Pakistan J. Zool., pp 1-14, 2024. **DOI**: https://dx.doi.org/10.17582/journal.pjz/20230912053433

The Complete Mitochondrial Genome of *Sphaeniscus atilius* (Walker, 1849) (Diptera: Tephritidae) and Implication for the Phylogeny of Tephritidae

Shibao Guo[1], Junhua Chen[1], Nan Song[2], Fangmei Zhang[1]*

[1]*Xinyang Agriculture and Forestry University, Xinyang 464000, China*
[2]*Henan Agriculture University, Zhengzhou 450002, China*

Article Information
Received 12 September 2023
Revised 09 December 2023
Accepted 26 December 2023
Available online 04 April 2024
(early access)

Authors' Contribution
Data curation: SG, NS. Formal analysis: FZ, NS. Investigation: JC. Methodology: FZ, NS. Funding acquisition: SG. Supervision: JC. Writing-original draft: SG, FZ. Writing-review and editing: SG, FZ, NS.

Key words
Tephritidae, *Sphaeniscus atilius*, Mitochondrial genome, Phylogeny

ABSTRACT

The complete mitochondrial genome of *Sphaeniscus atilius* was characterized and annotated in this study. The mitogenome was 16,854 bp in length and encoded 37 typical mitochondrial genes, including 13 protein-coding genes, 22 tRNA genes, 2 ribosomal RNA genes, and 1 control regions. The total length of the 13 PCGs was 11,140 bp, and the AT content was 79.8%. There were five types of start codons, ATT (*nad2*, *nad3*, *nad5*, and *nad6*), ATG (*cox2*, *cox3*, *atp6, nad4*, *nad4l*, and *cob*), CGA (*cox1*), as well as ATC (*atp8*) and ATA (*nad1*). Most of the PCGs had typical TAA stop codons, except *nad5* which terminated with incomplete forms T-. Ile, Phe, Leu and Asn were the most frequently used amino acids in mitochondrial PCGs. Most tRNA genes could be folded into the typical cloverleaf structure, except *trnS1* and *trnT* which lacked the dihydrouridine (DHU) and TΨC arms, respectively. Phylogenetic analyses based on 13 protein-coding genes among the available sequenced species of family Tephritidae by maximum likelihood and bayesian inference methods suggested the genus relationship of Tephritidae: ((*Bactrocera*, *Dacus*, *Zeugodacus*), *Felderimyia*, *Anastrepha*), (*Acrotaeniostola*, (*Neoceratitis*, *Ceratitis*), *Euleia*, *Rivellia*), (*Procecidochares*, (*Tephritis*, *Sphaenisscus*))))). Our results presented the first mitogenome from *Sphaeniscus* and provide insights into the species identification, taxonomy and phylogeny of *S. atilius*.

Online First Article

INTRODUCTION

Tephritidae, one of the largest families of Diptera, consists of more than 500 genera and almost 5,000 named species and predominantly distributes throughout the temperate and tropical areas of the world (Pape *et al*., 2009; Aluja and Norrbom, 2000; Mazzon *et al*., 2021). This family is also referred to as true fruit flies, with hundreds of fruit-eating species accounting for about 40% of the species. It has been reported to attack a great variety of fruit plants, bamboo culms, vegetables, flowers, and seeds (Korneyev, 1999; Dohm *et al*., 2014). In practice, some of the fruit-eating species in the *Anastrepha*, *Bactrocera*, *Ceratitis*, *Dacus*, and *Rhagoletis* genera have been considered serious agricultural pests due to their significant economic impact on the production of fruit crops and stored fruit (White and Elson-Harris, 1992; Aluja and Mangan, 2008). Melon fly, *Zeugodacus cucurbitae* (Siderhurst and Jang, 2010), medfly, *Ceratitis fasciventris* (Drosopoulou *et al*., 2017), together with *Bactrocera latifrons* (Yong *et al*., 2016), are well-known examples.

The insect mitochondrial genome has been regarded as a useful molecular marker in studies of phylogenetic and evolutionary analysis, genetic diversity, and species delimitation at the genus or species level, due to its small size, high copy numbers, maternal inheritance, unambiguous orthologous genes, conserved gene composition, and high evolutionary rate (Cameron, 2014; Song *et al*., 2016; Wilson and Xu, 2012). Generally, the typical insect mitochondrial genome is a highly conserved circular molecule ranging in size from approximately 14 to 40 kbp, encoding a fixed set of 37 genes, including 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes, 2 ribosomal RNA (rRNA) genes, and a control region (CR) or the A+T-rich region (Du *et al*., 2021; Wolstenholme, 1992), with only a few exceptions. For example, long gene intergenic spacers, gene rearrangements, and gene loss have also been reported in

* Corresponding author: zhangfm@xyafu.edu.cn
0030-9923/2024/0001-0001 $ 9.00/0

Copyright 2024 by the authors. Licensee Zoological Society of Pakistan.
This article is an open access ∂ article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

different orders of insects (Du *et al*., 2017; Gong *et al*., 2018; Yan *et al*., 2022).

Partial mitochondrial gene sequences have become a preferred approach for inferring phylogenetic and molecular systematic studies in several insect groups, such as *Arma custos* and *Picromerus lewisi* (Hemiptera) (Mu *et al*., 2022), *Episymploce splendens* (Blattodea) (Yan *et al*., 2022), and *Haematopinus tuberculatus* (Psocodea) (Fu *et al*., 2022), *Coomaniella copipes*, *Coomaniella dentata*, and *Dicerca corrugate* (Coleoptera) (Huang *et al*., 2022), including the Tephritidae family (*Ceratitis fasciventris* (Drosopoulou *et al*., 2017), *Bactrocera carambolae* (Drosopoulou *et al*., 2019), *Bactrocera biguttula* (Teixeira *et al*., 2019), *Zeugodacus cucurbitae* (Zhou *et al*., 2020), and *Lepidotrigona flavibasis* (Wang *et al*., 2021). The aforementioned studies of mitochondrial gene sequences explored the origin and evolution of insects, explained the species and evolution of the system, and revealed the geographical distribution of intraspecific polymorphism.

Sphaeniscus atilius (Diptera: Tephritidae: *Sphaeniscus*), which is distributed from India to Russia, Korea, Japan, Australasian and Oceanian regions, can be morphologically distinguished from any other tephritid species based on clear diagnostic morphological features, including an almost entirely dark brown body, the number of orbital frontal setae, and an almost perpendicular base of the discal band of the wing (Han *et al*., 2010). To date, the complete mitochondrial genomes (mitogenomes) of 46 species, belonging to 14 genera of Tephritidae, are available in GenBank (https://www.ncbi.nlm.nih.gov/nucleotide/). All of these species belong to 5 subfamilies, Sciomyzidae, Sepsidae, Lauxaniidae, Celyphidae, Platystomatidae, and Tephritidae, respectively. However, there are no reports on the molecular phylogeny studies of the mitochondrial genome information in *S. atilius*, which limits our comprehensive understanding of the evolutionary and phylogenetic relationships of *S. atilius*.

In the current study, we sequenced, annotated, and described the complete mitogenome of *S. atilius* using next-generation sequencing, which is the first complete mitogenome sequence reported in the genus *Sphaeniscus*. We predicted and analyzed the gene organization, base composition, PCGs, codon usage, and the structure of the tRNAs and rRNAs of its mitochondrial genome. Additionally, we carried out phylogenetic analyses based on maximum likelihood (ML) and Bayesian Inference (BI) methods to assess the phylogenetic position of *S. atilius*. These results will be greatly helpful for clarifying the phylogenetic status and relationships between different species of Tephritidae.

MATERIALS AND METHODS

Taxon sampling and DNA extraction

Specimens of *S. atilius* were collected in Mount Jigong, Xinyang, Henan Province, China (31°48′43″N, 114°05′43″E), in June 2020. Specimens were preserved in 100% ethanol, and stored at -20°C. After morphological identification, total genomic DNA was extracted from muscle tissue of pre-thoraxes using Tissue DNA kit (TIANGEN Biotech, Beijing, China) according to the manufacture's protocol. The impurities and concentration were detected by agarose (1%) electrophoresis and Nanodrop spectrophotometer (ThermoFisher Scientific, Waltham, MA), respectively.

Sequencing and assembling of mitochondrial genome

Genome sequencing was performed on an Illumina HiSeq 2500 platform (150 bp paired-end reads). The TruSeq library was prepared with an insert size of 400 bp. Quantity of sequencing data for each sample was at least 20 Gb. Sequencing was performed at Beijing Novogene Bioinformatics Technology Co., Ltd, China. A total of 2 Gb raw paired reads were generated. Data filtering was conducted using NGS QC-Toolkit v2.5 (Patel and Jain, 2012). After removing the connector and the unmatched, short, and poor-quality reads, the high-quality reads (Q20 > 90% and Q30 > 80%) were used for genome assembly.

De novo assembly was performed with IDBA-UD v. 1.1.1 (Peng *et al*., 2012). The parameter settings were as follows: 200 minimum sizes of contig, 41 as the minimum k-mer size, 10 as an iteration size, and 91 as the maximum k-mer size. The pre sequenced mitochondrial *cox1* gene was used to bait the mitogenome from the assembled contigs.

Mitochondrial genome annotation and analysis

Preliminary annotation of mitogenome was conducted using MITOS (http://mitos.bioinf.uni-leipzig.de/index.py) (Bernt *et al*., 2013). The gene boundaries were refined by blasting against closely related species. The secondary structures of tRNA genes were predicted in MITOS. The structure map of the mitogenome was drawn using OGDRAW v1.3.1 (Greiner *et al*., 2019). The annotated mitogenome sequence of *S. atilius* was deposited at GenBank (accession number OQ909100).

The mitogenome nucleotide composition and relative synonymous codon usage (RSCU) were computed using MEGA 7.0 (Kumar *et al*., 2016). AT and GC skews were calculated following the formula: AT Skew = (A − T)/(A + T), GC Skew = (G − C)/(G + C) (Perna and Kocher, 1995). The nucleotide diversity (Pi) and nonsynonymous (Ka)/synonymous (Ks) mutation rate ratios were calculated by

Online First Article

DnaSP v5.10.01 (Librado and Rozas, 2009).

Phylogenetic analysis

The new mitogenome sequence of *S. atilius* was merged with the existing dipteran mitogenome sequences. The amino acid sequences of 13 protein coding genes of 38 dipteran species were aligned individually using MAFFT with default parameters (Katoh and Standley, 2013). We chose Pachycerina (*Pachycerina decemlineata*) and Sciomyza (*Sciomyza simplex*) as outgroups. The alignments were trimmed by trimAl (Capella-Gutierrez *et al*., 2009). The resulting alignments were concatenated using SequenceMatrix 1.8 (Vaidya *et al*., 2011).

The phylogenetic analysis was conducted using ML and BI methods. ML tree was reconstructed using IQ-TREE v2.2.3 (Minh *et al*., 2020). The best-fitting model for the alignment was estimated by ModelFinder (Kalyaanamoorthy *et al*., 2017). Branch support was assessed by 10,000 ultrafast bootstrap replicates (Hoang *et al*., 2018). MrBayes 3.1 was used to reconstruct the BI tree, and four independent Markov chain runs were performed for 1,000,000 metropolis–coupled (MCMC) generations, sampling a tree every 100 generations.

RESULTS

Genome organization

The length of the *S. atilius* mitogenome was 16,854 bp in length, which consisted of the typical 37 genes including PCGs, 22 tRNA genes, and two rRNA genes (Fig. 1, Table I). In addition, a major non-coding region known as the control region (CR) or A+T-rich region was found between *rrnS* and *trnI*. The heavy chain (H-strand) encoded 23 genes (nine PCGs and 14 tRNAs). The remaining 4 PCGs, 8 tRNAs and 2 rRNAs were transcribed in the light chain (L-strand).

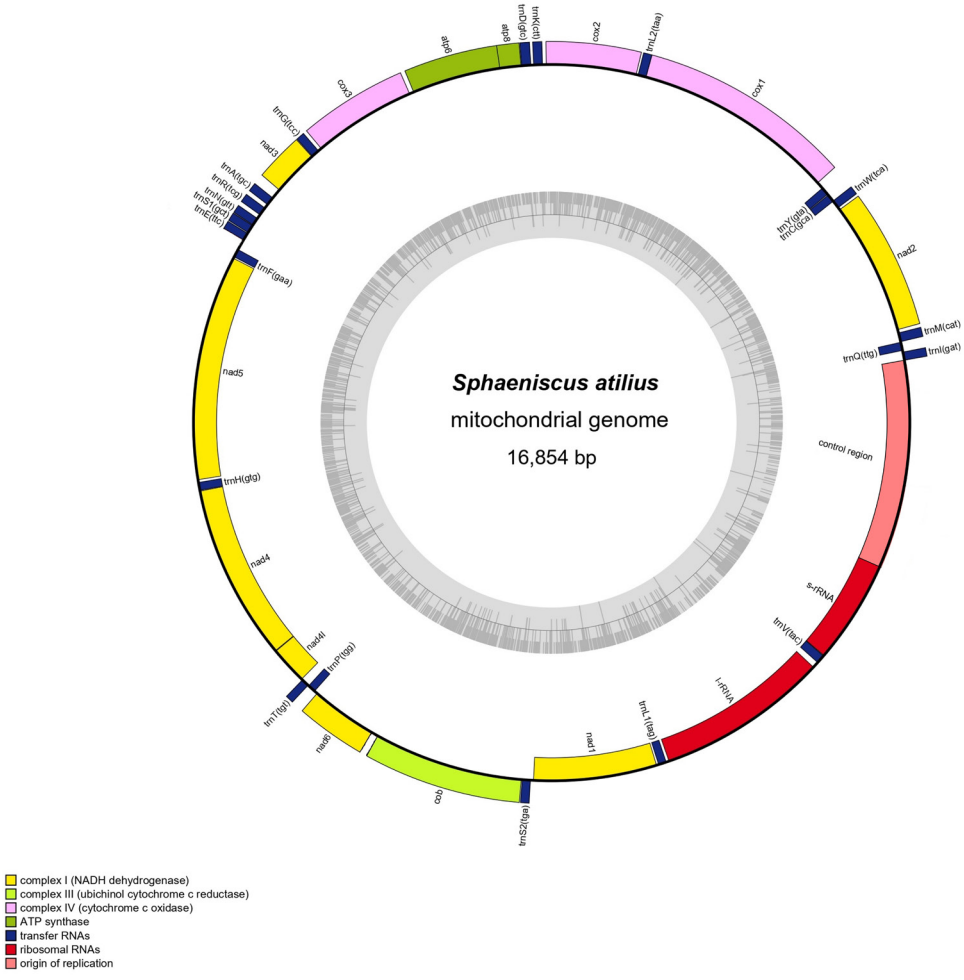

Fig. 1. The mitochondrial genome of *S. atilius* showing the protein coding genes, tRNAs, rRNAs and non-coding regions.

Table I. Summary of the mitogenome of *S. atilius.*

Gene	Strand	Location	Size (bp)	Anti codon	Start codon	Stop codon	Intergenic nucleotides
trnI	H	1-67	67	GAT			7
trnQ	L	67-135	69	TTG			-1
trnM	H	157-228	72	CAT			21
nad2	H	243-1250	1008		ATT	TAA	14
trnW	H	1242-1309	68	TCA			-8
trnC	L	1311-1376	66	GCA			1
trnY	L	1377-1443	67	GTA			47
cox1	H	1491-3026	1536		CGA	TAA	-5
trnL2	H	3021-3086	66	TAA			10
cox2	H	3110-3793	684		ATG	TAA	23
trnK	H	3810-3880	71	CTT			16
trnD	H	3881-3954	74	GTC			0
atp8	H	3947-4111	165		ATC	TAA	-7
atp6	H	4240-4917	678		ATG	TAA	28
cox3	H	4940-5725	786		ATG	TAA	22
trnG	H	5726-5791	66	TCC			0
nad3	H	5864-6217	354		ATT	TAA	72
trnA	H	6245-6313	69	TGC			27
trnR	H	6342-6408	68	TCG			28
trnN	H	6409-6471	65	GTT			0
trnS1	H	6474-6540	67	GCT			2
trnE	H	6532-6997	66	TTC			-8
trnF	L	7058-7125	68	GAA			60
nad5	L	7141-8854	1714		ATT	T(AA)	15
trnH	L	8855-8920	66	GTG			0
nad4	L	8919-10259	1341		ATG	TAA	-1
nad4L	L	10262-10552	291		ATG	TAA	2
trnT	H	10553-10616	64	TGT			0
trnP	L	10637-10702	66	TGG			20
nad6	H	10700-11206	507		ATT	TAA	-2
cob	H	11265-11401	1137		ATG	TAA	58
trnS2	H	11419-11485	67	TGA			17
nad1	L	11496-12423	939		ATA	TAA	10
trnL1	L	12424-12505	66	TAG			16
16sRNA	L	12547-13855	1309				41
trnV	L	13854-13925	72	TAC			-1
12sRNA	L	13913-14711	799				-12
CR		14712-16854	2143				0

Note: Strand of the genes is presented as L for majority and H for minority strand. In the column for intergenic length, a positive sign indicates the interval in base pairs between genes, while the negative sign indicates overlapping base pairs between genes.

Online First Article

Table II. Composition and skewness of the *S. atilius* mitogenome.

S. atilius	Size (bp)	A (%)	T (%)	G (%)	C (%)	A+T (%)	G+C (%)	AT-skew	GC-skew
Genome	16,854	41.80	39.92	7.67	10.61	81.72	18.28	0.02	-0.16
Protein-coding genes	11,140	34.15	45.67	10.30	9.88	79.82	20.18	-0.14	0.02
tRNA genes	1,490	41.41	39.06	10.94	8.59	79.47	19.53	0.03	0.12
rRNA genes	2108	39.28	42.88	11.48	6.36	82.16	17.84	-0.04	0.28
Detected CR	1674	45.65	42.60	1.61	7.46	88.25	9.07	0.03	0.64

CR, control region

There were five intergenic overlapping regions totaling 29 bp, with varying lengths of 1-8 bp, which ware mainly present in the tRNA genes. The two longest overlapping regions, both with a length of 8 bp, occured between *nad2* and *trnW* and between *trnS1* and *trnE*. Sixteen intergenic spacer regions were identified, totaling 555 bp in length, with the longest spacer sequence (72 bp) located between *trnG* and *nad3*, followed by a 60 bp spacer between *trnE* and *trnF*. There were also three regions without gene overlaps or intergenic spacers.

The mitogenome nucleotide composition was 41.80% for A, 39.92% for T, 7.67% for G, and 10.61% for C, respectively. The mitogenome was significantly biased to A+T (81.72%). The whole mitogenome of *S. atilius* exhibited a positive AT skew (0.02) and a negative GC skew (-0.16) (Table II).

Protein-coding genes

The total length of the 13 PGGs of in the *S. a*tilius mitogenome was 11,140 bp, accounting for 66.10% of the whole mitogenome sequence, and encoding a total of 3,612 codons. Among them, *nad5* (1,714 bp) was found to be the longest sequence, and *nad4L* (291 bp) was the shortest (Table I). Nine PCGs (*nad2*, *cox1*, *cox2*, *atp8*, *atp6*, *cox3*, *nad3*, *nad6*, and *cob*) were coded on the H-strand, while the remaining four PCGs (*nad5*, *nad4*, *nad4L*, and *nad1*) were located on the L-strand. The content of AT and GC was 79.82% and 20.18% in the 13 PCGs, exhibiting a highly AT bias (Table II). The AT skew was negative (-0.14), while the GC skew was positive (0.02) in the PCGs. Additionally, six PCGs (*cox2*, *cox3*, *atp6*, *nad4*, *nad4L*, and *cob*) initiated with an ATG start codons, four PCGs (*nad2*, *nad3*, *nad5* and *nad6*) used ATT as the start codon, nad1 used ATA, *atp8* used ATC, and *cox1* used CGA. The termination codons of 12 PCGs were TAA. Only the *nad5* used an incomplete stop codon T.

The amino acid usage of the 13 PCGs and the relative synonymous codon usage (RSCU) frequency are shown in Figure 2 and Table III. The most frequently used amino acids in mitochondrial PCGs were Ile, Phe, Leu, and Asn, accounting for 16.68%, 12.51%, 10.86%, and 8.13% of the total amino acids, respectively (Fig. 2A). The most frequently used codons were UUU (436), AUU (436), UUA (342), and AAU (277). The relatively scarce amino acids were Met (0.62%), Trp (0.67%), Arg1 (0.93%), and Asp (1.23%). The most frequent synonymous codons were UUA, AGA, and GUU, with UUA having the highest frequency of relative synonymous codons (RSCU = 4.86) (Fig. 2B).

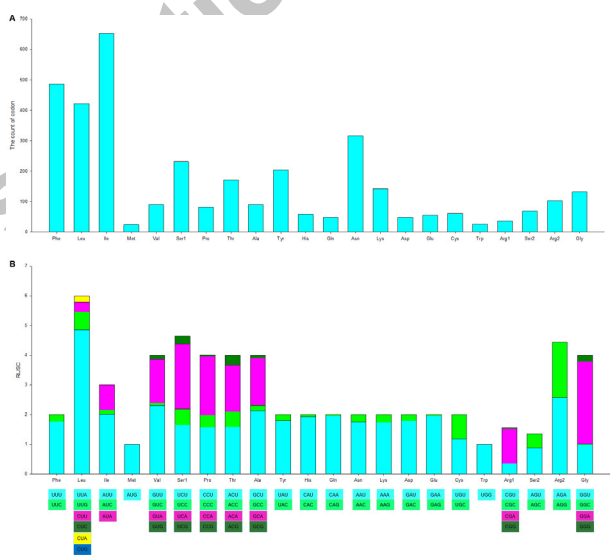

Fig. 2. Usage of amino acids of 13 PCGs in the *S. atilius* mitogenome. (A) The total numbers of codon families; (B) Relative synonymous codon usage (RSCU) of codon families.

The nucleotide diversity (Pi) of the PCGs among 46 species was calculated, ranging from 0.14 to 0.22 (Fig. 3). Among them, *nad6* (Pi=0.22) showed the most diverse nucleotide variability among all PCGs, followed by *nad2* (Pi=0.21), *nad3* (Pi=0.19), and *atp8* (Pi=0.18). The *nad5* (Pi=0.14), *nad1* (Pi=0.15), and *cox3* (0.15) genes exhibited relatively low values of nucleotide variability. The ratio of Ka/Ks was calculated for each gene of the 13 PCGs (Fig. 3). The value of the *cox3* gene (Ka/Ks=1.26) was higher than others. Meanwhile, the ratio of Ka/Ks of other 12 PCGs were all significantly less than 1, with the value of *cox1* gene being the lowest (Ka/Ks=0.06).

Online First Article

Table III. Codon number and RSCU in *S. atilius* mitochondrial PCGs.

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	431	1.77	UCU(S)	82	1.65	UAU(Y)	183	1.8	UGU(C)	36	1.18
UUC(F)	55	0.23	UCC(S)	27	0.54	UAC(Y)	20	0.20	UGC(C)	25	0.82
UUA(L)	342	4.86	UCA(S)	109	2.19	UAA(*)	60	1.05	UGA(*)	93	1.62
UUG(L)	42	0.60	UCG(S)	13	0.26	UAG(*)	19	0.33	UGG(W)	26	1.00
CUU(L)	23	0.33	CCU(P)	32	1.58	CAU(H)	56	1.93	CGU(R)	8	0.35
CUC(L)	0	0	CCC(P)	8	0.40	CAC(H)	2	0.07	CGC(R)	0	0
CUA(L)	15	0.21	CCA(P)	40	1.98	CAA(Q)	47	1.96	CGA(R)	27	1.17
CUG(L)	0	0	CCG(P)	1	0.05	CAG(Q)	1	0.04	CGG(R)	1	0.04
AUU(I)	436	2.01	ACU(T)	68	1.59	AAU(N)	277	1.75	AGU(S)	44	0.88
AUC(I)	33	0.15	ACC(T)	22	0.51	AAC(N)	39	0.25	AGC(S)	24	0.48
AUA(I)	183	0.84	ACA(T)	67	1.57	AAA(K)	123	1.73	AGA(R)	59	2.57
AUG(M)	24	1.00	ACG(T)	14	0.33	AAG(K)	19	0.27	AGG(R)	43	1.87
GUU(V)	52	2.31	GCU(A)	48	2.13	GAU(D)	43	1.79	GGU(G)	32	0.98
GUC(V)	2	0.09	GCC(A)	4	0.18	GAC(D)	5	0.21	GGC(G)	1	0.03
GUA(V)	33	1.47	GCA(A)	36	1.60	GAA(E)	54	1.96	GGA(G)	91	2.78
GUG(V)	3	0.13	GCG(A)	2	0.09	GAG(E)	1	0.04	GGG(G)	7	0.21

tRNAs and rRNAs

The 22 tRNA genes of the mitogenome of *S. atilius* had a total of 1,490 bp in length, 9.51% of the entire mitogenome, ranging from 65 bp (*trnN*) to 74 bp (*trnD*) (Table I). Among these, 14 genes (*trn1, trnM, trnW, trnL2, trnK, trnD, trnG, trnA, trnR, trnN, trnS1, trnE, trnT*, and *trnS2*) were located on the H-strand and the remaining eight genes (*trnQ*, *trnC*, *trnY*, *trnF*, *trnH*, *trnP*, *trnL1*, and *trnV*) were located on the L-strand. Through the analysis of the secondary structure of the tRNAs (Fig. 4), most tRNA genes could be folded into the typical cloverleaf secondary structure, while *trnS1* and *trnT* lacked the dihydrouridine (DHU) and TΨC arms, respectively. In the secondary structures of tRNAs of *S. atilius* (Fig. 3), three or four base pairs in the DHU arms, and four or five base pairs in the TΨC arms. Except the classic base pairs (A-U and C-G), fourteen wobble base pairs (G-U) were detected in nine genes (*trnA*, *trnC*, *trnF*, *trnG*, *trnH*, *trnP*, *trnQ*, *trnT*, and *trnV*), which occured in the amino acid-accepting arms, anticodon arms, TψC arms or DHU arms. Of them, *trnH* had the highest rate (three pairs each). Besides, five pairs of U-U base mismatches in *trnA*, *trnG*, *trnL1*, *trnR*, *trnV*, and *trnW* and one mismatches base in *trnN* were found in the TψC arms. The AT and GC content were 79.47% and 19.53% in the 22 tRNA genes, respectively, with a positive AT skew (0.03) and GC skew (0.12) (Table II).

There were two rRNAs in the mitogenome of *S. atilius*: a 1,309 bp *16S rRNA* (*rrnL*) and a 799 bp *12S rRNA* (*rrnS*) (Table I). The *16S rRNA* gene was located between *trnL1* and *trnV*, while the *12S rRNA* gene was located between *trnV* and the control region. Both *16S rRNA* and *12S rRNA* were embedded in the L-strand. The AT content was 82.16%, with a negative AT skew (-0.04) and positive GC skew (0.28) (Table II).

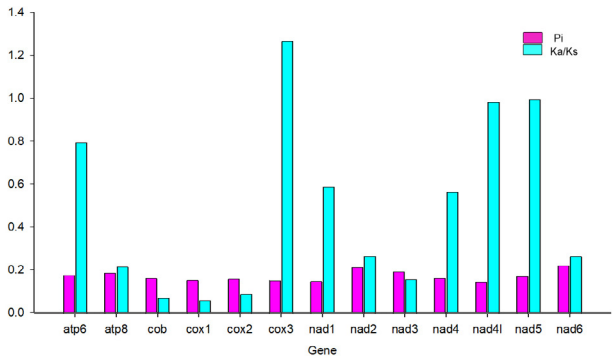

Fig. 3. The nucleotide diversity (Pi) and nonsynonymous (Ka) /synonymous (Ks) substitution rate ratios of 13 PCGs of Tephritidae species.

Control region

The mitogenome of *S. atilius* contained one large putative control region (CR), and the length of the CR was 2143 bp. It was located between the *rrnS* and *trnI*, *trnE* and *trnF,* nad6 and *cob*, and *rrnS* and *trnI*, respectively. The AT content of non-coding regions (88.25%) was obviously higher than other regions. Additionally, the positive AT skew (0.03) and GC skew (0.64) were detected in the control region (Table II).

Online First Article

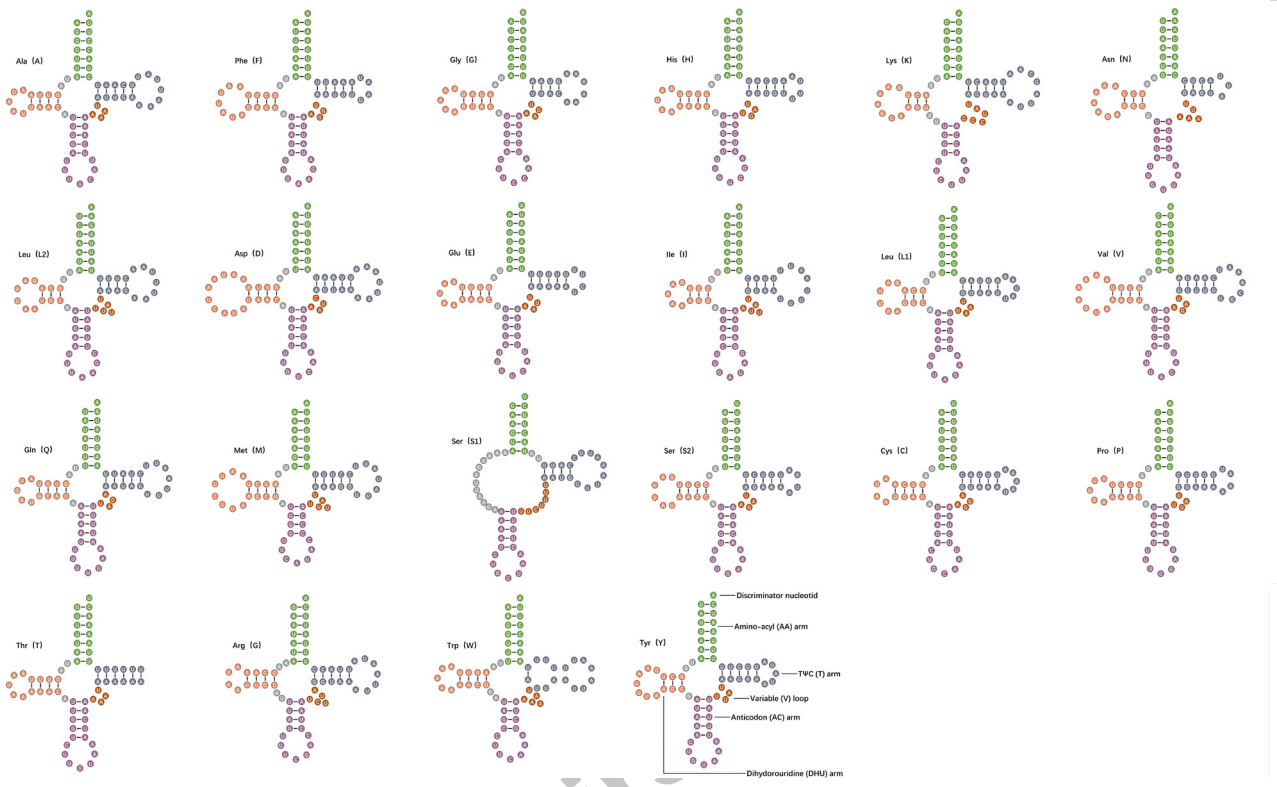

Fig. 4. Secondary structures of the transfer RNA genes (tRNAs) in *S. atilius* mitogenome.

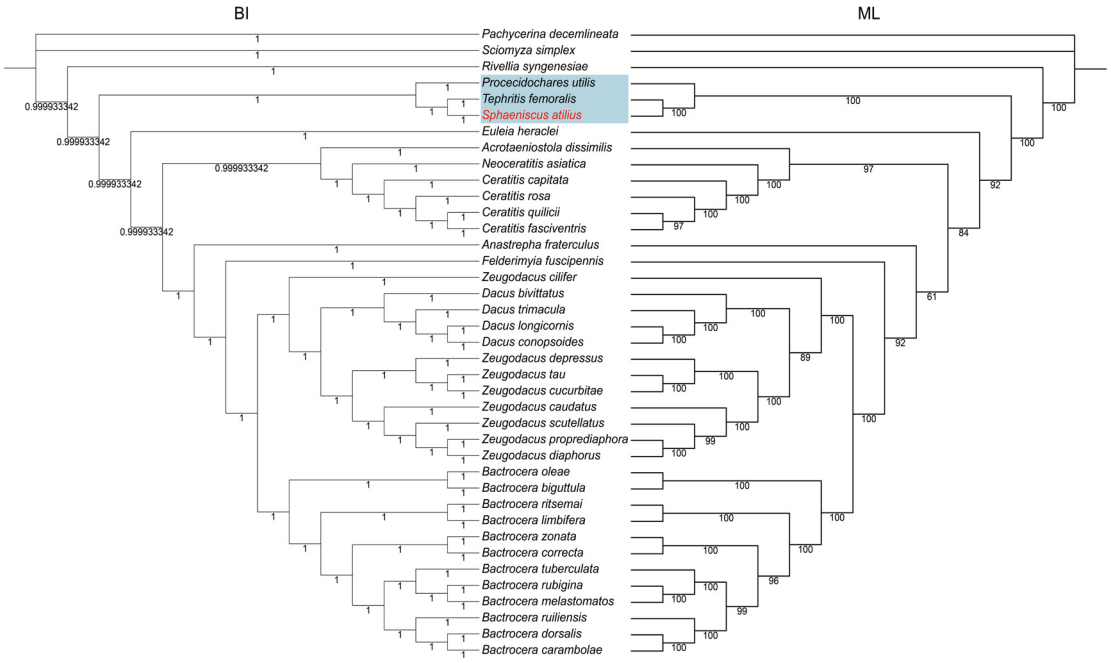

Fig. 5. Phylogenetic reconstruction of 39 PCGs sequences was performed using the maximum likelihood and bayesian inference methods analysis. The species sequenced in this study are indicated with red. Pachycerina (*Pachycerina Decemlineata*) and Sciomyza (*Sciomyza simplex*) were used as outgroups.

Phylogenetic analysis

The same phylogenetic relationships of *S. atilius* and other Tephritidae species were produced by ML and BI methods. The result revealed that the thirteen genera of *Tephritidae* species followed the following monophyletic relationships: ((*Bactrocera*, *Dacus*, *Zeugodacus*), *Felderimyia*, *Anastrepha*), (*Acrotaeniostola*, (*Neoceratitis*, *Ceratitis*), *Euleia*, *Rivellia*), (*Procecidochares*, (*Tephritis*, *Sphaenisscus*))))). Of them, *Bactrocera* (12 exemplars) formed a separate clade at the top of phylogenetic tree, and formed the sister group of a clade including *Dacus* (4 exemplars), *Zeugodacus* (8 exemplars), *Felderimyia* and *Anastrepha*. *Ceratitis* (4 exemplars) formed a monophyletic group and was sister to a clade comprising *Acrotaeniostola* and *Neoceratit*. *Euleia* and *Rivellia* formed the separate clades, respectively. *Sphaenisscus* and *Tephritis* clustered together showing a high statistical support value (PP=1, BS=100), and formed a sister group to *Procecidochares* (1 exemplar). *S. atilius* and *T. femoralis* were closely related and formed a sister group to *Procecidochares utilis*.

DISCUSSION

The complete mitogenome of *S. atilius* was a circular, double-stranded DNA molecule with a total length of 16,854 bp, which was similar to that of other Tephritoidea insects analyzed, ranging from 15,117 bp (*Tephritis femoralis*) to 16,739 bp (*Anastrepha fraterculus*) (Table IV). It has the typical organization and composition of an insect mitochondrion, including 13 PCGs, 22 tRNA genes, and 2 rRNA genes, which was consistent with the existing mitogenomes of Tephritoidea, such as *Bactrocera carambolae* (Drosopoulou *et al*., 2019), *Zeugodacus cucurbitae* (Zhou *et al*., 2020), *Dacus haikouensis* (Wang *et al*., 2022). The nucleotide composition of all regions showed a strong AT bias (Nguyen *et al*., 2020), as seen in other insects (Yan *et al*., 2022; Mu *et al*., 2022; Lv *et al*., 2021). The AT-skew of the entire mitogenome was positive (0.02), while the GC skew is negative (-0.16), indicating that the content of bases C was higher than that of G, and A was higher than T in the whole genome.

The mitogenome commonly exhibited compact arrangement, such as small intergenic spacers or overlapping genes (Ojala *et al*., 1981). In the present study, intergenic overlapping regions ranging from 1 to 8 bp, with a total length of 29 bp. Overlapping regions of similar size were common among Tephritoidea insects (Drosopoulou *et al*., 2017, 2019; Yong *et al*., 2016; Teixeira *et al*., 2019), while their positions varied across species. For example, the longest overlaps were found between *atp8* and *atp6* in *B. carambolae* (Drosopoulou *et al*., 2019), and between *nad2* and *trnW*, and between *trnS1* and *trnE* in *S. atilius*. Most of the gene overlaps occurred in tRNA genes due to the lower evolutionary constraints of these genes (Yuan *et al*., 2021).

The non-coding intergenic spacers, which were composed of less than 10 non-coding nucleotides in the mitochondria of most animals, which contributed to species identification and the evolution of insect mitochondrial genomes (Yan *et al*., 2022). Sixteen intergenic spacer regions were examined, with a total length of 555 bp, and the longest spacer sequence (72 bp) was located between *trnG* and *nad3*. Larger intergenic spacers in mitogenomes had been reported in Tephritoidea insects, such as 94 bp in *Bactrocera latifrons*, 82 bp in *Bactrocera melastomatos* and 79 bp in *Bactrocera umbrosa* (Yong *et al*., 2016). Generally, the duplication/random loss model and slipped-strand mispairing can be used to explain the origin of mitogenome intergenic spacers (Du *et al*., 2017; Cheng *et al*., 2016). Whether these long spacer regions being functional was controversial (Yan *et al*., 2022).

The 13 PCGs of the *S. atilius* mitogenome were found to be 11,140 bp in length and used a variety of start codons, including ATG for *cox2*, *cox3*, *atp6*, *nad4*, *nad4L*, and *cob*; ATT for *nad2*, *nad3*, *nad5*, and *nad6*; ATA for *nad1*; CGA for *cox1*; and ATC for *atp8*, as reported for *Bactrocera biguttula* (Teixeira *et al*., 2019) and *Bactrocera carambolae* PCGs (Drosopoulou *et al*., 2019). Alternative start codons had also been found in other insects, such as TTG in *Arma custos* and *Picromerus lewisi* (Mu *et al*., 2022) and GTG in *Nisia fuliginosa* (Lv *et al*., 2021). The *cox1* gene in the *S. atilius* mitogenome used CGA as the starting codon, consistent with other known insects (Yang *et al*., 2019). However, the starting codon of *cox1* was not always uniform, for example, TTG in *Episymploce splendens* (Yan *et al*., 2022) and ATA in *Anastatus fulloi* (Yi *et al*., 2022). The typical termination codons TAA was employed in 12 PCGs, which was common among metazoans (Yan *et al*., 2022; Yi *et al*., 2022), with one exception of an incomplete stop codon T for *nad5*. This exception was commonly observed in arthropod mitogenomes (Huang *et al*., 2022; Yi *et al*., 2022), and might be attributed to post-transcriptional modification during the mRNA maturation process (Boore, 1999; Lv *et al*., 2021).

The four most frequently used codons, UUU (Phe), AUU (Ile), UUA (Leu), and AAU (Asn) were observed in the *S. atilius* mitogenome, which was similar to other insect mitogenomes, such as those of Ephemeroptera (Li *et al*., 2021), Coleoptera (Zeng *et al*., 2021) and Hemiptera (Nguyen *et al*., 2020). Meanwhile, the RSCU analysis of the PCGs also indicated that A and U were the components that contributed to the high A+T bias of the full mitogenome. The nucleotide diversity (Pi) and the ratio of *Ka/Ks* of the PCGs among 43 Tephritoidea species

Online First Article

Table IV. List of taxa used for phylogenetic analysis.

Superfamily	Family	Genus	Species	GenBank accession	Length (bp)
Tephritoidea	Tephritidae	*Acrotaeniostola*	*Acrotaeniostola dissimilis*	MH900079	15,384
		Anastrepha	*Anastrepha fraterculus*	KX926433	16,739
			Bactrocera biguttula	MK293875	15,829
			Bactrocera carambolae	EF014414	15,915
			Bactrocera correcta	JX456552	15,936
			Bactrocera dorsalis	DQ845759	15,915
			Bactrocera limbifera	MG566056	15,860
			Bactrocera melastomatos	KT881557	15,945
			Bactrocera oleae	GU108463	15,821
			Bactrocera ritsemai	MF668132	15,927
			Bactrocera rubigina	MN714223	15,285
			Bactrocera ruiliensis	MN477221	15,870
			Bactrocera tuberculata	MT196006	15,273
			Bactrocera zonata	KP296150	15,935
		Ceratitis	*Ceratitis capitata*	AJ242872	15,980
			Ceratitis fasciventris	KY436396	16,017
			Ceratitis quilicii	MT998948	16,035
			Ceratitis rosa	MT997010	16,047
		Dacus	*Dacus bivittatus*	MG962404	15,833
			Dacus conopsoides	MH351199	15,852
			Dacus longicornis	NC_032690	16,253
			Dacus trimacula	MK940811	15,851
		Euleia	*Euleia heraclei*	MT410819	15,514
		Felderimyia	*Felderimyia fuscipennis*	MT702879	16.536
		Neoceratitis	*Neoceratitis asiatica*	MF434829	15,481
		Procecidochares	*Procecidochares utilis*	KC355248	15,922
		Sphaeniscus	*Sphaeniscus atilius*	OQ909100	16,285
		Zeugodacus	*Zeugodacus caudatus*	KT625491	15,866
			Zeugodacus cilifer	MT702880	15,843
			Zeugodacus cucurbitae	JN635562	15,825
			Zeugodacus depressus	KY131831	15,832
			Zeugodacus diaphorus	KT159730	15,890
			Zeugodacus propredіaphora	MN688227	15,829
			Zeugodacus scutellatus	KP722192	15,915
			Zeugodacus tau	KP711431	15,687
	Platystomatidae	*Rivellia*	*Rivellia syngenesiae*	MT410799	17,835
Lauxanioidea	Lauxaniidae	*Pachycerina*	*Pachycerina decemlineata*	NC_034923	16,286
Sciomyzoidea	Sciomyzidae	*Sciomyza*	*Sciomyza simplex*	MT410781	16,553

Online First Article

were calculated. The results showed that *nad6* exhibited the most diverse nucleotide variability among all PCGs, while *nad5*, *nad1*, and *cox3* exhibited a relatively low variation rate and were the most conserved genes. The overall ratios of *Ka/Ks* for most PCG genes were significantly less than 1, which suggests that these PCGs were under purifying selection. The *cox1* gene had the lowest *Ka/Ks* ratio (*Ka/Ks*=0.06), indicating that this gene had a relatively slow evolutionary rate (Hurst, 2002). This phenomenon occurred in almost all animals (Xiao *et al*., 2019), and had been subjected to species identification and evolutionary analysis in various arthropod species in Tephritoidea, as well as like in other insects (Demari-Silva *et al*., 2015; Zhou *et al*., 2020).

As in other insects, most tRNA genes in *S. atilius* could be folded into the typical clover-leaf secondary structure, while *trnS1* lacked the dihydrouridine (DHU) arm and *trnT* lacked the TΨC arm. This feature of *trnS1* had been observed in many other insect mitogenomes (Huang *et al*., 2022; Lv *et al*., 2021; Yuan *et al*., 2021; Yi *et al*., 2022; Soumia *et al*., 2022). However, the feature of *trnT* was less found in other insects. Other tRNAs, *trnA* also lacked the TΨC arm as described in *Chrysodeixis acuta* (Soumia et al., 2022). So, we speculated that this phenomenon could cause by the loss of gene in the process of evolution. We have added reason in the discussion. In addition, 14 wobble base pairs (G-U) and 5 pairs of U-U base mismatches in the tRNA genes of the *S. atilius* mitogenome were observed. Previous reports had suggested that wobble and mismatched pairs, which commonly occurred in insect tRNAs, were usually corrected through the editing process and sustain the transport function (Varani and McClain, 2000; Lavrov *et al*., 2000). The length, location, and base composition of the two rRNA genes were similar to those of other Tephritoidea insects, such as *Bactrocera arecae* (Yong et al., 2015) and *Bactrocera biguttula* (Teixeira *et al*., 2019).

In addition, the *S. atilius* mitogenome contained one large non-coding regions of which, the one located between *rrnS* and *trnI* was supposed to act as the origin of genome replication and gene transcription (Wolstenholme, 1992; Boore, 1999). The control region of insect mitogenomes ranged in size from tens of to several thousands of base pairs (Zhang *et al*., 1995; Lewis *et al*., 1995; Inohira *et al*., 1997). The control region was a source of length variation in the mitogenomes (Zhang *et al*., 1995; Lewis *et al*., 1995; Inohira *et al*., 1997). In the *S. atilius* mitogenome, 2143 bp in length was sequenced, which was more than the longest control region of 1,141 bp in Tephritidae insects (Mu *et al*., 2022). The control region was located between *rrnS* and *trnI* genes and also had a higher AT content (88.25%), compared to other Tephritidae insects, such as *Bactrocera arecae* (86.0%) (Yong *et al*., 2015), *Bactrocera melastomatos* (89.0%) (Yong *et al*., 2016), and *Ceratitis fasciventris* (90.24%) (Drosopoulou *et al*., 2017).

Phylogenetic analyses were based on the nucleotide sequences of the 13 PCGs used ML and BI methods to construct phylogenetic trees from the mitogenomes of 38 species of Diptera to elaborate phylogenetic relationship. The same phylogenetic relationships of *S. atilius* and other Tephritidae species were produced by two methods. In the present study, the phylogenetic relationships within Tephritidae could be presented as follows: ((*Bactrocera*, *Dacus*, *Zeugodacus*), *Felderimyia*, *Anastrepha*), (*Acrotaeniostola*, (*Neoceratitis*, *Ceratitis*), *Euleia*, *Rivellia*), (*Procecidochares*, (*Tephritis*, *Sphaenisscus*))))), in concordance with the findings of previous phylogenetic studies (Drosopoulou *et al*., 2019; Teixeira *et al*., 2019; Jia *et al*., 2019; Yang *et al*., 2020). The results showed that four genera (*Felderimyia*, *Zeugodacus*, *Dacus* and *Bactrocera*) were closely related and formed a sister group. *Dacus* and *Zeugodacus* constituted a sister to *Bactrocera*. This pattern had been demonstrated in the previous studies (Krosch *et al*., 2012; Virgilio *et al*., 2015; Jiang *et al*., 2016; San *et al*., 2018), which further be verified by our results. Besides, *S. atilius* and *T. femoralis* were closely related, and both of them formed the sister group of *P. utilis*. In this study, the phylogenetic placement of *S. atilius* was firstly investigated. Although the phylogenetic analyses on *S. atilius* were still limited, this result contributed to a molecular basis for the classification and phylogeny of *S. atilius* within Tephritidae.

Online First Article

CONCLUSION

In conclusion, the mitogenome of *S. atilius*, was the first one from the genus *Sphaenisscus*. This mitogenome showed high conservation in terms of gene size, organization, AT bias and secondary structures of tRNAs. The phylogenetic placement *S. atilius* was investigated and clarified, revealing that *S. atilius* and *T. femoralis* were closely related, and both of them formed the sister group of *P. utilis*. These results provided a framework for further studies of the phylogenetics and evolution of *S. atilius*.

DECLARATIONS

Funding

This study was supported by Special funds for Henan Provinces Scientific and Technological Development Guided by the Central Government (Z20221341063), Natural Science Foundation of Henan Province (No. 212300410229), Key Project for University Excellent Young Talents of Henan Province (No. 2020GGJS260),

the Project of Science and Technology Innovation Team (No. XNKJTD-007 and KJCXTD-202001). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Statement of conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

Aluja, M. and Mangan, R.L., 2008. Fruit fly (Diptera: Tephritidae) host status determination: Critical conceptual, methodological, and regulatory considerations. *Annu. Rev. Ent.*, **53**: 473-502. https://doi.org/10.1146/annurev.ento.53.103106.093350

Aluja, M. and Norrbom, A.L., 2000. *Fruit flies (Tephritidae): phylogeny and evolution of behavior*. CRC Press, Boca Raton. https://doi.org/10.1201/9781420074468

Bernt, M., Donath, A., Juhling, F., Externbrink, F., Florentz, C., Fritzsch, G., Putz, J., Middendorf, M. and Stadler, P.F., 2013. MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.*, **69**: 313-319. https://doi.org/10.1016/j.ympev.2012.08.023

Boore, J.L., 1999. Animal mitochondrial genomes. *Nucl. Acids Res.*, **27**: 1767-1780. https://doi.org/10.1093/nar/27.8.1767

Cameron, S.L., 2014. Insect mitochondrial genomics: Implications for evolution and phylogeny. *Annu. Rev. Ent.*, **59**: 95-117. https://doi.org/10.1146/annurev-ento-011613-162007

Capella-Gutierrez, S., Silla-Martinez, J.M. and Gabaldon, T., 2009. TrimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, **25**: 1972-1973. https://doi.org/10.1093/bioinformatics/btp348

Cheng, X.F., Zhang, L.P., Yu, D.N., Storey, K.B. and Zhang, J.Y., 2016. The complete mitochondrial genomes of four cockroaches (Insecta: Blattodea) and phylogenetic analyses within cockroaches. *Gene*, **586**: 115-122. https://doi.org/10.1016/j.gene.2016.03.057

Demari-Silva, B., Foster, P.G., de Oliveira, T.M., Bergo, E.S., Sanabani, S.S., Pessôa, R. and Sallum, M.A.M., 2015. Mitochondrial genomes and comparative analyses of *Culex camposi*, *Culex coronator*, *Culex usquatus* and *Culex usquatissimus* (Diptera:Culicidae), members of the coronator group. *BMC Genomics*, **16**: 831. https://doi.org/10.1186/s12864-015-1951-0

Dohm, P., Kovac, D., Freidberg, A., Rull, J. and Aluja, M., 2014. Basic biology and host use patterns of tephritid flies (Phytalmiinae: Acanthonevrini, Dacinae: Gastrozonini) breeding in bamboo (Poaceae: Bambusoidea). *Ann. entomol. Soc. Am.*, **107**: 184-203. https://doi.org/10.1603/AN13083

Drosopoulou, E., Pantelidou, C., Gariou-Papalexiou, A., Augustinos, A.A., Chartomatsidou, T., Kyritsis, G.A., Bourtzis, K., Mavragani-Tsipidou, P. and Zacharopoulou, A., 2017. The chromosomes and the mitogenome of *Ceratitis fasciventris* (Diptera: Tephritidae): Two genetic approaches towards the Ceratitis FAR species complex resolution. *Sci. Rep.*, **7**: 4877. https://doi.org/10.1038/s41598-017-05132-3

Drosopoulou, E., Syllas, A., Goutakoli, P., Zisiadis, G.A., Konstantinou, T., Pangea, D., Sentis, G., Sauers-Muller, A., Wee, S.L., Augustinos, A.A., Zacharopoulou, A. and Bourtzis, K., 2019. Tauhe complete mitochondrial genome of *Bactrocera carambolae* (Diptera: Tephritidae): Genome description and phylogenetic implications. *Insects*, **10**: 429. https://doi.org/10.3390/insects10120429

Du, C., Zhang, L., Lu, T., Ma, J., Zeng, C.,Yue, B.S. and Zhang, X.Y., 2017. Mitochondrial genomes of blister beetles (Coleoptera, Meloidae) and two large intergenic spacers in *Hycleus* genera. *BMC Genomics*, **18**: 698. https://doi.org/10.1186/s12864-017-4102-y

Du, Z., Wu, Y., Chen, Z., Cao, L., Ishikawa, T., Kamitani, S., Sota, T.J., Song, F., Tian, L., Cai, W.Z. and Li, H., 2021. Global phylogeography and invasion history of the spotted lanternfly revealed by mitochondrial phylogenomics. *Evol. Appl.*, **14**: 915-930. https://doi.org/10.1111/eva.13170

Fu, Y.T., Suleman, Yao, C.Q., Wang H.M., Wang W. and Liu, G.H., 2022. A novel mitochondrial genome fragmentation pattern in the buffalo louse *Haematopinus tuberculatus* (Psocodea: Haematopinidae). *Int. J. mol. Sci.*, **23**: 13092. https://doi.org/10.3390/ijms232113092

Gong, R.Y., Guo, X., Ma, J.N., Song, X.H., Shen, Y.M., Geng, F.N., Price, M.G., Zhang, X.Y. and Yue, B.S., 2018. Complete mitochondrial genome of *Periplaneta brunnea* (Blattodea: Blattidae) and phylogenetic analyses within Blattodea. *J. Asia Pac. Ent.*, **21**: 885-895. https://doi.org/10.1016/j.aspen.2018.05.006

Greiner, S., Lehwark P. and Bock, R., 2019. Organellar genome DRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucl. Acids Res.*, **47**: W59-W64. https://doi.org/10.1093/nar/gkz238

Online First Article

Han, H.Y. and Kwon, Y.J., 2010. A list of North Korean tephritoid species (Diptera: Tephritoidea) deposited in the Hungarian natural history museum. *Anim. Syst. Evol. Diver.*, **26**: 251-260. https://doi.org/10.5635/KJSZ.2010.26.3.251

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q. and Vinh, L.S., 2018. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol*., **35**: 518-522. https://doi.org/10.1093/molbev/msx281

Huang, X., Chen, B., Wei, Z. and Shi, A., 2022. First report of complete mitochondrial genome in the tribes Coomaniellini and Dicercini (Coleoptera: Buprestidae) and phylogenetic implications. *Genes* (Basel), **13**: 1074. https://doi.org/10.3390/genes13061074

Hurst, L.D., 2002. The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet*., **18**: 486. https://doi.org/10.1016/S0168-9525(02)02722-1

Inohira, K., Hara, T. and Matsuura, E.T., 1997. Nucleotide sequence divergence in the A+T-rich region of mitochondrial DNA in *Drosophila simulans* and *Drosophila mauritiana*. *Mol. Biol. Evol*., **14**: 814-822. https://doi.org/10.1093/oxfordjournals.molbev.a025822

Jia, P.F., Liu, J.H. and Dan, W.L., 2019. Complete mitochondrial genome of the bamboo-shoot fruit fly, *Acrotaeniostola dissimilis* (Diptera: Tephritidae) and its phylogenetic relationship within family Tephritidae. *Mitochondrial DNA B Resour*., **5**: 106-107. https://doi.org/10.1080/23802359.2019.1694455

Jiang, F., Pan, X.B., Li, X.K., Yu, Y.X., Zhang, J.H., Jiang, H.S., Dou, L.D. and Zhu, S.F., 2016. The first complete mitochondrial genome of *Dacus longicornis* (Diptera: Tephritidae) using next-generation sequencing and mitochondrial genome phylogeny of Dacini tribe. *Sci. Rep*., **6**: 36426. https://doi.org/10.1038/srep36426

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

Katoh, K. and Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol*. **30**: 772-780. https://doi.org/10.1093/molbev/mst010

Korneyev, V.A., 1999. *Phylogenetic relationships among higher groups of Tephritidae*. In: *Fruit flies* (Tephritidae) (ed. M.A.N. Aluja). Crc Press, Boca Raton. pp. 73-113. https://doi.org/10.1201/9781420074468.sec2

Krosch, M.N., Schutze, M.K., Armstrong, K.F., Graham, G.C., Yeates, D.K. and Clarke, A.R., 2012. A molecular phylogeny for the tribe Dacini (Diptera: Tephritidae): Systematic and biogeographic implications. *Mol. Phylogenet. Evol*., **64**: 513-523. https://doi.org/10.1016/j.ympev.2012.05.006

Kumar, S., Stecher, G. and Tamura, K., 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol*., **33**: 1870-1874. https://doi.org/10.1093/molbev/msw054

Lavrov, D.V., Brown, W.M. and Boore, J.L., 2000. A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede, *Lithobius forficatus*. *Proc. natl. Acad. Sci. USA*, **97**: 13738-13742. https://doi.org/10.1073/pnas.250402997

Lewis, D.L., Farr, C.L. and Kaguni L.S., 1995. Drosophila melanogaster mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons. *Insect mol. Biol*., **4**: 263-278. https://doi.org/10.1111/j.1365-2583.1995.tb00032.x

Li, R., Ma, Z. and Zhou, C., 2021. The first two complete mitochondrial genomes of Neoephemeridae (Ephemeroptera): Comparative analysis and phylogenetic implication for Furcatergalia. *Genes* (Basel), **12**: 1875. https://doi.org/10.3390/genes12121875

Librado, P. and Rozas, J., 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**: 1451-1452. https://doi.org/10.1093/bioinformatics/btp187

Lv, S.S., Zhang, Y.J., Gong, N. and Chen, X.S., 2021. Characterization and phylogenetic analysis of the mitochondrial genome sequence of *Nisia fuliginosa* (Hemiptera: Fulgoroidea: Meenoplidae). *J. Insect Sci*., **21**: 8. https://doi.org/10.1093/jisesa/ieab050

Mazzon, L., Whitmore, D., Cerretti, P. and Korneyev, V.A., 2021. New and confirmed records of fruit flies (Diptera, Tephritidae) from Italy. *Biodivers. Data J*., **9**: e69351. https://doi.org/10.3897/BDJ.9.e69351

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. and Lanfear, R., 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol*., **37**: 1530-1534. https://doi.org/10.1093/molbev/msaa015

Mu, Y.L., Zhang, C.H., Zhang, Y.J., Yang, L. and Chen, X.S., 2022. Characterizing the complete mitochondrial genome of *Arma custos* and

Online First Article

Picromerus lewisi (Hemiptera: Pentatomidae: Asopinae) and conducting phylogenetic analysis. *J. Insect Sci*., **22**: 6. https://doi.org/10.1093/jisesa/ieab105

Nguyen, D.T., Wu, B., Xiao, S. and Hao, W., 2020. Evolution of a record-setting AT-rich genome: Indel mutation, recombination, and substitution bias. *Genome Biol. Evol*., **12**: 2344-2354. https://doi.org/10.1093/gbe/evaa202

Ojala, D., Montoya, J. and Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature*, **29**: 470-474. https://doi.org/10.1038/290470a0

Pape, T., Bickel, D. and Meier, R., 2009. *Diptera diversity: Status, challenges and tools*. Brill, Leiden. https://doi.org/10.1163/ej.9789004148970.I-459

Patel, R.K. and Jain, M., 2012. NGS QC Toolkit: A toolkit for quality control of next generation sequencing data. *PLoS One*, **7**: e30619. https://doi.org/10.1371/journal.pone.0030619

Peng, Y., Leung, H.C., Yiu, S.M. and Chin, F.Y., 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.T. and Kocher, T.D., 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. mol. Evol*., **41**: 353-358. https://doi.org/10.1007/BF01215182

San Jose, M., Doorenweerd, C., Leblanc, L., Barr, N., Geib, S. and Rubinoff, D., 2018. Incongruence between molecules and morphology: A seven-gene phylogeny of Dacini fruit flies paves the way for reclassification (Diptera: Tephritidae). *Mol. Phylogenet. Evol*., **121**: 139-149. https://doi.org/10.1016/j.ympev.2017.12.001

Siderhurst, M.S. and Jang, E.B., 2010. Cucumber volatile blend attractive to female melon fly, *Bactrocera cucurbitae* (Coquillett). *J. chem. Ecol*., **36**: 699-708. https://doi.org/10.1007/s10886-010-9804-4

Song, S.N., Tang, P., Wei, S.J. and Chen, X.X., 2016. Comparative and phylogenetic analysis of the mitochondrial genomes in basal hymenopterans. *Sci. Rep*., **6**: 20972. https://doi.org/10.1038/srep20972

Soumia, P.S, Shirsat, D.V., Krishna, R., Pandi G, G.P., Choudhary, J.S., Naaz, N., Karuppaiah, V., Gedam, P.A., Anandhan, S. and Singh, M., 2022. Unfolding the mitochondrial genome structure of green semilooper (*Chrysodeixis acuta* Walker): an emerging pest of onion (*Allium cepa* L.). *PLoS One*, **17**: e0273635. https://doi.org/10.1371/journal.pone.0273635

Soumia, P.S., Shirsat, D.V., Krishna, R., G, G.P.P., Choudhary, J.S., Naaz, N.V.K., Gedam, P.A. and Singh, M., 2022. Unfolding the mitochondrial genome structure of green semilooper (*Chrysodeixis acuta* Walker): An emerging pest of onion (*Allium cepa* L.). *PLoS One*, **17**: e0273635. https://doi.org/10.1371/journal.pone.0273635

Teixeira da Costa, L., Powell, C., van Noort, S., Costa, C., Sinno, M., Caleca, V., Rhode, C., Kennedy, R.J., van Staden, M. and van Asch, B., 2019. The complete mitochondrial genome of *Bactrocera biguttula* (Bezzi) (Diptera: Tephritidae) and phylogenetic relationships with other Dacini. *Int. J. Biol. Macromol*., **126**: 130-140. https://doi.org/10.1016/j.ijbiomac.2018.12.186

Vaidya, G., Lohman, D.J. and Meier, R., 2011. Sequence Matrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, **27**: 171-180. https://doi.org/10.1111/j.1096-0031.2010.00329.x

Varani, G. and McClain, W.H., 2000. The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep*., **1**: 18-23. https://doi.org/10.1093/embo-reports/kvd001

Virgilio, M., Jordaens, K., Verwimp, C., Whit,e I.M. and De Meyer, M., 2015. Higher phylogeny of frugivorous flies (Diptera, Tephritidae, Dacini): Localised partition conflicts and a novel generic classification. *Mol. Phylogenet. Evol*., **85**: 171-179. https://doi.org/10.1016/j.ympev.2015.01.007

Wang, C.Y., Zhao, M., Wang, S.J., Xu, H.L., Yang, Y.M., Liu, L.N. and Feng, Y., 2021. The complete mitochondrial genome of *Lepidotrigona flavibasis* (Hymenoptera: Meliponini) and high gene rearrangement in *Lepidotrigona mitogenomes. J. Insect Sci*., **21**: 10. https://doi.org/10.1093/jisesa/ieab038

Wang, T., Cai, B., Wang, X.N. and Ren, Y.L., 2022. Mitochondrial genome of *Dacus haikouensis* Wang and Cheng 2002 using next-generation sequencing from China and its phylogenetic implication. *Mitochond.DNA B Resour*., **7**: 317-319. https://doi.org/10.1080/23802359.2021.2017365

White, I.M. and Elson-Harris, M.M., 1992. *Fruit flies of economic significance: Their identification and bionomics*. CAB International in association with ACIAR, Wallingford, Oxon, UK. pp. 601.

Wilson, A.J. and Xu, J., 2012. Mitochondrial inheritance: diverse patterns and mechanisms with an emphasis

Online First Article

on fungi. *Mycology*, **3**: 158-166.

Wolstenholme, D.R., 1992. Animal mitochondrial DNA: Structure and evolution. *Int. Rev. Cytol.*, **141**: 173-216. https://doi.org/10.1016/S0074-7696(08)62066-5

Xiao, L.F., Zhang, S.D., Long, C.P., Guo, Q.Y., Xu, J.S., Dai, X.H. and Wang, J.G., 2019. Complete mitogenome of a leaf-mining buprestid beetle, *Trachys auricollis*, and its phylogenetic implications. *Genes* (Basel), **10**: 992. https://doi.org/10.3390/genes10120992

Yan, L., Hou, Z.Z., Ma, J.N., Wang, H.M., Gao, J., Zeng, C.J., Chen, Q., Yu, B.S. and Zhang, X.Y., 2022. Complete mitochondrial genome of *Episymploce splendens* (Blattodea: Ectobiidae): A large intergenic spacer and lacking of two tRNA genes. *PLoS One*, **17**: e0268064. https://doi.org/10.1371/journal.pone.0268064

Yang, M., Zhang, H., Song, L., Shi, Y. and Liu, X., 2019. The complete mitochondrial genome of *Mahanta tanyae* compared with other zygaenoid moths (Lepidoptera: Zygaenoidea). *J. Asia-Pac. Entomol.*, **22**: 513-521. https://doi.org/10.1016/j.aspen.2019.03.010

Yang, M.J., Liu, J.H., Wan, X.S., Zhang, Q.L., Fu, D.Y., Wang, X.B., Dan, W.L. and Zhou, X.H., 2020. Complete mitochondrial genome of the black-winged fly, *Felderimyia fuscipennis* (Diptera: Tephritidae) and its phylogenetic relationship within family Tephritidae. *Mitochond. DNA B Resour.*, **5**: 3638-3639. https://doi.org/10.1080/23802359.2020.1831984

Yi, J.Q., Wu, H., Liu, J.B., Li, J.H., Lu, Y.L., Zhang, Y.F., Cheng, Y.J., Guo, Y., Li, D.S. and An, Y.X., 2022. Novel gene rearrangement in the mitochondrial genome of *Anastatus fulloi* (Hymenoptera Chalcidoidea) and phylogenetic implications for Chalcidoidea. *Sci. Rep.*, **12**: 1351. https://doi.org/10.1038/s41598-022-05419-0

Yong, H.S., Song, S.L., Lim, P.E., Chan, K.G., Chow, W.L. and Eamsobhana, P., 2015. Complete mitochondrial genome of *Bactrocera arecae* (Insecta: Tephritidae) by next-generation sequencing and molecular phylogeny of Dacini tribe. *Sci. Rep.*, **5**: 15155. https://doi.org/10.1038/srep15155

Yong, H.S., Song, S.L., Lim, P.E., Eamsobhana, P. and Suana, I.W., 2016. Complete mitochondrial genome of three *Bactrocera* fruit flies of *Subgenus Bactrocera* (Diptera: Tephritidae) and their phylogenetic implications. *PLoS One*, **11**: e0148201. https://doi.org/10.1371/journal.pone.0148201

Yuan, L.L., Ge, X.Y., Xie, G.L., Liu, H.Y. and Yang, Y.X., 2021. First complete mitochondrial genome of Melyridae (Coleoptera, Cleroidea): Genome description and phylogenetic implications. *Insects*, **12**: 87. https://doi.org/10.3390/insects12020087

Zeng, L.Y., Pang, Y.T., Feng, S.Q., Wang, Y.N., Stejskal, V., Aulicky, R., Zhang, S.F. and Li, Z.H., 2021. Comparative mitochondrial genomics of five Dermestid beetles (Coleoptera: Dermestidae) and its implications for phylogeny. *Genomics*, **113**: 927-934. https://doi.org/10.1016/j.ygeno.2020.10.026

Zhang, B., Nardi, F., Hull-Sanders, H., Wan, X. and Liu, Y., 2014. The complete nucleotide sequence of the mitochondrial genome of *Bactrocera minax* (Diptera: Tephritidae). *PLoS One*, **9**: e100558. https://doi.org/10.1371/journal.pone.0100558

Zhang, D.X. and Hewitt, G.M., 1997. Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. *Biochem. Syst. Ecol.*, **25**: 99-120. https://doi.org/10.1016/S0305-1978(96)00042-7

Zhang, D.X., Szymura, J.M. and Hewitt, G.M., 1995. Evolution and structural conservation of the control region of insect mitochondrial DNA. *J. mol. Evol.*, **40**: 382-391. https://doi.org/10.1007/BF00164024

Zhou, X.H., Liu, J.H., Zhang, Q.L., Wan, X.S., Fu, D.Y., Wang, X.B., Dan, W.L. and Yang, M.J., 2020. Complete mitochondrial genome of melon fly, *Zeugodacus cucurbitae* (Diptera: Tephritidae) from Kunming, Southwest China and the phylogeny within subfamily Dacinae. *Mitochond. DNA B Resour.*, **5**: 2828-2829. https://doi.org/10.1080/23802359.2020.1790318

Online First Article