

Mitochondrial Genomes of Two Barklice, *Psococerastis albimaculata* and *Longivalvus hyalospilus* (Psocoptera: Psocomorpha): Contrasting Rates in Mitochondrial Gene Rearrangement between Major Lineages of Psocodea

Hu Li¹✉, Renfu Shao^{2*}✉, Fan Song¹, Xuguo Zhou³, Qianqian Yang¹, Zhihong Li¹, Wanzhi Cai^{1*}

1 Department of Entomology, China Agricultural University, Beijing, China, **2** GeneCology Research Centre, Faculty of Science, Education and Engineering, University of the Sunshine Coast, Maroochydore, Queensland, Australia, **3** Department of Entomology, University of Kentucky, Lexington, Kentucky, United States of America

Abstract

The superorder Psocodea has ~10,000 described species in two orders: Psocoptera (barklice and booklice) and Phthiraptera (parasitic lice). One booklouse, *Liposcelis bostrychophila* and six species of parasitic lice have been sequenced for complete mitochondrial (mt) genomes; these seven species have the most rearranged mt genomes seen in insects. The mt genome of a barklouse, lepidopsocid sp., has also been sequenced and is much less rearranged than those of the booklouse and the parasitic lice. To further understand mt gene rearrangements in the Psocodea, we sequenced the mt genomes of two barklice, *Psococerastis albimaculata* and *Longivalvus hyalospilus*, the first representatives from the suborder Psocomorpha, which is the most species-rich suborder of the Psocodea. We found that these two barklice have the least rearranged mt genomes seen in the Psocodea to date: a protein-coding gene (*nad3*) and five tRNAs (*trnN*, *trnS1*, *trnE*, *trnM* and *trnC*) have translocated. Rearrangements of mt genes in these two barklice can be accounted for by two events of tandem duplication followed by random deletions. Phylogenetic analyses of the mt genome sequences support the view that Psocoptera is paraphyletic whereas Phthiraptera is monophyletic. The booklouse, *L. bostrychophila* (suborder Troctomorpha) is most closely related to the parasitic lice. The barklice (suborders Trogiomorpha and Psocomorpha) are closely related and form a monophyletic group. We conclude that mt gene rearrangement has been substantially faster in the lineage leading to the booklice and the parasitic lice than in the lineage leading to the barklice. Lifestyle change appears to be associated with the contrasting rates in mt gene rearrangements between the two lineages of the Psocodea.

Citation: Li H, Shao R, Song F, Zhou X, Yang Q, et al. (2013) Mitochondrial Genomes of Two Barklice, *Psococerastis albimaculata* and *Longivalvus hyalospilus* (Psocoptera: Psocomorpha): Contrasting Rates in Mitochondrial Gene Rearrangement between Major Lineages of Psocodea. PLoS ONE 8(4): e61685. doi:10.1371/journal.pone.0061685

Editor: Joseph Schacherer, University of Strasbourg, France

Received: January 21, 2013; **Accepted:** March 13, 2013; **Published:** April 22, 2013

Copyright: © 2013 Li et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research is supported by grants from the National Natural Science Foundation of China (Nos. 30825006, 31061160186, 31111140015), the Natural Science Foundation of Beijing (No. 6112013) to Wanzhi Cai; the Australian Research Council (DP120100240), the Australian Government (Australia-China Science & Research Fund Group Mission grant ACSRF00980) to Renfu Shao; and College of Agriculture, University of Kentucky. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rshao@usc.edu.au (RS); caiwz@cau.edu.cn (WC)

✉ These authors contributed equally to this work.

Introduction

The organization of mitochondrial (mt) genome is highly conserved in insects, as in most other bilateral animals [1,2]. With few exceptions, the mt genomes of insects consist of 13 protein-coding genes, two rRNA genes, 22 tRNA genes and a large non-coding region (also called the control region) on a single circular chromosome [1–3]. Arrangement of genes in mt genomes is usually stable; most insects known retained exactly the ancestral pattern of mt gene arrangement or have minor changes from the ancestral pattern of mt gene arrangement [1,4–8].

One group of insects that stands out and has major changes in the organization of mt genome is the superorder Psocodea. Psocodea has ~10,000 described species in two orders: Psocoptera (barklice and booklice) and Phthiraptera (parasitic lice) [9–12]. Complete mt genomes have been sequenced for two species of the Psocoptera and six species of the Phthiraptera [13–20]. Compared

to other insects, species of the Psocodea have three unusual features in their mt genomes that have not been found in any other insects. First, all of the eight species that have been sequenced have rearranged mt genomes. The booklouse, *Liposcelis bostrychophila*, and the six species of parasitic lice have the most rearranged mt genomes seen in insects - they differ at nearly every gene boundary from the putative ancestor of insects. Second, the booklouse, *L. bostrychophila* (Psocoptera: suborder Troctomorpha) and the parasitic lice in the suborder Anoplura have multipartite mt genomes. The mt genome of the booklouse, *L. bostrychophila*, has two mt chromosomes; each chromosome is 7–9 kb and has 16 to 22 genes [20]. The mt genomes of the human body louse, *Pediculus humanus*, and the human head louse, *Pediculus capitis*, have 20 minichromosomes; each minichromosome is 3–4 kb in size and contains one to three genes [14,16]. The mt genome of the human pubic louse, *Phthirus pubis*, has at least 14 minichromosomes; each minichromosome is 1.8–2.7 kb in size and contains one to five genes [16].

Table 1. Species of insects used in the phylogenetic analyses in the present study.

Order/suborder	Family	Species	Accession number	Reference
Psocoptera				
Trogiomorpha	Lepidopsocidae	lepidopsocid sp.	NC_004816	[13]
Psocomorpha	Psocidae	<i>Psococerastis albimaculata</i>	JQ910989	present study
		<i>Longivalvus hyalospilus</i>	JQ910986	present study
Troctomorpha	Liposcelidae	<i>Liposcelis bostrychophila</i>	JN645275 JN645276	[20]
Phthiraptera				
Ischnocera	Phlopteridae	<i>Bothriometopus macrocnemis</i>	NC_009983	[18]
		<i>Campanulotes bidentatus compar</i>	NC_007884	[17]
		<i>Ibidoecus bisignatus</i>	NC_015999	[19]
Amblycera	Boopidae	<i>Heterodoxus macropus</i>	NC_002651	[15]
Anoplura	Pediculidae	<i>Pediculus humanus</i>	FJ499473–90	[14]
		<i>Pthirus pubis</i>	EU219987–95, HM241895–8	[16]
Hemiptera				
	Nabidae	<i>Alloeorhynchus bakeri</i>	NC_016432	[28]
	Pentatomidae	<i>Halyomorpha halys</i>	NC_013272	[27]
Neuroptera				
	Chrysopidae	<i>Chrysoperla nipponensis</i>	NC_015093	[29]
Coleoptera				
	Carabidae	<i>Calosoma</i> sp.	NC_018339	[30]

doi:10.1371/journal.pone.0061685.t001

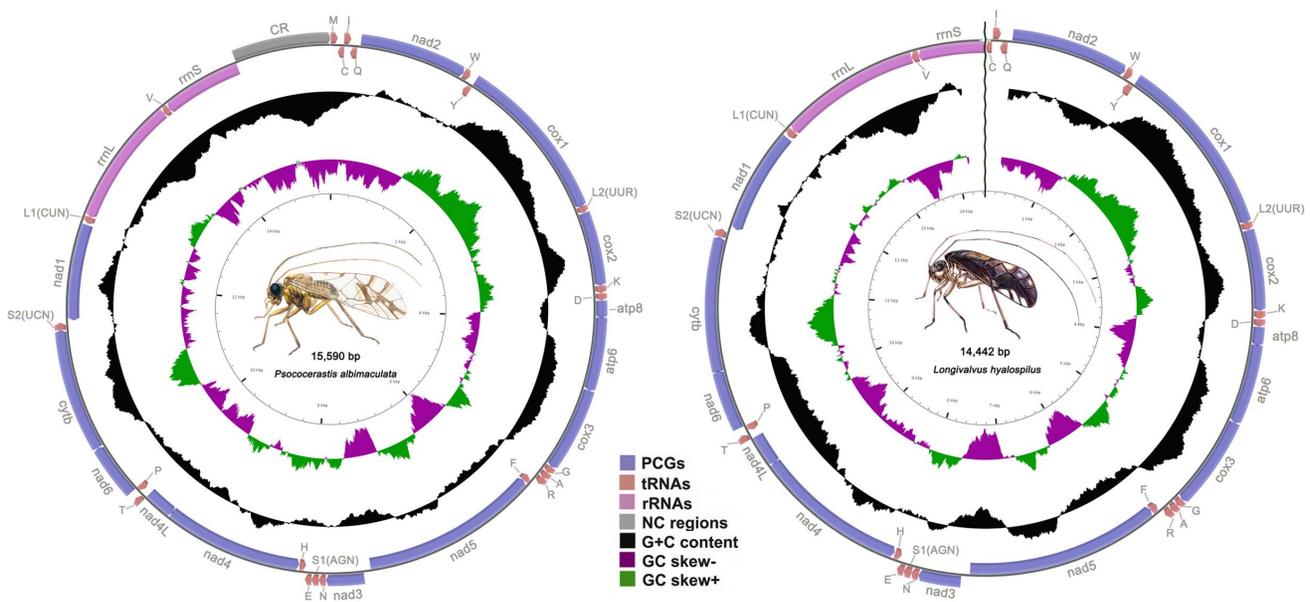


Figure 1. Mitochondrial genomes of the barklice, *Psococerastis albimaculata* and *Longivalvus hyalospilus*. Circular maps were drawn with CGView [42]. Arrows indicate the orientation of gene transcription. Protein-coding genes are shown as blue arrows, rRNA genes as purple arrows, tRNA genes as brown arrows and large non-coding regions (>100 bp) as grey rectangle. Abbreviations of gene names are: *atp6* and *atp8* for ATP synthase subunits 6 and 8, *cox1–3* for cytochrome oxidase subunits 1–3, *cytb* for cytochrome *b*, *nad1–6* and *nad4L* for NADH dehydrogenase subunits 1–6 and 4L, *rnl* and *rns* for large and small rRNA subunits. tRNA genes are indicated with their one-letter corresponding amino acids; the two tRNA genes for leucine and serine have different anticodons: *L1* (anticodon TAG), *L2* (TAA), *S1* (TCT) and *S2* (TGA). The GC content is plotted using a black sliding window, as the deviation from the average GC content of the entire sequence. GC-skew is plotted as the deviation from the average GC-skew of the entire sequence. The inner circle indicates the location of genes in the mt genome.

doi:10.1371/journal.pone.0061685.g001

Table 2. Codon usage of the protein-coding genes in the mitochondrial genomes of *Psococerastis albimaculata* and *Longivalvus hyalospilus*.

Codon	<i>P. albimaculata</i>		<i>L. hyalospilus</i>	
	count	%	count	%
A+T-rich codons ^a	1482	40.0	1509	41.0
G+C-rich codons ^b	473	12.8	480	13.0
Codon ratio ^c	0.13		0.13	

^aA+T-rich codons are those encoding amino acids Asn, Ile, Lys, Met, Phe, and Tyr.

^bG+C-rich codons are those encoding amino acids Ala, Arg, Gly, and Pro.

^cCodon ratio is G+C-rich codons/A+T-rich codons.

doi:10.1371/journal.pone.0061685.t002

Third, in the screamer louse, *Bothriometopus macrocnemis* (suborder Ischnocera), all mt genes are encoded by the same strand [18]. In the pigeon louse, *Campanulotes bidentatus compar* (suborder Ischnocera), all mt genes except *tmQ* are on the same strand [17].

The order Psocoptera has three suborders: Troctomorpha, Trogiomorpha and Psocomorpha. In addition to the booklouse, *L. bostrychophila* (suborder Troctomorpha), the complete mt genome of a barklouse, lepidopsocid sp. (suborder Trogiomorpha), has also been sequenced. The mt genome of the lepidopsocid sp. is much less rearranged than those of the booklouse and the parasitic lice; nevertheless, eight genes including a protein-coding gene have rearranged [13]. Prior to the present study, nothing is known about the mt genomes for species in the suborder Psocomorpha, which is the largest suborder of the Psocoptera, containing 25 of the 39 extant families and ~4,000 of the ~5,000 described species of the Psocoptera [10,21,22]. To further understand mt gene rearrangements and changes in mt genome organization in the Psocodea, we sequenced the mt genomes of two barklice, *Psococerastis albimaculata* and *Longivalvus hyalospilus*, the first representatives from the suborder Psocomorpha. We found that these barklice have the least rearranged mt genomes seen in the Psocodea to date. We show that there are contrasting rates in mt gene rearrangement between the two major lineages of the Psocodea.

Materials and Methods

Ethics Statement

No specific permits were required for the insects collected for this study. The insect specimens were collected from roadside vegetation by sweeping. The field collections did not involve

endangered or protected species. The species in the family of Psocidae are common insects and are not included in the “List of Protected Animals in China”.

Samples and DNA Extraction

Specimens of *P. albimaculata* and *L. hyalospilus* were collected in Kuankuoshui, Suiyang, Guizhou, China, in June 2010. Specimens were initially preserved in 95% ethanol in the field, and transferred to -20°C for long-term storage at the China Agricultural University (CAU). For each species, the genomic DNA was extracted from one male adult's muscle tissues of the thorax using the DNeasy DNA Extraction kit (Qiagen).

PCR Amplification and Sequencing

The mt genome was amplified by PCR in overlapping fragments with universal insect mt primers [23], and species-specific primers designed from sequenced fragments (Table S1). Short PCRs (<1.5 kb) were with Taq DNA polymerase (Qiagen); the cycling conditions were: 5 min at 94°C , followed by 35 cycles of 50 s at 94°C , 50 s at $48-55^{\circ}\text{C}$, 1–2 min at 72°C depending on the size of amplicons, and a final elongation step at 72°C for 10 min. Long PCRs (>1.5 kb) were with Long Taq DNA polymerase (New England BioLabs); the cycling conditions were: 30 s at 95°C , followed by 40 cycles of 10 s at 95°C , 50 s at $48-55^{\circ}\text{C}$, 3–6 min at 68°C depending on the size of amplicons, and a final elongation step at 68°C for 10 min. The concentration and size of PCR products were measured by spectrophotometry and agarose gel electrophoresis. PCR fragments were ligated into the pGEM-T Easy Vector (Promega); the resulting plasmid DNAs were isolated using the TIANprp Midi Plasmid Kit (Qiagen). All fragments were sequenced in both directions with an ABI 3730XL Genetic Analyzer, using the BigDye Terminator Sequencing Kit (Applied Biosystems) with two vector-specific primers and internal primers for primer walking.

Assembly, Annotation and Bioinformatics Analysis

Sequence reads from the mt genome of each barklouse species were assembled into contigs with Sequencher (Gene Codes). Protein-coding genes and rRNA genes were identified by BLAST searches in GenBank and then confirmed by alignment with homologous genes from other insects. tRNA genes were identified with tRNAscan-SE v.1.21 [24]. *tmR* and *tmSI*, which could not be identified by tRNAscan-SE, were determined by sequence similarity comparison with tRNA genes of other insects. The base composition, codon usage, and nucleotide substitution were analyzed with Mega 5.0 [25]. Secondary structures of stem-loop in control region were folded using Mfold [26].

Table 3. Nucleotide composition at each codon position of the protein-coding genes in the mitochondrial genomes of *Psococerastis albimaculata* and *Longivalvus hyalospilus*.

Codon	AT %		GC %		AT-skew		GC-skew	
	Pa	Lh	Pa	Lh	Pa	Lh	Pa	Lh
All codon	75.0	72.3	25.0	27.7	-0.14	-0.14	-0.03	-0.05
1st position	69.8	68.2	30.2	31.8	-0.009	0.005	0.20	0.16
2nd position	68.3	68.0	31.7	32.0	-0.39	-0.40	-0.11	-0.13
3rd position	87.0	80.6	13.0	19.4	-0.05	-0.03	-0.35	-0.25

Pa, *P. albimaculata*; Lh, *L. hyalospilus*; AT-skew = $(A - T)/(A + T)$; GC-skew = $(G - C)/(G + C)$.

doi:10.1371/journal.pone.0061685.t003

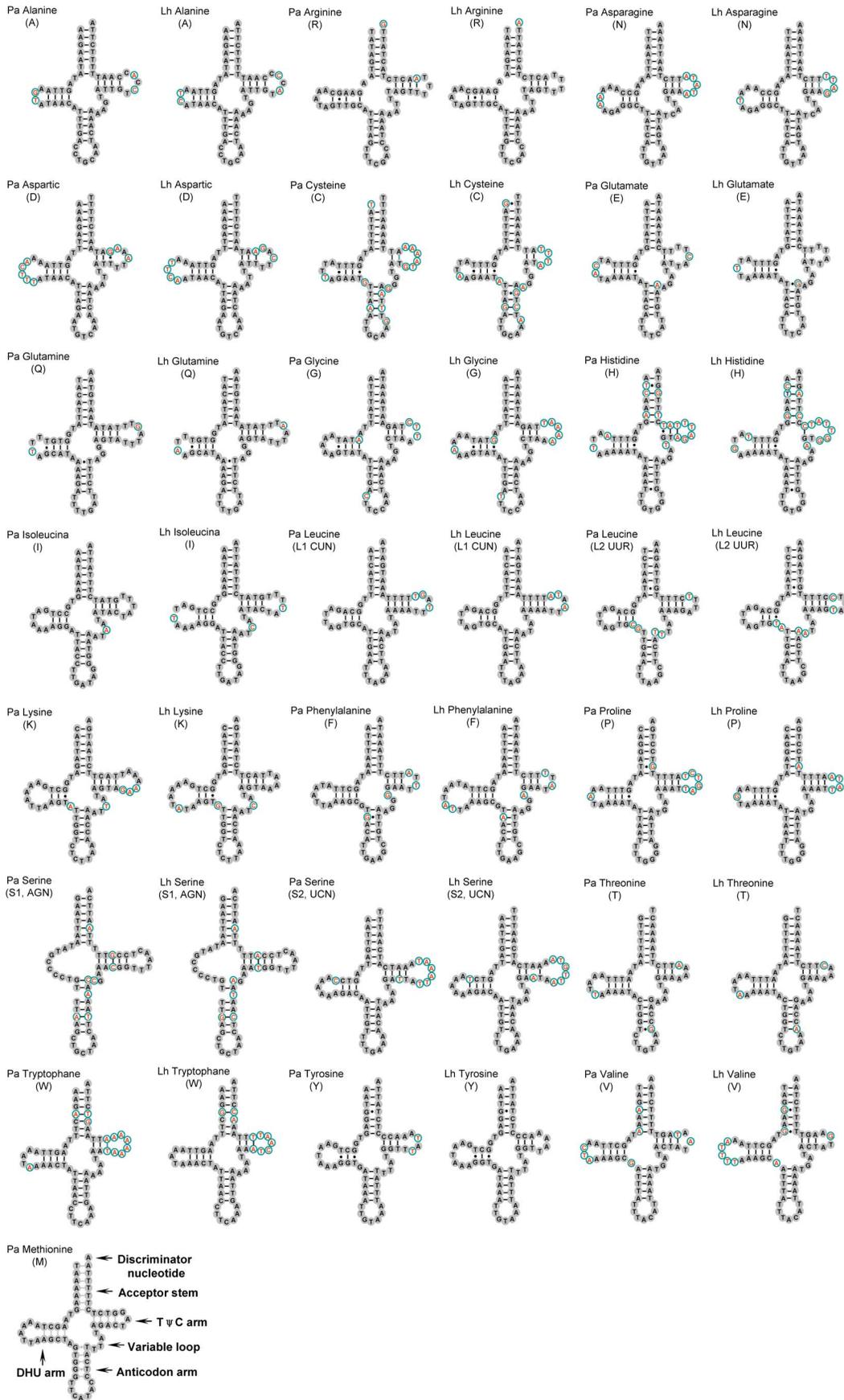


Figure 2. Inferred secondary structure of the mitochondrial tRNAs of the barklice, *Psococerastis albimaculata* and *Longivalvus hyalospilus*. tRNAs are labeled with the abbreviations of their corresponding amino acids. Lines indicate Watson-Crick bonds; dots indicate GU bonds. Green circle with red letter inside highlight variations of nucleotides in tRNA between two barklice. doi:10.1371/journal.pone.0061685.g002

Sequence Alignment

Four species from the Psocoptera and six species from the Phthiraptera were included in our phylogenetic analyses (Table 1). These species are: 1) three barklice, *P. albimaculata*, *L. hyalospilus* and lepidopsocid sp. [13]; 2) a booklouse, *L. bostrychophila* [20]; 3) four chewing lice, *Bothriometopus macrocnemis* [18], *Campanulotes bidentatus compar* [17], *Ibidocus bisignatus* [19], and *Heterodoxus macropus* [15]; 4) the human body louse, *Pediculus humanus* [14]; and 5) the human pubic louse, *Phthirus pubis* [16]. Two true bugs, *Alloeorhynchus bakeri* and *Halyomorpha halys* (Hemiptera) [27,28], the lacewing, *Chrysoperla nipponensis* (Neuroptera) [29], and the ground beetle, *Calosoma* sp. (Coleoptera) [30], were used as outgroups.

Sequences of all mt protein-coding genes and rRNA genes except *nad4* were used in phylogenetic analyses; *nad4* was excluded because it was not identified in the human pubic louse, *P. pubis* [16]. Segments of identical sequences (26–127 bp long) shared between five pairs of mt genes in the human body louse, *P. humanus*, and the human pubic louse, *P. pubis* [14,16], were also excluded to ensure only homologous regions of the mt genes were aligned and used in subsequent phylogenetic analyses. Alignment of the nucleotide sequences of each protein-coding gene and its putative amino acid sequence was with MUSCLE [31], adjusted to preserve the reading frame. Sequences of each rRNA gene were aligned with the GUIDANCE algorithm [32,33], adjusted to its RNA secondary structure [34]. The alignments of individual genes were concatenated after removing poorly aligned sites using Gblocks 0.91 [35].

Phylogenetic Analysis

Three alignments were used for phylogenetic analyses: 1) a concatenated nucleotide sequence alignment of protein-coding genes and two rRNA genes (PCG123R); 2) a concatenated nucleotide sequence alignment of the first and the second codon positions of protein-coding genes and two rRNA genes (PCG12R); and 3) a concatenated amino acid sequence alignment of protein-coding genes (AA). Partitioned ML and Bayesian analyses were run with PCG123R, PCG12R and AA matrix, using RAxML 7.0.3 [36] and MrBayes 3.2.1 [37]. The best-fit model for the amino acid sequence alignment was determined with ProtTest [38], and the jModelTest 0.1.1 [39] was used for the nucleotide sequence of each gene, according to the Akaike Information Criterion (AIC). For the ML analyses, GTRMIX option for

nucleotide sequence and MtREV model for amino acid sequence were used to optimize the topology. For the combined dataset, 1,000 independent runs from random starting trees were performed to find the highest scoring replicate. Node support was calculated by acquiring bootstrap values from heuristic searches of 1000 resampled datasets, using the rapid bootstrap feature of RAxML [40].

For Bayesian analyses, the most appropriate substitution model was GTR+I+G model for each protein-coding gene, 1st and 2nd codon positions of protein-coding genes, and *rrnL* gene; GTR+G model for *rrnS* gene; and MtREV model for amino acid sequence of each protein-coding genes. Two simultaneous runs of 10 million generations were conducted for the matrix and trees were sampled every 1,000 generations, with the first 25% discarded as burn-in. Stationarity was considered to be reached when the average standard deviation of split frequencies was below 0.01 [41].

Results and Discussion

Mitochondrial Genomes of the Barklice, *Psococerastis Albimaculata* and *Longivalvus Hyalospilus*

The mt genome of *P. albimaculata* contains the entire set of 37 genes (13 protein-coding genes, 22 tRNA genes, and two rRNA genes; Figure 1 and Table S2) and a putative control region that are usually present in animal mt genomes [1,3]. We found the same set of mt genes in *L. hyalospilus*, except the control region and the adjacent *trnM* gene due to our unsuccess to amplify this region (Figure 1 and Table S3). These two barklice have the same arrangement of mt genes to each other but differs from the putative ancestor of insects: a protein-coding gene (*nad3*) and five tRNA genes (*trnN*, *trnS1*, *trnE*, *trnM* and *trnC*) have translocated. Genes are encoded by both strands in the mt genomes of these two barklice: 14 genes on one strand whereas the rest on the other strand (Table S2 and S3). Thirteen pairs of adjacent genes in the mt genomes of these two barklice overlap by 1 to 16 bp. All of the protein-coding genes of the two barklice start with ATN codons and stop with TAA/TAG codons or truncated codons TA or T.

The nucleotide compositions of the mt genomes of the two barklice are biased toward A and T. The nucleotide skew statistics for the entire majority strand indicate moderate A skew and obvious C skew, and the coding strand of protein-coding genes display a moderate T skew and slight C skew (Table 2). The A+T-

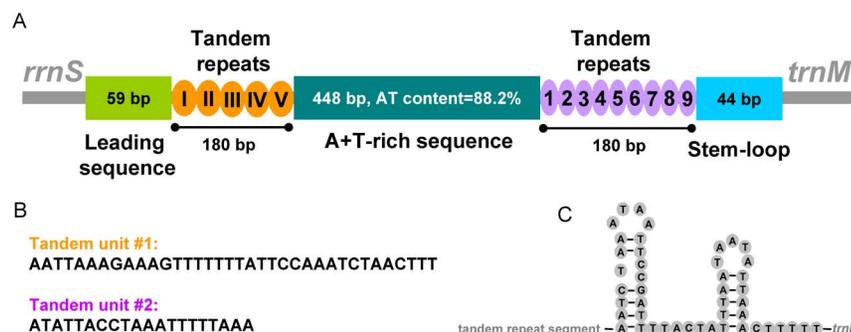


Figure 3. The control region in the mitochondrial genome of the barklouse, *Psococerastis albimaculata*. A) Structural elements in the control region; B) sequences of the tandem-repeat units; C) predicted stem-loop structures. doi:10.1371/journal.pone.0061685.g003

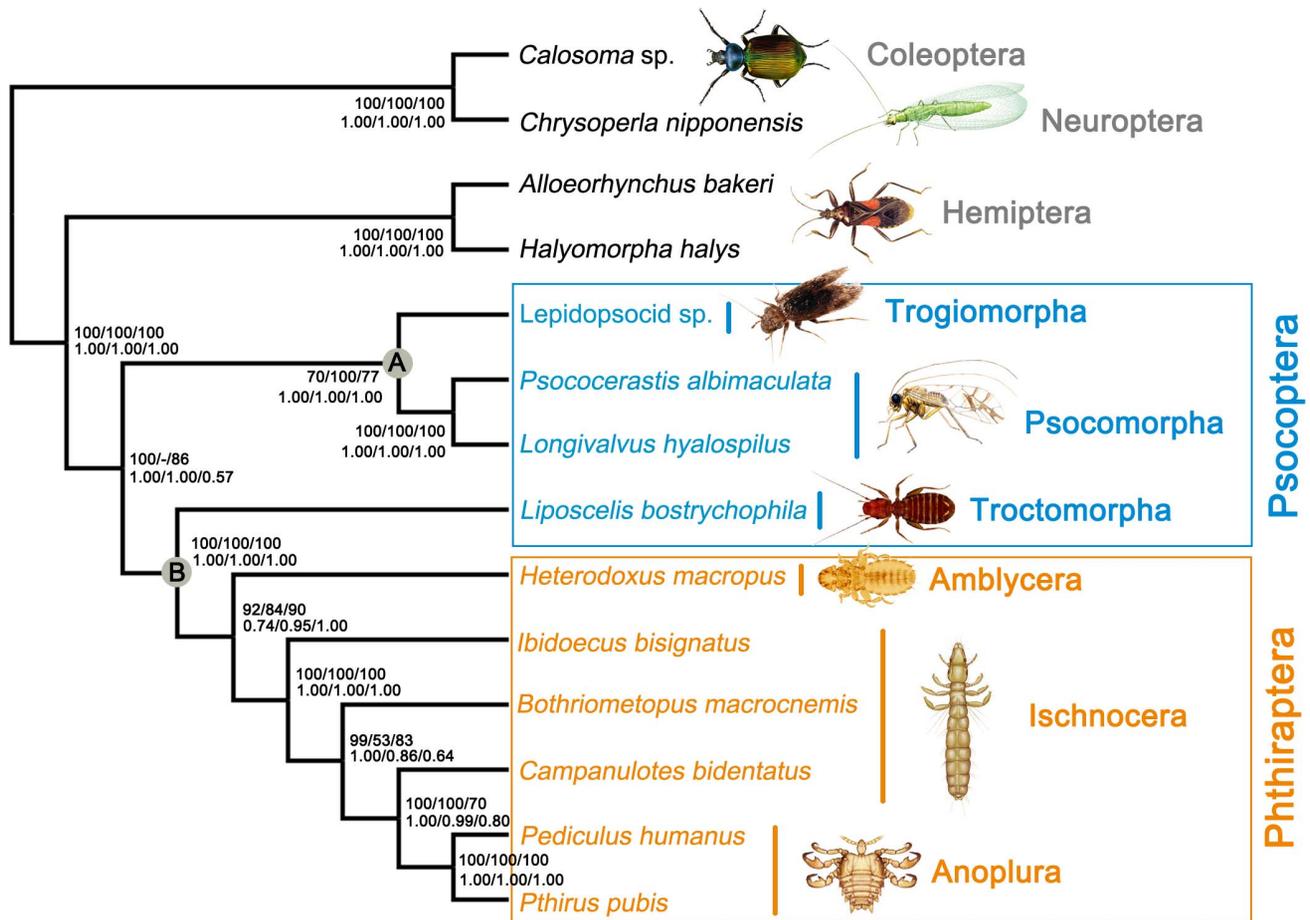


Figure 4. Phylogenetic relationships among major lineages of the Psocodea inferred from mitochondrial genome sequences. Numbers close to the branching points are ML bootstrap support values (above) and Bayesian posterior probabilities (below) in percentages; only support above 50% is shown. Numbers from left to right are from PCG123R, PCG12R and AA alignments respectively. doi:10.1371/journal.pone.0061685.g004

richness of mt genomes of these two barklice is reflected further in the codon usage (Table S4). The overall ratio of G+C-rich codons to A+T-rich codons is 0.13 in the two barklice and there is a strong bias to A+T at the third codon positions of the protein-coding genes (Table 3).

The multiple alignment of 21 tRNA genes (excluded *tmM*) extends over 1,366 positions and contains 1,229 conserved nucleotides (90.0%) between the two barklice. Nucleotides at the stems and anticodon loops were conserved; variations were largely restricted to the loop of T Ψ C and DHU arm (Figure 2). The nucleotide sequences of the mt genomes of two barklice, *P. albimaculata* and *L. hyalospilus*, have been deposited in GenBank under accession numbers JQ910989 and JQ910986.

The Control Region in the Mitochondrial Genome of the Barklouse, *Psococerastis albimaculata*

The putative control region (911 bp) was flanked by *rmS* and *tmM* in the mt genome of *P. albimaculata*. The control region is highly AT-rich (86.9%; majority strand) and can be divided into five parts (Figure 3A): 1) 59 bp leading sequence; 2) 180 bp tandem repeat sequences consisting of five 36 bp repeat units (Figure 3B); 3) 448 bp A+T-rich sequences (A+T 88.2%); 4) 180 bp tandem repeated sequences consisting of nine 20 bp repeat units (Figure 3B); and 5) 44 bp stem-loop structure (Figure 3C). Stem-loops are common in the control regions of animal mt

genomes and may have roles in the initiation of gene transcription and DNA replication [43–47]. The pattern of two tandem repeated sequences with an A+T-rich sequence in between is also present in the control region of the other barklouse, lepidopsocid-RS (Trogiomorpha), but not in the booklouse, *L. bostrychophila* (Troctomorpha), nor in any parasitic lice.

Outside the control region, there are 146 bp non-coding sequences in 13 intergenic regions of *P. albimaculata*, and 178 bp non-coding sequences in 12 intergenic regions in *L. hyalospilus*. Most non-coding sequences are scattered in short runs (1–16 bp in *P. albimaculata* and 1–26 bp in *L. hyalospilus*). However, two of these non-coding sequences are longer in length and locate in the same intergenic regions in the two barklice: 1) between *tmQ* and *nad2* (29 bp in *P. albimaculata* and 38 bp in *L. hyalospilus*); and 2) between *nad5* and *nad3* (71 bp in *P. albimaculata* and 73 bp in *L. hyalospilus*). These two non-coding sequences are in the regions where gene rearrangement occurred and are likely generated from events of tandem duplication followed by random deletions (See below).

Phylogenetic Relationships among Major Lineages of the Psocodea Inferred from Mitochondrial Genome Sequences

The Psocoptera (booklice and barklice) and the Phthiraptera (parasitic lice) have traditionally been recognized as two separate orders [12,48,49]. The Psocoptera (booklice and barklice) are free-

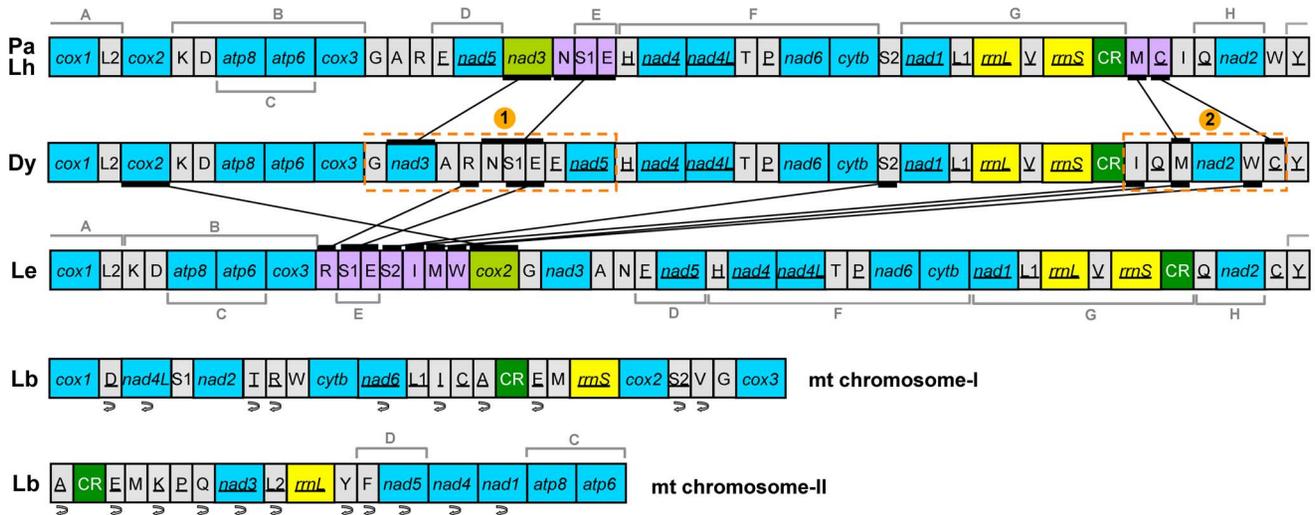


Figure 5. Comparison of mitochondrial gene arrangement between barklice, booklice and the hypothetical ancestor of insects. Abbreviations of species names are: Pa for *P. albimaculata*, Lh for *L. hyalospilus*, Dy for *Drosophila yakuba*, Le for lepidopsocid, and Lb for *L. bostrychophila*. Abbreviations of gene names follow Figure 1. Genes are transcribed from left to right except those underlined, which have the opposite transcriptional orientation. Pa and Lh have the same set of mt genes and gene arrangement, except the control region and the adjacent *trnM* gene due to our unsuccessful attempt to amplify this region in Lh. Shared gene boundaries are indicated by brackets and letter codes: A) *trnY-cox1-trnL2*; B) from *trnK* to *cox3*; C) *atp8-atp6*; D) *trnF-nad5*; E) *trnS1-trnE*; F) from *trnH* to *cytb*; G) from *nad1* to *CR*; and H) *trnQ-nad2*. Two ancestral gene blocks highlighted by the orange; broken lines indicate the inferred blocks that had duplicated in tandem. The circling arrows indicate inversions of mt genes relative to the hypothetical ancestor of insects. doi:10.1371/journal.pone.0061685.g005

living insects and consist of over 5,000 species with a world-wide distribution, and are divided into three suborders: Trogiomorpha, Troctomorpha and Psocomorpha [10,22]. Members of the order Phthiraptera (~4,900 species) are wingless insects, parasitic on birds and mammals. There are four recognized suborders in the Phthiraptera: Amblycera, Ischnocera, Anoplura and Rhynchophthirina [9,11,50]. Both morphological and molecular analyses indicate a close relationship between parasitic lice (Phthiraptera) and booklice (family Liposcelididae); the order Psocoptera is therefore paraphyletic [9,12,20,51–55]. The monophyly of the Phthiraptera, however, has been controversial. A number of shared morphological and physiological characters support the monophyly of the Phthiraptera [9,48,49]. Analyses of mt 12S and 16S rDNA and nuclear 18S rDNA sequences, however, indicate that the parasitic lice are paraphyletic: the suborder Amblycera is more closely related to the booklouse family Liposcelididae than to the other three suborders of the parasitic lice [12,51,54].

We tested the phylogenetic relationships among the major lineages of the Psocodea with the mt genome sequences of *P. albimaculata* and *L. hyalospilus*, and eight other species of the Psocodea. We used three different datasets: 1) a concatenated nucleotide sequence alignment of protein-coding genes and two rRNA genes (PCG123R); 2) a concatenated nucleotide sequence alignment of the first and the second codon positions of protein-coding genes and two rRNA genes (PCG12R); and 3) a concatenated amino acid sequence alignment of protein-coding genes (AA). All three datasets were run with partitioned ML and Bayesian analyses based on best-fit models. We recovered two major clades in the Psocodea with strong support regardless the dataset and the method we used: 1) species of barklice in the suborders Psocomorpha and Trogiomorpha were clustered together (Clade A, Figure 4); and 2) the booklouse, *L. bostrychophila* (suborder Troctomorpha), formed a clade with the parasitic lice (Clade B, Figure 4). The parasitic lice (Phthiraptera) are

monophyletic with strong support; within the parasitic lice, the suborder Ischnocera, however, is paraphyletic.

Contrasting Rates in Mitochondrial Gene Rearrangement between Two Major Lineages of the Psocodea

Seven species from the Clade B above have been sequenced for complete mt genomes previously, i.e., the booklouse, *L. bostrychophila* (suborder Troctomorpha) [20] and six parasitic lice [14–19]. All of these seven species have extremely rearranged mt genomes, having little similarity in gene arrangement with each other, nor with any other insects. Both translocations and inversions occurred in these species relative to the ancestral gene arrangement of insects. Only four ancestral gene boundaries of insects (*trnG-nad3*, *trnL1-trnL*, *nad4-nad4L*, and *atp8-atp6*) were found in the Clade B species, of which only *atp8-atp6* was retained by all of the seven species in the Clade B. Apparently, mt gene rearrangement has been occurring much more often in the Clade B species than in other insects.

In contrast, species from the Clade A investigated to date have much less rearranged mt genomes and retained most of the ancestral gene arrangements of insects [13]. In particular, the two barklice we sequenced in the present study, *P. albimaculata* and *L. hyalospilus*, have the least rearranged mt genomes seen in the Psocodea: a protein-coding gene (*nad3*) and five tRNAs (*trnV*, *trnS1*, *trnE*, *trnM* and *trnC*) have translocated relative to that of the ancestor of insects (Figure 5). Gene rearrangement in these two barklice occurred in the two “hot spot” regions for gene rearrangement [56]: 1) between *cox3* and *trnH*, and 2) between *CR* and *trnY* (Figure 5). Two events of tandem duplication followed by random deletions could account for, straightforwardly, the rearrangement of mt genes observed in *P. albimaculata* and *L. hyalospilus* (Figure 6). The number of mt genes that have rearranged in *P. albimaculata* and *L. hyalospilus* is even less than in the other barklouse, lepidopsocid-RS (*cox2* and seven tRNAs rearranged) [13]. Together, 20 gene boundaries (Figure 5, seven

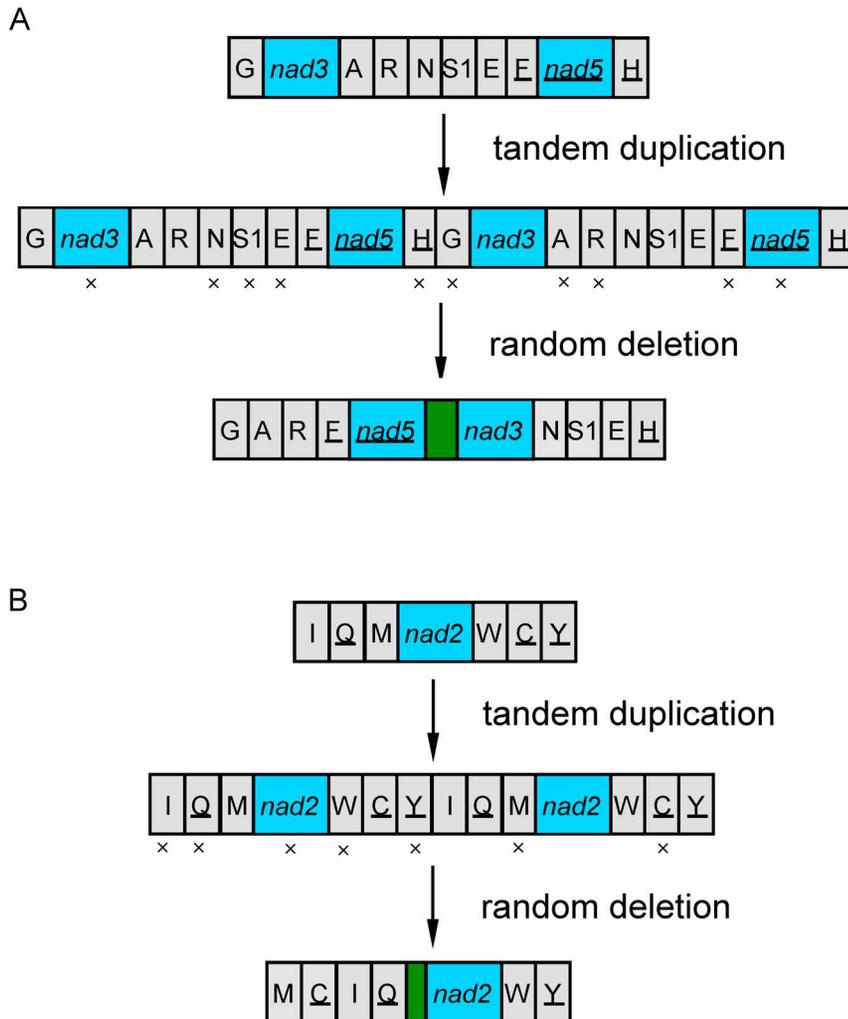


Figure 6. Inferred tandem-duplication-and-random-deletion events that account for the mitochondrial gene rearrangements in the barklice, *Psococerastis albimaculata* and *Longivalvus hyalospilus*. A) genes between *cox3* and *nad4*; B) genes between *CR* and *cox1*. Two longer non-coding sequences are highlighted by green color.
doi:10.1371/journal.pone.0061685.g006

gene blocks: A, B, D, E, F, G, and H) are shared and 32 ancestral gene boundaries are retained in these three barklice.

Tandem duplication followed by random deletion model has been proposed to account for mt gene rearrangement [8,57–59]. This model can explain the mt gene rearrangements observed in the two barklice, *P. albimaculata* and *L. hyalospilus* (Figure 6). The rearranged mt gene blocks of these two barklice, one from *tmG* to *tmH* and another from *tmM* to *tmY* can be generated by a single event of tandem duplication of the ancestral gene block from *tmG* to *tmH* (Figure 6A) and from *tmI* to *tmY* (Figure 6B), respectively, and followed by random deletion of excess genes. The non-coding sequences present in the two rearranged gene boundaries (between *tmQ* and *nad2*, and between *nad5* and *nad3*) are likely traces of the random deletion. The mt gene rearrangement in the other barklouse lepidopsocid-RS (suborder Trogiomorpha) is more complicated than in *P. albimaculata* and *L. hyalospilus*; *cox2* gene and four tRNA genes have translocated from long distance and cannot be accounted for alone by tandem duplication followed by random deletion model. The mt gene rearrangements in the booklice and parasitic lice are much more complicated than in the

barklice; multiple mechanisms and frequent rearrangement events are likely involved [16,60,61].

What caused the substantial difference between the two clades of the Psocodea in the rates of mt gene rearrangement? An obvious difference between the two clades is the lifestyle. The barklice in the Clade A are entirely free-living insects, which often feed on fungal spores; whereas Clade B (booklice and parasitic lice) is a mixture of short-term commensal and parasitic. The booklouse, *L. bostrychophila*, is mainly an inhabitant of households and a major pest to stored grains world-wide; moreover, there are many records of various species of booklice in the plumage of birds and the pelage of mammals, as well as in their nests [62,63]. This association is believed to be a short-term commensal (non-parasitic) relationship, which may have given rise to a parasitic and permanent association [64]. All parasitic lice (Phthiraptera) feed on the skin, skin debris or blood of their vertebrate hosts and spend their entire life cycle on the body of the host [12]. The lifestyle change in the Clade B appears to be associated with an increased rate of mt gene rearrangement. However, why they are linked and exactly what biological and lifestyle factors contributed to the contrasting rates in mt gene rearrangement between the two major

lineages of the Psocodea are not yet clear to us and remains to be determined.

Supporting Information

Table S1 Primers used in the present study.
(DOC)

Table S2 Genes in the mitochondrial genome of the barklouse, *Psococerastis albimaculata*.
(DOC)

Table S3 Genes in the mitochondrial genome of the barklouse, *Longivalvus hyalospilus*.
(DOC)

References

- Boore JL (1999) Animal mitochondrial genomes. *Nucleic Acids Res* 27: 1767–1780.
- Lavrov DV (2007) Key transitions in animal evolution: a mitochondrial DNA perspective. *Integr Comp Biol* 47: 734–743.
- Wolstenholme DR (1992) Animal mitochondrial DNA: structure and evolution. *Int Rev Cytol* 141: 173–216.
- Dowton M, Castro LR, Austin AD (2002) Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome ‘morphology’. *Invertebr Syst* 16: 345–356.
- Shao R, Barker SC (2007) Mitochondrial genomes of parasitic arthropods: implications for studies of population genetics and evolution. *Parasitology* 134: 153–167.
- Cameron SL, Sullivan J, Song H, Miller KB, Whiting MF (2009) A mitochondrial genome phylogeny of the Neuropterida (lace-wings, alderflies and snakeflies) and their relationship to the other holometabolous insect orders. *Zool Scr* 38: 575–590.
- Beckenbach AT (2012) Mitochondrial genome sequences of Nematocera (lower Diptera): evidence of rearrangement following a complete genome duplication in a winter crane fly. *Genome Biol Evol* 4: 89–101.
- Li H, Liu H, Shi AM, Štys P, Zhou XG, et al. (2012) The complete mitochondrial genome and novel gene arrangement of the unique-headed bug *Stenopirates* sp. (Hemiptera: Enicocephalidae). *PLoS ONE* 7: e29419.
- Lyal CHC (1985) Phylogeny and classification of the Psocodea, with particular reference to the lice (Psocodea: Phthiraptera). *Syst Entomol* 10: 145–165.
- Lienhard C, Smithers CN (2002) Psocoptera: world catalogue and bibliography. Geneva: Muséum d’Histoire Naturelle Press.
- Barker SC, Whiting M, Johnson KP, Murrell A (2003) Phylogeny of the lice (Insecta: Phthiraptera) inferred from small subunit rRNA. *Zool Scr* 32: 407–414.
- Johnson KP, Yoshizawa K, Smith VS (2004) Multiple origins of parasitism in lice. *Proc R Soc Lond B* 271: 1771–1776.
- Shao R, Dowton M, Murrell A, Barker SC (2003) Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. *Mol Biol Evol* 20: 1612–1619.
- Shao R, Kirkness EF, Barker SC (2009) The single mitochondrial chromosome typical of animals has evolved into 18 minichromosomes in the human body louse, *Pediculus humanus*. *Genome Res* 19: 904–912.
- Shao R, Campbell NJH, Barker SC (2011) Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). *Mol Biol Evol* 18: 858–865.
- Shao R, Zhu XQ, Barker SC, Herd K (2012) Evolution of extensively fragmented mitochondrial genomes in the lice of humans. *Genome Biol Evol* 4: 1088–1101.
- Covacin C, Shao R, Cameron S, Barker SC (2006) Extraordinary number of gene rearrangements in the mitochondrial genomes of lice (Phthiraptera: Insecta). *Insect Mol Biol* 15: 63–68.
- Cameron SL, Johnson KP, Whiting MF (2007) The mitochondrial genome of the screamer louse *Bothriometopus* (Phthiraptera: Ischnocera): effects of extensive gene rearrangements on the evolution of the genome. *J Mol Evol* 65: 589–604.
- Cameron SL, Yoshizawa K, Mizukoshi A, Whiting MF, Johnson KP (2011) Mitochondrial genome deletions and minicircles are common in lice (Insecta: Phthiraptera). *BMC Genomics* 12: 394.
- Wei DD, Shao R, Yuan ML, Dou W, Barker SC, et al. (2012) The multipartite mitochondrial genome of *Liposcelis bostrychophila*: insights into the evolution of mitochondrial genomes in bilateral animals. *PLoS ONE* 7: e33973.
- Yoshizawa K (2002) Phylogeny and higher classification of suborder Psocomorpha (Insecta: Psocoda: ‘Psocoptera’). *Zool J Linn Soc* 136: 371–400.
- Johnson KP, Mockford EL (2003) Molecular systematics of Psocomorpha (Psocoptera). *Syst Entomol* 28: 409–440.
- Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT (2006) Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annu Rev Ecol Syst* 37: 545–579.
- Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25: 955–964.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
- Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31: 3406–3415.
- Lee W, Kang J, Jung C, Hoelmer K, Lee SH, et al. (2009) Complete mitochondrial genome of brown marmorated stink bug *Halyomorpha halys* (Hemiptera: Pentatomidae), and phylogenetic relationships of hemipteran suborders. *Mol Cells* 28: 155–165.
- Li H, Liu HY, Cao LM, Shi AM, Yang HL, et al. (2012) The complete mitochondrial genome of the damselfly bug *Alloerhynchus bakeri* (Hemiptera: Nabidae). *Int J Biol Sci* 8: 93–107.
- Haruyama N, Mochizuki A, Sato Y, Naka H, Nomura M (2011) Complete mitochondrial genomes of two green lacewings, *Chrysoperla nipponensis* (Okamoto, 1914) and *Apochrysa matsumurae* Okamoto, 1912 (Neuroptera: Chrysopidae). *Mol Biol Rep* 38: 3367–3373.
- Song H, Sheffield NC, Cameron SL, Miller KB, Whiting MF (2010) When phylogenetic assumptions are violated: base compositional heterogeneity and among-site rate variation in beetle mitochondrial phylogenomics. *Syst Entomol* 35: 429–448.
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792–1797.
- Penn O (2010) GUIDANCE: a web server for assessing alignment confidence scores. *Nucleic Acids Res* 38: W23–W28.
- Penn O, Privman E, Landan G, Graur D, Pupko T (2010) An alignment confidence score capturing robustness to guide-tree uncertainty. *Mol Biol Evol* 27: 1759–1767.
- Li H, Liu H, Song F, Shi A, Zhou X, et al. (2012) Comparative mitogenomic analysis of damselfly bugs representing three tribes in the family Nabidae (Insecta: Hemiptera). *PLoS ONE* 7: e45925.
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17: 540–552.
- Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
- Abascal F, Zardoya R, Posada D (2007) ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics* 24: 1104–1105.
- Posada D (2008) jModelTest: Phylogenetic model averaging. *Mol Biol Evol* 25: 1253–1256.
- Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 57: 758–771.
- Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294: 2310–2314.
- Stothard P, Wishart DS (2005) Circular genome visualization and exploration using CGView. *Bioinformatics* 21: 537–539.
- Clayton DA (1982) Replication of animal mitochondrial DNA. *Cell* 28: 693–705.
- Clayton DA (1991) Replication and transcription of vertebrate mitochondrial DNA. *Annu Rev Cell Biol* 7: 453–478.
- Clary DO, Wolstenholme DR (1987) *Drosophila* mitochondrial DNA: Conserved sequences in the A+T-rich region and supporting evidence for a secondary structure model of the small ribosomal RNA. *J Mol Evol* 25: 116–125.
- Zhang DX, Szymura JM, Hewitt GM (1995) Evolution and structural conservation of the control region of insect mitochondrial DNA. *J Mol Evol* 40: 382–391.

Table S4 Codon usage in the protein-coding genes of the mitochondrial genomes of the barklice, *Psococerastis albimaculata* and *Longivalvus hyalospilus*.
(DOC)

Acknowledgments

We thank Prof. Fasheng Li and Dr. Luxi Liu (China Agricultural University) for their help in the insect identification.

Author Contributions

Conceived and designed the experiments: HL WZC. Performed the experiments: HL FS. Analyzed the data: HL RS FS. Contributed reagents/materials/analysis tools: RS WZC. Wrote the paper: HL RS XGZ QQY ZHL WZC.

47. Li H, Gao JY, Liu HY, Liu H, Liang AP, et al. (2011) The architecture and complete sequence of mitochondrial genome of an assassin bug *Agriosphodrus dohmi* (Hemiptera: Reduviidae). *Int J Biol Sci* 7: 792–804.
48. Hennig W (1966) Phylogenetic systematics. Urbana: University of Illinois Press.
49. Jamieson BGM, Dallai R, Afzelius BA (1999) Insects: Their spermatozoa and phylogeny. Enfield: Science Publisher.
50. Clay T (1970) The Amblycera (Phthiraptera: Insecta). *Bull Br Mus (Nat Hist Entomol* 25: 73–98.
51. Yoshizawa K, Johnson KP (2003) Phylogenetic position of Phthiraptera (Insecta: Paraneoptera) and elevated rate of evolution in mitochondrial 12S and 16S rDNA. *Mol Phylogenet Evol* 29: 102–114.
52. Yoshizawa K, Johnson KP (2006) Morphology of male genitalia in lice and their relatives and phylogenetic implications. *Syst Entomol* 31: 350–361.
53. Yoshizawa K, Johnson KP (2010) How stable is the 'Polyphyly of Lice' hypothesis (Insecta: Psocodea)? A comparison of phylogenetic signal in multiple genes. *Mol Phylogenet Evol* 55: 939–951.
54. Murrell A, Barker SC (2005) Multiple origins of parasitism in lice: phylogenetic analysis of SSU rDNA indicates that the Phthiraptera and Psocoptera are not monophyletic. *Parasitol Res* 97: 274–280.
55. Yoshizawa K, Lienhard C (2010) In search of the sister group of the true lice: a systematic review of booklice and their relatives, with an updated checklist of Liposcelididae (Insecta: Psocodea). *Arthropod Syst Phyl* 68: 181–195.
56. Dowton M, Austin AD (1999) Evolutionary dynamics of a mitochondrial rearrangement 'Hot Spot' in the Hymenoptera. *Mol Biol Evol* 16: 298–309.
57. Moritz C, Brown WM (1986) Tandem duplication of D-loop and ribosomal RNA sequences in lizard mitochondrial DNA. *Science* 308: 1425–1427.
58. Boore JL (2000) The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animal. In: Sankoff D, Nadeau JH, editor. *Comparative genomics*. Dordrecht: Kluwer Academic Press. p.133–147.
59. Shao R, Barker SC, Mitani H, Takahashi M, Fukunaga M (2006) Molecular mechanisms for the variation of mitochondrial gene content and gene arrangement among chigger mites of the genus *Leptotrombidium* (Acari: Acariformes). *J Mol Evol* 63: 251–261.
60. Dowton M, Campbell NJH (2001) Intramitochondrial recombination – is it why some mitochondrial genes sleep around? *Trends Ecol Evol* 16: 269–271.
61. Kraysberg Y, Schwartz M, Brown TA, Ebralidse K, Kunz WS, et al. (2004) Recombination of human mitochondrial DNA. *Science* 304: 981.
62. Pearlman JV (1960) Some African Psocoptera found on rats. *Entomologist* 93: 246–250.
63. Mockford EL (1967) Some Psocoptera from the plumage of birds. *Proc Entomol Soc Wash* 69: 307–309.
64. Hopkins GHE (1949) The host associations of the lice of mammals. *Proc Zool Soc Lond* 119: 387–604.