesis proposed by Yeates (1994), and also consistent with the single evolution of the endoparasitic larval stage in the clade Gerontini + Systropodini (Yeates and Greathead, 1997).

Sections I and III. Our analyses strongly support the topology of the early branching lineages of Bombyliidae as Heterotropinae + (Mythicomyiinae + (((Usiinae + Crocidiinae) + Ecliminae stat. rev.) + remaining Bombyliidae)). The topology of section I is consistent with the phylogenetic hypotheses presented in Trautwein et al. (2011), except that study proposed that Usiinae and Crocidiinae formed a grade between Mythicomyiinae and the remaining Bombyliidae, while

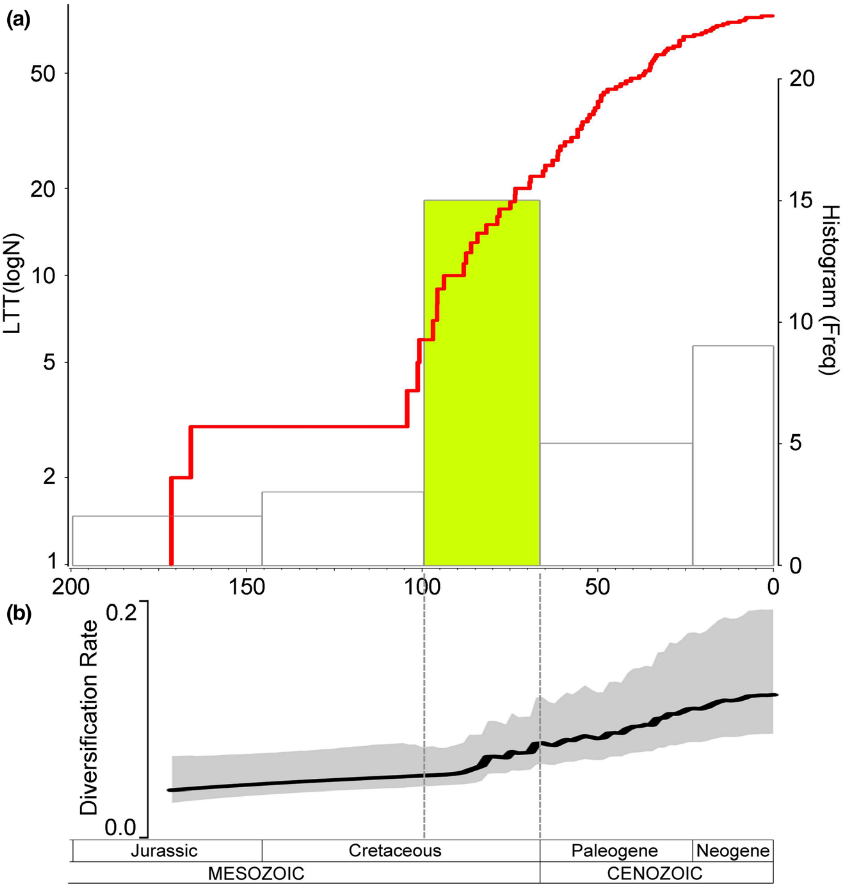


Fig. 5. Diversification dynamics of bee flies. A. Proportional lineages through time (LTT) plot based on the mean values for bee flies and distribution of subfamily and higher-level clade ages for bee flies. B. Global pattern of bee fly species diversification rate based on BAMM analyses.

our analyses propose a monophyletic clade of Ecliminae stat. rev. + (Usiinae + Crocidiinae) as sister to the remaining Bombyliidae.

Section III was well-supported in our analyses and the topology was consistent across all five analyses. Section III consists of a monophyletic lineage of Lordotinae + (Bombyliinae (Bombyliini, Acrophthalmydini) + ((*Neosardus* + Antoniinae) + Anthracinae)). Some regions of section III are similar to the topology of Trautwein et al. (2011), which proposed an unresolved clade including *Neosardus*, Bombyliini, Antoniinae and Anthracinae.

Five of the six tribes of Anthracinae were sampled in our study, and we recovered a relationship of Aphoebantini + ((Xeramoebini + Anthracini) + (Villini + -Exoprosopini)), with *Pachyanthrax* François nested in the Exoprosopini rather than the Villini. Our topology was similar to previous phylogenetic results using morphological data (Lambkin et al., 2003), but placed *Pseudopenthes* Roberts and *Atrichochira* Hesse within the Exoprosopini, unlike the combined molecular and morphological study where these two genera were separated from the remaining Exoprosopini by Villini (Lambkin and Yeates, 2003).

Endoparasitic larvae have been recorded from genera of Villini (*Villa* Lioy, *Poecilanthrax* Osten Sacken, and *Exechohypopion* Evenhuis; Yeates and Greathead, 1997) and Anthracini (*Thraxan* Yeates and Lambkin; Li et al., 2019). Our understanding of the distribution of endoparasitism in the Anthracinae is in its infancy, as is our knowledge of the relationships between the included tribes and genera.

The evolution of key morphological characters

The sand chamber of the female terminalia and the shape of the postcranium are the two most important characters supporting the backbone of the morphological phylogeny of the family (Yeates, 1994). In addition, the shape of the hind margin of the eye, together with the shape of the postcranium, were used to define the “Tomophthalmae” and “Homeophthalmae” (Bezzi, 1924). Our molecular analysis shows that their evolution was not as simplistic as previously assumed, and

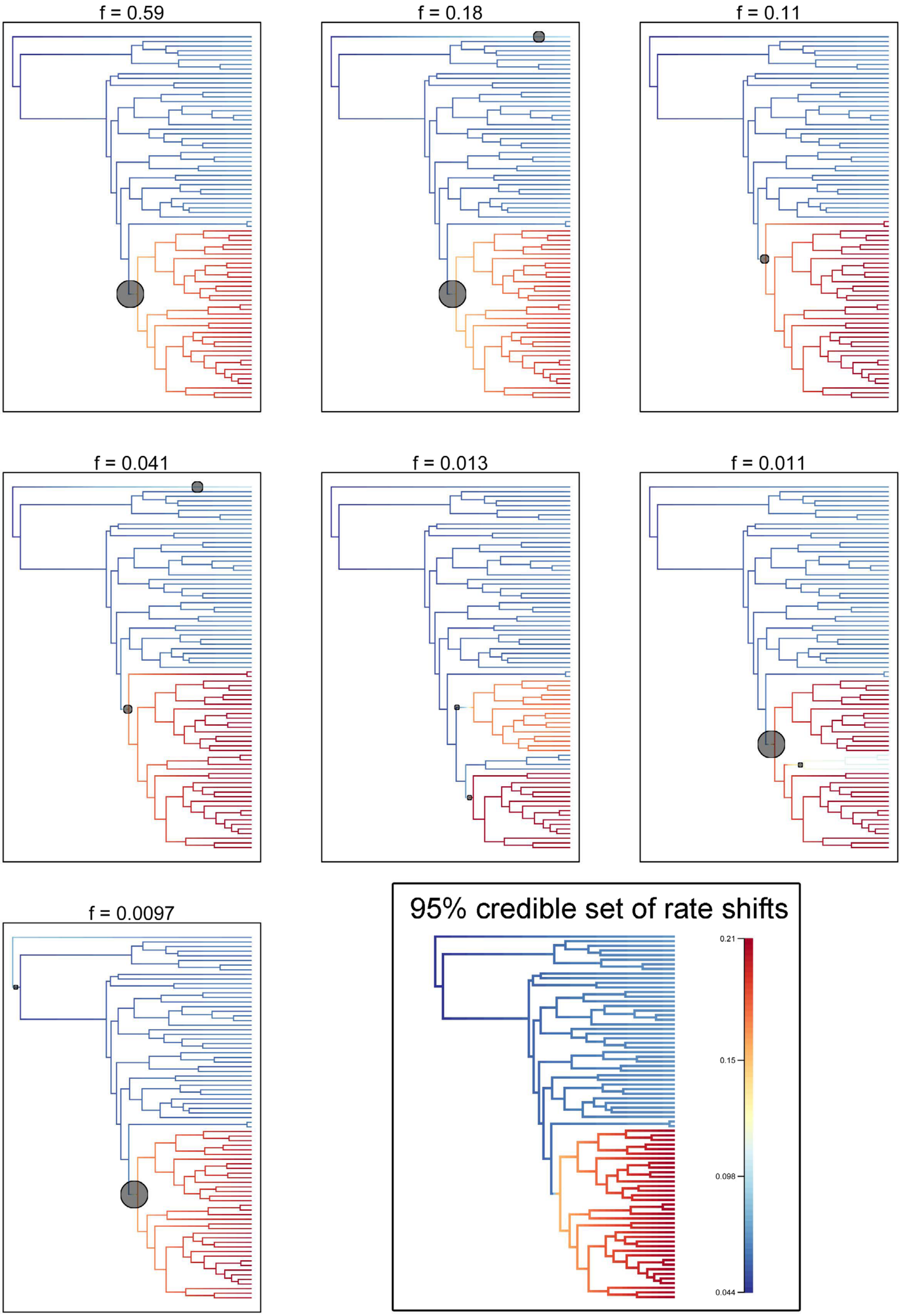

f = 0.59
f = 0.18
f = 0.11
f = 0.041
f = 0.013
f = 0.011
f = 0.0097
95% credible set of rate shifts
0.21
0.15
0.098
0.044

Fig. 6. Set of distinct diversification rate-shift configurations sampled by BAMM during simulation of the posterior probabilities. The seven most commonly sampled configurations are shown. Warm colors indicate high diversification rates and cold colors indicate low diversification rates. Grey dots indicate diversification rate shifts. Larger dots indicate larger diversification rate shifts. The sampling frequency of each diversification scheme is shown above each plot.

character reconstructions suggest that all of these key traits exhibit complex patterns of evolution across the phylogeny.

The evolution of the sand chamber. Sand chamber evolution is correlated with unusual oviposition behaviour: females brush sand into their sand chamber, and subsequently eggs are coated and protected by sand grains before the female oviposits while hovering. This complex structure involves modification of the tergites and sternites of several terminal segments of the female genitalia including sternite 8, tergite 8, and a fused tergite 9 + 10.

The bombyliid sand chamber most likely evolved once before the mid-Cretaceous, on the branch between the Mythicomyiinae and the remaining Bombyliidae, but was lost six times independently in the Usiinae, Toxophorinae, Phthiriinae stat. rev., Lordotinae, *Zentamyia* and *Eristalopsis* (Fig. 3).

The evolution of the postcranium. The postcranium was considered to be flat or slightly tumid in the outgroup, and evolved to form well-developed posteriorly directed lobes through the intermediate state of moderate posterior extensions (Yeates, 1994). The development of the postcranium caused the occipital foramen to retreat into a deep cup-shaped depression.

The cup-shaped postcranium has evolved twice, in the clade Tomomyzinae + Lomatiinae and the clade of (*Neosardus* + Antoniinae) + Anthracinae. A moderately produced postcranium may have evolved a minimum of four times, in *Enica* + Mariobezziinae, Oniromyiinae, *Pantarbes*, and *Neosardus* (Fig. 3).

The evolution of the hind eye margin. Indentation of the hind eye margin most likely evolved twice, separately, in Lomatiinae and the clade of (*Neosardus* + Antoniinae) + Anthracinae) (Fig. 3), and been lost in *Neosardus* and Prorostomatini, although the latter was not sampled in the present analyses. These complex homoplasious changes result in major differences between the morphological interpretation of Bombyliidae phylogeny and our phylogenomic one.

Divergence and diversification dynamics of bee flies

Based on our phylogenetic results, bee flies are divided into 14 monophyletic subfamilies and two broadly polyphyletic subfamilies, Bombyliinae and Cythereinae, each split into at least three clades. Although diverse at the subfamily level, over half of the bee fly generic and specific diversity is constrained to two subfamilies: Anthracinae and Bombyliinae. Within the phylogenetic framework, we show the radiation of angiosperms and historical warm climates might have shaped the present-day diversity of bee flies.

Divergence times of bee flies. The origin of bee flies was estimated by Wiegmann et al. (2003) at approximately 162–90 Ma in the Jurassic-Cretaceous, and by Wiegmann et al. (2011) at approximately 170 Ma. Our estimate of an origin of Bombyliidae between 165–194 Ma is consistent with the hypotheses that the family originated during the radiation of the lower Brachycera, associated with the recovery following the Jurassic extinction event. This period was marked by a gradual cooling and drying of terrestrial habitats, driving larvae to evolve mechanisms to resist desiccation such as parasitism (Wiegmann et al., 2011; Lambkin et al., 2013). Additionally, estimates of internal dates are also consistent with previous studies. For example, the split between *Megapalpus* Macquart and *Corsomyza* Wiedemann in the Mariobezzinae, estimated at approximately 35 Ma here, is consistent with the estimate of De Jager and Ellis (2017) based on a COI molecular clock. Moreover, the origin of *Systropus* is estimated at approximately 35 Ma here, which is later than the origin of the crown group of its mimicry model, Sphecidae, estimated by Peters et al. (2017) at approximately 83 Ma.

Origin of bee fly subfamilies with angiosperms radiation during a hot geological period. The origin of bee fly subfamilies and the divergence of higher clades mostly occurred during the late Cretaceous (100–66 Ma) (Fig. 5A), one of the most remarkable global hothouse periods in geological history (Scotese, 2016). These flower-visiting flies with long proboscides for nectar feeding, radiated in conjunction with the angiosperms (Barba-Montoya et al., 2018).

Our analyses sampled 94 species across the family, including 14 out of 17 subfamilies, accounting for 82.3% of the higher lineages. Therefore, the LTT plot and histogram of the higher-level diversification of bee flies provide sufficient evidence to demonstrate the increased diversification rate (Fig. 5A). A similar increased level of diversification during this hot geological period has been detected in spiders (Shao and Li, 2018), ants (Moreau et al., 2006), and more importantly, in fellow pollinators, the bees (Danforth et al., 2013).

Diversification dynamics of bee flies. The diversification rate-shift analyses indicate that Bombyliidae did not diversify at a constant rate throughout their history. BAMM suggested a remarkable increase in the species diversification rate during the late Cretaceous hothouse period (Scotese, 2016) (Fig. 5B), most likely because of a rate shift in the clade of Bombyliidae + ((*Neosardus* + Antoniinae) + Anthracinae) (Fig. 6). This clade includes two of the most species-rich subfamilies of bee flies, Bombyliinae *sensu stricto* and Anthracinae, both of which originated near the K-Pg boundary. The prior evolution of parasitic larvae may have allowed the rapid radiation of these subfamilies following the K-Pg impact, as that adaptation provided resistance to desiccation in the drier terrestrial environments (Wiegmann et al., 2011; Lambkin et al., 2013).

The subfamily Anthracinae includes over 2000 described species, and is the single largest subfamily of Bombyliidae. Larvae of four anthracine genera: *Poecilanthrax* Osten Sacken, *Villa* Lioy, *Exechohypopion* Evenhuis, and *Thraxan* Yeates and Lambkin, are endoparasitoids (Yeates and Greathead, 1997; Li et al., 2019). It is possible that diversity in this large endoparasitoid lineage is driven by high host specificity as the parasitoid species are probably highly adapted to the internal environment of their host.

Bombyliinae *sensu stricto* comprises the second most diverse bee fly subfamily in terms of genera and species. Most of the species diversity is contained in three genera: *Bombylius* Linnaeus, *Systoechus* and *Anastoechus* Osten Sacken, but the causes of such species richness are unknown. We might surmise that this subfamily of active pollinators co-evolved with the angiosperms during the late Cretaceous taking advantage of the hothouse conditions that both groups are well adapted to.

The bee fly sand chamber evolved prior to these major diversification events. Climate shifts, in conjunction with major changes in host insect and symbiotic plant communities, offered the opportunity for bee flies to exploit these resources with morphological adaptations already present in their populations. However, the evolution of head and eye shape occurred multiple times separately during the major diversification period. These convergences might be explained by the adaption to the new environments.

Conclusion

We generated the first phylogenomic data set and reconstructed evolutionary relationships in the second largest parasitic dipteran family, Bombyliidae. In total, 94 bee fly species from 76 genera were analyzed, representing 14 out of 17 subfamilies. Data sets with 216 and 550 homologous genes were used with different analytical methods to yield a hypothesis of Bombyliidae evolutionary history.

The monophyly of 11 out of 14 sampled subfamilies was recovered, and two tribes were reinstated as subfamilies: Phthiriinae stat. rev. and Ecliminae stat. rev., which bringing the total number of bee fly subfamilies to 19. The *incertae sedis* genus *Sericosoma* was placed in the subfamily Oniromyiinae. A modified key to the subfamilies of Bombyliidae is presented in Appendix S9. The relationships within Bombyliidae are largely resolved, with low support at a minority of nodes.

Both “sand chamber subfamilies” and “Tomophthalmae” are polyphyletic in our phylogenetic hypotheses. Our phylogeny suggests a single origin of the sand chamber before the mid-Cretaceous, but with independent losses of this feature in multiple lineages. The cup-shaped postcranium evolved twice independently, in the Tomomyzinae + Lomatiinae and (*Neosardus* + Antoniinae) + Anthracinae. Endoparasitoidism is found in two distantly related lineages of the tree, and evolved at least twice, once in the ancestor of the clade Gerontini + Systropodini, and at least once in the Anthracinae (Yeates and Greathead, 1997).

Higher-level relationships within Bombyliidae are largely resolved in our analyses, while several phylogenetic questions require future study. The relationships within section II need to be more firmly resolved, and two polyphyletic subfamilies, Cythereinae and Bombyliinae, need redefinition. These challenges may be addressed with denser taxon sampling, especially focused on Cythereinae, Bombyliinae (except Bombyliini and Acrophthalmydini), and Mariobezziinae. The position of Oligodraninae, Xenoprosopinae, and Phthiraxinae need to be considered in light of molecular data when available, and the monophyly of Crocidiinae remains to be tested. The tribes that should be focused on as the highest priorities for inclusion in future analyses are Hallidiini, Adelidiini, Nothoschistini, Peringueyimyiini and Prorostomatini. For the branching of the backbone of section II that experienced rapid diversification, denser taxon sampling may resolve the conflicting phylogenetic signal in our data matrices and improve the resolution of the tree.

Divergence time reconstruction indicated a Jurassic origin for Bombyliidae, and most of the higher-level diversification occurred during the late Cretaceous, perhaps caused by the diversification of plant and other insect lineages which created an increased diversity of adult food and larval host niches. In this study, phylogenomic data demonstrate their great power to resolve the evolutionary history of an ancient and diverse family.

Acknowledgements

This research was supported by the Schlinger Trust Endowment to the ANIC. Thanks to James Lumbers (Canberra), Chris Cohen (Greenville), Ben Parslow (Adelaide), David Ferguson (Canberra), Ding Yang (Beijing), Tingting Zhang (Taian) and Gang Yao (Jinhua) for collecting specimens. Thanks to Brian Cassel (Raleigh) for his help in the molecular lab. Thanks to Juanita Rodriguez (Canberra) for her help with the LTT plot and BAMM analyses. Thanks to Ondrej Hlinka for his help with the CSIRO supercomputer, and Bui Quang Minh for his help with IQ-TREE. Thanks to Rudolf Meier (Singapore), Torsten Dikow (Washington D.C.), Andrew Young (Guelph) and another anonymous reviewer for providing important suggestions to improve the paper. Lambkin thanks Queensland Museum for supporting her systematic studies. Thanks also to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the grant to Carlos Lamas (Proc. No. 302751/2019-0). A.G.E was supported by the National Research Foundation, South Africa (NRF:FBIP grant #110440).

Conflict of interest

None declared.

Data availability

Supplementary materials are also available on DRYAD (https://doi.org/10.5061/dryad.q2bvq83h9).

References

Ababneh, F., Jermiin, L.S., Ma, C. and Robinson, J., 2006. Matched-pairs tests of homogeneity with applications to homologous nucleotide sequences. Bioinformatics 22, 1225–1231.

Armstrong, J., 1979. Biotic pollination mechanisms in the Australian flora-a review. N.Z. J. Bot. 17, 467–508.

Barba-Montoya, J., dos Reis, M., Schneider, H., Donoghue, P.C. and Yang, Z., 2018. Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous Terrestrial Revolution. New Phytol. 218, 819–834.

Becker, T., 1913. Genera Bombyliidarum. Ezheg. Zool. Muz. 17, 421–502.

Bezzi, M., 1924. The Bombyliidae of the Ethiopian Region. London: British Museum (Natural History).

Borowiec, M.L., 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4, e1660.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A. and Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537.

Bowden, J., 1974. Studies in African Bombyliidae. VIII. On the Geroninae. J. Entomol. Soc. South. Afr. 37, 87–108.

Bowden, J., 1975. Studies in African Bombyliidae. IX. On *Hyperusia* Bezzi and the tribe Corsomyzini. J. Entomol. Soc. South. Afr. 38, 99–107.

Bowden, J., 1980. Bombyliidae. In: Crosskey, R.W., Adrian, P.C., Kenneth, S.G.V., Oldroyd, H. (Eds.), Catalogue of the Diptera of the Afrotropical Region. British Museum (Natural History), London, pp. 381–430.

Bowden, J., 1985. The tribal classification of the Bombyliinae with particular reference to the Bombyliini and Dischistini, and the description of a new genus from South America (Dipt., Bombyliidae). Entomol. Mon. Mag. 121, 99–107.

Bowker, A.H., 1948. A test for symmetry in contingency tables. J. Am. Stat. Assoc. 43, 572–574.

Bushnell, B. 2014. BBMap: A Fast, Accurate, Splice-aware Aligner (No. LBNL-7065E). Lawrence Berkeley National Lab. (LBNL), Berkeley, CA.

Bushnell, B., Rood, J. and Singer, E., 2017. BBMerge–accurate paired shotgun read merging via overlap. PLoS One 12, e0185056.

Cerretti, P., O'Hara, J.E., Wood, D.M., Shima, H., Inclan, D.J. and Stireman, J.O. III (2014) Signal through the noise? Phylogeny of the Tachinidae (Diptera) as inferred from morphological evidence. Syst. Entomol. 39, 335–353.

Danforth, B.N., Cardinal, S., Praz, C., Almeida, E.A. and Michez, D., 2013. The impact of molecular data on our understanding of bee phylogeny and evolution. Annu. Rev. Entomol. 58, 57–78.

De Jager, M.L. and Ellis, A.G., 2017. Evolutionary history of a keystone pollinator parallels the biome occupancy of angiosperms in the Greater Cape Floristic Region. Mol. Phylogenet. Evol. 107, 530–537.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. and Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Evenhuis, N.L., 1990. Systematics and Evolution of the Genera in the Subfamilies Usiinae and Phthiriinae (Diptera: Bombyliidae) of the World (Vol. 3). Brill Archive, Honolulu, HI.

Evenhuis, N.L., 1991. World Catalog of Genus-Group Names of Bee Flies (Diptera: Bombyliidae). Bishop Mus. Bull. Ent., Honolulu, HI.

Evenhuis, N.L., 2002. Catalog of the Mythicomyiidae of the World (Insecta: Diptera). Bishop Mus. Bull. Ent., Honolulu, HI.

Evenhuis, N.L., 2014. Family Mythicomyiidae. In: Catalog of the fossil files of the world (Insecta: Diptera) website. Available at: http://hbs.bishopmuseum.org/fossilcat/fossmythico.html.

Evenhuis, N.L., 2017. Family Bombyliidae. In: Catalog of the fossil files of the world (Insecta: Diptera) website. Available at: http://hbs.bishopmuseum.org/fossilcat/fossbomb.html.

Evenhuis, N.L. and Greathead, D.J., 1999. World Catalog of Bee Flies (Diptera: Bombyliidae). Backhuys Publishers, Leiden.

Gillung, J.P., Winterton, S.L., Bayless, K.M., Khouri, Z., Borowiec, M.L., Yeates, D., Kimsey, L.S., Misof, B., Shin, S., Zhou, X., Mayer, C., Petersen, M. and Wiegmann, B.M., 2018. Anchored phylogenomics unravels the evolution of spider flies (Diptera, Acroceridae) and reveals discordance between nucleotides and amino acids. Mol. Phylogenet. Evol. 128, 233–245.

Goloboff, P.A. and Catalano, S.A., 2016. TNT version 1.5, including a full implementation of phylogenetic morphometrics. Cladistics 32, 221–238.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. and Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotech. 29, 644–652.

Greathead, D.J., 1988. The relationships of *Tillyardomyia* Tonnoir with a redefinition of the subfamily Ecliminae (Diptera: Bombyliidae). N.Z. Ent. 11, 14–21.

Grimaldi, D.A., 2016. Diverse orthorrhaphan flies (Insecta: Diptera: Brachycera) in amber from the Cretaceous of Myanmar: Brachycera in Cretaceous amber, Part VII. Bull. Am. Mus. Nat. Hist. 408, 1–132.

Hall, J.C., 1969. Review of the subfamily Cylleniinae with a world revision of the genus *Thevenetimyia* Bigot (*Eclimus* auct.) (Diptera: Bombyliidae). Univ. Calif. Publ. Ent. 56, 1–85.

Hall, J.C. 1976. The Bombyliidae of Chile (Diptera: Bombyliidae). University of California Press, London.

Heard, T., Vithanage, V. and Chacko, E., 1990. Pollination biology of cashew in the Northern Territory of Australia. Aust. J. Agric. Res. 41, 1101–1114.

Heath, T.A., Huelsenbeck, J.P. and Stadler, T., 2014. The fossilized birth-death process for coherent calibration of divergence-time estimates. Proc. Natl Acad. Sci. USA 111, E2957–E2966.

Hesse, A.J., 1956. A revision of the Bombyliidae (Diptera) of Southern Africa. Part III. Ann. S. Afr. Mus. 35, 465–972.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q. and Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.

Hull, F.M., 1973. Bee Flies of the World: The Genera of the Family Bombyliidae. Smithsonian Institution Press, Washington, DC.

Jeffroy, O., Brinkmann, H., Delsuc, F. and Philippe, H., 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22, 225–231.

Jermiin, L., Ho, S.Y.W., Ababneh, F., Robinson, J. and Larkum, A.W., 2004. The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. Syst. Biol. 53, 638–643.

Kalugina, N.S. and Kovalev, V.G., 1985. Dipterous Insects of Jurassic Siberia. Paleontological Institute, Academy of Sciences, Moscow, 198 pp. [In Russian.]

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., von Haeseler, A. and Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.

Katoh, K. and Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Kendall, D.G., 1948. On the generalized “birth-and-death” process. Ann. Math. Stat. 19, 1–15.

Klopfstein, S., Massingham, T. and Goldman, N., 2017. More on the best evolutionary rate for phylogenetic analysis. Syst. Biol. 66, 769–785.

Kuang, T., Tornabene, L., Li, J., Jiang, J., Chakrabarty, P., Sparks, J.S., Naylor, G.J.P. and Li, C., 2018. Phylogenomic analysis on the exceptionally diverse fish clade Gobioidei (Actinopterygii: Gobiiformes) and data-filtering based on molecular clocklikeness. Mol. Phylogenet. Evol. 128, 192–202.

Kück, P., Meusemann, K., Dambach, J., Thormann, B., von Reumont, B.M., Wägele, J.W. and Misof, B., 2010. Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. Front. Zool. 7, 10.

Kumar, S., Filipski, A.J., Battistuzzi, F.U., Kosakovsky Pond, S.L. and Tamura, K., 2011. Statistics and truth in phylogenomics. Mol. Biol. Evol. 29, 457–472.

Lamas, C.J.E. and Couri, M.S., 2004. Cladistic analysis of the Crocidiinae (Diptera, Bombyliidae). Studia Dipt. 11, 513–523.

Lamas, C.J.E. and Nihei, S.S., 2007. Biogeographic analysis of Crocidiinae (Diptera, Bombyliidae): finding congruence among morphological, molecular, fossil and paleogeographical data. Rev. Bras. Ent. 51, 267–274.

Lambkin, C.L., Sinclair, B.J., Pape, T., Courtney, G.W., Skevington, J.H., Meier, R., Yeates, D.K., Blagoderov, V. and Wiegmann, B.M., 2013. The phylogenetic relationships among infraorders and superfamilies of Diptera based on morphological evidence. Syst. Ent. 38, 164–179.

Lambkin, C.L. and Yeates, D.K., 2003. Genes, morphology and agreement: congruence in Australian anthracine bee flies (Diptera: Bombyliidae: Anthracinae). Invert. Syst. 17, 161–184.

Lambkin, C.L., Yeates, D.K. and Greathead, D.J., 2003. An evolutionary radiation of bee flies in semi-arid Australia: Systematics of the Exoprosopini (Diptera: Bombyliidae). Invert. Syst. 17, 735–891.

Larsson, A., 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30, 3276–3278.

Le, S.Q., Dang, C.C. and Gascuel, O., 2012. Modeling protein evolution with several amino acid replacement matrices depending on site rates. Mol. Biol. Evol. 29, 2921–2936.

Le, S.Q. and Gascuel, O., 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. 25, 1307–1320.

Li, X. and Yeates, D.K., 2017. Challenging the definition of the subfamily Usiinae: Phthiraxini, an unusual new tribe of bee flies (Diptera: Bombyliidae), based on *Phthiraxia bowdeni* gen. and sp. nov. from Western Australia. Aust. Ent. 56, 302–310.

Li, X. and Yeates, D.K., 2018. Morphological phylogeny of the Australian genera of the bee fly subfamily Bombyliinae (Diptera: Bombyliidae) with description of four new genera. Invert. Syst. 32, 504.

Li, X. and Yeates, D.K., 2019. Phylogeny, classification and biogeography of bombyliine bee flies (Diptera, Bombyliidae). Syst. Entomol. 44, 862–885.

Li, X., Yuan, D., Rodriguez, J. and Yeates, D.K., 2019. Sand wasp (Hymenoptera: Crabronidae) parasites emerging from mud wasp nests (Hymenoptera: Sphecidae)—a reliable host record of *Thraxan* Yeates & Lambkin (Diptera: Bombyliidae: Anthracinae) with description of the pupal exuviae of three *Thraxan* species. Zootaxa 4609, 149–159.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. and DePristo, M.A., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303.

Meyer, B., Meusemann, K. and Misof, B., 2011. MARE: MAtrix REduction—A Tool to Select Optimized Data Subsets from Supermatrices for Phylogenetic Inference. Zentrum für molekulare Biodiversitätsforschung (zmb) am ZFMK, Bonn, Germany.

Miller, M.A., Pfeiffer, W. and Schwartz, T., 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop, New Orleans.

Minh, B.Q., Nguyen, M.A.T. and von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195.

Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G. et al, 2014) Phylogenomics resolves the timing and pattern of insect evolution. Science 346, 763–767.

Misof, B. and Misof, K.A., 2009. Monte Carlo approach successfully identifies randomness of multiple sequence alignments: a more objective approach of data exclusion. Syst. Biol. 58, 21–34.

Moreau, C.S., Bell, C.D., Vila, R., Archibald, S.B. and Pierce, N.E., 2006. Phylogeny of the ants: diversification in the age of angiosperms. Science 312, 101–104.

Mühlenberg, M., 1970. Besonderheiten im Bau der Receptacula seminis von parasitischen Fliegen (Diptera, Bombyliidae). Z. Morphol. Ökol. Tiere. 87, 471–479.

Mühlenberg, M., 1971. Phylogenetisch-systematische Studien an Bombyliiden (Diptera). Z. Morphol. Ökol. Tiere. 70, 73–102.

Myskowiak, J., Garrouste, R. and Nel, A., 2016. A new genus and species of micro bee fly from the Earliest Eocene French amber (Diptera: Mythicomyiidae: Psiloderoidinae). Zootaxa 4114, 583–586.

Nel, A., 2008. The oldest bee fly in the French Paleocene (Diptera: Bombyliidae). C.R. Palevol. 7, 401–405.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A. and Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.

Pape, T., Blagoderov, V. and Mostovski, M.B., 2011. Order Diptera Linnaeus, 1758. In: Zhang, Z.-Q. (Ed.), Animal Biodiversity: An Outline of Higher-Level Classification and Survey of Taxonomic Richness. Zootaxa, Magnolia Press, Vol. 3148, pp. 222–229.

Paradis, E., 2012. Analysis of Phylogenetics and Evolution with R. Springer Science & Business Media, New York, NY.

Pauli, T., Burt, T.O., Meusemann, K., Bayless, K., Donath, A., Podsiadlowski, L., Mayer, C., Kozlov, A., Vasilikopoulos, A., Liu, S. and Zhou, X., 2018. New data, same story: phylogenomics does not support Syrphoidea (Diptera: Syrphidae, Pipunculidae). Syst. Entomol. 43, 447–459.

Pease, J.B., Brown, J.W., Walker, J.F., Hinchliff, C.E. and Smith, S.A., 2018. Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. Am. J. Bot. 105, 385–403.

Peters, R.S., Krogmann, L., Mayer, C., Donath, A., Gunkel, S., Meusemann, K., Kozlov, A., Podsiadlowski, L., Petersen, M., Lanfear, R. et al, 2017) Evolutionary history of the Hymenoptera. Curr. Biol. 27, 1013–1018.

Petersen, M., Meusemann, K., Donath, A., Dowling, D., Liu, S., Peters, R.S., Podsiadlowski, L., Vasilikopoulos, A., Zhou, X., Misof, B. and Niehuis, O., 2017. Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. BMC Bioinf. 18, 111.

Plummer, M., Best, N., Cowles, K., Vines, K., Sarkar, D. and Almond, R., 2013. coda: Output Analysis and Diagnostics for MCMC. R News 6, 7–11.

Rabosky, D.L., Donnellan, S.C., Grundler, M. and Lovette, I.J., 2014. Analysis and visualization of complex macroevolutionary dynamics: an example from Australian scincid lizards. Syst. Biol. 63, 610–627.

Ranwez, V., Douzery, E.J., Cambon, C., Chantret, N. and Delsuc, F., 2018. MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. Mol. Biol. Evol. 35, 2582–2584.

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

Rohdendorf, B., (1974) The historical development of Diptera. (Vol. 100), pp. 360. University of Alberta Press Edmonton, Alberta, Canada.

Sabrosky, C.W., 1999. Family-group names in Diptera. An annotated catalog. Myia 10, 1–360.

Scotese, C.B., 2016. A New Global Temperature Curve for the Phanerozoic. GSA Annual Meeting in Denver, CO. Abstracts with Programs 48, 7.

Shao, L. and Li, S., 2018. Early Cretaceous greenhouse pumped higher taxa diversification in spiders. Mol. Phylogenet. Evol. 127, 146–155.

Shin, S., Bayless, K.M., Winterton, S.L., Dikow, T., Lessard, B.D., Yeates, D.K., Wiegmann, B.M. and Trautwein, M.D., 2018. Taxon sampling to address an ancient rapid radiation: a supermatrix phylogeny of early brachyceran flies (Diptera). Syst. Entomol. 43, 277–289.

Sinclair, B.J., Cumming, J.M. and Brooks, S.E., 2013. Male terminalia of Diptera (Insecta): a review of evolutionary trends, homology and phylogenetic implications. Insect Syst. Evol. 44, 373–415.

Smith, M.A., Wood, D.M., Janzen, D.H., Hallwachs, W. and Hebert, P.D., 2007. DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera, Tachinidae) are not all generalists. Proc. Natl Acad. Sci. USA 104, 4967–4972.

Stireman, J.O. III, Cerretti, P., O'Hara, J.E., Blaschke, J.D. and Moulton, J.K., 2019. Molecular phylogeny and evolution of world Tachinidae (Diptera). Mol. Phylogenet. Evol. 139, 106358.

Stireman, J.O. III, O'Hara, J.E. and Wood, D.M., 2006. Tachinidae: evolution, behavior, and ecology. Annu. Rev. Entomol. 51, 525–555.

Teasdale, L.C., Köhler, F., Murray, K.D., O'Hara, T. and Moussalli, A., 2016. Identification and qualification of 500 nuclear, single-copy, orthologous genes for the Eupulmonata (Gastropoda) using transcriptome sequencing and exon capture. Mol. Ecol. Resour. 16, 1107–1123.

Theodor, O., 1983. Genitalia of Bombyliidae (Diptera). Israel Academy of Sciences and Humanities, Jerusalem.

Trautwein, M.D., Wiegmann, B.M. and Yeates, D.K., 2010. A multigene phylogeny of the fly superfamily Asiloidea (Insecta): taxon sampling and additional genes reveal the sister-group to all higher flies (Cyclorrhapha). Mol. Phylogenet. Evol. 56, 918–930.

Trautwein, M.D., Wiegmann, B.M. and Yeates, D.K., 2011. Overcoming the effects of rogue taxa: Evolutionary relationships of the bee flies. PLoS Curr. 3, RRN1233.

Waterhouse, R.M., Tegenfeldt, F., Zdobnov, E.M. and Kriventseva, E.V., 2013. OrthoDB: a hierarchical catalog of animal fungal and bacterial orthologs. Nucleic Acids Res. 41, 358–365.

Wedmann, S. and Yeates, D.K., 2008. Eocene records of bee flies (Insecta, Diptera, Bombyliidae, Comptosia): Their palaeobiogeographic implications and remarks on the evolutionary history of bombyliids. Palaeontology 51, 231–240.

Wiegmann, B.M., Thorne, J.L., Yeates, D.K. and Kishino, H., 2003. Time flies: a new molecular time-scale for fly evolution without a clock. Syst. Biol. 52, 745–756.

Wiegmann, B.M., Trautwein, M.D., Winkler, I.S., Barr, N.B., Kim, J., Lambkin, C., Bertone, M.A., Cassel, B.K., Heimberg, A.M. et al (2011) Episodic radiations in the fly tree of life. Proc. Natl Acad. Sci. USA 108, 5690–5695.

Winkler, I.S., Blaschke, J.D., Davis, D.J., Stireman, J.O. III, O'Hara, J.E., Cerretti, P. and Moulton, J.K., 2015. Explosive radiation or uninformative genes? Origin and early diversification of tachinid flies (Diptera: Tachinidae). Mol. Phylogenet. Evol. 88, 38–54.

Woodley, N.E., 1989. Phylogeny and classification of the "Orthorrhaphous" Brachycera. In: McAlpine, J.F. and Wood, D.M. (coord.) (Eds.), Manual of Nearctic Diptera, Chapter 115, Vol. 3, Monograph No. 32, pp. 1371–1395. Research Branch, Agriculture Canada, Ottawa, ON.

Woodley, N.E., Borkent, A. and Wheeler, T.A., 2009. Phylogeny of the Diptera. In: Brown, B.V., Borkent, A., Cumming, J.M., Wood, D.M., Woodley, N.E., Zumbado, M.A. (Eds.), Manual of Central American Diptera Volume 1, pp. 79–94. NRC Research Press, Ottawa, ON.

Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J., Liu, S., Huang, W., He, G., Gu, S., Li, S., Zhou, X., Lam, T., Liu, Y., Xu, X., Wong, G.K. and Wang, J., 2014. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinformatics 30, 1660–1666.

Yamaguchi, C., Lamas, C.J.E. and Evenhuis, N.L., 2017. A new Neotropical genus of bee flies (Diptera, Bombyliidae, Phthiriinae). Zootaxa 4363, 521–534.

Ye, X., Yao, G., Shih, C., Ren, D. and Wang, Y., 2019. New bee flies from the mid-Cretaceous Myanmar amber (Brachycera: Asiloidea: Bombyliidae). Cretac. Res. 100, 5–13.

Yeates, D.K., 1994. Cladistics and classification of the Bombyliidae (Diptera: Asiloidea). Bull. Am. Mus. Nat. Hist. 219, 191.

Yeates, D.K., 2002. Relationships of extant lower Brachycera (Diptera): a quantitative synthesis of morphological characters. Zool. Scr. 31, 105–121.

Yeates, D.K. and Greathead, D.J., 1997. The evolutionary pattern of host use in the Bombyliidae (Diptera): a diverse family of parasitoid flies. Biol. J. Linn. Soc. 60, 149–185.

Yeates, D.K. and Lambkin, C.L., 1998. Cryptic species diversity and character congruence: review of the tribe Anthracini (Diptera: Bombyliidae) in Australia. Invert. Syst. 12, 977–1078.

Young, A.D., Lemmon, A.R., Skevington, J.H., Mengual, X., Ståhls, G., Reemer, M., Jordaens, K., Kelso, S., Lemmon, E.M., Hauser, M. et al, 2016) Anchored enrichment dataset for true flies (order Diptera) reveals insights into the phylogeny of flower flies (family Syrphidae). BMC Evol. Biol. 16, 143.

Yu, Y., Harris, A.J., Blair, C. and He, X., 2015. RASP (reconstruct ancestral state in phylogenies): a tool for historical biogeography. Mol. Phylogenet. Evol. 87, 46–49.

Zaitsev, V.F., 1992. Contribution to the phylogeny and systematics of the superfamily Bombylioidea (Diptera). Entomol. Rev. 71, 94–114.

Zwick, A., Regier, J.C. and Zwickl, D.J., 2012. Resolving discrepancy between nucleotides and amino acids in deep-level arthropod phylogenomics: differentiating serine codons in 21-amino-acid models. PLoS One 7, e47450.

Supporting Information

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

Appendix S1. Taxa sampled in the present study, with number of loci retained after all cleaning steps.

Appendix S2. Average pairwise distance of 916 amino acid loci, showing selected loci with moderate substitution rate.

Appendix S3. SymTest and AliStat heat maps of pairwise sequence comparisons for the data sets.

Appendix S4. ML and MP trees from different data sets and partitioning schemes.

Appendix S5. Topology generated by 550AA, with node numbers.

Appendix S6. QS scores, BS values, divergence times and fossil information from different analyses for the preferred topology from 550AA.

Appendix S7. Relevant sampling fractions of different clades for diversification analysis.

Appendix S8. Partitioning scheme and model selection of 216AA_MF+MERGE, 550NT, and 216NT.

Appendix S9. Key to subfamilies of Bombyliidae (modified from Yeates (1994)).