PENSOFT
SENCKENBERG world of biodiversity
Leibniz Association

Mitochondrial genomes provide new insights into the phylogeny and evolution of Anthomyiidae (Insecta: Diptera)

He-Nan Li[1,*], Wen-Ya Pei[1,*], Ming-Fu Wang[2], Bang-Qing Chen[3], Hong-Lin Peng[3], Rong-Jun Cao[3], Ming-Teng Zhao[4], Jun Yang[4], Xiao-Chen Zhang[4], Dong Zhang[1]

1 *School of Ecology and Nature Conservation, Beijing Forestry University, Qinghua East Road 35, Beijing, 100083, China*

2 *College of Life Science, Shenyang Normal University, Huanghe North Street 253, Shenyang, 110034, China*

3 *Dalaoling Nature Reserve Administration of Yichang Three Gorges, Yichang 443000, Xiling Second Road 13, Yichang, 443000, China*

4 *Management Office of Beijing Baihua Mountain National Nature Reserve, Baihua Mountain Road 102, Beijing, 102311, China*

* *These two authors contributed equally to this work.*

https://zoobank.org/05F07EDD-7E3F-4FB4-A635-1FA072BCC08C

Corresponding author: Dong Zhang (zhangdong_bjfu@bjfu.edu.cn)

Received 13 May 2023
Accepted 3 September 2023
Published 20 December 2023

Academic Editors Brian Wiegmann, Anna Hundsdörfer

Citation: Li H-N, Pei W-Y, Wang M-F, Chen B-Q, Peng H-L, Cao R-J, Zhao M-T, Yang J, Zhang X-C, Zhang D (2023) Mitochondrial genomes provide new insights into the phylogeny and evolution of Anthomyiidae (Insecta: Diptera). Arthropod Systematics & Phylogeny 81: 1051–1062. https://doi.org/10.3897/asp.81.e106356

Abstract

Anthomyiidae is a cosmopolitan and diverse family of Calyptratae, and is routinely considered to play key roles in both ecology and agriculture. The higher-level phylogenetic classification of Anthomyiidae has been highly controversial, necessitating further molecular data for precise reconstruction of phylogenetic relationships. In this study, we successfully acquired and annotated 18 new mitogenomes of anthomyiids. Moreover, the mitogenomes of the following genera *Eustalomyia*, *Hyporites*, *Leucophora*, *Emmesomyia* and *Eutrichota* are reported for the first time. The 18 mitogenomes are compared with confamilial species to assess genetic variation and to better understand evolutionary relationships within the family Anthomyiidae. In comparisons among 13 mitochondrial protein coding genes (PCG), the calculation of evolutionary rate exhibited *nad1* as the fastest evolving gene in Anthomyiidae. Among the anthomyiids investigated, *cox2* and *nad4* had the lowest genetic distance across the 13 PCGs, suggesting a high degree of conservation for these two genes. Herein, we conducted phylogenetic analyses of the newly sequenced mitogenomes along with 11 known anthomyiids to investigate the interrelationships of Anthomyiidae. Our results indicate that Anthomyiidae is a monophyletic lineage and sister group to Scathophagidae, confirming prior findings based on morphological and molecular analyses. We recovered two subfamilies as monophyletic (Myopininae, Pegomyinae) while Anthomyiinae was polyphyletic. The great species diversity of anthomyiid flies limits the availability of mitogenomes for accurately resolving the phylogeny of Anthomyiidae. Nonetheless, our study provides novel insight into the molecular taxonomy, evolution, and phylogeny of the family Anthomyiidae.

Key words

Calyptratae, evolutionary rate, mitogenome, molecular analysis, Muscoidea, phylogenetics

1. Introduction

Anthomyiidae (Diptera: Calyptratae, Anthomyiidae) are the second-most speciose family in a grade of flies called the the Muscoidea (Kutty et al. 2008), comprising approximately 40 genera and 2,000 species worldwide. The species diversity of Holarctic Anthomyiidae is extremely rich, accounting for nearly one third of the known global fauna, but it remains inadequately researched (Wang et al. 2014). Larvae of some genera of Anthomyiidae are economically important as phytophagous pests on diverse crops of commercial interest, with the best-known pests, *Delia* Robineau-Desvoidy and *Strobilomyia* Michelsen, inflicting substantial damage to both agricultural and forest plants (Hao et al. 2016; Sachet et al. 2006). Adults are found in humid, cool forests and some are active pollinators, while others are drawn to decaying plants or dung (Grisales et al. 2016). Anthomyiids exhibit a rich diversity in appearance, anatomy, ecology and behavior, and whether serving as pollinators or pests, they have a have a substantial impact on human society (Córdova-García et al. 2023; Moretti et al. 2021; Wang et al. 2014).

Taxonomy of Anthomyiidae is challenging due to a reliance on male genitalia for most morphological diagnoses. A systematic classification for Anthomyiidae is currently deficient and no comprehensive experiments have been conducted using rigorous cladistic argumentation to systematize this family (Michelsen et al. 2010). The phylogenetic relationships of Anthomyiidae are still contentious, and lack a universally accepted classification system (Michelsen 1991, Xue and Chao 1998), Michelsen (2000) tentatively erected four major subgroups, the subfamilies Anthomyiinae, Myopininae, and Pegomyinae, and *Phaonantho* Albuquerque genus-group (Michelsen 2000), based on morphological cladistic analysis.

Notwithstanding the economic and ecological significance, only few molecular studies have treated the Anthomyiidae (Gomes et al. 2021; Kutty et al. 2008, 2010, 2019). In recent years, several researchers have investigated the internal relationships among diverse species of Calyptratae. Mitochondrial and nuclear rDNA genes have been used for phylogenetic analysis that included representatives of Anthomyiidae (Ding et al. 2015; Zhang et al. 2015; Li et al. 2022). Nonetheless, the limited sampling of anthomyiids precludes a thorough testing of classification and phylogeny within the family. Additionally, the use of partial genes in prior investigations also failed to conclude reliable phylogenetic relationships within Anthomyiidae (Kutty et al. 2008, 2010). Consequently, phylogenetic relationships within the family remain ambiguous, highlighting the need for more comprehensive phylogenetic information derived from longer DNA sequences such as complete mitochondrial genomes.

Mitochondrial genomes have been shown to supply an increase in molecular information content as compared to individual genes, making them conducive to investigations of phylogeny and evolution across a broad diversity of insects (Cameron 2014). Characteristics such as coding gene conservation, maternal inheritance, rare recombination and rapid evolutionary rate make mtDNA an appropriate marker for species identification and molecular evolutionary studies of Anthomyiidae (Ding et al. 2015; Zhang et al. 2015; Li et al. 2022). Meanwhile, diverse levels of genetic pattern and rate variation, for instance, nucleotide composition, codon usage and nucleotide substitution (Gibson et al. 2004; Jia and Higgs 2007), have also been extensively utilized for comparative and phylogenetic analyses. Still relatively few studies employ mitogenomes to reconstruct the phylogeny of Anthomyiidae. The number of mitogenomes from Anthomyiidae deposited in GenBank has increased gradually over time. As of May 2023, only 11 complete Anthomyiidae mitogenomes had been reported on GenBank, representing three subfamilies, with subfamilies Myopininae and Pegomyinae represented by only a single sequenced species.

To expand the available coverage of anthomyiid mitogenomes for comparison and analysis across various taxonomic levels, we sequenced multiple newly sampled anthomyiid mitogenomes to compare these with publicly available sequences. We used a method of next-generation sequencing of multispecies pooled genomic DNA to acquire mitogenomes for 18 anthomyiids, belonging to three subfamilies: Anthomyiinae (eleven species), Pegomyinae (six species) and Myopininae (one species). Additionally, we constructed phylogenetic relationships using maximum likelihood (ML) and Bayesian inference (BI) methods, to investigate higher-level phylogeny within Anthomyiidae. This approach provides novel insights into the phylogenetics and classification of Anthomyiidae and can be used to support their morphological identification.

2. Materials and Methods

2.1. Sampling Collection and Identification

All anthomyiids were captured by malaise traps in the Baihua Mountain (39°50′11.04″N, 115°34′41.52″E) and Dalaoling National Natural Reserve (31°4′35.6″N and 110°56′11.6″E), from 2017 to 2019 in China. All experimental materials were preserved in absolute ethanol and cryopreserved at –20°C until further processing in the Museum of Beijing Forestry University (BFU), Beijing, China. Specimens of Anthomyiidae were initially identified by Mingfu Wang using available taxonomic keys (Xue and Chao 1998), and identifications were confirmed using DNA barcodes (*cox1*) obtained from the assembled mitogenomes held in public databases (i.e., BOLD, NCBI) and confirmed by BLAST search to the genus level (Michelsen 2011). All Anthomyiidae mitogenome data from NCBI were downloaded and employed in comparative mitogenomic analyses with the 18 new mitogenomes in this study (Table 1).

Table 1. Taxonomic information and GenBank accession numbers of mitochondrial genomes used in the study. *Species documented in this study.

Family	Subfamily	Species	Accession No.
Anthomyiidae	Anthomyiinae	*Anthomyia confusanea*	OP616801*
		Anthomyia illocata	MW296030
		Anthomyia oculifera	OP616786*
		Anthomyia pluvialis	OP616785*
		Anthomyia procellaris	MT584110
		Botanophila fugax	MT410801
		Botanophila sp.	OP616795*
		Delia antiqua	NC028226
		Delia longitheca	OP616787*
		Delia platura	MT483617
		Delia takizawai	OP616791*
		Eustalomyia hilaris	OP616792*
		Eustalomyia vittipes	OP616796*
		Fucellia costalis	MH823369
		Hydrophoria lancifer	OP616790*
		Hydrophoria linogrisea	MT483657
		Hylemya vagans	MT410822
		Hylemyza partita	MT584149
		Hyporites sp.	OP616793*
		Leucophora shanxiensis	OP616797*
	Myopininae	*Pegoplata annulata*	OP616788*
		Pegoplata infirma	MT410786
	Pegomyinae	*Emmesomyia oriens*	OP616789*
		Eutrichota similis	OP616798*
		Pegomya bicolor	MT410802
		Pegomya exilis	OP616794*
		Pegomya flaviprecoxa	OP616799*
		Pegomya quadrivittata	OP616784*
		Pegomya sp.	OP616800*
Outgroups			
Calliphoridae	Luciliinae	*Lucilia sericata*	AJ422212
Drosophilidae	Drosophilinae	*Drosophila mercatorum*	MK575470
Fanniidae		*Fannia scalaris*	MT017706
Muscidae	Muscinae	*Musca domestica*	NC024855
Sarcophagidae	Sarcophaginae	*Sarcophaga crassipalpis*	NC026667
Scathophagidae	Scathophaginae	*Scathophaga inquinata*	MT483619
		Scathophaga stercoraria	KM200724
Tachinidae	Phasiinae	*Subclytia rotundiventris*	MN199029

2.2. DNA Extraction, Mitogenomes Sequencing, Assembly and Annotation

We used the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany) according to manufacturer's protocol for DNA extracted from individual adult flies. Qubit 3.0 was used to quantify the concentration of the DNA samples. To enhance sequencing efficiency and minimize resource waste, hybrid libraries were adopted (Gillett et al. 2014). Subsequently, the genomic DNA was pooled and then sequenced on the Illumina Novaseq 6000 platform (PE150, Illumina, San Diego, CA). Raw reads were trimmed using Trimmomatic (Bolger et al. 2014), with each library yielding approximately 5 Gb of clean data. These were assembled de novo using IDBA-1.1.1 (Peng et al. 2012). To identify mitogenomes, two sequence fragments of mtDNA (*cox1* and *cytb*) (Crampton-Platt et al. 2015; Yan et al. 2019) were amplified as bait sequences to acquire the best-fitting mitochondrial scaffolds using Basic Local Alignment Search Tool (BLAST) with a similarity threshold of 98% (Altschul et al. 1990). The 13 protein-coding genes (PCGs) and two ribosomal RNA genes (rRNAs) were annotated using Geneious v2020.0.2 by alignment to other reported Calyptratae flies for each orthologous gene (Kearse et al. 2012). Positional annotation of 22 transfer RNA genes (tRNAs) was achieved using the online MITOS tool (Bernt et al. 2013). Complete mitochondrial genomes were submitted to NCBI under the accession numbers of OP616784-OP616801. The

associated SRA, BioProject, and Bio-Sample numbers are SRR25463435-SRR25463439, PRJNA1000204, and SAMN36763070-SAMN36763087, respectively.

2.3. Sequence Analyses

Sequence comparisons were carried out in PhyloSuite software (Zhang et al. 2020) to estimate nucleotide composition and relative synonymous codon usage (RSCU) among the 18 newly sequenced mitochondrial genomes. Base composition skewness analysis was calculated on all available anthomyiid mitogenomes using the specific formulas: AT-skew = (A - T) / (A + T) and GC-skew = (G - C) / (G + C) (Perna et al. 1995). Nucleotide divergence (Pi) value of three subfamilies was computed through DnaSP v6. (Rozas et al. 2017). Additionally, the ratios of Ka (nonsynonymous substitutions)/Ks (synonymous substitutions) based on 13 aligned PCGs were also measured with DnaSP v6 to compare substitution rate (Rozas et al. 2017). The Kimura 2-parameter model in MEGA 5 was used in calculations of mean genetic distances among the three subfamilies (Tamura et al. 2011). DAMBE 7.0 was applied to assess the substitution saturation (*Iss*) of each codon position based on all PCGs under the GTR model (Xia 2018).

2.4. Phylogenetic Analyses

The 29 complete mitogenomes from three subfamilies of Anthomyiidae were chosen to construct the phylogenetic tree, including 18 new mitogenomes documented in this study. Eight outgroups were selected to represent seven outgroup families of Diptera, with the placed between Drosophilidae (*Drosophila mercatorum*) and all other sampled flies. Phylogenetic relationships were inferred from analyses of a dataset of the 13 mitochondrial PCGs. To construct this dataset, each PCG of 37 mitogenomes was individually aligned using MAFFT (Katoh and Standley 2013). The optimal partitioning schemes and best-fitting model for each PCG were obtained by PartitionFinder 2 (Lanfear et al. 2017). Phylogenetic analyses (ML and BI) were performed on a concatenated 13 PCG dataset using the online CIPRES Science Gateway (Miller et al. 2010). For ML analysis, the node support values were inferred by ultrafast bootstrap resampling (BP) with 1000 replicates in IQ-TREE. Two separate Markov chain Monte Carlo (MCMC) chains were carried out for BI analyses, spanning 10 million generations simultaneously, with sampling occurring every 1000 iterations. In Bayesian analyses, posterior probabilities (PPs) were calculated after discarding the initial 25% samples as burn-in. Convergence was assessed by confirming that the average standard deviation of split frequencies was less than 0.01 in MrBayes 3.2.6 and effective sample size (ESS) was greater than 200 in Tracer (Ronquist et al. 2012; Rambaut et al. 2018). Phylograms were modified and visualized using FigTree v 1.4.

3. Results and discussion

3.1. Mitogenome organization

Our newly sequenced mitogenomes of Anthomyiidae show some variation in genome size, ranging from 15,635 bp to 21,098 bp in length. They are compact circular, double-stranded molecules, and are composed of the core 37 genes and a control region. The majority strand (J-strand) encoded 23 genes (9 PCGs, and 14 tRNAs), while the remaining genes were transcribed on the minority strand (N-strand) (4 PCGs, 8 tRNAs, and 2 rRNAs). All newly sequenced Anthomyiidae mitogenomes were conserved in gene order and orientation, consistent with previously published Muscoidea mitogenomes (Ren et al. 2019; Oliveira et al. 2008; Li et al. 2016). All PCGs began with a typical start codon (ATN), except for the *cox1* initiated with TCG. Furthermore, most PCGs ended with the termination codons TAA/TAG, the occurrence of the TAA is more frequently observed than TAG, while three PCGs (*cox2*, *nad4* and *nad5*) terminated with T, which is a common phenomenon in Calyptratae (Ren et al. 2019; Oliveira et al. 2008; Li et al. 2016; Li et al. 2020; Yan et al. 2021b; Zhao et al. 2013).

Relative synonymous codon usage (RSCU) values of all three subfamilies are illustrated in Figure 1. All PCGs encoded 22 standard amino acids. The most continually encoded amino acids in Anthomyiidae are Ala, Arg, Gly, Leu2, Pro, Ser2, Thr and Val, with Ser2 in possession of the highest value across all three subfamilies. UUA (Leu2) is the most frequently utilized codon among all sampled anthomyiids.

3.2. High degree of nucleotide heterogeneity and contrasting rates of evolution

The AT content of PCGs from all anthomyiids varies between 76.7% (Anthomyiinae) and 77.7% (Pegomyinae) with a mean value of 77%, whereas the basic composition of all PCGs is homogeneous. Third codon positions possess a substantially higher AT content than first and second codon positions, according to analyses of the average base composition at each codon position (Table 2). By contrasting the AT content of each PCG across all Anthomyiidae, *nad6* (84.16%) shows the highest mean value, followed by *atp8* (83.24%) and *nad4L* (82.16%). On the other hand, the average value of *cox1* (70.85%) and *cox3* (71.76%) is the lowest. Across all the 13 PCGs of Anthomyiidae, the AT-skew is negative, but is highest in Anthomyiinae (-0.152) and lowest in Myopininae (-0.157); whereas the GC-skew is positive, ranging from 0.028 to 0.038, and remains consistent across both Anthomyiinae and Myopininae (Fig. 2). This pattern resembles that observed in most Calyptratae mitogenomes, and suggests that strand bias may be a dissymmetric mutation occurring during DNA replication (Ren et al. 2019; Oliveira et al. 2008; Li et al. 2016; Li et al. 2020; Yan et al. 2021b; Zhao et al. 2013).

Table 2. Nucleotide composition of mitochondrial genomes of anthomyiid flies at subfamily level.

Regions	Feature	Anthomyiinae	Myopininae	Pegomyinae
Whole genome	A+T(%)	78.6	78.4	80.5
	AT-Skew	0.016	0.007	0.021
	GC-Skew	–0.166	–0.154	–0.137
PCGs	A+T(%)	76.7	77.2	77.7
	AT-Skew	–0.152	–0.157	–0.154
	GC-Skew	0.028	0.028	0.038
1st codon	A+T(%)	70.9	71.1	71.5
	AT-Skew	–0.125	–0.137	–0.134
	GC-Skew	0.187	0.193	0.205
2nd codon	A+T(%)	71.0	71.0	71.4
	AT-Skew	–0.305	–0.303	–0.306
	GC-Skew	–0.135	–0.141	–0.134
3rd codon	A+T(%)	88.3	89.5	90.2
	AT-Skew	–0.050	–0.057	–0.049
	GC-Skew	0.037	0.039	0.054

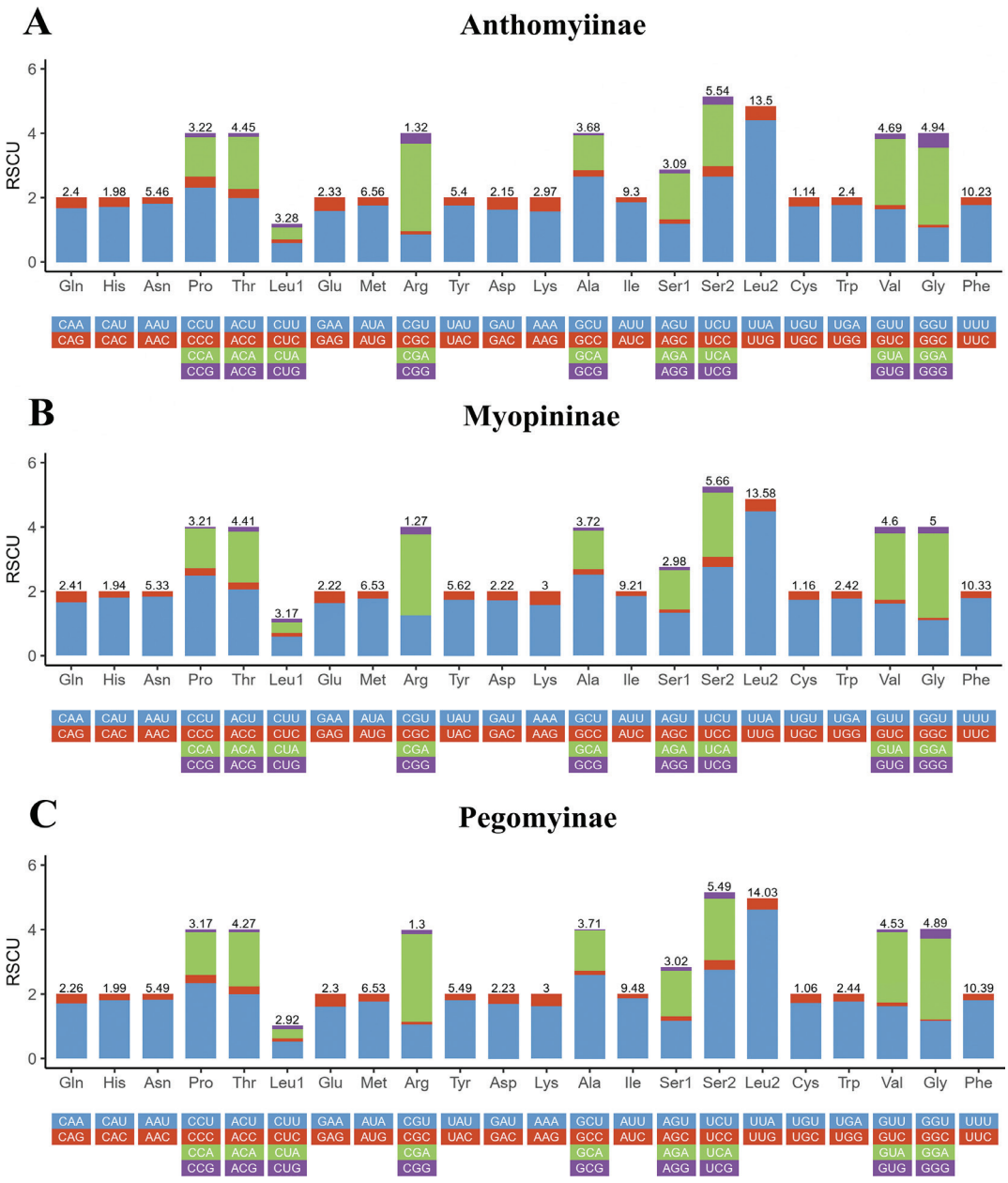

Figure 1. Relative synonymous codon usage (RSCU) in the mitogenomes of three subfamilies.

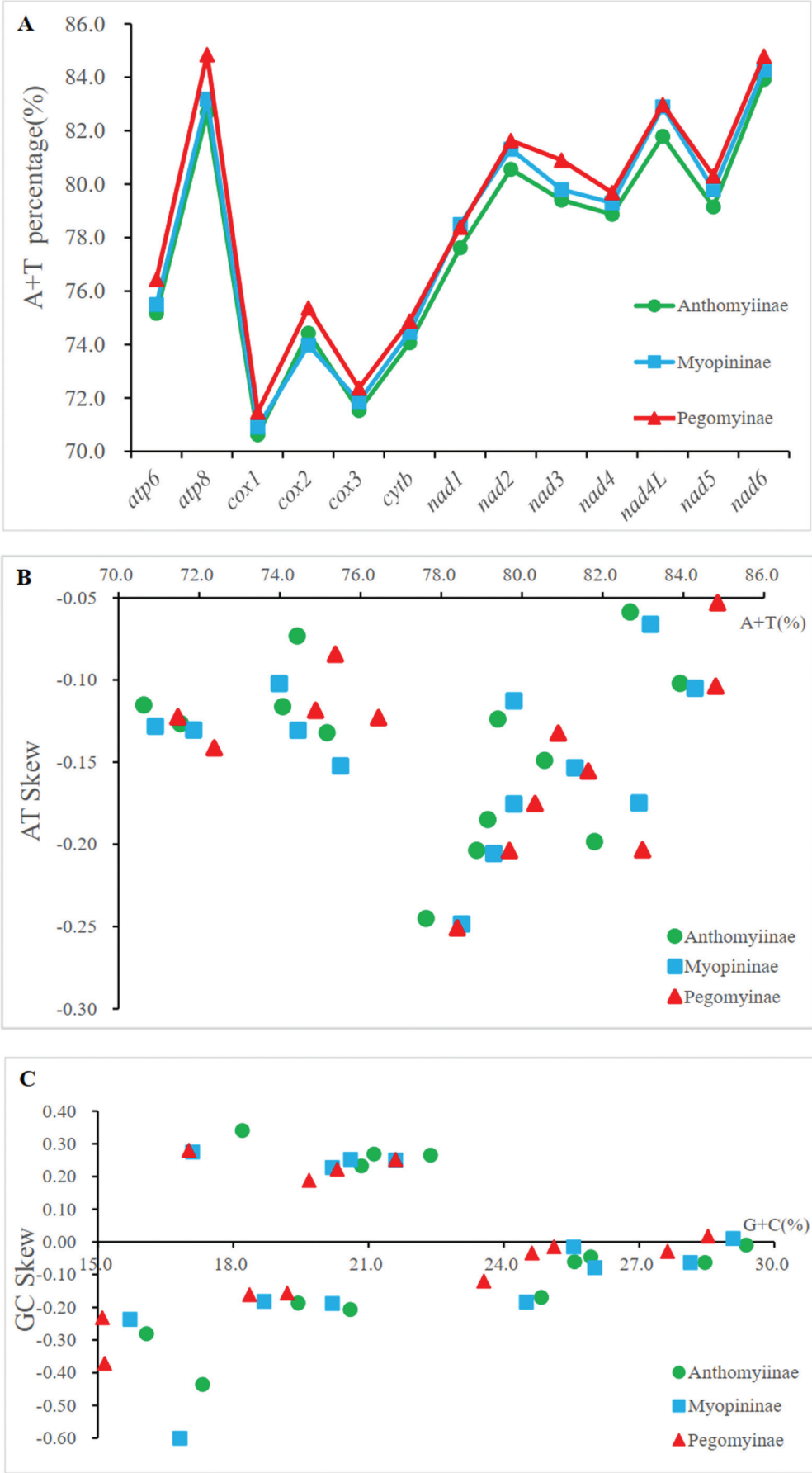

Figure 2. Nucleotide composition analysis of mitochondrial genomes from three subfamilies of Anthomyiidae: A + T percentage of the 13 protein-coding genes (**A**) and the corrections between A + T% vs. AT skew (**B**) and G + C% vs. GC skew (**C**) in the 13 protein-coding genes.

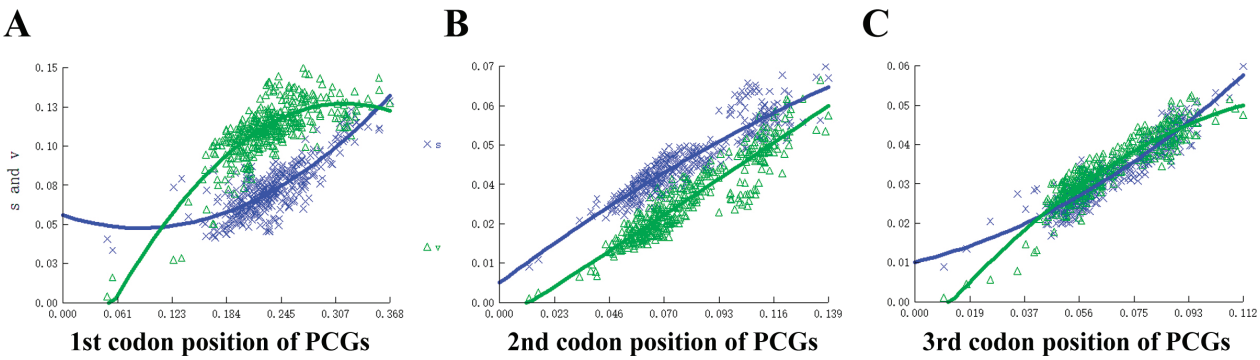

Figure 3. Nucleotide substitution saturation plots of all 13 mitochondrial protein-coding genes. **A** The 1st codon positions. **B** The 2nd codon positions. **C** The 3rd codon positions. Plots in blue and green indicate transition and transversion, respectively.

Additionally, saturation plots reveal that only the third codon locations in all of the PCGs exhibit notable heterogeneity, implying that levels of heterogeneity in the PCG123 are exceedingly low (Fig. 3). There is also no evidence of saturation from sequence comparisons including the more freely evolving third codon position (Yan et al. 2019), thereby supporting the suitability of nucleotide analyses for phylogenetic reconstruction at this level.

To explore sequence evolution among the 13 PCGs sampled in anthomyiids, the values of Ka, Ks, and Ka/Ks (ω) for each PCG were computed, respectively (Fig. 4). The range of Ka for the typical gene-specific substitution rates was 0.034 (*cox2*) to 0.375 (*cox1*). The gene (*nad1*) showed the highest evolutionary rate (ω = 1.001) of all the PCGs, indicating that it is likely undergoing positive or relaxed selection pressure. In contrast, *cox2* displayed the lowest value (ω = 0.108), reflecting that may be subject to strong purifying selection. It is feasible that weak or sporadic positive selection may occur in this environment with strong purifying selection when lifestyle changes result in increased energy demands or reduced oxygen availability. As a result, phylogenetic reconstruction could take advantage of all PCGs. Besides, the model of evolution among 13 PCGs was mostly in line with the previous literature (Yan et al. 2021b).

Nucleotide diversity among the 13 PCGs is shown in Figure 5. The four PCGs with marked variation were *cox1* (Pi = 0.39), *nad3* (Pi = 0.354), *nad1* (Pi = 0.333), and *nad5* (Pi = 0.306), while *cox2* (Pi = 0.098), *nad4* (Pi = 0.099), *cox3* (Pi = 0.106), and *cytb* (Pi = 0.109) exhibited relatively low Pi values, demonstrating that they are the most conserved genes among the 13 PCGs. Genetic distance analyses also show an analogous tendency (Fig. 5). The mean value of genetic distances within 29 mitogenomes shows that *cox1* (mean value = 0.842), *nad3* (0.704) and *nad1* (0.693) have experienced a comparatively rapid evolution. Inversely, *cox2* (0.106), *nad4* (0.106) and *cox3* (0.115) with lower measured distances are evolving relative slowly.

Overall, we find that the *cox1* gene evolves at a considerably higher rate and under comparatively relaxed purifying selection among anthomyiids, manifesting that *nad1* gene could be a suitable candidate marker for clarifying the phylogenetic relationships among taxa with morphological traits that are difficult to interpret.

3.3. Phylogenetic analyses

Phylogenetic analyses were carried out on mitogenomes from 29 anthomyiids and eight outgroups, including our 18 newly sequenced Anthomyiidae mitogenomes. Both ML and BI methods were conducted on 13 PCGs and produced completely resolved trees with identical topologies and with most branches receiving strong support. The muscoids were confirmed as a non-monophyletic group or grade, with Scathophagidae plus Anthomyiidae placed as sister to the clade Oestroidea ((Sarcophagidae + Calliphoridae) + Tachinidae), congruent with previous studies (Kutty et al. 2010; Yan et al. 2021a). While the number of clusters within Calyptratae appear to be reliably established, there are still many partial subordinate taxa with weak support or unstable nodes. It is widely known that limited taxon sampling can lead to phylogenetically biased results, and rapid radiation can make resolution of relationships more difficult to resolve (Wiegmann et al. 2011), and this may explain the challenges posed in resolving the muscoid radiation. In our trees, Anthomyiidae was recovered as monophyletic and placed as sister group to the Scathophagidae (BP = 86, PP = 1.00; Fig. 6), in agreement with a recent molecular phylogenetic analyses (Gomes et al. 2021). This finding contrasts with a phylogenomic study using nuclear markers in which Scathophagidae was nested within Anthomyiidae (Buenaventura et al. 2020). In all of these studies, including our own, the Scathophagidae have been represented by limited taxon sampling.

Anthomyiidae and Scathophagidae are clearly very closely related, with various forms of evidence supporting either a sister-group relationship or placement of the latter family within a more broadly defined Anthomyiidae. They are morphologically similar, with both families possessing a long anal vein, usually reaching wing edge at least as a fold (Buck et al. 2009). Reliance on male genitalia to support taxonomy and classification of anthomyiids has made classification of the group technically challenging. Several researchers have provided morphological evidence for the monophyly of Anthomyiidae (Michelsen, 1991; Xue and Chao, 1998), while an increasing number of molecular phylogenetic analyses have found Anthomyiidae to be paraphyletic, containing Scathophagidae (Kutty et al. 2010, 2019). In the Kutty

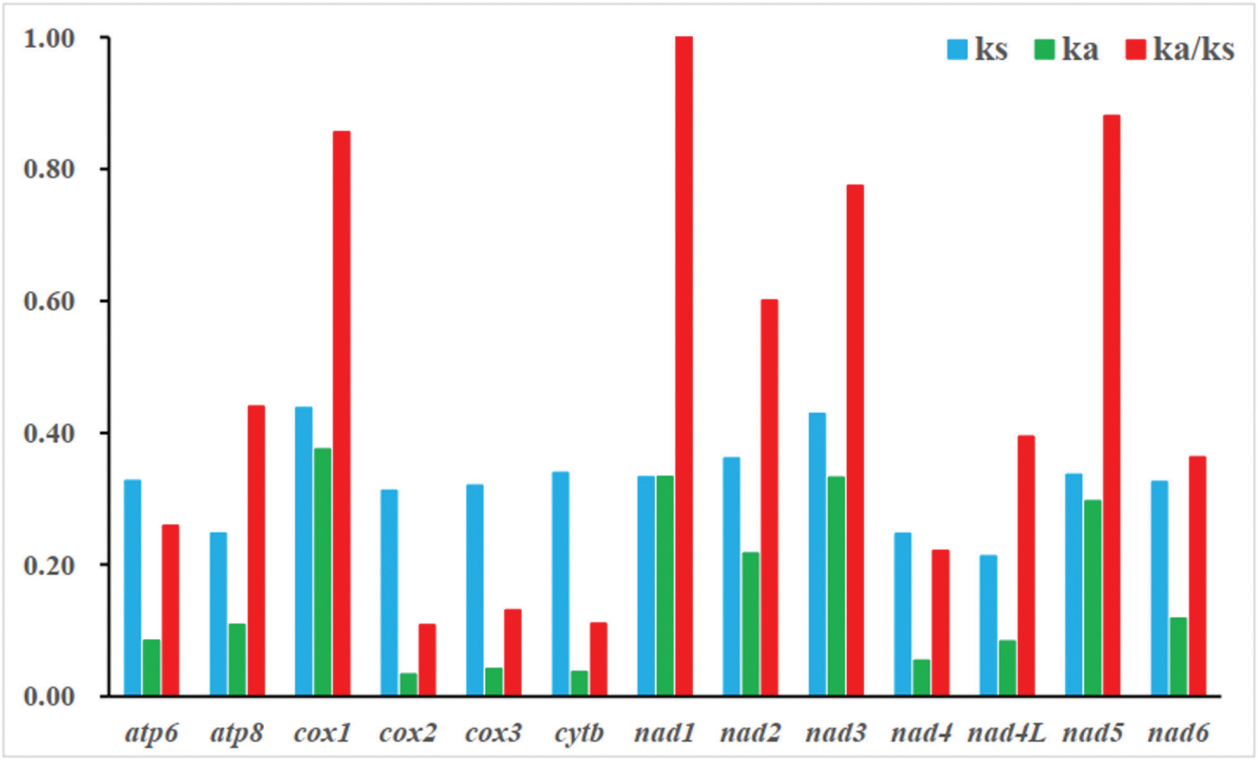

Figure 4. Evolutionary rates of anthomyiid mitogenomes. The non-synonymous substitutions rate (Ka), the synonymous substitutions rate (Ks), and the ratio of the rate of non-synonymous substitutions to the rate of synonymous substitutions (Ka/Ks) for each PCG.

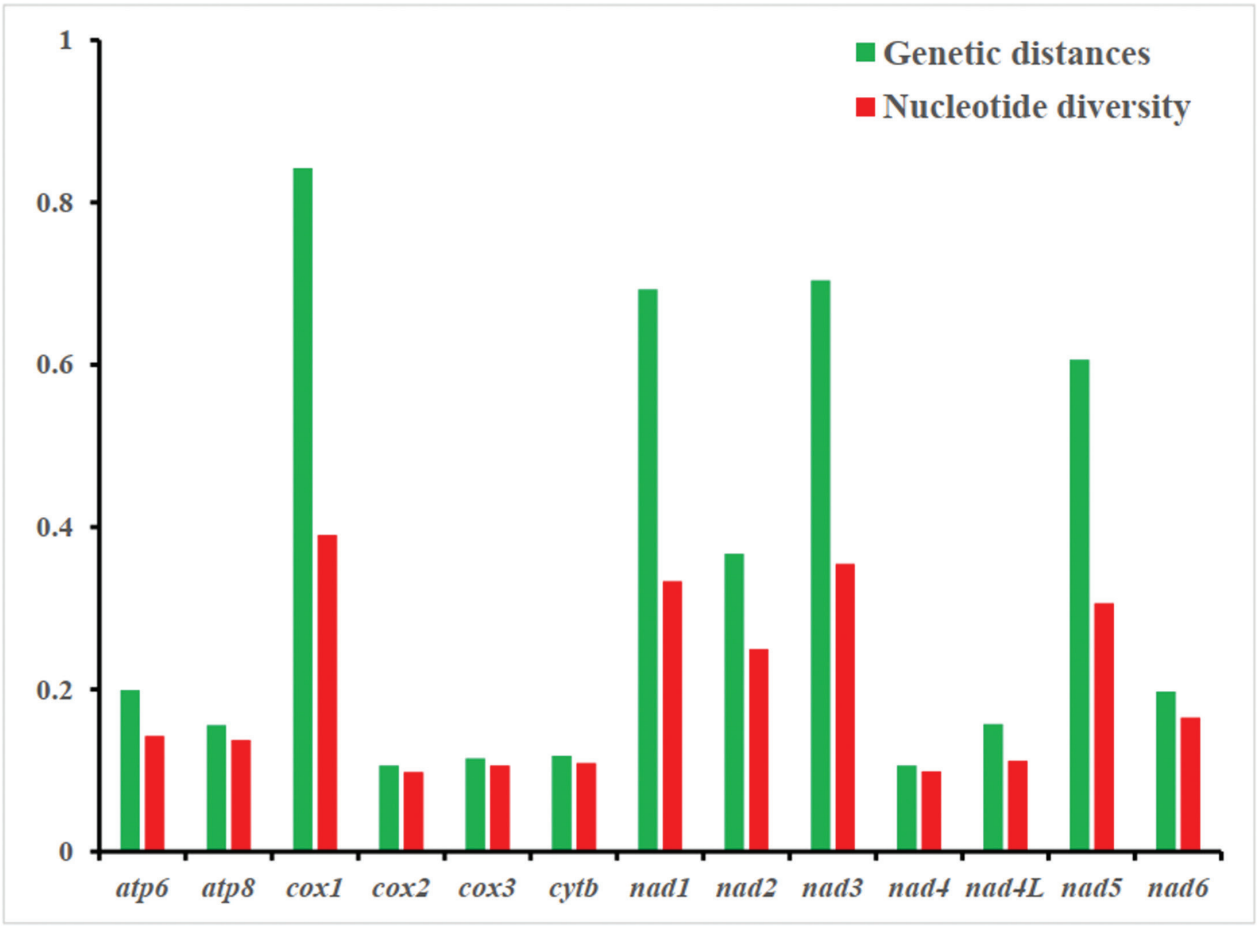

Figure 5. Nucleotide diversity (Pi) and genetic distances of the 13 protein-coding genes of Anthomyiidae.

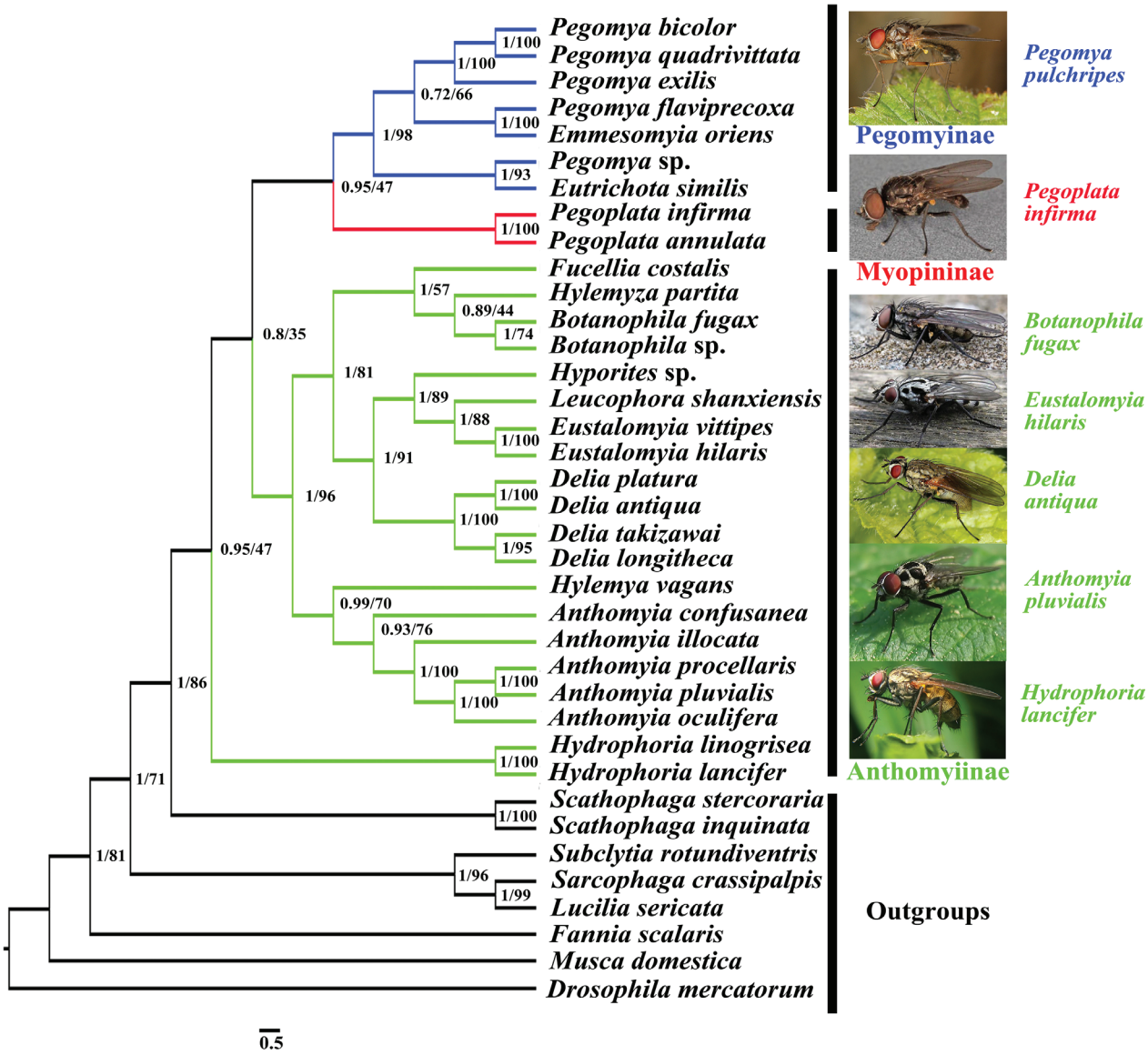

Figure 6. Inferred phylogenetic tree from ML and BI methods based on the concatenated 13 protein-coding genes among 29 Anthomyiidae species. Node values nodes are posterior probabilities (PP) / bootstrap support values (BS) based on 1000 replicate searches.

et al. (2010) analysis, the subfamily Anthomyiinae was found to be paraphyletic, comprising 20 species belonging to 10 genera. In addition, Myopininae and Pegomyinae were recovered as monophyletic sister taxa and placed as the sister group of Anthomyiinae, excluding *Hydrophoria*. *Hydrophoria* formed a distinct early diverging branch, sister to all other anthomyiids, but with low support (BP = 47, PP = 0.95). Relationships among anthomyiid genera, for example (((*Eustalomyia*+*Leucophora*)+*Hyporites*)+*Delia*), were strongly supported and as the sister group to the clade ((*Botanophila*+*Hylemyza*)+*Fucellia*) (Kutty et al. 2010). More recently, Gomes et al. (2021) also recovered *Botanophila* as a sister group of the genus *Hylemyza*. The specimen labelled *Botanophila* sp. was difficult to determine morphologically, and here it grouped as sister group to *B. fugax*. Multiple *Pegomya* species were polyphyletic, and *Emmesomyia oriens* emerged as sister to *P. flaviprecoxa*, with robust support (BP = 100, PP = 1.00). Griffiths (1982) proposed that the condition of a bilobate pregonite is a synapomorphy uniting *Pegomya*, *Emmesomyia* and *Eutrichota*, and this was supported by molecular data (Kutty et al. 2010). It is still noteworthy that several anthomyiid taxa, including *Anthomyia*, *Botanophila*, and *Delia*, are extremely diverse and presumably paraphyletic (Kutty et al. 2008, 2010). Nevertheless, only a single species in each of these genera was used in these studies, thus increased sampling will be necessary to adequately resolve relationships among diverse Anthomyiidae genera.

4. Conclusions

In this study, we provide a systematic analysis of 18 mitogenomes representing three subfamilies of Anthomyiidae. This is the first report of mitogenomes from the three

genus *Eustalomyia*, *Hyporites* and *Leucophora* of the subfamily Anthomyiinae, and two genus *Emmesomyia* and *Eutrichota* of the subfamily Pegomyinae. Our study reveals conserved traits among anthomyiid mitogenomes, including strongly biased A + T richness, a more rapidly evolving *nad1* gene and a positive GC skew among the 13 PCGs. Both ML and BI phylogenetic trees using the 13 PCGs yield an identical topology, with most divergences possessing strong bootstrap and posterior probability support. These results provide fundamental information on mitogenome organization and reinforce an increased understanding of phylogenetic relationships within the family Anthomyiidae.

Author Contributions

Conceptualization, D.Z.; Methodology, H.L., W.P. and M.W.; Software, H.L. and W.P.; Validation, H.L.; Formal Analysis, H.L. and M.W.; Investigation, B.C., H.P., R.C., M.Z., J.Y. and X.Z.; Resources, B.C., H.P., R.C., M.Z., J.Y. and X.Z.; Data Curation, H.L., W.P. and M.W.; Writing—Original Draft Preparation, H.L. and W.P.; Writing—Review and Editing, H.L.; Visualization, W.P. and M.W.; Supervision, D.Z.; Project Administration, D.Z.; Funding Acquisition, D.Z. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in the study are deposited in the NCBI database repository, accession numbers: OP616784-OP616801. The associated SRA, BioProject, and Bio-Sample numbers are SRR25463435-SRR25463439, PRJNA1000204, and SAMN36763070-SAMN36763087, respectively.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgements

Prof. Brian Wiegmann and two anonymous reviewers are acknowledged for their contribution in improving this manuscript. This research was funded by the National Natural Science Foundation of China (No. 32170450, 31872964), the Beijing Forestry University Outstanding Young Talent Cultivation Project (No. 2019JQ0318) and the Beijing Forestry University Outstanding Postgraduate Mentoring Team Award.

References

Altschul SF, Gish W, Miller W, Myers EW. Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler P (2013) MITOS: Improved de novo metazoan mitochondrial genome annotation. Molecular Phylogenetics and Evolution 69: 313–319. https://doi.org/10.1016/j.ympev.2012.08.023

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

Buck M, Woodley NE, Borkent A, Wood DM, Pape T, Vockeroth J, Michelsen V, Marshall SA (2009) Kye to Diptera families – adults. In: Brown BV, Borkent A, Cumming JM, Wood DM, Woodley NE, Zumbado MA (eds), Manual of Central American Diptera, Vol. 1. Ottawa: NRC Research Press: 95–144.

Buenaventura E, Lloyd MW, Perilla López JM, González VL, Thomas-Cabianca A, Dikow T (2020) Protein-encoding ultraconserved elements provide a new phylogenomic perspective of Oestroidea flies (Diptera: Calyptratae). Systematic Entomology 46: 5–27. https://doi.org/10.1111/syen.12443

Cameron SL (2014) Insect mitochondrial genomics: implications for evolution and phylogeny. Annual Review of Entomology 59: 95–117. https://doi.org/10.1146/annurev-ento-011613-162007

Córdova-García G, Navarro-de-la-Fuente L, Pérez-Staples D, Williams T, Lasa R (2023) Biology and Ecology of *Delia planipalpis* (Stein) (Diptera: Anthomyiidae), an Emerging Pest of Broccoli in Mexico. Insects 14: 659–674. https://doi.org/10.3390/insects14070659

Crampton-Platt A, Timmermans M, Gimmel M, Kutty S, Cockerill T, Khen C, Vogler T (2015) Soup to tree: the phylogeny of beetles inferred by mitochondrial metagenomics of a bornean rainforest sample. Molecular Biology and Evolution 32: 2302–2316. https://doi.org/10.1093/molbev/msv111

Ding S, Li X, Wang N, Cameron SL, Mao M, Wang Y, Xi Y, Yang D (2015) The phylogeny and evolutionary timescale of muscoidea (Diptera: Brachycera: Calyptratae) inferred from mitochondrial genomes. Plos One 10: e0134170. https://doi.org/10.1371/journal.pone.0134170

Gibson A, Gowri–Shankar VG, Higgs PG, Rattray MA (2004) Comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods. Molecular Biology and Evolution 22: 251–264. https://doi.org/10.1093/molbev/msi012

Gillett CPDT, Crampton-Platt A, Timmermans MJTN, Jordal BH, Emerson BC, Vogler AP (2014) Bulk de novo mitogenome assembly from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionoidea). Molecular Biology and Evolution 31: 2223–2237. https://doi.org/10.1093/molbev/msu154

Gomes LRP, Souza DdS, Carvalho CJB (2021) First Insights into the Evolution of Neotropical Anthomyiid Flies (Diptera: Anthomyiidae). Systematics and Biodiversity 19: 724–737. https://doi.org/10.1080/14772000.2021.1914765

Griffiths GCD (1984) Anthomyiidae [part]. Flies of the Nearctic Region: Cyclorrhapha II Schizophora: Calyptratae. Vol. 8, pp: 289–408.

Grisales D, Lopes AC, Carvalho CJBD (2016) Family Anthomyiidae. Zootaxa 4122: 803–806. https://doi.org/10.11646/zootaxa.4122.1.68

Hao YJ, Zhang YJ, Si FL, Fu DY, He ZB, Chen B (2016) Insight into the possible mechanism of the summer diapause of *Delia antiqua* (Diptera: Anthomyiidae) through digital gene expression analysis. Insect Science 23: 438–451. https://doi.org/10.1111/1744-7917.12323

Jia W, Higgs PG (2007) Codon usage in mitochondrial genomes: distinguishing contextdependent mutation from translational selection. Molecular Biology and Evolution 25: 339–351. https://doi.org/10.1093/molbev/msm259

Katoh K, Standley D (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30(4): 772–780. https://doi.org/10.1093/molbev/mst010

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kutty SN, Pape T, Pont A, Wiegmann BM, Meier R (2008) The Muscoidea (Diptera: Calyptratae) are paraphyletic: Evidence from four mitochondrial and four nuclear genes. Molecular Phylogenetics and Evolution 49: 639–652. https://doi.org/10.1016/j.ympev.2008.08.012

Kutty SN, Pape T, Wiegmann BM, Meier R (2010) Molecular phylogeny of the Calyptratae (Diptera: Cyclorrhapha) with an emphasis on the superfamily Oestroidea and the position of Mystacinobiidae and McAlpine's fly. Systematic Entomology 35(4): 614–635. https://doi.org/10.1111/j.1365-3113.2010.00536.x

Kutty SN, Meusemann K, Bayless K, Marinho M, Pont A, Zhou X, Misof B, Wiegmann B, Yeates D, Cerretti P (2019) Phylogenomic analysis of Calyptratae: resolving the phylogenetic relationships within a major radiation of Diptera. Cladistics 35(6): 605–622. https://doi.org/10.1111/cla.12375

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

Li H, Yan L, Pei W, Hu Y, Wang A, Wang Z, Zhang D (2022) First Mitogenome of *Anthomyia illocata* (Diptera, Anthomyiidae) Yielded by next-Generation Sequencing. Mitochondrial DNA Part B 7: 875–877. https://doi.org/10.1080/23802359.2022.2076626

Li X, Wang Y, Su S, Yang D (2016) The Complete Mitochondrial Genomes of *Musca domestica* and *Scathophaga stercoraria* (Diptera: Muscoidea: Muscidae and Scathophagidae). Mitochondrial DNA 27: 1435–1436. https://doi.org/10.3109/19401736.2014.953080

Li XY, Yan LP, Pape T, Gao YY, Zhang D (2020) Evolutionary insights into bot flies (Insecta: Diptera: Oestridae) from comparative analysis of the mitochondrial genomes. International Journal of Biological Macromolecules 149: 371–380. https://doi.org/10.1016/j.ijbiomac.2020.01.249

Michelsen (1991) Revision of the aberrant New World genus Coenosopsia (Diptera: Anthomyiidae), with a discussion of anthomyiid relationships. Systematic Entomology 16: 85–104.

Michelsen V (2000) Oldest authentic record of a fossil calyptrate fly (Diptera): a species of Anthomyiidae from early Coenozoic Baltic amber. Studia dipterologica 7: 11–18.

Michelsen V (2010) Anthomyiidae (Anthomyiid Flies). In: Brown BV, Borkent A, Cumming JM, Wood DM, Woodley NE, Zumbado M. (Eds.), Manual of Central American Diptera. Vol. 2. National Research Council Press, Ottawa: 1271–1276.

Michelsen V (2011) Anthomyiidae. In: Pape T (Ed.) Diptera Brachycera. Fauna Europaea ver. 2.4. [Ver. 2.6.2 available from http://www.faunaeur.org]

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the 2010 Gateway Computing Environments Workshop (GCE), New Orleans, LA, USA, 14 November; pp: 1–8.

Moretti E, Wickings K, Nault B (2021) Environmental factors and crop management that affect *Delia antiqua* damage in onion fields. Agriculture, Ecosystems and Environment 314: 107420. https://doi.org/10.1016/j.agee.2021.107420

Oliveira MT, Barau JG, Junqueira ACM, Feijão PC, Rosa Acda, Abreu CF, Azeredo-Espin AML, Lessinger AC (2008) Structure and Evolution of the Mitochondrial Genomes of Haematobia Irritans and Stomoxys Calcitrans: The Muscidae (Diptera: Calyptratae) Perspective. Molecular Phylogenetics and Evolution 48: 850–857. https://doi.org/10.1016/j.ympev.2008.05.022

Peng Y, Leung HCM, Yiu SM, Chin FYL (2012) IDBA-UD: A de novo assembler for single cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28, 1420–1428. https://doi.org/10.1093/bioinformatics/bts174

Perna N, Kocher T (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. Journal of Molecular Evolution 41: 353–358. https://doi.org/10.1007/BF012-15182

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

Ren L, Shang Y, Yang L, Shen X, Chen W, Wang Y, Cai J, Guo Y (2019) Comparative Analysis of Mitochondrial Genomes among Four Species of Muscid Flies (Diptera: Muscidae) and Its Phylogenetic Implications. International Journal of Biological Macromolecules 127: 357–364. https://doi.org/10.1016/j.ijbiomac.2019.01.063

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard M, Huelsenbeck J (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61: 539–542. https://doi.org/10.1093/sysbio/sys029

Rozas J, Ferrer-Mata A, Sanchez-DelBarrio J, Guirao-Rico S, Librado P, Ramos-Onsins S, Sanchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. Molecular Biology and Evolution 34: 3299–3302. https://doi.org/10.1093/molbev/msx248

Sachet JM, Roques A, Despres L (2006) Linking patterns and processes of species diversification in the cone flies *Strobilomyia* (Diptera: Anthomyiidae). Molecular Biology and Evolution 41: 606–621. https://doi.org/10.1016/j.ympev.2006.06.005

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28: 2731–2739. https://doi.org/10.1093/molbev/msr121

Wang M, Michelsen V, Li K, Zhu W (2014) Supplementary Catalogue of the Anthomyiidae (Diptera) of China. Zookeys 453: 71–109. https://doi.org/10.3897/zookeys.453.8282

Wang QK, Yang YZ, Liu MQ, Zhang D (2014) Fine structure of *Delia platura* (Meigen) (Diptera: Anthomyiidae) revealed by scanning electron microscopy. Microscopy Research and Technique 77, 619–630. https://doi.org/10.1002/jemt.22380

Wiegmann BM, Trautwein MD, Winkler IS, Barr NB, Kim JW, Lambkin C, Bertone MA, Cassel BK, Bayless KM, Heimberg AM, Wheeler BM, Peterson KJ, Pape T, Sinclair BJ, Skevington JH, Blagoderov V, Caravas J, Kutty SN, Schmidt-Ott U, Kampmeiern GE, Thompsono FC, Grimaldip DA, Beckenbachq AT, Courtneyr GW, Friedrichk M, Meier R, Yeates DK (2011) Episodic radiations in the fly tree of life. Proceedings of the National Academy of Sciences of the United States of America 108: 5690–5695. https://doi.org/10.1073/pnas.1012675108

Xia X (2018) DAMBE7: new and improved tools for data analysis in molecular biology and evolution. Molecular Biology and Evolution 35: 1550–1552. https://doi.org/10.1093/molbev/msy073

Xue WQ, Chao CM (Eds) (1998) Flies of China. Volume 1. Liaoning Science and Technology Press, Shenyang: 1–1365.

Yan LP, PapeT, Elgar M, Gao YY, Zhang D (2019) Evolutionary history of stomach bot flies in the light of mitogenomics. Systematic Entomology 44: 797–809. https://doi.org/10.1111/syen.12356

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

Yan LP, Xu WT, Zhang D, Li JQ (2021b) Comparative analysis of the mitochondrial genomes of flesh flies and their evolutionary implication. International Journal of Biological Macromolecules 174: 385–391. https://doi.org/10.1016/j.ijbiomac.2021.01.188

Zhao Z, Su T, Chesters D, Wang S, Ho S, Zhu C, Chen X, Zhang CT (2013) The mitochondrial genome of *Elodia flavipalpis* Aldrich (Diptera: Tachinidae) and the evolutionary timescale of tachinid flies. Plos One 8(4): e61814. https://doi.org/10.1371/journal.pone.0061814

Zhang NX, Yu G, Li TJ, He QY, Zhou Y, Si FL, Ren S, Chen B (2015) The Complete mitochondrial genome of *Delia antiqua* and its implications in Dipteran phylogenetics. Plos One 10: e0139736. https://doi.org/10.1371/journal.pone.0139736

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li W, Wang G (2020) PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular Ecology Resources 20: 348–355. https://doi.org/10.1111/1755-0998.13096