

The Complete Mitochondrial Genome of the Chicken Body Louse, *Menacanthus cornutus*, and Evolutionary Patterns of Extensive Gene Rearrangements in the Mitochondrial Genomes of Amblycera (Psocodea: Phthiraptera)

Siyu Gong, Ye Xu, Shiwen Xu, Yanxin Liang, Li Tian, Wanzhi Cai, Hu Li and Fan Song *

Department of Entomology and MOA Key Lab of Pest Monitoring and Green Management, College of Plant Protection, China Agricultural University, Beijing 100193, China; gsy1207@cau.edu.cn (S.G.); xuyezdk@163.com (Y.X.); xusw2019@126.com (S.X.); yusisama@163.com (Y.L.); ltian@cau.edu.cn (L.T.); caiwz@cau.edu.cn (W.C.); tigerleecau@hotmail.com (H.L.)

* Correspondence: fansong@cau.edu.cn; Tel.: +86-10-62734842

Citation: Gong, S.; Xu, Y.; Xu, S.; Liang, Y.; Tian, L.; Cai, W.; Li, H.; Song, F. The Complete Mitochondrial Genome of the Chicken Body Louse, *Menacanthus cornutus*, and Evolutionary Patterns of Extensive Gene Rearrangements in the Mitochondrial Genomes of Amblycera (Psocodea: Phthiraptera). *Genes* **2022**, *13*, 522. <https://doi.org/10.3390/genes13030522>

Academic Editor: Mark C. Field

Received: 25 February 2022

Accepted: 12 March 2022

Published: 16 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. 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/>).

Abstract: Animal mitochondrial (mt) genomes are typically double-strand circular DNA molecules, but diverse structural variations have been widely found in multiple groups. In parasitic lice (Phthiraptera), the structure of mt genomes varies remarkably across all five suborders. In this study, we reported the complete mt genome of a chicken body louse, *Menacanthus cornutus*, which has a typical single circular mt chromosome and drastic mt gene rearrangements. This mt genome is 15,693 bp in length, consisting of 13 protein-coding genes, 23 tRNA genes, 2 rRNA genes, and a control region. A comparison with a typical insect mt genome suggested that two highly similar *trnM* are present in the mt genome of *M. cornutus*. Moreover, almost every single gene was rearranged, and over half of mt genes were inverted. Phylogenetic analyses inferred from the mt genome sequences supported the monophyly and position of Amblycera. Mapped over the phylogenetic relationships of Amblycera, we identified two inversion events for the conserved gene blocks in Boopidae and Menoponidae. The inverted *ND4L-ND4* was likely a synapomorphic rearrangement in Menoponidae. Our study demonstrated the importance of sequencing mt genomes for additional taxa to uncover the mechanism underlying the structural evolution of the mt genome in parasitic lice.

Keywords: *Menacanthus cornutus*; mitochondrial genome; gene rearrangement; evolution

1. Introduction

The mitochondrion is an organelle with independent genetic material that is closely related to cellular energy metabolism and acts as a power bank in the life activities of eukaryotes [1–3]. Characterized by its stable gene composition, conserved gene order, maternal inheritance, and fast evolutionary rate [2,4], the mitochondrial (mt) genome has been extensively applied as a molecular marker for species delimitation, molecular phylogeny, population genetics and so forth [5–8]. Typically, an animal mt genome is a 14–20 kb double-strand circular DNA molecule containing 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs), two ribosomal RNA genes (rRNAs), and a control region (CR) regulating the replication and transcription of the mt genome [2,4,9]. The majority of insect mt genomes present a relatively conservative gene order, such as that in *Drosophila yakuba*, which is considered as the putative ancestral pattern of the insect gene arrangement [10]. Yet, with a rapid increase in insect mt genome sequencing in recent years, structural variations with novel gene arrangement patterns have been discovered, especially in the order Phthiraptera [11–15], whose gene rearrangement is the most intense, with almost every single gene rearranged.

Phthiraptera, also known as parasitic lice, are mostly harmful parasitic insects. There are approximately 5000 species in five suborders, of which about 4000 and 1000 species parasitize birds and mammals, respectively [16,17]. In addition to the high degree of gene rearrangements, the typical single chromosome of the mt genome was found to be replaced by several circular minichromosomes in some parasitic lice, a phenomenon commonly known as mt genome fragmentation [18–23]. Mt genome fragmentation was first discovered in the human body louse, *Pediculus humanus*, whose typical single chromosome was fragmented into 20 minichromosomes [18,19]. The fragmented mt genomes were indicated to be an ancestral trait in a clade of Eutherian mammal lice (Trichodectera, Rhynchophthirina, and Anoplura) [20]. However, more recent studies suggest that mt genome fragmentation is more widespread among lice, as full or partial mt genomes with multiple chromosomes have been reported from Ischnocera [21] and Amblycera [22]. Moreover, there is a high degree of intra-genus variation in the patterns of gene rearrangements and genome fragmentation in lice [20,21], making this order an ideal group for exploring the evolutionary mechanism underpinning the structural evolution of the mt genome.

To date, a total of 37 complete or nearly complete mt genomes of parasitic lice have been published in the GenBank database. As the suborder with the most diverse host, which is parasitic on both birds and mammals, Amblycera plays a non-negligible role in the evolution of parasitic lice and may contain valuable evidence for the mt genome structural evolution. However, for Amblycera, which contains six families, only three complete mt genomes from two families (Boopidae and Menoponidae) were reported [11,20,22]. In this study, we applied high-throughput sequencing to obtain the mt genome of the chicken body louse *Menacanthus cornutus*, a pathogenic hematophagous species in Amblycera that may cause anaemia, heavy multi-focal lesion, or even the death of infested birds [24,25]. We analyzed the patterns of gene rearrangement and genome fragmentation in this species, and we used the mt gene data to recover the phylogenetic relationships of Phthiraptera. Based on the phylogenetic tree, we assessed the evolution of gene arrangements in Amblycera. Our study enriched the mt genome data of Amblycera and provided more information to the diversity and evolution of louse mt genomes.

2. Materials and Methods

2.1. Sample Collection, Genomic DNA Extraction, and Sequence Amplification

Louse specimens were collected from chickens, *Gallus gallus* in Chongqing Jinyun Mountain, which acted as the host of *Menacanthus cornutus*. The specimens were preserved in 100% ethanol at −20 °C after removal from their host until DNA extraction. Total genomic DNA was extracted from a whole individual louse using the DNeasy Blood and Tissue Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions. The genomic DNA was stored below −20 °C at the Entomological Museum of China Agricultural University (Beijing, China).

The fragments of two mt genes (*COI*, ~610 bp and *srRNA*, ~530 bp) were amplified by polymerase chain reaction (PCR). The PCR products were sequenced by Sanger sequencing at Tsingke Biotechnology (Beijing, China).

2.2. Mitochondrial Genome Sequencing, Assembly, and Annotation

An Illumina TruSeq library was prepared with an average insert size of 350 bp and was sequenced on the Illumina NovaSeq 6000 platform at Berry Genomics (Beijing, China). A total of 8 Gb of clean data (150 bp paired-end reads) was obtained for the species. The raw reads were trimmed of adapters using Trimmomatic [26] and low-quality and short reads were removed using Prinseq [27]. High-quality reads were used in *de novo* assembly using IDBA-UD [28] with the parameters: similarity threshold 98%, minimum *k* value 45, and maximum *k* value 145. To identify the mt genome sequence, the assembled contigs were searched with the *COI* and *srRNA* partial sequences using BLAST [29] with

at least 98% similarity. The contigs were then used as reference sequences to fish sequencing reads in Geneious Prime v2020.0.5 using the “map to reference” option. The assembly parameters were set as no gaps, minimum overlap 100 bp, minimum overlap identity 98%, maximum mismatches per read 2%, and maximum ambiguity 2 until the two ends of the contig overlapped and ensured the graph of coverage was smooth. Finally, we obtained a complete circular mt genome with an average coverage depth of 2901×.

tRNA genes were annotated using the MITOS web server [30] and the tRNAscan-SE search server [31] with the invertebrate mt code. The locations of PCGs and rRNA genes were identified by alignment with homologous genes of other parasitic lice available in the GenBank. The nucleotide composition of the mt genome was analyzed with MEGA 11 [32]. To measure the nucleotide compositional differences of the whole mt genome, we also used the following formulas: (1) AT skew = $(A - T)/(A + T)$ and (2) GC skew = $(G - C)/(G + C)$ to calculate the AT- and GC-skews [33].

2.3. Phylogenetic Analyses

A total of 32 mt genomes of parasitic lice were used for the phylogenetic analyses, including the newly sequenced mt genome of *Menacanthus cornutus*. Seven mt genomes (one Hemiptera species and six Psocoptera species) were used as outgroups (Table 1).

Eleven PCGs and two rRNA genes were used for phylogenetic analyses. The *ND2* and *ND6* genes were not found in the elephant louse (*Haematomyzus elephantis*) and opossum louse (*Cummingsia maculata*), and, thus, these two PCGs were excluded from the phylogenetic analyses. Each PCG was aligned individually based on codons for amino acids using the MAFFT algorithm [34], implemented in the TranslatorX online platform with the L-INS-i strategy and default setting [35]. Two rRNA genes were individually aligned using the MAFFT v7.0 online server with the G-INS-i strategy [36]. After removing poorly aligned sites using GBLOCKS v0.91b [37], the alignment of the individual gene was concatenated into two datasets: (i) the PCGRNA matrix, including eleven PCGs and two rRNA genes (7815 bp in total) and (ii) the PCG12RNA matrix, including the first and second codon positions of the eleven PCGs and two rRNA genes (5689 bp in total).

Table 1. Taxa used in this study.

Order	Suborder	Family	Species	GenBank Accession Number
Outgroup				
Hemiptera	Heteroptera	Nabidae	<i>Alloeorhynchus bakeri</i>	NC_016432
Psocoptera	Trogiomorpha	Lepidopsocidae	<i>Lepidopsocid</i> sp.	NC_004816
	Psocomorpha	Psocidae	<i>Psococerastis albimaculata</i>	NC_021400
			<i>Longivalvus hyalospilus</i>	JQ910986
	Troctomorpha	Liposcelidae	<i>Liposcelis decolor</i>	NC_023839
			<i>Liposcelis bostrychophila</i>	JN645275-645276
			<i>Liposcelis paeta</i>	KF649225-649226
Ingroup				
Phthiraptera	Amblycera	Boopidae	<i>Heterodoxus macropus</i>	AF270939
			<i>Heterodoxus spiniger</i>	MW199168
		Menoponidae	<i>Menacanthus cornutus</i> ^a	OM718871
			<i>Colpocephalum griffoneae</i>	NC_039530
			<i>Amyrsidea minuta</i>	MH001227
			<i>Osborniella crotophagae</i>	MW199175
			<i>Myrsidea</i> sp.	MW199172-199174
		Ricinidae	<i>Ricinus</i> sp.	MW199176

Ischnocera	Trimenoponidae	<i>Cummingsia maculata</i>	MW199177-199179
	Trichodectidae	<i>Bovicola bovis</i>	MH001189-001200
		<i>Bovicola ovis</i>	MH001201-001212
		<i>Bovicola caprae</i>	MH001176-001188
		<i>Trichodectes canis</i>	MH001213-001224 and MH823541
	Philopteridae	<i>Ibidoecus bisignatus</i>	NC_015999
		<i>Bothriometopus mac-rocnemis</i>	EU183542
		<i>Campanulotes compar</i>	MH001225
		<i>Falcolipeurus quadri-pustulatus</i>	MH001226
		<i>Falcolipeurus suturalis</i>	MW696813
		<i>Columbicola passerinae</i>	MT094266-094282
		<i>Haematomyzus elephan-tis</i>	KF933032-933041
Rhyncophthirina	Haematomyzidae		
Anoplura	Polyplacidae	<i>Polyplax spinulosa</i>	KF647762-647772
		<i>Polyplax asiatica</i>	KF647751-647761
	Hoplopleuridae	<i>Hoplopleura</i> sp.	MT792483-792494
	Microthoraciidae	<i>Microthoradus praelongi-ceps</i>	KX090378-090389
			KF939318, KF939322, KF939324, KF939326, and KJ434034-38
	Haematopinidae	<i>Haematopinus asini</i>	KC814611-814619
		<i>Haematopinus apri</i>	KC814602-814610
		<i>Haematopinus suis</i>	EU219987-219995, HM241895-241898, JQ976018, and MT721740
	Pthiridae	<i>Pthirus pubis</i>	
	Pediculidae	<i>Pediculus schaeffi</i>	KC241882-241897
		<i>Pediculus humanus</i>	FJ499473-499490
		<i>Pedicinus obtusus</i>	MT792495-792506
		<i>Pedicinus badii</i>	MT721726-721739

^a Mitochondrial genome sequenced in the present study.

Many previous studies have revealed that the site-heterogeneous mixture model (CAT+GTR) implemented in PhyloBayes can avoid the false grouping of unrelated taxa with a similar base composition and an accelerated evolutionary rate in the reconstruction of the high-level phylogeny of many insect orders based on mt genome sequences [6,7,38,39]. Therefore, both datasets were analyzed under the CAT+GTR model using PhyloBayes MPI v1.8c [40]. Two independent Markov chain Monte Carlo chains were run after the removal of constant sites from the alignment and were stopped after the two runs had satisfactorily converged (maxdiff < 0.1). A consensus tree was computed from the remaining trees that were combined from two runs after the initial 25% trees of each run were discarded as burn-in.

3. Results and Discussion

3.1. The General Features of the *Menacanthus cornutus* mt Genome

The complete mt genome of *Menacanthus cornutus* is a single circular chromosome that is 15,693 bp in size, including 38 coding genes (13 PCGs, 23 tRNA genes, and 2 rRNA genes) and a CR (Figure 1, Table S1). In total, 12 genes were transcribed from the minority strand (N-strand) while the other 26 genes were encoded on the majority strand (J-strand). Six gene overlaps were observed, ranging from 1 to 7 bp in length, and the longest occurred between *ATP6* and *ATP8*, which was common among the mt genomes of insects [4]. The CR was located between *trnC* and *trnK*, with a size of 871 bp. Apart from the CR, we also identified 23 non-coding regions (NCRs) comprising a total of 582 bp, with the longest being 88 bp between *trnC* and *trnF*.

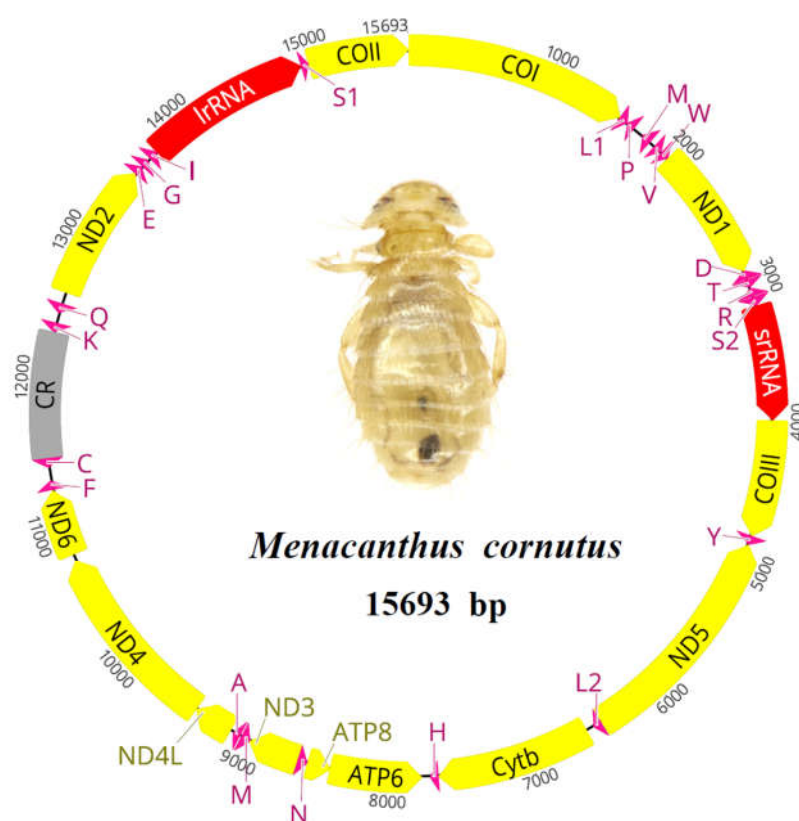


Figure 1. The mitochondrial genome of *Menacanthus cornutus*. The arrows indicate the direction of gene transcription. The yellow, pink, red, and gray colors indicate the regions of PCGs, tRNA genes, rRNA genes, and control regions, respectively. Abbreviations: *ATP6* and *ATP8*, adenosine triphosphate (ATP) synthase subunits 6 and 8; *COI–COIII*, cytochrome C oxidase subunits I–III; *Cytb*, cytochrome b; *ND1–ND6* and *ND4L*, nicotinamide adenine dinucleotide hydrogen (NADH) dehydrogenase subunits 1–6 and 4L; *16S* and *23S*, large and small rRNA subunits; The tRNA genes are indicated with their one-letter corresponding amino acids; CR, control region.

All the PCGs were initiated by ATN (eight by ATT, two by ATA, one by ATA, and one by ATC), except *ND1* used TTG as a start codon. A complete TAA or TAG was used by all the PCGs as the termination codon (Table S1). The 23 detected tRNA genes ranged from 54–70 bp, and all the secondary structures could be folded into the typical clover-leaf structure. Different from the typical 22 tRNAs in insect mt genomes, 2 *trnM* were found

in the *Menacanthus cornutus* mt genome. The sequences of the two *trnM* were highly identical, with differences occurring on only six nucleotide positions in the DHU arm and the TΨC arm (Figure S1). Two copies of *trnM* were also detected in *Cummingsia maculata* (Amblycera: Trimenoponidae), whereas the sequence similarity between the two *trnM* was lower than *Menacanthus cornutus* (16 nucleotide positions were different). The large rRNA subunit (*lrRNA*) and small rRNA subunit (*srRNA*) were encoded on the J strand; however, these two rRNA genes were transcribed from the N strand in the ancestral insect mt genome [4,41]. The two rRNA genes were in the locations *trnI* to *trnS1* and *trnS2* to *COIII*, with lengths of 1163 bp and 800 bp respectively.

The nucleotide composition of the *Menacanthus cornutus* mt genome is significantly biased toward adenine (A) and thymine (T), with an A+T content of 74.1% (Table S2). In contrast to most published insect mt genomes, the nucleotide composition shows a negative AT-skew (−0.055) and a positive GC-skew (0.023) [42]. The AT and GC skews of several amblyceran lice are all negative, which may result from the gene rearrangement events, especially the inversion of genes [22]. Within the whole mt genome, CR shows the most biased A+T content (76.9%), while the A+T content is 73.8% in the PCGs, 73.1% in the tRNAs, and 73.7% in the rRNAs, respectively.

3.2. Phylogenetic Analyses of the Suborder Amblycera and Other Parasitic Lice

The suborder-level phylogeny of Phthiraptera was reconstructed using Bayesian inference based on the two concatenated datasets (PCGRNA and PCG12RNA) (Figure 2). The two phylogenetic trees showed identical topology. The monophyly of the Phthiraptera was well supported (BPPs = 1), and the sister relationship between Phthiraptera and the family Liposcelididae was highly supported (BPPs = 1). Amblycera was monophyletic with strong support (BPPs = 1) and was relatively distantly related to the other monophyletic clade that included all four other suborders. The topology of the phylogenetic tree also supported the paraphyly of Ischnocera, as *Columbicola passerinae* and *Bothriometopus macrocnemis* were not nested inside other ischnoceran species. However, this bird louse clade was shown to be monophyletic by studies using nuclear genes [43,44]. Trichodectera, which was previously regarded as the family Trichodectidae in Ischnocera, was clustered with Rhynchophthirina and Anoplura, which are parasitic on eutherian mammals, with strong support in our study (BPPs = 1). This mammal-infesting clade, comprising Trichodectera, Rhynchophthirina and Anoplura, was also supported to exist by recent studies using transcriptome data [43], shared derived features of mt genome fragmentation [20], and orthologous nuclear genes [44]. Within the suborder Amblycera, *Menacanthus cornutus* is indicated to be a sister lineage with *Amyrsidea minuta* (BPPs = 1), and a monophyletic Menoponidae was also well supported (BPPs = 1). Ricinidae was the most distantly related of the Amblycera species examined (BPPs = 1), and the three mammal lice (*Cummingsia maculata*, *Heterodoxus macropus*, and *Heterodoxus spiniger*) were clustered together as the sister group to the family Menoponidae (BPPs = 0.99).

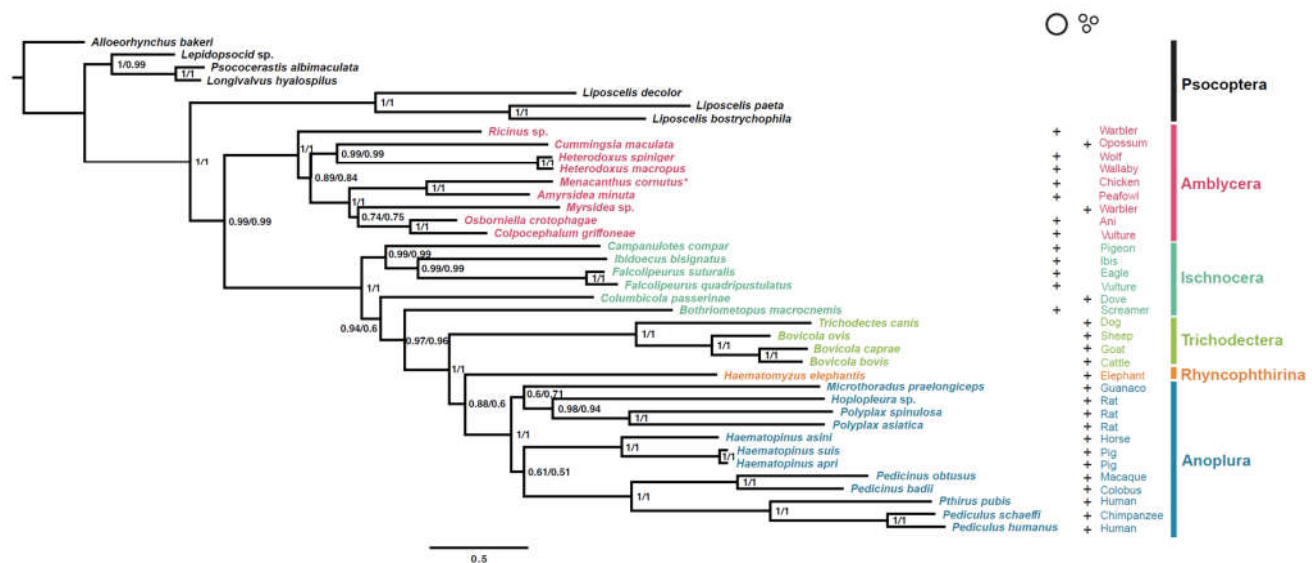


Figure 2. Phylogenetic trees of Phthiraptera inferred from the PCGRNA and PCG12RNA datasets using PhyloBayes under the site-heterogeneous mixture CAT + GTR model. Supports at nodes (from left to right) are Bayesian posterior probabilities (BPPs) for PCGRNA and PCG12RNA. The mitochondrial genome sequence of *M. cornutus* is highlighted with an asterisk (*). The genome architecture (typical or fragmented) is indicated by a plus sign (+). The bird and mammal hosts are indicated after the names of the parasitic lice.

The published mt genomes of Trichodectera, Rhynchophthirina, and Anoplura lice, which are parasitic on eutherian mammals, all have extensive mt genome fragmentation (9–20 minichromosomes) [20,23]. The pigeon and dove louse *Columbicola passerinae*, from the family Philopteridae, was fragmented into 15–17 minichromosomes [21]. In a recent study, fragmented mt genomes of Amblycera (*Cummingsia maculata* and *Myrsidea* sp.) were assembled from the sequencing data in the NCBI SRA database [22]. According to the phylogenetic trees in our study, the fragmented mt genomes evolved from single chromosome mt genomes more than once in Phthiraptera. Even in the suborder Amblycera, mt genome fragmentation evolved independently at least two times, as the *Cummingsia maculata* and *Myrsidea* sp. which have fragmented mt genomes were not nested in one clade.

3.3. Evolution of Gene Rearrangement in the Mitochondrial Genome of Amblycera

Previous studies show that the mt genome of the order Phthiraptera is characterized by drastic mt gene rearrangement [11–15]. As one of the suborders of Phthiraptera, Amblycera also have variable and highly rearranged mt genomes [22]. For the newly sequenced mt genome of *Menacanthus cornutus*, nearly all genes were rearranged (Figure 3). No conserved gene clusters could be found in the tRNAs, and, in total, 12 tRNAs were inverted. Even for the PCGs and rRNAs, which rarely rearranged in insect mt genomes, only two conserved gene blocks could be detected (*ATP8-ATP6* and *ND4-ND4L*). However, the coding direction of the two gene blocks was inverted in the mt genome of *Menacanthus cornutus* compared to the putative ancestral type of *D. yakuba*.

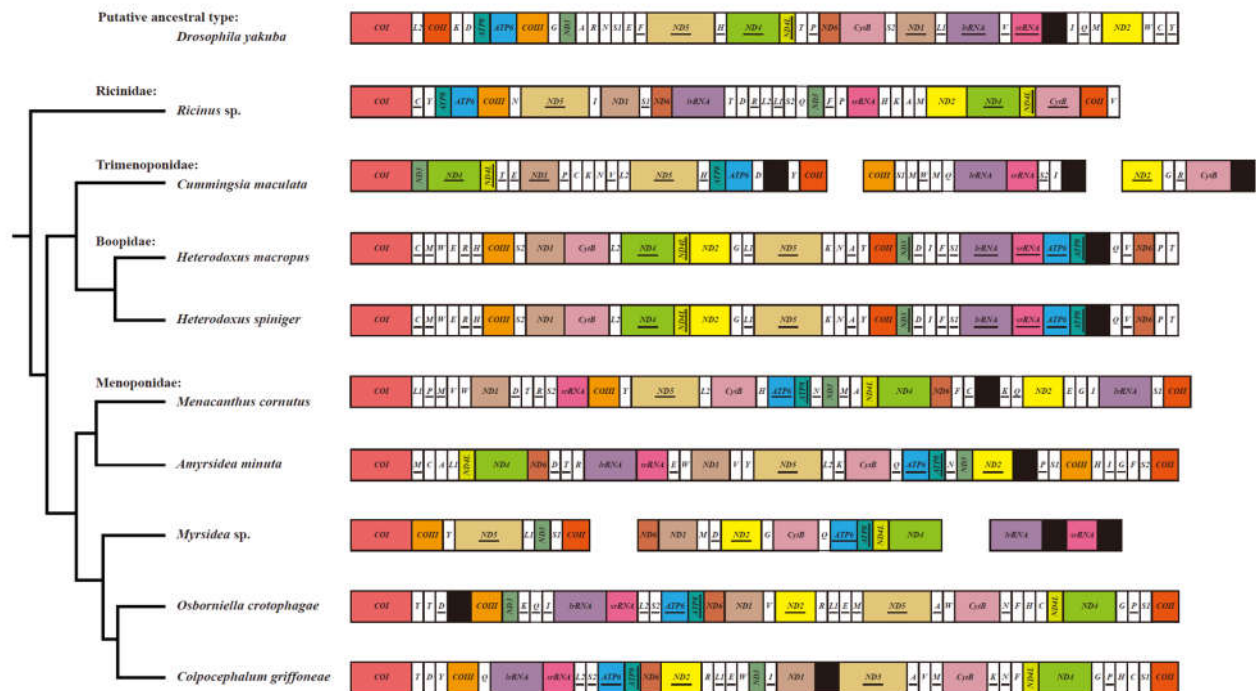


Figure 3. Gene arrangements of the Amblycera mitochondrial genomes. The mt genomes have been linearized for ease of comparison and arbitrarily begin with *COI* when possible. Different PCGs and rRNA genes are shown in different colors, with tRNA genes in white and control or non-coding intergenic regions in black. The gene names are the standard abbreviations used in this study. Underlined labels indicate the gene was transcribed from the minority strand.

To better understand the evolution of extensive gene rearrangements in the mt genomes of Amblycera, gene orders of published amblyceran mt genomes were summarized and mapped on an estimated phylogenetic tree (Table 2, Figure 3). The patterns of gene rearrangement are highly variable across Amblycera, and the gene blocks *ATP8-ATP6* and *ND4-ND4L* are the only region shared by all nine species. However, even in this region, a slight variation in the coding direction can be found between some species. In contrast to the other three families, an inverted *ND4L-ND4* is shared by all five Menoponidae species, and this gene rearrangement pattern may be synapomorphic across Menoponidae. The gene blocks *ATP8-ATP6* were not changed in Ricinidae and Trimenoponidae but were inverted in Boopidae and Menoponidae species. Such gene rearrangement patterns agree with the recovered phylogenetic relationships, as the outgroup, Ricinidae, and some of the ingroup, Trimenoponidae, had the same coding direction of *ATP8-ATP6* and *ND4-ND4L* as that found in *D. yakuba*, and two inversion events happened in Boopidae (*ATP8-ATP6* to *ATP6-ATP8*) and Menoponidae (*ND4-ND4L* to *ND4L-ND4*) during the evolution of Amblycera. At the family level, we found that the gene blocks *trnS1-COII* and *lrRNA-srRNA* were present in most of Menoponidae species but disappeared in other families. In Boopidae, *Heterodoxus macropus*, and *Heterodoxus spiniger* have a completely identical gene order, indicating this gene arrangement may be conserved within the genus of Amblycera. The structure of the mt genome is more conserved in Menoponidae. *Colpocephalum griffoneae* and *Osborniella crotophagae* share identical gene arrangements in two major gene (PCGs and rRNAs) clusters *lrRNA-srRNA-ATP6-ATP8-ND6* and *ND5-Cytb-ND4L-ND4-COII-COI-COIII*.

Table 2. The derived mitochondrial gene blocks shared by parasitic lice in Amblycera.

Family	Species	Fragmented mt Genome	<u>ND4-</u> <u>ND4L</u>	ND4L- ND4	ATP8- ATP6	<u>ATP6-</u> <u>ATP8</u>	<i>trnS1-</i> <i>COII</i>	<i>lrRNA-</i> <i>srRNA</i>
Ricinidae	<i>Drosophila yakuba</i> (Putative ancestral type)	–	+	–	+	–	–	–
	<i>Ricinus</i> sp.	–	+	–	+	–	–	–
Trimenoponidae	<i>Cummingsia maculata</i>	+	+	–	+	–	–	+
Boopidae	<i>Heterodoxus macropus</i>	–	+	–	–	+	–	–
	<i>Heterodoxus spiniger</i>	–	+	–	–	+	–	–
Menoponidae	<i>Menacanthus cornutus</i>	–	–	+	–	+	+	–
	<i>Amyrsidea minuta</i>	–	–	+	–	+	–	+
	<i>Myrsidea</i> sp.	+	–	+	–	+	+	+
	<i>Osborniella crotophagae</i>	–	–	+	–	+	+	+
	<i>Colpocephalum griffoneae</i>	–	–	+	–	+	+	+

Note: Plus (+) indicates presence; minus (–) indicates absence. Underlined labels indicate the gene was transcribed from the minority strand.

Overall, our study added more data on the mt genomes of parasitic lice and enhanced our understanding of the structural diversity of the mt genome of lice. More mt genome sequencing and comparative studies are needed to uncover the mechanisms underlying the structural evolution of the mt genome in parasitic lice.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/article/10.3390/genes13030522/s1, Table S1: Structure of the *Menacanthus cornutus* mitochondrial genome, Table S2: Nucleotide composition of the *Menacanthus cornutus* mitochondrial genome, Figure S1: Secondary structures and sequence similarity of *trnM1* and *trnM2*. Inferred Watson-Crick bonds are illustrated by lines and GU bonds are illustrated by dots. The identical nucleotides between *trnM1* and *trnM2* are labeled with green circles and the variable sites are in pink.

Author Contributions: Conceptualization, F.S.; methodology, S.G., Y.X., and F.S.; investigation, S.G., S.X., and Y.L.; formal analysis, S.G., L.T., H.L., and W.C.; writing—original draft preparation, S.G. and F.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (Nos. 31900343, 32170474, 31922012) and the Young Elite Scientist Sponsorship Program by CAST (No. YESS20200106).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The mitochondrial genome sequences of *Menacanthus cornutus* which used in this study have been deposited into GenBank (accession numbers: OM718871).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Wallace, D.C. Structure and evolution of organelle genomes. *Microbiol. Rev.* **1982**, *46*, 208–240. <https://doi.org/10.1128/membr.46.2.208-240.1982>.
- Boore, J.L. Animal mitochondrial genomes. *Nucleic Acids Res.* **1999**, *27*, 1767–1780. <https://doi.org/10.1093/nar/27.8.1767>.
- McBride, H.M.; Neuspiel, M.; Wasiak, S. Mitochondria: More than just a powerhouse. *Curr. Biol.* **2006**, *16*, R551–R560. <https://doi.org/10.1016/j.cub.2006.06.054>.
- Cameron, S.L. Insect mitochondrial genomics: Implications for evolution and phylogeny. *Annu. Rev. Entomol.* **2014**, *59*, 95–117. <https://doi.org/10.1146/annurev-ento-011613-162007>.
- Hebert, P.D.N.; Cywinska, A.; Ball, S.L.; deWaard, J.R. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B* **2003**, *270*, 313–321. <https://doi.org/10.1098/rspb.2002.2218>.

6. Song, F.; Li, H.; Jiang, P.; Zhou, X.; Liu, J.; Sun, C.; Vogler, A.P.; Cai, W. Capturing the phylogeny of Holometabola with mitochondrial genome data and Bayesian site-heterogeneous mixture models. *Genome Biol. Evol.* **2016**, *8*, 1411–1426. <https://doi.org/10.1093/gbe/evw086>.
7. Li, H.; Leavengood, J.M.; Chapman, E.G.; Burkhardt, D.; Song, F.; Jiang, P.; Liu, J.; Zhou, X.; Cai, W. Mitochondrial phylogenomics of Hemiptera reveals adaptive innovations driving the diversification of true bugs. *Proc. R. Soc. B* **2017**, *284*, 20171223. <https://doi.org/10.1098/rspb.2017.1223>.
8. Du, Z.; Hasegawa, H.; Cooley, J.R.; Simon, C.; Yoshimura, J.; Cai, W.; Sota, T.; Li, H. Mitochondrial genomics reveals shared phylogeographic patterns and demographic history among three periodical cicada species groups. *Mol. Biol. Evol.* **2019**, *36*, 1187–1200. <https://doi.org/10.1093/molbev/msz051>.
9. Zhang, D.X.; Szymura, J.M.; Hewitt, G.M. Evolution and structural conservation of the control region of insect mitochondrial DNA. *J. Mol. Evol.* **1995**, *40*, 382–391. <https://doi.org/10.1007/BF00164024>.
10. Clary, D.O.; Wolstenholme, D.R. The mitochondrial DNA molecular of *Drosophila yakuba*: Nucleotide sequence, gene organization, and genetic code. *J. Mol. Evol.* **1985**, *22*, 252–271. <https://doi.org/10.1007/bf02099755>.
11. Shao, R.; Campbell, N.J.H.; Barker, S.C. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). *Mol. Biol. Evol.* **2001**, *18*, 858–865. <https://doi.org/10.1093/oxfordjournals.molbev.a003867>.
12. Shao, R.; Campbell, N.J.H.; Schmidt, E.R.; Barker, S.C. Increased rate of gene rearrangement in the mitochondrial genomes of three orders of hemipteroid insects. **2001**, *18*, 1828–1832. <https://doi.org/10.1093/oxfordjournals.molbev.a003970>.
13. Covacin, C.; Shao, R.; Cameron, S.; Barker, S.C. Extraordinary number of gene rearrangements in the mitochondrial genomes of lice (Phthiraptera: Insecta). *Insect Mol. Biol.* **2006**, *15*, 63–68. <https://doi.org/10.1111/j.1365-2583.2005.00608.x>.
14. Cameron, S.L.; Johnson, K.P.; Whiting, M.F. The mitochondrial genome of the screamer louse *Bothriometopus* (Phthiraptera: Ischnocera): Effects of extensive gene rearrangements on the evolution of the genome. *J. Mol. Evol.* **2007**, *65*, 589–604. <https://doi.org/10.1007/s00239-007-9042-8>.
15. Nie, Y.; Fu, Y.T.; Zhang, Y.; Deng, Y.P.; Wang, W.; Tu, Y.; Liu, G.H. Highly rearranged mitochondrial genome in *Falcolipeurus* lice (Phthiraptera: Philopteridae) from endangered eagles. *Parasite. Vector.* **2021**, *14*, 269. <https://doi.org/10.1186/s13071-021-04776-5>.
16. Kim, K.C. *Evolutionary Parallelism in Anoplura and Eutherian Mammals*; Oxford University Press: Oxford, UK, 1988.
17. Price, R.D.; Hellenthal, R.A.; Palma, R.L. *World Checklist of Chewing Lice with Host Associations and Keys to Families and Genera*; Illinois Natural History Survey Special Publication: Illinois, USA, 2003.
18. Shao, R.; Kirkness, E.F.; Barker, S.C. The single mitochondrial chromosome typical of animals has evolved into 18 minichromosomes in the human body louse, *Pediculus humanus*. *Genome Res.* **2009**, *19*, 904–912. <https://doi.org/10.1101/gr.083188.108>.
19. Shao, R.; Zhu, X.Q.; Barker, S.C.; Herd, K. Evolution of extensively fragmented mitochondrial genomes in the lice of humans. *Genome Biol. Evol.* **2012**, *4*, 1088–1101. <https://doi.org/10.1093/gbe/evs088>.
20. Song, F.; Li, H.; Liu, G.; Wang, W.; James, P.; Colwell, D.D.; Tran, A.; Gong, S.; Cai, W.; Shao, R. Mitochondrial genome fragmentation unites the parasitic lice of eutherian mammals. *Syst. Biol.* **2019**, *68*, 430–440. <https://doi.org/10.1093/sysbio/syy062>.
21. Sweet, A.D.; Johnson, K.P.; Cameron, S.L. Mitochondrial genomes of *Columbicola* feather lice are highly fragmented, indicating repeated evolution of minicircle-type genomes in parasitic lice. *PeerJ* **2020**, *8*, e8759. <https://doi.org/10.7717/peerj.8759>.
22. Sweet, A.D.; Johnson, K.P.; Cao, Y.; de Moya, R.S.; Skinner, R.K.; Tan, M.; Herrera, S.V.; Cameron, S.L. Structure, gene order, and nucleotide composition of mitochondrial genomes in parasitic lice from Amblycera. *Gene* **2021**, *768*, 145312. <https://doi.org/10.1016/j.gene.2020.145312>.
23. Fu, Y.T.; Dong, Y.; Wang, W.; Nie, Y.; Liu, G.H.; Shao, R. Fragmented mitochondrial genomes evolved in opposite directions between closely related macaque louse *Pedicinus obtusus* and colobus louse *Pedicinus badii*. *Genomics* **2020**, *112*, 4924–4933. <https://doi.org/10.1016/j.ygeno.2020.09.005>.
24. Audi, A.H.; Asmau, A.M. Prevalence of Bird Louse, *Menacanthus Cornutus* (Phthiraptera- Amblycera) in four selected Poultry Farms in Kano State, Nigeria. *Bajopas* **2014**, *7*, 142–146. <http://dx.doi.org/10.4314/bajopas.v7i1.26>.
25. Kumar, A.; Kumar, R. Seasonal changes in the population of *Menacanthus cornutus* (Phthiraptera: Amblycera). *J. Parasit. Dis.* **2014**, *38*, 307–310. <https://doi.org/10.1007/s12639-013-0249-x>.
26. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
27. Schmieder, R.; Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **2011**, *27*, 863–864. <https://doi.org/10.1093/bioinformatics/btr026>.
28. Peng, Y.; Leung, H.C.M.; Yiu, S.M.; Chin, F.Y.L. IDBA-UD: A *de novo* assembler for single cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **2012**, *28*, 1420–1428. <https://doi.org/10.1093/bioinformatics/bts174>.
29. Altschul, S.; Madden, T.; Schäffer, A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.

30. Bernt, M.; Donath, A.; Jühling, F.; Externbrink, F.; Florentz, C.; Fritzsch, G.; Pütz, J.; Middendorf, M.; Stadler, P.F. MITOS: Improved *de novo* metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* **2013**, *69*, 313–319. <https://doi.org/10.1016/j.ympev.2012.08.023>.
31. Chan, P.P.; Lowe, T.M. tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. *Methods Mol. Biol.* **2019**, *1962*, 1–14. https://doi.org/10.1007/978-1-4939-9173-0_1.
32. Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027. <https://doi.org/10.1093/molbev/msab120>.
33. Perna, N.T.; Kocher, T.D. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol.* **1995**, *41*, 353–358. <https://doi.org/10.1007/BF01215182>.
34. Katoh, K.; Kuma, K.; Toh, H.; Miyata, T. MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **2005**, *33*, 511–518. <https://doi.org/10.1093/nar/gki198>.
35. Abascal, F.; Zardoya, R.; Telford, M.J. TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* **2010**, *38*, W7–W13. <https://doi.org/10.1093/nar/gkq291>.
36. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. <https://doi.org/10.1093/molbev/mst010>.
37. Talavera, G.; Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **2007**, *56*, 564–577. <https://doi.org/10.1080/10635150701472164>.
38. Schwentner, M.; Richter, S.; Rogers, D.C.; Giribet, G. Tetraconatan phylogeny with special focus on Malacostraca and Branchiopoda: Highlighting the strength of taxon-specific matrices in phylogenomics. *Proc. R. Soc. B* **2018**, *285*, 20181524. <http://dx.doi.org/10.1098/rspb.2018.1524>.
39. Song, N.; Li, X.; Yin, X.; Li, X.; Yin, J.; Pan, P. The mitochondrial genomes of palaeopteran insects and insights into the early insect relationships. *Sci. Rep.* **2019**, *9*, 17765. <https://doi.org/10.1038/s41598-019-54391-9>.
40. Lartillot, N.; Rodrigue, N.; Stubbs, D.; Richer, J. PhyloBayes MPI: Phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* **2013**, *62*, 611–615. <https://doi.org/10.1093/sysbio/syt022>.
41. Feng, Z.; Wu, Y.; Yang, C.; Gu, X.; Wilson, J.J.; Li, H.; Cai, W.; Yang, H.; Song, F. Evolution of tRNA gene rearrangement in the mitochondrial genome of ichneumonoid wasps (Hymenoptera: Ichneumonoidea). *Int. J. Biol. Macromol.* **2020**, *164*, 540–547. <https://doi.org/10.1016/j.ijbiomac.2020.07.149>.
42. Wei, S.J.; Shi, M.; Chen, X.X.; Sharkey, M.J.; van Achterberg, C.; Ye, G.Y.; He, J.H. New views on strand asymmetry in insect mitochondrial genomes. *PLoS ONE* **2010**, *5*, e12708. <https://doi.org/10.1371/journal.pone.0012708>.
43. Johnson, K.P.; Nguyen, N.; Sweet, A.D.; Boyd, B.M.; Warnow, T.; Allen, J.M. Simultaneous radiation of bird and mammal lice following the K-Pg boundary. *Biol. Lett.* **2018**, *14*, 20180141. <https://doi.org/10.1098/rsbl.2018.0141>.
44. de Moya, R.S.; Yoshizawa, K.; Walden, K.K.O.; Sweet, A.D.; Dietrich, C.H.; Johnson, K.P. Phylogenomics of Parasitic and Non-parasitic Lice (Insecta: Psocodea): Combining Sequence Data and Exploring Compositional Bias Solutions in Next Generation Data Sets. *Syst. Biol.* **2021**, *70*, 719–738. <https://doi.org/10.1093/sysbio/syaa075>.