Bulletin of Entomological Research

cambridge.org/ber

The complete mitochondrial genome of *Brachytarsina amboinensis* (Diptera: Hippoboscoidea: Streblidae) provides new insights into phylogenetic relationships of Hippoboscoidea

Jinting Yang[1], Wang Yujuan[2], Huijuan Yang[3] and Xiaobin Huang[1]

[1]Yunnan Provincial Key Laboratory for Zoonosis Control and Prevention, Institute of Pathogens and Vectors, Dali University, Dali, Yunnan, China; [2]Jilin Provincial Key Laboratory of Animal Resource Conservation and Utilization, Northeast Normal University, Changchun, China and [3]Department of Pathogen Biology, Institute of Tropical Medicine, School of Public Health, Southern Medical University, Guangzhou, China

Research Paper

Cite this article: Yang J, Wang Y, Yang H, Huang X (2024). The complete mitochondrial genome of *Brachytarsina amboinensis* (Diptera: Hippoboscoidea: Streblidae) provides new insights into phylogenetic relationships of Hippoboscoidea. *Bulletin of Entomological Research* 1–11. https://doi.org/10.1017/S0007485324000762

Received: 10 July 2023
Revised: 15 November 2023
Accepted: 30 October 2024

Keywords:
bat flies; *Brachytarsina amboinensis*; comparative analysis; Hippoboscoidea; mitogenome; Streblidae

Corresponding author:
Xiaobin Huang;
Email: huangxb633@nenu.edu.cn

Abstract

The family Streblidae is a significant grouping of dipteran insects within the superfamily Hippoboscoidea, which parasitizes the body surface of bats. With the global spread of bat-related pathogens in recent years, Streblidae has gained increasing attention due to its potential for pathogen transmission. A sample of *Brachytarsina amboinensis* was sequenced on the *B. amboinensis* were obtained, compared with available Streblidae mitogenomes, and the phylogeny of Hippoboscoidea was reconstructed. The results indicate that the mitochondrial genome of *B. amboinensis* exhibits a relatively high degree of conservation, with an identical gene count, arrangement, and orientation as the ancestral insect's genome. Base composition analysis revealed a strong bias towards A and T in the base composition. Selection pressure analysis indicated strong purifying selection acting on *cox1*. Pairwise genetic distance analysis showed that *cox1* evolved at a relatively slow rate. Regarding phylogenetic relationships, the constructed phylogenetic trees using Bayesian inference and Maximum Likelihood methods supported the monophyly of the Hippoboscoidea, Glossinidae, Hippoboscidae, and Nycteribiidae clades, with high nodal support values. Our research confirmed the paraphyly of the families Streblidae. In the familial relations between Nycteribiidae and Streblidae, New World Streblidae share a closer kinship with Nycteribiidae. This contrasts with prior findings which indicated that Old World Streblidae share a closer kinship with Nycteribiidae. This study not only enhances the molecular database for bat flies but also provides a valuable reference for the identification and phylogenetic analysis of Streblidae.

Introduction

The Streblidae is a group of dipteran insects belonging to the superfamily Hippoboscoidea, comprising five subfamilies, 33 genera, and 229 species as of current knowledge (Dick *et al.*, 2016). These dipteran flies are important specialized ectoparasites of bats and feed on bat blood. They are commonly referred to as bat flies, along with Nycteribiidae (da Silva *et al.*, 2023). Bat flies exhibit a range of unique and significant morphological and physiological adaptations as a result of their long-term parasitic lifestyle. One of the most notable is their adenotrophic viviparity reproductive mode, wherein larvae develop in the female oviduct and are nourished by secretions from accessory glands until maturation into third instar larva (Yang *et al.*, 2023*b*). Nycteribiidae have wholly vestigial wings that are flattened dorsally and ventrally. They are also inserted dorsally on legs, with the head folded backwards and placed on the thorax, giving them a resemblance to spiders. In contrast, Streblidae have distinctively shaped wings and extensive membranous abdomens (Han *et al.*, 2022). Recent global outbreaks of bat-associated pathogens, such as Ebola viruses, severe acute respiratory syndrome coronavirus (SARS-CoV), SARS-CoV-2, and Middle East respiratory coronavirus (MERS-CoV), have brought bats into the spotlight (Letko *et al.*, 2020). A substantial body of research has detected human and animal pathogens in bat flies, including the méjal virus, Amate virus (Ramîrez-Martínez *et al.*, 2021), Hemoplasmas (Wang et al., 2023), *Bartonella* (Yang *et al.*, 2024) and others. Additionally, Kuang *et al.* (2023) successfully isolated Nelson Bay reovirus (NBV) from bat flies collected in Yunnan Province, with the S1, S2, and M1 segments showing high similarity to human pathogens. Streblidae have been known to occasionally bite humans (Szentiványi *et al.*, 2023), and their flight ability increases the possibility of pathogen spillover, which has generated widespread interest among researchers. However, conducting research in this area remains challenging due

to the small size of Streblidae, the absence of distinct morphological characteristics for identification, as well as morphological identification reference material.

Mitochondrial genomes are circular double-stranded DNA molecules that typically consist of 13 protein-coding genes (PCGs), 22 transfer RNAs (tRNAs), two ribosomal RNAs (rRNAs), and a control region (CR) (Clary and Wolstenholme, 1985). They are the only extranuclear genetic information carrier in animals and are characterized by maternal inheritance, conserved gene content, small molecular weight (approximately 14–21 kb), high mutation rates, and rapid evolution. Therefore, mitochondrial DNA (mtDNA) is often employed in comparative evolutionary genomics and phylogenetic analysis research (Crozier and Crozier, 1993). However, the GenBank database only currently contains three complete mitochondrial genomes of Streblidae. This is significantly disproportionate to the diversity of Streblidae.

In this study, we report on the first complete mitochondrial genome of *Brachytarsina amboinensis*. We provide a detailed analysis of the features of the new mitochondrial genome, including nucleotide composition and each tRNA's secondary structure. We also compare the differences in nucleotide composition, codon usage, relative synonymous codon usage (RSCU), and AT-skew and GC-skew among different Streblidae mitochondrial genomes. Finally, we investigate the phylogenetic relationships between 17 species in the Hippoboscoidea, 2 species in the Muscoidea, and 2 species in the Oestroidea, based on PCGs. The findings not only enrich the molecular database for bat flies but also provide a reference for the identification and phylogenetic analysis of Streblidae.

Materials and methods

Sample collection and processing

In July 2022, a total of 22 Streblidae specimens were collected from the surface of the eastern bent-wing bat (*Miniopterus fuliginosus)* in Binchuan County (100.58′E, 25.83′N), Dali Bai Autonomous Prefecture, Yunnan Province, China, and were subsequently preserved in 95% anhydrous ethanol in the laboratory. The samples were identified as *B. amboinensis* using a SZ2-ILST dissection microscope (Olympus, Tokyo, Japan) and following the identification characteristics described in Wenzel (1976). One of the samples was sent to Novogene Co., Ltd. (Beijing, China) for DNA extraction and sequencing, while the remaining specimens were stored in a −20°C freezer at the Institute of Pathogens and Vectors, Dali University. Upon completion of DNA extraction, a portion of the DNA is reserved for the purpose of validating CR. An Illumina paired-end (PE) library was prepared and DNA sequencing was performed on the Illumina Novoseq 6000 sequencing platform, generating a total data volume of 3 G (150 bp reads).

Sequence assembly, annotations and analysis

The raw sequence data of *B. amboinensis* were assembled using NOVOPLASTY 4.2.1 (Dierckxsens *et al.*, 2017), and the CR was verified by Sanger sequencing. The specific primers were designed to amplify the CR sequence based on the *16s rRNA* and *nad2* genes in the genome. The forward primer sequence is AACAGGCGAACATTATATTTGCCG, and the reverse primer sequence is AATCAGTAATGAAACGGAGCAG. The PCR amplification system outlined in the instructions of TaKaRa LA Taq (TaKaRa, Dalian, China) was employed, with the PCR reaction conditions involving an initial denaturation at 94°C for 4 min, followed by 37 cycles of denaturation at 94°C for 30 s, annealing at 50°C for 30 s, extension at 72°C for 1 min, and a final extension at 72°C for 10 min. The resulting amplified products were analyzed by electrophoresis on a 1% agarose gel, and then sent to Sangon Biotech (Shanghai) Co., Ltd. for bidirectional sequencing. The Sanger trace files were viewed using Geneious Primer 2.2.0, followed by removal of the poorly sequenced parts and manual editing of the sequence. The edited sequence was then aligned with the assembled complete mitochondrial genome. The assembled CR was found to be a perfect match with the CR obtained by PCR. Annotation of the PCGs, tRNAs, and rRNAs of the assembled circular sequence was performed using the MITOS Web Server (Donath *et al.*, 2019), and the annotated tRNAs were validated using the tRNA-scan Web Server (Chan and Lowe, 2019). The 13 PCGs were compared with known sequences in the database and manually corrected. The starting positions of the genes cox1, *nad1*, *nad5*, and *nad6* were modified, while the ending positions of *cox2*, *nad4*, and *nad5* were altered. The secondary structure of tRNAs was predicted using MITOS, and Adobe Illustrator CC 2019 was used for image processing. The circular mitochondrial genome was visualized using the online tool GENOMEVX (Conant and Wolfe, 2008). The mitochondrial genome data has been deposited in GenBank with the accession number OQ247894.

Comparative analysis

MEGA 11 (Tamura *et al.*, 2021) was utilized for calculating nucleotide composition statistics, pairwise genetic distances, and Relative Synonymous Codon Usage (RSCU) for four Streblidae sequences (GenBank accession numbers: MK896865, MK896866, OQ301747, and OQ247894). RSCU was calculated by excluding incomplete codons of PCGs. Moreover, AT-skew and GC-skew were computed using the following formulae: AT-skew = (A − T)/(A + T) and GC-skew = (G − C)/(G + C). The Ka/Ks ratio of nonsynonymous (Ka) and synonymous (Ks) substitutions between *B. amboinensis* and the other three Streblidae (*Paradyschiria parvula*, *Paratrichobius longicrus*, and *Raymondia* sp.) was analyzed using the mlwl-based Ka/Ks calculator 2.0 (Wang *et al.*, 2010). Additionally, the nucleotide diversity (Pi) of PCGs in the four Streblidae was examined using DnaSP 6.0 (Rozas *et al.*, 2017) with a sliding window of 100 bp and a step size of 20 bp.

Phylogenetic analyses

We selected a total of 17 species, representing all four families of the superfamily Hippoboscoidea, two species from the superfamily Muscoidea, and two species from the superfamily Oestroidea, with *Chironomus tepperi* and *Dixella aestivalis* as outgroups (table 1). The 13 PCGs of the mitochondrial genome were extracted for building the phylogenetic trees using the PhyloSuite platform (Zhang *et al.*, 2020). The PCGs were aligned using MAFFT 7 (Rozewicki *et al.*, 2019), and the fuzzy alignments were removed using Gblocks 0.91b (Castresana, 2000). The aligned sequences were manually concatenated using PhyloSuite. Maximum Likelihood (ML) and Bayesian inference (BI) analyses were performed using IQ-TREE 2.2.0 (Minh *et al.*, 2020) and MrBayes 3.2.7a (Ronquist *et al.*, 2012), respectively. The best partitioning schemes and corresponding nucleotide substitution models were inferred using ModelFinder (Kalyaanamoorthy *et al.*, 2017) with the corrected Akaike information criterion (AICC) and greedy search algorithm. The best-fit partitioning scheme and corresponding models used on the BI phylogenetic tree are shown in table S1.

https://doi.org/10.1017/S0007485324000762 Published online by Cambridge University Press

Table 1. Mitochondrial genome information used in this study

Superfamily	Family	Species	Length (bp)	GenBank accession No.	References
Hippoboscoidea	Glossinidae	*Glossina austeni*	17,449	MZ826152	Porter *et al.* (2022)
		Glossina brevipalpis	17,751	MZ826153	Porter *et al.* (2022)
	Hippoboscidae	*Melophagus ovinus*	15,573	KX870852	Liu *et al.* (2017)
		Lipoptena sp.	16,953	MT679542	Wang *et al.* (2021)
		Ornithomya biloba	18,654	MZ379837	Li *et al.* (2022)
	Streblidae	*Paradyschiria parvula*	14,588	MK896865	Trevisan *et al.* (2019)
		Paratrichobius longicrus	16,296	MK896866	Trevisan *et al.* (2019)
		Raymondia sp.	16,514	OQ301747	Poon *et al.* (2023)
		Brachytarsina amboinensis	15,693	OQ247894	This study
	Nycteribiidae	*Basilia ansifera*	16,964	MZ826150	Porter *et al.* (2022)
		Dipseliopoda setosa	19,164	MZ826151	Porter *et al.* (2022)
		Nycteribia parvula	16,060	OP442519	Yang *et al.* (2023*d*)
		Nycteribia allotopa	15,161	OQ270755	Yang *et al.* (2023*c*)
		Nycteribia formosana	15,107	OQ675011	unpublished
		Phthiridium szechuanum	14,896	OP459298	Zhang *et al.* (2023)
		Penicillidia dufourii	15,354	OQ658726	unpublished
		Phthiridium sp.	16,155	OQ301748	Poon *et al.* (2023)
Oestroidea	Tachinidae	*Biomeigenia flava*	16,333	OP650038	unpublished
		Blepharipa latigena	15,826	OP650039	unpublished
Muscoidea	Anthomyiinae	*Hylemya nigrimana*	16,746	ON099429	unpublished
	Muscidae	*Coenosia pumila*	15,214	MT872669	unpublished
Outgroups	Dixidae	*Dixella aestivalis*	16,465	KT878382	Briscoe *et al.* (2017)
	Chironomidae	*Chironomus tepperi*	15,652	JN861749	Beard *et al.* (1993)

The ML analysis was evaluated using the ultrafast bootstrap approximation approach for 1000 replicates. BI analyses were performed with two Markov chain Monte Carlo (MCMC) chains runs of 10,000,000 generations, with sampling every 1000 generations. Convergence was considered to be reached when the average standard deviation of the split frequencies was lower than 0.01. The first 25% of generations from each MCMC chain run to account for potential lack of convergence were discarded as burn-in, and the remaining samples were used to generate the majority consensus trees and to estimate the posterior probabilities. The generated phylogenetic trees were viewed and edited using FigTree 1.44 (http://tree.bio.ed.ac.uk/software/figtree/).

Results

Genomic organization and base compositions

The mitochondrial genome of *B. amboinensis* is circular and double-stranded, with a length of 15,693 bp. It contains 13 PCGs, 22 tRNAs, two rRNAs, and one CR (fig. 1 and table 2), exhibiting a relatively high degree of conservation, with an identical gene count, arrangement, and orientation as the ancestral insect's genome (Boore, 1999). The nucleotide composition of *B. amboinensis* mitochondrial DNA is 42.46% A, 38.67% T, 7.25% G, and 11.62% C, with an overall AT content of 81.13%.

The protein coding genes

The 13 PCGs of *B. amboinensis* range in length from 156 to 1696 bp, with *nad5* being the longest and *atp8* being the shortest. The total length of the 13 PCGs is 11,083 bp, and the AT content is 80.46%. All PCGs start with ATN (5 ATG, 6 ATT) start codons, except for *cox1* (starting with TCG) and *nad1* (starting with TTG), which use non-standard start codons. All PCGs, except *cox1*, *nad4*, and *nad5*, end with TAN stop codons (7 TAA, 3 TAG), while *cox1*, *nad4*, and *nad5* have single T stop codons. The use of TCG as a start codon for *cox1* is common in Diptera (Ma *et al.*, 2022), as non-standard start codons for *cox1* are frequently observed in holometabolous insects (Li *et al.*, 2020). Incomplete stop codons are fairly common in the Diptera mitogenomes and can be converted into a potential stop codon via polyadenylation to TAA (Nelson *et al.*, 2012).

The tRNAs, rRNAs and control region

The complete mitochondrial genome of *B. amboinensis* contains 22 tRNAs, ranging in length from 62–72 bp (table 2). Nineteen of them can be folded into typical cloverleaf structures (fig. 2), while *trnS1* lacks the dihydrouridine arm (DHU), and *trnG* and *trnT* lack the thymidine–pseudouridine–cytidine (TΨC) loop.

https://doi.org/10.1017/S0007485324000762 Published online by Cambridge University Press

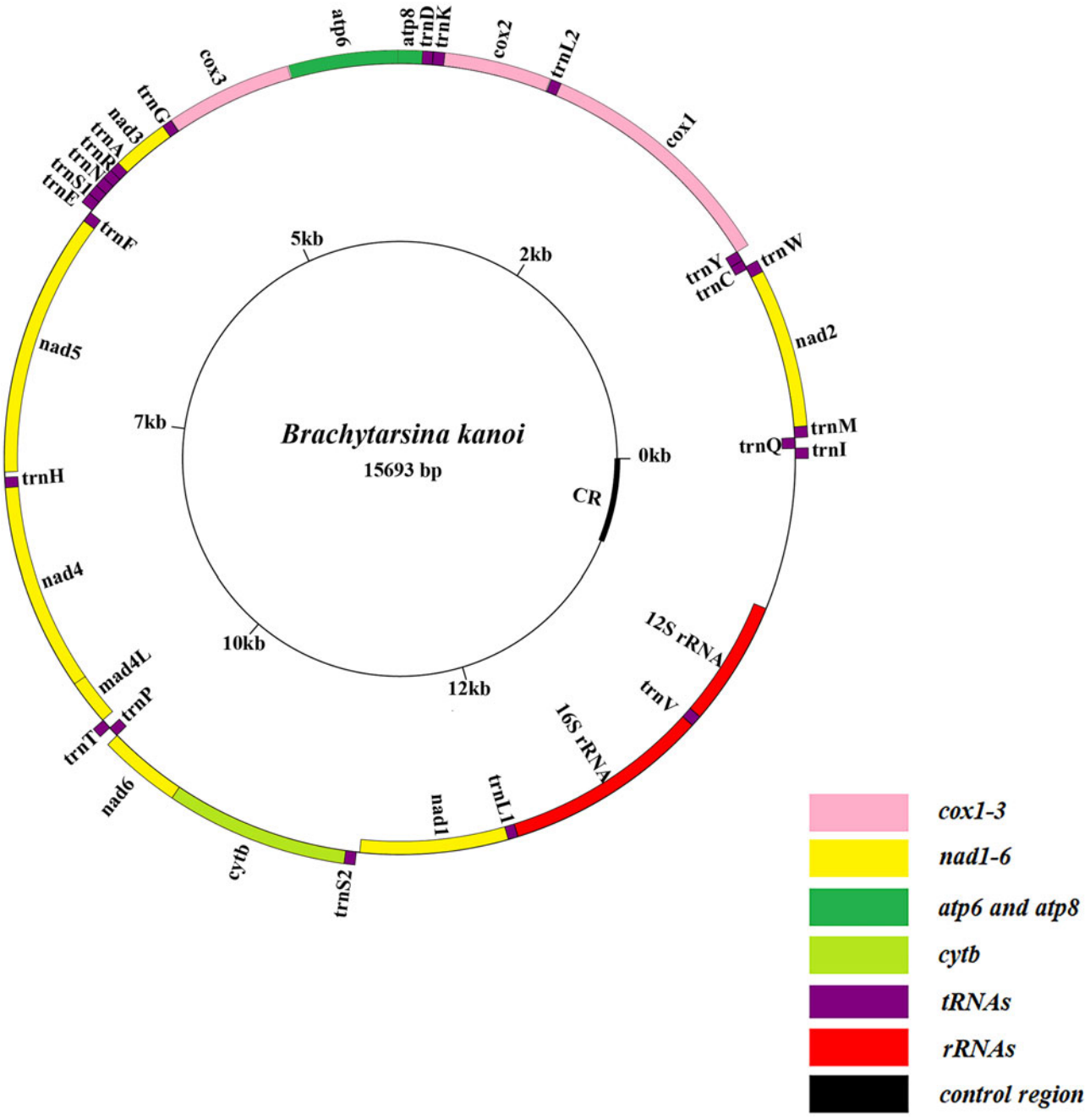


Figure 1. The circular mitochondrial genome map of *Brachytarsina amboinensis.*

There are 17 non-matching base pairs in the tRNAs, mostly consisting of GU mismatches (15 in total). Additionally, there are 2 UU mismatches in the acceptor stem of *trnA* and *trnL2* (fig. 2).

The combined length of the two rRNAs is 2150 bp, with an overall A + T content of 82.22%. The *16S rRNA* is 1367 bp in length and located between *trnL* and *trnV*, while the *12S rRNA* is 783 bp in length and located between *trnV* and the CR.

The length of the CR is 968 bp, and is located between the *12S rRNA* and the *trnI*. It possesses a very high AT content of 85.43%. This region is significant in the mitochondrial genome, as it plays a critical role in DNA replication and the initiation of transcription (Zhang and Hewitt, 1997).

Comparative analysis

It is commonly recognized that the analysis of base composition and strand asymmetry, including AT- and GC-skew, provides insights into base composition bias in mitochondria (fig. 3). Based on the analysis of the four Streblidae mitochondrial genomes, an AT content ranging from 78.63 to 81.13% was observed, indicative of a preference for A and T nucleotides, which is typical of insect mitochondrial genomes (Clary and Wolstenholme, 1985). Furthermore, the AT content in CR of the four Streblidae is notably higher, ranging from 83.09 to 92.3%, compared to the entire genome. The observed strand asymmetry pattern in the Streblidae mitochondrial genomes

https://doi.org/10.1017/S0007485324000762 Published online by Cambridge University Press

Table 2. Mitochondrial genome organization of *Brachytarsina amboinensis*

		Position			Codon	
Gene	Coding strand	Start	Stop	Length (bp)	Start	Stop
trnI	J	1	66	66		
trnQ	N	64	132	69		
trnM	J	132	200	69		
nad2	J	201	1190	990	ATT	TAA
trnW	J	1189	1256	68		
trnC	N	1249	1310	62		
trnY	N	1311	1376	66		
cox1	J	1375	2908	1534	TCG	T
trnL2	J	2909	2973	65		
cox2	J	2980	3646	667	ATT	TAA
trnK	J	3647	3717	71		
trnD	J	3719	3784	66		
atp8	J	3785	3940	156	ATT	TAA
atp6	J	3934	4611	678	ATG	TAA
cox3	J	4620	5405	786	ATG	TAA
trnG	J	5409	5470	62		
nad3	J	5471	5824	354	ATT	TAG
trnA	J	5823	5888	66		
trnR	J	5888	5950	63		
trnN	J	5950	6015	66		
trnS1	J	6016	6082	67		
trnE	J	6083	6150	68		
trnF	N	6168	6235	68		
nad5	N	6236	7964	1696	ATT	T
trnH	N	7965	8030	66		
nad4	N	8031	9363	1333	ATG	T
nad4L	N	9357	9647	291	ATG	TAA
trnT	J	9650	9713	64		
trnP	N	9714	9778	65		
nad6	J	9780	10,292	513	ATT	TAA
cytb	J	10,292	11,428	1137	ATG	TAG
trnS2	J	11,427	11,494	68		
nad1	N	11,516	12,463	948	TTG	TAG
trnL1	N	12,464	12,527	64		
16S rRNA	N	12,528	13,848	1321		
trnV	N	13,849	13,920	72		
12S rRNA	N	13,921	14,725	805		

aligns with the characteristic strand bias of dipteran insects, with an AT-skew range of 0.02 to 0.08 and a GC-skew range of −0.28 to −0.23 (Ren *et al.*, 2019). These skewed strand compositions are likely attributed to mutational and selective pressures (Kono *et al.*, 2018), with the GC-skew value associated with the direction of replication within the genome (Chen *et al.*, 2016).

For evaluating synonymous codon bias, the Relative Synonymous Codon Usage (RSCU) serves as a key parameter, and the obtained values directly reflect the codon usage preference

https://doi.org/10.1017/S0007485324000762 Published online by Cambridge University Press

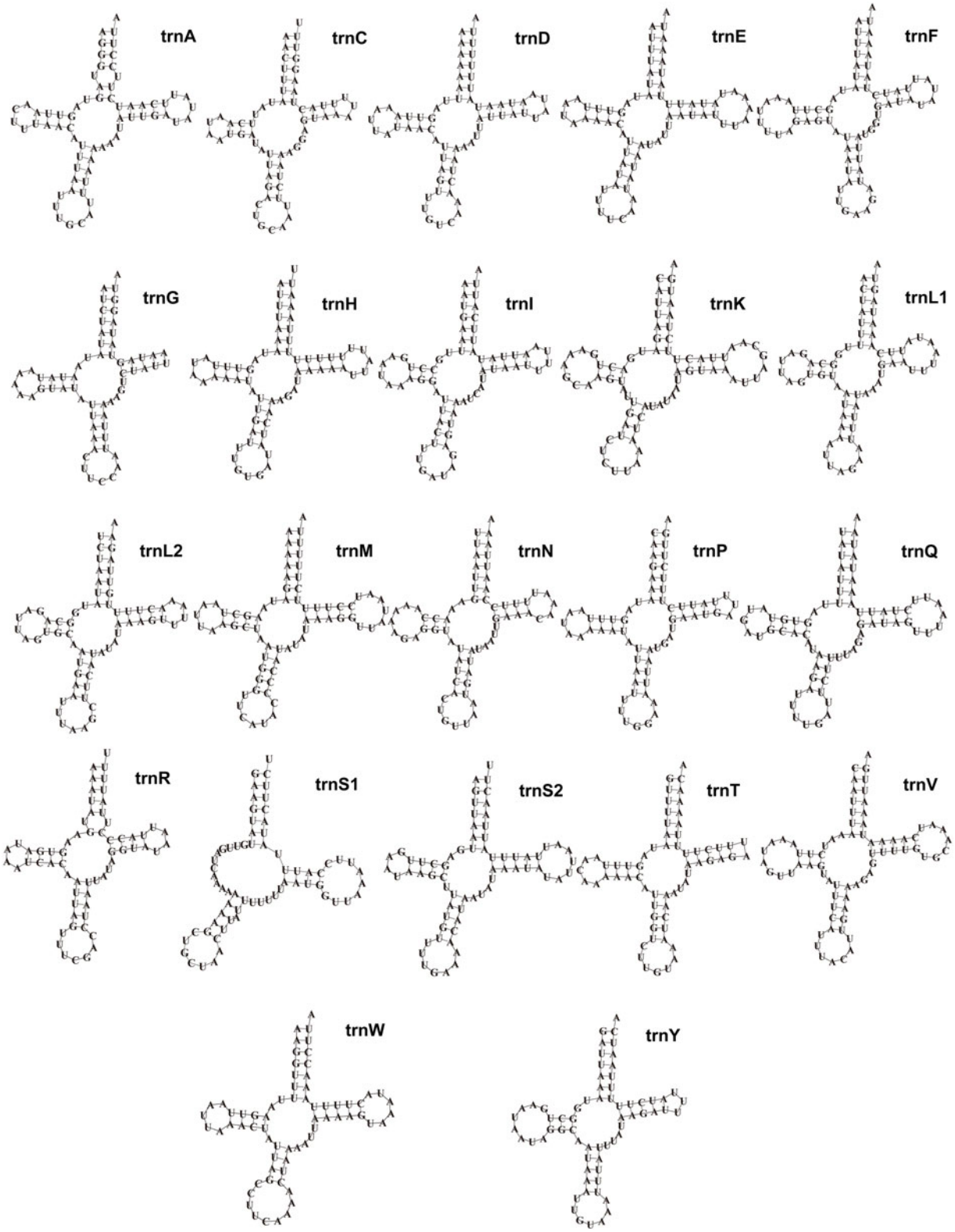


Figure 2. Secondary structure of the tRNA genes from the mitochondrial genome of *Brachytarsina amboinensis*.

(fig. 4). In *B. amboinensis*' complete mitochondrial genome, there are a total of 3684 codons encoding proteins, excluding the stop codons. The most commonly used codons are UUA, AUU, UUU, AUA, and AAU, whereas the least commonly used codons are CUC, CGC, AGG, and UGG. Among all 62 codons, 26 of them are considered to have a usage preference, as these synonymous codons have a positive codon bias (RSCU value > 1). The codon usage preference among Streblidae is similar, with the top five most commonly used codons across all four species being UUA, AUU, UUU, AUA, and AAU (fig. 4). Additionally, the

https://doi.org/10.1017/S0007485324000762 Published online by Cambridge University Press

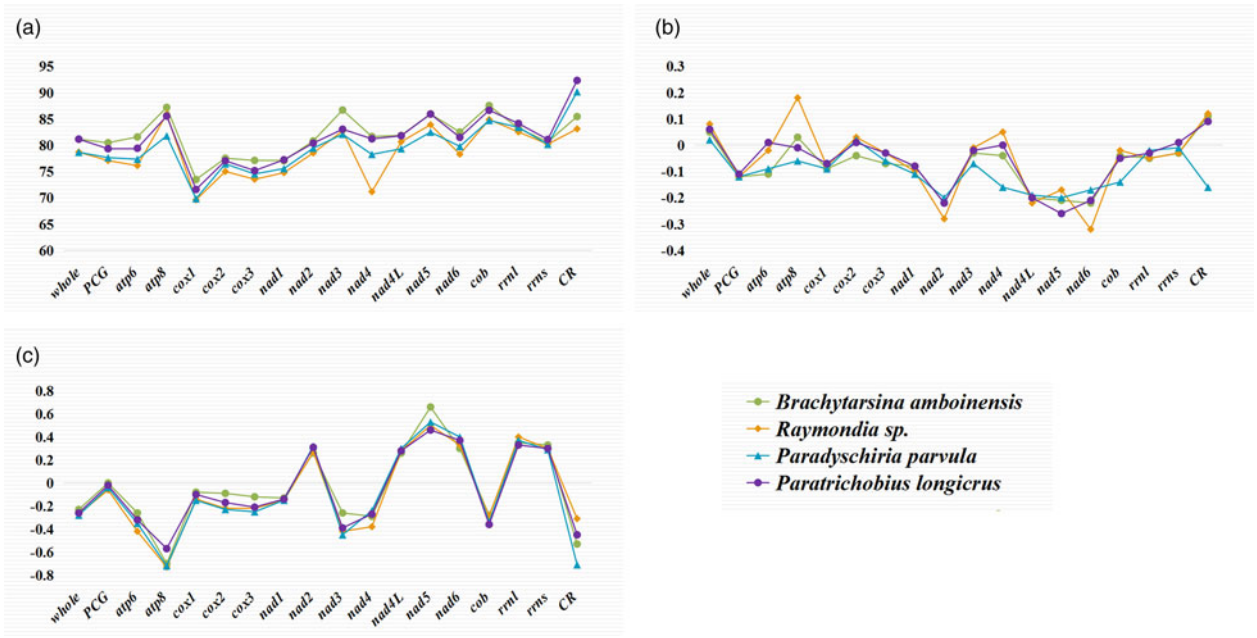


Figure 3. AT content (A), AT-skew (B) and GC-skew (C) of the 13 PCGs of four Streblidae.

RSCU values of the third position nucleotides being A or T are significantly higher than those of G or C, indicating that the high A + T content in the mitochondrial genome leads to codon usage bias.

The evolutionary patterns for each PCG were assessed using pairwise Ka/Ks and genetic distances (fig. 5). The highest Ka/Ks values were observed for *nad5* (0.59) and *nad6* (0.59), suggesting that they may have undergone faster evolutionary rates compared to other PCGs. This implies that *nad5* and *nad6* may have experienced more relaxed selection constraints and accumulated a greater number of mutations over time. In contrast, *cox1* (0.12) displayed the smallest Ka/Ks value, indicating a slower evolutionary rate. This suggests that *cox1* faces greater evolutionary pressure and evolves at a slower rate compared to the other genes. The Ka/Ks values of all 13 PCGs were less than 1, indicating that purifying selection may have dominated the evolution of the mitochondrial genome (Hurst, 2002; Chang *et al.*, 2020). Genetic distances showed similar results: *cox1* (0.21) was evolving comparatively slowly, while *nad2* (0.43), *nad3* (0.42), and *nad6* (0.43) were evolving comparatively quickly. The Pi values of the 13 PCGs in the four Streblidae ranged from 0.178 to 0.328 (fig. 6). Among these genes, *nad2* showed the highest variability (0.328), followed by *nad6* (0.313), while *cox1* (0.178) showed the lowest variability.

Phylogenetic analysis

Using BI and ML methods, the phylogenetic relationships of the Hippoboscoidea were analyzed based on 13 PCGs (fig. 7). The constructed BI and ML trees both supported the monophyly of Hippoboscoidea, Glossinidae, Hippoboscidae, and Nycteribiidae, and had high node support values. Streblidae were found to be paraphyletic, with differing relationships to the Nycteribiidae in the two trees. In the BI tree, New World Streblidae were grouped with Nycteribiidae with a high posterior probability (BPP = 1), providing strong support for this relationship. In the ML tree, Old World Streblidae were grouped with Nycteribiidae with moderate bootstrap support (BS = 70), indicating that this relationship is less certain. The Hippoboscidae was found to be the sister group to the bat flies (Streblidae and Nycteribiidae), while the Glossinidae was sister to a clade comprising all three of these families.

Discussion

Designing universal markers is indeed crucial for resolving species identification issues. Universal markers, such as specific DNA sequences or genetic loci that are conserved across related species, are valuable tools for distinguishing between different organisms. The *cox1* gene has historically been widely used as a universal DNA marker for insect species identification, but its high conservation within the insect mitogenome can limit its efficiency (Ožana *et al.*, 2022; Yang *et al.*, 2023*a*). Due to the conservativeness of the cox1 gene, it may not be possible to reach the critical threshold for species identification of closely related species. Before determining the species identity, it may only be possible to assign it to a higher taxonomic group, such as a phylum or order. However, the nad2 and nad6 genes display significant nucleotide variations and a high Ka/Ks ratio, which can be utilized for distinguishing species with close phylogenetic relationships, thus establishing them as potential markers for taxonomic categorization within the family Streblidae. Furthermore, combining *cox1* with other genes has been demonstrated to enhance the accuracy and reliability of species identification in some studies (Alberfkani *et al.*, 2022). However, it is important to emphasize that further experimental validation and cross-validation are necessary to ensure the effectiveness and stability of newly designed markers. This validation process is critical for establishing the reliability and utility of these markers for species identification and delimitation.

The formation of the concept of the superfamily Hippoboscoidea was a lengthy process. Initially, it only included three core families, which are the 'Pupipara' (Nycteribiidae, Streblidae, and Hippoboscidae). The current understanding of Hippoboscoidea includes the Glossinidae and the Pupipara, and this was proposed by Hennig (1971) in 1971, and authors such as McAlpine (1989) and Ziegler (2003) have since defended this view. But Nycteribiidae and Streblidae have even been classified

https://doi.org/10.1017/S0007485324000762 Published online by Cambridge University Press

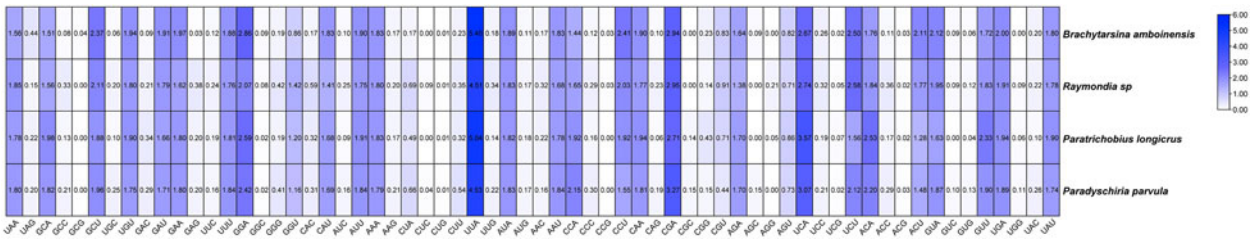
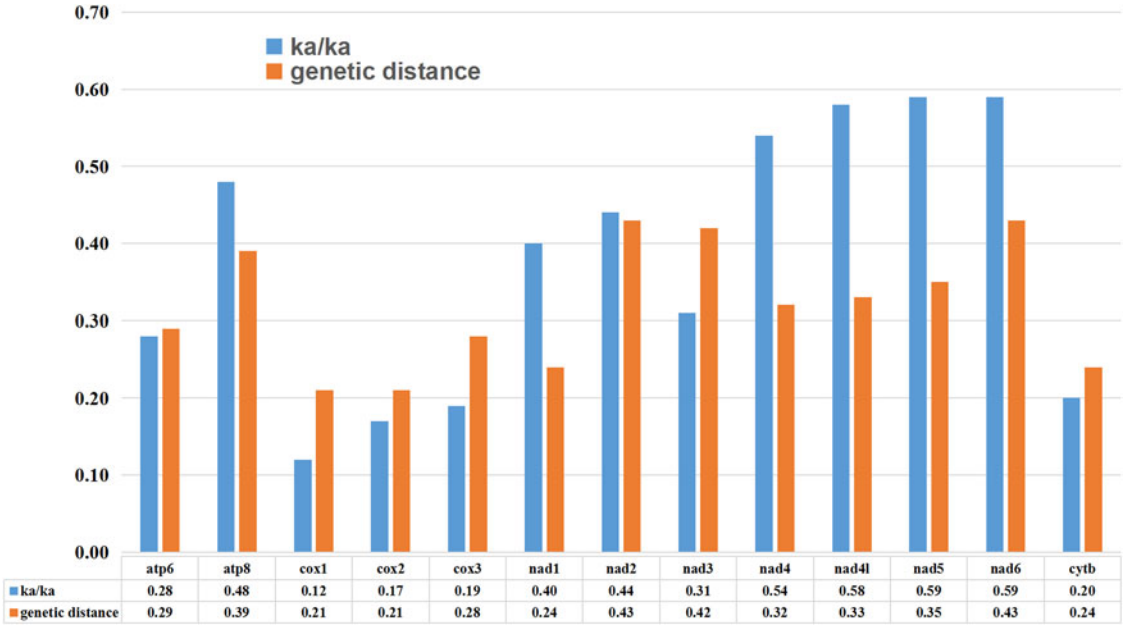


Figure 4. Heat map of RSCU value of each codon among four species of Streblidae.



	atp6	atp8	cox1	cox2	cox3	nad1	nad2	nad3	nad4	nad4l	nad5	nad6	cytb
ka/ka	0.28	0.48	0.12	0.17	0.19	0.40	0.44	0.31	0.54	0.58	0.59	0.59	0.20
genetic distance	0.29	0.39	0.21	0.21	0.28	0.24	0.43	0.42	0.32	0.33	0.35	0.43	0.24

Figure 5. Substitution ratios in the individual PCG of mitochondrial genomes of four Streblidae (the mean value of Ka/Ks and genetic distance under figure).

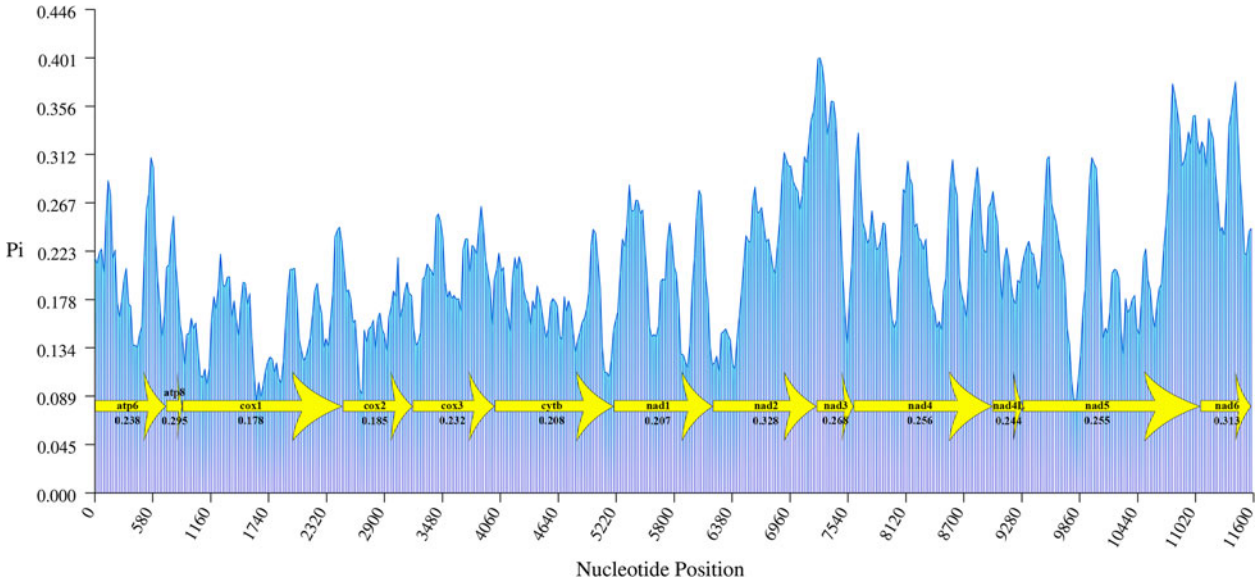


Figure 6. The nucleotide diversity (Pi) of 13 PCGs of five Streblidae.

https://doi.org/10.1017/S0007485324000762 Published online by Cambridge University Press

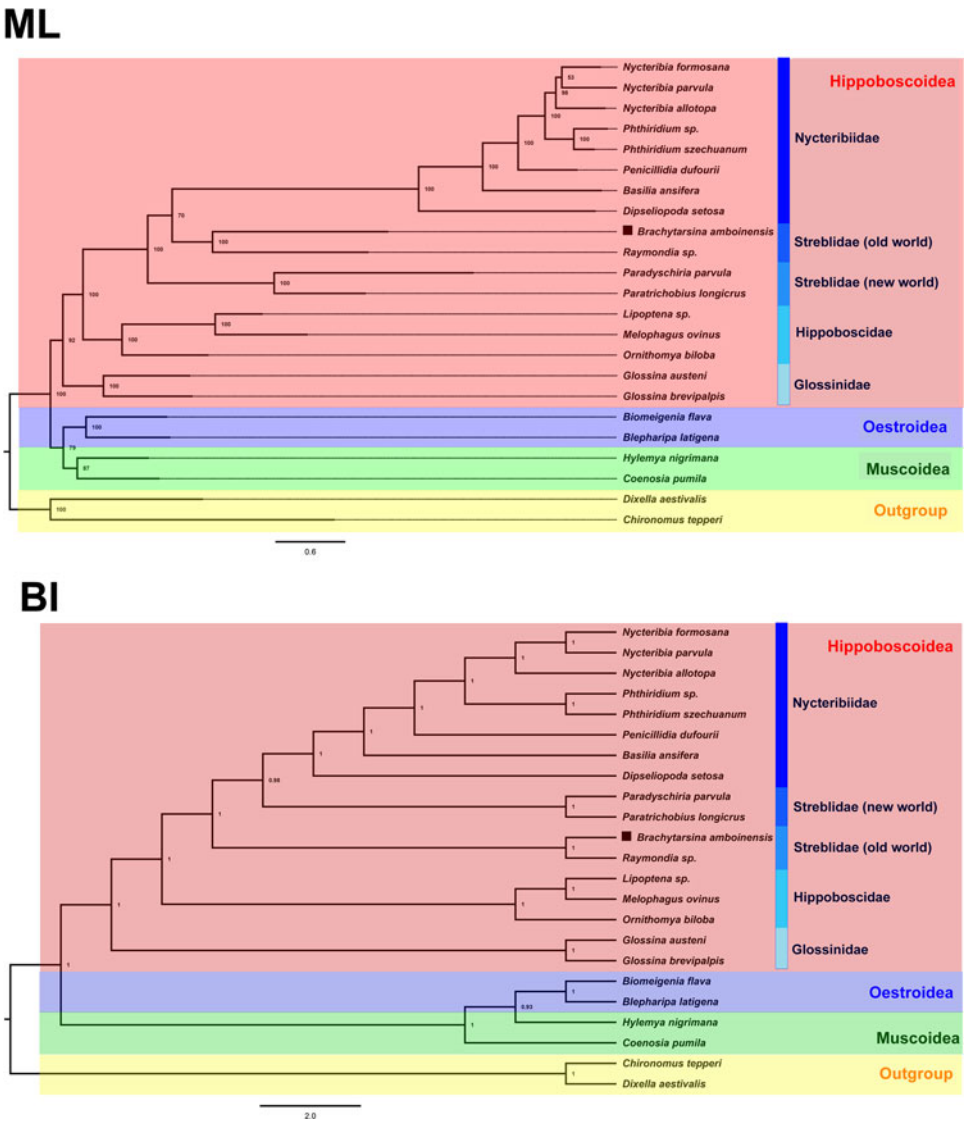


Figure 7. Phylogenetic analysis based on nucleotide sequences of 13 PCGs from the mitogenome. Numbers at nodes indicate bootstrap (ML) and posterior probability (BI) support values, respectively. '■' indicates the specie in this study.

under the family Hippoboscidae at times (McAlpine, 1989). At the same time, different opinions and conclusions existed regarding the evolutionary relationships within the Hippoboscoidea. Hennig (1971) suggested that Glossinidae forms the sister group of the other three families within Hippoboscoidea, and that the families Nycteribiidae and Streblidae constitute a monophyletic group. Conversely, McAlpine (1989) supported the concept that Glossinidae and Hippoboscidae together form the sister group of Nycteribiidae and Streblidae. Molecular data and ML analyses by Nirmala *et al.* (2001) and Dittmar *et al.* (2006) supported the latter view, but the Streblidae were found to be a paraphyletic group. Subsequent research by Petersen *et al.* (2007) conducted a phylogenetic analysis using combined sequence data from *CAD*, *cox1*, *16S*, and *28S*, and supported the view that Glossinidae, Hippoboscidae, and Nycteribiidae were monophyletic. However, upon analyzing Old and New World Streblidae together, the Streblidae displayed paraphyly, with those from the Old World clustering with the Nycteribiidae. Poon *et al.* (2023) constructed a phylogenetic tree based on the *cox1* gene and demonstrated that the Streblidae from the Old World were closely related to the Nycteribiidae and formed a clade which was sister to the New World Streblidae. Kutty *et al.* (2010) conducted a phylogenetic analysis based on PCGs, including *cox1*, *cytb*, *Ef1α*, and *CAD*, and showed that Glossinidae, along with three other families within the superfamily Hippoboscoidea, formed a sister group. Nycteribiidae and Streblidae formed a clade, with Nycteribiidae being a monophyletic group, and Streblidae being paraphyletic. Moreover, the previous studies on the Hippoboscoidea predominantly used partial PCGs, which reflects the relatively limited phylogenetic relationships (Cameron *et al.*, 2007).

The present study employed all 13 PCGs of the mitogenomes and encompassing all major lineages to investigate the systematic relationships within the Hippoboscoidea. Both the ML and BI trees indicated that Glossinidae, Hippoboscidae, and bat flies (Nycteribiidae and Streblidae) were monophyletic groups. The Glossinidae were observed to form a sister group with the other three families, while Hippoboscidae formed a sister group with the bat flies. Interestingly, the New World Streblidae and

https://doi.org/10.1017/S0007485324000762 Published online by Cambridge University Press

Nycteribiidae formed a clade, representing a divergence from previous studies and providing new insights into the evolutionary relationships of the Hippoboscoidea. However, the limited number of complete mitochondrial genomes for this group still poses a challenge in understanding their systematic evolution, highlighting the need for additional genomic data to facilitate a comprehensive understanding of their relationships.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0007485324000762.

Author contributions. J. T. Y., X. B. H., Y. J. W. and H. J. Y. designed the study, J. T. Y., X. B. H., Y. J. W. and H. J. Y. collated the data, J. T. Y. analysed the data and wrote the first draft of manuscript, X. B. H. completed the final draft of the manuscript and all authors contributed to the final version.

Financial support. This work was supported by the National Natural Science Foundation of China [No. 32001096 to Xiaobin Huang] and the Science and Technology Planning Project of the Yunnan Provincial Department of Science and Technology (No. 202001AT070025 to Xiaobin Huang).

Competing interests. None.

Ethical standards. No specific permits were required for the insects collected for this study. The capture of bats according to the standards and procedures set by the Animal Ethics Committee of Dali University. (Name: Dali University Ethics Committee; approval ID: MECDU-202104-27).

References

Alberfkani MI, Albarwary AJ, Jaafar GM, Zubair AI and Abdullah RY (2022) Molecular characterization and phylogenetic analysis of *cox*1 and ITS 1 gene fragments of Moniezia species isolated from sheep. *Pakistan Veterinary Journal* **42**, 566–570.

Beard C, Hamm DM and Collins F (1993) The mitochondrial genome of the mosquito Anopheles gambiae: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. *Insect Molecular Biology* **2**, 103–124.

Boore JL (1999) Animal mitochondrial genomes. *Nucleic Acids Research* **27**, 1767–1780.

Briscoe AG, Sivell D and Harbach RE (2017) The complete mitochondrial genome of *Dixella aestivalis* (Diptera: Nematocera: Dixidae). *Mitochondrial DNA Part A* **28**, 83–84.

Cameron SL, Lambkin CL, Barker SC and Whiting MF (2007) A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high precision. *Systematic Entomology* **32**, 40–59.

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**, 540–552.

Chan PP and Lowe TM (2019) tRNAscan-SE: searching for tRNA genes in genomic sequences. *Gene Prediction: Methods and Protocols* **1962**, 1–14.

Chang H, Qiu Z, Yuan H, Wang X, Li X, Sun H, Guo X, Lu Y, Feng X and Majid M (2020) Evolutionary rates of and selective constraints on the mitochondrial genomes of Orthoptera insects with different wing types. *Molecular Phylogenetics and Evolution* **145**, 106734.

Chen W-H, Lu G, Bork P, Hu S and Lercher MJ (2016) Energy efficiency trade-offs drive nucleotide usage in transcribed regions. *Nature Communications* **7**, 11334.

Clary DO and Wolstenholme DR (1985) The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *Journal of Molecular Evolution* **22**, 252–271.

Conant GC and Wolfe KH (2008) GenomeVx: simple web-based creation of editable circular chromosome maps. *Bioinformatics* **24**, 861–862.

Crozier RH and Crozier YC (1993) The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization. *Genetics* **133**, 97–117.

da Silva SG, Ferreira FF, Hrycyna G, Eriksson A, Graciolli G and Canale GR (2023) Determinants of the composition of ectoparasitic flies of bats (Diptera: Streblidae, Nycteribiidae) in the Amazon and Cerrado landscape scales and ecotonal areas. *Parasitology Research* **122**, 1851–1861.

Dick CW, Graciolli G and Guerrero R (2016) Family streblidae. *Zootaxa* **4122**, 784–802.

Dierckxsens N, Mardulyn P and Smits G (2017) NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research* **45**, e18.

Dittmar K, Porter ML, Murray S and Whiting MF (2006) Molecular phylogenetic analysis of nycteribiid and streblid bat flies (Diptera: Brachycera, Calyptratae): implications for host associations and phylogeographic origins. *Molecular Phylogenetics and Evolution* **38**, 155–170.

Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, Middendorf M and Bernt M (2019) Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. *Nucleic Acids Research* **47**, 10543–10552.

Han HJ, Li ZM, Li X, Liu JX, Peng QM, Wang R, Gu XL, Jiang Y, Zhou CM and Li D (2022) Bats and their ectoparasites (Nycteribiidae and Spinturnicidae) carry diverse novel Bartonella genotypes, China. *Transboundary and Emerging Diseases* **69**, e845–e858.

Hennig W (1971) *Neue untersuchungen ueber die familien der Diptera Schizophora (Diptera: Cyclorrhapha)*. Stuttgart: The Germany: Staatliches Museum für Naturkunde.

Hurst LD (2002) The Ka/Ks ratio: diagnosing the form of sequence evolution. *TRENDS in Genetics* **18**, 486–487.

Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A and Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**, 587–589.

Kono N, Tomita M and Arakawa K (2018) Accelerated laboratory evolution reveals the influence of replication on the GC skew in *Escherichia coli*. *Genome Biology and Evolution* **10**, 3110–3117.

Kuang G, Xu Z, Wang J, Gao Z, Yang W, Wu W, Liang G, Shi M and Feng Y (2023) Nelson bay reovirus isolated from bats and blood-sucking arthropods collected in Yunnan Province, China. *Microbiology Spectrum* **11**, e05122.

Kutty SN, Pape T, Wiegmann BM and 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**, 614–635.

Letko M, Seifert SN, Olival KJ, Plowright RK and Munster VJ (2020) Bat-borne virus diversity, spillover and emergence. *Nature Reviews Microbiology* **18**, 461–471.

Li XY, Yan LP, Pape T, Gao YY and 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.

Li X, Wang L and Yang D (2022) The complete mitochondrial genome of *Ornithomya biloba* (Diptera, Hippoboscidae). *Mitochondrial DNA Part B* **7**, 856–858.

Liu Z-Q, Kuermanali N, Li Z, Chen S-J, Wang Y-Z, Tao H and Chen C-F (2017) The complete mitochondrial genome of the parasitic sheep ked *Melophagus ovinus* (Diptera: Hippoboscidae). *Mitochondrial DNA Part B* **2**, 432–434.

Ma XX, Wang FF, Wu TT, Li Y, Sun XJ, Wang CR and Chang QC (2022) First description of the mitogenome and phylogeny: *Aedes vexans* and *Ochlerotatus caspius* of the Tribe Aedini (Diptera: Culicidae). *Infection, Genetics and Evolution* **102**, 105311.

McAlpine J (1989) Phylogeny and classification of the Muscomorpha. *Manual of Nearctic Diptera* **3**, 1397–1518.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A and Lanfear R (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**, 1530–1534.

Nelson L, Lambkin CL, Batterham P, Wallman JF, Dowton M, Whiting MF, Yeates DK and Cameron SL (2012) Beyond barcoding: a mitochondrial genomics approach to molecular phylogenetics and diagnostics of blowflies (Diptera: Calliphoridae). *Gene* **511**, 131–142.

Nirmala X, Hypša V and Žurovec M (2001) Molecular phylogeny of Calyptratae (Diptera: Brachycera): the evolution of 18S and 16S ribosomal rDNAs in higher dipterans and their use in phylogenetic inference. *Insect Molecular Biology* **10**, 475–485.

Ožana S, Dolný A and Pánek T (2022) Nuclear copies of mitochondrial DNA as a potential problem for phylogenetic and population genetic studies of Odonata. *Systematic Entomology* **47**, 591–602.

Petersen FT, Meier R, Kutty SN and Wiegmann BM (2007) The phylogeny and evolution of host choice in the Hippoboscoidea (Diptera) as reconstructed using four molecular markers. *Molecular Phylogenetics and Evolution* **45**, 111–122.

Poon ESK, Chen G, Tsang HY, Shek CT, Tsui WC, Zhao H, Guénard B and Sin SYW (2023) Species richness of bat flies and their associations with host bats in a subtropical east Asian region. *Parasites & Vectors* **16**, 1–15.

Porter ML, Lutz H, Steck M and Chong RA (2022) The complete mitochondrial genomes and phylogenetic analysis of two Nycteribiidae bat flies (Diptera: Hippoboscoidea). *Mitochondrial DNA Part B* **7**, 1486–1488.

Ramírez-Martínez MM, Bennett AJ, Dunn CD, Yuill TM and Goldberg TL (2021) Bat flies of the family streblidae (Diptera: Hippoboscoidea) host relatives of medically and agriculturally important 'bat-associated' viruses. *Viruses* **13**, 860.

Ren L, Shang Y, Yang L, Shen X, Chen W, Wang Y, Cai J and 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.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA and Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**, 539–542.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE and Sánchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34**, 3299–3302.

Rozewicki J, Li S, Amada KM, Standley DM and Katoh K (2019) MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Research* **47**, W5–W10.

Szentiványi T, Heintz AC, Markotter W, Wassef J, Christe P and Glaizot O (2023) Vector-borne protozoan and bacterial pathogen occurrence and diversity in ectoparasites of the Egyptian Rousette bat. *Medical and Veterinary Entomology* **37**, 189–194.

Tamura K, Stecher G and Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* **38**, 3022–3027.

Trevisan B, Alcantara DM, Machado DJ, Marques FP and Lahr DJ (2019) Genome skimming is a low-cost and robust strategy to assemble complete mitochondrial genomes from ethanol preserved specimens in biodiversity studies. *PeerJ* **7**, e7543.

Wang D, Zhang Y, Zhang Z, Zhu J and Yu J (2010) KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics, Proteomics & Bioinformatics* **8**, 77–80.

Wang M, Wang J, Guo Y, Zheng Q, Nouhoum D and Meng F (2021) Complete mitochondrial genome of a potential vector louse fly, *Lipoptena grahami* (Diptera, Hippoboscidae). *Mitochondrial DNA Part B* **6**, 1752–1753.

Wang R, Li ZM, Peng QM, Gu XL, Zhou CM, Xiao X, Han HJ and Yu XJ (2023) High prevalence and genetic diversity of hemoplasmas in bats and bat ectoparasites from China. *One Health* **16**, 100498.

Wenzel RL (1976) The streblid batflies of Venezuela (Diptera: Streblidae). *Brigham Young University Science Bulletin, Biological Series* **20**, 1.

Yang H, Chen T and Dong W (2023*a*) Comparative analysis of the mitochondrial genome of *Dermacentor steini* from different regions in China. *Parasitology* **150**, 195–205.

Yang JT, Huang XB, Wang YJ, Yang HJ, Zhang XZ and Zheng XY (2023*b*) Complete mitochondrial genome of *Penicillidia jenynsii* (Diptera: Hippoboscoidea: Nycteribiidae) and phylogenetic relationship. *Parasitology* **150**, 623–630.

Yang JT, Huang XB, Wang YJ, Yang HJ, Zhang XZ and Zheng XY (2023*c*) Complete mitogenome of *Nycteribia allotopa* Speiser, 1901 (Diptera, Hippoboscoidea, Nycteribiidae) and comparative analysis of mitochondrial genomes of Nycteribiidae. *Parasitology International* **96**, 102769.

Yang JT, Huang XB, Wang YJ, Yang HJ, Zhang XZ and Zheng XY (2023*d*) The complete mitochondrial genome of *Nycteribia parvula* Speiser, 1901 (Diptera, Nycteribiidae). *Mitochondrial DNA Part B* **8**, 276–280.

Yang JT, Wang YJ, Yang HJ, Zhang XZ, Zheng XY and Huang XB (2024) Infection status and molecular detection of pathogens carried by ectoparasites of *Miniopterus fuliginosus* bats in Yunnan, China. *Parasitology International* **98**, 102823.

Zhang DX and Hewitt GM (1997) Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. *Biochemical Systematics and Ecology* **25**, 99–120.

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX and Wang GT (2020) PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Molecular Ecology Resources* **20**, 348–355.

Zhang X, Huang X, Wang Y, Yang H, Yang J and Zheng X (2023) The complete mitochondrial genome of *Phthiridium szechuanum* (Nycteribiidae, Diptera). *Mitochondrial DNA Part B* **8**, 211–214.

Ziegler J (2003) Ordnung Diptera, Zweiflügler (Fliegen und Mücken). *Lehrbuch der Speziellen Zoologie. Begründet von Alfred Kaestner* **2**, 756–860.

https://doi.org/10.1017/S0007485324000762 Published online by Cambridge University Press