Systematics and evolution of predatory flower flies (Diptera Syrphidae) based on exon-capture sequencing

Mengual, Ximo

2023-04

Mengual , X , Mayer , C , Burt , T O , Moran , K M , Dietz , L , Nottebrock , G , Pauli , T , Young , A D , Brasseur , M V , Kukowka , S , Kelso , S , Etzbauer , C , Bot , S , Hauser , M , Jordaens , K , Miranda , G F G , Ståhls , G , Van Steenis , W , Peters , R S & Skevington , J H 2023 , ' Systematics and evolution of predatory flower flies (Diptera Syrphidae) based on exon-capture sequencing ' , Systematic Entomology , vol. 48 , no. 2 , pp. 250-277 . https://doi.org/10.1111/syen.1257

http://hdl.handle.net/10138/356580
https://doi.org/10.1111/syen.12573

Received: 29 June 2022 | Accepted: 27 September 2022

DOI: 10.1111/syen.12573

ORIGINAL ARTICLE

Systematic Entomology

Royal Entomological Society

Systematics and evolution of predatory flower flies (Diptera: Syrphidae) based on exon-capture sequencing

Ximo Mengual[1] | Christoph Mayer[1] | Trevor O. Burt[1] | Kevin M. Moran[2,3] | Lars Dietz[1] | Gaby Nottebrock[1] | Thomas Pauli[1,4] | Andrew D. Young[5] | Marie V. Brasseur[1] | Sandra Kukowka[1] | Scott Kelso[2] | Claudia Etzbauer[1] | Sander Bot[6] | Martin Hauser[7] | Kurt Jordaens[8] | Gil F. G. Miranda[2] | Gunilla Ståhls[9] | Wouter van Steenis[10] | Ralph S. Peters[1] | Jeffrey H. Skevington[2,3]

[1]Zoologisches Forschungsmuseum Alexander Koenig, Leibniz-Institut zur Analyse des Biodiversitätswandels, Bonn, Germany

[2]Agriculture and Agri-Food Canada, Canadian National Collection of Insects, Arachnids and Nematodes, Ottawa, Ontario, Canada

[3]Department of Biology, Carleton University, Ottawa, Ontario, Canada

[4]Institute of Medical Bioinformatics and Systems Medicine, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany

[5]School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada

[6]Haren, The Netherlands

[7]California Department of Food and Agriculture, Sacramento, California, USA

[8]Royal Museum for Central Africa, Invertebrates Unit, Tervuren, Belgium

[9]University of Helsinki, Finnish Museum of Natural History Luomus, Helsinki, Finland

[10]Breukelen, The Netherlands

Correspondence
Ximo Mengual, Zoologisches Forschungsmuseum Alexander Koenig, Leibniz-Institut zur Analyse des Biodiversitätswandels, Adenauerallee 127, 53113 Bonn, Germany.
Email: x.mengual@leibniz-lib.de

Funding information
Agriculture and Agri-Food Canada; H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 642241

Abstract

Flower flies (Diptera: Syrphidae) are one of the most species-rich dipteran families and provide important ecosystem services such as pollination, biological control of pests, recycling of organic matter and redistributions of essential nutrients. Flower fly adults generally feed on pollen and nectar, but their larval feeding habits are strikingly diverse. In the present study, high-throughput sequencing was used to capture and enrich phylogenetically and evolutionary informative exonic regions. With the help of the BAITFISHER software, we developed a new bait kit (SYRPHIDAE1.0) to target 1945 CDS regions belonging to 1312 orthologous genes. This new bait kit was successfully used to exon capture the targeted loci in 121 flower fly species across the different subfamilies of Syrphidae. We analysed different amino acid and nucleotide data sets (1302 loci and 154 loci) with maximum likelihood and multispecies coalescent models. Our analyses yielded highly supported similar topologies, although the degree of the SRH (global stationarity, reversibility and homogeneity) conditions varied greatly between amino acid and nucleotide data sets. The sisterhood of subfamilies Pipizinae and Syrphinae is supported in all our analyses, confirming a common origin of taxa feeding on soft-bodied arthropods. Based on our results, we define Syrphini **stat.rev.** to include the genera *Toxomerus* and *Paragus*. Our divergence estimate analyses with BEAST inferred the origin of the Syrphidae in the Lower Cretaceous (125.5–98.5 Ma) and the diversification of predatory flower flies around the K–Pg boundary (70.61–54.4 Ma), coinciding with the rise and diversification of their prey.

KEYWORDS
hover flies, phylogeny, Syrphidae, Syrphinae, target DNA enrichment, tribal classification

INTRODUCTION

Syrphidae (Insecta: Diptera) is a charismatic family with over 6300 described species worldwide (Skevington et al., 2019) (Figure 1). Syrphids are popular among citizen-scientists and dedicated amateurs for their conspicuous nature and attractive coloration, while simultaneously misunderstood by the broader public due to frequent confusion with stinging Hymenoptera. Most syrphid adults have the ability to hang suspended in the air and visit flowers to fulfil their diet requirements of pollen and nectar, and also use them as mating sites, leading to the common names hover and flower flies (Figure 1). Compared with the adults, syrphid larvae are very variable in structure, habits and feeding modes (Inouye et al., 2015; Larson et al., 2001; Rotheray & Gilbert, 2011; Willmer, 2011). Larval feeding strategies comprise phytophagous species and pollen feeders (Ricarte et al., 2017; Stuke, 2000; Weng & Rotheray, 2009; Zuijen & Nishida, 2011); fungivorous larvae feeding on fungal fruiting bodies; saprophagous species (including saproxylic) feeding on dung, decaying wood and filter feeding in water bodies of several types (Rotheray & Gilbert, 1999, 2011), and predatory taxa feeding on brood of social Hymenoptera and on several groups of soft-bodied arthropods (Reemer, 2003; Rojo et al., 2003). Kleptoparasitic or parasitoid feeding strategies are known as well (Fleischmann et al., 2016; Pérez-Lachaud et al., 2014).

Syrphids provide many important ecosystem services (Dunn et al., 2020), defined as the benefits to human well-being provided by organisms interacting in natural ecosystems (Daily, 1997). Adults are key pollinators in natural ecosystems and agricultural areas (Cook et al., 2020; Hodgkiss et al., 2018; Inouye et al., 2015; Jauker & Wolters, 2008; Pérez-Bañón et al., 2007; Ssymank et al., 2008; Ssymank & Kearns, 2009) and visit 52% of global crop plants (Rader et al., 2016, 2020). The ecological pollination service provided by the flower flies is estimated to be worth around US $360 billion per year (Doyle et al., 2020) and is vital for human food safety (Klein et al., 2007; Potts et al., 2016). Moreover, their larvae act as biological control agents of pests on a large variety of arthropods (Arcaya et al., 2017; Bellefeuille et al., 2019; Moerkens et al., 2021; Nelson et al., 2012; Tenhumberg, 1995), but also on certain weeds (Grosskopf, 2005; Rizza et al., 1988; Sheppard et al., 1995), and as decomposers of organic matter (Lardé, 1989, 1990; Morales & Wolff, 2010). Besides these well-known ecosystem services, the long-range migratory movements of flower flies (Aubert & Goeldlin de Tiefenau, 1981; Finch & Cook, 2020; Menz et al., 2019) redistribute essential nutrients and transport billions of pollen grains between distant areas, maintaining the genetic diversity of the visited plants. These migrations represent a considerable net transport of biomass and energy, and also facilitate pollen transfer and the regulation of crop pests (Wotton et al., 2019). Thus, the life history of Syrphidae and their abundance make them vital players in maintaining ecological networks and providing beneficial services. Flower flies do not only provide this large variety of ecosystem services, but can also be used as bioindicators to assess biodiversity loss and the efficiency of restoration and conservation policies (Ricarte et al., 2011; Sommaggio, 1999; Tscharntke et al., 2005). Unlike some other important beneficial insects, evidence of syrphid diversity declines is limited (Biesmeijer et al., 2006; Wotton et al., 2019), but recent studies show strong declines in their abundance (Barendregt et al., 2022; Gatter et al., 2020). Nevertheless, their ecosystem services may prove to be increasingly essential in the Anthropocene. Finally, flower flies have a direct economic and medical significance to human kind as insects of forensic importance (Heo et al., 2020; Magni et al., 2013) and as myiasis agents (Pérez-Bañón et al., 2020).

Syrphidae are currently divided into four subfamilies: Microdontinae, Eristalinae, Pipizinae and Syrphinae (Mengual et al., 2015). All subfamilies are recovered as monophyletic groups in recent phylogenetic studies with the exception of Eristalinae (Hippa & Ståhls, 2005; Mengual et al., 2015; Moran et al., 2022; Mullens et al., 2022; Pauli et al., 2018; Young, Lemmon, et al., 2016). Microdontine immatures live inside ant nests and feed on ant brood or parasitize them (Pérez-Lachaud et al., 2014; Reemer, 2013), whereas eristaline larvae are mostly saprophagous (Aracil et al., 2019; Pérez-Bañón et al., 2003; Rotheray, 1993), but there are also predatory immatures in phytotelmata and in wasp and bee nests (Rotheray, 2003; Rupp, 1989) and phytophagous species that may be agricultural pests (Brunel & Cadou, 1994; Edwards & Bevan, 1951; Ricarte et al., 2017; Souba-Dols et al., 2020; Tompsett, 2002). The other two subfamilies, Pipizinae and Syrphinae, have primarily predaceous larvae feeding mostly on soft-bodied Hemiptera like aphids, scale insects, psyllids and white flies, but also on other arthropods such as thrips and larvae of other insects (Rojo & Marcos-García, 1997; Rojo et al., 2003 and references therein; Downes et al., 2017). Some New World species of the Syrphinae have secondarily phytophagous larvae that feed on pollen, mine leaves, or bore plant stems (Dumbardon-Martial, 2016; Mengual et al., 2008a; Nishida et al., 2002; Reemer & Rotheray, 2009; Weng & Rotheray, 2009; Zuijen & Nishida, 2011).

Syrphinae comprises approximately 1800 described species and represents the largest radiation with primarily predatory larvae within the family. The subfamily has classically been divided into four tribes, that is, Bacchini, Paragini, Syrphini and Toxomerini (Dušek & Láska, 1967; Vockeroth, 1969, 1992). Molecular studies pointed out the need of a revision for the tribal classification of Syrphinae (Mengual, 2015; Mengual et al., 2008b, 2012, 2015; Mengual & Thompson, 2011; Pauli et al., 2018; Young, Marshall, & Skevington, 2016) since Bacchini was never resolved as monophyletic and the tribes Paragini and Toxomerini rendered Syrphini paraphyletic. After discussing some external adult morphological characteristics and male genitalia, Mengual (2020), using limited molecular data, divided the bacchines into two more inclusive tribes, Bacchini and Melanostomini.

The tribe Paragini is monotypic and has been recognized as a distinct taxon almost since the erection of the genus *Paragus* Latreille (Vujić et al., 2008), although the tribe was formally erected by Glumac (1960) based on male genitalia characters (previous combinations of suprageneric rank were not adopted; see Sabrosky, 1999). The current tribal concept is based on larval and adult morphology (Vujić et al., 2008) and its phylogenetic relationship with other Syrphini

FIGURE 1 Images of some genera and species included in the present study. (a) *Pipiza femoralis*, by Scott King (https://creativecommons.org/licenses/by-nc/4.0/); (b) *Doros aequalis*, by Stephen A. Marshall (https://creativecommons.org/licenses/by-nc/4.0/); (c) *Baccha elongata*, by carnifex (https://creativecommons.org/licenses/by/4.0/); (d) *Dideopsis aegrota*, by suchihfen (https://creativecommons.org/licenses/by-nc/4.0/); (e) *Melangyna lasiophthalma*, by Michael Knapp (https://creativecommons.org/licenses/by/4.0/); (f) *Melangyna collatus*, by Bruce Tardif (https://creativecommons.org/licenses/by-nc/4.0/); (g) *Allobaccha* sp. (*picta*-group), by Lauren Steyn (https://creativecommons.org/licenses/by-nc-nd/4.0/); (h) *Allobaccha* sp., by © Colin Hutton; (i) *Allobaccha amphithoe*, by Nikolai Vladimirov (https://creativecommons.org/licenses/by-nc/4.0/); (j) *Salpingogaster punctifrons*, by Wayne Fidler (https://creativecommons.org/licenses/by-nc/4.0/); (k) *Asarkina* sp., by Marcus F.C. Ng (https://creativecommons.org/licenses/by-nc/4.0/); (l) *Asarkina* sp., by David Kohl (https://creativecommons.org/licenses/by-nc/4.0/); (m) *Pelecinobaccha costata*, by skitterbug (https://creativecommons.org/licenses/by/4.0/); (n) *Austroscaeva occidentalis*, by Andrea Cocucci (https://creativecommons.org/licenses/by-nc/4.0/); (o) *Antillus ascitus*, by © Franklin Howley-Dumit Serulle.

genera remained uncertain until recently (Mengual, 2015; Mengual et al., 2015; Mengual, Ståhls, et al., 2018; Young, Lemmon, et al., 2016). Earlier molecular phylogenetic studies using parsimony as the optimality criterion (Mengual, 2015; Mengual et al., 2008b) resolved *Allobaccha* Curran as sister to Paragini, whereas the same (Mengual, 2015) or similar (Mengual, Ståhls, et al., 2018) data set under maximum likelihood placed *Paragus* as sister taxon of the *Eupeodes-Scaeva* clade within Syrphini. Mengual et al. (2015) hypothesized that the sister relationship between *Paragus* and *Allobaccha* was due to long-branch attraction (Bergsten, 2005; Felsenstein, 1978).

Enderlein (1938) established the tribe Toxomerini for the single genus *Toxomerus* Macquart and nine other genera (a mix of taxa from Syrphini, Melanostomini and Toxomerini), but it was Vockeroth (1969) who recognized this tribe as monogeneric. All recent phylogenetic studies resolved *Toxomerus* and the genus *Eosalpingogaster* Hull embedded within the large Neotropical radiation of the old generic concept of *Ocyptamus* Macquart (Mengual, 2015, 2020; Mengual et al., 2008b, 2012, 2015, 2021; Mengual & Thompson, 2011; Miranda et al., 2016). Recent taxonomic revisions have divided the old generic concept of *Ocyptamus* into 17 different genera (Mengual, Miranda, & Thompson, 2018; Miranda et al., 2014, 2020), but the phylogenetic relationships among these genera and *Toxomerus* and *Eosalpingogaster* are still not well understood.

High-throughput sequencing (HTS) enables genome-scale, cost-effective data collection and facilitates a larger number of taxa to be processed (Andermann et al., 2020; Lemmon & Lemmon, 2013). Genome reduction methods, also known as genomic partitioning (Turner et al., 2009) or genome-subsampling, are applied to generate genomic subsets suitable for phylogenetic inference (Bravo et al., 2019; McCormack et al., 2013) due to the prohibitive cost and computational demands of whole-genome sequencing, annotation and analysis. Among these techniques, sequence-capture methods (also termed target enrichment or targeted sequencing; Mamanova et al., 2010) are used to enrich sequence libraries for specific regions of interest from a genome (Faircloth et al., 2012; Gnirke et al., 2009; Lemmon et al., 2012; Lemmon & Lemmon, 2013). There are several strategies to enrich target DNA using hybrid-capture methods, which target different genomic loci. Frequently used in phylogenomics are ultra-conserved elements capture (UCEs; Faircloth et al., 2012; Faircloth, 2017) and anchored hybrid enrichment (AHE; Lemmon et al., 2012) that use (ultra)conserved regions in the genome across taxa as binding sites to enrich more variable flanking regions. Both strategies have important benefits (Bossert & Danforth, 2018; Young & Gillung, 2020; Zhang et al., 2019), but also drawbacks (Bank et al., 2017; Eberle et al., 2020) such as the sequencing of phylogenetically uninformative sequence sections as they serve as anchors for the enrichment, uncertain orthology of flanking regions and the low probability that the flanking regions contain coding DNA (but see Bossert & Danforth, 2018). Here, the original exon-capture method proposed by Gnirke et al. (2009) was used to enrich exons, independently of whether they are highly conserved among species or not. The BAITFISHER software (Mayer et al., 2016) was developed to design hybrid enrichment baits from multiple sequence alignments for a wide range of scenarios, including exon capture of less similar sequences, with the aim to minimize the number of required baits for the taxonomic group of interest. The strategy to enrich orthologous single-copy protein-coding genes has successfully been applied to several plant families (Li et al., 2017) and to different metazoan taxa such as stony corals (Anthozoa: Scleractinia; Quek et al., 2020), sea spiders (Pycnogonida: Pantopoda; Dietz et al., 2019), isopods (Malacostraca: Isopoda; Stringer et al., 2021), wasps (Insecta: Hymenoptera; Bank et al., 2017; Klopfstein et al., 2019; Maletti et al., 2021; Mayer et al., 2016; Pauli et al., 2021), butterflies and moths (Insecta: Lepidoptera; Call et al., 2021; Mayer et al., 2021), and cockroaches (Insecta: Blattodea; Evangelista et al., 2021). Moreover, it has also been successfully used in freshwater macrozoobenthos metabarcoding (Gauthier et al., 2020) and environmental DNA studies (Giebner et al., 2020).

Although most phylogenetic studies in Syrphidae are based on morphology and/or a few molecular markers, different genome reduction methods have already been used as well. Young, Lemmon, et al. (2016) presented the first use of AHE in insect phylogenetics with 343 analysed loci (from 559 targeted loci) and a data set containing 30 flower fly species. Later, Pauli et al. (2018) used 3145 genes derived from transcriptomes (including 10 syrphids) to study, and ultimately reject, the old concept of a monophyletic Syrphoidea. Data type and taxon sampling were different in these two studies, but results from both are comparable and echo those from studies with few loci. Despite efforts to construct a robust evolutionary framework and phylogenetic hypothesis for the Syrphinae genera, there still are taxa with ambiguous phylogenetic relationships or whose phylogenetic placement has not been studied. Interpreting the evolutionary history of the group relies on a stable classification and a better understanding of their phylogenetic relationships. In particular, uncovering the phylogeny of the subfamily Syrphinae will provide insight into the origin of adult migration, the ecological basis of predation and the potential use as biological control agents of certain taxa, and the multiple switches between different larval feeding modes and their pollination patterns, such as flower plants visited and time.

In the present study, we aimed to infer a robust phylogeny of the predatory flower flies (subfamilies Syrphinae and Pipizinae) and to address the problems of the current tribal classification in Syrphinae. A secondary aim was to infer the phylogenetic placement of *Allobaccha* and *Asarkina* Macquart, both genera in need of a taxonomic revision, as part of the project 'BIG4 – Biosystematics, Informatics and Genetics of the big 4 insect groups: training tomorrow's researchers and entrepreneurs' (http://big4-project.eu/ and https://cordis.europa.eu/project/id/642241) (see Burt & Mengual, 2017). To accomplish our goals, we designed a new target DNA enrichment bait kit (SYRPHIDAE1.0) to capture and phylogenetically analyse more than 1300 single-copy protein-coding genes from 123 flower fly species belonging to all four subfamilies. Based on the resulting topology, we perform a dating analysis for the family Syrphidae based on carefully selected flower fly fossils to gain insight into the evolutionary timing and pattern of these charismatic and speciose group of Diptera.

MATERIALS AND METHODS

Taxon sampling and specimen documentation

Taxa were selected to cover as much generic and subgeneric diversity as possible within the subfamilies Pipizinae and Syrphinae with special emphasis on two genera, *Allobaccha* and *Asarkina*. Several genera from Microdontinae and Eristalinae were selected as outgroup taxa based on Young, Lemmon, et al. (2016). All flower flies were collected with a hand-net or a Malaise trap and preserved and stored in 96% ethanol at −20 or −80°C until extraction. We used our new bait kit SYRPHIDAE1.0 (see below) to enrich genomic DNA for the loci of interest for 123 flower fly species, but two taxa failed and were not further used in our study (see Table S1). These taxa represent three Microdontinae genera, 14 genera and subgenera in Eristalinae, 5 of 8 genera of Pipizinae, 9 of 10 genera and subgenera of Bacchini, 7 of 8 genera and subgenera of Melanostomini, the genus *Toxomerus* (Toxomerini), all 4 subgenera of *Paragus* (Paragini), and 65 of 81 genera and subgenera of Syrphini (see Table S1). Furthermore, we also mined the loci of interest in 14 previously assembled dipteran transcriptomes from Pauli et al. (2018), comprising 10 syrphid species, *Sapromyza sciomyzina* Schiner (Lauxaniidae), *Nephrocerus atrapilus* Skevington (Pipunculidae), *Platypeza anthrax* Loew (Platypezidae) and *Meroplius fasciculatus* (Brunetti) (Sepsidae) (see Table S1).

We followed the recommendations by Meier (2017) and Packer et al. (2018) on species names, taxon concepts, identification methods and literature, and voucher depository (summarized in Table S1). Specimen information (locality, date, collector, identifier and unique identifier) is accessible via the GenBank Accession Numbers (Table S1) and in BOLD (https://www.boldsystems.org/) under the data set DS-SYRPHTE (https://doi.org/10.5883/DS-SYRPHTE). Moreover, entire specimens or remnants of the studied specimens were preserved and properly labelled as DNA voucher specimens, and deposited at the Royal Museum for Central Africa (KMMA), the Canadian National Collection of Insects, Arachnids and Nematodes (CNC), and the Zoological Research Museum Alexander Koenig (ZFMK), as listed in Table S1.

Gene selection

For the present study, we used the same genetic markers as Pauli et al. (2018). We searched the OrthoDB 7 database for genes that are single copy and present in all of the five reference species (i.e., *Bombyx mori* Linnaeus [Lepidoptera: Bombycidae], *Danaus plexippus* [Linnaeus] [Lepidoptera: Nymphalidae], *Aedes aegypti* [Linnaeus] [Diptera: Culicidae], *Drosophila melanogaster* Meigen [Diptera: Drosophilidae] and *Glossina morsitans* Westwood [Diptera: Glossinidae]), which span a wide range of Mecopterida. This search yielded 3145 single-copy orthologous genes (OGs), which were used as candidate genes for the Bait design in this study.

In order to generate a marker set that has an overlap with previous studies, which included only a small number of genomic loci, we included baits for another 10 coding loci. Based on available DNA sequences of different flower fly species of all subfamilies (Moran et al., 2022; J.H. Skevington, unpublished), we designed baits also for three protein-coding genes, that is, the 5′-end of the alanyl-tRNA synthetase (AATS), *period* (PER), and the 5′-end (CAD1) and 3′-end (CAD4/5) of the carbamoylphosphate synthase region of CAD (carbamoyl-phosphate synthetase 2, aspartate transcarbamylase and dihydroorotase). These genes have been used to infer phylogenetic relationships in Diptera (Gibson et al., 2011; Mengual et al., 2017; Moran et al., 2022). Moreover, we also designed new baits for six loci from the AHE data set by Young, Lemmon, et al. (2016), namely, the nuclear loci L35, L44 (RNA-binding Protein 15 or RBP-15), L54 (also known as the 5′-end of the Tubby-like Protein or TULP), L113, L133 and L272. Besides the newly designed baits for the mentioned genes (AATS, PER, CAD1, CAD4/5, L35, RBP-15, TULP, L113, L133 and L272), we also included in our final bait kit the original 10 baits for the locus L352 (also known as the 5′-end of the Casein kinase 1 or CK1) from Young, Lemmon, et al. (2016).

Bait design for target enrichment of genomic DNA

Template alignments for the bait design of the 3145 OGs were extracted from alignments generated in the study by Pauli et al. (2018). For the 3145 multiple sequence alignments (MSAs) of the OGs generated in Pauli et al. (2018), we extracted the sequences of the 10 syrphid species belonging to the four subfamilies (see Table S1). Template alignments for the present study were obtained by removing from the reduced alignments all positions consisting only of gaps and/or ambiguous nucleotides.

We designed hybrid enrichment baits for the 3145 OGs using the software BAITFISHER version 1.2.7 and BAITFILTER version 1.0.5 (Mayer et al., 2016) as follows: we used the gene models (official gene set version 3.2) of the common fruit fly *D. melanogaster* and the corresponding genome assembly (version 6.12; Hoskins et al., 2015) to allow BAITFISHER to excise individual exons from the template alignments. BAITFISHER now designed baits for every possible starting position of a bait region (called candidate bait regions) in all exons for a tiling design of three baits of length 120 bp with an offset of 20 bp between consecutive baits. This implies a bait region size of 160 bp. Considering the phylogenetic relations and branch lengths inferred by Pauli et al. (2018), we required that the nucleotide sequence of at least one flower fly species and the nucleotide sequence of each non-Syrphidae species had to be present in full length in each 160 bp long bait region. For the BAITFISHER runs, we specified a clustering threshold of 0.15 for the hierarchical clustering, and used the heuristic algorithm to determine the Hamming 1-centre sequence for each cluster (Mayer et al., 2016).

For the 10 coding loci that we added for backward compatibility with previous studies, we obtained template alignments for the coding regions from Young, Lemmon, et al. (2016), Moran et al. (2022) and J.H. Skevington (unpublished sequences). BAITFISHER was used with the

same settings as above to design baits for all possible candidate bait regions, except that no exons were excised in this case.

After designing baits for all possible candidate bait regions, we used BAITFILTER to assess whether baits in a candidate bait region could fail binding full length on flower fly genomic DNA because the target region is interrupted by an intron not present in the template alignment of transcriptomic sequences, or whether baits might bind to multiple regions in the genome. We used BAITFILTER to conduct a BLAST search (Altschul et al., 1990) of all baits in all candidate bait regions against an unpublished draft genome of the drone fly *Eristalis tenax* (Linnaeus) (Diptera: Syrphidae). We kept only those candidate bait regions for which (i) all baits had not more than one very good BLAST hit to the unpublished *E. tenax* draft genome (BAITFILTER options '--blast-l --blast-first-hit-evalue 0.000001 --blast-second-hit-evalue 0.000001') and for which (ii) at least one of the bait sequences of each bait stack mapped at least 84% of its length to this genome (BAITFILTER option '--blast-min-hit-coverage-of-baits-in-tiling-stack 0.84'; same settings as in Bank et al., 2017). Finally, we again used the BAITFILTER program to select among all remaining candidate bait regions for each exon the bait region with the highest sequence coverage in the template alignment. In case of ties, we picked the positional (relative to the start codon) first bait region. If exons were too short for harbouring a full bait region, or if the BLAST filter removed all candidate bait regions for an exon, the exon was not included in the kit, which reduced the number of exons that we targeted. The final bait kit SYRPHIDAE1.0 contained 24,166 baits, targeting 1945 exons belonging to 1312 different OGs (see Data availability).

Baits were synthesized by Agilent Technologies (Waldbronn, Germany) and the SureSelect^{XT2} kit (Agilent Technologies) was used for conducting the hybrid enrichment.

DNA barcoding

For the selected taxa for enrichment, the 5′-end of the mitochondrial cytochrome oxidase *c* subunit I (COI) gene was sequenced for each specimen using Sanger sequencing in order to act as a surrogate voucher (Young, Lemmon, et al., 2016). DNA primers, as well as amplification, purification, sequencing protocols and edition were carried out as described in Gibson et al. (2010) for specimens sequenced at CNC, or as described in Rozo-Lopez and Mengual (2015) for specimens sequenced at ZFMK.

COI sequences from transcriptomes were extracted as follows. The COI sequences generated at CNC for the exon-capture taxa were aligned using MAFFT v7.394 (Katoh & Standley, 2013), and the hmmbuild program from the HMMER software package version 3.1b2 (http://hmmer.org/; Wheeler & Eddy, 2013) was used to build hidden Markov models for the COI sequences. The assembled transcriptomes were searched against the COI-HMM with the nhmmer program and the best hit, that is, the transcript matching best the COI-HMM, was extracted and the COI sequence was excised from this transcript with the information provided by nhmmer. Finally, all COI sequences were aligned using MAFFT v7.394 with the '--globalpair --maxiterate 1000' options, which resulted in an alignment without gaps. The COI alignment was then translated to amino acids to ensure there were no stop codons. The COI data matrix contained a total of 654 nucleotide characters, which we included in our phylogenetic analyses. Sequences were submitted to BOLD and uploaded from there to GenBank (see Table S1 for GenBank Accession Numbers).

Molecular laboratory analyses

Genomic DNA (gDNA) was extracted from a single leg (large specimens) or whole body (small specimens) following the protocol detailed by Bank et al. (2017), with some minor modifications. gDNA was extracted using Qiagen DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany) including the RNase-digestion step and elution in water. We then checked the quantity and quality of the extracted gDNA as in Bank et al. (2017).

We followed the SureSelect^{XT2} Target Enrichment System Protocol for Illumina Paired-End Multiplexed Sequencing during the library preparation. For each sample, we used 100 ng gDNA and fragmented the DNA into 250–350 bp fragments with Bioruptor PICO sonicator (Diagenode s.a., Seraing, Belgium). Samples with less than 100 ng of gDNA were also used with further modifications in the protocol, for example, number of PCR cycles. The results of the fragmentation step were assessed with a Fragment Analyser (Advanced Analytical Technologies GmbH, Heidelberg, Germany) run.

For 'End Repair' and 'A-tailing' we used the chemicals that were included in the Agilent Sure Select^{XT2} Library Prep Kit (Agilent Technologies Inc., Santa Clara, CA, U.S.A.) and followed the same protocol as Bank et al. (2017). A quantity check with Quantus Fluorometer (Promega, Fitchburg, WI, U.S.A.) after 'A-tailing' was necessary to evaluate the adaptor concentration for the next steps. We used NEB kits [NEBNext Quick Ligation Module; NEBNext Multiplex Oligos for Illumina (Dual Index Set1) Kit; NEBNext Q5 HotStart HiFi PCR Master] (NEB, Ipswich, U.S.A.) for adaptor-ligation and library PCR to dual-index our samples, which decreases the chances of cross-contamination (Costello et al., 2018). Adaptor-ligation was done according to the manufacturer protocol and we used a 1:25 dilution of adaptor. Next, adaptor-ligated DNA was dual-indexed through amplification following the PCR program: initial denaturation temperature of 98°C for 30 s, followed by 10 cycles at 98°C for 10 s and at 65°C for 75 s, followed by a 5 min final extension at 65°C and a cool-down to 4°C without end. We purified the amplicons with AMPure XP beads in a ratio of 1:0.9. Then, we checked for quality and quantity with a Fragment Analyser and a Quantus Fluorometer. Later, we used the SureSelect^{XT2} Pre-Capture ILM Module Box 2 Kit and the SureSelect Custom Baits 0.5–2.9 Mb (Agilent Technologies Inc., Santa Clara, CA, U.S.A.) and followed the Agilent protocol for hybridization 'capture library size smaller than 3.0 MB', with a reduction of the volumes by 50%.

For hybridization, we combined eight libraries in equimolar absolute amounts of 93 ng in one pool. The total volume of one pool was

then evaporated to 3.5 μl with a SpeedVac R SPD 111 V (ThermoFisher Scientific, Waltham, MA, U.S.A.), pipetted to 4.5 μl of Sure Select^{XT2} Blocking Mix and incubated under the following conditions: 5 min at 95°C, followed by 5 min at 65°C, and a final step at 65°C without time limit. Then, we prepared the bait mix with 2.5 μl of RNase Block (1:10), 1 μl of Baits and 18.5 μl of SureSelect^{XT2} Hybridization Buffer for each pool; the bait mix was added to the incubated libraries. Subsequently, the mix of libraries and baits were incubated for about 48 h at 65°C in a GeneAmp PCR System 2720. After the hybridization time, the enriched libraries were captured with 50 μl Dynabeads MyOne Streptavidin T1 per pool and were washed using buffers from SureSelect^{XT2} Pre-Captured Box 1 (Agilent Technologies Inc.). The post-capture libraries were then amplified using the following PCR program: initial denaturation temperature of 98°C for 2 min, followed by 12 cycles at 98°C for 30 s, 60°C for 30 s, and 72°C for 60 s, followed by a 10 min final extension at 72°C. We purified the amplicons with AMPure XP beads in a ratio of 1:0.75 to remove oligonucleotide primer dimers. Each of the processed library pools was eluted in 40 μl nuclease-free water and checked for quality and quantity with a Fragment Analyser and a Quantus Fluorometer.

Finally, we sent the finished pools to StarSEQ GmbH (Mainz, Germany) for sequencing on an Illumina Nextseq 500 System (Illumina Inc., San Diego, CA, U.S.A.) with a read length of 150 bp and an estimated output of 6.25 million reads per pool.

De novo assembly of reads and orthology prediction

As mentioned, the final bait kit SYRPHIDAE1.0 targets 1945 exons belonging to 1312 different OGs. In the present phylogenetic study, we excluded the 10 coding loci from previous studies (AATS, PER, CAD1, CAD4/5, CK1, TULP, L35, L113, L133, L272 and L352). Thus, we only used the raw data for the 1302 single-copy OGs.

For the data analysis, we mostly followed Bank et al. (2017). First, we trimmed adapters and low quality bases from paired-end reads with TRIMMOMATIC version 0.36 (Bolger et al., 2014) using the parameters 'LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN: 36', we checked the data quality with FASTQC version 0.11.7 (Andrews, 2010) and de novo-assembled the paired-end reads with IDBA-UD version 1.1.1_320 (Peng et al., 2012) in two steps. Then, we merged paired-end reads with the 'fq2fa --merge --filter" program from the IDBA package and assembled the reads using idba_ud with the options '--mink 20 --step 5 --maxk 120 --num_threads 1'.

The 3145 candidate single-copy OGs were mined with ORTHOGRAPH version 0.6.3 (Petersen et al., 2017) using as a reference the official gene sets of the five reference species named above. The 3145 candidate OGs were specified in an ORTHOGRAPH set file. For the ORTHOGRAPH analysis and reporting step default parameters were specified in the configuration file. In the resulting fasta output files, headers were changed to the HaMStRAD format using *orthograph2hamstrad.pl* included in the ORTHOGRAPH package. Finally, the *summarize_orthograph_results.pl* script was used to combine the sequences of the 1302 OGs of interest (see below) for the different taxa to create fasta files for each OG.

Multiple sequence alignment and data matrices

Each OG was aligned individually with MAFFT v7.394, using the L-INS-i algorithm on the amino acid level. In these alignment files, sequence outliers were identified in two steps as in Misof et al. (2014). In the first step, outliers were identified (879 sequences in 224 OGs). These outliers were removed and re-added to the alignment file using MAFFT with the '–add' option. After this refinement step, 805 sequences in 200 OG alignments remained outliers and were subsequently removed from the alignments. Next, the five reference taxa that were included in the OG files were removed, together with gap-only positions resulting from their removal. We then generated nucleotide alignments corresponding to the amino acid alignments using a modified version of the PAL2NAL software (see Misof et al., 2014).

ALISCORE version 2.0 (Misof & Misof, 2009) was used to search for sequence segments that are ambiguously aligned in the amino acid gene alignments. ALISCORE was run with default parameters except for using the '-e' option to cope with sequence alignments containing many gaps and the '–r 10^27' option to compare all sequence pairs in each sliding window. Sequence segments identified by ALISCORE were removed in the amino acid alignments using ALICUT (Kück et al., 2010) and corresponding segments were removed in the nucleotide alignments. After removing alignment positions, sequences might contain only gaps and ambiguous nucleotides. These sequences were also removed from the gene alignments.

We also searched for genes with no information content with the software MARE version 0.1.2-rc (Misof et al., 2013). One gene (EOG7MDG91) was removed in this step and 1301 OGs plus COI were included, resulting in 1302 genes in the final data set.

The software FASCONCAT (Kück & Meusemann, 2010) was used to create concatenated sequence files as well as partition files based on gene boundaries. Before concatenating the gene alignments, terminal gaps were replaced by Ns and Xs in the nucleotide and amino acid sequence alignments, respectively.

In addition, we generated a restrictive data set, in which only OGs with at least partial sequence information present for all 135 taxa (121 enriched species and 14 transcriptomes) were kept. Only 154 loci (153 OGs plus COI) were included in this restrictive data set. We constructed two data matrices, namely AA-154 and NT-154, for the 154 loci data set with corresponding amino acid and nucleotide gene alignments, respectively. Furthermore, we produced a reduced nucleotide sequence matrix containing the second codon positions only, NT2-154. Similarly for the 1302 loci data set (1301 OGs plus COI), we created three corresponding amino acid and nucleotide alignments, namely AA-1302, NT-1302 and NT2-1302. Altogether, we constructed six matrices.

Phylogenetic analyses

We conducted phylogenetic analyses for all data sets using the concatenated data matrix approach as well as the multispecies coalescent model (Rannala & Yang, 2003). In all the analyses, *P. anthrax* was used to root the inferred trees.

Concatenated data matrix approach

For all six matrices, we used MODELFINDER (Kalyaanamoorthy et al., 2017) implemented in IQ-TREE (Nguyen et al., 2015) to determine an optimal partitioning scheme and evolutionary model, that is, we merged genes into larger partitions if modelling their evolution under one model is favoured by the Bayesian Information Criterion (BIC). Finally the best model was determined for each meta-partition. This was done with IQ-TREE version 2.0-rc1 (Minh et al., 2020) with the options '-spp gene-boundaries-file.nex -m MFP+MERGE -cptime 4000 -rclusterf 10 –safe', which conduct partition merging in ModelFinder using the rcluster algorithm (Lanfear et al., 2014) and allow each partition to have a different substitution model and overall evolutionary rate.

Utilizing the resulting partitioning scheme and selected model, we conducted full tree reconstruction analyses using the maximum likelihood (ML) criterion and, in the same run, computed 10,000 ultrafast bootstrap (UFB) replicates (Hoang et al., 2018; Minh et al., 2013) with the command line options "-p best-scheme-file.nex -B 10000 -bnni -cptime 4000 -pre bootstrap -safe --sampling GENESITE" for all data matrices. For the matrix AA-1302 we also determined (a) the Shimodaira-Hasegawa-like approximate likelihood ratio test (SH-aLRT) (Anisimova et al., 2011; Guindon et al., 2010; Simmons & Kessenich, 2020) to obtain another set of support values with the aid of 1000 replicates, as well as (b) 300 standard nonparametric bootstrap (BS) replicates (Felsenstein, 1985) with the goal to conduct a rogue taxon analysis. We chose the matrix AA-1302 for these analyses to illustrate Figure 1.

The rogue taxon analyses were carried out using the web interface of the ROGUENAROCK software (https://rnr.h-its.org/) (Aberer et al., 2013) on the basis of the 300 nonparametric bootstrap replicates of the AA-1302 as well as the best inferred tree for the AA-1302 data set. The identification of a rogue taxon in ROGUENAROCK is based on the improvement of branch support in the tree once a taxon is excluded. We specified the majority rule consensus (MR) as well as the extended majority rule consensus (MRE) algorithm with a maximum dropout size of 1 and 3 (maximum number of taxa pruned at the same time) when running ROGUENAROCK. Finally, two more tree reconstructions have been conducted using the AA-1302 data set after removing the single taxon identified to be a potential rogue taxon by ROGUENAROCK by again using the same partitioning and model selection result as above. The two analyses were conducted without (AA-1302r) and with the option -symtest-remove-bad (AA-1302rs), which automatically removes data partitions identified by the SYMTEST software (Ababneh et al., 2006; Jermiin et al., 2004) to violate the assumption of a symmetric substitution process with a p value <0.05. Both analyses were repeated 50 times using random starting trees and the tree with the highest likelihood among the 50 replicates was chosen as the best tree. Apart from the symtest option, the following IQ-TREE parameters were specified: -B 1000 -alrt 1000 -bnni -safe.

Multispecies coalescent tree estimation

In the presence of incomplete lineage sorting (ILS), the ML method has been shown to be statistically inconsistent for tree reconstruction using concatenated gene sequences (Roch & Steel, 2015). By contrast, the multispecies coalescent (MSC) method (see, e.g., Edwards, 2009; Degnan & Rosenberg, 2009) as implemented in ASTRAL (Mirarab et al., 2014) is still consistent (Warnow, 2015) under a sequence evolution governed by the MSC model (Knowles & Kubatko, 2010); hence, it can model discordance between gene trees and species trees caused by ILS (Liu et al., 2015). Therefore, we also used ASTRAL-III (Zhang et al., 2018) to reconstruct species trees from gene trees that have been inferred using ML.

We computed gene trees for all gene alignments of the AA-154, AA-1302, NT-154, NT-1302, NT2-154 and NT2-1302 data sets using IQ-TREE version 1.6.8 with the following command line parameters: '-m MFP -seed 1 -nt AUTO -AICc -bb 1000 -bnni -wbt'. Then, ASTRAL-III version 5.6.3 was used to infer species trees from the gene trees of the six mentioned data sets with default parameters, except the Java program parameter '-Xmx40000M' to provide ASTRAL sufficient RAM. As suggested by several authors (Bossert et al., 2021; Mirarab, 2019; Zhang et al., 2018), we collapsed nodes in each gene tree if their ultrafast bootstrap support value was less than 10 before ASTRAL computes quartet-based support (QBS) (Sayyari & Mirarab, 2016). For comparison, we also collapsed nodes in gene trees for which support was less than 70. Furthermore, we conducted a multi-locus bootstrap (MLB) analysis (Seo, 2008) with 500 replicates for the AA-154, AA-1302, NT-154 and NT-1302 data sets using the '-r 500' option in ASTRAL.

Finally, inferred phylogenies were visualized in Archaeopteryx 0.9928 beta (Han & Zmasek, 2009) and FIGTREE v1.4.4 (Rambaut, 2009), and edited with ADOBE® Illustrator CS5.

Data set exploration

Using the SYMTEST software version 2.0.47 (https://github.com/ottmi/symtest) (Ababneh et al., 2006; Jermiin et al., 2004) we performed pairwise comparisons using Bowker's matched-pairs tests of symmetry (Bowker, 1948) for each gene and for each data set to assess whether they contain sequences that violate the assumption of global stationary, reversibility, and homogeneity (SRH conditions) (see Supplemental Data and Supplementary files for the generated heat maps based on the inferred p-values, using default window and step sizes). We also performed the same pairwise comparisons using SYMTEST for the six data sets without transcriptomic data.

Divergence time estimation

Although more than 100 described flower fly fossils exist (Evenhuis, 1994; Nidergas et al., 2018; Hadrava et al., 2020; Ngô-Muller & Nel, 2020), less than one third belong to Syrphinae and Pipizinae. Due to the lack of a recent taxonomic revision of syrphid fossils, we only included the oldest fossils that can be assigned to a genus or family without doubt based on the original descriptions. For time calibration of the tree we used six fossil calibration points (one for Platypezidae, one for Pipunculidae and four for Syrphidae) following the best-practice recommendations by Parham et al. (2012). *Prosyrphus thompsoni* Grimaldi is the oldest described fossil (minimum age of 98.17 Ma) that belongs to the stem group of Syrphidae (Grimaldi, 2018). We place this fossil at the root of the family to calibrate the crown syrphids (see Table S2 for more information).

To reduce computational burden, amino acid alignments of the 1302 genes (AA-1302 data set) were merged into a smaller number of partitions before BEAST analysis (Bouckaert et al., 2014, 2019). The best partitioning scheme was selected using MODELFINDER as implemented in IQ-TREE 2.1, using the relaxed hierarchical cluster algorithm examining only the top 10% of partitioning schemes. This reduced the number of partitions from 1302 to 75. The XML file for the BEAST analysis was created with BEAUTI 2.6.6 (Bouckaert et al., 2019). Clock models and trees were linked for all partitions, and the substitution model JTT with four gamma site categories was used. As a starting tree, we used the tree obtained by IQ-TREE. Branch lengths were multiplied by 400 to fit calibration ages and tree node heights were automatically adjusted to make the tree ultrametric. All tree rearrangement operators were set to zero to fix the tree topology during the analysis. A calibrated Yule model was used for the tree prior. For the root, a uniform prior with a minimum of 98.17 Ma (age of oldest known stem syrphid fossil) and a maximum of 180 Ma was used; for all other calibrated nodes we assumed log-normal priors with the age of the oldest known fossil as offset (see Table S2 for dates). The maximum of 180 Ma was based on the dated analysis by Wiegmann et al. (2011, fig. S3), who considered *Protoreogeton admirabilis* Mostovski (from the Jurassic of Mongolia, 171.6 to 164.7 Ma) as a crown member of Eremoneura (see also Wiegmann et al., 2003 and Supplementary Material).

We assumed a mean of 3 and standard deviation of 1.25, which produces a distribution of divergence dates that is not overly narrow but within the realm of plausibility. The MCMC analysis was run with BEAST 2.6.4 for 10 million iterations, and trees and parameters were saved every 1000 generations. After the run, the model parameters, likelihood and node height values were checked for convergence with TRACER 1.7.2 (Rambaut et al., 2018) and the first 50% of the Markov chain was discarded as burn-in. A maximum clade credibility tree was then calculated with TREEANNOTATOR 2.6.4 (Bouckaert et al., 2019). Finally, we plotted the dated phylogeny with geological timescale using the packages *phyloch* (Heibl, 2008) and *strap* (Bell & Lloyd, 2015) in R (R Core Team, 2021).

Another set of calibrations was performed with MCMCTREE (Yang & Rannala, 2006) as implemented in PAML 4.8 (Yang, 2007). As log-normal priors are not implemented in MCMCTREE, we used skew-t priors that give a similar distribution. We used the minimum ages of the fossils as location parameters and, to make an age smaller than that of the fossil very unlikely, we chose a shape parameter of 100,000. To replicate the log-normal distribution as closely as possible, we chose parameters of scale and degrees of freedom that gave the same mean and variance as our log-normal distribution according to the formulas given in Arellano-Valle and Azzalini (2013). With 100 Ma as the time unit, as recommended in the MCMCTREE tutorial (dos Reis et al., 2017), this resulted in a scale parameter of 0.3453 and degrees of freedom of 2.298. MCMCTREE was then run for one million generations, sampling every 50 generations, with a burn-in of 100,000 generations.

Determining enrichment efficiency

We determined the efficiency of our kit to enrich the OGs of interest by mapping the quality trimmed reads back onto the OGs. OG sequences have undergone multiple filtering steps removing sequence segments that would hamper the mapping. Thus, in order to be able to achieve an unbiased result, we mapped the reads onto the OGs obtained from ORTHOGRAPH after outlier detection, but before Aliscore and minimum coverage filter steps. We mapped the read onto the OGs with BWA-MEM2 version 2.1 (Vasimuddin et al., 2019) with default parameters. Resulting SAM files were sorted, converted to the bam format and the read coverage was estimated with samtools version 1.10 (Li et al., 2009). Mean species coverage values were obtained by computing a weighted mean of all OGs sequence coverage values, using the lengths of the CDS regions as weights. From the BWA output files we also determined the number of reads that mapped or did not map to the target region, that is, the proportion of reads that are on or off target. The mapping and coverage estimation was implemented as a Snakemake workflow using SNAKEMAKE version 5.323.0 (Köster & Rahmann, 2012).

Data availability

Specimen collection data, systematics, identification method and unique identifiers can be obtained from Table S1 (Supplementary files). Voucher specimens are located at the Royal Museum for Central Africa (KMMA), the Canadian National Collection of Insects, Arachnids and Nematodes (CNC), and the Zoological Research Museum Alexander Koenig (ZFMK).

The Supplemental Data are available at the Mendeley Data repository (http://doi.org/10.17632/ynghdr2msn.1; data contain the 1Kite_Mecopterida sqlite database, the config files for the ORTHOGRAPH runs, the SYRPHIDA1.0 bait kit, analyses scripts, assembled DNA contig sequences, finalized alignments for each data set and for individual genes, obtained phylogenetic trees, and individual gene trees for the MSC analyses). Unprocessed sequence data have been deposited in the NCBI Sequence Read Archive under BioProject PRJNA837085 (http://www.ncbi.nlm.nih.gov/bioproject/837085).

See the Supplemental Data and Supplementary files under the Supporting Information.

RESULTS

Data capture

We targeted 1312 genes, with 1945 CDS regions, with the bait kit SYRPHIDAE1.0 and we successfully enriched all targeted genes. Due to difficulties with mapping some of them onto the unpublished draft genome of the drone fly *E. tenax*, we did not include in our phylogenetic analyses the three protein-coding genes obtained from the pre-omic studies (AATS, PER, CAD1 and CAD4/5) and the seven loci selected from Young et al. (Young, Lemmon, et al., 2016; Young, Marshall, & Skevington, 2016), that is, L35, CK1, TULP, L113, L133, L272 and L352. The OG EOG7MDG91 was removed by MARE, since its information content was zero. Thus, we based our molecular inference on 1302 genes (1301 OGs + COI).

Targeted loci were successfully enriched in all flower fly species ($n = 121$), except for two taxa that were excluded after DNA extraction due to low DNA quality (Table S1). The number of enriched genes per species varied from 845 (64.90% for *Microdon devius* (Linnaeus)) to 1284 (98.62% for *Afrosyrphus schmuttereri* Mengual et al.) (Table S3). The mean length of the targeted gene was 438.43 bp (SD = 200.28 bp; median = 412 bp) excluding the transcriptomic data (Table S5) and 1,211,632 reads per species on average were obtained, assembled by IDBA-UD in 8795 contigs on average, and these had a mean contig length of 666 bp (SD = 106.21 bp) (Table S4). For additional sequencing and assembly information, see Supplementary information.

Alignment statistics

Final alignments for each gene and for each data set can be found in Supplementary information. The complete data set used for phylogenetic analyses contained 135 taxa, 1302 loci and 268,411 amino acids (AA-1302) and 805,233 nucleotides (NT-1302). The restrictive data set comprised 135 taxa, 154 loci and 36,629 amino acids (AA-154), and 109,887 nucleotides (NT-154). The amount of completeness was 70.9% (29.1% of missing data) for the complete data set (1302 loci) and 77.37% (22.63% of missing data) for the restrictive data set (154 loci). The total amount of completeness in our data sets is similar to those data sets of Dietz et al. (2019) (average 75.5%), Li et al. (2021) (60.45% in the 216 loci data set), Quek et al. (2020) (67.57%–69.14%), or Mayer et al. (2021) (AA: 59%, AA_red: 74%).

Phylogenetic analyses

Our concatenated and multispecies coalescence analyses recovered almost identical topologies for Syrphidae, with a few incongruences within defined clusters. For all 14 major phylogenetic analyses (data sets AA-154, AA-1302, AA-1302r, AA-1302rs, NT-154, NT-1302, NT2-154 and NT2-1302 using ML and MSC) highly congruent topologies were inferred with high support values for most nodes (see Figures S1–S14). To illustrate our results we follow the suggestion of Cohen et al. (2021) and use the inferred topology from the largest data set (Figure 2). Microdontinae was resolved as sister to all other flower flies within a monophyletic Syrphidae, whereas Pipizinae and Syrphinae (both resolved as very well-supported clades sister to each other) were nested inside a paraphyletic Eristalinae (Figure 2). Three major lineages with tribal rank were resolved within Syrphinae: Melanostomini, Bacchini and Syrphini, the latter including the two monotypic tribes Paragini and Toxomerini. Moreover, the large cluster of Syrphini (with Paragini and Toxomerini) is divided into two clades with maximum support (Figure 2) or very high support (Figures S1–S13). The exception is the ASTRAL analyses with the NT2-1302 and NT2-154 data sets, where these two Syrphini clades have a moderate to low support (Figures S10 and S13).

Despite the very congruent topologies for the different data sets, the branch support values (Table 1) indicate that more nodes within Syrphidae (outgroups excluded) have high support (UFB ≥90) for the complete data sets (1302 loci) compared with the equivalent restrictive data sets (154 loci), either using the concatenated or the MSC approach. Within each group of data sets (1302 loci and 154 loci), the nucleotide data set (NT) has more branches with high support than the amino acid (AA) or second codon position (NT2) data sets, again independently of the approach used. In the concatenated approach, the highest global UFB support was found in the NT-1302 data set, but the highest percentage of highly supported (UFB ≥90) branches and the lowest number of branches with moderate to low support (UFB <90) was recovered in the AA-1302r data set, the original AA-1302 data set without the rogue taxon (Table 1).

One taxon was indicated to be rogue, *Melangyna* (*Melangyna*) *lasiophthalma* (Zetterstedt) (Figure 1e). Its phylogenetic placement varies to a notable degree between analyses (Figures S1–S14). The removal of this taxon (AA-1302r) increased the stability and support values in one of the two major Syrphini clades, where the species was resolved (Table 1; compare Figures S1 and S7). When we removed data partitions that violate the assumption of a symmetric substitution process (identified by the option-symtest-remove-bad) (AA-1302rs), support values also increased but to a minor degree, in general (Figure S8).

In the MSC analyses, no major differences were found when we collapsed branches in gene trees for which support was less than 10 or less than 70. Final QBS values calculated in ASTRAL were similar and inferred topologies were highly congruent. As in other comparisons, minor incongruences were found within defined clusters due to low branch support. Thus, all our comments refer to ASTRAL phylogenies based on gene trees where we collapsed branches with bootstrap support less than 10 (Figures S9–S14).

The three support measures (SH-aLRT, UFB and BS) used in the IQ-TREE analysis using the AA-1302 data set identified the same branches with moderate to low support (Figure S1). Most discrepancies

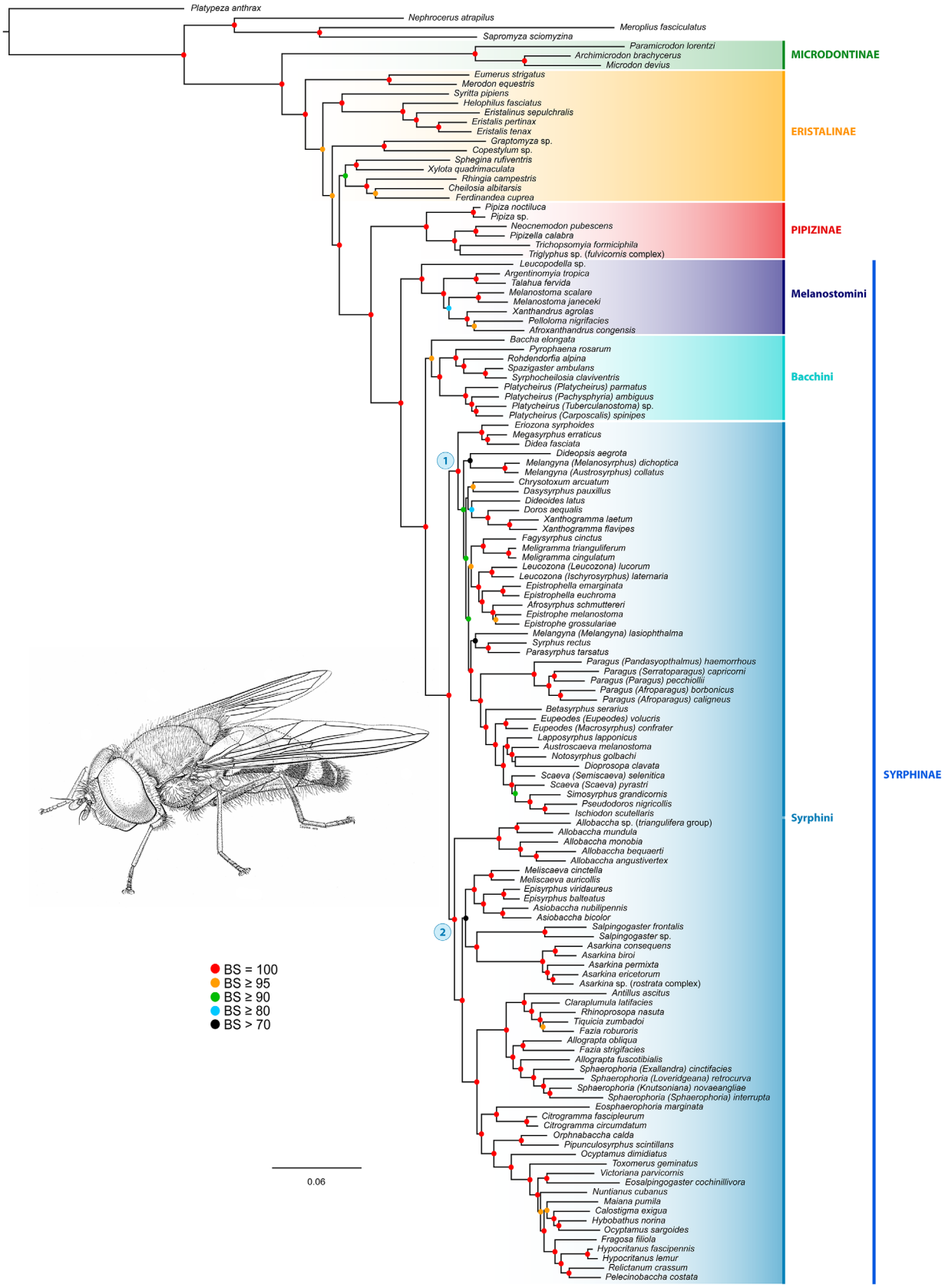

FIGURE 2 ML phylogeny of Syrphidae inferred in IQ-TREE using the AA-1302 concatenated matrix. Nonparametric bootstrap values are depicted on the branches. Numbers 1 and 2 within a blue circle indicate the two major clades within Syrphini. Line drawing adapted from Vockeroth and Thompson (1987): *Syrphus torvus* Osten Sacken, 1875.

TABLE 1 Support values for the data sets under different approaches within Syrphidae, excluding the non-syrphid taxa

Concatenated supermatrix	Global UFB support	Percentage of nodes with high support (UFB ≥90)	Number of nodes with UFB <90
AA-1302	95.26	86.92	17/130
NT-1302	**98.49**	93.08	9/130
NT2-1302	94.02	85.38	19/130
AA-154	89.76	73.85	34/130
NT-154	94.18	88.46	15/130
NT2-154	84.8	68.46	41/130
AA-1302r	97.38	**94.62**	**7/130**
AA-1302rs	98.38	91.54	11/130
Multispecies coalescence	**Global MLB support**	**Percentage of nodes with high support (MLB ≥90)**	**Number of nodes with MLB <90**
AA-1302	94.57	91.54	11/130
NT-1302	**97.44**	**95.38**	**6/130**
NT2-1302	–	–	–
AA-154	87.68	72.31	36/130
NT-154	96.22	90.77	12/130
NT2-154	–	–	–

Note: Global bootstrap support (GBS), as defined by Buenaventura et al. (2021), was calculated by averaging all ultrafast bootstrap (UFB) support values on a given tree for each concatenated data set, or averaging all multi-locus bootstrap (MLB) support values in the case of MSC analyses. Bold values indicate the highest percentage in the first two columns, and the lowest number of nodes with UFB<90 in the third column.

between inferred phylogenies using different approaches and data sets are found around branches with moderate to low support (UFB <90 and MLB <90) and include the placement of the same taxa. Discrepancies related to the position of a single taxon, usually in a derived clade, leaving the composition of the clade otherwise identical, were named minor incongruences. We find minor incongruences when the placement variation of a certain taxon is limited to a very inclusive clade. For instance, the sisterhood of *Triglyphus* Loew and *Trichopsomyia* Williston, the placement of *Maiana pumila* (Austen) within the large radiation of genera previously placed under *Ocyptamus*, or the relationships within the clade comprising *Dioprosopa* Hull, *Austroscaeva* Láska, Mazánek & Mengual and *Notosyrphus* Vockeroth. Notable incongruences consist of the placement of a taxon changing the composition of a clade, or related to the relationship between larger clades. In our 14 inferred phylogenies notable incongruences are the placement of *Baccha* Fabricius (Figures S4, S8, S12 and S14) or the sisterhood of two well-supported clades within Syrphini, one including *Meliscaeva* Frey, *Asiobaccha* Violovitsch and *Episyrphus* Matsumura & Adachi, and the other cluster formed by *Asarkina* and *Salpingogaster* Schiner (Figures S6, S7, S8, S9, and S11). The alternative placement of *Baccha* and the clade comprising *Asarkina* and *Salpingogaster* (when compared with their placements in Figure 2) were always recovered with low support (UBF <75 or MLB <55) and using reduced data sets, either using AA-1302rs, NT2 or 154 loci data sets.

Compositional heterogeneity test

SYMTEST results indicate a remarkably different behaviour of the data sets either based on AAs, all nucleotides or second codon positions only (Figures S15–S26). Amino acids (AA) and second codon positions (NT2) data sets have a similar large percentage of pairs of sequences that seem to not violate the SRH conditions, implying that the sequences are compositionally more homogeneous compared to the nucleotide data sets (NT-1302 and NT-154), which exhibit a presence of a large proportion of sequences with compositional heterogeneity.

The removal of transcriptomic data produced heat maps with larger proportions of sequences that seem to not violate the SRH conditions. Restrictive data sets (154 loci) are more compositionally homogeneous than their respective complete data sets (1302 loci) (compare Figures S15 and S21).

Per gene the SRH test did not indicate any violations of the SRH conditions.

Divergence time estimation

BEAST and MCMCTREE analyses differed in the inferred dates for the phylogeny based on the AA-1302 data set (Table 2; Figures S27 and S28). Following Warnock et al. (2017) we refer to the inferred 95% highest posterior density (HPD) intervals instead of means or medians. Most of the inferred 95% HPD intervals by MCMCTREE do not overlap with the 95% HPD intervals provided by BEAST and they are usually narrower than the credibility intervals inferred by BEAST. It looks like the Bayesian inference using MCMCTREE (Figure S28; Table 2) pushed all dates towards the maximum age of the root (set at 180 Ma) and we will not discuss this analysis further.

The BEAST inference dated the origin of the Syrphidae in the Lower Cretaceous, between the Aptian and Albian Ages (95%

TABLE 2 Inferred 95% highest posterior density (HPD) intervals and median (in Ma) for the BEAST and MCMCTREE analyses, using six calibration points (Table S2)

	BEAST		MCMCTREE	
	Median	**95% HPD**	**Median**	**95% HPD**
Syrphidae origin	105.3	125.5–98.5	176.6	179.6–169.5
Diversification of Syrphidae	80.2	95.7–73.1	130.8	133.1–125.6
Pipizinae–Syrphinae clade origin	59.1	70.6–54.4	120.0	122.1–115.1
Diversification of Pipizinae	42.9	51.4–39.8	59.5	60.9–57.0
Diversification of Syrphinae	55.1	65.8–50.4	112.6	114.6–108.1
Diversification of Melanostomini	51.2	61.0–46.9	112.6	114.6–108.1
Bacchini–Syrphini clade origin	50.7	60.4–45.6	92.8	94.4–89.1
Diversification of Bacchini	44.2	52.7–40.4	92.7	94.4–89.1
Diversification of Syrphini	42.6	50.9–39.4	70.8	72.0–68.0
Diversification of *Allobaccha*	32.9	39.4–29.9	34.1	35.0–32.7
Diversification of *Asarkina*	17.2	20.6–15.8	22.8	23.4–21.8

HPD = 125.5–98.5 Ma). Pipizinae and Syrphinae originated more recently, between the Maastrichtian (Upper Cretaceous) and the Ypresian (Eocene) Ages around the K–Pg boundary (95% HPD = 70.61–54.4 Ma). Melanostomini and Bacchini branched off very soon after the origin of the Syrphinae during the Palaeocene and Early Eocene (Table 2).

DISCUSSION

Systematics

Our reconstruction of the subfamilial and tribal classification of flower flies is in agreement with previous studies. The big-headed fly species (Pipunculidae) is placed as sister to Schizophora, rendering the old concept of Syrphoidea obsolete, in agreement with previous molecular studies (Bayless et al., 2021; Pauli et al., 2018; Wiegmann et al., 2011; Young, Lemmon, et al., 2016). Microdontinae, also known as ant flies, is resolved as monophyletic and sister to all other flower flies. The inferred divergence time in the BEAST analysis places the origin of this lineage in the Upper Cretaceous (95.7–73.1 Ma). The phylogenetic position of microdontines, in concordance with previous works (Mengual et al., 2015; Pauli et al., 2018; Rotheray & Gilbert, 2008; Skevington & Yeates, 2000; Ståhls et al., 2003; Thompson, 1969; Young, Lemmon, et al., 2016) continues the discussion about its systematic rank. Some authors argue that their morphological and ecological specifications deserve a family status (Speight, 1987, 2020; Thompson, 1972, 2020), whereas most researchers consider them a subfamily of Syrphidae (Reemer & Ståhls, 2013; Skevington et al., 2019; Ssymank et al., 2021). The systematic rank of the microdontines seems a matter of personal choice (Reemer & Ståhls, 2013). If the family rank is preferred for the microdontines, our suggestion is to group the Microdontidae with the flower flies under an old name but with a new concept, the superfamily Syrphoidea.

Accumulated evidence does not sustain the current, outdated concept of Eristalinae based mostly on adult morphology. Phylogenetic studies during last decades using molecular and morphological data recover a non-monophyletic Eristalinae (Mengual et al., 2015; Moran et al., 2022; Mullens et al., 2022; Pauli et al., 2018; Skevington & Yeates, 2000; Ståhls et al., 2003; Young, Lemmon, et al., 2016). Our results echo previous results, where several lineages branch off before the Syrphinae-Pipizinae split and the first is that comprising the genera *Eumerus* Meigen and *Merodon* Meigen (see Moran et al., 2022; Pauli et al., 2018; Young, Lemmon, et al., 2016). The systematics of these flower flies deserves further study and a new data set using the bait kit SYRPHIDAE1.0 is being assembled (Moran et al., in prep.)

The present study corroborates the phylogenetic placement of Pipizinae as the sister group to Syrphinae (Mengual et al., 2015; Moran et al., 2022; Pauli et al., 2018; Rotheray & Gilbert, 2008; Ståhls et al., 2003; Young, Lemmon, et al., 2016). Pipizinae (Figure 1a) is the smallest subfamily of flower flies (Mengual et al., 2015; Vujić et al., 2013) and our BEAST analysis inferred a divergence from its sister clade Syrphinae around the K–Pg boundary (70.61–54.4 Ma). A minor discrepancy among the different data sets and approaches is due to the very low support of the cluster comprising *Trichopsomyia* and *Triglyphus* (Figures 2 and S1–S14). We lack three pipizine genera in our analyses, namely *Cryptopipiza* Mutin, *Heringia* Rondani and *Claussenia* Vujić & Ståhls in Vujić et al., which might bring instability to this particular branch. All previous studies using molecular data resolved the pipizine relationships with low support (Mengual et al., 2015; Vujić et al., 2013); thus, generic relationships within the monophyletic Pipizinae remain unclear.

The phylogenetic placement and monophyly of Syrphinae are unambiguously supported in our analyses. This subfamily is further divided into four well-supported clades: Melanostomini, Bacchini and two large clades within Syrphini (including the monotypic Toxomerini and Paragini). Our inferred phylogeny for these predatory flower flies

(Figures 2 and S1–S14) agrees with recent phylogenetic studies recognizing the tribes Melanostomini and Bacchini (Mengual, 2015, 2020; Mengual et al., 2021; Mengual, Ståhls, et al., 2018; Moran et al., 2022; Pauli et al., 2018; Ståhls et al., 2003; Young, Lemmon, et al., 2016). Our findings parallel those of Mengual (2020) resolving *Leucopodella* Hull as sister to the remaining melanostomines, placing *Pelloloma* Vockeroth within Melanostomini, and indicating a close relationship between *Argentinomyia* Lynch Arribálzaga, *Talahua* Fluke and *Melanostoma* Schiner. Furthermore, our results support the generic status of *Afroxanthandrus* Kassebeer, in agreement with Mengual (2020) but contradicting Goeldlin de Tiefenau and Thompson (2019).

The placement of *Baccha* (Figure 1c) as an independent lineage, not as the sister group to all other bacchines, is recovered in four analyses always with low support: AA-154 data set with the ML criterion (UFB = 74; Figure S4); AA-1302rs data set with the ML criterion (SH-aLRT = 62.6, UFB = 73; Figure S8); AA-154 data set with the MSC method (QBS = 0 and MLB = 33 for Bacchini without *Baccha*; Figure S12); and NT2-154 using ASTRAL (QBS = 0.75 for Bacchini without *Baccha*; Figure S14). We assume that these exceptions (with significant low support) to the general result of *Baccha* as sister to all other bacchines may be due to the reduced data sets and that the complete data set (1302 loci) is needed to recover this placement. We have recovered slightly different relationships among the other bacchines compared to Mengual (2020), with *Pyrophaena* Schiner being sister to (*Rohdendorfia* Smirnov + (*Spazigaster* Rondani + *Syrphocheilosia* Stackelberg)) with all branches having a nonparametric bootstrap value of 100 (Figure 2). No species of *Eocheilosia* Hull was available for the present study.

In its current concept Syrphini is paraphyletic as it does not include the genera *Toxomerus* and *Paragus*. Thus, we define the tribe Syrphini **stat.rev.** to include the species of *Toxomerus* and *Paragus*, and consider Toxomerini and Paragini no longer valid tribes. These two monotypic taxa have been resolved within the large radiation of Syrphini in the recent literature (Mengual, 2015, 2020; Mengual et al., 2008a, 2008b, 2012, 2015, 2021; Mengual, Ståhls, et al., 2018; Mengual & Thompson, 2011; Moran et al., 2022; Young, Lemmon, et al., 2016). *Toxomerus* is resolved within a large radiation of American genera (mostly Neotropical species but some Nearctic), whereas *Paragus* is placed as sister to the *Eupeodes-Scaeva* radiation (sensu Mengual, Ståhls, et al., 2018).

Inside the hypandrium (male genitalia), flower flies have the aedeagus fused to the parameral sheath forming the phallus, which can be moved thanks to the phallapodeme (Cumming & Wood, 2017). Based on our new definition, members of Syrphini have the phallus (named aedeagus in older literature) divided into basiphallus and distiphallus, whereas taxa in Bacchini and Melanostomini have a simple, unsegmented phallus. Other adult morphological characters used in the past, such as face and scutellum entirely black for Bacchini and Melanostomini, do not always apply. This definition of Syrphini makes necessary a revision of the male genitalia of *Paragus* from an anatomical point of view. Vockeroth (1969) stated that *Paragus* males have the phallus undivided, with a short, rather enlarged, tubular median portion with a complex lateral lobe arising at the base of either side. This argument was accepted by subsequent authors but never tested (Tot et al., 2020; Vujić et al., 2008). The origin of the lateral lobes and their homologous structures in other Syrphini need further study. As mentioned above, the phallapodeme assists in moving the base of the phallus and the postgonites (Cumming et al., 1995; Cumming & Wood, 2017). If the so-called lateral lobes of the aedeagus sensu Vujić et al. (2008) are basally in contact with the phallapodeme, then they should be part of the phallus, that is, possibly the basiphallus. At the same time, the so-called aedeagus sensu Vujić et al. (2008) (equivalent to the median portion of the aedeagus sensu; Vockeroth, 1969) could be homologous with the distiphallus in other syrphines. In other words, if we can consider the lateral lobes as part of the phallus, consequently the phallus in *Paragus* is divided into two structures as well.

Two major clades within Syrphini receive high support in most analyses, mirroring previous studies based on molecular and morphological data (Mengual, 2015, 2020; Mengual et al., 2008b, 2015, 2021; Moran et al., 2022; Young, Lemmon, et al., 2016). In Figure 2, these two clusters are indicated with numbers 1 and 2 within a blue circle, and the word Syrphini is located between the two clusters to facilitate their visualization. No particular morphological characteristic in adults or larvae is known to diagnose these two clades; thus, we decline to give names until we can characterize them based on morphological features. One of these major clades (number 2) has *Allobaccha* as sister to a group of genera, including *Asarkina* and the *Sphaerophoria* and *Ocyptamus* lineages (sensu Mengual et al., 2021), among others. Within this major grouping (number 2), inferred relationships are similar to those recovered by Mengual et al. (2021) with the Indomalayan and Australian *Eosphaerophoria* Frey and *Citrogramma* Vockeroth sister to the American *Ocyptamus* lineage (represented by 14 genera in our analyses; see Figure 1m for one representative of the lineage). Our results corroborate the taxonomic status of many genera previously placed under *Ocyptamus*, confirm the *Incertae sedis* placement of *Syrphus sargoides* Macquart (Mengual, Miranda, & Thompson, 2018; Miranda et al., 2014, 2020), and support the relationships inferred by Mengual et al. (2021) for the *Sphaerophoria* lineage, with the genera *Fazia* Shannon and *Allograpta* Osten Sacken polyphyletic. Minor differences with previous hypotheses are the relationships between the *Sphaerophoria* subgenera and the sisterhood of *Tiquicia* Thompson and *Fazia roburoris* (Fluke).

A major incongruence between the ML and MSC analyses for this Syrphini clade number 2 is the branch clustering the monophyletic group of three genera, namely *Meliscaeva*, *Episyrphus* and *Asiobaccha*, and the clade with *Asarkina* and *Salpingogaster* Schiner (Figure 2). This particular branch receives a large disparity of support values, but the alternative topology with the two clades branching separately also receives uneven support values (Figures S1–S14). The ML and MSC analyses with the AA-1302 data set recover the branch with very low support (Figures S1 and S9), but the ML and MSC analyses using the NT-1302 matrix resolved it with high support. From the morphological and biogeographical points of view, the sisterhood of *Asarkina* and *Salpingogaster* is interesting. There is no obvious morphological character that may suggest their sisterhood and the general habitus of each is quite distinct: *Salpingogaster*, endemic of the New World, has a

petiolate abdomen and a metafemur with rows of spinose setae on the apical half (Figure 1j), whereas *Asarkina* (Figure 1k,l) has a very broad and flat abdomen with unarmored hind legs and is found mostly in the Afrotropical, Indomalayan and Australian realms, with one species reaching the Russian Far East. Although we did not sample the South African subgenus *Asarkina* (*Achoanus*) Munro, we did include two Afrotropical species and two Indomalayan *Asarkina*. The four taxa were resolved together, including *Asarkina biroi* Bezzi with its black face (an earlier diagnostic character of the tribes Bacchini and Melanostomini). Published phylogenies using a small number of molecular markers obtained with Sanger sequencing always recovered *Salpingogaster* as sister to at least one member of the *Ocyptamus* and/or *Sphaerophoria* lineages when *Asarkina* was not sampled in those studies (Mengual, 2020; Mengual et al., 2015; Mengual, Ståhls, et al., 2018; Moran et al., 2022); conversely, *Asarkina* was always recovered as sister to the clade comprising *Meliscaeva*, *Episyrphus* and *Asiobaccha* when no *Salpingogaster* species was included in the survey (Mengual et al., 2008a, 2008b). Nevertheless, other previously inferred phylogenies that sampled both genera placed *Asarkina* as sister to the cluster with *Meliscaeva*, *Episyrphus* and *Asiobaccha*, and at the same time, placed *Sapingogaster* near the *Ocyptamus* and *Sphaerophoria* lineages (Mengual, 2015; Mengual et al., 2021). The phylogenetic relationships of *Salpingogaster* and *Asarkina* require further studies where more species should be sampled.

Allobaccha (Figures 1g–i) was resolved as the sister group to the rest of the members of the Syrphini clade number 2, in agreement with the latest inferred ML phylogenies (Mengual, 2015; Mengual et al., 2021). Published cladograms inferred using parsimony as optimality criterion grouped *Allobaccha* and *Paragus* together (Mengual, 2015; Mengual et al., 2008a, 2008b), although Mengual et al. (2015) pointed out that this could be the effect of a long-branch attraction. We sampled an undescribed species belonging to a recently described subgenus, *Allobaccha* (*Petioleomyia*) Thompson, here identified as *Allobaccha* sp. (*triangulifera* group). As noted by Thompson (2017), this group of species differs from other *Allobaccha* species in several morphological characters, but our inferred topologies placed this taxon within the clade of *Allobaccha*, together with two African species and two other taxa from Australia and New Guinea. Among the sampled *Allobaccha* species, we have included a several morphologically confounding species. *Allobaccha mundula* (Wulp) possesses a pilose postpronotum, a diagnostic character of the subfamily Eristalinae not present in any other syrphine. We also sampled one member of the Afrotropical *picta* species group (similar to Figure 1g), that is, *Allobaccha angustivertex* Enderlein, which differs remarkably from other *Allobaccha* species by their largely infuscated, broad wings. In our results, all *Allobaccha* species were grouped together, and the recognition of the subgenus *Petioleomyia* would require division of the genus into many other subgenera, which would need to be defined morphologically. At present, our results do not justify the erection of the subgenus *Petioleomyia*. Consequently, we synonymize *Petioleomyia* **syn.n.** under *Allobaccha*.

This mentioned major Syrphini clade (number 2) is vital to understand the evolution of the larval biology in flower flies. The prey diversity of the predatory members of this clade is peerless among flower flies (Mengual et al., 2012; Mengual, Miranda, & Thompson, 2018; Rotheray et al., 2000; Ureña & Hanson, 2010). Moreover, all the known phytophagous and pollen-feeding larvae within Syrphinae are found in this clade (Mengual et al., 2008a, 2012, 2021). It seems that among the two large American radiations within the *Sphaerophoria* and *Ocyptamus* lineages, some species had a secondary transition to phytophagy in their larval feeding mode, and this happened at least three times independently in different genera (Dumbardon-Martial, 2016; Nishida et al., 2002; Reemer & Rotheray, 2009; Weng & Rotheray, 2009; Zuijen & Nishida, 2011).

The other major clade within Syrphini (number 1) inferred in all our analyses comprises a large number of genera as well. Among them, it contains the *Eupeodes-Scaeva* lineage (as defined in Mengual, Ståhls, et al., 2018) with a well-supported Neotropical radiation comprising *Dioprosopa*, *Notosyrphus* and *Austroscaeva* (Figure 1n). The phylogenetic relationships among these three Neotropical genera are not resolved as the branch support is low (BS <40; Figures 2 and S1), but *Lapposyrphus* Dušek & Láska was resolved as the sister taxon to this Neotropical cluster with high branch support (BS = 100; UFB = 100; Figures 2 and S1). Our results corroborate the results by Mengual, Ståhls, et al. (2018) and support *Simosyrphus* Bigot, *Ischiodon* Sack, *Dioprosopa* and *Pseudodoros* Becker as being valid genera. All the subgenera of *Paragus* were grouped together, as sister to the *Eupeodes-Scaeva* lineage, mirroring previously inferred phylogenies using ML as optimality criterion (Mengual, 2015; Mengual et al., 2021; Young, Lemmon, et al., 2016; see above comments on *Allobaccha*).

A clade with three genera is resolved as sister to all other genera placed in this major Syrphini clade (number 1), namely *Eriozona* Schiner, *Megasyrphus* Dušek & Láska and *Didea* Macquart. The phylogenetic position and relationship of these three genera is congruent with previous studies (Mengual, 2015; Mengual et al., 2008b; Mengual, Ståhls, et al., 2018; Rotheray & Gilbert, 1989). Another major discrepancy among analyses is found in this major clade due to the low support for the deeper branches in the clade. Three consecutive branches have moderate to low support between the clade of (*Eriozona* + (*Megasyrphus* + *Didea*)) and the branch where *Paragus* and the *Eupeodes-Scaeva* lineage are sister taxa. Several factors may explain this lack of support, for example, the short length of the branches, which might indicate a rapid radiation, the presence of rogue taxa as indicated before, or missing genera that would help resolve the branching pattern within this clade.

Other taxa with changing phylogenetic placement among data sets and approaches are *Dideoides* Brunetti and *Dideopsis* Matsumura (Figure 1d). The concatenated ML analyses with the AA-1302 (Figure 2) and AA-1302r (Figure S7) data sets resolved *Dideoides* sister to (*Doros* Meigen + *Xanthogramma* Schiner) and placed *Dideopsis* sister to the Australasian *Melangyna* species. Vockeroth (1969) stated that *Dideoides* is a morphologically well-defined genus without evident close relatives, and Láska et al. (2000) concluded that the puparium of *Dideopsis* is not morphologically similar to any other known puparia. The placement of these two commonly sampled but enigmatic genera

is a novelty, which will help to investigate further relationships with the above-mentioned taxa.

Within this major clade of Syrphini (number 1) there are relationships found in all our analyses using different data sets and under concatenated and MSC approaches. One example is the sisterhood of *Fagisyrphus* Dušek & Láska and *Meligramma* Frey, which are resolved as sister to a clade with *Leucozona* Schiner, *Epistrophella* Dušek & Láska, *Afrosyrphus* Curran and *Epistrophe* Walker (Figures S1–S13). Our results support the generic rank of *Epistrophella* and the close relationship between *Afrosyrphus* and *Epistrophe*, which was already suggested based on morphological characters of the male genitalia (Vockeroth, 1969). Other sisterhoods corroborate highly similar adult morphologies, for example, *Doros* (Figure 1b) and *Xanthogramma*, or *Syrphus* Fabricius and *Parasyrphus* Matsumura in Matsumura & Adachi. All our analyses resolved *Chrysotoxum* Meigen as sister to *Dasysyrphus* Enderlein with very high support, confirming the decision of Vockeroth (1992) to include *Chrysotoxum* into Syrphini.

Our results resolved the current concept of *Melangyna* Verrall, 1901 as non-monophyletic. Vockeroth (1969) established four *Melangyna* subgenera, including *Meligramma*. The other three subgenera were never grouped together in any of our analyses with strong support. Although *Melangyna* (*Austrosyrphus*) Vockeroth (Figure 1f) and *Melangyna* (*Melanosyrphus*) Vockeroth were grouped together in all our analyses, *Melangyna* (*Melangyna*) (Figure 1e) was identified as a rogue taxon by RogueNaRock and its phylogenetic placement is unstable. In his masterpiece, Vockeroth (1969) suggested the possibility that his new taxa *Melanosyrphus* and *Austrosyrphus* were distinct genera and pointed out the unusual geographic distribution of his *Melangyna* concept. More species of these taxa need to be sampled to better understand their phylogenetic relationships, but it seems plausible that the Australian taxa are not congeneric with Palaearctic *Melangyna*.

Compositional heterogeneity, nodal support and topological concordance

SYMTEST results indicate that all data sets violate, at least to some degree, the SRH conditions (Figures S15–S26). As pointed out by Bank et al. (2017) and Gillung et al. (2018) the SRH violations measured by SymTest may not impact our results as all inferred topologies are highly congruent. Our data sets based on amino acids and the second codon position (AA and NT2) are more compositionally homogeneous than the nucleotide data sets, in agreement with Gillung et al. (2018). Conversely, within each group of data sets (1302 loci and 154 loci), the nucleotide data set (NT) has more branches with high support (UFB ≥90) than the amino acid (AA) or second codon position (NT2) data sets, independent of the approach used.

At the same time, restrictive data sets (154 loci) are much more homogenous than complete data sets (1302 loci) and transcriptomes usually show high compositional heterogeneity. Contrarily, more branches within Syrphidae (outgroups excluded) have high support for the complete data sets (1302 loci) compared with the equivalent restrictive data sets (154 loci), either using the concatenated or the MSC approach (Table 1).

We believe that the discrepancies in the nodal support, that is, more compositional heterogeneity in NT data sets but more branches with high support, and more homogeneous behaviour in restrictive data sets but higher percentage of branches with high support in complete data sets, may be related to the presence of a rogue taxon. *Melangyna* (*Melangyna*) *lasiophthalma* was identified as a rogue taxon by ROGUENAROCK, although its phylogenetic placement received moderate support using the concatenated approach with the complete amino acid data set (AA-1302) (see Figures 2 and S1). In our inferred phylogeny using the concatenated approach with the AA-1302r (Figure S7) or the AA-1302rs (Figure S8) data sets, where *M. lasiophthalma* was excluded, all the internal branches of this major clade (Syrphini number 1) received higher support values (e.g., Figure S7: SH-aLRT = 100 and UFB = 100; except one branch with SH-aLRT = 98 and UFB = 63). In addition, the inferred topology with the AA-1302r has more branches and a higher percentage of branches with high support (UFB ≥90) than the NT-1302 (Table 1). Altogether, the exclusion of taxa identified as rogue certainly improves the BS in our study, as suggested by Aberer et al. (2013) and in agreement with Trautwein et al. (2011) and Bayless et al. (2021).

Table 1 shows that the global bootstrap support (GBS) as defined by Buenaventura et al. (2021) does not always correlate well with the number, or the percentage, of branches with high support. Subsampling methods can influence the perceived success or failure of the ML or MSC methods (Edwards, 2016). Thus, we calculated several support measurements, like the QBS in ASTRAL analyses, as concatenation-based analyses do not survey the relative support for alternative topologies (Vasilikopoulos et al., 2021). Overall, support values (SH-aLRT, QBS, UFB and BS) among different data sets using ML and MSC pointed out the same branches with moderate to low support (Figures 2 and S1).

All our inferred phylogenies were highly supported and congruent, independent of the data set and approach used in the analysis. Within Syrphini, the above-mentioned incongruences between data sets and methods are related to short branches in the inferred phylograms: three consecutive branches receive moderate to low support (SH-aLRT: 69.3–96.4; UFB: 36–71; BS: 92.3–93.3) in the Syrphini major clade number 1, and the branch supporting *Salpingogaster* + *Asarkina* together with (*Meliscaeva* + (*Episyrphus* + *Asiobaccha*)) has very low support (SH-aLRT: 2.5; UFB: 36; BS: 75.2) (see Figures 2 and S1). The support values for the consecutive branches increased a lot with the removal of the identified rogue taxon (Figures S7 and S8), but not for the node involving *Salpingogaster* and *Asarkina*. These short branches or low stemminess (Fiala & Sokal, 1985) might suggest closely spaced cladogenetic events (Rokas et al., 2005). The inferred ages for these cladogenetic events (middle Eocene) are overlapping. In other words, the inferred branching pattern may be explained by a rapid evolutionary diversification, especially for the deeper nodes in the Syrphini major clade number 1.

The fact that ML and MSC approaches resulted in virtually the same topology suggests that the minor incongruences between

inferred phylogenies of different data sets were neither due to high levels of ILS or ancient introgression (Vasilikopoulos et al., 2019), nor that there was a substitution-model misspecification (Klopfstein et al., 2019). Other published data sets show discordance between nucleotides and AAs (Bayless et al., 2021; Buenaventura, 2021; Gillung et al., 2018; Rota-Stabelli et al., 2013), between concatenated ML approach and MSC estimation (Brunke et al., 2021; Cao et al., 2022; Cohen et al., 2021; Hosner et al., 2016; Kim et al., 2020; Kohli et al., 2021; Maletti et al., 2021; Smith et al., 2015), or between larger/complete data sets and smaller/restrictive data sets (Buenaventura, 2021; Yan et al., 2021). This is not the case in the present study, where the inferred topologies are also highly congruent with previous published phylogenies using only a handful of molecular markers obtained through Sanger sequencing (Mengual, 2015, 2020; Mengual et al., 2008a, 2008b, 2012, 2015, 2021; Mengual, Ståhls, et al., 2018; Moran et al., 2022). Interestingly, these 'traditional' or 'bad' genes (sensu Winkler et al., 2015) seem to recover the same phylogenetic relationships as our new SYRPHIDAE1.0 bait kit. As mentioned by Philippe et al. (2005, p. 552) 'the increase in resolution obtained by analyzing larger data sets is not in itself a guarantee of accuracy. Conversely, the agreement between the methods does not mean that the obtained tree is correct'. On top, Brower (2019, p. 350) reminded that 'any measure of support or stability is only an estimate of how well a given data set supports a particular topology and not a measure of how likely to be true said topology might be'. To sum up, despite the criticism that phylogenetic relationships in Diptera based on molecular characters are volatile and inconsistent as they vary over the years as more data become available (Borkent, 2018, 2020), we constantly recover the same Syrphini relationships using data from HTS, Sanger sequencing or a combination of Sanger and morphological characters. Our results support the statement of Whitfield & Kjer (2008, p. 454) that '[f]or shallower phylogenetic depths, such as species within genera and recently diverged genera within families [at least within Syrphidae], data from as few as three to four genes spanning the mitochondrial and nuclear genomes … may suffice for good phylogenetic resolution and support'.

Estimated dates

In Bayesian dating analyses, the fossil calibration densities (temporal constraint on a node) can be the main driver of accuracy and precision in the posteriors or inferred age distributions (Bromham et al., 2018; Brown & Smith, 2018; Warnock et al., 2012, 2017). In our BEAST analysis, we used log-normal distributions to summarize palaeontological information, which we tried to copy in our MCMCTREE analysis using skew-t priors (Ho & Phillips, 2009; Marshall, 2019). It is not easy to select objectively a value for the required parameters (mean, standard deviation, and hard minimum bound) for these distributions (Ho & Phillips, 2009) or to overcome all sources of uncertainty (Bromham, 2019; Bromham et al., 2018). The consequences are different inferred ages from our analyses (Figures S27 and S28; Table 2). A possible explanation is that the used skew-t priors are less precise than the log-normal priors (Carruthers & Scotland, 2020).

We did not study any flower fly fossils and used only syrphid fossils for calibration, whose identification to genus level could be inferred from the original description. The affiliation uncertainty (stem or crown) may impact the divergence estimates (Ware & Barden, 2016). In addition, most fossils are poorly constrained with a relatively large age uncertainty (Table S2) (dos Reis et al., 2016; Marshall, 2019).

According to our BEAST analyses, flower flies emerged in the Lower Cretaceous, concurrent with a peak diversification of insect families in the Aptian (Schachat et al., 2019). Previous inferred ages for the family are found within our inferred 95% HPD interval (e.g., Bertone & Wiegmann, 2009; Caravas & Friedrich, 2013; Wiegmann et al., 2003) or are a little younger (95–82 Ma) (e.g., Ding et al., 2015; Grimaldi & Cumming, 1999; Grimaldi & Engel, 2005; Wiegmann et al., 2011). Only representatives of the stem group are known from the Mesozoic (Grimaldi, 2018), but the fossil fauna of Syrphidae is rich in the Eocene and Oligocene (Evenhuis, 1994; Ngô-Muller & Nel, 2020). Syrphini nodes with short branches have an inferred age between Middle and Upper Eocene (47.78–35.49 Ma) (Figure S27), which might indicate a rapid radiation during this period of time. Notably, two large Syrphini clades with mostly Neotropical members originated in the Late Paleogene (mostly in the Oligocene), namely the *Sphaerophoria* (33.95–25.99 Ma) and *Ocyptamus* (34.24–26.27 Ma) lineages; but also the Neotropical radiation in the *Eupeodes-Scaeva* lineage (35.27–27.12 Ma). The origin of these Neotropical clades predates the uplift in the Northern Andes, when Amazonia was dominated by a river system with northwest drainage (Hoorn et al., 2010), but the diversification of these flower flies likely occurred during the large-scale changes due to the Andean mountain formation.

The origin of Syrphidae is estimated at the time when angiosperms were still rare, before or at the beginning of their rapid diversification (Coiro et al., 2019; Sauquet et al., 2022; Schachat et al., 2019; van der Kooi & Ollerton, 2020), and flower fly early diversification concurred with the expansion of temperate zones due to the Pangea breakup (Chaboureau et al., 2014) and a peak in global temperature (Li et al., 2021), under humid and hot conditions. The origin of the pollinator-plant interaction for Syrphidae is estimated as 60 Ma during the Palaeocene (Ollerton, 1999), when the diversification of predatory flower flies occurred (Table 2; Figure 3). However, the plant reproductive structures visited by flower flies have their geochronological appearance during the Lower to Middle Cretaceous (Labandeira, 2005). The branching of the Microdontinae (95.7–73.1 Ma) is younger than the inferred age of the origin of ants in the Lower Cretaceous (103–123 Ma) and closer to the origin of the poneroid and formicoid clades (113.7–98.2 Ma) (Borowiec et al., 2019), whose members have been found in association with microdontines (Reemer, 2013). The diversification estimates of predatory syrphids (Pipizinae and Syrphinae) postdate the origin of their major prey group, Hemiptera (Grimaldi & Engel, 2005). Sternorrhyncha is the hemipteran suborder with the largest number of syrphid prey (Rojo et al., 2003), already present during the Carboniferous

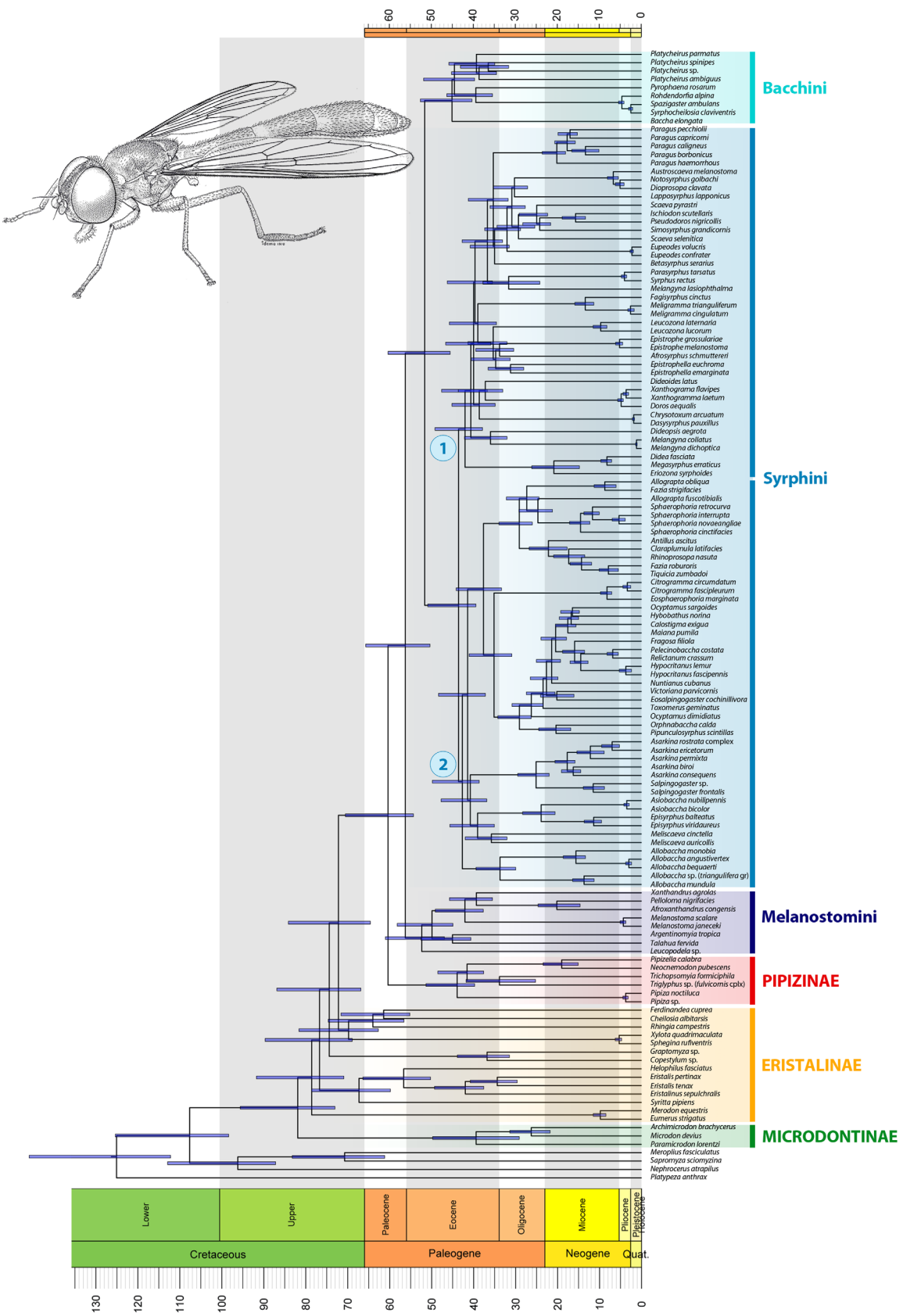

FIGURE 3 Ultrametric and time-calibrated cladogram of Syrphidae inferred using BEAST. The phylogenetic tree is identical to the one shown in Figure 2, but additionally shows the 95% highest posterior density of estimated divergence times (purple bars) (see also Figure S26). Numbers 1 and 2 within a blue circle indicate the two major clades within Syrphini. Line drawing adapted from Vockeroth and Thompson (1987): *Baccha cognata* Loew, 1863.

(Drohojowska et al., 2020; Garrouste et al., 2019; Johnson et al., 2018; Wang et al., 2016). Nevertheless, a rapid diversification of Aphidini aphids, common prey of Syrphinae larvae, occurred during the Eocene (Kim et al., 2011; Monnin et al., 2020) coinciding with the diversification of Syrphinae based on fossil evidence and our age estimate.

The radiation of angiosperm plants is suggested to trigger the diversification of other insect groups, especially herbivorous lineages (Kawahara et al., 2019; Misof et al., 2014; Peters et al., 2017; Sabatinelli et al., 2020; Sann et al., 2018; Suvorov et al., 2021). Our study indicates that the major diversification driver is the larval prey and their radiation, rather than the rise to dominance of the angiosperms.

CONCLUSIONS

In this study, we have designed 24,166 hybrid-capture baits to target 1945 CDS regions belonging to 1312 OGs. Our laboratory work proves that our bait kit SYRPHIDAE1.0 is highly specific, enriching all loci. We were able to effectively capture targeted loci for 121 flower fly species from 88 genera, with no taxonomic bias in the complete (1302 loci) and the restrictive (154 loci) data sets. Ongoing studies verify the efficacy of the bait kit SYRPHIDAE1.0 to enrich targeted loci in Microdontinae and Eristalinae, as well as in other dipteran families (Moran et al., in prep.). We believe that our SYRPHIDAE1.0 bait kit lays between a 'made to measure' and a 'one size fits all' tool (Kadlec et al., 2017), which efficiently captures CDS regions to resolve dipteran phylogenetic relationships and is applicable from population level to deep-level phylogenomics (McCormack et al., 2013).

Different data sets including transcriptomic data and phylogenetic reconstructions methodologies inferred highly congruent and supported topologies, with minor discrepancies related to short branch length and possibly rapid cladogenetic events. Removal of rogue taxa improved the support values for some deeper nodes within Syrphini and, in general, nucleotide data sets have a larger proportion of sequences with compositional heterogeneity than AA data sets.

Our inferred phylograms agree with previous phylogenetic relationships inferred with Sanger sequencing, supporting the sisterhood of Pipizinae and Syrphinae and a single origin of predatory larvae feeding on soft-bodied hemipterans. Based on our results and current evidence, *Allobaccha* is a senior synonym of *Petioleomyia* **syn.n.** and Syrphinae is divided into three tribes: Bacchini, Melanostomini and Syrphini **stat.rev.**, which includes the genera *Toxomerus* and *Paragus*.

Our divergence time estimation analysis with BEAST inferred the origin of the Syrphidae in the Lower Cretaceous, between the Aptian and Albian Ages (95% HPD = 125.5–98.5 Ma), and the diversification of predatory flower flies (Pipizinae and Syrphinae) occurred around the K–Pg boundary (95% HPD = 70.61–54.4 Ma). The Andean uplift may have played a major role in the diversification of Neotropical lineages, and the Eocene seems an important setting in generating the generic diversity of Syrphinae.

The application on Syrphidae of the exon-capture method assisted by BAITFISHER was successful in capturing all targeted loci and inferring well-supported and highly congruent phylogenies. The SYRPHIDAE1.0 bait kit is a reliable tool for future research in phylogenomics and future investigations in flower flies. The use of this sequence-capture method will foster future in-depth phylogenomic analyses in several flower fly lineages, crucial to understanding the evolution and the ecosystem services provided by this important pollinator group. Moreover, a detailed and dated phylogeny will provide a stable framework for other research fields, including biological control, comparative genomics, the evolution of larval feeding modes and inquiline–host associations, origin and biogeography of the different lineages, migration behaviour or mimicry development.

AUTHOR CONTRIBUTIONS

Ximo Mengual: conceptualization (lead); writing – original draft (lead); funding acquisition (equal); data curation (supporting); supervision; visualization (lead); resources (equal); writing – original draft (lead); writing – review and editing (equal). **Christoph Mayer:** data curation (supporting); formal analysis (lead); methodology (supporting); visualization (supporting); writing – original draft (supporting); writing – review and editing (equal). **Trevor O. Burt:** data curation (supporting); writing – original draft (supporting); writing – review and editing (equal). **Kevin M. Moran:** data curation (supporting); resources (equal); writing – original draft (supporting); writing – review and editing (equal). **Lars Dietz:** formal analysis (supporting); writing – original draft (supporting); writing – review and editing (equal). **Gaby Nottebrock:** data curation (supporting). **Thomas Pauli:** data curation (supporting); methodology (lead); writing – review and editing (equal). **Andrew D. Young:** data curation (supporting); resources (equal); writing – review and editing (equal). **Marie V. Brasseur:** data curation (supporting); formal analysis (supporting); writing – review and editing (equal). **Sandra Kukowka:** data curation (lead); writing – original draft (supporting); writing – review and editing (equal). **Scott Kelso:** data curation (supporting). **Claudia Etzbauer:** data curation (supporting). **Sander Bot:** resources (equal); writing – review and editing (equal). **Martin Hauser:** resources (equal); writing – review and editing (equal). **Kurt Jordaens:** resources (equal); writing – review and editing (equal). **Gil F. G. Miranda:** resources (equal); writing – review and editing (equal). **Gunilla Ståhls:** resources (equal); writing – review and editing (equal). **Wouter van Steenis:** resources (equal); writing – review and editing (equal). **Ralph S. Peters:** conceptualization (supporting); writing – review and editing (equal). **Jeffrey H. Skevington:** conceptualization (supporting); funding acquisition (equal); resources (equal); writing – original draft (supporting); writing – review and editing (equal).

ACKNOWLEDGMENTS

Ximo Mengual and Trevor O. Burt have received funding from the European Union's Horizon 2020 Research and Innovation Programme under Marie Sklodowska-Curie grant agreement no. 642241. Funding to Jeffrey H. Skevington was provided by Agriculture and Agri-Food Canada A-base and Biomob grants. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the Mendeley Data repository at http://doi.org/10.17632/ynghdr2msn.1. Unprocessed sequence data have been deposited in the NCBI Sequence Read Archive under BioProject PRJNA837085 (http://www.ncbi.nlm.nih.gov/bioproject/837085). Specimen information (locality, date, collector, identifier, and unique identifier) is accessible via the GenBank Accession Numbers (Table S1) and in BOLD (https://www.boldsystems.org/) under the Data set DS-SYRPHTE (https://doi.org/10.5883/DS-SYRPHTE).

ORCID

Ximo Mengual https://orcid.org/0000-0002-6185-9404
Christoph Mayer https://orcid.org/0000-0001-5104-6621
Trevor O. Burt https://orcid.org/0000-0003-3805-0997
Kevin M. Moran https://orcid.org/0000-0002-9460-4619
Thomas Pauli https://orcid.org/0000-0002-8381-6624
Marie V. Brasseur https://orcid.org/0000-0003-2044-4867
Sander Bot https://orcid.org/0000-0002-5504-4559
Martin Hauser https://orcid.org/0000-0002-6368-3529
Kurt Jordaens https://orcid.org/0000-0003-4321-5944
Gil F. G. Miranda https://orcid.org/0000-0001-7919-2639
Gunilla Ståhls https://orcid.org/0000-0003-0505-0691
Ralph S. Peters https://orcid.org/0000-0001-7784-9203
Jeffrey H. Skevington https://orcid.org/0000-0002-1445-9870

REFERENCES

Ababneh, F., Jermiin, L.S., Ma, C. & Robinson, J. (2006) Matched-pairs tests of homogeneity with applications to homologous nucleotide sequences. *Bioinformatics*, 22, 1225–1231. https://doi.org/10.1093/bioinformatics/btl064

Aberer, A.J., Krompass, D. & Stamatakis, A. (2013) Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. *Systematic Biology*, 62, 162–166. https://doi.org/10.1093/sysbio/sys078

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andermann, T., Torres Jiménez, M.F., Matos-Maraví, P., Batista, R., Blanco-Pastor, J.L., Gustafsson, A. et al. (2020) A guide to carrying out a phylogenomic target sequence capture project. *Frontiers in Genetics*, 10, 1407. https://doi.org/10.3389/fgene.2019.01407

Andrews, S. (2010) *FastQC: A quality control tool for high throughput sequence data*. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [Accessed 20 January 2022].

Anisimova, M., Gil, M., Dufayard, J.F., Dessimoz, C. & Gascuel, O. (2011) Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Systematic Biology*, 60(5), 685–699. https://doi.org/10.1093/sysbio/syr041

Aracil, A., Pérez-Bañón, C., Mengual, X., Radenković, S., Ståhls, G., Vujić, A. et al. (2019) New information about the pre-imaginal morphology of genus *Graptomyza* (Diptera, Syrphidae, Volucellini): description of third-instar larva and re-description of puparium of *G. signata* (Walker, 1860). *African Invertebrates*, 60(1), 15–30. https://doi.org/10.3897/afrinvertebr.60.31521

Arcaya, E., Pérez-Bañón, C., Mengual, X., Zubcoff-Vallejo, J.J. & Rojo, S. (2017) Life table and predation rates of the syrphid fly *Allograpta exotica*, a control agent of the cowpea aphid *Aphis craccivora*. *Biological Control*, 115, 74–84. https://doi.org/10.1016/j.biocontrol.2017.09.009

Arellano-Valle, R.B. & Azzalini, A. (2013) The centred parameterization and related quantities of the skew-t distribution. *Journal of Multivariate Analysis*, 113, 73–90. https://doi.org/10.1016/j.jmva.2011.05.016

Aubert, J. & Goeldlin De Tiefenau, P. (1981) Observations sur les migrations de Syrphides (Dipt.) dans les Alpes de Suisse occidentale. *Mitteilungen der Schweizerischen Entomologischen Gesellschaft*, 54, 377–388. https://doi.org/10.5169/seals-402013

Bank, S., Sann, M., Mayer, C., Meusemann, K., Donath, A., Podsiadlowski, L. et al. (2017) Transcriptome and target DNA enrichment sequence data provide new insights into the phylogeny of vespid wasps (Hymenoptera: Aculeata: Vespidae). *Molecular Phylogenetics and Evolution*, 116, 213–226. https://doi.org/10.1016/j.ympev.2017.08.020

Barendregt, A., Zeegers, T., van Steenis, W. & Jongejans, E. (2022) Forest hoverfly community collapse: abundance and species richness drop over four decades. *Insect Conservation and Diversity.*, 15, 510–521. https://doi.org/10.1111/icad.12577

Bayless, K.M., Trautwein, M.D., Meusemann, K., Shin, S., Petersen, M., Donath, A. et al. (2021) Beyond *Drosophila*: resolving the rapid radiation of schizophoran flies with phylotranscriptomics. *BMC Biology*, 19, 23. https://doi.org/10.1186/s12915-020-00944-8

Bell, M.A. & Lloyd, G.T. (2015) Strap: an R package for plotting phylogenies against stratigraphy and assessing their stratigraphic congruence. *Palaeontology*, 58, 379–389. https://doi.org/10.1111/pala.12142

Bellefeuille, Y., Fournier, M. & Lucas, E. (2019) Evaluation of two potential biological control agents against the Foxglove aphid at low temperatures. *Journal of Insect Science*, 19, 2. https://doi.org/10.1093/jisesa/iey130

Bergsten, J. (2005) A review of long-branch attraction. *Cladistics*, 21, 163–193. https://doi.org/10.1111/j.1096-0031.2005.00059.x

Bertone, M.A. & Wiegmann, B.M. (2009) True flies (Diptera). In: Hedges, S. B. & Kumar, S. (Eds.) *The Timetree of life*. Oxford: University Press, pp. 270–277.

Biesmeijer, J.C., Roberts, S.P., Reemer, M., Ohlemuller, R., Edwards, M., Peeters, T. et al. (2006) Parallel declines in pollinators and insect-pollinated plants in Britain and The Netherlands. *Science*, 313(5785), 351–354. https://doi.org/10.1126/science.1127863

Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borkent, A. (2018) The state of phylogenetic analysis: narrow visions and simple answers-examples from the Diptera (flies). *Zootaxa*, 4374(1), 107–143. https://doi.org/10.11646/zootaxa.4374.1.7

Borkent, A. (2020) Shrinking biodiversity, dwindling taxonomy and building a broader science. *Megataxa*, 1(1), 53–58. https://doi.org/10.11646/megataxa.1.1.11

Borowiec, M.L., Rabeling, C., Brady, S.G., Fisher, B.L., Schultz, T.R. & Ward, P.S. (2019) Compositional heterogeneity and outgroup choice influence the internal phylogeny of the ants. *Molecular Phylogenetics and Evolution*, 134, 111–121. https://doi.org/10.1016/j.ympev.2019.01.024

Bossert, S. & Danforth, B.N. (2018) On the universality of target-enrichment baits for phylogenomic research. *Methods in Ecology and Evolution*, 9(6), 1453–1460. https://doi.org/10.1111/2041-210x.12988

Bossert, S., Murray, E.A., Pauly, A., Chernyshov, K., Brady, S.G. & Danforth, B.N. (2021) Gene tree estimation error with ultraconserved elements: an empirical study on *Pseudapis* bees. *Systematic Biology*, 70(4), 803–821. https://doi.org/10.1093/sysbio/syaa097

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D. et al. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A. et al. (2019) BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 15(4), e1006650. https://doi.org/10.1371/journal.pcbi.1006650

Bowker, A.H. (1948) A test for symmetry in contingency tables. *Journal of the American Statistical Association*, 43, 572–574. https://doi.org/10.1080/01621459.1948.10483284

Bravo, G.A., Antonelli, A., Bacon, C.D., Bartoszek, K., Blom, M.P.K., Huynh, S. et al. (2019) Embracing heterogeneity: coalescing the tree of life and the future of phylogenomics. *PeerJ*, 7, e6399. https://doi.org/10.7717/peerj.6399

Bromham, L. (2019) Six impossible things before breakfast: assumptions, models, and belief in molecular dating. *Trends in Ecology & Evolution*, 34(5), 474–486. https://doi.org/10.1016/j.tree.2019.01.017

Bromham, L., Duchêne, S., Hua, X., Ritchie, A.M., Duchêne, D.A. & Ho, S.Y.W. (2018) Bayesian molecular dating: opening up the black box. *Biological Reviews of the Cambridge Philosophical Society*, 93(2), 1165–1191. https://doi.org/10.1111/brv.12390

Brower, A.V.Z. (2019) "Maximum support" = 100% BS. *Cladistics*, 35(3), 349–350. https://doi.org/10.1111/cla.12356

Brown, J.W. & Smith, S.A. (2018) The past sure is tense: on interpreting phylogenetic divergence time estimates. *Systematic Biology*, 67(2), 340–353. https://doi.org/10.1093/sysbio/syx074

Brunel, E. & Cadou, D. (1994) Syrphid larvae (Diptera: Syrphidae) mining the roots of artichoke (*Cynara scolymus* L.) in Brittany. *Dipterists Digest*, 1, 69–71.

Brunke, A.J., Hansen, A.K., Salnitska, M., Kypke, J.L., Predeus, A.V., Escalona, H. et al. (2021) The limits of Quediini at last (Staphylinidae: Staphylininae): a rove beetle mega-radiation resolved by comprehensive sampling and anchored phylogenomics. *Systematic Entomology*, 46, 396–421. https://doi.org/10.1111/syen.12468

Buenaventura, E. (2021) Museomics and phylogenomics with protein-encoding ultraconserved elements illuminate the evolution of life history and phallic morphology of flesh flies (Diptera: Sarcophagidae). *BMC Ecology and Evolution*, 21, 70. https://doi.org/10.1186/s12862-021-01797-7

Buenaventura, E., Lloyd, M.W., Perilla López, J.M., González, V.L., Thomas-Cabianca, A. & Dikow, T. (2021) Protein-encoding ultraconserved elements provide a new phylogenomic perspective of Oestroidea flies (Diptera: Calyptratae). *Systematic Entomology*, 46, 5–27. https://doi.org/10.1111/syen.12443

Burt, T.O. & Mengual, X. (2017) Origin and diversification of hoverflies: a revision of the genera *Asarkina* and *Allobaccha*—a BIG4 consortium PhD project. *Research Ideas and Outcomes*, 3, e19860. https://doi.org/10.3897/rio.3.e19860

Call, E., Mayer, C., Twort, V., Dietz, L., Wahlberg, K. & Espeland, M. (2021) Museomics: phylogenomics of the moth family Epicopeiidae (Lepidoptera) using target enrichment. *Insect Systematics and Diversity*, 5, 6. https://doi.org/10.1093/isd/ixaa021

Cao, Y., Dietrich, C.H., Zahniser, J.N. & Dmitriev, D.A. (2022) Dense sampling of taxa and characters improves phylogenetic resolution among deltocephaline leafhoppers (Hemiptera: Cicadellidae: Deltocephalinae). *Systematic Entomology*, 1–15, 430–444. https://doi.org/10.1111/syen.12540

Caravas, J. & Friedrich, M. (2013) Shaking the Diptera tree of life: performance analysis of nuclear and mitochondrial sequence data partitions. *Systematic Entomology*, 38(1), 93–103. https://doi.org/10.1111/j.1365-3113.2012.00657.x

Carruthers, T. & Scotland, R.W. (2020) Insights from empirical analyses and simulations on using multiple fossil calibrations with relaxed clocks to estimate divergence times. *Molecular Biology and Evolution*, 37(5), 1508–1529. https://doi.org/10.1093/molbev/msz310

Chaboureau, A.C., Sepulchre, P., Donnadieu, Y. & Franc, A. (2014) Tectonic-driven climate change and the diversification of angiosperms. *Proceedings of the National Academy of Sciences of the United States of America*, 111(39), 14066–14070. https://doi.org/10.1073/pnas.1324002111

Cohen, C.M., Noble, K., Jeffrey Cole, T. & Brewer, M.S. (2021) The phylogeny of robber flies (Asilidae) inferred from ultraconserved elements. *Systematic Entomology*, 46, 812–826. https://doi.org/10.1111/syen.12490

Coiro, M., Doyle, J.A. & Hilton, J. (2019) How deep is the conflict between molecular and fossil evidence on the age of angiosperms? *The New Phytologist*, 223(1), 83–99. https://doi.org/10.1111/nph.15708

Cook, D.F., Voss, S.C., Finch, J.T.D., Rader, R.C., Cook, J.M. & Spurr, C.J. (2020) The role of flies as pollinators of horticultural crops: an Australian case study with worldwide relevance. *Insects*, 11, 341. https://doi.org/10.3390/insects11060341

Costello, M., Fleharty, M., Abreu, J., Farjoun, Y., Ferriera, S., Holmes, L. et al. (2018) Characterization and remediation of sample index swaps by non-redundant dual indexing on massively parallel sequencing platforms. *BMC Genomics*, 19, 332. https://doi.org/10.1186/s12864-018-4703-0

Cumming, J.M., Sinclair, B.J. & Wood, D.M. (1995) Homology and phylogenetic implications of male genitalia in Diptera - Eremoneura. *Insect Systematics and Evolution*, 26, 120–151. https://doi.org/10.1163/187631295X00143

Cumming, J.M. & Wood, D.M. (2017) Adult morphology and terminology. Manual of Afrotropical Diptera, Volume 1. Introductory chapters and keys to Diptera families (ed. by A.H. Kirk-Spriggs & B.J. Sinclair), pp. 89–133. Suricata 4. South African National Biodiversity Institute, Pretoria.

Daily, G.C. (1997) Introduction: what are ecosystem services? In: Daily, G.C. (Ed.) *Nature's services: societal dependence on natural ecosystems*. Washington, D.C.: Island Press, pp. 1–10.

Degnan, J.H. & Rosenberg, N.A. (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24(6), 332–340. https://doi.org/10.1016/j.tree.2009.01.009

Dietz, L., Dömel, J.S., Leese, F., Mahon, A.R. & Mayer, C. (2019) Phylogenomics of the longitarsal Colossendeidae: the evolutionary history of an Antarctic Sea spider radiation. *Molecular Phylogenetics and Evolution*, 136, 206–214. https://doi.org/10.1016/j.ympev.2019.04.017

Ding, S., Li, X., Wang, N., Cameron, S.L., Mao, M., Wang, Y. et al. (2015) The phylogeny and evolutionary timescale of Muscoidea (Diptera: Brachycera: Calyptratae) inferred from mitochondrial genomes. *PLoS One*, 10(7), e0134170. https://doi.org/10.1371/journal.pone.0134170

dos Reis, M., Álvarez-Carretero, S. & Yang, Z. (2017) *MCMCTree tutorials* [WWW document]. Available from: http://abacus.gene.ucl.ac.uk/software/MCMCtree.Tutorials.pdf [Accessed 8 February 2022].

dos Reis, M., Donoghue, P. & Yang, Z. (2016) Bayesian molecular clock dating of species divergences in the genomics era. *Nature Reviews Genetics*, 17(2), 71–80. https://doi.org/10.1038/nrg.2015.8

Downes, M.F., Skevington, J.H. & Thompson, F.C. (2017) A new ant inquiline flower fly (Diptera: Syrphidae: Pipizinae) from Australia. *Australian Entomologist*, 44(1), 29–38.

Doyle, T., Hawkes, W.L.S., Massy, R., Powney, G.D., Menz, M.H.M. & Wotton, K.R. (2020) Pollination by hoverflies in the Anthropocene. *Proceedings of the Royal Society B: Biological Sciences*, 287, 20200508. https://doi.org/10.1098/rspb.2020.0508

Drohojowska, J., Szwedo, J., Żyła, D., Huang, D.-Y. & Müller, P. (2020) Fossils reshape the Sternorrhyncha evolutionary tree (Insecta, Hemiptera). *Scientific Reports*, 10, 11390. https://doi.org/10.1038/s41598-020-68220-x

Dumbardon-Martial, E. (2016) Pollen feeding in the larva of *Toxomerus pulchellus* (Diptera, Syrphidae). *Bulletin de la Société entomologique de France*, 121, 413–420. https://doi.org/10.1080/00222930802610576

Dunn, L., Lequerica, M., Reid, C.R. & Latty, T. (2020) Dual ecosystem services of syrphid flies (Diptera: Syrphidae): pollinators and biological control agents. *Pest Management Science*, 76, 1973–1979. https://doi.org/10.1002/ps.5807

Dušek, J. & Láska, P. (1967) Versuch zum Aufbau eines natürlichen Systems mitteleuropäischer Arten der Unterfamilie Syrphinae (Diptera). *Acta Scientiarum Naturalium Academiae Scientiarum Bohemicae Brno*, 1, 349–390.

Eberle, J., Ahrens, D., Mayer, C., Niehuis, O. & Misof, B. (2020) A plea for implementing a standardized set of nuclear markers in DNA taxonomy. *Trends in Ecology and Evolution*, 35(4), 336–345. https://doi.org/10.1016/j.tree.2019.12.003

Edwards, S.V. (2009) Is a new and general theory of molecular systematics emerging? *Evolution*, 63, 1–19. https://doi.org/10.1111/j.1558-5646.2008.00549.x

Edwards, S.V. (2016) Phylogenomic subsampling: a brief review. *Zoologica Scripta*, 45, 63–74. https://doi.org/10.1111/zsc.12210

Edwards, E.E. & Bevan, W.J. (1951) On the Narcissus flies, *Merodon equestris* (F.) and *Eumerus tuberculatus* (Rond.) and their control. *Bulletin of Entomological Research*, 41(3), 593–598. https://doi.org/10.1017/S000748530002784X

Enderlein, G. (1938) Beitrag zur Kenntnis der Syrphiden. *Sitzungsberichte der Gesellschaft Naturforschender Freunde zu Berlin*, 1937, 192–237.

Evangelista, D., Simon, S., Wilson, M.M., Kawahara, A.Y., Kohli, M.K., Ware, J.L. et al. (2021) Assessing support for Blaberoidea phylogeny suggests optimal locus quality. *Systematic Entomology*, 46, 157–171. https://doi.org/10.1111/syen.12454

Evenhuis, N.L. (1994) *Catalogue of the fossil flies of the world (Insecta: Diptera)*. Leiden: Backhuys Publishers.

Faircloth, B.C. (2017) Identifying conserved genomic elements and designing universal bait sets to enrich them. *Methods in Ecology and Evolution*, 8(9), 1103–1112. https://doi.org/10.1111/2041-210x.12754

Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T. & Glenn, T.C. (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, 61, 717–726. https://doi.org/10.1093/sysbio/sys004

Felsenstein, J. (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, 27, 401–410. https://doi.org/10.2307/2412923

Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4), 783–791. https://doi.org/10.2307/2408678

Fiala, K.L. & Sokal, R.R. (1985) Factors determining the accuracy of cladogram estimation: evaluation using computer simulation. *Evolution*, 39, 609–622. https://doi.org/10.1111/j.1558-5646.1985.tb00398.x

Finch, J.T.D. & Cook, J.M. (2020) Flies on vacation: evidence for the migration of Australian Syrphidae (Diptera). *Ecological Entomology*, 45, 896–900. https://doi.org/10.1111/een.12856

Fleischmann, A., Rivadavia, F., Gonella, P.M., Pérez-Bañón, C., Mengual, X. & Rojo, S. (2016) Where is my food? Brazilian flower fly steals prey from carnivorous sundews in a newly discovered plant-animal interaction. *PLoS One*, 11, e0153900. https://doi.org/10.1371/journal.pone.0153900

Garrouste, R., Schubnel, T., Oudard, J., Roques, P. & Nel, A. (2019) The insect Konservat-Lagerstätte of the upper carboniferous of Avion (France): an exceptional geoheritage. In: Nascimbene, P.C. (Ed.) *Abstracts of the 8th international conference on fossil insects, Arthropods & Amber, Santo Domingo 2019*. Amber World Museum: Santo Domingo, pp. 43–44.

Gatter, W., Ebenhöh, H., Kima, R., Gatter, W. & Scherer, F. (2020) 50-jährige Untersuchungen an migrierenden Schwebfliegen, Waffenfliegen und Schlupfwespen belegen extreme Rückgänge (Diptera: Syrphidae, Stratiomyidae; Hymenoptera: Ichneumonidae). *Entomologische Zeitschrift*, 130, 131–142.

Gauthier, M., Konecny-Dupré, L., Nguyen, A., Elbrecht, V., Datry, T., Douady, C. et al. (2020) Enhancing DNA metabarcoding performance and applicability with bait capture enrichment and DNA from conservative ethanol. *Molecular Ecology Resources*, 20, 79–96. https://doi.org/10.1111/1755-0998.13088

Gibson, J.F., Kelso, S., Jackson, M.D., Kits, J.H., Miranda, G.F.G. & Skevington, J.H. (2011) Diptera-specific polymerase chain reaction amplification primers of use in molecular phylogenetic research. *Annals of the Entomological Society of America*, 104(5), 976–997. https://doi.org/10.1603/AN10153

Gibson, J.F., Skevington, J.H. & Kelso, S. (2010) Placement of the Conopidae (Diptera) within the Schizophora based on ten mtDNA and nrDNA gene regions. *Molecular Phylogenetics and Evolution*, 56, 91–103. https://doi.org/10.1016/j.ympev.2010.03.026

Giebner, H., Langen, K., Bourlat, S.J., Kukowka, S., Mayer, C., Astrin, J.J. et al. (2020) Comparing diversity levels in environmental samples: DNA sequence capture and metabarcoding approaches using 18S and COI genes. *Molecular Ecology Resources*, 20, 1333–1345. https://doi.org/10.1111/1755-0998.13201

Gillung, J.P., Winterton, S.L., Bayless, K.M., Khouri, Z., Borowiec, M.L., Yeates, D. et al. (2018). Anchored phylogenomics unravels the evolution of spider flies (Diptera, Acroceridae) and reveals discordance between nucleotides and amino acids. *Molecular Phylogenetics and Evolution*, *128*, 233–245. https://doi.org/10.1016/j.ympev.2018.08.007

Glumac, S. (1960) Phylogenetical system of the syrphid-flies (Syrphidae, Diptera) based upon the male genitalia structure and the type of the larvae with characteristics of the family and tribes. *Glasnik prirod Muzeja Srpske Zemlje*, 16, 82–103.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E.M., Brockman, W. et al. (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature Biotechnology*, 27, 182–189. https://doi.org/10.1038/nbt.1523

Goeldlin de Tiefenau, P. & Thompson, F.C. (2019) A new flower fly from the Afrotropics (Diptera: Syrphidae). *Entomologist's Monthly Magazine*, 155, 159–162. https://doi.org/10.31184/M00138908.1553.3967

Grimaldi, D.A. (2018) Basal Cyclorrhapha in amber from the cretaceous and tertiary (Insecta: Diptera), and their relationships: Brachycera in cretaceous amber part IX. *Bulletin of the American Museum of Natural History*, 423, 1–97. https://doi.org/10.1206/0003-0090-423.1.1

Grimaldi, D.A. & Cumming, J. (1999) Diptera in cretaceous amber. *Bulletin American Museum of Natural History*, 239, 1–123 http://hdl.handle.net/2246/1583

Grimaldi, D.A. & Engel, M.S. (2005) *Evolution of insects*. New York: Cambridge University Press.

Grosskopf, G. (2005) Biology and life history of Cheilosia urbana (Meigen) and Cheilosia psilophthalma (Becker), two sympatric hoverfies approved for the biological control of hawkweeds (Hieracium spp.) in New Zealand. *Biological Control*, *35*, 142–154. https://doi.org/10.1016/j.biocontrol.2005.06.013

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321. https://doi.org/10.1093/sysbio/syq010

Hadrava, J., Nidergas, V., Daňková, K., Pecharová, M., Nel, A. & Prokop, J. (2020) *Blera miocenica*: a new species of early Miocene hoverfly (Diptera: Syrphidae) from The Czech Republic and its palaeoenvironmental significance. *Insect Systematics & Evolution*, 51(5), 811–819. https://doi.org/10.1163/1876312X-00001043

Han, M.V. & Zmasek, C.M. (2009) phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics*, 10, 356. https://doi.org/10.1186/1471-2105-10-356

Heibl, C. (2008) *PHYLOCH: R language tree plotting tools and interfaces to diverse phylogenetic software packages*. Available from: http://www.christophheibl.de/Rpackages.html.

Heo, C.C., Rahimi, R., Mengual, X., Isa, M.S., M., Zainal, S., Khofar, P.N. et al. (2020) *Eristalinus arvorum* (Fabricius, 1787) (Diptera: Syrphidae) in human skull: a new fly species of forensic importance. *Journal of Forensic Sciences*, 65, 276–282. https://doi.org/10.1111/1556-4029.14128

Hippa, H. & Ståhls, G. (2005) Morphological characters of adult Syrphidae: descriptions and phylogenetic utility. *Acta Zoologica Fennica*, 215, 1–72.

Ho, S.Y.W. & Phillips, M.J. (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, 58(3), 367–380. https://doi.org/10.1093/sysbio/syp035

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q. & Vinh, L.S. (2018) UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35, 518–522. https://doi.org/10.1093/molbev/msx281

Hodgkiss, D., Brown, M.J.F. & Fountain, M.T. (2018) Syrphine hoverflies are effective pollinators of commercial strawberry. *Journal of Pollination Ecology*, 22, 55–66. https://doi.org/10.26786/1920-7603(2018)five

Hoorn, C., Wesselingh, F.P., ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J. et al. (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, 330(6006), 927–931. https://doi.org/10.1126/science.1194585

Hoskins, R.A., Carlson, J.W., Wan, K.H., Park, S., Mendez, I., Galle, S.E. et al. (2015) The release 6 reference sequence of the *Drosophila melanogaster* genome. *Genome Research*, 25(3), 445–458. https://doi.org/10.1101/gr.185579.114

Hosner, P.A., Faircloth, B.C., Glenn, T.C., Braun, E.L. & Kimball, R.T. (2016) Avoiding missing data biases in phylogenomic inference: an empirical study in the Landfowl (Aves: Galliformes). *Molecular Biology and Evolution*, 33(4), 1110–1125. https://doi.org/10.1093/molbev/msv347

Inouye, D.W., Larson, B.M.H., Ssymank, A. & Kevan, P.G. (2015) Flies and flowers III: ecology of foraging and pollination. *Journal of Pollination Ecology*, 16, 115–133. https://doi.org/10.26786/1920-7603(2015)15

Jauker, F. & Wolters, V. (2008) Hover flies are efficient pollinators of oilseed rape. *Oecologia*, 156, 819–823. https://doi.org/10.1007/s00442-008-1034-x

Jermiin, L., Ho, S.Y.W., Ababneh, F., Robinson, J. & Larkum, A.W. (2004) The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. *Systematic Biology*, 53, 638–643. https://doi.org/10.1080/10635150490468648

Johnson, K.P., Dietrich, C.H., Friedrich, F., Beutel, R.G., Wipfler, B., Peters, R.S. et al. (2018) Phylogenomics and the evolution of hemipteroid insects. *Proceedings of the National Academy of Sciences of the United States of America*, 115(50), 12775–12780. https://doi.org/10.1073/pnas.1815820115

Kadlec, M., Bellstedt, D.U., Le Maitre, N.C. & Pirie, M.D. (2017) Targeted NGS for species level phylogenomics: "made to measure" or "one size fits all"? *PeerJ*, 5, e3569. https://doi.org/10.7717/peerj.3569

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A. & Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14, 587–589. https://doi.org/10.1038/nmeth.4285

Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780. https://doi.org/10.1093/molbev/mst010

Kawahara, A.Y., Plotkin, D., Espeland, M., Meusemann, K., Toussaint, E., Donath, A. et al. (2019) Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proceedings of the National Academy of Sciences of the United States of America*, 116(45), 22657–22663. https://doi.org/10.1073/pnas.1907847116

Kim, M.J., Kim, I. & Cameron, S.L. (2020) How well do multispecies coalescent methods perform with mitochondrial genomic data? A case study of butterflies and moths (Insecta: Lepidoptera). *Systematic Entomology*, 45, 857–873. https://doi.org/10.1111/syen.12431

Kim, H., Lee, S. & Jang, Y. (2011) Macroevolutionary patterns in the Aphidini aphids (Hemiptera: Aphididae): diversification, host association, and biogeographic origins. *PLoS One*, 6(9), e24749. https://doi.org/10.1371/journal.pone.0024749

Klein, A.-M., Vaissière, B.E., Cane, J.H., Steffan-Dewenter, I., Cunningham, S.A., Kremen, C. et al. (2007) Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences*, 274, 303–313. https://doi.org/10.1098/rspb.2006.3721

Klopfstein, S., Langille, B., Spasojevic, T., Broad, G.R., Cooper, S.J.B., Austin, A.D. et al. (2019) Hybrid capture data unravel a rapid radiation of pimpliform parasitoid wasps (Hymenoptera: Ichneumonidae: Pimpliformes). *Systematic Entomology*, 44, 361–383. https://doi.org/10.1111/syen.12333

Knowles, L.L. & Kubatko, L.S. (2010) *Estimating species trees: practical and theoretical aspects*. Hoboken, NJ: Wiley-Blackwell.

Kohli, M., Letsch, H., Greve, C., Béthoux, O., Deregnaucourt, I., Liu, S. et al. (2021) Evolutionary history and divergence times of Odonata (dragonflies and damselflies) revealed through transcriptomics. *iScience*, 24, 103324. https://doi.org/10.1016/j.isci.2021.103324

Köster, J. & Rahmann, S. (2012) Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics*, 28, 2520–2522. https://doi.org/10.1093/bioinformatics/bts480

Kück, P. & Meusemann, K. (2010) FASconCAT: convenient handling of data matrices. *Molecular Phylogenetics and Evolution*, 56, 1115–1118. https://doi.org/10.1016/j.ympev.2010.04.024

Kück, P., Meusemann, K., Dambach, J., Thormann, B., von Reumont, B.M., Wägele, J.W. et al. (2010) Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. *Frontiers in Zoology*, 7, 10. https://doi.org/10.1186/1742-9994-7-10

Labandeira, C.C. (2005) Fossil history and evolutionary ecology of Diptera and their associations with plants. In: Yeates, D.K. & Wiegmann, B.M. (Eds.) *The evolutionary biology of flies*. New York: Columbia University Press, pp. 217–273.

Lanfear, R., Calcott, B., Kainer, D., Mayer, C. & Stamatakis, A. (2014) Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, 14, 82. https://doi.org/10.1186/1471-2148-14-82

Lardé, G. (1989) Investigation on some factors affecting larval growth in a coffee-pulp bed. *Biological Wastes*, 30, 11–19. https://doi.org/10.1016/0269-7483(89)90139-0

Lardé, G. (1990) Growth of *Ornidia obesa* (Diptera: Syrphidae) larvae on decomposing coffee pulp. *Biological Wastes*, 34, 73–76. https://doi.org/10.1016/0269-7483(90)90144-H

Larson, B.M.H., Kevan, P.G. & Inouye, D.W. (2001) Flies and flowers: taxonomic diversity of anthophiles and pollinators. *The Canadian Entomologist*, 133, 439–466. https://doi.org/10.4039/Ent133439-4

Láska, P., Bičík, V., Dušek, J., Mazánek, L. & Holinka, J. (2000) Description of the puparia of *Afrosyrphus varipes* and *Dideopsis aegrotus* (Diptera, Syrphidae). *Acta Universitatis Palackianae Olomucensis, Facultas Rerum Naturalium Biologica*, 38, 79–83.

Lemmon, A.R., Emme, S.A. & Lemmon, E.M. (2012) Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61, 727–744. https://doi.org/10.1093/sysbio/sys049

Lemmon, E.M. & Lemmon, A.R. (2013) High-throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, 44(1), 99–121. https://doi.org/10.1146/annurev-ecolsys-110512-135822

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N. et al. (2009) The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, X., Hao, B., Pan, D. & Schneeweiss, G.M. (2017) Marker development for phylogenomics: the case of Orobanchaceae, a plant family with contrasting nutritional modes. *Frontiers in Plant Science*, 8, 1973. https://doi.org/10.3389/fpls.2017.01973

Li, X., Teasdale, L.C., Bayless, K.M., Ellis, A.G., Wiegmann, B.M., Lamas, C. et al. (2021) Phylogenomics reveals accelerated late cretaceous diversification of bee flies (Diptera: Bombyliidae). *Cladistics*, 37(3), 276–297. https://doi.org/10.1111/cla.12436

Liu, L., Xi, X. & Davis, C.C. (2015) Coalescent methods are robust to the simultaneous effects of long branches and incomplete lineage sorting. *Molecular Biology and Evolution*, 32, 791–805. https://doi.org/10.1093/molbev/msu331

Magni, P., Pérez-Bañón, C., Borrini, M. & Dadour, I. (2013) *Syritta pipiens* (Diptera: Syrphidae), a new species associated with human cadavers. *Forensic Science International*, 231, e19–e23. https://doi.org/10.1016/j.forsciint.2013.05.023

Maletti, S., Niehuis, O., Mayer, C., Sann, M., Klopfstein, S., Nottebrock, G. et al. (2021) Phylogeny, taxonomics, and ovipositor length variation of the *Pteromalus albipennis* species group (Hymenoptera: Chalcidoidea: Pteromalidae: Pteromalinae). *Journal of Zoological Systematics and Evolutionary Research*, 59, 349–358. https://doi.org/10.1111/jzs.12433

Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A. et al. (2010) Target-enrichment strategies for next-generation sequencing. *Nature Methods*, 7(2), 111–118. https://doi.org/10.1038/nmeth.1419

Marshall, C.R. (2019) Using the fossil record to evaluate timetree timescales. *Frontiers in Genetics*, 10, 1049. https://doi.org/10.3389/fgene.2019.01049

Mayer, C., Dietz, L., Call, E., Kukowka, S., Martin, S. & Espeland, M. (2021) Adding leaves to the Lepidoptera tree: capturing hundreds of nuclear genes from old museum specimens. *Systematic Entomology*, 46, 649–671. https://doi.org/10.1111/syen.12481

Mayer, C., Sann, M., Donath, A., Meixner, M., Podsiadlowski, L., Peters, R. S. et al. (2016) BaitFisher: a software package for multispecies target DNA enrichment probe design. *Molecular Biology and Evolution*, 33, 1875–1886. https://doi.org/10.1093/molbev/msw056

McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C. & Brumfield, R.T. (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, 66(2), 526–538. https://doi.org/10.1016/j.ympev.2011.12.007

Meier, R. (2017) Citation of taxonomic publications: the why, when, what and what not. *Systematic Entomology*, 42, 301–304. https://doi.org/10.1111/syen.12215

Mengual, X. (2015) The systematic position and phylogenetic relationships of *Asiobaccha* Violovitsh (Diptera, Syrphidae). *Journal of Asia-Pacific Entomology*, 18, 397–408. https://doi.org/10.1016/j.aspen.2015.03.010

Mengual, X. (2020) Phylogenetic relationships of the bacchine flower flies (Diptera: Syrphidae) based on molecular characters, with a description of a new species of *Melanostoma* (Schiner, 1860). *Contributions to Zoology*, 89, 210–244. https://doi.org/10.1163/18759866-20191410

Mengual, X., Kerr, P., Norrbom, A.L., Barr, N.B., Lewis, M.L., Stapelfeldt, A. M. et al. (2017) Phylogenetic relationships of the tribe Toxotrypanini (Diptera: Tephritidae) based on molecular characters. *Molecular Phylogenetics and Evolution*, 113, 84–112. https://doi.org/10.1016/j.ympev.2017.05.011

Mengual, X., Miranda, G.F.G. & Thompson, F.C. (2018) Unraveling *Ocyptamus* and the *Baccha* legacy (Diptera: Syrphidae): redefinition of groups and new species descriptions. *Zootaxa*, 4461(1), 1–44. https://doi.org/10.11646/zootaxa.4461.1.1

Mengual, X., Ståhls, G., Láska, P., Mazánek, L. & Rojo, S. (2018) Molecular phylogenetics of the predatory lineage of flower flies *Eupeodes-Scaeva* (Diptera: Syrphidae), with the description of the Neotropical genus *Austroscaeva* gen.nov. *Journal of Zoological Systematics and Evolutionary Research*, 56, 148–169. https://doi.org/10.1111/jzs.12212

Mengual, X., Ståhls, G. & Rojo, S. (2008a) Molecular phylogeny of *Allograpta* (Diptera, Syrphidae) reveals diversity of lineages and non-monophyly of phytophagous taxa. *Molecular Phylogenetics and Evolution*, 49, 715–727. https://doi.org/10.1016/j.ympev.2008.09.011

Mengual, X., Ståhls, G. & Rojo, S. (2008b) First phylogeny of predatory flower flies (Diptera, Syrphidae, Syrphinae) using mitochondrial COI and nuclear 28S rRNA genes: conflict and congruence with the current tribal classification. *Cladistics*, 24, 543–562. https://doi.org/10.1111/j.1096-0031.2008.00200.x

Mengual, X., Ståhls, G. & Rojo, S. (2012) Is the mega-diverse genus *Ocyptamus* (Diptera, Syrphidae) monophyletic? Evidence from molecular characters including the secondary structure of 28S rRNA. *Molecular Phylogenetics and Evolution*, 62, 191–205. https://doi.org/10.1016/j.ympev.2011.09.014

Mengual, X., Ståhls, G. & Rojo, S. (2015) Phylogenetic relationships and taxonomic ranking of pipizine flower flies (Diptera: Syrphidae) with implications for the evolution of aphidophagy. *Cladistics*, 31, 491–508. https://doi.org/10.1111/cla.12105

Mengual, X., Ståhls, G. & Skevington, J.H. (2021) Life on an Island: the phylogenetic placement of *Loveridgeana* and Afrotropical *Sphaerophoria* (Diptera: Syrphidae) inferred from molecular characters. *Systematics and Biodiversity*, 19, 22–53. https://doi.org/10.1080/14772000.2020.1795743

Mengual, X. & Thompson, F.C. (2011) Carmine cochineal killers: the flower fly genus *Eosalpingogaster* Hull (Diptera: Syrphidae) revised. *Systematic Entomology*, 36(4), 713–731. https://doi.org/10.1111/j.1365-3113.2011.00588.x

Menz, M.H.M., Brown, B.V. & Wotton, K.R. (2019) Quantification of migrant hoverfly movements (Diptera: Syrphidae) on the west coast of North America. *Royal Society Open Science*, 6, 190153. https://doi.org/10.1098/rsos.190153

Minh, B.Q., Nguyen, M.A.T. & von Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*, 30(5), 1188–1195. https://doi.org/10.1093/molbev/mst024

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. et al. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Miranda, G.F.G., Marshall, S.A. & Skevington, J.H. (2014) Revision of the genus *Pelecinobaccha* Shannon, description of *Relictanum* gen. nov., and redescription of *Atylobaccha flukiella* (Curran, 1941) (Diptera: Syrphidae). *Zootaxa*, 3819(1), 1–154. https://doi.org/10.11646/zootaxa.3819.1.1

Miranda, G.F.G., Skevington, J.H. & Marshall, S.A. (2020) New generic concepts for orphaned lineages formerly treated as part of the genus *Ocyptamus* Macquart, 1834 (Diptera, Syrphidae). *Zootaxa*, 4822(2), 151–174. https://doi.org/10.11646/zootaxa.4822.2.1

Miranda, G.F.G., Skevington, J.H., Marshall, S.A. & Kelso, S. (2016) The genus *Ocyptamus* Macquart (Diptera: Syrphidae): a molecular phylogenetic analysis. *Arthropod Systematics and Phylogeny*, 74, 161–176.

Mirarab, S. (2019) Species tree estimation using ASTRAL: practical considerations. *arXiv* (preprint), arXiv:1904.03826v2.

Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S. & Warnow, T. (2014) ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30(17), i541–i548. https://doi.org/10.1093/bioinformatics/btu462

Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C. et al. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346(6210), 763–767. https://doi.org/10.1126/science.1257570

Misof, B., Meyer, B., von Reumont, B.M., Kück, P., Misof, K. & Meusemann, K. (2013) Selecting informative subsets of sparse supermatrices increases the chance to find correct trees. *BMC Bioinformatics*, 14, 348. https://doi.org/10.1186/1471-2105-14-348

Misof, B. & Misof, K. (2009) A Monte Carlo approach successfully identifies randomness in multiple sequence alignments: a more objective means of data exclusion. *Systematic Biology*, 58, 21–34. https://doi.org/10.1093/sysbio/syp006

Moerkens, R., Boonen, S., Wäckers, F.L. & Pekas, A. (2021) Aphidophagous hoverflies reduce foxglove aphid infestations and improve seed set and fruit yield in sweet pepper. *Pest Management Science*, 77(6), 2690–2696. https://doi.org/10.1002/ps.6342

Monnin, D., Jackson, R., Kiers, E.T., Bunker, M., Ellers, J. & Henry, L.M. (2020) Parallel evolution in the integration of a co-obligate aphid symbiosis. *Current Biology*, 30(10), 1949–1957.e6. https://doi.org/10.1016/j.cub.2020.03.011

Morales, G.E. & Wolff, M. (2010) Insects associated with the composting process of solid urban waste separated at the source. *Revista Brasileira de Entomologia*, 54(4), 645–653. https://doi.org/10.1590/S0085-56262010000400017

Moran, K.M., Skevington, J.H., Kelso, S., Mengual, X., Jordaens, K., Young, A.D. et al. (2022) A multigene phylogeny of the eristaline flower flies (Diptera: Syrphidae), with emphasis on the subtribe Criorhinina. *Zoological Journal of the Linnean Society*, 194, 120–135. https://doi.org/10.1093/zoolinnean/zlab006

Mullens, N., Sonet, G., Virgilio, M., Goergen, G., Janssens, S.B., De Meyer, M. et al. (2022) Systematics of Afrotropical Eristalinae (Diptera: Syrphidae) using mitochondrial phylogenomics. *Systematic Entomology*, 47, 315–328. https://doi.org/10.1111/syen.12532

Nelson, E.H., Hogg, B.N., Mills, N.J. & Daane, K.M. (2012) Syrphid flies suppress lettuce aphids. *BioControl*, 57, 819–826. https://doi.org/10.1007/s10526-012-9457-z

Ngô-Muller, V. & Nel, A. (2020) A new hoverfly genus and species from the middle Eocene Baltic amber (Diptera: Syrphidae). *Palaeoentomology*, 3(1), 68–75. https://doi.org/10.11646/palaeoentomology.3.1.10

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

Nidergas, V., Hadrava, J., Garrouste, R., Prokop, J., Schubnel, T. & Nel, A. (2018) The first pipizine hoverfly genus and species from the Oligocene of France (Diptera: Syrphidae). *Acta Palaeontologica Polonica*, 63(3), 539–548. https://doi.org/10.4202/app.00500.2018

Nishida, K., Rotheray, G. & Thompson, F.C. (2002) First non-predaceous syrphine flower-fly (Diptera: Syrphidae): a new leaf-mining *Allograpta* from Costa Rica. *Studia Dipterologica*, 9, 421–436.

Ollerton, J. (1999) La evolución de las relaciones polinizador-planta en los artrópodos. *Boletín de la Sociedad Entomológica Aragonesa*, 26, 741–758.

Packer, L., Monckton, S.K., Onuferko, T.M. & Ferrari, R.R. (2018) Validating taxonomic identifications in entomological research. *Insect Conservation and Diversity*, 11(1), 1–12. https://doi.org/10.1111/icad.12284

Parham, J.F., Donoghue, P.C., Bell, C.J., Calway, T.D., Head, J., Holroyd, P.A. et al. (2012) Best practices for justifying fossil calibrations. *Systematic Biology*, 61(2), 346–359. https://doi.org/10.1093/sysbio/syr107

Pauli, T., Burt, T., Meusemann, K., Bayless, K., Donath, A., Podsiadlowski, L. et al. (2018) New data, same story: phylogenomics does not support Syrphoidea (Diptera: Syrphidae, Pipunculidae). *Systematic Entomology*, 43(3), 447–459. https://doi.org/10.1111/syen.12283

Pauli, T., Meusemann, K., Kukowka, S., Sann, M., Donath, A., Mayer, C. et al. (2021) Analysis of RNA-Seq, DNA target enrichment, and sanger nucleotide sequence data resolves deep splits in the phylogeny of cuckoo wasps (Hymenoptera: Chrysididae). *Insect Systematics and Diversity*, 5, 1. https://doi.org/10.1093/isd/ixaa018

Peng, Y., Leung, H.C.M., Yiu, S.M. & Chin, F.Y.L. (2012) IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, 28, 1420–1428. https://doi.org/10.1093/bioinformatics/bts174

Pérez-Bañón, C., Petanidou, T. & Marcos-García, M.Á. (2007) Pollination in small islands by occasional visitors: the case of *Daucus carota* subsp. *commutatus* (Apiaceae) in the Columbretes archipelago, Spain. *Plant Ecology*, 192, 133–151. https://doi.org/10.1007/s11258-006-9233-1

Pérez-Bañón, C., Rojas, C., Vargas, M., Mengual, X. & Rojo, S. (2020) A world review of reported myiases caused by flower flies (Diptera: Syrphidae), including the first case of human myiasis from *Palpada scutellaris* (Fabricius, 1805). *Parasitology Research*, 119, 815–840. https://doi.org/10.1007/s00436-020-06616-4

Pérez-Bañón, C., Rotheray, G., Hancock, E., Marcos-García, M.Á. & Zumbado, M.A. (2003) Immature stages and breeding sites of some Neotropical saprophagous syrphids (Diptera: Syrphidae). *Annals of the Entomological Society of America*, 96, 458–471. https://doi.org/10.1603/0013-8746(2003)096[0458:ISABSO]2.0.CO;2

Pérez-Lachaud, G., Jervis, M.A., Reemer, M. & Lachaud, J.-P. (2014) An unusual, but not unexpected, evolutionary step taken by syrphid flies: the first record of true primary parasitoidism of ants by Microdontinae. *Biological Journal of the Linnean Society*, 111, 462–472. https://doi.org/10.1111/bij.12220

Peters, R.S., Krogmann, L., Mayer, C., Donath, A., Gunkel, S., Meusemann, K. et al. (2017) Evolutionary history of the hymenoptera. *Current Biology*, 27(7), 1013–1018. https://doi.org/10.1016/j.cub.2017.01.027

Petersen, M., Meusemann, K., Donath, A., Dowling, D., Liu, S., Peters, R.S. et al. (2017) Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. *BMC Bioinformatics*, 18, 111. https://doi.org/10.1186/s12859-017-1529-8

Philippe, H., Delsuc, F., Brinkmann, H. & Lartillot, N. (2005) Phylogenomics. *Annual Review of Ecology, Evolution, and Systematics*, 36(1), 541–562. https://doi.org/10.1146/annurev.ecolsys.35.112202.130205

Potts, S., Imperatriz-Fonseca, V., Ngo, H., Aizen, M.A., Biesmeijer, J.C., Breeze, T.D. et al. (2016) Safeguarding pollinators and their values to human well-being. *Nature*, 540, 220–229. https://doi.org/10.1038/nature20588

Quek, R.Z.B., Jain, S.S., Neo, M.L., Rouse, G.W. & Huang, D. (2020) Transcriptome-based target-enrichment baits for stony corals (Cnidaria: Anthozoa: Scleractinia). *Molecular Ecology Resources*, 20, 807–818. https://doi.org/10.1111/1755-0998.13150

R Core Team. (2021) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available from: https://www.R-project.org/. [Accessed 10 February 2022]

Rader, R., Bartomeus, I., Garibaldi, L.A., Garratt, M.P.D., Howlett, B.G., Winfree, R. et al. (2016) Non-bee insects are important contributors to global crop pollination. *Proceedings of the National Academy of Sciences of the United States of America*, 113, 146–151. https://doi.org/10.1073/pnas.1517092112

Rader, R., Cunningham, S.A., Howlett, B.G. & Inouye, D.W. (2020) Non-bee insects as visitors and pollinators of crops: biology, ecology, and

management. *Annual Review of Entomology*, 65(1), 391–407. https://doi.org/10.1146/annurev-ento-011019-025055

Rambaut, A. (2009) *FigTree v. 1.4.4: Tree Figure Drawing Tool*. Available from: https://github.com/rambaut/figtree [Accessed 10 February 2022].

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior summarisation in Bayesian phylogenetics using tracer 1.7. *Systematic Biology*, 67(5), 901–904. https://doi.org/10.1093/sysbio/syy032

Rannala, B. & Yang, Z. (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, 164(4), 1645–1656. https://doi.org/10.1093/genetics/164.4.1645

Reemer, M. (2013) Review and phylogenetic evaluation of associations between Microdontinae (Diptera: Syrphidae) and ants (Hymenoptera: Formicidae). *Psyche*, 2013, 1–9. https://doi.org/10.1155/2013/538316

Reemer, R. & Rotheray, G.E. (2009) Pollen feeding larvae in the presumed predatory syrphine genus *Toxomerus* Macquart (Diptera, Syrphidae). *Journal of Natural History*, 43, 939–949. https://doi.org/10.1080/00222930802610576

Reemer, M. & Ståhls, G. (2013) Phylogenetic relationships of Microdontinae (Diptera: Syrphidae) based on molecular and morphological characters. *Systematic Entomology*, 38, 661–688. https://doi.org/10.1111/syen.12020

Ricarte, A., Marcos-García, M.Á. & Moreno, C.E. (2011) Assessing the effects of vegetation type on hoverfly (Diptera: Syrphidae) diversity in a Mediterranean landscape: implications for conservation. *Journal of Insect Conservation*, 15(6), 865–877. https://doi.org/10.1007/s10841-011-9384-9

Ricarte, A., Souba-Dols, G.J., Hauser, M. & Marcos-García, M.Á. (2017) A review of the early stages and host plants of the genera *Eumerus* and *Merodon* (Diptera: Syrphidae), with new data on four species. *PLoS One*, 12, e0189852. https://doi.org/10.1371/journal.pone.0189852

Rizza, A., Campobasso, G., Dunn, P.H. & Stazi, M. (1988) *Cheilosia corydon* (Diptera: Syrphidae), a candidate for the biological control of musk thistle in North America. *Annals of the Entomological Society of America*, 81, 225–232. https://doi.org/10.1093/aesa/81.2.225

Roch, S. & Steel, M. (2015) Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. *Theoretical Population Biology*, 100, 56–62. https://doi.org/10.1016/j.tpb.2014.12.005

Rojo, S., Gilbert, F., Marcos-García, M.Á., Nieto, J.M. & Mier, M.P. (2003) *A world review of predatory hoverflies (Diptera, Syrphidae: Syrphinae) and their prey*. Alicante: Ediciones CIBIO (Centro Iberoamericano de la Biodiversidad).

Rojo, S. & Marcos-García, M.Á. (1997) Syrphid predators (Dipt.: Syrphidae) of gall forming aphids (Hom.: Aphididae) in Mediterranean areas: implications for biological control of fruit trees pests. *Entomophaga*, 42, 269–276. https://doi.org/10.1007/BF02769903

Rokas, A., Krüger, D. & Carroll, S.B. (2005) Animal evolution and the molecular signature of radiations compressed in time. *Science*, 310(5756), 1933–1938. https://doi.org/10.1126/science.1116759

Rota-Stabelli, O., Lartillot, N., Philippe, H. & Pisani, D. (2013) Serine codon-usage bias in deep phylogenomics: Pancrustacean relationships as a case study. *Systematic Biology*, 62(1), 121–133. https://doi.org/10.1093/sysbio/sys077

Rotheray, G.E. (1993) *Colour guide to hoverfly larvae (Diptera, Syrphidae) in Britain and Europe*. Sheffield: Dipterists Digest.

Rotheray, G.E. (2003) The predatory larvae of two *Nepenthosyrphus* species living in pitcher plants (Diptera, Syrphidae). *Studia Dipterologica*, 10, 219–226.

Rotheray, G.E. & Gilbert, F.S. (1989) The phylogeny and systematics of European predacious Syrphidae (Diptera) based on larval and puparial stages. *Zoological Journal of the Linnean Society*, 95, 29–70. https://doi.org/10.1111/j.1096-3642.1989.tb02222.x

Rotheray, G.E. & Gilbert, F. (1999) Phylogeny of Palaearctic Syrphidae (Diptera): evidence from larval stages. *Zoological Journal of the Linnean Society*, 127(1), 1–112. https://doi.org/10.1111/j.1096-3642.1999.tb01305.x

Rotheray, G.E. & Gilbert, F. (2008) Phylogenetic relationships and the larval head of the lower Cyclorrhapha (Diptera). *Zoological Journal of the Linnean Society*, 153, 287–323. https://doi.org/10.1111/j.1096-3642.2008.00395.x

Rotheray, G.E. & Gilbert, F. (2011) *The natural history of hoverflies*. Ceredigion: Forrest Text.

Rotheray, G.E., Zumbado, M., Hancock, E.G. & Thompson, F.C. (2000) Remarkable aquatic predators in the genus *Ocyptamus* (Diptera, Syrphidae). *Studia dipterologica*, 7(2), 385–389.

Rozo-Lopez, P. & Mengual, X. (2015) Mosquito species (Diptera, Culicidae) in three ecosystems from the Colombian Andes: identification through DNA barcoding and adult morphology. *ZooKeys*, 513, 39–64. https://doi.org/10.3897/zookeys.513.9561

Rupp, L. (1989) *Die mitteleuropäischen Arten der Gattung Volucella (Diptera, Syrphidae) als Kommensalen und Parasitoide in den Nestern von Hummeln und sozialen Wespen: Untersuchungen zur Wirtsfindung, Larvalbiologie und Mimikry*. PhD dissertation, Albert-Ludwigs University, Freiburg-im-Breisgau.

Sabatinelli, G., Eberle, J., Fabrizi, S. & Ahrens, D. (2020) A molecular phylogeny of Glaphyridae (Coleoptera: Scarabaeoidea): evolution of pollination and association with 'poppy guild' flowers. *Systematic Entomology*, 45, 838–848. https://doi.org/10.1111/syen.12429

Sabrosky, C.W. (1999) Annotated catalog of family-group names, alphabetical by type genera. *Myia*, 10, 35–326.

Sann, M., Niehuis, O., Peters, R.S., Mayer, C., Kozlov, A., Podsiadlowski, L. et al. (2018) Phylogenomic analysis of Apoidea sheds new light on the sister group of bees. *BMC Evolutionary Biology*, 18(1), 71. https://doi.org/10.1186/s12862-018-1155-8

Sauquet, H., Ramírez-Barahona, S. & Magallón, S. (2022) The age of flowering plants is unknown. *EcoEvoRxiv*. https://doi.org/10.32942/osf.io/n4v6b

Sayyari, E. & Mirarab, S. (2016) Fast coalescent-based computation of local branch support from quartet frequencies. *Molecular Biology and Evolution*, 33(7), 1654–1668. https://doi.org/10.1093/molbev/msw079

Schachat, S.R., Labandeira, C.C., Clapham, M.E. & Payne, J.L. (2019) A cretaceous peak in family-level insect diversity estimated with mark–recapture methodology. *Proceedings of the Royal Society B: Biological Sciences*, 286, 20192054. https://doi.org/10.1098/rspb.2019.2054

Seo, T.K. (2008) Calculating bootstrap probabilities of phylogeny using multilocus sequence data. *Molecular Biology and Evolution*, 25(5), 960–971. https://doi.org/10.1093/molbev/msn043

Sheppard, A.W., Aeschlimann, J.-P., Sagliocco, J.-L. & Vitou, J. (1995) Below-ground herbivory in *Carduus nutans* (Asteraceae) and the potential for biological control. *Biocontrol Science and Technology*, 5(3), 261–270. https://doi.org/10.1080/09583159550039729

Simmons, M.P. & Kessenich, J. (2020) Divergence and support among slightly suboptimal likelihood gene trees. *Cladistics*, 36, 322–340. https://doi.org/10.1111/cla.12404

Skevington, J.H., Locke, M.M., Young, A.D., Moran, K., Crins, W.J. & Marshall, S.A. (2019) *Field guide to the flower flies of northeastern North America*. Princeton: Princeton University Press.

Skevington, J.H. & Yeates, D.K. (2000) Phylogeny of the Syrphoidea (Diptera) inferred from mtDNA sequences and morphology with particular reference to classification of the Pipunculidae (Diptera). *Molecular Phylogenetics and Evolution*, 16(2), 212–224. https://doi.org/10.1006/mpev.2000.0787

Smith, S.A., Moore, M.J., Brown, J.W. & Yang, Y. (2015) Analysis of phylogenomic datasets reveals conflict, concordance, and gene

duplications with examples from animals and plants. *BMC Evolutionary Biology*, 15, 150. https://doi.org/10.1186/s12862-015-0423-0

Sommaggio, D. (1999) Syrphidae: can they be used as environmental bioindicators? *Agriculture, Ecosystems & Environment*, 74, 343–356. https://doi.org/10.1016/S0167-8809(99)00042-0

Souba-Dols, G.J., Ricarte, A., Hauser, M., Speight, M. & Marcos-García, M.Á. (2020) What do *Eumerus* Meigen larvae feed on? New immature stages of three species (Diptera: Syrphidae) breeding in different plants. *Organisms Diversity & Evolution*, 20, 267–284. https://doi.org/10.1007/s13127-020-00437-0

Speight, M.C.D. (1987) External morphology of adult Syrphidae (Diptera). *Tijdschrift voor Entomologie*, 130, 141–175 https://biostor.org/reference/49928

Speight, M.C.D. (2020) Species accounts of European Syrphidae, 2020. In: *Syrph the net, the database of European Syrphidae (Diptera)*, Vol. 104. Dublin: Syrph the Net publications, pp. 1–314.

Ssymank, A., Jordaens, K., De Meyer, M., Reemer, M. & Rotheray, G.E. (2021) Syrphidae (flower flies or hoverflies). In: Kirk-Spriggs, A.H. & Sinclair, B.J. (Eds.) *Manual of Afrotropical Diptera. Volume 3. Brachycera—Cyclorrhapha, excluding Calyptratae*. Pretoria: South African National Biodiversity Institute, pp. 1439–1491.

Ssymank, A. & Kearns, C. (2009) Flies-pollinators on two wings. In: Ssymank, A., Hamm, A. & Vischer-Leopold, M. (Eds.) *Caring for pollinators – safeguarding agro-biodiversity and wild plant diversity*. Bundesamt für Naturschutz: Bonn, pp. 39–52.

Ssymank, A., Kearns, C.A., Pape, T. & Thompson, F.C. (2008) Pollinating flies (Diptera): a major contribution to plant diversity and agricultural production. *Biodiversity*, 9, 86–89.

Ståhls, G., Hippa, H., Rotheray, G., Muona, J. & Gilbert, F. (2003) Phylogeny of Syrphidae (Diptera) inferred from combined analysis of molecular and morphological characters. *Systematic Entomology*, 28, 433–450. https://doi.org/10.1046/j.1365-3113.2003.00225.x

Stringer, D.N., Bertozzi, T., Meusemann, K., Delean, S., Guzik, M.T., Tierney, S.M. et al. (2021) Development and evaluation of a custom bait design based on 469 single-copy protein-coding genes for exon capture of isopods (Philosciidae: *Haloniscus*). *PLoS One*, 16(9), e0256861. https://doi.org/10.1371/journal.pone.0256861

Stuke, J.-H. (2000) Phylogenetische Rekonstruktion der Verwandtschaftsbeziehungen innerhalb der Gattung *Cheilosia* Meigen, 1822 anhand der Larvenstadien (Diptera: Syrphidae). *Studia Dipterologica Supplement*, 8, 1–118.

Suvorov, A., Scornavacca, C., Fujimoto, M.S., Bodily, P., Clement, M., Crandall, K.A. et al. (2021) Deep ancestral introgression shapes evolutionary history of dragonflies and damselflies. *Systematic Biology*, 71, 526–546. https://doi.org/10.1093/sysbio/syab063

Tenhumberg, B. (1995) Estimating predatory efficiency of *Episyrphus balteatus* (Diptera: Syrphidae) in cereal fields. *Environmental Entomology*, 24(3), 687–691. https://doi.org/10.1093/ee/24.3.687

Thompson, F.C. (1969) A new genus of Microdontine flies (Diptera: Syrphidae) with notes on the placement of the subfamily. *Psyche*, 76, 74–85. https://doi.org/10.1155/1969/62102

Thompson, F.C. (1972) A contribution to a generic revision of the Neotropical Milesinae (Diptera: Syrphidae). *Arquivos de Zoologia, Sao Paulo*, 23, 73–215. https://doi.org/10.11606/issn.2176-7793.v23i2p73-215

Thompson, F.C. (2017) Two new flower fly groups from the orient (Diptera: Syrphidae) with the description of a new species. *Entomologist's monthly Magazine*, 153, 171–178.

Thompson, F.C. (2020) Ant and flower flies: two different groups. *Fly Times*, 64, 24–25.

Tompsett, A. (2002) Narcissus: investigations into the control of large narcissus Fly (*Merodon equestris* [F]) using non-chemical methods. *Acta Horticulturae*, 570, 391–394. https://doi.org/10.17660/ActaHortic.2002.570.55

Tot, T., Radenković, S., Nedeljković, Z., Likov, L. & Vujić, A. (2020) Descriptions of two new species of the genus *Paragus* Latreille (Diptera: Syrphidae), with a key to males of all South African species. *Zootaxa*, 4780(2), 341–355. https://doi.org/10.11646/zootaxa.4780.2.7

Trautwein, M.D., Wiegmann, B.M. & Yeates, D.K. (2011) Overcoming the effects of rogue taxa: evolutionary relationships of the bee flies. *PLoS currents*, 3, RRN1233. https://doi.org/10.1371/currents.RRN1233

Tscharntke, T., Klein, A.M., Kruess, A., Steffan-Dewenter, I. & Thies, C. (2005) Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. *Ecology Letters*, 8(8), 857–874. https://doi.org/10.1111/j.1461-0248.2005.00782.x

Turner, E.H., Ng, S.B., Nickerson, D.A. & Shendure, J. (2009) Methods for genomic partitioning. *Annual Review of Genomics and Human Genetics*, 10, 263–284. https://doi.org/10.1146/annurev-genom-082908-150112

Ureña, O. & Hanson, P. (2010) A fly larva (Syrphidae: *Ocyptamus*) that preys on adult flies. *Revista de Biología Tropical*, 58(4), 1157–1163.

van der Kooi, C.J. & Ollerton, J. (2020) The origins of flowering plants and pollinators. *Science*, 368(6497), 1306–1308. https://doi.org/10.1126/science.aay3662

Vasilikopoulos, A., Balke, M., Beutel, R.G., Donath, A., Podsiadlowski, L., Pflug, J.M. et al. (2019) Phylogenomics of the superfamily Dytiscoidea (Coleoptera: Adephaga) with an evaluation of phylogenetic conflict and systematic error. *Molecular Phylogenetics and Evolution*, 135, 270–285. https://doi.org/10.1016/j.ympev.2019.02.022

Vasilikopoulos, A., Gustafson, G.T., Balke, M., Niehuis, O., Beutel, R.G. & Misof, B. (2021) Resolving the phylogenetic position of Hygrobiidae (Coleoptera: Adephaga) requires objective statistical tests and exhaustive phylogenetic methodology: a response to Cai et al. (2020). *Molecular Phylogenetics and Evolution*, 162, 106923. https://doi.org/10.1016/j.ympev.2020.106923

Vasimuddin, M., Misra, S., Li, H. & Aluru, S. (2019) Efficient architecture-aware acceleration of BWA-MEM for multicore systems. *IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 2019, 314–324. https://doi.org/10.1109/IPDPS.2019.00041

Vockeroth, J.R. (1969) A revision of the genera of the Syrphini (Diptera, Syrphidae). *Memoirs of the Entomological Society of Canada*, 101(Supplement 62), 1–176. https://doi.org/10.4039/entm10162fv

Vockeroth, J.R. (1992) *The flower flies of the subfamily Syrphinae of Canada, Alaska, and Greenland (Diptera: Syrphidae)*. Ottawa: The Insects and Arachnids of Canada, Part 18. Centre for Land and Biological Resources Research.

Vockeroth, J.R. & Thompson, F.C. (1987) Family Syrphidae. In: McAlpine, J.F. (Ed.) *Manual of Nearctic Diptera*, Vol. 2. Quebec: Research Branch, Agriculture Canada Monograph 28, pp. 713–743.

Vujić, A., Ståhls, G., Ačanski, J., Bartsch, H., Bygebjerg, R. & Stefanović, A. (2013) Systematics of Pipizini and taxonomy of European *Pipiza* Fallén: molecular and morphological evidence (Diptera, Syrphidae). *Zoologica Scripta*, 42(3), 288–305. https://doi.org/10.1111/zsc.12005

Vujić, A., Ståhls, G., Rojo, S., Radenković, S. & Šimić, S. (2008) Systematics and phylogeny of the tribe Paragini (Diptera: Syrphidae) based on molecular and morphological characters. *Zoological Journal of the Linnean Society*, 152(3), 507–536. https://doi.org/10.1111/j.1096-3642.2007.00353.x

Wang, Y.H., Engel, M.S., Rafael, J.A., Wu, H., Rédei, D., Xie, Q. et al. (2016) Fossil record of stem groups employed in evaluating the chronogram of insects (Arthropoda: Hexapoda). *Scientific Reports*, 6, 38939. https://doi.org/10.1038/srep38939

Ware, J.L. & Barden, P. (2016) Incorporating fossils into hypotheses of insect phylogeny. *Current Opinion in Insect Science*, 18, 69–76. https://doi.org/10.1016/j.cois.2016.10.003

Warnock, R.C.M., Yang, Z. & Donoghue, P.C.J. (2012) Exploring uncertainty in the calibration of the molecular clock. *Biology Letters*, 8, 156–159. https://doi.org/10.1098/rsbl.2011.0710

Warnock, R., Yang, Z. & Donoghue, P. (2017) Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. *Proceedings of the Royal Society. Biological Sciences*, 284(1857), 20170227. https://doi.org/10.1098/rspb.2017.0227

Warnow, T. (2015) Concatenation analyses in the presence of incomplete lineage sorting. *PLoS currents*, 7. https://doi.org/10.1371/currents.tol.8d41ac0f13d1abedf4c4a59f5d17b1f7

Weng, J.L. & Rotheray, G.E. (2009) Another non-predaceous syrphine flower fly (Diptera: Syrphidae): pollen feeding in the larva of *Allograpta micrura*. *Studia Dipterologica*, 15(2008), 245–258.

Wheeler, T.J. & Eddy, S.R. (2013) Nhmmer: DNA homology search with profile HMMs. *Bioinformatics*, 29, 2487–2489. https://doi.org/10.1093/bioinformatics/btt403

Whitfield, J.B. & Kjer, K.M. (2008) Ancient rapid radiations of insects: challenges for phylogenetic analysis. *Annual Review of Entomology*, 53, 449–472. https://doi.org/10.1146/annurev.ento.53.103106.093304

Wiegmann, B.M., Trautwein, M.D., Winkler, I.S., Barr, N.B., Kim, J.-W., Lambkin, C. et al. (2011) Episodic radiations in the fly tree of life. *Proceedings of the National Academy of Sciences*, 108(14), 5690–5695. https://doi.org/10.1073/pnas.1012675108

Wiegmann, B.M., Yeates, D.K., Thorne, J.L. & Kishino, H. (2003) Time flies, a new molecular time-scale for brachyceran fly evolution without a clock. *Systematic Biology*, 52(6), 745–756. https://doi.org/10.1093/sysbio/52.6.745

Willmer, P. (2011) *Pollination and floral ecology*. Princeton: Princeton University Press.

Winkler, I.S., Blaschke, J.D., Davis, D.J., Stireman, J.O., III, O'Hara, J.E., Cerretti, P. et al. (2015) Explosive radiation or uninformative genes? Origin and early diversification of tachinid flies (Diptera: Tachinidae). *Molecular Phylogenetics and Evolution*, 88, 38–54. https://doi.org/10.1016/j.ympev.2015.03.021

Wotton, K.R., Gao, B., Menz, M.H.M., Morris, R.K.A., Ball, S.G., Lim, K.S. et al. (2019) Mass seasonal migrations of hoverflies provide extensive pollination and crop protection services. *Current Biology*, 29(13), 2167–2173. https://doi.org/10.1016/j.cub.2019.05.036

Yan, L., Buenaventura, E., Pape, T., Kutty, S.N., Bayless, K.M. & Zhang, D. (2021) A phylotranscriptomic framework for flesh fly evolution (Diptera, Calyptratae, Sarcophagidae). *Cladistics*, 37, 540–558. https://doi.org/10.1111/cla.12449

Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24(8), 1586–1591. https://doi.org/10.1093/molbev/msm088

Yang, Z. & Rannala, B. (2006) Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution*, 23(1), 212–226. https://doi.org/10.1093/molbev/msj024

Young, A.D. & Gillung, J.P. (2020) Phylogenomics—principles, opportunities and pitfalls of big-data phylogenetics. *Systematic Entomology*, 45(2), 225–247. https://doi.org/10.1111/syen.12406

Young, A.D., Lemmon, A.R., Skevington, J.H., Mengual, X., Ståhls, G., Reemer, M. et al. (2016) Anchored enrichment dataset for true flies (order Diptera) reveals insights into the phylogeny of flower flies (family Syrphidae). *BMC Evolutionary Biology*, 16, 143. https://doi.org/10.1186/s12862-016-0714-0

Young, A.D., Marshall, S.A. & Skevington, J.H. (2016) Revision of *Platycheirus* Lepeletier and Serville (Diptera: Syrphidae) in the Nearctic north of Mexico. *Zootaxa*, 4082(1), 1–317. https://doi.org/10.11646/zootaxa.4082.1.1

Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. (2018) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, 19, 153. https://doi.org/10.1186/s12859-018-2129-y

Zhang, Y.M., Williams, J.L. & Lucky, A. (2019) Understanding UCEs: a comprehensive primer on using ultraconserved elements for arthropod phylogenomics. *Insect Systematics and Diversity*, 3(5), 3. https://doi.org/10.1093/isd/ixz016

Zuijen, M.P. & Nishida, K. (2011) Description of life history and immature stages of phytophagous flower fly, *Allograpta zumbadoi* Thompson (Diptera: Syrphidae: Syrphinae). *Studia Dipterologica*, 17, 37–51.

SUPPORTING INFORMATION

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

Figures S1–S8. Syrphidae topologies inferred using different data sets and methods.

Figures S9–S14. Syrphidae topologies inferred using different data sets and methods.

Figures S15–S26. SYMTEST heat maps of the Bowker's test using different data sets.

Figures S27–S28. Dated phylogenies using BEAST and MCMCTREE.

Table S1. Information about the specimens studied in the present work.

Table S2. Fossils used as calibration points in the dated analyses.

Table S3. Enrichment information and statistics for the samples included in this study.

Table S4. Assembly information for the Syrphidae specimens studied in the present work.

Table S5. Statistics for the OG included in this study, with and without transcriptomic data.

How to cite this article: Mengual, X., Mayer, C., Burt, T.O., Moran, K.M., Dietz, L., Nottebrock, G. et al. (2023) Systematics and evolution of predatory flower flies (Diptera: Syrphidae) based on exon-capture sequencing. *Systematic Entomology*, 48(2), 250–277. Available from: https://doi.org/10.1111/syen.12573