scientific reports

Check for updates

OPEN

The first mitogenome of the subfamily Stenoponiinae (Siphonaptera: Ctenophthalmidae) and implications for its phylogenetic position

Xiaoxia Lin, Ju Pu & Wenge Dong✉

Fleas are the most important insect vectors that parasitize warm-blooded animals and are known vectors of zoonotic pathogens. A recent study showed that *Stenoponia polyspina* parasitizing *Eospalax baileyi* in Zoige County have carried *Bartonella* spp. and *Spotted fever group Rickettsia* (SFGR). Accurate identification and differentiation of fleas are essential for prevention and control of zoonotic pathogens. To understand phylogenetic relationship of the subfamily Stenoponiinae, we described morphological characteristics of *S. polyspina* and sequenced its mitogenome with 14,933 bp in size and high A+T content (~ 79%). The *S. polyspina* mitogenome retained the ancestral pattern of mitochondrial gene arrangement of arthropods without rearrangement. The start codons of 13 protein-coding genes (PCGs) are traditional ATN and the stop codons are TAA or TAG. Anticodon loops of all tRNA genes were 7 bp except for $trnL_2$ and *trnD* had anticodon loops with 9 bp and the abnormal anticodon loops may be associated with frameshifting mutation. Genetic distance and Ka/Ks ratios indicated that all 13 PCGs of *S. polyspina* were subjected to purifying selection, with *cox1* at the slowest rate and *atp8* at the fastest rate. The mitogenomes of 24 species representing 7 families in the order Siphonaptera were selected to reconstruct phylogenetic tree based on concatenated nucleotide sequences of two datasets (PCGRNA matrix and PCG12RNA matrix) using Bayesian inference (BI) and Maximum likelihood (ML) methods. Phylogenetic tree supported that the superfamilies Ceratophylloidea, Vermipsylloidea, Pulicoidea were monophyletic and the superfamily Hystrichopsylloidea was paraphyletic. The family Ctenophthalmidae was monophyletic in PCGRNA-ML (codon partition) and paraphyletic in the remain trees. *S. polyspina* belongs to the subfamily Stenoponiinae was closely more related to the subfamily Rhadinopsyllinae. This paper explored phylogenetic position of diverse clades within the order Siphonaptera based on morphological and mitogenome data of *S. polyspina*. Our research enriched NCBI database of the order Siphonaptera.

Keywords *Stenoponia polyspina*, Monophyletic, Mitogenome, Stenoponiinae

Fleas (Arthropoda: Insecta: Siphonaptera) are holometabolous insects with body flattened, legs developed, climb and jump well, prick-sucking mouthparts and generally parasitize birds and mammals[1]. Fleas are important insect vectors that transmit a wide range of zoonotic agents by feeding on blood of host, including *Bartonella* spp., *Yersinia pestis*, and *Rickettsia* spp[2,3]. A recent study showed that *S. polyspina* parasitizing *Eospalax baileyi* in Zoige County carried *Bartonella* spp. and *Spotted fever group Rickettsia* (SFGR)[4]. Fleas are highly sensitive to host's temperature. They will leave their original host and look for a new host when the host's temperature is too high or descend. It will increase new or recurring risk of plague and other infectious diseases[5]. With 16 species, the subfamily Stenoponiinae is primarily found in the Palaearctic and is distinguished by its larger body size and dark brown with striking genal comb spanning most lateral portion of head[6]. Accurate differentiation and identification of fleas has been performed by morphological characteristics, such as the shape and structure of

Institute of Pathogens and Vectors, Yunnan Provincial Key Laboratory for Zoonosis Control and Prevention, Dali University, Dali 671000, Yunnan, China. ✉email: dongwenge2740@sina.com

complex genitalia and distribution of setae, spines and ctenidia of body[7–9]. However, there are still challenges in identifying related species and cryptic species with similar morphological characteristics. Therefore, Phylogenetic position was inferred in diverse families within the order Siphonaptera based on morphological characteristics and molecular data of *S. polyspina*. These results will provide technical support for accurate diagnosis and improved prevention and treatment of zoonotic parasitic diseases.

Sampling locality was Zoige County in China, which was geographically unique in the intersection of Sichuan, Qinghai, and Gansu Provinces. Some of its neighboring counties have been plagued by plague, i.e., Xiahe and Luqu Counties in Gansu Province, Maqu and Jiuzhi Counties in Qinghai Province, as well as Chengdu in Sichuan Province, which were listed as natural plague foci[10]. Additionally, Zoige County is abundant in natural resources and animals and frequent human-animal contact, so there was potential plague risk. Research on fleas in Zoige County may provide some perspective for prevention and control of flea-borne diseases.

To date, most researches on the order Siphonaptera have focused predominantly on morphology and single-gene molecules, with the taxonomic status of the family Ctenophthalmidae remaining controversial. Medvedev, 1998 placed the family Ctenophthalmidae under the family Hystrichopsyllidae and assigned it a subfamily status based on morphology[11]. This is the opposite of Lewis, 1993[12]. Whiting et al., 2008 first performed formal branch analysis of the order Siphonaptera based on *28S*, *18S*, *COII* and *EF1-α*, and reported that the family Ctenophthalmidae has highly paraphyletic[13]. Acosta et al., 2013 performed phylogenetic analyses that placed the tribe *Phalacropsyllini* based on *18S* and *28S* gene sequences. The taxonomic status of the genus *Catallagia*, previously assigned to the tribe *Phalacropsyllini*, and now assigned to the tribe *Neopsyllini*[14]. Zurita et al., 2015 assessed the monophyly of the subfamily Stenoponiinae based on *ITS1*, *ITS2*, *cox1*, *18S* gene sequences and suggested promoted to the family level[6]. Therefore, the family Ctenophthalmidae is currently in a rather confused position at the family, subfamily and genus levels. The *S. polyspina* mitogenome from the subfamily Stenoponiinae was reported for the first time in this study, which fills gap in the subfamily Stenoponiinae mitogenomes. Additionally, we selected the newly sequencing *S. polyspina* and twenty-three species of the order Siphonaptera mitogenomes downloaded from the NCBI for phylogenetic analyses, based on concatenated nucleotide sequences of two datasets (PCGRNA matrix and PCG12RNA matrix) to study the controversial phylogenetic status of the family Ctenophthalmidae within the order Siphonaptera.

Materials and methods

Sample collection and morphological identification

Stenoponia polyspina was collected from *Eospalax baileyi* (No. 133; Date: October 2021; Sampling location: Zoige County, Sichuan Province, China) for morphological and molecular data analysis. Collected specimens were preserved in 95% ethanol and stored at − 80 °C refrigerator until use. Morphological identification at the species level was performed and photographed using a SZ2-ILST dissecting microscope (Olympus, Tokyo, Japan) according to the identifying features described in the *Fauna Sinica Insecta Siphonaptera*[1]. Specimens were deposited at the Institute of Pathogen and Vector Biology, Dali University. This study was approved by the Animal Ethics Committee of Dali University (Approval No. MECDU-201912–20).

Mitogenome sequencing, assembly and annotation

We extracted total genomic DNA from flea tissue using the DNeasy Tissue Kit (Qiagen, Germany), following the manufacturer's protocol. Total DNA was subsequently sent to Shanghai Winnerbio Technology Co., Ltd. (Shanghai, China) for sequencing. All amplicons were sequenced with Illumina Novoseq 6000 platform[15]. To make assembly sequence more accurate, raw data were quality controlled and spliced using Trimmomatic v. 0.35[16] to remove low-quality and splice contaminated reads to obtain clean reads. Clean reads were assembled de novo using MitoZ 2.3[17]. To ensure accuracy of mitogenome, sequencing depth of assembled sequences should generally be ≥ 100x. The sequencing depth of *S. polyspina* was 206.157X, which ensured accuracy of single bases in the *S. polyspina* mitogenome.

The MITOS server[18] was used for initial annotation. However, the MITOS server often could not correctly identify the start and stop codons, so these were manually annotated by aligning the sequences with the annotated mitogenome data of the other fleas downloaded from GenBank. Annotation results were imported into Geneious Prime, compared with the reference dataset, and the start codon and stop codon positions of protein-coding genes were adjusted. Twenty-two tRNA genes were predicted using ARWEN[19], tRNAscan-SE 2.0[20] and MITOS[18]. Two ribosomal RNA (rRNA) genes were confirmed by comparison with homologous genes from other published flea species[21,22]. The *S. polyspina* mitogenome was deposited in GenBank under accession number OR834393.

Nucleotide composition and codon usage were calculated using MEGA11[23], and the Kimura-2-parameter model was used to calculate genetic distance. KAKS_Calculator 2.0[24] was used to calculate the non-synonymous substitution rate (Ka) and synonymous substitution rate (Ks).

Phylogenetic analysis

Phylogenetic tree of 24 species representing 7 families of the order Siphonaptera was constructed and *Boreus elegans* as the outgroup by combining the newly sequencing *S. polyspina* and twenty-three species of the order Siphonaptera mitogenomes downloaded from the NCBI (Table S1)[21,22,25–35]. Sequences were aligned in batches using MAFFT v7.313[36]. Ambiguously aligned positions were trimmed using Gblocks[37] and alignment of individual genes was concatenated into two datasets: (i) the PCGRNA matrix, which included thirteen PCGs and two rRNA genes (13,025 bp in total) and (ii) the PCG12RNA matrix, which included the first and second codon positions of the thirteen PCGs and two rRNA genes (9,269 bp in total). Using ModelFinder[38], two datasets were identified based on BIC criterion to construct best substitution model for ML tree and BI tree (Tables S2 and S3). Maximum likelihood (ML) and Bayesian analyses (BI) were inferred using IQ–TREE[39] and MrBayes[40],

respectively. The ML tree was selected by an ultrafast bootstrap approximation approach with 5,000 replicates. The BI tree was inferred with four independent Markov chains (MCMC) run for 1,000,000 generations and sampled every 1000 generations, with the first 25% discarded as burn-in. FigTree v.1.4.4 (http://tree.bio.ed.ac.uk/ software/figtree/) was used to visualize the phylogenetic trees, in which PCG12RNA-ML (gene partition), PCG12RNA-ML (codon partition) and PCG12RNA-BI (gene partition) were used to obtain consensus trees, and finally we obtained six phylogenetic trees.

Results

Morphological identification

Male: The genal comb with 14 spines, eyes vestigial, occipital row bristles in order 4, 6, 9, pronotal comb with 40 spines. The abdominal comb I 32 spines, tergum II-VI apical spinelet in order 4, 2, 5, 0, 0, four antepygidial bristles. There are four long bristles and four short bristles on posterior margin of sternum VIII. The movable process is longer, 4.5–5.0 times as long as wide. End of immovable process bluntly rounded almost as high as movable process. The end of the apical arm of sternum IX extends distinctly posteriorly with bristles of approximately 40, and the proximal part of the anterior margin usually projects upward (Fig. 1A).

Female: The genal comb with 17 spines, eyes vestigial, occipital row in order 4, 6, 10, pronotal with comb 40 spines. The abdominal comb I 42 bristles, tergum II-VI apical spinelet in order 6, 4, 3, 4, 0. The tergum VIII has a small inner concavity at the posterior margin, with coarse bristles in a row behind the six antepygidial bristles. The posterior margin of sternum VII has a deeper and wider concavity forming a wide dorsal lobe and a narrower ventral lobe, but there is greater variability in shape (Fig. 1B).

Mitogenome structure and organization

The *S. polyspina* mitogenome was 14,933 bp in size, except for non-coding region (incomplete sequenced in this study). Gene arrangement patterns of *S. polyspina* are consistent with that of the putative arthropod ancestor (*Drosophila yakuba*)[41], which contained 37 typical mitochondrial genes (13 PCGs, 22 tRNAs and 2 rRNAs). Among them, 23 genes were located on the majority strand (J-strand), including 9 protein-encoding genes and 14 transfer RNA genes, and the remaining 14 genes were located on the minority strand (N-strand) (Fig. 2, Table 1). Nucleotide composition of *S. polyspina* was A (39.2%), T (39.6%), C (13.0%), G (8.1%), and AT content (78.8%) was higher than GC content (21.2%) with significant AT bias. Additionally, the *S. polyspina* mitogenome presented negative AT-skew (-0.005) and GC-skew (-0.231) (Table 2).

S. polyspina has an extremely compact mitochondrial (mt) genome with 37 genes on a single circular chromosome. Overlapping regions and intergenic spacer regions with different sizes were found. There were 15 gene overlaps ranging from 1–8 bp in size. The longest overlap sequence was 8 bp between *trnW* and *trnC*. There were 10 gene intergenic sequences ranging from 1 to 10 bp in size. The longest intergenic sequence was 10 bp between $trnS_2$ and *nad1* (Table 1).

Protein-coding genes and codon usage

The total length of 13 PCGs was 11,124 bp, accounting for 74.5% of the *S. polyspina* mitogenome. AT content of 13 PCGs with ranging from *cox1* (71.3%) to *nad6* (86.5%) (Table 2). The start codons of all PCGs are ATN. The stop codons are TAA or TAG (Table 1). The relative synonymous codon usage (RSCU) of the *S. polyspina* mitogenome was calculated. The most frequently used codons were UUA (4.76). Twenty-eight codons were preference codons (RSCU > 1). Three GC-rich codons, CUG, ACG, and GUG, were unused among 64 codons encoded invertebrate mitochondrion (Fig. 3). Most codons with high RSCU values ending with A or T and had significant AT preference (Fig. 3). Leucine (16.2%), Isoleucine (11.6%), Serine (10.0%), Phenylalanine (9.7%),

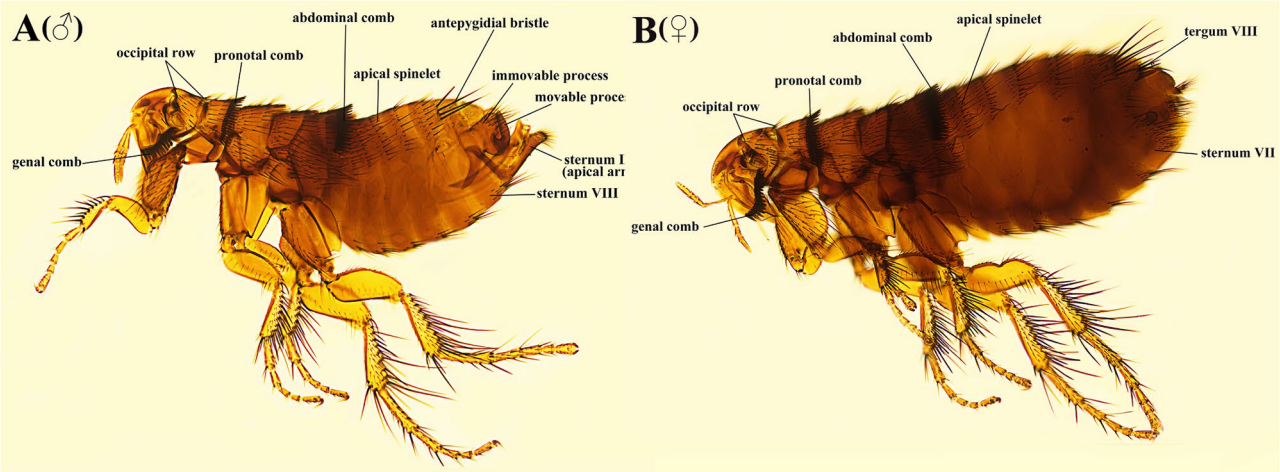

Figure 1. Morphological characteristics of 1(**A**) male and 1(**B**) female of the *Stenoponia polyspina*.

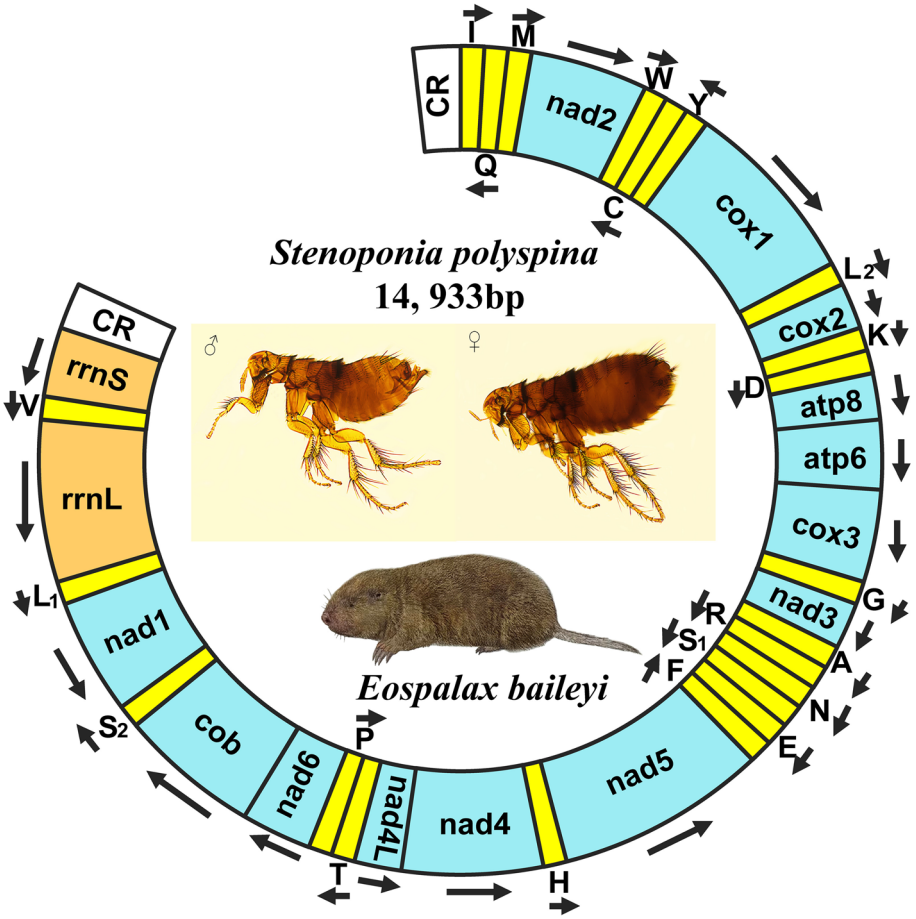

Figure 2. The circular map of the *Stenoponia polyspina* mitogenome (note: The morphological figure of *Eospalax baileyi* from Handbook of the mammals of the world: vol. 7: rodents II)[42].

Methionine (9.3%), and Asparagine (6.5%) were more abundant than the other amino acids and accounted for more than half of all amino acids (63.3%) (Fig. 4).

Ka/Ks and genetic distance

The evolutionary rates of 13 PCGs of *S. polyspina* were analyzed. Ka/Ks ratios ranged from 0.0220 to 0.2315. Ka/Ks ratios were less than 1. It indicated that purifying selection was dominant in the *S. polyspina* mitogenome. The lowest Ka/Ks ratios for *cox1* gene indicated that the gene has been subjected to stronger functional constraints during evolution and had the lowest mutation. The highest Ka/Ks ratios for *atp8* gene indicated that the gene has been subjected to relatively weaker selection pressure and has stronger non-synonymous mutation (Fig. 5). Additionally, Kimura-2 parameter (K2P) distances among 13 PCGs in the *S. polyspina* mitogenome was calculated, indicating that *cox1* gene had the smallest distance, and *atp8* gene had the largest distance (Fig. 5). In conclusion, *atp8* was the most variable gene, while *cox1* was the most conserved gene.

tRNA and rRNA genes

We identified 22 tRNA genes typical of bilateral animals in the *S. polyspina* mitogenome. These genes ranged from 61 to 69 bp in length, with a total length 1405 bp (Table 1). The inferred tRNA genes all presented typical cloverleaf secondary structure, except for $trnS_1$ which lacked the D-arm (Fig. 6). A total of 18 G-U mismatches were showed in 22 tRNA genes. The anticodon loops were 7 bp for all tRNAs except for the 9 bp anticodon loops of $trnL_2$ and *trnD* (Fig. 6).

Both the large subunits rRNA gene (*rrnL*) and small subunits rRNA gene (*rrnS*) were on the N-strand. The *rrnL* gene with 1,299 bp was located between $trnL_1$ and *trnV*, while the *rrnS* gene with 782 bp was located between *trnV* and control region (*CR*). Total length of two rRNA genes in *S. polyspina* mitogenome was 2081 bp, with an AT content of 81.9% and strong AT bias (Table 2).

Phylogenetic analysis

The ML and BI trees constructed using nucleotide sequences in this study, it revealed that the superfamilies Ceratophylloidea, Vermipsylloidea and Pulicoidea were monophyletic, but the superfamily Hystrichopsylloidea was paraphyletic (Figs. 7 and 8). The family Ctenophthalmidae was monophyletic in PCGRNA-ML (codon partition)

Gene	Strand	Location	Size	Anticodon	Start/stop codon	Intergenic Nucleotides
tRNA-Ile(I)	H	150–212	63	GAU		
tRNA-Gln(Q)	L	212–280	69	UUG		− 1
tRNA-Met(M)	H	284–347	64	CAU		3
nad2	H	351–1364	1014		ATT/TAA	3
tRNA-Trp(W)	H	1363–1427	65	UCA		− 2
tRNA-Cys(C)	L	1420–1480	61	GCA		− 8
tRNA-Tyr(Y)	L	1481–1543	63	GUA		
cox1	H	1541–3076	1536		ATC/TAA	− 3
tRNA-Leu(L2)	H	3079–3142	64	UAA		2
cox2	H	3143–3823	681		ATG/TAA	
tRNA-Lys(K)	H	3825–3893	69	CTT		1
tRNA-Asp(D)	H	3894–3956	63	GUC		
atp8	H	3957–4118	162		ATT/TAA	
atp6	H	4112–4786	675		ATG/TAA	− 7
cox3	H	4786–5568	783		ATG/TAA	− 1
tRNA-Gly(G)	H	5569–5633	65	UCC		
nad3	H	5634–5984	351		ATT/TAG	
tRNA-Ala(A)	H	5983–6045	63	UGC		− 2
tRNA-Arg(R)	H	6045–6105	61	UCG		− 1
tRNA-Asn(N)	H	6110–6174	65	GUU		4
tRNA-Ser(S1)	H	6175–6242	68	UCU		
tRNA-Glu(E)	H	6243–6304	62	UUC		
tRNA-Phe(F)	L	6303–6365	63	GAA		− 2
nad5	L	6365–8080	1716		ATG/TAA	− 1
tRNA-His(H)	L	8082–8142	61	GUG		1
nad4	L	8142–9479	1338		ATG/TAA	− 1
nad4L	L	9473–9760	288		ATG/TAG	− 7
tRNA-Thr(T)	H	9763–9825	63	UGU		2
tRNA-Pro(P)	L	9826–9889	64	UGG		
nad6	H	9892–10,398	507		ATC/TAA	2
Cytb	H	10,398–11,531	1134		ATG/TAG	− 1
tRNA-Ser(S2)	H	11,530–11,592	63	UGA		− 2
nad1	L	11,603–12,541	939		ATG/TAA	10
tRNA-Leu(L1)	L	12,543–12,605	63	UAG		1
rrnL	L	12,606–13,904	1299			
tRNA-Val(V)	L	13,905–13,967	63	UAC		
rrnS	L	13,967–14,748	782			− 1
incomplete CR		1–149;14,749–14,933	149; 185			

Table 1. Organization of the *Stenoponia polyspina* mitogenome.

and was paraphyletic in the remaining trees, including the subfamilies Rhadinopsyllinae, Ctenophthalminae, Neopsyllinae and Stenoponiinae, but was poorly supported (Boot = 23–65; Bpp = 0.507–0.857). The subfamily Stenoponiinae was closely related to the subfamily Rhadinopsyllinae, with strongly support for both the ML and BI trees (Boot = 94–97; Bpp = 0.857–1).

Discussion

The order Siphonaptera (fleas) are extremely specialized insects. To adapt to their hosts, fleas developed a variety of combs and backward-facing bristles that made it difficult for the hosts to get rid of it. Some research showed the adaptive characteristics of fleas and suggested that the super-heightened development of setae, spines and combs was often related to host characteristics—these structures were related to life and death of parasitic fleas[43]. In addition, some researchers have studied the number, thickness, length, position, and size of inter-spine spaces of pronotal comb, as well as the overlapping of certain combs, and thought that they were not only related to the nature of host's plumage and the size of diameter, but also to ecological habits and environment of host and flea species itself, which could be categorized as adaptive features[44]. Although *E. baileyi* has a large amount of hair, it is short, thin and soft. *S. polyspina* has developed abundant combs to grasp its host.

In the complete mitogenome of *Ctenocephalides felis felis*, we found that the CR region is very long (up to 6 Kb), and there are many repeated sequences[27]. This study based on next-generation sequencing (NGS) with

Gene	A%	T%	C%	G%	AT%	AT-skew	GC-skew
nad1	32.3	46.3	6.8	14.6	78.6	− 0.179	0.363
nad2	35.6	46.3	11.7	6.4	81.9	− 0.13	− 0.293
nad3	32.5	48.4	11.7	7.4	80.9	− 0.197	− 0.224
nad4	33.9	45.7	7.2	13.2	79.6	− 0.147	0.289
nad4L	36.1	47.6	4.2	12.2	83.7	− 0.137	0.489
nad5	35.2	45	6.8	13	80.2	− 0.123	0.316
nad6	39.6	46.9	8.7	4.7	86.5	− 0.084	− 0.294
cox1	31.7	39.6	15.6	13.2	71.3	− 0.111	− 0.084
cox2	35.1	40.2	14.8	9.8	75.3	− 0.068	− 0.202
cox3	30.5	42.1	15.3	12	72.6	− 0.16	− 0.121
atp6	34.7	42.1	13.8	9.5	76.8	− 0.097	− 0.185
atp8	38.3	43.8	14.2	3.7	82.1	− 0.068	− 0.586
cob	32.6	42.5	14.5	10.4	75.1	− 0.131	− 0.163
PCGs	33.9	43.9	11.1	11.1	77.8	− 0.128	0.002
rRNAs	40.5	41.4	6.1	12.1	81.9	− 0.011	0.328
tRNAs	41	38.6	8.3	12	79.6	0.029	0.182
Whole genome	39.2	39.6	13	8.1	78.8	− 0.005	− 0.231

Table 2. Nucleotide composition and skewness of the *Stenoponia polyspina* mitogenome.

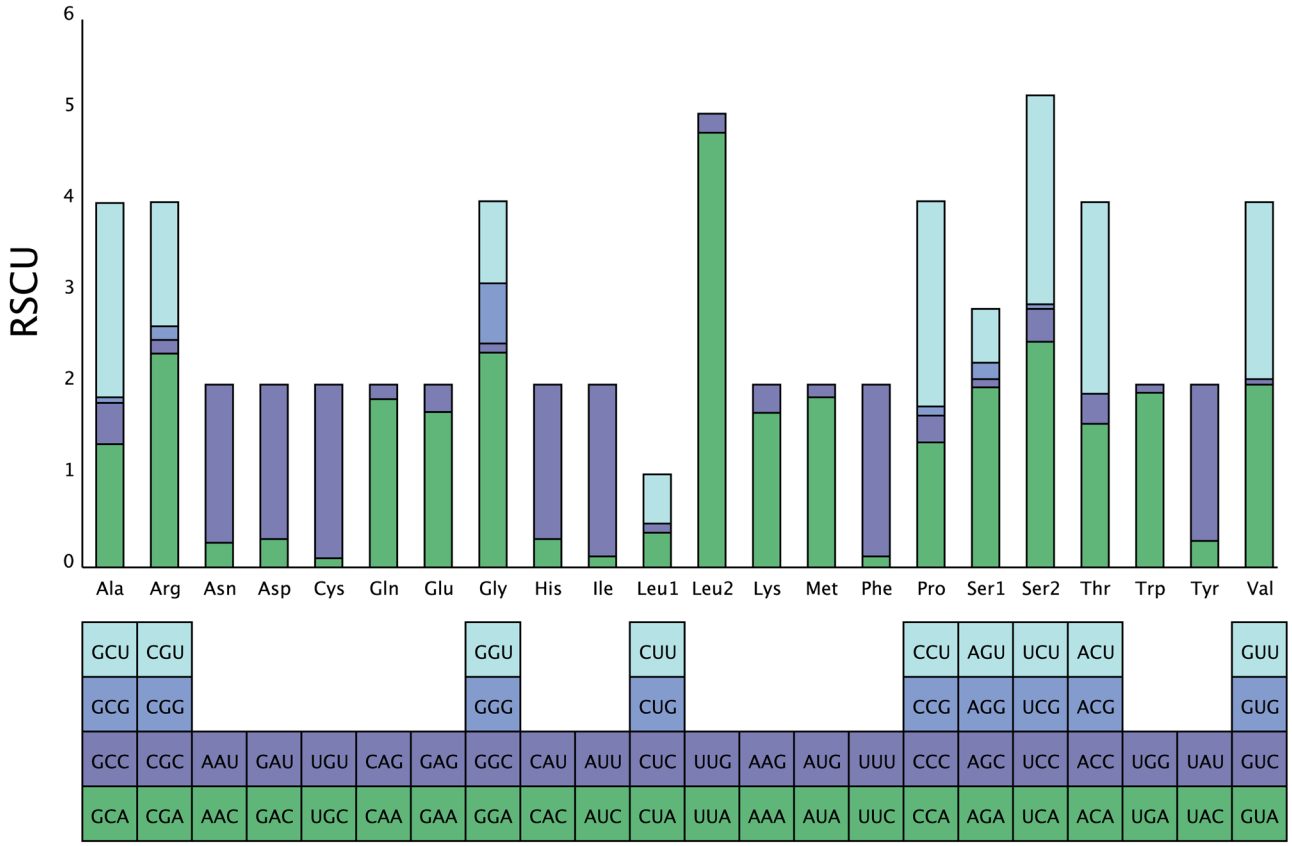

Figure 3. Codon distribution and RSCU of the *Stenoponia polyspina* mitogenome.

a short reading length. The presence of repeated sequences, stem-loop structure and AT-rich motifs in the CR region makes it difficult to obtain the complete mitogenome of *S. polyspina*. We sequenced only the partial control region of the *S. polyspina* mitogenome, and AT content reached 78.8%. If control region of the *S. polyspina* mitogenome was sequenced completely, the AT content would have been greater than 80%. A study of holometabolous insects has indicated that sequences of the orders Lepidoptera, Trichoptera, and Siphonaptera were more A + T rich than those of the orders Diptera and Mecoptera[45]. We speculated that fleas require more

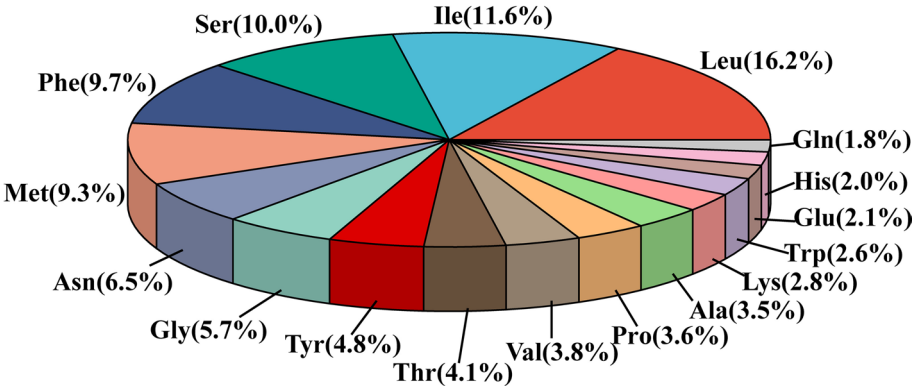

Figure 4. The proportion of each amino acid of the *Stenoponia polyspina* mitogenome.

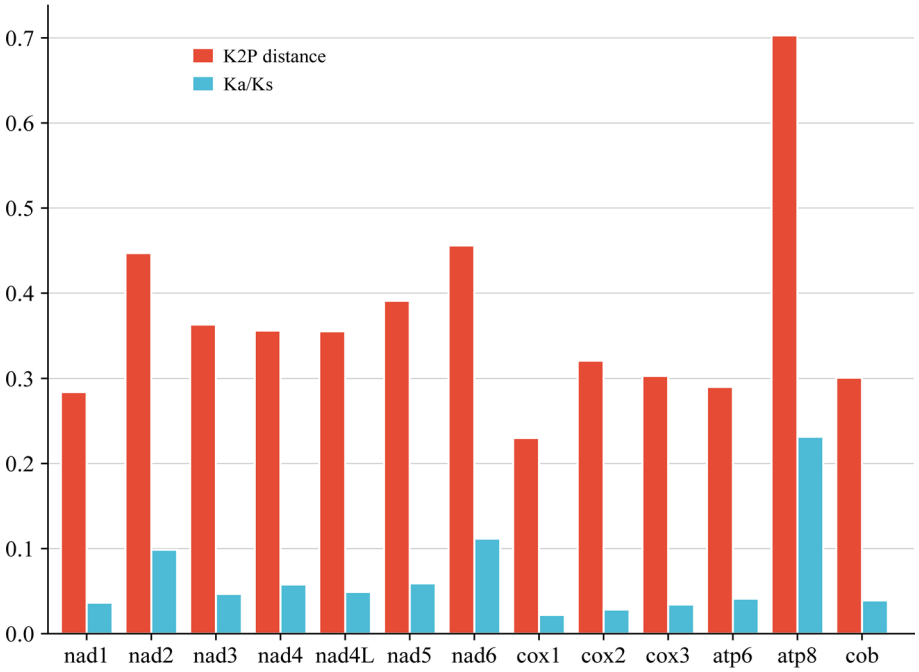

Figure 5. Ratios of Ka/Ks and K2P distances of each protein-coding gene among *Stenoponia polyspina*.

energy for jumping, and that fleas specialize in sucking blood and must have adenosine triphosphate (ATP) in their blood to do so[1], which result in their high AT content.

We found that the *S. polyspina* mitogenome is highly conserved and its gene arrangement patterns are consistent with that of the putative arthropod ancestor (*D. yakuba*)[41]. There are 28 codons with RSCU values > 1, which is consistent with other holometabolous insects[46]. The 22 tRNA genes all presented typical cloverleaf secondary structure except for $trnS_1$ which lacked the D-arm, which is typical feature of the metazoan mitogenome[47]. Additionally, we calculated Ka/Ks ratios for 13 PCGs of *S. polyspina*, and all ratios < 1, indicating that they were subjected to purifying selection and had the slower evolutionary rate. The slower evolutionary rate of *S. polyspina* might be the following reasons: (1) Due to long-term interactions between parasites and hosts, some features of parasites are closely related to biological characteristics of their hosts[48]. The slower evolutionary rate of *S. polyspina* might be related to its hosts. *S. polyspina* only parasitize *E. baileyi* as a single host and has a small range of activity, mostly living in burrows and activity time was shorter outside the nest, lasting only a few minutes to a few tens of minutes[49]. Like the lazy behavior of the *E. baileyi*, it may make the *S. polyspina* relatively conservative in genetics and slow in evolution. (2) Some research indicated that the evolutionary rate of the bilaterian mitogenome was: endoparasites > ectoparasites with reduced locomotory capacity > free-living lineages with low locomotory capacity > parasitoids > ectoparasites with high locomotory capacity > micropredators and strongly locomotor free-living animals[50]. Compared with suck lice, which parasitized on rodents, *S. polyspina* with high locomotory capacity (jumps well), parasitized only adult, while suck lice are ectoparasites with lower locomotory capacity permanently parasitized host. Therefore, differences in life history and locomotor capacity may lead to a conserved mitogenome of *S. polyspina* and extensive fragmented mitogenomes of suck lice. (3) Generally,

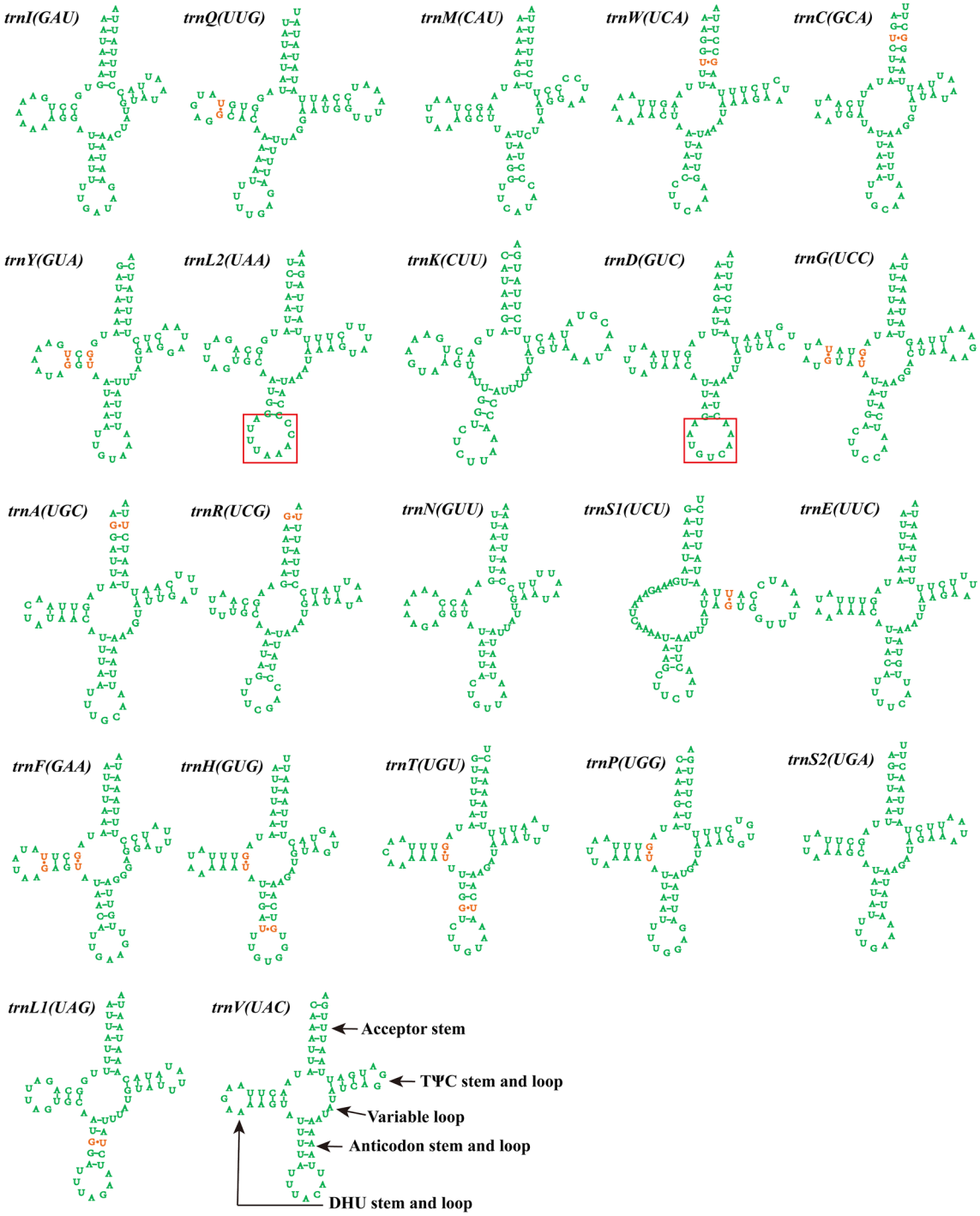

Figure 6. Secondary structure of 22 tRNA genes from the *Stenoponia polyspina* mitogenome.

the anticodon loops had 7 bases, whereas the anticodon loops of $trnL_2$ and *trnD* of *S. polyspina* were 9 bp. This variant was also found in tRNAs of other arthropod mitogenomes, such as *Dermatophagoides pteronyssinus* ($trnL_2$)[51] and *Camelus bactrianus ferus* ($trnS_1$)[52]. This indicate that tRNAs with 8-base or 9-base anticodon loops were associated with frameshifting[53]. Therefore, mutation might be related to occasional frameshifting during translation, resulting from defective in mitochondrial protein synthesis, which influenced on growth, development and slow evolution of species.

Phylogenetic tree supported that the superfamilies Ceratophylloidea, Vermipsylloidea, Pulicoidea were monophyletic, but the superfamily Hystrichopsylloidea was paraphyletic, consistent with previous studies[22]. The family Ctenophthalmidae was rich encompassed nearly a quarter of flea species, and its taxonomic status remains confused[13,54]. The family Ctenophthalmidae comprises nine subfamilies, with mitogenomes sequencing for only six species, representing four subfamilies (Ctenophthalminae, Stenoponiinae, Rhadinopsyllinae, Neopsyllinae).

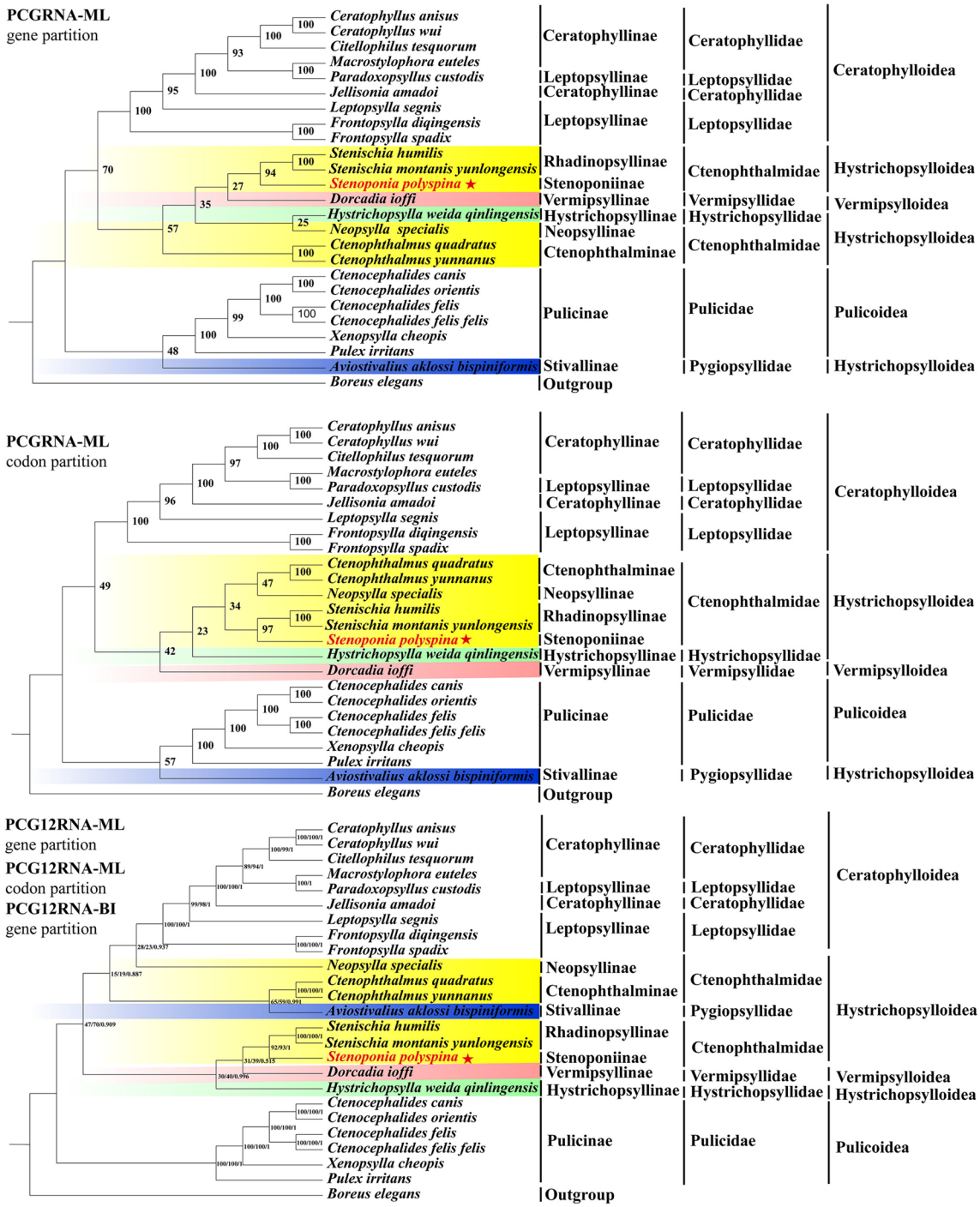

Figure 7. Phylogenetic tree of 24 species of the order Siphonaptera was constructed by maximum likelihood (ML).

Notably, the family Ctenophthalmidae was monophyletic in the PCGRNA-ML (codon partition) phylogenetic tree, which has never been proposed in previous studies. Moreover, the relationships of the family Ctenophthalmidae with the families Vermipsyllidae, Hystrichopsyllidae and Pygiopsyllidae were inconsistent and poorly supported in different datasets (PCGRNA matrix and PCG12RNA matrix) and different methods (ML and BI). This implies that more research is necessary to fully understand the taxonomic status of the family Ctenophthalmidae, as the current findings are insufficient. Additionally, some studies have assessed the monophyletic of the subfamily Stenoponiinae based on *cox1* and *18S* genes and suggested that the subfamily Stenoponiinae should be elevated to the family level. At present study, the only one mitogenomes of the subfamily Stenoponiinae was sequenced, which is insufficiently rich in data. Therefore, more species mitogenomes from diverse families within the order Siphonaptera will be sequenced in the future, and the phylogenetic position at the controversial family level will be accurately assessed.

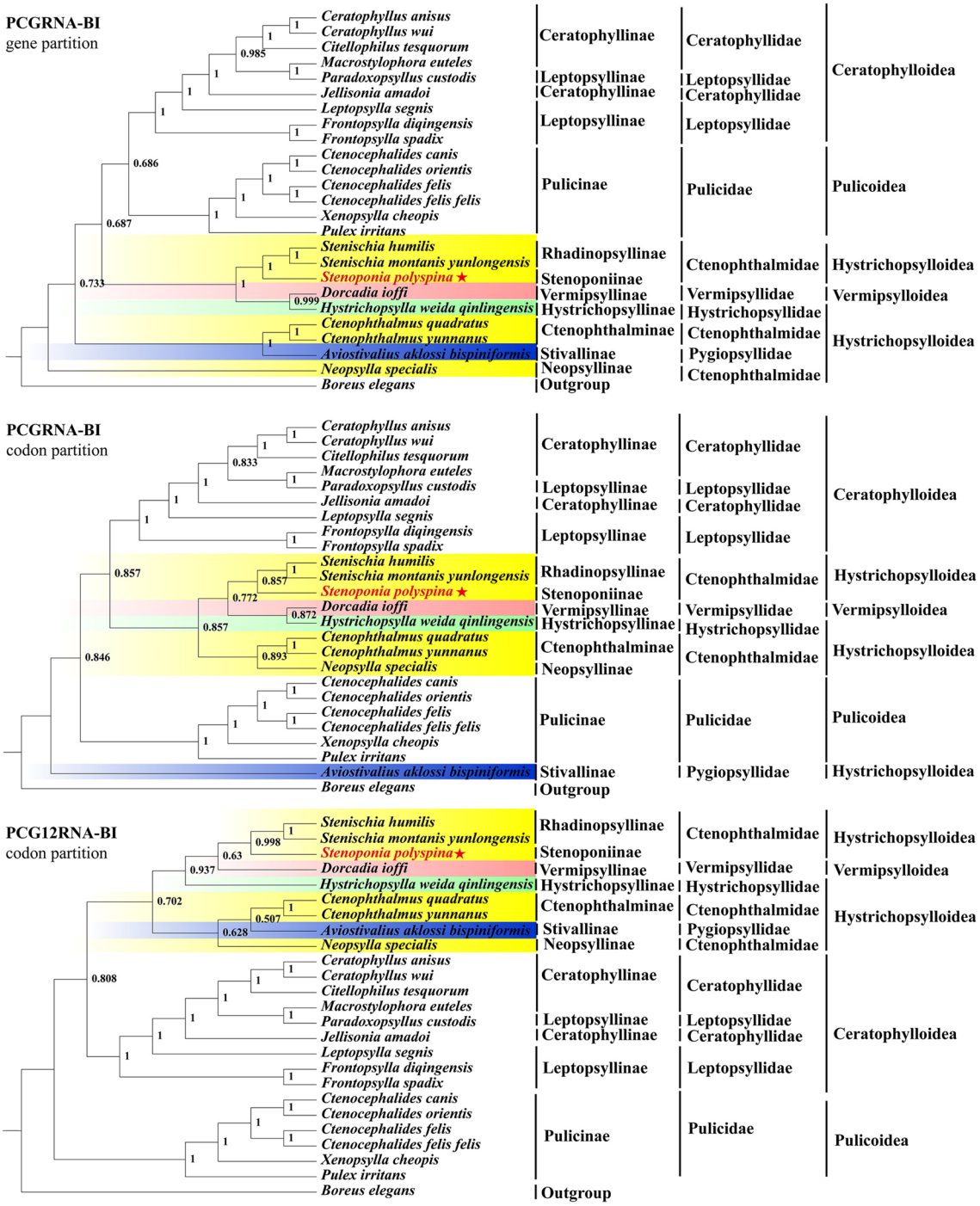

Figure 8. Phylogenetic tree of 24 species of the order Siphonaptera was constructed by Bayesian inference (BI).

Conclusion

The *S. polyspina* mitogenome was reported for the first time in this study and retained ancestral mitogenome organization of arthropods, which was in stark contrast to the extensive fragmented mitogenomes of parasitic lice. This phenomenon might be closely related to host, life history and locomotory capacity. Phylogenetic tree supported that the superfamilies Ceratophylloidea, Vermipsylloidea and Pulicoidea were monophyletic, but the superfamily Hystrichopsylloidea was paraphyletic. The phylogenetic relationships among families Ctenophthalmidae, Vermipsyllidae, Hystrichopsyllidae and Pygiopsyllidae have small variances in different datasets (PCGRNA matrix and PCG12RNA matrix) and different methods (ML and BI). *S. polyspina* belongs to the subfamily Stenoponiinae and was closely more related to the subfamily Rhadinopsyllinae. Due to insufficient mitogenomes data of the order Siphonaptera, the taxonomic status of the family Ctenophthalmidae will need to be explored in the future. Our results provide a significant advance in evolution process of the order Siphonaptera and an important perspective for understanding evolution of the family Ctenophthalmidae. These mitogenomes

will provide novel molecular markers for studying the taxonomy and phylogeny of the order Siphonaptera in the future.

Statement

All methods were performed in accordance with the relevant guidelines and regulations.

Data availability

The nucleotide sequences of the S. polyspina mitogenome were deposited in GenBank (https://www.ncbi.nlm.nih.gov/) under accession number OR834393.

Received: 24 February 2024; Accepted: 1 August 2024
Published online: 06 August 2024

References

1. Wu, H. Y. *Fauna Sinica Insecta Siphonaptera* (Science Press, 2007).
2. Bitam, I., Dittmar, K., Parola, P., Whiting, M. F. & Raoult, D. Fleas and flea-borne diseases. *Int. J. Infect. Dis.* **14**, e667-676 (2010).
3. Hamzaoui, B. E., Zurita, A., Cutillas, C. & Parola, P. Fleas and flea-borne diseases of North Africa. *Acta Trop.* **211**, 105627 (2020).
4. Xiang, Y. *Detection and Genetic Analysis of Vector-borne Pathogens Carried by Yak and Plateau zokor and Their Ectoparasites in Zoige County* (Southwest Minzu University, 2022).
5. Kreppel, K. S., Telfer, S., Rajerison, M., Morse, A. & Baylis, M. Effect of temperature and relative humidity on the development times and survival of *Synopsyllus fonquerniei* and *Xenopsylla cheopis*, the flea vectors of plague in Madagascar. *Parasit Vectors* **9**, 82 (2016).
6. Zurita, A., Callejón, R., De Rojas, M., Gómez López, M. S. & Cutillas, C. Molecular study of *Stenoponia tripectinata tripectinata* (Siphonaptera: Ctenophthalmidae: Stenoponiinae) from the Canary Islands: taxonomy and phylogeny. *Bull. Entomol. Res.* **105**, 704–711 (2015).
7. Hastriter, M. W., Miller, K. B., Svenson, G. J., Martin, G. J. & Whiting, M. F. New record of a phoretic flea associated with earwigs (Dermaptera, Arixeniidae) and a redescription of the bat flea *Lagaropsylla signata* (Siphonaptera, Ischnopsyllidae). *ZooKeys* **657**, 67–79 (2017).
8. López Berrizbeitia, M. F., Hastriter, M. W. & Díaz, M. M. A New Flea Species of the Genus *Cleopsylla* (Siphonaptera: Stephanocircidae) from Northwestern Argentina. *J. Parasitol.* **102**, 514–519 (2016).
9. Linardi, P. M. & Santos, J. L. *Ctenocephalides felis felis* vs. *Ctenocephalides canis* (Siphonaptera: Pulicidae): Some issues in correctly identify these species. *Rev. Bras. Parasitol. Vet.* **21**, 345–354 (2012).
10. Wu, C. X. *et al.* Plague epidemic surveillance in Ruoergai county, 2006–2010. *Modern Prevent. M Edicine* **28**, 461–464 (2012).
11. Medvedev, S. G. Classification of fleas (Order Siphonaptera) and its theoretical foundations. *Entomol. Rev.* **78**, 1080–1093 (1998).
12. Lewis, R. E. Classification of fleas (Order Siphonaptera) and its theoretical foundations Part 8. New taxa described between 1984 and 1990, with a current classification of the order. *Entomol. Soc. Am.* **30**, 239–256 (1993).
13. Whiting, M., Whiting, A., Hastriter, M. & Dittmar, K. A molecular phylogeny of fleas (Insecta: Siphonaptera): origins and host associations. *Cladistics* **24**, 677–707 (2008).
14. Acosta, R. & Morrone, J. J. Phylogenetics of the tribe Phalacropsyllini (Siphonaptera: Ctenophthalmidae: Neopsyllinae) based on molecular and morphological evidence. *Zootaxa* **3630**, 333 (2013).
15. Modi, A., Vai, S., Caramelli, D. & Lari, M. The Illumina Sequencing Protocol and the NovaSeq 6000 System. *Methods Mol. Biol.* **2242**, 15–42 (2021).
16. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
17. Meng, G., Li, Y., Yang, C. & Liu, S. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucl. Acids Res.* **47**, e63 (2019).
18. Bernt, M. *et al.* MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* **69**, 313–319 (2013).
19. Laslett, D. & Canbäck, B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* **24**, 172–175 (2008).
20. Chan, P. P., Lin, B. Y., Mak, A. J. & Lowe, T. M. tRNAscan-SE 2.0: Improved detection and functional classification of transfer RNA genes. *Nucl. Acids Res.* **49**, 9077–9096 (2021).
21. Liu, Y. F. *et al.* Analysis of complete mitogenomes and phylogenetic relationships of Frontopsylla spadix and Neopsylla specialis. *Front. Veterinary Sci.* https://doi.org/10.3389/fvets.2023.1250381 (2023).
22. Chen, B. *et al.* Complete mitochondrial genome of Ctenophthalmus quadratus and Stenischia humilis in China provides insights into fleas phylogeny. *Front. Veterinary Sci.* https://doi.org/10.3389/fvets.2023.1255017 (2023).
23. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
24. Wang, D., Zhang, Y., Zhang, Z., Zhu, J. & Yu, J. KaKs_Calculator 2.0: A toolkit incorporating gamma-series methods and sliding window strategies. *Genomics Proteomics Bioinf.* **8**, 77–80 (2010).
25. Beckenbach, A. T. Mitochondrial genome sequences of representatives of three families of scorpionflies (Order Mecoptera) and evolution in a major duplication of coding sequence. *Genome* **54**, 368–376 (2011).
26. Driscoll, T. P. *et al.* A chromosome-level assembly of the cat flea genome uncovers rampant gene duplication and genome size plasticity. *BMC Biol.* **18**, 70 (2020).
27. Zhang, Y., Nie, Y., Deng, Y. P., Liu, G.-H. & Fu, Y. T. The complete mitochondrial genome sequences of the cat flea Ctenocephalides felis felis (Siphonaptera: Pulicidae) support the hypothesis that C. felis isolates from China and USA were the same C. f. felis subspecies. *Acta Trop.* **217**, 217 (2021).
28. Zhang, Y. *et al.* Mitochondrial phylogenomics provides insights into the taxonomy and phylogeny of fleas. *Parasit. Vectors* **15**, 223 (2022).
29. Wei, F. *et al.* The complete mitochondrial genome of Xenopsylla cheopis (Siphonaptera: Pulicidae). *Mitochondrial DNA. Part B, Resources* **7**, 170–171 (2022).
30. Beliavskaia, A. *et al.* Metagenomics of culture isolates and insect tissue illuminate the evolution of Wolbachia, Rickettsia and Bartonella symbionts in Ctenocephalides spp. fleas. *Microb. Genomics* https://doi.org/10.1099/mgen.0.001045 (2023).
31. Tan, L. *et al.* The complete mitochondrial genome of the flea Hystrichopsylla weida qinlingensis (Siphonaptera: Hystrichopsylla). *Mitochondrial DNA. Part B, Resources* **8**, 501–503 (2023).
32. Xiang, H. T., Wen, F. Q. & Wang, G. L. The complete nucleotide sequence of the mitochondrial genome of Dorcadia ioffi (Siphonaptera: Vermipsyllidae). *Mitochondrial DNA. Part B, Resources* **2**, 389–390 (2017).
33. Tan, L., Guan, X., Zhang, L., Zhu, F. & Lei, C. The complete mitochondrial genome of the flea Ceratophyllus wui (Siphonaptera: Ceratophyllidae). *Mitochondrial DNA. Part B, Resources* **3**, 401–402 (2018).

34. Cameron, S. L. The complete mitochondrial genome of a flea, Jellisonia amadoi (Siphonaptera: Ceratophyllidae). *Mitochondrial DNA* **26**, 289–290 (2015).
35. Liu, Y. F. *et al.* Complete mitogenomes characterization and phylogenetic analyses of Ceratophyllus anisus and Leptopsylla segnis. *Front. Veterinary Sci.* https://doi.org/10.3389/fvets.2023.1218488 (2023).
36. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
37. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000).
38. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
39. Nguyen, L. T., Schmidt, H. A., Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
40. Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Boil.* **61**, 539–542 (2012).
41. Clary, D. O. & Wolstenholme, D. R. The mitochondrial DNA molecular of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *J. Mol. Evol.* **22**, 252–271 (1985).
42. Wilson, D. E., Lacher, T. E. & Mittermeier, R. A. *Handbook of the mammals of the world rodents II* (Lybx Edicions, 2017).
43. Traub, R. Co-evolution of fleas and mammals. *Ent Kyoto*, 324 (1980).
44. Traub, R. The relationship between the spines, combs and other skeletal features of fleas(Siphonaptera) and the vestiture, affinities and habits of their hosts. *J. Med. Ent.* **9**, 601 (1972).
45. Song, F. *et al.* Capturing the Phylogeny of Holometabola with Mitochondrial Genome Data and Bayesian Site-Heterogeneous Mixture Models. *Genome Biol. Evol.* **8**, 1411–1426 (2016).
46. Sang, L. Q. *Codon preference and evolution of insect mitochondrial gene Type* (Shaanxi Normal University, 2019).
47. Wang, Y., Cao, J. J., Li, N., Ma, G. Y. & Li, W. H. The first mitochondrial genome from Scopuridae (Insecta: Plecoptera) reveals structural features and phylogenetic implications. *Ann. Entomol. Soc. Am.* **122**, 893–902 (2019).
48. Louhi, K. R., Karvonen, A., Rellstab, C. & Jokela, J. Is the population genetic structure of complex life cycle parasites determined by the geographic range of the most motile host? *Infect. Genet. Evol.* **10**, 1271–1277 (2010).
49. Wei, W., Wang, Q., Zhou, W., Fan, N. The population dynamics and dispersal of plateau zokor after removing. *Shou lei xue bao = Acta theriologica sinica* **17**, 53–61, (1997).
50. Jakovlić, I. *et al.* Mitogenomic evolutionary rates in bilateria are influenced by parasitic lifestyle and locomotory capacity. *Nat. Commun.* **14**, 6307 (2023).
51. Dermauw, W., Van Leeuwen, T., Vanholme, B. & Tirry, L. The complete mitochondrial genome of the house dust mite *Dermatophagoides pteronyssinus* (Trouessart): A novel gene arrangement among arthropods. *BMC Genomics* **10**, 107 (2009).
52. Cui, P. *et al.* A complete mitochondrial genome sequence of the wild two-humped camel (*Camelus bactrianus ferus*): An evolutionary history of camelidae. *BMC Genomics* **8**, 241 (2007).
53. Tuohy, T. M., Thompson, S., Gesteland, R. F. & Atkins, J. F. Seven, eight and nine-membered anticodon loop mutants of tRNA(2Arg) which cause +1 frameshifting. Tolerance of DHU arm and other secondary mutations. *J. Mol. Biol.* **228**, 1042–1054 (1992).
54. Zhu, Q., Hastriter, M. W., Whiting, M. F. & Dittmar, K. Fleas (Siphonaptera) are Cretaceous, and evolved with Theria. *Mol. Phylogenet. Evol.* **90**, 129–139 (2015).

Acknowledgements

We thank Xian-guo Guo and Cheng-fu Zhao for their help during specimen collection. In addition, we would like to thank the flea mitogenomes available in the NCBI database provided of Xing Yang, Dali University, Yunnan, China; Yu Zhang, Hunan Agricultural University, Hunan, China; Liang-fei Tan, Huazhong Agricultural University, Hubei, China; Fu-rong Wei, Chinese Center for Disease Control and Prevention, Beijing, China; Hai-Tao Xiang, Gansu Agricultural University, Gansu, China; Victoria I. Verhoeve, University of Maryland School of Medicine, Baltimore, MD, USA; Stephen L. Cameron, Queensland University of Technology, Brisbane, QLD, Australia; Alexandra Beliavskaia, University of Liverpool, UK. This work was supported by the National Natural Science Foundation of China [No.32260152 to Wenge Dong].

Author contributions

Wen-ge Dong and Xiao-xia Lin designed and conducted the research. Wen-ge Dong and Xiao-xia Lin contributed reagents and materials. Wen-ge Dong, Xiao-xia Lin and Ju Pu analyzed the data. Xiao-xia Lin and Wen-ge Dong wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-69203-y.

Correspondence and requests for materials should be addressed to W.D.