ORIGINAL ARTICLE

Integration of mitogenomic and morphological data disentangles the systematics of *Pollenia* and establishes a revised phylogenetic hypothesis for the Polleniidae

Nikolas P. Johnston[1,2,3] | Marcin Piwczyński[2] | Paulina Trzeciak[2] |
Kinga Walczak[2] | Krzysztof Szpila[2]

[1]School of Life Sciences, University of Technology Sydney, Ultimo, New South Wales, Australia

[2]Department of Ecology and Biogeography, Faculty of Biological and Veterinary Sciences, Nicolaus Copernicus University in Toruń, Toruń, Poland

[3]Centre for Sustainable Ecosystem Solutions, School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, New South Wales, Australia

Correspondence
Nikolas P. Johnston and Krzysztof Szpila, Department of Ecology and Biogeography, Faculty of Biological and Veterinary Sciences, Nicolaus Copernicus University in Toruń, Lwowska 1, Toruń, Poland.
Email: nikolas.johnston@uts.edu.au, szpila@umk.pl

Funding information
Polish National Science Centre

Abstract

The Polleniidae (Diptera) are a family of flies best known for species of the genus *Pollenia*, which overwinter inside human dwellings. Previously divided across the Calliphoridae, Tachinidae and Rhinophoridae, the polleniid genera have only recently been united. Several studies have utilized molecular data to analyse polleniid phylogenetic relationships, although all have suffered from low taxon sampling or insufficient phylogenetic signal in molecular markers. To alleviate these problems, we utilized two automated organellar genome extraction software, GetOrganelle and MitoFinder, to assemble mitogenomes from genome skimming data from 22 representatives of the polleniid genera: *Dexopollenia*, *Melanodexia*, *Morinia*, *Pollenia* and *Xanthotryxus*. From these analyses, we provide 14 new mitogenomes for the Polleniidae and perform phylogenetic analyses of 13 protein-coding mitochondrial genes using both maximum likelihood and Bayesian inference. Subfamilial phylogenetic relationships within the Polleniidae are interrogated and *Pollenia* is found to form a monophyletic clade sister to *Melanodexia*, *Morinia* and *Dexopollenia*, providing no evidence for the synonymisation of any of these genera. Our topology conflicts with previous morphology-based cladistic interpretations, with the *amentaria*, *griseotomentosa*, *semicinerea* and *viatica* species-groups resolving as non-monophyletic. We provide support for our topology through analysis of adult morphology and male and female terminalia, while identifying new diagnostic characters for some of the clades of the *Pollenia*. To test the validity of the current diagnostic morphology in the Polleniidae, newly assembled cytochrome C oxidase subunit 1 (COI) data are combined with a polleniid COI barcode reference library and analysed using the species delimitation software ASAP. COI barcodes support the current morphologically defined species within the *Pollenia*.

KEYWORDS
DNA barcoding, genome skimming, phylogenomics, *Pollenia*, Polleniidae

INTRODUCTION

Members of the oestroid fly family Polleniidae (Diptera) are known as 'cluster flies' due to the habit of some species to overwinter *en masse* in dark enclosed spaces such as attics, cupboards, or window frames. Polleniid flies are abundant, broadly distributed, and important plant pollinators (Howlett et al., 2016; Rader et al., 2020; Vezsenyi et al., 2022). The life history of the preimaginal stages of many polleniid species is unknown, however several species have been recognized as parasitoids of earthworms (El-Husseini, 2019; Szpila, 2003a; Tawfik & El-Husseini, 1971; Yahnke & George, 1972). Some polleniid species (i.e., *P. ibalia* Séguy and *P. vagabunda* [Meigen]) are also thought to be associated with caterpillars and other lepidopteran larvae, but this has yet to be formally documented (Rognes, 1991a, 2010).

The most recent checklist of the global Polleniidae (Gisondi et al., 2020) summarized 147 polleniid species across nine genera: *Pollenia* Robineau-Desvoidy (95 spp.), *Alvamaja* Rognes (1 sp.), *Dexopollenia* Townsend (21 spp.), *Melanodexia* Williston (8 spp.), *Morinia* Robineau-Desvoidy (13 spp.), *Nesodexia* Villeneuve (1 sp.), *Xanthotryxus* Aldrich (7 spp.) and the two genera tentatively assigned to the Polleniidae: *Anthracomyza* Malloch (1 sp.) and *Wilhelmina* Villeneuve (1 sp.). An additional species of *Pollenia* from Austria has been described since the publication of the checklist (Schluesslmayr & Sivell, 2021).

The polleniid genera were historically divided across the Calliphoridae, Rhinophoridae and Tachinidae, until the phylogenetic study of Cerretti et al. (2019), who used morphological data and three genetic loci to unify all genera and to establish the Polleniidae family status. The phylogenetic hypothesis of Cerretti et al. (2019) placed Polleniidae as sister to the Tachinidae, a relationship corroborated by several other molecular phylogenetic studies (Cerretti et al., 2017, Blaschke et al., 2018, Kutty et al., 2019, Stireman et al., 2019, Buenaventura et al., 2020, Yan et al., 2021).

While the family status of the Polleniidae is now established, research evaluating the subfamilial phylogenetic relationships is limited. Recent phylogenetic hypotheses for the family have suffered from limited taxon sampling and, in most cases, phylogenetic trees contain only a small representation of species from a few genera. For example, the family-level polleniid phylogeny of Cerretti et al. (2019) contained only six species from three genera (*Morinia* – 4 spp., *Alvamaja* – 1 sp. and *Pollenia* – 1 sp.), of which only one species each of *Morinia* and *Pollenia* were represented by molecular data. This study is the most comprehensive polleniid phylogenetic analysis to date, and the positions of all polleniid genera, except *Morinia*, *Pollenia* and *Alvamaja*, are yet to be tested. Furthermore, while there is some morphological evidence for the unification of all polleniid genera into a family (Cerretti et al., 2019), the current family level diagnosis is weak and based on a combination of over ten morphological characters states, including several taken from the male and female terminalia (Gisondi et al., 2020). It should also be noted that Gisondi et al. (2020) highlighted the uncertainty of some of these diagnostic characters (i.e., 'size' or 'ground colour') by including broad character states (i.e., 'small- to medium-sized' and 'yellow to black', respectively).

Pollenia species-group system

In contrast to the Polleniidae as a whole, phylogenetic relationships within the genus *Pollenia* are better understood. *Pollenia*, especially the West Palearctic representatives, has been the subject of several revisionary works (Jacentkovskỳ, 1941, 1942; Zumpt 1956; Lehrer 1963, 1967; Mihályi 1976; Rognes, 1987a, 1987b, 1988, 1991b, 1992a, 1992b and 2010), which resulted in a species-group classification system as well as two comprehensive phylogenetic studies: one using morphological data (Rognes, 1992a) and most recently using molecular data (Szpila et al., 2022). The *Pollenia* species-group system, established for West Palearctic species, is based on 11 species-groups (composed of 1–9 species per group) based on shared morphology (summarized in Table 1), developed by Rognes across a series of revisions (Rognes, 1987a, 1987b, 1988, 1991b, 1992a, 1992b and 2010). Within this system, because of the taxonomic complexity of *Pollenia* (Dear, 1986; Rognes, 1992a), most species-groups were defined by Rognes based on a combination of many morphological characters (up to 40), including those of the male and female terminalia. Rognes' also produced a cladistic analysis of the *Pollenia* species-groups based on shared morphological characters (Figure 1a), with each species-group represented as a monophyletic clade (Rognes, 1992a).

Szpila et al. (2022) tested Rognes' *Pollenia* species-group system through the application of molecular phylogenetics. Their phylogenetic study utilized three molecular markers (one mitochondrial and two nuclear) to reconstruct relationships among 18 West Palearctic species of *Pollenia*, representing eight of the ten species-groups. The results of this study only partly aligned with the morphology-based cladistic analysis of Rognes (1992a), with three species-groups—*rudis*, *tenuiforceps* and *viatica*—resolved as monophyletic, while two, *amentaria* and *griseotomentosa*, appeared to be paraphyletic. Other relationships between species-groups were consistent between the two analyses, except regarding the sister group to *vagabunda* (*viatica* in Rognes, 1992a and *amentaria* in Szpila et al. [2022]) and the position of *P. venturii* Zumpt (sister to *viatica* + *vagabunda* + *amentaria* + *haeretica* in Rognes [1992a] and sister to *viatica* in Szpila et al. [2022]).

Despite demonstrating molecular support for some infrageneric relationships in *Pollenia*, the phylogenetic hypothesis of Szpila et al. (2022) had limitations. First, it only used a low number of genetic loci and hence was unable to produce robust support for most clades. Second, it had limited taxon sampling: several species-groups were only represented by only a single species and no representatives of the *semicinerea*- and *haeretica*- groups nor any other polleniid genera were included (except for *Morinia doronici* [Scopoli], which was used as an outgroup).

An expanded phylogeny incorporating additional taxa and more robust genetic data would alleviate some of the limitations of Szpila et al., (2022) and provide a more robust framework to formally reassess the morphological characteristics originally used by Rognes (1992a) to delimit species and species-groups within the Polleniidae.
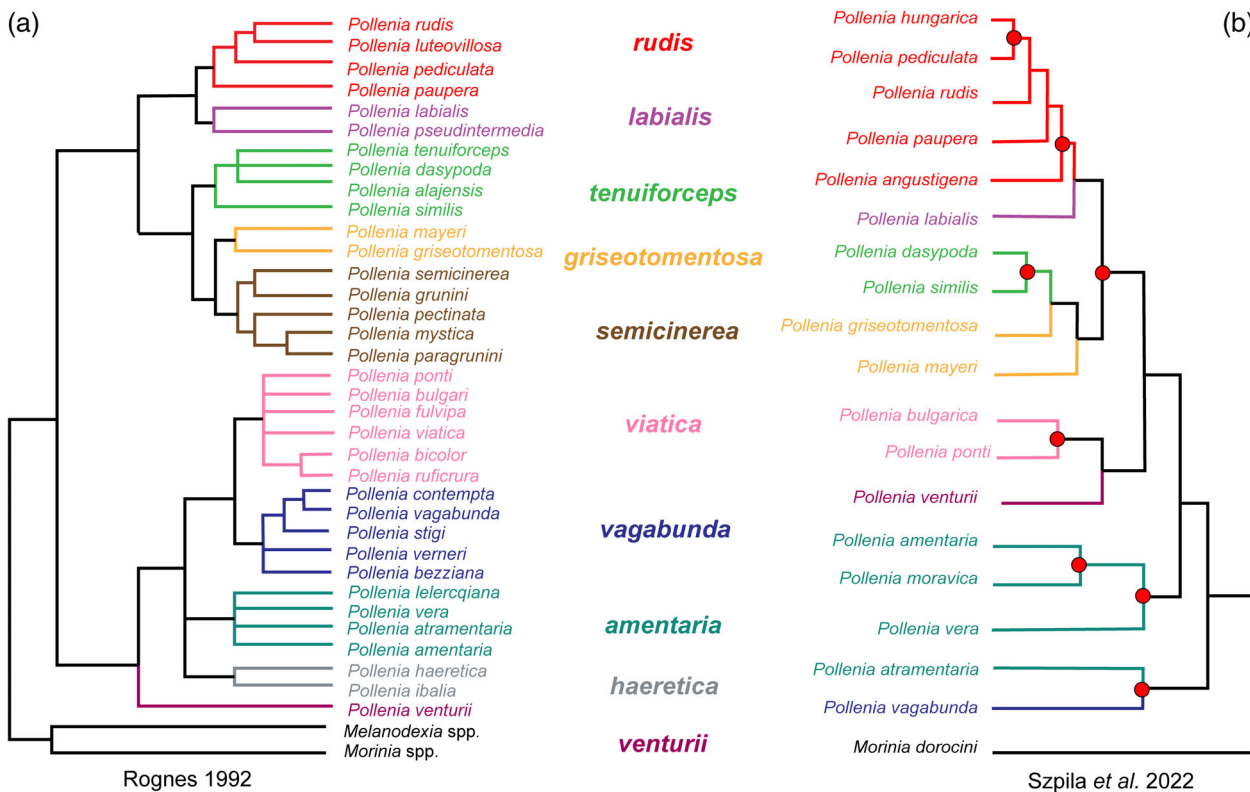
TABLE 1 Summary of the *Pollenia* species-group system.

Species-group	Constituent species	Relevant taxonomic citations
amentaria species-group	*P. amentaria* (Scopoli, 1763) *P. atramentaria* (Meigen, 1826) *P. leclercqiana* (Lehrer, 1978) *P. vera* Jacentkovský, 1936	Rognes (1992a), Szpila and Draber-Mońko (2008)
griseotomentosa species-group	*P. griseotomentosa* (Jacentkovský, 1944), *P. margarita* Schluesslmayr & Sivell, 2021 *P. mayeri* Jacentkovsky, 1941	Rognes (1988, 1992a), Schluesslmayr and Sivell (2021)
haeretica species-group	*P. haeretica* Séguy, 1928 *P. ibalia* Séguy, 1930	Rognes (1992a, 2010)
japonica species-group	*P. japonica* Kano & Shinonaga, 1966	Rognes (1992a)
labialis species-group (originally called *intermedia* species-group)	*P. labialis* Robineau-Desvoidy, 1863 *P. pseudointermedia* Rognes, 1987	Rognes (1987a)
rudis species species-group	*P. angustigena* Wainwright, 1940 *P. bartaki* Rognes, 2016 *P. hungarica* Rognes, 1987 *P. luteovillosa* Rognes, 1987 *P. paupera* Rondani, 1862 *P. pediculata* Macquart, 1834 *P. rudis* (Fabricius, 1794)	Rognes (1987b, 1992a, 2016)
semicinerea species-group	*P. agneteae* Rognes, 2019 *P. grunini* Rognes, 1988 *P. mystica* Rognes, 1988 *P. paragrunini* Rognes, 1988 *P. pectinata* Grunin, 1966 *P. semicinerea* Villeneuve, 1911	Rognes (1988, 2019)
tenuiforceps species-group	*P. alajensis* Rohdendorf, 1926, *P. dasypoda* Portschinsky, 1881, *P. similis* (Jacentkovskỳ, 1941), *P. tenuiforceps* Séguy, 1928	Rognes (1988, 2002), Szpila (2003b)
vagabunda species-group	*P. bezziana* Rognes, 1992 *P. contempta* Robineau-Desvoidy, 1863 *P. stigi* Rognes, 1992 *P. vagabunda* (Meigen, 1826) *P. verneri* Rognes, 1992	Rognes (1992b)
venturii species-group	*P. venturii* Zumpt, 1956	Rognes (1992a)
viatica species-group	*P. bicolor* Robineau-Desvoidy, 1830 *P. bulgarica* Jacentkovský, 1939 *P. fulvipalpis* Macquart, 1835 *P. mediterranea* Grunin, 1966 *P. mesopotamica* Mawlood & Abdul-Rassoul, 2009 *P. ponti* Rognes, 1991 *P. ruficrura* Rondani, 1862 *P. rufifemorata* Rognes & Baz, 2008 *P. viatica* Robineau-Desvoidy, 1830	Rognes (1991a, 1991b), Rognes and Baz (2008), Mawlood & Abdul-Rassoul (2009), Szpila et al. (2022)

Polleniid species delimitation

Rognes' morphological framework for the *Pollenia* is based on 43 characters, of which 22 referred to male and female terminalia (12 male and 10 female). Within this framework, the morphology of phallus can be seen as the best source of information for reliable definitions of new species, species-groups and larger clades in *Pollenia*. The diagnostic power of the male terminalia was also identified by Rognes' predecessors (Zumpt 1956, Lehrer 1963), as evidenced by previous infrageneric classifications being based primarily on the shape of the median hypophallic lobe and the lateral hypophallic lobes ('vesicae' sensu Zumpt, 'lobes hypophalliques ventraux' sensu Lehrer). Rognes also identified several other taxonomically important structures of the distiphallus, including the ventral plate (ventral sclerotization of the distiphallus), mesohypophallus and paraphallic processes (Rognes, 1992a). In the female terminalia, the shape and level of sclerotization of spermathecae and lateral sacs and the spinulation of ovipositor were also considered to be taxonomically important

INTEGRATIVE SYSTEMATICS OF POLLENIIDAE

Systematic
Entomology

Royal
Entomological
Society

299

FIGURE 1 Previous hypotheses for relationships between *Pollenia* species-groups. (a) Cladogram based on the morphological data of Rognes (1992a) generated using a maximum parsimony method. (b) Cladogram based on the three-locus data set of Szpila et al. (2022) generated using a maximum likelihood method. Red circles indicate branches with >85% bootstrap support. Species-group names are indicated between cladograms in bold italics. Taxa and corresponding branches in each cladogram are colour-coded by species-group.

(Lobanov, 1976; Rognes, 1992a). To test the validity of using these complex characters for the definition of species in addition to their phylogenetic analysis, Szpila et al. (2022) developed a COI barcode reference library for the Palearctic *Pollenia* by combining COI data from 18 *Pollenia* species with an additional 5405 polleniid COI sequences extracted from barcode of life database (BOLD). Szpila et al. (2022) utilized this combined reference library to test whether morphological species within *Pollenia* were also supported by molecular data. This analysis successfully delimited all included species, except for *P. moravica* (Jacentkovský), which was subsequently synonymized back with *P. amentaria* (Scopoli). While this analysis did provide molecular support for some of the morphologically defined species within the Polleniidae, it only included species from the genus *Pollenia* and did not include any Pollenia from outside the Palearctic.

Technical limitations to producing a robust and comprehensive polleniid phylogeny

A considerable limitation on the development of a robust phylogeny of the Polleniidae is the scarcity of high-quality preserved material, particularly of the less common or monotypic genera outside the well-sampled Palearctic region. Furthermore, the majority of polleniid species are difficult to sample in substantial numbers because they do not

readily come to baits or traps and their life history is unknown (Gisondi et al., 2020). The difficulty of finding fresh material has led to a bias in recent phylogenetic studies to include only the relatively common Palearctic polleniid species, usually representing the genera *Morinia* or *Pollenia* (i.e., Cerretti et al., 2019; Kutty et al., 2019; Szpila et al., 2022).

Resolving complex relationships in large organismal groups should rely on as broad sampling as possible from all biogeographic regions and the inclusion of species that are rare and often difficult to freshly collect. In this regard, museums and other preserved insect collections act as invaluable sources of biological material, harbouring specimens of many species that would otherwise be impossible to reliably collect for the purposes of genetic research (Yeates et al., 2016). A drawback of using this material for genetic purposes is that it is often of poor quality, due, at least in part, to storage conditions (i.e., long storage times and the use of DNA damaging preservatives) (Zimmermann et al., 2008; Espeland et al., 2010; Yeates et al., 2016). Despite the complications of using museum material, various sequencing techniques are available that can retrieve genomic data from low-quality or fragmented DNA (summarized in Table 2 of Raxworthy & Smith, 2021). One such technique is genome skimming, which involves sequencing genomic DNA at a low resolution with high-copy number genomic regions (such as mitochondrial genes) sequenced at a higher coverage than single-copy regions. The benefits of this

TABLE 2 Assemble species by automatic partitioning (ASAP) delimitation results.

Number of morpho species	Number of delimited species	ASAP-score	P-val (rank)	W (rank)	Threshold distance
31	33	2.50	5.75 e-03 (3)	2.50 e-04 (2)	0.019386
	33	4.50	4.29 e-02 (4)	7.82 e-05 (5)	0.021277
	38	5.00	4.55 e-03 (2)	6.78 e-05 (8)	0.007780
	37	5.50	2.44 e-01 (8)	9.40 e-05 (3)	0.009489
	29	6.00	1.23 e-03 (1)	3.14 e-05 (11)	0.038129
	36	8.00	1.66 e-01 (7)	6.21 e-05 (9)	0.010923

Note: ASAP-score – A summary statistic incorporating both *p*-val and W used to rank species partitions. *p*-val – The probability that each partition in the analysis contains a single species. W – The relative barcode gap width between the proposed partitioning scheme and the previous scheme. Threshold distance – The midpoint between the current distance that resulted in merging of partitions and the midpoint distance used in the previous partitioning scheme. *p*-val and W values are also ranked in their respective columns as indicated in parentheses.

technique are two-fold. First, by only sequencing at a low resolution, the cost and time associated with sequencing and assembly are dramatically reduced. Second, by aiming for only high-copy number regions, this technique is particularly robust against issues related to the use of museum material (i.e., DNA fragmentation and/or low concentrations of DNA).

The main aim of the present study was to utilize mitogenomes to infer phylogenetic relationships within the family Polleniidae. Automatic organellar assemblers were used to extract mitogenomes from genome skimming data from 24 species representing five polleniid genera. Protein-coding loci were then extracted from these mitogenomes and used to produce a phylogenetic hypothesis for the Polleniidae. This phylogeny was then used in combination with morphological data to validate both Rognes' *Pollenia* species-group system and the topology of the three loci molecular phylogeny of Szpila et al. ([2022](#)). Finally, freshly extracted COI barcode sequences were combined with the *Pollenia* COI reference library of Szpila et al. ([2022](#)) and analysed using the species delimitation software ASAP to test whether molecularly delimited species were congruent with those identified using morphology.

MATERIALS AND METHODS

Taxon sampling

Representative taxa of the Nearctic and Palearctic Polleniidae were selected for DNA extraction based on their availability in preserved museum material and to represent a broad diversity of genera (Table S1). In addition to this material, mitogenomes for *Pollenia angustigena* Wainwright (MT628554) and *Morinia doronici* (MT410784) were obtained from GenBank. The final taxa selection for phylogenetic analysis included 32 ingroup taxa (from 5 genera: 1 sp. *Dexopollenia*, 1 sp. *Melanodexia*, 2 spp. *Morinia*, 25 spp. *Pollenia*, 1 sp. *Xanthotryxus* and one representative from each of the families Calliphoridae and Tachinidae) and 1 outgroup taxon (a representative from the Sarcophagidae). For the purposes of the species delimitation analysis, COI sequences for an additional two unidentified polleniid

species were also obtained: a *Melanodexia* species (ASILO563-17) from GenBank and a *Pollenia* species from the dataset of Szpila et al., ([2022](#)) (KEIB_DIP_00037). The final COI dataset for delimitation analysis thus included a total of 32 representative polleniid species (from 5 genera: 1 sp. *Dexopollenia*, 2 spp. *Melanodexia*, 2 spp. *Morinia*, 26 spp. *Pollenia* 1 sp. *Xanthotryxus*).

DNA extraction

Genomic DNA was extracted from the dried legs of museum specimens or, in the case of *Lucilia sericata* (Meigen), an entire larva, using a DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA, U.S.A.) following the manufacturer's instructions. The isolated DNA was quantified with a Qubit 3.0 fluorometer using a dsDNA High Sensitivity Assay Kit (Life Technologies, Inc., Carlsbad, CA, U.S.A.). To control for the degree of DNA fragmentation, samples were analysed via gel electrophoresis using a 1% agarose gel stained with GelRed (Biotium, Darmstadt, Germany). To test the utility of the genome skimming method to generate data from poor quality samples, samples with a DNA concentration below the Qubit detection limit were not eliminated.

Library preparation and sequencing

Prior to library preparation, extracted genomic DNA was sheared to 100–400 bp with a Covaris sonicator (Covaris, Inc., Woburn, MA, U.S.A.) and then combined with 1X TE buffer to obtain a final volume of 50 μl. Library preparation was completed using a NEBNext Ultra II DNA Library Prep Kit (New England BioLabs Inc., Ipswich, MA, U.S.A.) following the manufacturer's instructions. Briefly, hairpin-loop NEBNext adaptors were ligated to the blunt-ended DNA fragments. In this step, we used three variants of adaptor dilution according to the manufacturer's recommendation: 15 μM of working adaptor concentration for 101–1000 ng, 1.5 μM for 5–100 ng and 0.6 μM for less than 5 ng of DNA. Afterwards, size selection using AMPure XP (Beckman Coulter, Carlsbad, CA, U.S.A.) was completed for samples with an input greater than 50 ng. The number of beads used for size selection was

chosen based on an assumed 200 bp insert size. A polymerase chain reaction (PCR) master mix was then prepared containing 15 μl of adaptor-ligated DNA fragments, 25 μl of NEBNext Ultra II Q5 Master Mix and 10 μl of NEBNext Multiplex Oligos for Illumina (New England BioLabs Inc., Ipswich, MA, U.S.A.). To ensure enough product was amplified for each library, each PCR reaction was modified by adjusting the number of cycles (12–30) to correspond with the concentration of adaptor-ligated DNA. PCR products were then purified using AMPure XP (0.9 × ratio of bead to sample volume) and the DNA concentration quantified using a Qubit 3.0 fluorometer. All 33 libraries were normalized by dilution with 0.1X TE buffer to achieve the same DNA concentration across all libraries.

All libraries were then pooled, and a final quantification of DNA concentration was completed using a Qubit 3.0 fluorometer and 2100 Bioanalyzer with a High Sensitivity DNA Analysis Kit (Agilent Technologies Inc., Santa Clara, CA, U.S.A.). To remove library fragments containing only adapters, we used Pippin Prep (Sage Science, Beverly, MA, U.S.A.) to select for fragments ranging from 170 to 400 bp. The final product was then purified with AMPure XP (1.8 × ratio of bead to sample volume) and re-suspended in 1X TE buffer. The final pooled library was then commercially sequenced (Macrogen) using one lane of an Illumina HiSeq 2500 to generate PE reads of 100 bp. Low coverage sequencing was performed at an approximate depth of 1X based on an estimated genome size of ∼450–550Mbp as previously determined for other calyptrate Diptera (*Phormia regina* Robineau-Desvoidy ∼550Mbp; Andere 2016, *Lucilia cuprina* [Wiedemann] ∼450Mbp; Andstead 2015).

Following sequencing, raw demultiplexed paired-end reads were assessed for quality using FastQC (Andrews, 2010). Trimmomatic v.0.36 (Bolger et al., 2014) was then used to remove adaptors, trim sequences and filter low quality reads, using the following parameters as recommended for paired end sequences (Bolger et al., 2014; supplementary data): leading base cut-off threshold = 3, trailing base cut-off threshold = 3, Sliding window trimming [size: quality threshold] = 4:15, Minimum sequence length = 36. Following trimming, forward and reverse paired read outputs from Trimmomatic were used for subsequent bioinformatic analyses.

Mitogenome extraction and assembly

Studies have shown that mitogenome completeness and accuracy can be affected by extraction and assembly methods (Allio et al., 2020; Jin et al., 2020; Timbó et al., 2017). To ensure the accuracy of our mitogenomes and to compare the efficacy of alternative automated mitogenome extraction and assembly software, three alternative approaches were used to process raw reads: (1) MitoFinder (Allio et al., 2020) with the metaSPADES assembler (Nurk et al., 2017), (2) Mitofinder using the MegaHit assembler (Li et al., 2016) and (3) GetOrganelle (Jin et al., 2020) using the SPADES assembler (Bankevich et al., 2012).

For mitogenome extraction and assembly using MitoFinder v.1.4.1, filtered reads for each species were assembled and annotated using an unpublished polleniid mitogenome obtained from GenBank

(MT017729) as a reference (identified as *P. pediculata* Macquart but with a COI barcode region most closely matching *P. rudis* [Fabricus]). Mitogenomes were assembled using both the metaSPADES and MegaHit to identify the most effective assembler for our data (using the --metaspades or --megahit options respectively). For both the metaSPADES and MegaHit assemblies, assembler parameters were left in their default configurations provided through the MitoFinder pipeline (k-mer sizes: 21, 33, 55 for metaSPADES and 21, 29, 39, 59, 79, 99, 119 for MegaHit).

tRNA annotation was completed using the default annotator, MiTFi (Juhling et al., 2012). Other assembly and annotation parameters were not modified (default parameters available at: github.com/ RemiAllio/MitoFinder#detailed-options). Final circular mitogenome structure diagrams were created from annotated mitogenome data using OGDRAW (Greiner et al., 2019). As a comparison, GetOrganelle was also used to extract mitogenomes for each species. GetOrganelle uses an internal reference database of animal mitochondrial genomes and, as such, does not require a custom reference genome. Default parameters for animal mitochondria extraction and assembly parameters were used (automatically estimated word size and k-mer sizes: 21, 45, 65, 85, 105). As GetOrganelle does not allow for automated gene or tRNA annotation, mitogenomes extracted using GetOrganelle were annotated using the MITOS online webserver (Bernt et al., 2013; available at: http://mitos.bioinf.uni-leipzig.de). To compare between software packages and assemblers, three metrics were utilized: annotation success, number of protein-coding genes extracted and total length of the extracted mitogenome (BP).

No loci were able to be extracted for *Xanthotryxus mongol* Aldrich using the above methods. However, it was noted following trimming with Trimmomatic that *X. mongol* contained a greater number of R1 unpaired reads than R1 paired reads. To attempt to harvest data from these additional un-paired R1 reads, the MitoFinder with metaSPADES process was repeated using only the un-paired R1 reads (i.e., single-ended). This modified method allowed for the extraction of COI only. This method was repeated for other taxa for which we were unable to extract any loci, but we did not successfully retrieve any additional loci.

COI barcode reference library and species delimitation analysis

To test whether polleniid species can be delimited using molecular data, a COI barcode reference library was created by combining COI sequences extracted from our newly assembled mitogenomes, 87 COI barcodes (representing 18 species) from Szpila et al. (2022) and COI barcodes from the BOLD (available at https://www.boldsystems. org). BOLD barcodes were identified using the search term 'Polleniidae' and selected to represent the highest number of taxa. For species in which there were more than 10 representative barcode sequences (e.g., *P. rudis* and *P. pediculata*) available from BOLD, 10 representative sequences were chosen at random. A list of all species included in the COI dataset is available in Supplementary Table S3.

We utilized Assemble Species by Automatic Partitioning (ASAP) to test for the presence of multiple species in the final COI dataset (Puillandre et al., 2021). In brief, ASAP functions by applying ascending hierarchical clustering. Sequences are merged into partitions, which are subsequently merged further until only a single partition remains containing all sequences. At each clustering step, each partition is evaluated using two metrics: first, the probability that each partition in the scheme represents a single species, and second, the barcode gap width between the previous and new partitioning schemes. ASAP then uses these two metrics to rank all partitions and propose the optimal partitioning scheme (Puillandre et al., 2021). The combined COI barcode reference library was analysed using the ASAP web server (https://bioinfo.mnhn.fr/abi/public/asap/asapweb.html, accessed 1st November, 2021) with the split groups parameter set to 0.01. For all analyses, we utilized the simple distance method over the more parameterized Kimura-2P and Jukes-Cantor models for the analysis of barcoding sequences derived from closely related species, following the recommendation of Srivathsan and Meier (2012).

Multiple sequence alignment

Following assembly and annotation, mitogenomes were used to generate two multiple sequence alignments (MSAs), a high data coverage (a minimum of 10/13 loci represented, or greater than 75%, for each included species) and a second low data coverage MSA that allowed missing data (a minimum of 1/13 loci represented for each included species).

To create the high data coverage MSA, sequence data representing all 13 mitochondrial protein-coding genes (PCGs) (ATP6, ATP8, COI, COII, COIII, CYTB, ND1, ND2, ND3, ND4, ND4L, ND5 and ND6) were extracted from mitogenomes assembled using the MitoFinder with metaSPADES method. The data were then combined with PCGs extracted from the mitogenomes obtained from GenBank for *P. angustigena* (MT628554) and *M. doronici* (MT410784). Each individual PCG was then aligned in Geneious Prime v.2022.0.1 using the MAFFT plugin. MAFFT alignments were completed using the L-INSI algorithm with 1000 iterative refinement cycles (−maxiterate 1000) (Katoh & Standley, 2013). Individual PCG alignments were then concatenated in Geneious Prime to create the high data coverage MSA, which included 23 of the 32 ingroup taxa (4 genera: 1 sp. *Dexopollenia*, 1 sp. *Melanodexia*, 1 sp. *Morinia*, 18 spp. *Pollenia* and one representative from each of the families Calliphoridae and Tachinidae) and 1 outgroup taxon (one representative from the Sarcophagidae).

To enable the analysis of those taxa for which only fragmentary data were recovered, a second low data coverage MSA was also created. This MSA consisted of all the data from the high data coverage MSA and additional sequences for *M. argenticincta* (Senior-White) and *P. tenuiforceps* Jacentkovskỳ taken from the MitoFinder with MegaHit assembly, *P. leclercqiana* (Lehrer) data taken from the GetOrganelle assembly and *P. bulgarica* Jacentkovskỳ, *P. hungarica* Rognes, *P. paupera* Rognes, *P. ponti* Rognes and *P. similis* Jacentkovskỳ data obtained from GenBank. The final dataset included all 32 ingroup taxa

and 1 outgroup taxon (one representative from the Sarcophagidae). The loci included for each taxon and their corresponding GenBank accession numbers are available in Supplementary Table S5.
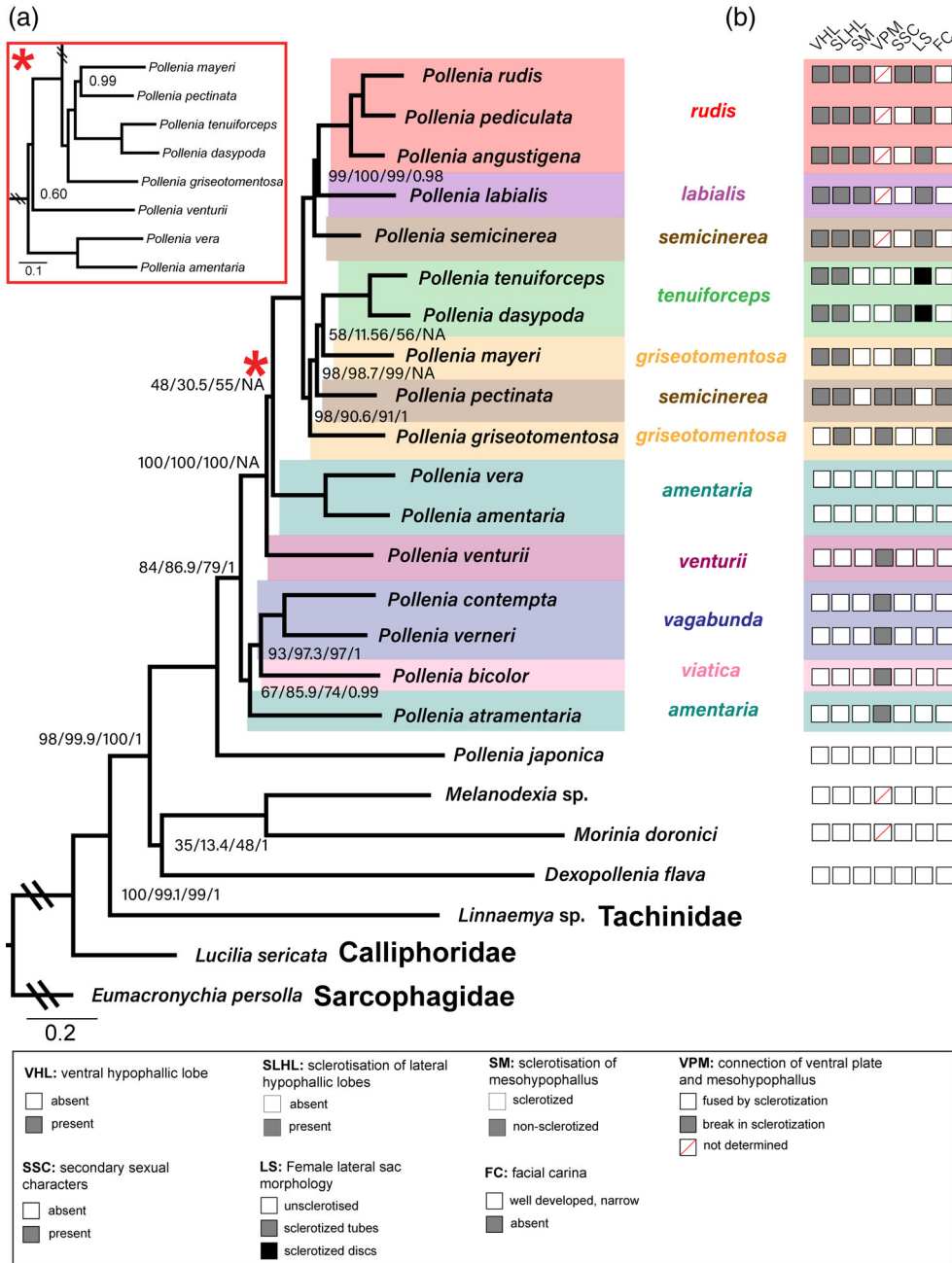
Phylogenetic analysis

The optimal partitioning scheme and substitution model for the maximum likelihood and Bayesian phylogenetic analysis of both MSAs were inferred using PartitionFinder2 v.2.2.1 (Lanfear et al., 2017) according to the Bayesian Information Criterion (BIC) and using the 'greedy' search algorithm (partitioning schemes available: Supplementary Table S4). We selected the 'greedy' search algorithm for our analyses, as the software manual recommends this algorithm be used for datasets composed of greater than 10 loci. The optimal partitioning schemes for both MSAs, as determined by PartitionFinder, were utilized for both maximum likelihood analysis using RAxML and Bayesian inference (BI) using ExaBayes (available in Supplementary Table S4). RAxML only supports the GTR substitution model. As such, this model, with the additional parameter +G, was used across all partitions. It should be noted that the results of PartitionFinder2 recommended using both the +I and + G options. We, however, preferred to use +G only, due to potential errors in tree estimation caused by the non-independence of the proportion of invariable sites (I) and the alpha parameter of the gamma distribution (G), particularly at low values of alpha (Nguyen et al. 2017). ExaBayes only allows the use of the nucleotide substitution model GTR + G, and, as such, this was used for the analysis of all partitions.

Maximum likelihood phylogenetic analyses were completed for each MSA using RAxML v.8.2.12 (Stamatakis, 2014). The program algorithm '−f a' was utilized to both complete a rapid bootstrap analysis with 1000 replicates and search for the best scoring maximum likelihood tree in a single run.

BI analyses were completed for each MSA using ExaBayes v.1.5 (Aberer et al., 2014). Eight simultaneous Bayesian analyses were completed using four Markov Chain Monte Carlo (MCMC) chains per run (three heated, one cold; temperature = 0.1) run for 1.5 million generations, sampling every 1000 generations with the first 25% of samples discarded as burn-in (a total of 8 [runs] × 4 [chains] × 1125 [samples – burn-in] = 36,000 samples across all runs). All priors remained unmodified and in their default state. Effective sampling of the priors and MCMC convergence was established by ensuring effective sample size (ESS) and potential scale reduction factors (PRSF) were > 200 and ∼ 1, respectively. A 50% majority rule extended consensus tree was then prepared from the resultant trees, with the first 25% of all trees discarded as burn-in using the in-built 'consense' software. The extended majority rule consensus tree, the default tree produced by ExaBayes, differs from a regular majority rule consensus tree in that non-trivial bipartitions that occur in less than 50% of the trees are also included in the final topology based on the frequency of their occurrence (Stamatakis & Izquierdo-Carrasco, 2011).

A second partition selection, model optimisation and maximum likelihood analysis were completed for each MSA using IQ-TREE v.1.6.12 (Nguyen et al., 2015) and the inbuilt ModelFinder software.

FIGURE 2 (a) – Phylogenetic hypothesis for the Polleniidae generated using a maximum likelihood method in the IQ-TREE. Bootstrap support values are shown for branches that have less than 99.5% support across all analyses (or 0.99.5 posterior probability for ExaBayes). Support values are in the following order from right to left: IQ-TREE bootstrap support, SH-alrt support, RAxML bootstrap support and ExaBayes posterior probability. The red asterisk indicates the part of the tree where Bayesian analysis provided an alternative topology that is shown in the corresponding red square. (b) – Visualization of the distribution of seven selected characters on the tree.

All possible partitioning schemes and substitution models were analysed ($-m$ MFP + MERGE) and the highest scoring partitioning scheme and models, according to the Bayesian Information Criteron (BIC), were chosen for subsequent analyses (partitioning scheme and models available in Supplementary Table S4). Node support was estimated using both 10,000 ultrafast bootstrap replications and SH-Like approximate likelihood ratio tests (Guindon et al., 2010) with 10,000 replicates ($-$alrt 10,000 -bb 10,000).

All phylogenetic analyses were completed using the University of Technology, Sydney eResearch High Performance Computing Cluster. Final tree graphics (summarizing RAxML, IQ-TREE and BI analyses) were produced using FigTree v.1.4.3 (available from: http://tree.bio. ed.ac.uk/software/figtree) and edited using Adobe Illustrator (available from: https://www.adobe.com/au/products/illustrator). For presentation of the statistical support values, the names of the analysis and subsequent tests of support have been shortened as follows:
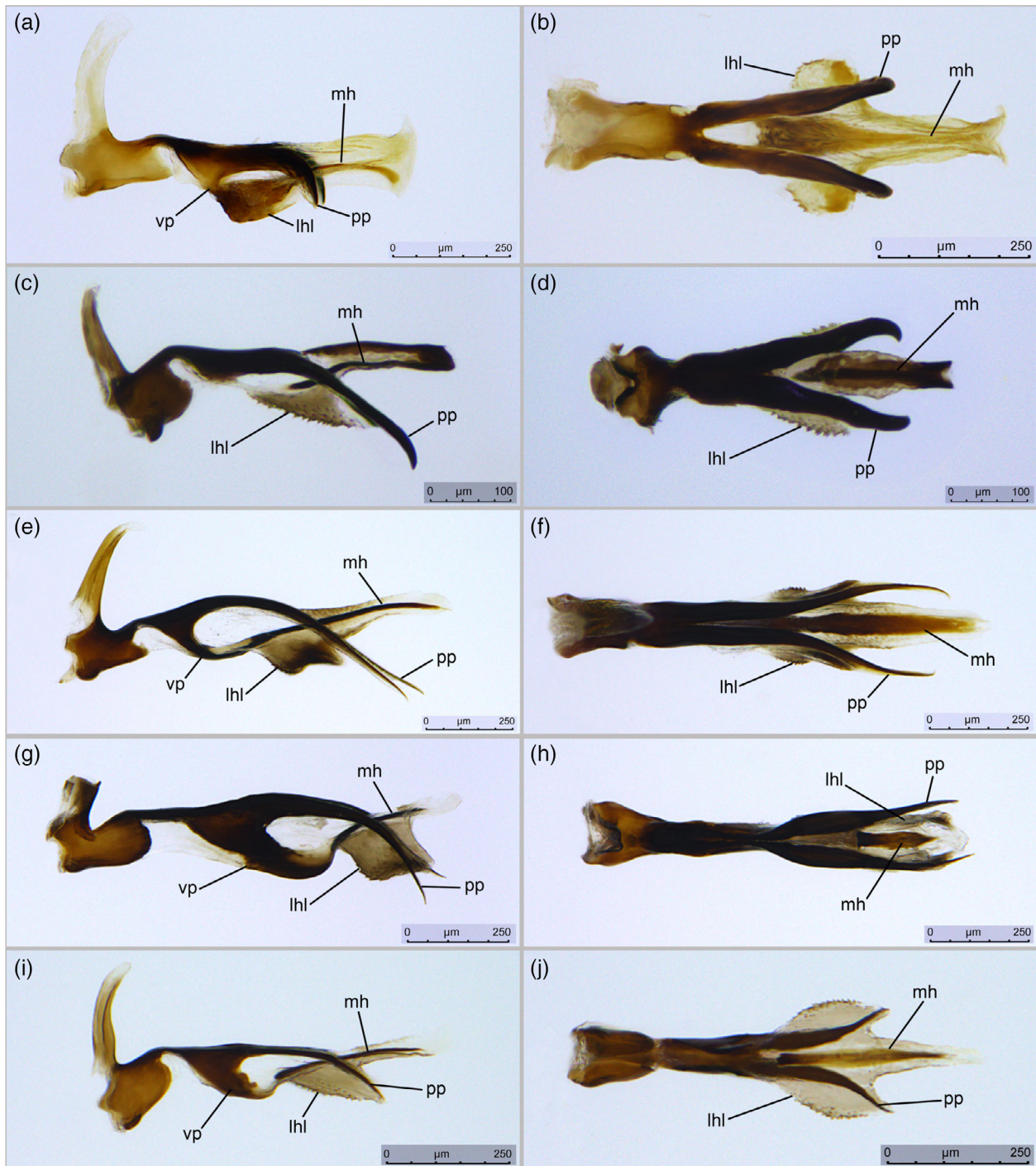
FIGURE 3 Polleniidae phallus morphology: (a, b) – *Dexopollenia flava*; (c, d) – *Morinia doronici*; (e, f) – *Pollenia amentaria*; (g, h) – *Pollenia atramentaria*; (i, j) – *Pollenia bicolor*. Abbreviations: lhl, lateral hypophallic lobe; mh, mesohypophallus; pp, paraphallic process; vp, ventral plate.

IQ-TREE ultrafast bootstrap support = UFBS, IQ-TREE SH-alrt branch support = SH, RAxML bootstrap support = RBS, ExaBayes posterior probability = PP. The final tree figure (Figure 2) summarizes support values for all analyses on the IQ-TREE topology. For simplicity, when explaining node support in the phylogenies, high support indicates clades with 90%–100% support for ML (inclusive of UFBS, SH and RBS) and PP = 0.9–1 for BI analyses, moderate with 85%–90% support and PP = 0.75–0.9 and low with <85%

and PP < 0.75. GenBank accession data for PCGs is available in Supplementary Table S5.

Morphology

As part of our analysis of the *Pollenia* species-group system, we reconciled the morphological characters used in the cladistic analysis of
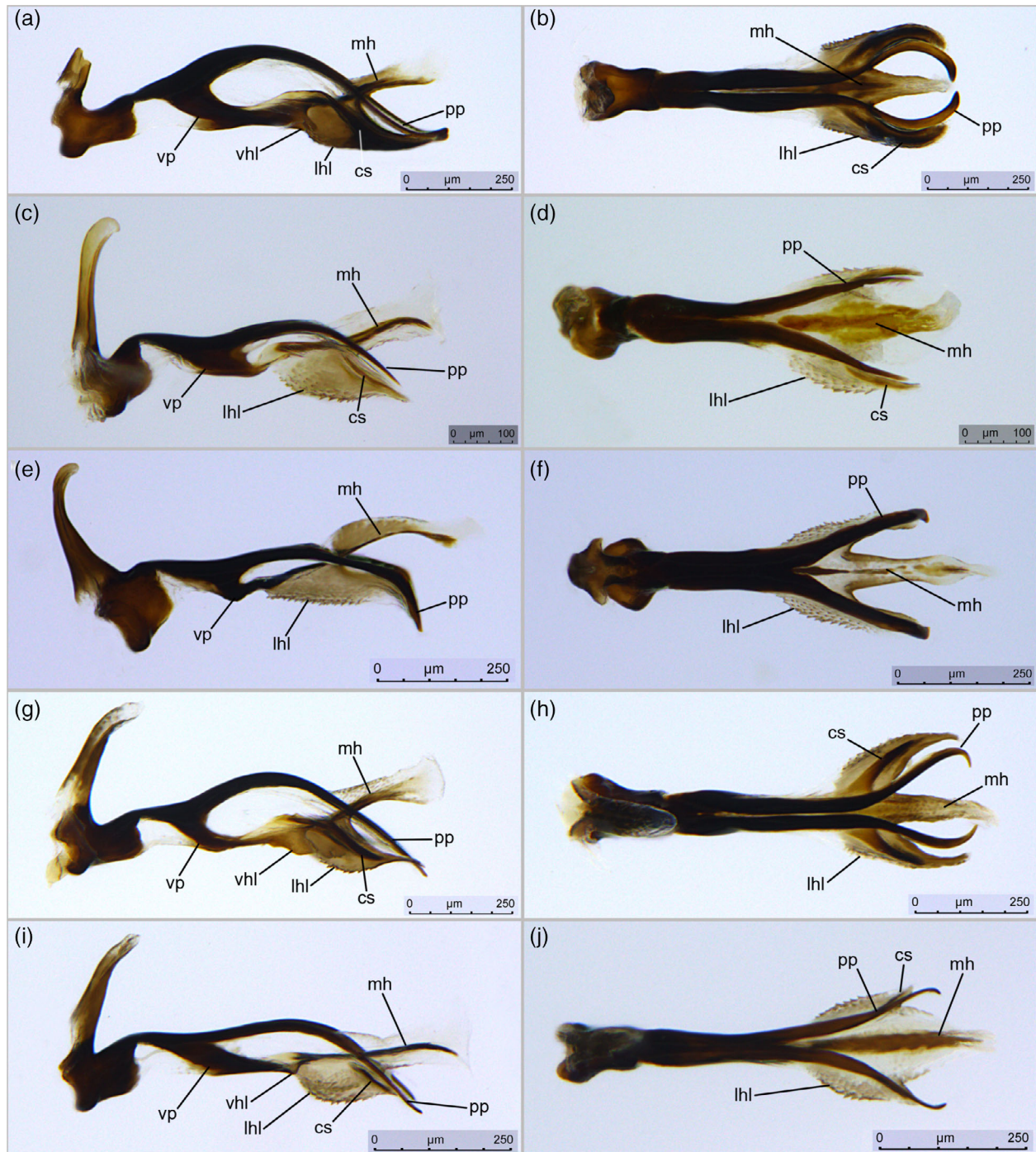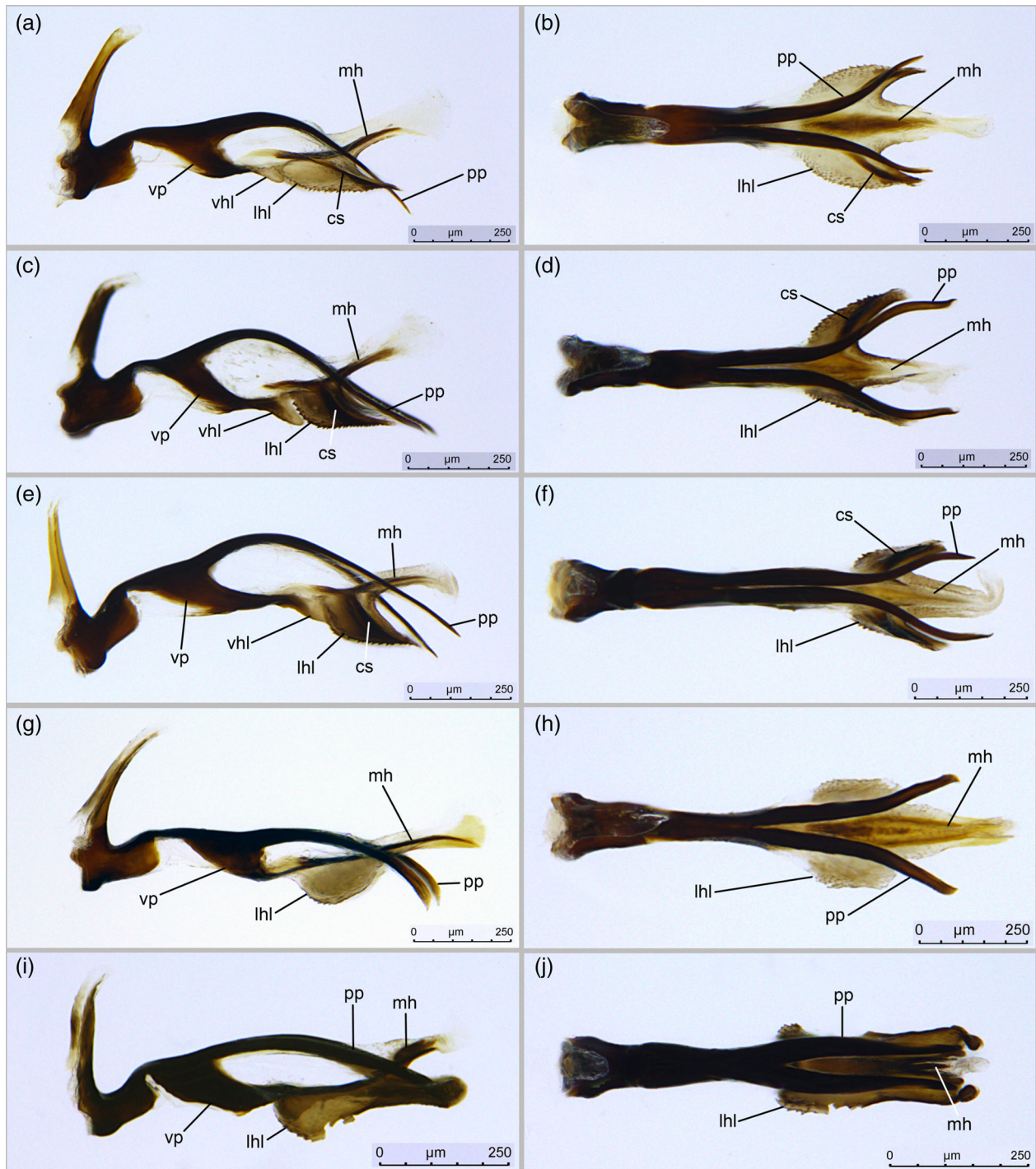
FIGURE 4 *Pollenia* phallus morphology: (a, b) – *Pollenia dasypoda*; (c, d) – *Pollenia griseotomentosa*; (e, f) – *Pollenia japonica*; (g, h) – *Pollenia labialis*; (i, j) – *Pollenia mayeri*. Abbreviations: cs, central sclerotization of lateral hypophallic lobe; lhl, lateral hypophallic lobe; mh, mesohypophallus; pp, paraphallic process; vhl, ventral hypophallic lobe; vp, ventral plate.

Rognes (1992a, Appendix 2). To verify the characters proposed by Rognes (1992a), terminalia from males of 22 species of Polleniidae were dissected, and illustrations of the phallus in lateral and dorsal view are provided for 15 species (Figures 3–5). Male terminalia were dissected from dry pinned specimens following the procedure proposed by Rognes (1991a). Whole abdomens were broken off, placed into 10% KOH, boiled for 5 min, and subsequently dissected in demineralized water. Dissected parts were transferred to glycerol for storage. Focus stacked images of phalluses were taken with a Leica DFC495 digital camera mounted on a Leica M205C stereomicroscope (Leica Camera AG, Germany). The terminology used to describe the terminalia follows Rognes (1992a) with some modifications proposed by Cerretti et al. (2019).

FIGURE 5 *Pollenia* phallus morphology: (a, b) – *Pollenia pectinata*; (c, d) – *Pollenia rudis*; (e, f) – *Pollenia semicinerea*; (g, h) – *Pollenia vagabunda*; (i, j) – *Pollenia venturii*. Abbreviations: cs, central sclerotization of lateral hypophallic lobe; lhl, lateral hypophallic lobe; mh, mesohypophallus; pp, paraphallic process; vhl, ventral hypophallic lobe; vp, ventral plate.

Following morphological analysis, seven characters were found to be diagnostically useful for *Pollenia* (listed in Table 3) and were subsequently used to attempt to morphologically define both the large clades and species-groups resolved in our phylogenetic tree. These seven characters were chosen because they were easily observable, unambiguous for interpretation, and not composite. Four of these characters are related to phallus morphology, one to female terminalia, one to the secondary

sexual dimorphism (Figure S2) and one to the head morphology (Table 3). Characters and character states for species were extracted from previous revisions (Jacentkovskỳ, 1941, 1942; Zumpt 1956; Lehrer 1963, 1967; Mihályi 1976; Rognes, 1987a, 1987b, 1988, 1991b, 1992a and 1992b), museum specimens, and, in the case of phallus morphology, also by dissection of male terminalia (Figures 3–5). Interpretation of character states (plesiomorphic/apomorphic) followed by Rognes (1988, 1992a). Two

INTEGRATIVE SYSTEMATICS OF POLLENIIDAE

Systematic
Entomology

Royal
Entomological
Society

307

TABLE 3 Morphological characters used in analysis.

Character	Matrix code	Character states	State type	Code
Male reproductive structures				
Ventral hypophallic lobe	VHL	Absent	Plesiomorphic	0
		Present	Apomorphic	1
Sclerotization of lateral hypophallic lobes	SLHL	Absent	Plesiomorphic	0
		Present	Apomorphic	1
Sclerotization of the mesohypophallus	SM	Sclerotized	Plesiomorphic	0
		Non-sclerotized	Apomorphic	1
Shape of the paraphallic processes	PP	Straight	Plesiomorphic	0
		Bent inward	Apomorphic	1
Connection of ventral plate and mesohypophallus	VPM	Fused by continuous sclerotisation	Plesiomorphic	0
		Break in sclerotization	Apomorphic	1
Secondary sexual characters	SSC	Absent	Plesiomorphic	0
		Present	Apomorphic	1
Female reproductive structures				
Lateral sacs	LS	Unsclerotized	Plesiomorphic	0
		Sclerotized tubes	Apomorphic	1
		Sclerotized discs	Apomorphic	2
Spermathecae	S	Spherical	Plesiomorphic	0
		Elongate	Apomorphic	1
Head morphology				
Facial carina	FC	Well developed, narrow	Plesiomorphic	0
		Absent	Apomorphic	1
		Well developed, broad	Apomorphic	2

characteristics, level of sclerotization of mesohypophallus and presence/absence of continued sclerotization between ventral plate and mesohypophallus, are proposed for the first time. Interpretation of character states in this case is based on the IQ-TREE topology (Figure 2, Table 3). Morphological character states for each analysed species were summarized adjacent to their names on our phylogenetic tree shown in Figure 2 and are detailed in Table 4.

RESULTS

Mitogenome extraction and structure

The application of a genome skimming approach to the Polleniidae was successful, with the number of raw reads obtained ranging from 2.4–6.4 million (for *Lucilia cuprina* and *Pollenia similis*, respectively; sequencing results summarized in Table S2) equating to ~0.5–1.2X coverage (assuming a genome size of ~550Mbp; Andere et al., 2016, Anstead et al., 2015). Both the MitoFinder and GetOrganelle software packages successfully extracted and annotated mitogenomes from the genome skimming data (comparative software assembler and annotation results summarized in Table S2). Using MitoFinder with MetaSPADES produced the best

results of all software assembler combinations, with 17 annotated mitogenomes (14 for Polleniidae and 3 from other families) produced and 22 species for which at least one protein-coding gene was able to be extracted (19 Polleniidae spp. and 3 from other families). Due to its efficacy with our dataset, we utilized this software and assembler combination for the creation of the final draft mitogenomes. Within the Polleniidae, mitogenomes ranged in size from 14,784–17,027 bp in length, the mitogenomes assembled from the families Calliphoridae, Sarcophagidae and Tachinidae were more consistent in size ranging from 16,060–16,398 bp. For species from all analysed families, nucleotide composition was generally AT-biased (ranging from 76%–78%). The genetic composition and arrangement of genes was identical across all analysed species and consisted of 13 PCGs, 2 rRNAs and 22 tRNAs. An example of the general polleniid mitogenome structure is summarized in Figure S3. Accession numbers for both mitogenome drafts and PCGs are available in Table S5.

In comparison to MitoFinder with MetaSPADES, using MitoFinder with the MegaHit assembler produced fewer annotated genomes (only 14) but was able to extract at least one protein-coding gene from a greater number of species—24 species in total (ranging from 1–13 PCGs). GetOrganelle combined with the MITOS online annotation software performed similarly to Mitofinder with MegaHit,

TABLE 4 Data matrix from morphological reanalysis. Morphological character acronyms and characters states are detailed in Table 3.

Species	VHL	SLHL	SM	PP[a]	VPM	SSC	LS	S[a]	FC
Dexopollenia flava	0	0	0	0	0	0	0	–	0
Melanodexia	0	0	0	0	–	0	0	–	0
Morinia argentincincta	0	0	0	0	–	0	0	–	0
Morinia doronici	0	0	0	0	–	0	0	–	0
Pollenia amentaria	0	0	0	0	0	0	0	0	0
Pollenia angustigena	1	1	1	0	–	0	1	1	0
Pollenia atramentaria	0	0	0	0	1	0	0	0	0
Pollenia bicolor	0	0	0	0	1	0	0	0	2
Pollenia bulgarica	0	0	0	0	0	0	0	0	2
Pollenia contempta	0	0	0	0	1	0	0	0	0
Pollenia dasypoda	1	1	0	1	0	1	2	0	0
Pollenia griseotomentosa	0	1	0	0	1	0	0	0	1
Pollenia hungarica	1	1	1	0	–	1	1	1	0
Pollenia japonica	0	0	0	0	0	0	0	0	0
Pollenia labialis	1	1	1	1	–	0	1	1	1
Pollenia leclercqiana	0	0	0	0	0	0	0	0	0
Pollenia mayeri	1	1	0	0	0	1	0	0	1
Pollenia paupera	1	1	1	0	–	0	1	1	2
Pollenia pectinata	1	1	0	0	1	1	0	0	1
Pollenia pediculata	1	1	1	0	–	0	1	1	0
Pollenia ponti	0	0	0	0	0	0	0	0	2
Pollenia rudis	1	1	1	0	–	1	1	0	0
Pollenia semicinerea	1	1	1	0	–	0	1	0	0
Pollenia similis	1	1	0	1	0	1	2	0	0
Pollenia tenuiforceps	1	1	0	1	0	0	2	0	0
Pollenia vagabunda	0	0	0	0	1	0	0	0	0
Pollenia venturii	0	0	0	0	1	0	0	0	0
Pollenia vera	0	0	0	0	0	0	0	0	0
Pollenia verneri	0	0	0	0	1	0	0	0	0
Xanthotryxus mongol	0	0	0	0	0	0	0	–	0

[a]Morphological characters not mapped on the phylogenetic tree, Figure 2.

producing 16 annotated mitogenomes and extracting PCGs from 23 species.

Species delimitation analysis

The ASAP delimitation analysis identified the best partition scheme for the COI data as 33 groups (ASAP-Score = 2.5; p = 0.0057; W-rank = 0.0002; and threshold distance = 0.019). The second highest ranked partition scheme was identical to the first and identified 33 groups (ASAP-Score = 4.5; p = 0.0043; W-rank = 0.0008; and threshold distance = 0.021). Of the 33 delimited species, 31 matched with described morphospecies. Two inconsistencies between described and delimited species were identified: (1) *Dexopollenia flava* (Aldrich, 1930), was separated from the BOLD sequence identified as *Pollenia* (sic) *flava* (Bold

reference ID: ASILO562-17) and (2) two individuals of *P. vagabunda* (KEIB_DIP_00045 & KEIB_DIP_00046) were separated from the rest of the *P. vagabunda* sequences. The summarized results of the ASAP analysis are available in Table 2.

Phylogenetic relationships among Polleniidae genera

In RAxML, IQ-TREE and BI analyses of the high data coverage MSA, the Polleniidae clustered together as a single clade sister to the Tachinid genus *Linnaemya* Robineau-Desvoidy with high support (UFBS = 100%, SH = 99.1%, RBS = 99%; PP = 1) (Figure 2). At the infrageneric level, two well supported clades were resolved (Figure 2). First, *Melanodexia* + *Morinia* grouped together, sister to *Dexopollenia* with low support (UFBS = 35%, SH = 13.4%, RBS = 48%; PP = 1). This

polleniid clade (*Melanodexia* + *Morinia* + *Dexopollenia*) was then sister to all *Pollenia* species with high support (UFBS = 98%, SH = 99.9%, RBS = 100%; PP = 1).

The addition of COI data for *Xanthotryxus* and an additional *Morinia* species in the low data coverage MSA (Figure S1) caused several changes in the overall phylogenetic tree. In RAxML, IQ-TREE and BI analyses, three major clades were resolved. As with the high data coverage MSA, the Polleniidae all clustered as a clade sister to *Linnaemya* with high support (UFBS = 98.6%, SH = 99%, RBS = 99%, PP = 1). *Xanthotryxus* + *Dexopollenia* was then sister to all other polleniid genera with high support (UFBS = 99.6%, SH = 99%, RBS = 98%; PP = 1). *Morinia* + *Melanodexia* resolved together in a clade sister to all *Pollenia* with poor support (UFBS = 29.2%, SH = 53%, RBS = 43%, PP = 0.72).

Phylogenetics of *Pollenia* species-groups

Within the genus *Pollenia*, all West Palearctic species formed a monophyletic clade sister to the far Eastern Palearctic species *P. japonica* Kano & Shinonaga (UFBS = 100%, SH = 93%, RBS = 90%; PP = 1). Within the West Palearctic *Pollenia* species, three major clades were resolved with moderate to high support: (1) *rudis*-group + *P. labialis* Robineau-Desvoidy (*labialis*-group) + *P. semicinerea* Villeneuve (*semicinerea*-group) (UFBS = 100%, SH = 100%, RBS = 100%, PP = 1), (2) *tenuiforceps*-group + *griseotomentosa*-group + *P. pectinata* (UFBS = 100%, SH = 100%, RBS = 100%, PP = 1), (3) *vagabunda*-group + *P. bicolor* Robineau-Desvoidy + *P. atramentaria* (Meigen) (*amentaria*-group in part) (UFBS = 84%, SH = 86.9%, RBS = 79%, PP = 1) (Figure 2). All species-groups except for *amentaria*, *griseotomentosa* and *semicinerea* were resolved as monophyletic in the RAxML, IQ-TREE and BI analyses. The monophyly of the *venturii*- and *viatica*- groups could not be assessed as only a single representative of each group was present in our analyses.

The *semicinerea*-group was split, with *P. semicinerea* placed in a clade with the *rudis*- and *labialis*- groups (UFBS = 100%, SH = 100%, RBS = 99%, PP = 1) and *P. pectinata* placed in a clade with the *tenuiforceps*- and *griseotomentosa*- groups (UFBS = 98%, SH = 98.7%, RBS = 99%, PP = NA). *Pollenia pectinata* was placed between *P. mayeri* Jacentkovskỳ and *P. griseotomentosa*, which resulted in the *griseotomentosa*-group resolving as not monophyletic. In contrast to the ML phylogenetic trees, BI analysis placed *P. pectinata* as sister to *P. mayeri* (PP = 0.99), then resolved *P. griseotomentosa* as sister to *P. pectinata* + *P. mayeri* (PP = 1).

Our analyses did not resolve the *amentaria*-group as monophyletic. Instead, *P. atramentaria* was placed as sister to *P. bicolor* (*viatica*-group) in the RAxML, IQ-TREE and BI analyses (UFBS = 67%, SH = 85.9%, RBS = 74%, PP = 0.99). Alternatively, the other *amentaria*-group species (*P. amentaria* and *P. vera* Jacentkovskỳ) resolved as sister to the clade *rudis* + *labialis* + *semicinerea* + *tenuiforceps* + *griseotomentosa* in ML analyses (UFBS = 48%, SH = 30.5%, RBS = 55%), whereas BI analysis grouped *P. amentaria* and *P. vera* as sister to the same clade with the addition of *P. venturii* (PP = 1).

Analysis of the low data coverage MSA recovered the same clades, albeit with lower support. The inclusion of additional COI sequences did not change any relationships, and the additional species included were placed as follows: *Pollenia hungarica* and *P. paupera* were both resolved within the *rudis*-group clade (UFBS = 99%, SH = 99.9%, RBS = 93%, PP = 1). *Pollenia similis* resolved as sister to *P. dasypoda* Portschinsky and *P. tenuiforceps* within the *tenuiforceps*-group (UFBS = 100%, SH = 98%, RBS = 97%, PP = 1). *Pollenia leclerqiana* resolved sister to *P. amentaria* within the *amentaria*-group (UFBS = 93%, SH = 91.4%, RBS = 99%, PP = 1). *Pollenia vagabunda* placed sister to *P. contempta* Robineau-Desvoidy within the *vagabunda*-group (UFBS = 100%, SH = 99.9%, RBS = 98%, PP = 1). *Pollenia ponti* and *P. bulgarica* did not place with the rest of the *viatica*-group, and alternatively resolved, with low support, as sisters to *amentaria* (in part) and outside of the *P. bicolor* + *vagabunda*-group + *amentaria*-group clade in the RAxML and IQ-TREE and BI analyses (UFBS = 54%, SH = 27.8%, RBS = 4%, PP = 0.03).

Morphology

Only the following clades were able to be defined by unambiguous shared apomorphic character states (Figure 2):

1. Clade *griseotomentosa*-group + *rudis*-group + *semicinerea*-group (in part) + *tenuiforceps*-group: median hypophallic lobe present (except *P. griseotomentosa*); central sclerotization of lateral hypophallic lobes present (described as absent in *P. mayeri* by Rognes 1992a, 1992b and 1988, but well-illustrated in Figures 4 and by Draber-Mońko (2004)).
2. Clade *P. semicinerea* + *P. labialis* + *rudis*-group: mesohypophallus not sclerotized (when observed in dorsal view); lateral sacs sclerotized and formed as tubes (except *P. labialis*).
3. Clade *P. dasypoda* + *P. tenuiforceps*: median hypophallic lobe fully sclerotized; lateral hypophallic lobes with blunt apex; paraphallic processes with thickened apex; paraphallic processes bent inwards (also in *P. labialis*); lateral sacs sclerotized and formed as discs.
4. Clade *P. amentaria* + *P. vera*: ventral plate short.
5. Clade *P. atramentaria* + *P. bicolor* + *vagabunda*-group: unsclerotized break between the ventral plate and the mesohypophallus present (also in *P. venturii*).
6. Clade *P. contempta* + *P. verneri*: male cerci apically widened and bent backwards; additional setae in front of three regular humeral setae.

DISCUSSION

Mitogenomics of the Polleniidae

Although material quality varied (i.e., some were stored for several decades in the presence of damaging chemicals such as paradichlorobenzene), genome skimming paired with automatic mitogenome

assembly and annotation was highly successful in this group. Draft mitogenomes were produced for 14 polleniid species, and PCGs were extracted from an additional 7 species (21 spp. total). Mitogenome structure was consistent between polleniid genera and species and followed the general organizational pattern of all schizophoran Diptera (13 PCGs, 2 rRNAs and 22 tRNAs) (Junqueira et al., 2016). Alternative assembler and annotation software combinations caused dramatic changes in the overall success, size and quality of mitogenome drafts. MitoFinder with the metaSPAdes assembler produced the highest number of annotated mitogenomes for our data set, however, all software combinations extracted a similar number of PCGs from all species. Our results are in line with other studies which have identified metaSPADES to consistently provide more loci, longer scaffolds and less missing data on test datasets than alternative assemblers (Allio et al., 2020; Nurk et al., 2017). MegaHit and metaSPADES are both metagenomic assemblers and have been designed to account for variance in sequencing coverage (Allio et al., 2020; Li et al., 2016; Nurk et al., 2017), which makes these assemblers particularly suitable for use in cases where data is fragmented or only has low coverage, such as with the sequences extracted from the museum material used in this study (particularly those stored under non-optimum conditions). We identified that choice of assembly and annotation software can cause changes in both the number of mitogenomes and the number of PCGs successfully extracted, and as such, we recommend future studies include an optimisation step testing several mitogenome assembly and annotation software combinations to identify the optimal software for each unique dataset.

The most variation between assembly methods was observed in the assembly of the control region of the mitochondria, which varied between methods by >1000 bp for some species. The control region is difficult to accurately assemble because it contains repeat regions (Bronstein et al., 2018; Formenti et al., 2021), can be heteroplasmic (Ludwig et al., 2000; Irwin et al., 2009) and can be entirely duplicated within the mitogenome of some species (Shao et al. 2004). These genetic features contribute to the difficulty of accurately assembling this region from short reads in our study. It is also likely that the variation in assembly of the control region is what contributed to the large variation in mitogenome size we observed between polleniid species in this study. Improving read length rather than depth would provide better resolution of these repetitive regions, and as such, using longer reads (100–150 bp) or long read sequencing, where reads can be over 10,000 bp long, would also produce good quality results.

Barcode reference library and species delimitations

Our species delimitation analysis expands upon the library created by Szpila et al. (2022) by including all available taxa from the entire Polleniidae family. All species from all genera (5 total) were successfully delimited using ASAP, with 33 species delimited from 31 input morphospecies in both the 1st and 2nd highest rated ASAP partitioning schemes. Of the two additional delimited groups, the first is a species from BOLD identified as '*Pollenia flava*', which is likely a misidentified

Dexopollenia species as it groups closely with our newly sequenced *Dexopollenia flava*. We are confident in the correct identification of our specimen as the shape of its terminalia (Figure 3) fits well with the illustration provided for this species by Kano (1968: fig. 18). The second additional delimited group involved two individuals of *P. vagabunda* which were resolved separate from the rest of the *P. vagabunda* sequences (KEIB_DIP_00045 & KEIB_DIP_00046). These specimens were collected from Puerto de la Morcuera, Spain, whereas all other representatives were collected from Poland. It is therefore possible that these two specimens represent either a morphologically cryptic species distinct from *P. vagabunda*, or alternatively, geographic isolation between the two populations has resulted in two genetically distinct groups. Either way, investigating *P. vagabunda* from a broad range of additional localities, particularly the Mediterranean regions of Europe, would enable intraspecific genetic variability within this species to be determined. Apart from these two additional delimited groups, our results agree with the results of the Automatic Barcode Gap Discovery (ABGD) and Bayesian Phylogenetics and Phylogeography (BPP) analyses performed by Szpila et al. (2022). Interestingly, despite the cryptic morphology of the Polleniidae, almost all previously identified species were able to be molecularly confirmed. This suggests that the approach of defining new species based on the morphology of the male terminalia, like the shape of particular parts of the distiphallus, is appropriate. The diagnostic power of terminalia morphology in the *Pollenia* is further highlighted by the intraspecific instability of other morphological characters previously considered to be taxonomically significant (Rognes, 1991a, 1992a), like the width of the frons of the male or the presence of the microtrichosity on the abdomen (Rognes, 2019; Szpila & Draber-Mońko, 2008). In addition to the morphology of the male terminalia, we identified the following characters as useful for definition of the new species inside particular species-groups of *Pollenia*: the shape of the facial carina, the presence of outer posthumeral seta, the base colour of the legs, the number of posteroventral setae of the fore tibia, the number of anterodorsal setae on the mid tibia, the number of marginal setae on the scutellum, and finally the development of the secondary sexually dimorphic vestiture on the ventral surface of the abdomen and hind legs (Figure S2).

Subfamilial relationships within the Polleniidae

The genus *Pollenia* was defined by Rognes (1992a) by the character state 'yellow crinkly hairs on scutum and pleura'. This definition worked well for North American and West Palearctic fauna, as *Melanodexia* and *Morinia* are devoid of these specific hairs. It is notable, however, that this character state is also present in *Dexopollenia* and *Xanthotryxus* (Fan et al., 1997; Kurahashi, 1967; Kano, 1968; Zumpt 1956). Species of the genus *Pollenia* are separated from *Dexopollenia* and *Xanthotryxus* by hairy parafacial plates, but this character state is also present in *Melanodexia* and is additionally interpreted by Rognes (1991a) as plesiomorphic.

Recent attempts to designate unambiguous character states for the definition of polleniid genera have been unsuccessful (Cerretti et al., 2019; Gisondi et al., 2020). Morphologically, there are no strong apomorphic character states which define each genus, instead much like the *Pollenia* species-groups, genera are defined based on a combination of characters. In the morphology-based key to genera of the world Polleniidae provided by Gisondi et al. (2020), the genera *Dexopollenia* and *Pollenia* are both divided into two separate groups each, located in different parts of the key. Furthermore, terminal taxa in this key are often characterized by the 'simultaneous' presence of sets of character states, without the designation of even a single genus-specific character state (plesio- or apomorphic). Despite the difficulties of morphologically defining the polleniid genera and the restriction of taxon sampling in our study to Palearctic species, our phylogenetic hypothesis placed *Pollenia* in a monophyletic group sister to all other genera (high data coverage phylogeny) or to clade *Melanodexia* + *Morinia* (low data coverage phylogeny). For this reason, we can reject the advocation of previous authors (Cerretti et al., 2019, Gisondi et al., 2020) to synonymize *Dexopollenia* and *Xanthotryxus* with *Pollenia*. It is also worth noting that none of the apomorphic character states diagnostic within *Pollenia* were present in the other genera we included in our analysis (as visualized on our phylogenetic tree, Figure 2). Indeed, this lack of diagnostic morphological characters is the reason why previous authors have had difficulty defining and phylogenetically placing the polleniid genera *Dexopollenia* and *Xanthotryxus* (Cerretti et al., 2019, Gisondi et al., 2020).

The genus *Pollenia*

Pollenia was resolved as monophyletic in the RaxML, IQ-TREE and BI analyses. Interestingly, all West Palearctic species formed a strongly supported clade sister to *P. japonica*, a species from Japan and South Korea (Gisondi et al., 2020; Verves & Khrokalo, 2006). *Pollenia japonica* also did not exhibit any of the apomorphic character states identified in the other *Pollenia* species-groups. Fan et al. (1997) in the key for Chinese fauna separated *P. japonica* (together with East Palearctic *P. shaanxiensis* Fan & Wu) from other *Pollenia* (six species including *P. pectinata* and *P. pediculata*) by the presence of the exclusively black setulae on the gena and postgena ('bucca'). The addition of numerous yellow setulae on the predominantly black vestiture of the genal area is a character state present not only in most of the West Palearctic *Pollenia* but also in *D. flava* and *X. mongol*. The alternative character state, the absence of numerous yellow setulae on the vesti-ture of the genal area, is present in *Melanodexia* and *Morinia* and at least three West Palearctic *Pollenia* species: *P. leclercqiana*, *P. semicinerea* and *P. venturii*. The phallus of *P. japonica* is somewhat similar to that of *Melanodexia* and *Morinia* (Hall, 1948; Rognes, 1991a) with these groups sharing narrow lateral hypophallic lobes and thick paraphallic processes (Figure 4). However, *P. japonica* can be distinguished from these genera through the presence of the well-defined ventral plate, which may be considered as an apomorphy for Palearctic *Pollenia*, although this structure is described as 'very short' in

species of *amentaria*-group (Rognes, 1992a) and is also present in species of *Xanthotryxus* (Zumpt 1956, Kano, 1968, Fan et al., 1997).

Pollenia from Australia and New Zealand are characterized as morphologically distinct (Dear, 1986) from Palearctic species, and they have never been analysed within a molecular phylogenetic framework (Gisondi et al., 2020). The phallus morphology of the New Zealand species is relatively well illustrated (lateral view of the phallus, only) (Dear, 1986). However, the comparison of this structure with the West-Palearctic species and *P. japonica* in a phylogenetic context is difficult because of the high morphological diversity of New Zealand *Pollenia* (Dear, 1986). We can only conclude that *Pollenia* species from the Oceanic region will not possess the apomorphic character states (lack of median hypophallic lobe and the central sclerotization of lateral hypophallic lobes) we have defined for the most derived *Pollenia* clade in our analysis. However, it is worth mentioning that New Zealand *Pollenia* display a diversity of shapes of the lateral hypophallic lobes which significantly exceeds what we observe in the Holarctic species. The phallus of some New Zealand *Pollenia* is also devoid (*P. advena* Dear) or almost devoid (*P. antipodea* Dear; *P. lativertex* Dear) of the ventral plate (Dear, 1986), which resembles the character state described for the genera *Melanodexia* and *Morinia*.

Phylomitogenomics and morphology do not support the current *Pollenia* species-group system

Our phylogeny provides an opportunity to re-evaluate the *Pollenia* species-group system established by Rognes (1992a). We identified that most of the characters utilized in the cladistic analysis of Rognes (characters detailed in appendix 2 and 3 of Rognes, 1992a) display a mosaic distribution, and therefore, only a few species-groups can be defined based on shared morphology: *tenuiforceps* (three apomorphies), *vagabunda* (two apomorphies), *viatica* (three apomorphies) and *amentaria* (a single apomorphy). Other species-groups are defined by a combination of character states whereby no single state is unique for any group. Some characters and subsequent character states identified by Rognes (1992a) may be characterized as uncertain—For example, they were defined broadly using the non-specific terminology 'species-group type' within each species-group. The most obvious example is the 'shape of the lateral hypophallic lobes.' This character had nine states for different species-groups, that is, '*rudis* type', '*tenuiforceps* type', '*vagabunda* type' and so forth. We found that these general 'species-group type' character states were not consistent within the groups: *amentaria*, *griseotomentosa*, *semicinerea* and *viatica*. Interestingly, these four groups appeared to be non-monophyletic in our phylogenetic hypothesis as well as in the phylogenetic hypothesis of Szpila et al. (2022, *semicinerea* group was not included).

Although the species-group system in *Pollenia* is generally not supported by our molecular results, the most derived part of our tree aligns well with the relationships resolved by Rognes (1992a) and Szpila et al. (2022) and unites *rudis*, *labialis*, *tenuiforceps*, *griseotomentosa* and *semicinerea* together into a large, strongly supported clade.

This clade is well corroborated by the morphology of the phallus and includes all species of *Pollenia* known as parasitoids of Lumbricidae (Opisthopora) as well as all species of Holarctic *Pollenia* with expressed secondary sexual dimorphism. Alternatively, in the more basally branching part of our tree, the current species-group classification system is not supported with the *amentaria*-group being split by the *venturii*- and *vagabunda*-groups, with moderate support. The inclusion of additional taxa in the supplementary phylogenetic analysis did not result in major changes to the overall topology. Only the *viatica*-group resolved differently between the primary and supplementary phylogenetic trees. The addition of *P. ponti* and *P. bulgarica* in the supplementary analysis resulted in a polyphyletic *viatica*-group, with (1) *P. ponti* and *P. bulgarica* placed, with poor support, as sisters to the *amentaria*-group (in part) + the large clade composed of *rudis*-, *labialis*-, *tenuiforceps*-, *griseotomentosa*- and *semicinerea*-groups and (2) *Pollenia bicolor* placed as sisters to the *vagabunda*-group (in agreement with the primary phylogenetic analysis).

Our morphological reanalysis of the species-groups defined sensu Rognes (1992a) shows further evidence against the current species-group system within *Pollenia*. For example, within the *viatica*-group, which was polyphyletic in our phylogeny, species were united by Rognes based on three autapomorphies of the female terminalia. However, detailed inspection of the keys, descriptions and data matrix provided by Rognes (1991b, 1992a) and Rognes & Baz (2008) revealed striking morphological differences between *P. bicolor*, *P. ruficrura* Rondani and the other species belonging to the *viatica*-group (i.e., *P. ponti* and *P. bulgarica*). Rognes (1991a, 1992a) described the hypophallic lobes for the entire group as 'low, elongate, square...', but the shape of the lateral hypophallic lobes in *P. bicolor* and *P. ruficrura* is not square/rectangular and appears less elongated than the lobes of other species (Figure 3i, j; Lehrer 1963; Blackith 1991; Rognes, 1991b; Szpila & Draber-Mońko, 2008). Furthermore, both *P. bicolor* and *P. ruficrura* also possess an unsclerotized area between the ventral plate and mesohypophallus, whereas in other members of the *viatica*-group, the ventral plate and mesohypophallus form a continuous sclerotized rod. *Pollenia bicolor* and *P. ruficrura* can also be separated from the rest of the *viatica*-group through the presence of one antero-dorsal seta on the mid tibia (2–4 in other *viatica*-group species) and three pairs of marginal setae on the scutellum (4–5 in other *viatica*-group species; Rognes, 1992a; Rognes & Baz 2008).

Our resolution of the *semicinerea*-group as non-monophyletic, with the separation of *P. pectinata* and *P. semicinerea*, is also morphologically supported by several character states. The close affinity of *P. semicinerea* to species of the *rudis*-group is supported by the shared 'sclerotized tube' form of the lateral sacs and weak sclerotization of the mesohypophallus. In comparison, *P. pectinata* possesses unsclerotized lateral sacs and a sclerotized mesohypophallus. Rognes (1988), in his preliminary phylogenetic hypothesis for the *Pollenia*, proposed a sister relationship between the *rudis*- and *semicinerea*-groups. This deduction was based on two-character states related to the morphology of the phallus and ovipositor morphology; however, these states are not suitable for the purposes of uniting these groups. The first character state, paraphallic processes with armed tips, is only present

in 3 of the 5 species of *semicinerea*-group and 4 of the 7 species of *rudis*-group. The second character state, the position of marginal setae on the tergite T8, is present in all species of *rudis*- and *semicinerea*-groups but is also widespread across the other *Pollenia* species-groups (Rognes, 1992a).

The placement of *Pollenia atramentaria* as sister to *vagabunda*-group + *P. bicolor* in our phylogeny also rendered the *amentaria*-group paraphyletic. A similar result was also obtained in the phylogenetic tree of Szpila et al. (2022), with *P. atramentaria* resolving as sister to *P. vagabunda*. The close relationship between *P. atramentaria* and *P. vagabunda* is supported by genital morphology with both species possessing an unsclerotized break between the ventral plate and the mesohypophallus (also present in *P. bicolor* and *P. ruficrura* representing the *viatica*-group) (Figure 3, Zumpt 1956, Draber-Mońko, 2004).

The placement of *P. pectinata* in our phylogenetic analysis also renders the *griseotomentosa*-group paraphyletic, however, we refrain from taking any taxonomic action to remedy this until more taxa from both *semicinerea*- and *griseotomentosa*-groups are included in future analyses.

Unfortunately, we were unable to add any representatives of the *haeretica*-group to our analysis. This group includes only two species, *P. haeretica* Séguy, which is known from Algeria, Italy and Tunisia and *P. ibalia*, which has only been observed from Morocco. Rognes (2010) was unable to identify close relationships between the *haeretica*-group and any of the other *Pollenia* species-groups, however, he did indicate that some features were shared between this group and *P. venturii* (i.e., completely black body and thin curled black setulae on the scutum). The inclusion of this species-group in future systematic studies should be a priority.

Overall, our results highlight that the species-group classification system proposed by Rognes (1992a) is not fully supported by our phylogenetic results nor morphological analysis. Although some morphological characters, such as the male and female terminalia, can be used to predictively identify relationships between species in *Pollenia*, the current species-group classification system in *Pollenia* needs to be re-evaluated, facilitated by greater taxon sampling and more robust genetic analyses.

Conclusions

The Polleniidae are a morphologically and systematically complex group, in which relationships at the family, generic and species level are only just beginning to be disentangled. Through the application of mitogenomic techniques, we produced mitogenomes for 14 polleniid species, and through ASAP analysis of extracted COI data successfully reconciled 31 polleniid species. Our phylogenetic analysis of mitogenomic data facilitated the identification of key relationships between polleniid genera and within the genus *Pollenia*, showing strong support for a monophyletic *Pollenia* and rejecting the synonymisation of the genera *Dexopollenia* and *Xanthotryxus* with the Palearctic *Pollenia*. Morphological analysis enabled us to identify key structures that support the topology of our phylogenetic tree,

INTEGRATIVE SYSTEMATICS OF POLLENIIDAE

Systematic
Entomology

Royal
Entomological
Society

313

of which male terminalia appear to provide the best support. We critically re-evaluated the species-group system within the genus *Pollenia*, highlighting that it is not supported by morphology nor our mitogenomic phylogeny. We identified several paraphyletic relationships that should be the focus of future systematic studies in this group, most notably the relationships between the most basally branching *Pollenia* species, the placement of *P. pectinata* within the *griseotomentosa*-group and the separation of *P. atramentaria* from the rest of the *amentaria*-group.

AUTHOR CONTRIBUTIONS

N.P. Johnston: Conceptualisation, Data Curation, Formal Analysis, Investigation, Methodology, Writing- Original Draft, Visualization; M. Piwczyński: Conceptualisation, Data Curation, Formal Analysis, Investigation, Methodology, Supervision, Resources, Writing - Review & Editing; P. Trzeciak: Investigation, Methodology; K. Walczak: Investigation, Methodology; K. Szpila: Conceptualisation, Data Curation, Formal Analysis, Investigation, Methodology, Supervision, Resources, Writing - Review & Editing.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the supplementary material of this article, except for molecular data which are stored on GenBank (https://www.ncbi.nlm.nih.gov/genbank/) and available using the accession numbers provided in the supplementary material. The NCBI Short Read Archive Bioproject number for this project is PRJNA880496. Other data including Tree files, multiple sequence alignments and the COI barcode sequence library are available on the online repository FigShare under the following DOI: https://doi.org/10.6084/m9.figshare.21100591.

ORCID

Nikolas P. Johnston https://orcid.org/0000-0002-4208-4666
Marcin Piwczyński https://orcid.org/0000-0002-1756-5580
Paulina Trzeciak https://orcid.org/0000-0003-4508-6584
Kinga Walczak https://orcid.org/0000-0002-7413-9828
Krzysztof Szpila https://orcid.org/0000-0002-3039-3146

REFERENCES

Aberer, A.J., Kobert, K. & Stamatakis, A. (2014) Exabayes: massively parallel bayesian tree inference for the whole-genome era. *Molecular Biology and Evolution*, 31, 2553–2556. https://doi.org/10.1093/molbev/msu236

Allio, R., Schomaker-Bastos, A., Romiguier, J., Prosdocimi, F., Nabholz, B. & Delsuc, F. (2020) MitoFinder: efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Molecular Ecology Resources*, 20, 892–905. https://doi.org/10.1111/1755-0998.13160

Andere, A.A., Platt, R.N., Ray, D.A. & Picard, C.K. (2016) Genome sequence of *Phormia regina* Meigen (Diptera: Calliphoridae): implications for medical, veterinary and forensic research. *BMC Genomics*, 17, 842. https://doi.org/10.1186/s12864-016-3187-z

Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Anstead, C., Korhonen, P., Young, N. et al. (2015) *Lucilia cuprina* genome unlocks parasitic fly biology to underpin future interventions. *Nature Communications*, 6, 7344 (2015). https://doi.org/10.1038/ncomms8344

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S. et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, 19, 455–477. https://doi.org/10.1089/CMB.2012.0021=

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G. et al. (2013) MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, 69(2), 313–319. https://doi.org/10.1016/J.YMPEV.2012.08.023

Blaschke, J.D., Stireman, J.O., III, O'Hara, J.E., Cerretti, P. & Moulton, J.K. (2018) Molecular phylogenetics and piercer evolution in the bug-killing flies (Diptera: Tachinidae: Phasiinae). *Systematic Entomology*, 43, 218–238. https://doi.org/10.1111/syen.12272

Blackith, R.E. (1991) *Pollenia bicolor* Robineau-Desvoidy, 1830; un Diptère énigmatique du Midi de la France (Dipt. Calliphoridae). *Bulletin de La Société Entomologique de France*, 96, 271–274. https://doi.org/10.3406/bsef.1991.17726

Bronstein, O., Kroh, A. & Haring, E. (2018) Mind the gap! The mitochondrial control region and its power as a phylogenetic marker in echinoids. *BMC Evolutionary Biology*, 18. https://doi.org/10.1186/s12862-018-1198-x

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

Buenaventura, E., Lloyd, M.W., Perilla López, J.M., González, V.L., Thomas-Cabianca, A. & Dikow, T. (2020) Protein-encoding ultraconserved elements provide a new phylogenomic perspective of Oestroidea flies (Diptera: Calyptratae). *Systematic Entomology*, 46(1), 5–27. Portico. https://doi.org/10.1111/syen.12443

Cerretti, P., Stireman, J.O., Badano, D., Gisondi, S., Rognes, K., Giudice, G. L. et al. (2019) Reclustering the cluster flies (Diptera: Oestroidea, Polleniidae). *Systematic Entomology*, 44, 957–972. https://doi.org/10.1111/syen.12369

Cerretti, P., Stireman, J.O., Pape, T., O'Hara, J.E., Marinho, M.A.T., Rognes, K. et al. (2017) First fossil of an oestroid fly (Diptera: Calyptratae: Oestroidea) and the dating of oestroid divergences. *PLoS One*, 12, 1–24. https://doi.org/10.1371/journal.pone.0182101

Dear, J.P. (1986) Calliphoridae (Insecta: Diptera). *Fauna of New Zealand*, 8, 1–86.

Draber-Mońko, A. (2004) Calliphoridae Plujki (Insecta: Diptera). In: *Fauna Poloniae 23*. Fundacja Natura Optima Dux/Muzeum i Instytut Zoologii. Warsaw: PAN. p. 662.

El-Husseini, M.M.M. (2019) Endo- or ecto-parasitism with the cluster fly, *Pollenia dasypoda* Portochisky [sic!] (Diptera: Calliphoridae), based on

the diameter of its host body, the earthworm *Allolobophora caligonosa* (Sav.). *Egyptian Journal of Biological Pest Control*, 29, 1–5. https://doi.org/10.1186/s41938-019-0150-8

Espeland, M., Irestedt, M., Johanson, K.A., Åkerlund, M., Bergh, J.E. & Källersjö, M. (2010) Dichlorvos exposure impedes extraction and amplification of DNA from insects in museum collections. *Frontiers in Zoology*, 7, 1–7. https://doi.org/10.1186/1742-9994-7-2/FIGURES/5

Fan, Z., Gan, Y., Chen, Z., Fang, J., Zheng, S., Gan, Y. et al. (1997) *Diptera: Calliphoridae. Fauna Sinica, Insecta*, Vol. 6. Beijing: Science Press, p. 707 [in Chinese with English description of the new taxa].

Formenti, G., Rhie, A., Balacco, J., Haase, B., Mountcastle, J., Fedrigo, O. et al. (2021) Complete vertebrate mitogenomes reveal widespread repeats and gene duplications. *Genome Biology*, 22. Available from: https://doi.org/10.1186/s13059-021-02336-9

Gisondi, S., Rognes, K., Badano, D., Pape, T. & Cerretti, P. (2020) The world Polleniidae (Diptera, Oestroidea): key to genera and checklist of species. *ZooKeys*, 971, 105–155. https://doi.org/10.3897/zookeys.971.51283

Greiner, S., Lehwark, P. & Bock, R. (2019) OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research*, 47, W59–W64.

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

Hall, D.G. (1948) *The blowflies of North America*. Lafayette, Indiana: The Thomas Say Fundation, p. 477.

Howlett, B.G., Davidson, M.M., Pattemore, D.E., Walker, M.K. & Nelson, W.R. (2016) Seasonality of calliphorid and sarcophagid flies across Canterbury arable farms requiring pollinators. *New Zealand Plant Protection*, 69, 290–295. https://doi.org/10.30843/nzpp.2016.69.5899

Irwin, J.A., Saunier, J.L., Niederstätter, H., Strouss, K.M., Sturk, K.A., Diegoli, T.M. et al. (2009) Investigation of heteroplasmy in the human mitochondrial DNA control region: a synthesis of observations from more than 5000 global population samples. *Journal of Molecular Evolution*, 68, 516–527. Available from: https://doi.org/10.1007/s00239-009-9227-4

Jacentkovskỳ, D. (1941) Kuklice (Tachinoidea, Diptera) Moravy a Slezska. *Acta Societatis Scientiarum Naturalium Moravicae*, 13, 1–64.

Jacentkovskỳ, D. (1942) Beitrag zur Kenntnis der Gattung Pollenia R.D. Abhandlungen der Russ. *Forschungsgesellschaft Prag*, 11, 193–225.

Jin, J.J., Yu, W.B., Yang, J.B., Song, Y., Depamphilis, C.W., Yi, T.S. et al. (2020) GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology*, 21, 241. https://doi.org/10.1186/s13059-020-02154-5

Juhling, F., Putz, J., Bernt, M., Donath, A., Middendorf, M., Florentz, C. et al. (2012) Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. *Nucleic Acids Research*, 40(7), 2833–2845.

Junqueira, A.C.M., Azeredo-Espin, A.M.L., Paulo, D.F., Marinho, M.A.T., Tomsho, L.P., Drautz-Moses, D.I. et al. (2016) Large-scale mitogenomics enables insights into Schizophora (Diptera) radiation and population diversity. *Scientific Reports*, 6(1), 1–13. https://doi.org/10.1038/srep21762

Kano R. 1968. Calliphoridae (Insecta: Diptera). Fauna Japonica, Biological Society of Japan, 188 pp + XXIII plates.

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

Kurahashi, H. (1967) Studies on the Calypterate Muscoid flies from Japan VI. Revision of the tribes Bengaliini and Polleniini of the subfamily Calliphorinae and the subfamilies Chrysomyiinae and Rhiniinae (Diptera, Calliphoridae). *The Science Reports of Kanazawa University*, 12, 255–302.

Kutty, S.N., Meusemann, K., Bayless, K.M., Marinho, M.A.T., Pont, A.C., Zhou, X. et al. (2019) Phylogenomic analysis of Calyptratae: resolving the phylogenetic relationships within a major radiation of Diptera. *Cladistics*, 6, 605–622. https://doi.org/10.1111/cla.12375

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2017) Partitionfinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34, 772–773. https://doi.org/10.1093/molbev/msw260

Lehrer, A.Z. (1963) Études sur les Diptères Calliphorides. I. – La classification des Polleniinae palearctiques et leur dispersion en Roumanie. *Bulletin et Annales de la Société Royale d'Entomologie de Belgique*, 99, 285–310.

Lehrer, A.Z. (1967) Études sur les Diptères Calliphorides. II. – Revision de la nomenclature de la tribu Polleniini. *Bulletin et Annales de la Société Royale d'Entomologie de Belgique*, 103, 255–259.

Li, D., Luo, R., Liu, C.M., Leung, C.M., Ting, H.F., Sadakane, K. et al. (2016) MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*, 102(6), 3–11.

Lobanov, A.M. (1976) Opredelitel' samok sem. Calliphoridae (Diptera) fauny SSSR po yaitsekladam. In: Opredeliteli po faunie SSSR, Izdavaemye Zoologicheskim Institutom Akademii Nauk SSSR, Vol. 109. Izadetelstvo "Nauka", Leningrad, p. 48.

Ludwig, A., May, B., Debus, L. & Jenneckens, I. (2000) Heteroplasmy in the mtDNA control region of sturgeon (Acipenser, Huso and Scaphirhynchus). *Genetics*, 156, 1933–1947. https://doi.org/10.1093/genetics/156.4.1933

Mawlood, N.A. & Abdul-Rassoul, M.S. (2009) A new species in the genus *Pollenia* Rob. -Desvoidy, 1830 (Diptera Calliphoridae). *Bulletin of the Iraq Natural History Museum*, 10, 59–65.

Mihályi, F. (1976) Contribution to the knowledge of the Genus *Pollenia* R.D. (Diptera: Calliphoridae). *Acta Zoologica Academiae Scientiarum Hungaricae*, 22, 327–333.

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

Nguyen, L.-T., von Haeseler, A. & Minh, B.Q. (2017) Complex models of sequence evolution require accurate estimators as exemplified with the invariable site plus gamma model. *Systematic Biology*, 67(3), 552–558. https://doi.org/10.1093/sysbio/syx092

Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P.A. (2017) metaSPAdes: a new versatile metagenomic assembler. *Genome Research*, 27(5), 824–834.

Puillandre, N., Brouillet, S. & Achaz, G. (2021) ASAP: assemble species by automatic partitioning. *Molecular Ecology Resources*, 21, 609–620. https://doi.org/10.1111/1755-0998.13281

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, 20.1–20.17. https://doi.org/10.1146/annurev-ento-011019-025055

Raxworthy, C.J. & Smith, B.T. (2021) Mining museums for historical DNA: advances and challenges in museomics. *Trends in Ecology & Evolution*, 36, 1049–1060. https://doi.org/10.1016/J.TREE.2021.07.009

Rognes, K. (1987a) The taxonomy of the *Pollenia rudis* species-group in the Holarctic region (Diptera: Calliphoridae). *Systematic Entomology*, 12, 475–502. https://doi.org/10.1111/j.1365-3113.1987.tb00219.x

Rognes, K. (1987b) A new species in the intermedia-group and a new synonymy in the genus *Pollenia* Robineau-Desvoidy, 1830 (Diptera:

Calliphoridae). *Systematic Entomology*, 12, 381–388. https://doi.org/10.1111/j.1365-3113.1987.tb00208.x

Rognes, K. (1988) The taxonomy and phylogenetic relationships of the *Pollenia semicinerea* species-group (Diptera: Calliphoridae). *Systematic Entomology*, 13, 315–345. https://doi.org/10.1111/j.1365-3113.1988.tb00245.x

Rognes, K. (1991a) *Blowflies (Diptera, Calliphoridae) of Fennoscandia and Denmark*. Leiden-New York-Kobenhavn-Koln: E.J. Brill/Scandinavian Science Press Ltd.

Rognes, K. (1991b) Revision of the cluster-flies of the *Pollenia viatica* species-group (Diptera: Calliphoridae). *Systematic Entomology*, 16, 439–498. https://doi.org/10.1111/j.1365-3113.1991.tb00678.x

Rognes, K. (1992a) Revision of the cluster-flies of the *Pollenia venturii* species-group, with a cladistic analysis of Palaearctic species of *Pollenia* Robineau-Desvoidy (Diptera: Calliphoridae). *Insect Systematics and Evolution*, 23, 233–248. https://doi.org/10.1163/187631292X00083

Rognes, K. (1992b) Revision of the cluster-flies of the *Pollenia vagabunda* species-group (Diptera: Calliphoridae). *Entomologica Scandinavica*, 23, 95–114. https://doi.org/10.1163/187631292X00056

Rognes, K. (2010) Revision of the cluster flies of the *Pollenia haeretica* species-group (Diptera, Calliphoridae). *Zootaxa*, 56, 39–56. https://doi.org/10.11646/zootaxa.2499.1.3

Rognes, K. (2016) A new species of *Pollenia* Robineau-Desvoidy, 1830 from Jordan (Diptera: Calliphoridae: Polleniinae). *Zootaxa*, 4067, 569–576. https://doi.org/10.11646/zootaxa.4067.5.3

Rognes, K. (2019) The Calliphoridae (Diptera) of Armenia. *Zootaxa*, 4576, 375–391. https://doi.org/10.11646/zootaxa.4576.2.11

Rognes, K. & Baz, A. (2008) A new species in the *Pollenia viatica* species-group from sierra de Guadarrama, Spain (Diptera: Calliphoridae). *Studia Dipterologica*, 14, 389–395.

Schluesslmayr, G. & Sivell, O. (2021) *Pollenia margarita*, a new species of *Pollenia* Robineau-Desvoidy, 1830 from Austria (Diptera, Polleniidae). *Linzer Biologische Beiträge*, 53, 413–419.

Shao, R., Barker, S.C., Mitani, H., Aoki, Y. & Fukunaga, M. (2004) Evolution of duplicate control regions in the mitochondrial genomes of Metazoa: a case study with Australasian *Ixodes* ticks. *Molecular Biology and Evolution*, 22, 620–629. https://doi.org/10.1093/molbev/msi047

Srivathsan, A. & Meier, R. (2012) On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. *Cladistics*, 28, 190–194. https://doi.org/10.1111/j.1096-0031.2011.00370.x

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stamatakis, A. & Izquierdo-Carrasco, F. (2011) Result verification, code verification and computation of support values in phylogenetics. *Briefings in Bioinformatics*, 12(3), 270–279. https://doi.org/10.1093/bib/bbq079

Stireman, J.O., Cerretti, P., O'Hara, J.E., Blaschke, J.D. & Moulton, J.K. (2019) Molecular phylogeny and evolution of world Tachinidae (Diptera). *Molecular Phylogenetics and Evolution*, 139, 106358. https://doi.org/10.1016/j.ympev.2018.12.002

Szpila, K. (2003a) Actual impact of the blowfly – parasitoids in earthworms (Oligochaeta: Lumbricidae) of cultivated fields, against the literature data. *Wiadomości Entomologiczne*, 22, 211–220.

Szpila, K. (2003b) First instar larvae of nine West-Palaearctic species of *Pollenia* Robineau-Desvoidy, 1830 (Diptera: Calliphoridae). *Entomologica Fennica*, 14. https://doi.org/10.33338/ef.84188

Szpila, K. & Draber-Mońko, A. (2008) *Pollenia moravica* (Jacentkovskỳ, 1941) (Diptera: Calliphoridae) recorded from Poland for the first time. *Fragmenta Faunistica*, 51, 139–142. https://doi.org/10.3161/00159301ff2008.51.2.139

Szpila, K., Piwczyński, M., Glinkowski, W., Lutz, L., Akbarzadeh, K., Baz, A. et al. (2022) First molecular phylogeny and species delimitation of

west Palaearctic *Pollenia* (Diptera: Polleniidae). *Zoological Journal of the Linnean Society*, zlac035, 1–16. https://doi.org/10.1093/ZOOLINNEAN/ZLAC035

Tawfik, M.F.S. & El-Husseini, M.M. (1971) Life-history of *Pollenia dasypoda* Portochisky [sic!] a parasite of the earth worm *Allolobophora caliginosa* (Sav.). *Bulletin de la Société Entomoloque d'Égypte*, 55, 275–287.

Timbó, R.V., Togawa, R.C., Costa, M.M.C., Andow, D.A. & Paula, D.P. (2017) Mitogenome sequence accuracy using different elucidation methods. *PLoS One*, 12(6), e0179971. https://doi.org/10.1371/JOURNAL.PONE.0179971

Verves, Y. & Khrokalo, L. (2006) Fam. Calliphoridae. In: Leley, A.S. (Ed.) *Key to the insects of Russian Far East. Vol. VI. Diptera and Siphonaptera. Pt 4*. Vladivostok: Dal'nauka, pp. 15–60 [in Russian].

Vezsenyi, K.A., Langer, S.V., Samkari, B.A. & Beresfrod D.V. (2022) The history and current state of cluster flies (Diptera: Polleniidae: *Pollenia*) in North America, with new Canadian provincial records. *Canadian Entomologists*, 154: e24.

Yahnke, W. & George, J.A. (1972) Rearing and immature stages of the cluster fly (*Pollenia rudis*) (Diptera: Calliphoridae) in Ontario. *Canadian Entomologist*, 104, 567–576. https://doi.org/10.4039/Ent104567-4

Yan, L., Pape, T., Meusemann, K., Kutty, S.N., Meier, R., Bayless, K. et al. (2021) Monophyletic blowflies revealed by phylogenomics. *BMC Biology*, 19, 230.

Yeates, D.K., Zwick, A. & Mikheyev, A.S. (2016) Museums are biobanks: unlocking the genetic potential of the three billion specimens in the world's biological collections. *Current Opinion in Insect Science*, 18, 83–88.

Zimmermann, J., Hajibabaei, M., Blackburn, D.C., Hanken, J., Cantin, E., Posfai, J. et al. (2008) DNA damage in preserved specimens and tissue samples: a molecular assessment. *Frontiers in Zoology*, 5, 1–13. https://doi.org/10.1186/1742-9994-5-18/FIGURES/9

Zumpt, F. (1956) 64i. Calliphorinae. In: Lindner, E. (Ed.) *Die Fliegen der Palaearktischen Region*. Band 11 [Lieferungen 190, 191, 193]. Stuttgart: Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), p. 140.

SUPPORTING INFORMATION

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

Figure S1. Low data coverage MSA phylogenetic hypothesis for the Polleniidae.

Figure S2. Sexually dimorphic characters in *Pollenia*.

Figure S3. Representative Polleniidae Mitogenome Structure, *Pollenia atramentaria*.

Table S1. Specimens included in mitogenome extraction and assembly.

Table S2. Mitogenome assembly results.

Table S3. COI Barcode library specimens and accession numbers.

Table S4. Partitioning schemes for phylogenetic analysis.

Table S5. MT genome and PCG accession numbers.