The Mitochondrial Genome of the Guanaco Louse, *Microthoracius praelongiceps*: Insights into the Ancestral Mitochondrial Karyotype of Sucking Lice (Anoplura, Insecta)

Renfu Shao[1,*], Hu Li[2], Stephen C. Barker[3], and Simon Song[1,*]

[1]GeneCology Research Centre, School of Science and Engineering, Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Maroochydore, Queensland, Australia

[2]Department of Entomology, China Agricultural University, Beijing, China

[3]Parasitology Section, School of Chemistry and Molecular Biosciences, University of Queensland, St Lucia, Queensland, Australia

*Corresponding authors: E-mails: rshao@usc.edu.au; depings@yahoo.com.

Accepted: January 31, 2017

Data deposition: The nucleotide sequence of the mitochondrial genome of the guanaco louse, *Microthoracius praelongiceps*, has been deposited at GenBank (accession numbers KX090378–KX090389).

Abstract

Fragmented mitochondrial (mt) genomes have been reported in 11 species of sucking lice (suborder Anoplura) that infest humans, chimpanzees, pigs, horses, and rodents. There is substantial variation among these lice in mt karyotype: the number of minichromosomes of a species ranges from 9 to 20; the number of genes in a minichromosome ranges from 1 to 8; gene arrangement in a minichromosome differs between species, even in the same genus. We sequenced the mt genome of the guanaco louse, *Microthoracius praelongiceps*, to help establish the ancestral mt karyotype for sucking lice and understand how fragmented mt genomes evolved. The guanaco louse has 12 mt minichromosomes; each minichromosome has 2–5 genes and a non-coding region. The guanaco louse shares many features with rodent lice in mt karyotype, more than with other sucking lice. The guanaco louse, however, is more closely related phylogenetically to human lice, chimpanzee lice, pig lice, and horse lice than to rodent lice. By parsimony analysis of shared features in mt karyotype, we infer that the most recent common ancestor of sucking lice, which lived ~75 Ma, had 11 minichromosomes; each minichromosome had 1–6 genes and a non-coding region. As sucking lice diverged, split of mt minichromosomes occurred many times in the lineages leading to the lice of humans, chimpanzees, and rodents whereas merger of minichromosomes occurred in the lineage leading to the lice of pigs and horses. Together, splits and mergers of minichromosomes created a very complex and dynamic mt genome organization in the sucking lice.

Key words: mitochondrial genome, genome fragmentation, minichromosome, guanaco, sucking lice.

Introduction

The suborder Anoplura contains 540 species of blood-sucking lice in 50 genera and 15 families; these lice parasitize ~840 species of eutherian mammals from 29 orders (Kim 1988; Durden and Musser 1994a, 1994b). Sucking lice are obligate ectoparasites that cannot survive without their hosts. To adapt to the microenvironment of the host body surface, sucking lice have evolved distinct morphology such as flattened bodies (0.3–8 mm long), elongated mouthparts, and claws for clinging to hairs (Kim and Ludwig 1978). Sucking lice are of medical and veterinary significance as parasites and vectors of disease agents (Nelson et al. 1970; Gibney et al. 1985; Otter et al. 2003; Hornok et al. 2010). Becuase the discovery of an extensively fragmented mitochondrial (mt) genome in the human body louse, *Pediculus humanus* (Shao et al. 2009), 10 other species of blood-sucking lice have been sequenced for mt genomes: 1) the human head louse, *Pediculus capitis*, and the human pubic louse, *Pthirus pubis* (Shao et al. 2012); 2) the chimpanzee louse, *Pediculus schaeffi* (Herd et al. 2015); 3) the domestic pig louse, *Haematopinus suis*, and the wild pig louse, *Haematopinus apri* (Jiang et al. 2013); 4) the horse louse, *Haematopinus asini* (Song et al. 2014); 5) the lice

of the greater bandicoot rat and the Asian house rat, *Polyplax asiatica* and *Polyplax spinulosa* (Dong et al. 2014b); and 6) the lice of the Cheverier's field mouse and the Bower's white-toothed rat, *Hoplopleura akanezumi* and *Hoplopleura kitti* (Dong et al. 2014a). All of these sucking lice have fragmented mt genomes that differ radically from the typical single-chromosome mt genome of animals (Boore 1999; Lavrov 2007). The mt genes of these lice are on 9–20 minichromosomes; each minichromosome is 1.5–4 kb in size and has 1–8 genes. There is substantial variation among these lice in mt karyotype, that is, the distribution of mt genes among minichromosomes. Indeed, even species of the same genus differ from one another in mt karyotype (Herd et al. 2015; Song et al. 2014; Dong et al. 2014a, 2014b).

Recently, the mt genome of the elephant louse, *Haematomyzus elephantis* (suborder Rhynchophthirina) was also sequenced; the 33 mt genes identified were on 10 minichromosomes (Shao et al. 2015). The elephant louse is a chewing louse in the suborder Rhynchophthirina, which is sister to Anoplura. Thus, mt genome fragmentation likely occurred in or before the most recent common ancestor (MRCA) of Anoplura and Rhynchophthirina, which lived ~92 million years ago (Light et al. 2010; Smith et al. 2011; Shao et al. 2015).

To establish the ancestral mt karyotype for the sucking lice and to understand the evolution of fragmented mt genomes in these lice, we sequenced the mt genome of the guanaco louse, *Microthoracius praelongiceps* (family Microthoraciidae). Microthoraciidae is a small family in Anoplura with a single genus and four species that infest camels and lamas (Kim and Ludwig 1978; Kim 1988; Durden and Musser 1994a). In the phylogeny proposed by Kim (1988) based on 39 morphological characters, the 15 families of sucking lice were divided into three clades in two major lineages. Microthoraciidae was grouped with two other families, Ratemiidae (lice of donkeys, horses, and zebras) and Echinophthiriidae (lice of sea lions and seals); together, these three families formed the clade microthoracoid. The other five families of sucking lice that have been studied to date for mt genomes are from the other two clades: Pediculidae (human head body louse, chimp louse), Pthiridae (human pubic louse) and Haematopinidae (pig lice and horse lice) from the pediculoid, whereas Hoplopleuridae (rodent lice) and Polyplacidae (rodent lice) from the polyplacoid (Kim 1988). We inferred the phylogeny of sucking lice with mt genome sequences and established the ancestral mt karyotype of the sucking lice. Our results indicate that both splits and mergers of minichromosomes have occurred in sucking lice in a lineage-specific manner, and are responsible for the evolution of the complex and dynamic mt genome organization in these lice.

Materials and Methods

Louse Collection, DNA Extraction, Mt Genome Amplification, and Sequencing

The guanaco lice, *M. praelongiceps*, were collected in Argentina from guanacos, *Lama guanicoe* (Camelidae) (sample no. B2539). Total DNA were extracted from individual lice with DNeasy Tissue kit (QIAGEN). Fragments of three mt genes, *rrnS* (340 bp), *rrnL* (316 bp) and *cox1* (605 bp), were amplified by polymerase chain reaction (PCR) with primer pairs 12SA–12SB, 16SF–Lx16SR and mtd6–mtd11 (supplementary table S1, Supplementary Material online). These three pairs of primers target the gene regions that are conserved among arthropods. The three gene fragments were sequenced directly using Sanger method at the Australian Genome Sequencing Facility (AGRF). Three pairs of specific primers for *M. praelongiceps*, 12SB2539F–12SB2539R, 16SB2539F–16SB2549R, and cox1B2539F–cox1B2539R, were designed from the sequences of *rrnS*, *rrnL*, and *cox1* fragments (supplementary table S1, Supplementary Material online). The two specific primers in each pair go outwards with 42, 13, and 33 bp in between, respectively. PCRs with these specific primers amplified three near full-length mt minichromosomes of *M. praelongiceps* (supplementary fig. S1, Supplementary Material online); each contains a non-coding region and *rrnS*, *rrnL*, and *cox1*, respectively. These amplicons (2.3, 2.6, and 2.9 kb in size) were sequenced using Sanger method at the AGRF. Another pair of primers specific to *M. praelongiceps*, B2539F–B2539R (supplementary table S1, Supplementary Material online), was designed from conserved sequences of the non-coding regions that flank the coding regions of the three minichromosomes. The PCR with the primer pair B2539F–B2539R produced a mixture of amplicons ranging from 0.5 to 1.8 kb in size, expected from the coding regions of the entire set of mt minichromosomes of *M. praelongiceps* (supplementary fig. S1, Supplementary Material online). These amplicons were sequenced with Illumina Hiseq 2000 platform at the BGI Hong Kong. The PCR strategy used in this study was developed from our observations in previous studies that each mt minichromosome has a distinct coding region but a well-conserved non-coding region (Shao et al. 2009, 2012; Jiang et al. 2013; Dong et al. 2014a, 2014b; Song et al. 2014).

Ex *Taq* (Takara) was used in the initial short PCRs with the following cycling conditions: 94 °C for 1 min; 35 cycles of 98 °C for 10 s, 45 °C for 30 s, 72 °C for 1 min; and a final extension of 72 °C for 2 min. LA *Taq* (Takara) was used in the long PCRs with the cycling conditions: 94 °C for 1 min; 35 cycles of 98 °C for 10 s, 55–65 °C (depending on primers) for 40 s, 68 °C for 4 min; and a final extension of 72 °C for 8 min. Negative controls were run with each PCR experiment. PCR amplicons were checked by agarose gel (1%) electrophoresis; the sizes of amplicons were estimated by comparing with DNA markers. PCR products were purified with Wizard SV Gel/PCR clean-up system (Promega).

Assembly of Illumina Sequence Reads, Gene Identification, and Verification of Individual Mt Minichromosomes

Purified PCR amplicons generated above with the primer pair B2539F–B2539R from the coding regions of the mt minichromosomes of *M. praelongiceps* were sequenced with Illumina Hiseq 2000 platform at the BGI: the insert size was 170 bp; 100 bp from each end of the insert was obtained (i.e., pair-end sequencing). Illumina sequence reads were assembled into contigs with Geneious 7.1.8 (Kearse et al. 2012). The assembly parameters were minimum overlap identity 98% and minimum overlap 50 bp. tRNA genes were identified using tRNAscan-SE (Lowe and Eddy 1997) and ARWEN (Laslett and Canback 2008). Protein-coding genes and rRNA genes were identified with Basic Local Alignment Search Tool (BLAST) searches of GenBank (Gish and States 1993; Altschul et al. 1997). Sequence alignments were with Clustal X (Larkin et al. 2007). The size and circular organization of each mt minichromosome of *M. praelongiceps* identified by sequence-read assembly were verified by PCR (supplementary fig. S1, Supplementary Material online) using outbound specific primers designed from the coding region of each minichromosome (supplementary table S1, Supplementary Material online). The forward primer and reverse primer in each pair were next to each other with a small gap in between. PCRs with these primers amplified each circular minichromosome in full or near full length; these amplicons were also sequenced with Illumina Hiseq 2000 platform as described above to obtain the full-length sequences of the non-coding regions of the minichromosomes. PCR set-up, cycling conditions, agarose gel electrophoresis and size measurement were the same as described above. Negative controls were run for each PCR test. The nucleotide sequence of the mt genome of *M. praelongiceps* was deposited in GenBank under accession numbers KX090378–KX090389.

Phylogenetic Analyses

We inferred the phylogenetic relationships among 22 species of insects using mt genome sequences: 1) the guanaco louse; 2) 11 other species of sucking lice; 3) five species of chewing lice; 4) four species of psocids; and 5) the damsel bug, *Alloeorhynchus bakeri*, which was used as the outgroup (table 1). Sequences of eight mt protein-coding genes (*atp6*, *atp8*, *cox1*, *cox2*, *cox3*, *cob*, *nad4L*, and *nad6*) and two rRNA genes (*rrnS* and *rrnL*) were aligned individually. Sequences of five protein-coding genes (*nad1*, *nad2*, *nad3*, *nad4*, and *nad5*) were not available to all of the 21 species and thus were excluded from our analysis. Protein-coding gene sequences (excluding stop codons) were aligned based on amino acid sequences using the MAFFT algorithm implemented in TranslatorX online platform (Abascal et al. 2010). rRNA genes were aligned using the MAFFT v7.0 online server with G-INS-i strategy (Katoh and Standley 2013). Ambiguously aligned sites were omitted using GBlocks v0.91b (Talavera and Castresana 2007) with default settings. Individual gene alignments were concatenated after removing poorly aligned sites using GBlocks v0.91b. Two concatenated alignments were used in subsequent phylogenetic analyses: 1) PCGRNA matrix, which contains all three codon positions of the eight protein-coding genes and the two rRNA genes (5,881 bp in total); and 2) PCG12RNA matrix, which contains only the first and the second codon positions of the eight protein-coding genes and the two rRNA genes (4,347 bp in total). Both matrices were analyzed using maximum likelihood (ML) and Bayesian methods with RAxML-HPC2 8.1.11 (Stamatakis 2006) and PhyloBayes MPI 1.4f (Lartillot et al. 2013), respectively. For ML analysis, separate partitions were created for each gene in the matrix and bootstrap analyses with 1,000 replicates were performed with the fast ML method implemented in RAxML using the GTRGAMMA model for each partition. For Bayesian analyses, the site-heterogeneous mixture model (CAT + GTR) was used. Two independent chains starting from a random tree were run for 30,000 cycles, with trees being sampled at every cycle. The initial 7,500 trees of each MCMC run were discarded as burn-in. A consensus tree was computed from the remaining 45,000 trees combined from two runs, and the two runs converged at maxdiff <0.3.

Reconstruction of the Ancestral Mt Karyotype of Sucking Lice

Among the software available for ancestral genome reconstruction, ANGES (Jones et al. 2012) is the most promising for reconstructing the ancestral mt karyotype of sucking lice. However, ANGES can only reconstruct ancestral karyotypes for multichromosomal linear genomes and unichromosomal circular genomes, but not for multichromosomal circular genomes such as the mt genomes of the sucking lice. On the other hand, no suitable evolution model is available for fragmented animal mt genomes due to very limited knowledge and experimental data available for these genomes, limiting probability-based approaches for the reconstruction of the ancestral mt karyotype of the sucking lice. We used a manual parsimony method, reported by Gordon et al (2009) (also see Neame 2009), to infer the mt karyotype of the MRCA of the sucking lice. We adopted the same parsimony principles of Gordon et al (2009) by inferring a genome character to be ancestral to all sucking lice: 1) if it is present in at least one of the two major lineages of sucking lice and also in the outgroup, the elephant louse, *H. elephantis* (suborder Rhynchophthirina); or 2) if it is present in both of the major lineages of sucking lice. Characters that are present only in one of the major lineages of sucking lice or only in the outgroup cannot be inferred to be ancestral to the sucking lice. We mapped the mt genome characters of the sucking lice on the phylogenetic tree inferred from the mt genome sequences to identify the shared characters. We then counted the

Table 1
Species of Insects Included in the Phylogenetic Analyses in This Study

Species	Common Name	GenBank Accession Number	Reference
Microthoradus praelongiceps	Guanaco louse	KX090378–89	This manuscript
Pediculus humanus	Human body louse	FJ499473–90	Shao et al. (2009)
Pediculus capitis	Human head louse	JX080388–407	Shao et al. (2012)
Pediculus schaeffi	Chimpanzee louse	KC241882–97, KR706168–69	Herd et al. (2015)
Pthirus pubis	Human pubic louse	EU219988–95, HM241895–8	Shao et al. (2012)
Haematopinus suis	Domestic pig louse	KC814602–10	Jiang et al. (2013)
Haematopinus apri	Wild pig louse	KC814611–19	Jiang et al. (2013)
Haematopinus asini	Horse louse	KF939318, KF939322, KF939324, KF939326, KJ434034–38	Song et al. (2014)
Hoplopleura kitti	Rodent louse	KJ648933–43	Dong et al. (2014a)
Hoplopleura akanezumi	Rodent louse	KJ648922–32	Dong et al. (2014a)
Polyplax asiatica	Rat louse	KF647751–61	Dong et al. (2014b)
Polyplax spinulosa	Rat louse	KF647762–72	Dong et al. (2014b)
Haematomyzus elephantis	Elephant louse	KF933032–41	Shao et al. (2015)
Bothriometopus macrocnemis	Screamer louse	NC_009983	Cameron et al. (2007)
Campanulotes bidentatus compar	Small pigeon louse	NC_007884	Covacin et al. (2006)
Ibidoecus bisignatus	Glossy ibis head louse	NC_015999	Cameron et al. (2011)
Heterodoxus macropus	Wallaby louse	NC_002651	Shao et al. (2001)
Liposcelis bostrychophila	Booklouse	JN645275-6	Wei et al. (2012)
Psococerastis albimaculata	Barklouse	NC_021400	Li et al. (2013)
Longivalvus hyalospilus	Barklouse	JQ910986	Li et al. (2013)
Lepidopsocid sp.	Barklouse	NC_004816	Shao et al. (2003)
Alloeorhynchus bakeri	True bug	NC_016432	Li et al. (2012)

changes required to explain the observed data if a character is inferred to be ancestral. If two or more characters are conflicting with one another, the character with the least changes is inferred to be ancestral. We consider protein-coding and rRNA genes separately from tRNA genes because tRNA genes are much more mobile than protein-coding and rRNA genes in terms of their chromosomal locations.

Results and Discussion

The Mt Genome of the Guanaco Louse, *M. praelongiceps*

We obtained 10,485,938 clean sequence reads with Illumina Hiseq platform from the PCR amplicons of the guanaco louse generated with the 12 pairs of minichromosome-specific primers (supplementary table S1, Supplementary Material online); each read was 100 bp long. We assembled the sequence reads into contigs and identified all of the 37 mt genes that were typical of bilateral animals; these genes were on 12 minichromosomes (fig. 1). The minichromosomes are circular and range from 2,194 to 2,940 bp in size (mean = 2,512; SD = 239); each minichromosome consists of a coding region and a non-coding region (fig. 1; table 2). The coding region of each minichromosome contains 2–5 genes, and varies in size from 465 bp for *trnG-nad3-trnW* minichromosome to 1,745 bp for *trnH-nad5-trnF* minichromosome (Note: minichromosomes are named after their genes hereafter). Nine of the 12 minichromosomes of *M. praelongiceps* have a single protein-coding or rRNA gene each; the other three minichromosomes have two protein-coding genes each. The 22 tRNA genes are on 11 of the 12 minichromosomes; each minichromosome has 1–4 tRNA genes except *atp8-atp6* minichromosome, which has no tRNA gene (fig. 1; supplementary fig. S2, Supplementary Material online). Each of the 37 mt genes in *M. praelongiceps* is present in only one minichromosome; there is no overlap in gene content between minichromosomes. All of the 37 mt genes of *M. praelongiceps* have the same orientation of transcription relative to the non-coding region except for *trnT*, *nad1*, and *trnQ*, which are in a cluster and have an opposite orientation (fig. 1).

We sequenced the full-length non-coding regions of all of the 12 minichromosomes. The non-coding regions range from 1,128 to 1,729 bp (table 2); the size variation is due to the additional sequences in the middle of the non-coding regions of *trnG-nad3-trnW*, *rrnS-trnC* and *atp8-atp6* minichromosomes (supplementary fig. S3, Supplementary Material online). It is noteworthy that these three minichromosomes have the shortest coding regions (465, 765, and 841 bp) but have the longest non-coding regions (1,729, 1,520 and 1,433 bp; table 2). In contrast, *trnH-nad5-trnF* minichromosome has the longest coding region (1,745 bp) but has the shortest non-coding region (1,128 bp). The size contrast between coding and non-coding regions may indicate a selective pressure in the guanaco louse for the overall size of

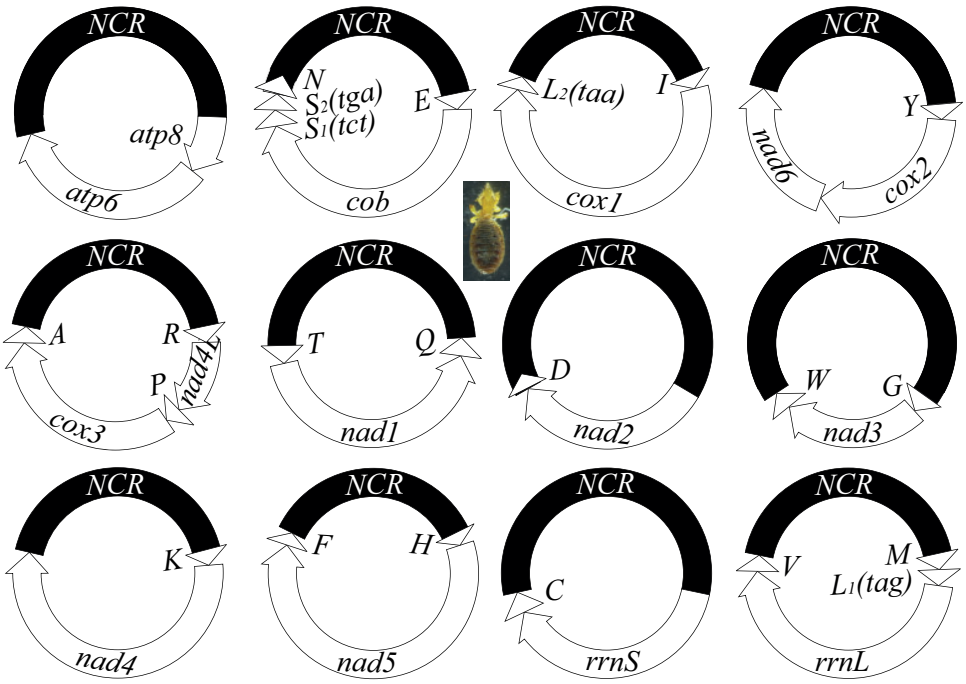


Fig. 1.—The mitochondrial genome of the guanaco louse, *Microthoracius praelongiceps*. Gene name, transcription orientation and length (bp) are indicated in the coding region; non-coding regions are in black. Gene names are: *atp6* and *atp8* (for ATP synthase subunits 6 and 8), *cox1-3* (for cytochrome c oxidase subunits 1–3), *cob* (for cytochrome b), *nad1-6* and *nad4L* (for NADH dehydrogenase subunits 1–6 and 4L), *rrnS* and *rrnL* (for small and large subunits of ribosomal RNA). tRNA genes are indicated with their single-letter abbreviations of the corresponding amino acids.

Table 2

Mitochondrial Minichromosomes of the Guanaco Louse, *Microthoradus praelongiceps*, Identified by Illumina Sequencing

Minichromosome	Minichromosome Size (bp)	Size of Coding Region (bp)	Size of Non-coding Region (bp)	Number of Illumina Sequence Reads	Mean Coverage	GenBank Accession Number
atp8-atp6	2,274	841	1,433	661,887	27,787	KX090378
E-cob-S$_1$-S$_2$-N	2,647	1,356	1,291	510,688	19,156	KX090379
I-cox1-L$_2$	2,940	1,661	1,279	114,200	3,501	KX090380
Y-cox2-nad6	2,496	1,197	1,299	1,234,961	46,868	KX090381
R-nad4L-P-cox3-A	2,529	1,234	1,295	2,660,977	99,812	KX090382
T-nad1-Q	2,310	1,020	1,290	122,203	5,332	KX090383
nad2-D	2,361	1,055	1,306	1,089,226	50,287	KX090384
G-nad3-W	2,194	465	1,729	828,915	37,867	KX090385
K-nad4	2,623	1,328	1,295	2,773,568	103,879	KX090386
H-nad5-F	2,873	1,745	1,128	136,991	4,767	KX090387
M-L$_1$-rrnL-V	2,606	1,305	1,301	258,821	8,632	KX090388
rrnS-C	2,285	765	1,520	93,501	3,406	KX090389
Total	30,138 (mean = 2,512; SD = 239)	13,972 (mean = 1,164; SD = 364)	16,166 (mean = 1,347; SD = 152)	10,485,938 (mean = 873,828; SD = 945,105)	(mean = 34,275; SD = 35808)	

minichromosomes. Excluding the additional sequences in the non-coding regions of *trnG-nad3-trnW*, *rrnS-trnC*, and *atp8-atp6* minichromosomes, the non-coding regions of the 12 minichromosomes have high sequence similarity to one another with conserved motives throughout. Overall, the non-coding regions have high C and G content (>50%) in one end but high A and T content (>70%) in the other end. Indeed, a CG-rich motif (90 bp, 65.6% C and G) is

immediately downstream the 3′-end of the coding regions and an AT-rich motif (62 bp, 98.4% A and T) is upstream the 5′-end of the coding regions, in all of the minichromosomes (supplementary fig. S3, Supplementary Material online). The contrasting nucleotide composition likely indicates different roles of the two ends of non-coding regions in minichromosome replication and gene transcription, for example, the transcription may initiate in the AT-rich end and terminate in the GC-rich end.

Shared Features in Mt Karyotype between the Guanaco Louse and Rodent Lice

Shao et al. (2015) proposed the expression of "mt karyotype" for sucking lice and the elephant louse whose mt genomes comprise multiple minichromosomes. A mt karyotype refers to the: 1) number of minichromosomes, 2) topology of minichromosomes (linear or circular), and 3) gene content and gene arrangement of each minichromosome. All of the 11 species of sucking lice and the elephant louse investigated to date have fragmented mt genomes. Among these lice, the human body louse and human head louse have identical mt karyotypes (Shao et al. 2009, 2012); so do the domestic pig louse and the wild pig louse (Jiang et al. 2013). This is expected as human body louse and human head louse are closely related sister species (or subspecies), so are domestic pig louse and wild pig louse. However, all other species of sucking lice and the elephant louse differ from one another in mt karyotypes, even for species in the same genus, for example, between *Po. asiatica* (louse of the greater bandicoot rat) and *Po. spinulosa* (louse of the Asian house rat) (Dong et al. 2014b), and between *Ho. akanezumi* (louse of the Cheverier's field mouse) and *Ho. kitti* (louse of the Bower's white-toothed rat) (Dong et al. 2014a).

The mt karyotype of the guanaco louse is distinct from those of other 11 species of sucking lice and the elephant louse. However, the guanaco louse shares many features with the rodent lice. *First*, two minichromosomes of the guanaco louse, *trnR-nad4L-trnP-cox3-trnA* and *trnM-trnL$_1$-rrnL-trnV*, are also present, and only present, in the rat louse, *Ho. kitti*; *trnM-trnL$_1$-rrnL-trnV* minichromosome is also present in the rat louse, *Po. spinulosa*, with a minor change (table 3). *Second*, when only protein-coding genes are considered, the *cox2-nad6* minichromosome of the guanaco louse is present, and only present, in *Polyplax* rat lice; the *nad4L-cox3* minichromosome of the guanaco louse is present, and only present, in all of the four rodent lice (fig. 2). *Third*, five mt gene clusters, each with three or more genes, of the guanaco louse are present, and only present, in the rodent lice (table 4).

Phylogeny of Sucking Lice Inferred from Mt Genome Sequences

In order to determine whether the features of mt karyotypes shared between the guanaco louse and the rodent lice are derived for these lice, or whether they represent the ancestral condition of all sucking lice, we inferred the phylogeny of sucking lice using the mt genome sequences available. Our datasets contain the guanaco louse, 11 other species of sucking lice, five species of chewing lice, four species of psocids, and the damsel bug (order Hemiptera) as the outgroup (table 1). We obtained four trees from two concatenated data matrices (PCGRNA: 5,881 bp of protein-coding and rRNA gene sequences; PCG12RNA: 4,347 bp, same as PCGRNA but the 3rd-codon sequences of protein-coding genes excluded) and two tree-construction methods (ML and Bayesian). These four trees have similar topologies to each other and consistently support the monophyly of sucking lice (suborder Anoplura) and the close relationship between Anoplura and Rhynchophthirina to the exclusion of two other suborders of chewing lice, Ischnocera and Amblycera (fig. 3*A*). Within the sucking lice, the four rodent lice of the genera *Hoplopleura* and *Polyplax* are closely related to each other and the guanaco louse is more closely related to human lice, chimpanzee lice, pig lice and horse lice than to the rodent lice with strong support (fig. 3*A*; supplementary figs. S4–S6, Supplementary Material online). The relationships revealed in our analysis are consistent with the family-level phylogeny of the sucking lice proposed by Kim (1988) based on 39 morphological characters, which divided the 15 families of the sucking lice (suborder Anoplura) into three clades in two major lineages, that is, polyplacoid, and microthoracoid + pediculoid.

Light et al. (2010) used the sequences of *18S*, *EF-1α*, and *cox1* genes to infer the phylogeny of 49 species of sucking lice from eight of the 15 families, including 28 species of rodent lice from five genera, three human lice, a chimpanzee louse and a domestic pig louse. Light et al. (2010) also divided the sucking lice into two major lineages and showed that the two most species-rich families of sucking lice, Hoplopleuridae and Polyplacidae, were polyphyletic. The guanaco louse, however, was not sampled in Light et al. (2010), nor other species in the family Microthoraciidae. The relationship among the five genera of sucking lice that were sampled in both the present study and Light et al. (2010) was consistent between the two studies except for the genus *Polyplax*. Light et al. (2010) had one *Polyplax* species, *Po. serrate*, and grouped it with the lice of humans and chimpanzees. We had two *Polyplax* species, *Po. asiatica* and *Po. spinulosa*, and grouped them with the two *Hoplopleura* species with strong support (fig. 3*A*; supplementary figs. S4–S6, Supplementary Material online).

Smith et al. (2011) used the sequences of *18S*, *EF-1α*, and *cox1* to construct the phylogeny for parasitic lice (Phthiraptera) including a subset of 21 species of sucking lice sampled in Light et al. (2010). These two studies are consistent in that they divide the sucking lice into two major lineages but they differ in the placement of species in three genera: *Echinophthirius*, *Neohaematopinus* and *Pterophthirus*. No *Polyplax* species was sampled in Smith et al. (2011). The relationship among the four genera of sucking lice that were

Table 3

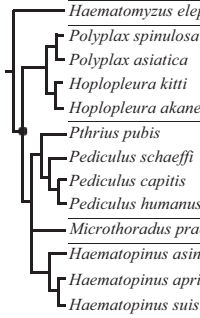
The Mitochondrial Minichromosomes (All Genes Included) Shared between the Guanaco Louse, *Microthoradus praelongiceps*, and Other Sucking Lice and the Elephant Louse

Species	Family	Host	*atp8-atp6*	*E-cob-S_1-S_2-N*	*I-cox1-L_2*	*Y-cox2-nad6*	*R-nad4L-P-cox3-A*	*T-nad1-Q*[a]	*nad2-D*	*G-nad3-W*	*K-nad4*	*H-nad5-F*	*rrnS-C*	*M-L_1-rrnL-V*
Microthoradus praelongiceps	Microthoraciidae	Guanaco	+	+	+	+	⊕	⊗	+	+	+	+	+	⊕
Polyplax asiatica	Polyplacidae	Greater bandicoot rat	+	–	–	–	–	–	–	–	–	–	+	⊕
Polyplax spinulosa	Polyplacidae	Asian house rat	+	–	–	–	–	–	–	–	+	+	–	⊕[b]
Hoplopleura akanezumi	Hoplopleuridae	Chevrier's field mouse	–	–	–	–	–	NA	–	NA	+	NA	+	–
Hoplopleura kitti	Hoplopleuridae	Bower's white-toothed rat	–	–	–	–	⊕	–	–	–	+	NA	–	⊕
Haematopinus suis	Haematopinidae	Domestic pig	–	–	–	-	-	-	-	-	-	+	+	-
Haematopinus apri	Haematopinidae	Wild pig	-	-	-	-	-	-	-	-	-	+	+	-
Haematopinus asini	Haematopinidae	Horse	-	-	-	-	-	-	-	-	-	-	-	-
Pediculus humanus	Pedicinidae	Human (body)	+	-	-	-	-	-	-	-	+	-	-	-
Pediculus capitis	Pediculidae	Human (head)	+	-	-	-	-	-	-	-	+	-	-	-
Pediculus schaeffi	Pedicinidae	Chimp	+	-	-	-	-	-	-	-	+	-	-	-
Pthrius pubis	Pthiridae	Human (pubic)	+	-	-	-	-	-	-	-	NA	-	-	-
Haematomyzus elephantis	Haematomyzidae	Elephant	-	-	-	-	-	⊗	NA	-	-	-	-	-

NOTE.—Plus (+) indicates presence; minus (-) indicates absence; plus in circle (⊕) indicates presence in only guanaco louse and rodent lice; cross in circle (⊗) indicates presence in only guanaco louse and non-rodent lice.

[a]Opposite orientation relative to other genes.

[b]L_2 instead of L_1.



Species	Suborder	Host	*atp8-atp6*	*cob*	*cox1*	*cox2-nad6*	*nad4L-cox3*	*nad1-nad3*	*nad2*	*nad4*	*nad5*	*rrnS*	*rrnL*	*cox2*	*nad6*	*nad1*	*nad3*
Haematomyzus elephantis	Rhynchophthirina	Elephant	-	-	+	-	-	-	N.A.	+	+	+	+	+	-	+	-
Polyplax spinulosa	Anoplura	Asian house rat	+	+	+	⊕	⊕	+	+	+	+	+	+	-	-	-	-
Polyplax asiatica	Anoplura	Greater bandicoot rat	+	+	+	⊕	⊕	+	+	+	+	+	+	-	-	-	-
Hoplopleura kitti	Anoplura	Bower's white-toothed rat	+	+	+	-	⊕	+	+	+	N.A.	+	+	+	+	-	-
Hoplopleura akanezumi	Anoplura	Chevrier's field mouse	+	+	+	-	⊕	N.A.	+	+	N.A.	+	+	+	+	N.A.	N.A.
Pthrius pubis	Anoplura	Human (pubic)	+	+	+	-	-	-	+	N.A.	+	+	+	+	+	+	+
Pediculus schaeffi	Anoplura	Chimp	+	+	+	-	-	-	+	+	+	+	+	+	+	+	+
Pediculus capitis	Anoplura	Human (head)	+	+	+	-	-	-	+	+	+	+	+	+	+	+	+
Pediculus humanus	Anoplura	Human (body)	+	+	+	-	-	-	+	+	+	+	+	+	+	+	+
Microthoradus praelongiceps	Anoplura	Guanaco	+	+	+	⊕	⊕	-	+	+	+	+	+	-	-	+	+
Haematopinus asini	Anoplura	Horse	-	+	-	-	-	+	-	-	-	-	+	-	-	-	-
Haematopinus apri	Anoplura	Wild pig	-	+	-	-	-	+	-	-	+	+	+	-	-	-	-
Haematopinus suis	Anoplura	Domestic pig	-	+	-	-	-	+	-	-	+	+	+	-	-	-	-
If ancestral to sucking lice, changes required to account the data			2	1	1	4	3	3	1	1	1	1	0	3	4	2	3
Ancestral to sucking lice (suborder Anoplura) by parsimony			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No

Fig. 2.—Inference of the ancestral mitochondrial karyotype (tRNA genes excluded) of the sucking lice (suborder Anoplura). Plus (+) indicates presence; minus (-) indicates absence; plus in circle (⊕) indicates presence in only guanaco louse and rodent lice. The phylogenetic tree was inferred from mt genome sequences (see fig. 3 for details); branches with bootstrap support values <50% were collapsed. *cox2-nad6* minichromosome is more parsimonious (four subsequence changes required) than *cox2* and *nad6* are on their own minichromosomes (seven subsequence changes required). *nad1-nad3* minichromosome is more parsimonious (three subsequence changes required) than *nad1* and *nad3* are on their own minichromosomes (five subsequence changes required).

sampled in both the present study and Smith et al. (2011) was consistent between the two studies (fig. 3*A*; supplementary figs. S4–S6, Supplementary Material online).

What Is the Ancestral Mt Karyotype of the Sucking Lice?

Mt karyotypes are available now for 12 species from all of the three clades of the two major lineages of sucking lice (polyplacoid, and microthoracoid + pediculoid). We mapped the mt karyotypes on the phylogeny inferred from the mt genome sequences of these lice, which is consistent with the phylogeny proposed from morphology by Kim (1988), to infer the ancestral mt karyotype of sucking lice. On the basis of parsimony analysis, we infer that: 1) the ancestral mt karyotype of sucking lice comprises 11 types of minichromosomes; and 2) each ancestral minichromosome has a single coding region with 1–6 genes and a single non-coding region (fig. 4). Our inference is on three grounds. *First*, the structure of a single coding region plus a single non-coding region in each minichromosome is seen in all of the 12 species of sucking lice and the elephant louse investigated to date. Slight variation in this structure is found only in two minichromosomes of *Hoplopleura* rodent lice (Dong et al. 2014a) and in one minichromosome of the horse louse (Song et al. 2014), in which an additional, short non-coding region is present. Thus, minichromosomes with one coding region plus one non-coding region can be inferred to be ancestral to sucking lice. *Second*, when only mt protein-coding genes and rRNA genes are considered, all of the 11 types of minichromosomes we inferred to be ancestral to sucking lice are present in both major lineages of the sucking lice; five of them are also present in the outgroup, the elephant louse (fig. 2). Inferring these 11 types of minichromosomes to be ancestral to sucking lice requires the fewest subsequent changes (0–4 for each minichromosome) to account for the variation observed among the 12 species of sucking lice and the outgroup, the elephant louse (fig. 2). In general, mt protein-coding and rRNA genes of animals are much more stable than tRNA genes in terms of their chromosomal locations; this is also the case for sucking lice. We, therefore, consider protein-coding and rRNA genes separately from tRNA genes. *Third*, the location of each of the 22 tRNA genes relative to its upstream or downstream protein-coding or rRNA genes in the inferred ancestral mt karyotype of sucking lice is present in both major lineages of the sucking lice; 10 of them are also present in the outgroup, the elephant louse (fig. 5). Inferring the shared location characters of these tRNA genes to be ancestral to sucking lice requires the fewest subsequent changes (0–5 for each tRNA gene) to account for the variation observed among the 12 species of sucking lice and the outgroup, the elephant louse (fig. 5).

In addition to parsimony-based approaches, probability-based approaches such as ML can also be used to reconstruct ancestral genomes. Probability-based approaches require models for genome evolution established on experimental evidence. Currently, there is no suitable evolution model yet for fragmented animal mt genomes. Although developing such a model is beyond the scope of the present study, we would suggest that four factors should be considered in developing evolution models for fragmented animal mt genomes: 1) splits and mergers of mt minichromosomes apparently occur at different rates; splits appear to occur much more often than mergers in the sucking lice; 2) the rate of minichromosome split can vary among different lineages, so can the rate of minichromosome merger, as observed in the sucking lice; 3) protein-coding and rRNA genes are much less likely to change

Table 4

The Mitochondrial Gene Clusters (Three or More Genes) Shared between the Guanaco Louse, *Microthoradus praelongiceps*, and Other Sucking Lice and the Elephant Louse

Louse Species	Family	Host	*E-cob-S_1-S_2*	*I-cox1-L_2*	*Y-cox2-nad6*	*R-nad4L-P-cox3-A*	*R-nad4L-P-cox3*	*T-nad1-Q*[a]	*G-nad3-W*	*H-nad5-F*	*M-L_1-rrnL-V*	*L_1-rrnL-V*
Microthoradus praelongiceps	Microthoraciidae	Guanaco	⊕	⊗	⊕	⊕	⊕	⊗	+	+	⊕	+
Polyplax asiatica	Polyplacidae	Greater bandicoot rat	-	-	⊕	⊕[b]	⊕[b]	-	+	-	⊕	+
Polyplax spinulosa	Polyplacidae	Asian house rat	-	-	⊕	-	⊕	-	+	+	⊕[c]	+[c]
Hoplopleura akanezumi	Hoplopleuridae	Chevrier's field mouse	⊕	-	-	⊕	⊕	NA	NA	NA	-	-
Hoplopleura kitti	Hoplopleuridae	Bower's white-toothed rat	⊕	-	-	⊕	⊕	-	-	NA	⊕	+
Haematopinus suis	Haematopinidae	Domestic pig	-	⊗	-	-	-	⊗	+	+	-	+[d]
Haematopinus apri	Haematopinidae	Wild pig	-	⊗	-	-	-	⊗	+	+	-	+[d]
Haematopinus asini	Haematopinidae	Horse	-	⊗	-	-	-	⊗	+	+	-	-
Pediculus humanus	Pedicinidae	Human (body)	-	-	-	-	-	-	-	-	-	-
Pediculus capitis	Pediculidae	Human (head)	-	-	-	-	-	-	-	-	-	-
Pediculus schaeffi	Pedicinidae	Chimp	-	-	-	-	-	-	-	-	-	-
Pthrius pubis	Pthiridae	Human (pubic)	-	-	-	-	-	-	-	-	-	-
Haematomyzus elephantis	Haematomyzidae	Elephant	-	-	-	-	-	⊗	-	-	-	+

NOTE.—Plus (+) indicates presence; minus (-) indicates absence; plus in circle (⊕) indicates presence in only guanaco louse and rodent lice; cross in circle (⊗) indicates presence in only guanaco louse and non-rodent lice.

[a]Opposite orientation relative to other genes

[b]Pseudo-*P* instead of *P*.

[c]L_2 instead of L_1.

[d]Pseudo-*V* instead of *V*.

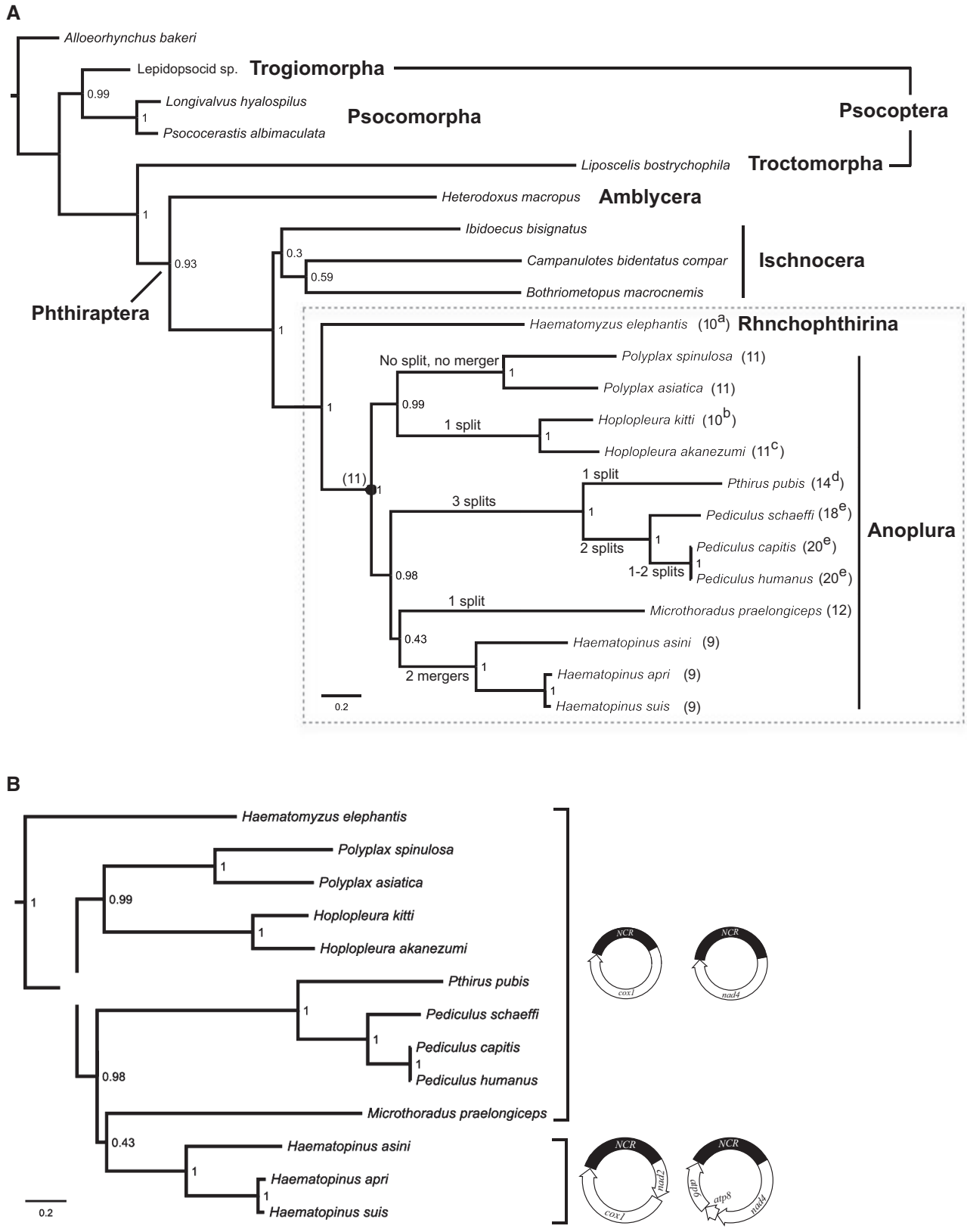


Fig. 3.—(*A*) Phylogenetic relationship of the guanaco louse to other parasitic lice, booklice and barklice. The tree was constructed using Bayesian method with concatenated sequence of eight mitochondrial (mt) protein-coding genes (*atp6*, *atp8*, *cox1*, *cox2*, *cox3*, *cob*, *nad4L*, and *nad6*) and two rRNA genes

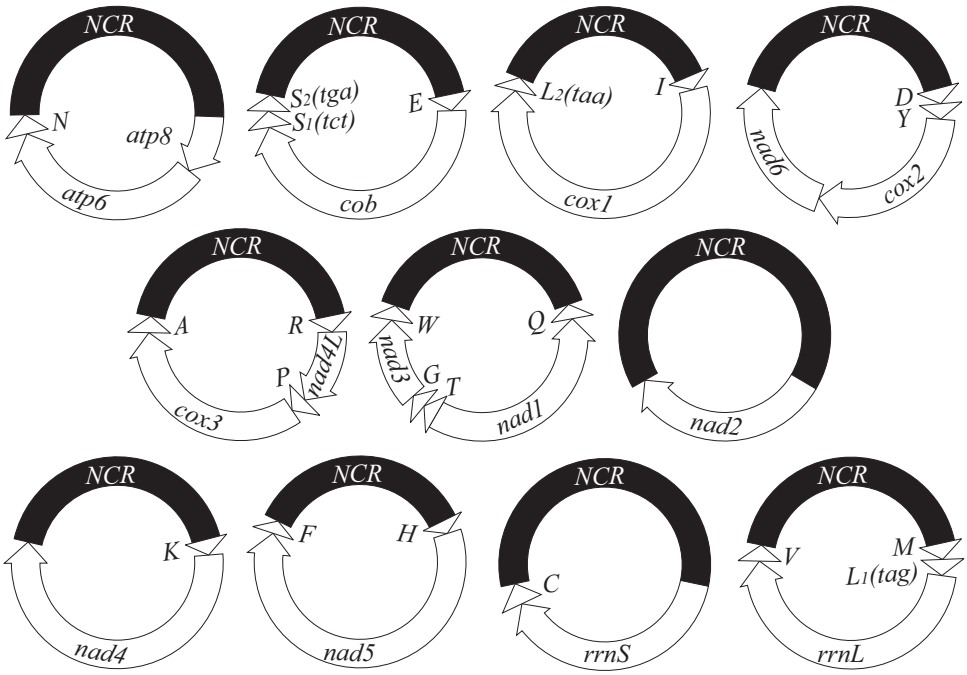


Fig. 4.—Inferred ancestral mitochondrial karyotype of the sucking lice (Anoplura). Gene name and transcription orientation are indicated in the coding region; non-coding regions (NCR) are in black. See fig. 1 legend for gene name abbreviation.

their chromosomal locations than tRNA genes; and 4) a small number of tRNA genes can swap their identities by point mutations at anticodons (Lavrov and Lang 2005; Shao et al. 2009, 2012; Dong et al. 2014b).

Splits and Mergers of Minichromosomes Shaped the Complex and Dynamic Mt Genome Organization of the Sucking Lice

Sucking lice evolved from chewing lice ~92 million years ago (Ma) and diversified into two major clades ~75 Ma (Light et al. 2010; Smith et al. 2011). When comparing the parsimony-based ancestral mt karyotype of sucking lice with those of the 12 species of sucking lice available now, it is clear that splits of minichromosomes have occurred many times since the MRCA of sucking lice; mergers of minichromosomes have also occurred but much less frequently than minichromosome splits (fig. 3*A*). Further, splits and mergers of minichromosome are rather uneven among the lineages leading to the six families of sucking lice investigated to date. In the lineage to the two *Polyplax* species (family Polyplacidae), neither splits nor mergers of minichromosomes occurred; thus, the number of minichromosomes remained unchanged at 11. Changes in mt karyotypes in these two *Polyplax* species were caused entirely by tRNA gene translocations between minichromosomes. Unlike in *Polyplax* species, changes in mt karyotypes in the other 10 species of sucking lice were caused by both tRNA gene translocations and splits or mergers of minichromosomes. In the lineage of the two *Hoplopleura* species (Hoplopleuridae), for which 29 and 34 genes were identified respectively, an ancestral minichromosome, *trnD-trnY-cox2-nad6*, split into two: one has *trnD-trnY-cox2* and the other has *nad6* (fig. 6*A*). In the lineage of the guanaco louse, *M. praelongiceps* (Microthoraciidae), another ancestral

Fig. 3.—Continued

(5,881 bp in total). Posterior probabilities for each grouping are indicated near the branch nodes. The tree was rooted with the true bug, *Alloeorhynchus bakeri* (order Hemiptera). In the broken-line boxed area, the numbers of split and/or merger events of mt minichromosomes in each lineage of sucking lice (suborder Anoplura) are indicated above the branches; the numbers of mt minichromosomes (in brackets) are indicated for the MRCA of sucking lice (the node with a black dot) and for each species (after species names). Note that the full set of 37 mt genes were not found in all species: [a]34 mt genes were found in *Haematomyzus elephantis*; [b]29 mt genes were found in *Hoplopleura kitti*; [c]34 mt genes were found in *Hoplopleura akanezumi*; and [d]34 mt genes were found in *Pthirus pubis*. [e]Point mutations at third anti-codon positions of $trnL_1$ and $trnL_2$ contributed to the increase of two minichromosomes in *Pediculus* species (Shao et al. 2012; Herd et al. 2015). (*B*) *cox1* and *nad4* have their mitochondrial minichromosomes in the elephant louse and all of the sucking lice except for the lice of horses and pigs. Note that tRNA genes are excluded in the minichromosomes, and *nad4* is not found in the human pubic louse. See tables 1 and 3 for the common names and the hosts of the lice. The phylogeny was extracted from part A of this figure.

Species	Suborder	Host	*cox3-A*	*rrnS-C*	*D-Y-cox2*	*E-cob*	*nad5-F*	*G-nad3*	*H-nad5*	*I-cox1*	*K-nad4*	*L1-rrnL*	*cox1-L2*	*M-L1-rrnL*	*atp6-N*	*P-cox3*	*nad1-Q*	*R-nad4L*	*cob-S1*	*S1-S2*	*cob-S1-S2*	*T-nad1*	*rrnL-V*	*nad3-W*
Haematomyzus elephantis	Rhynchophthir	Elephant	-	-	-	-	-	+	+	+	+	+	-	-	-	+	+	+	-	-	-	+	+	-
Polyplax spinulosa	Anoplura	Asian house rat	-	+	+	+	+	+	+	-	+	-	-	-	-	+[1]	-	+	-	+	-	+[2]	+	+
Polyplax asiatica	Anoplura	Greater bandicoot rat	+	+	+	+	-	+	+	-	+	+	+	+	-	+	-	+	-	+	-	+	+	+
Hoplopleura kitti	Anoplura	Bower's white-toothed rat	+	-	+	+	NA	+	NA	+	+	+	-	+	+	+	+	+	+	+	+	-	+	-
Hoplopleura akanezumi	Anoplura	Chevrier's field mouse	+	-	+	+	NA	NA	NA	+	+	-	-	-	+	+	NA	+	+	+	+	-	+	-
Pthirus pubis	Anoplura	Human (pubic)	+	-	-	-	-	+	-	-	NA	-	-	-	-	-	+	+	+	-	-	-	-	-
Pediculus schaeffi	Anoplura	Chimp	+	+	-	-	-	-	-	-	+	+	-	-	-	-	+	-	+	-	-	-	-	-
Pediculus capitis	Anoplura	Human (head)	+	+	-	-	-	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-
Pediculus humanus	Anoplura	Human (body)	+	+	-	-	-	-	-	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-
Microthoradus praelongiceps	Anoplura	Guanaco	+	+	-	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+
Haematopinus asini	Anoplura	Horse	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	-	+	-	+
Haematopinus apri	Anoplura	Wild pig	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	-	+	+[3]	+
Haematopinus suis	Anoplura	Domestic pig	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	-	+	+[3]	+
If ancestral to sucking lice, changes required to account the data			2	3	3	2	3	1	1	2	0	3	4	5	4	1	1	1	4	2	4	2	2	3
Ancestral to sucking lice (suborder Anoplura) by parsimony			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Fig. 5.—Inference of the ancestral location of the 22 mitochondrial tRNA genes of the sucking lice relative to their upstream or downstream protein-coding or rRNA genes. Plus (+) indicates presence; minus (-) indicates absence; [1]pseudo-P upstream *cox3*; [2]pseudo-T upstream *nad1*; [3]pseudo-V downstream *rrnL*. The phylogenetic tree was inferred from mt genome sequences (see fig. 3 for details); branches with bootstrap support values <50% were collapsed.

minichromosome, *trnQ-nad1-trnT-trnG-nad3-trnW* (note: genes underlined have an opposite orientation of transcription to those not underlined) split into two: one has *trnQ-nad1-trnT* and the other has *trnG-nad3-trnW* (fig. 6*A*). The split of this minichromosome increased the number of minichromosomes to 12 in the guanaco louse (fig. 1).

Compared with the rodent lice and the guanaco louse, the three human lice and the chimpanzee louse (Pediculidae and Pthiridae) have had many more minichromosome splits in their lineages. *First*, the two ancestral minichromosomes above that split in the *Hoplopleura* species and the guanaco louse also split, independently, in the lice of humans and chimpanzees (fig. 6*A*). *Second*, the ancestral minichromosome, *trnR-nad4L-trnP-cox3-trnA*, split into two in the lice of humans and chimpanzees: one has *trnR-nad4L-trnP* and the other has *cox3-trnA*. *Third*, in the human pubic louse, *P. pubis* (Pthiridae), another ancestral minichromosome, *rrnS-trnC*, split into two: one has *rrnS* and the other has *trnC*. *Fourth*, in the three *Pediculus* species (Pediculidae), that is, human head louse, human body louse and chimpanzee louse, two minichromosomes with tRNA genes only, *trnT-trnD-trnH* and *trnW-trnS2*, were likely generated via minichromosome split as these five tRNA genes are together with protein-coding genes in two minichromosomes in the human pubic louse (Shao et al. 2012). *Fifth*, in the human head louse and human body louse (Pediculidae), another two minichromosomes with only tRNA genes, *trnS1-trnN-trnE* and *trnM*, were generated by the split of *cob-trnS1-trnN-trnE-trnM* minichromosome via one or two split events, which occurred after the divergence of these two human lice from the chimpanzee louse ~6 Ma (Shao et al. 2009, 2012; Herd et al. 2015). Thus, in total, split of minichromosome have occurred 4 times in the human pubic louse, 5 times in the chimpanzee louse, and 6 or 7 times in the lineage to the human head louse and the human body louse since the MRCA of sucking lice. The mt genomes of the human head louse and the human body louse are the most fragmented among sucking lice to date.

Intriguingly, mt minichromosomes appear to have evolved in the opposite direction in the lice of pigs and horses (Haematopinidae). The mt genomes of the two pig lice and the horse louse comprise nine minichromosomes each and are the least fragmented among the sucking lice (Jiang et al. 2013; Song et al. 2014). Initially, it was thought that these least fragmented mt genomes might be ancestral to sucking lice or close to the ancestral condition (Jiang et al. 2013; Dong et al. 2014b). However, as more data become available, there is mixed evidence. On one hand, the horse and pig lice have the most slowly-evolving mt genomes among the sucking lice, judged by the branch length on the phylogenetic trees inferred from mt genome sequences (fig. 3; supplementary figs. S4–S6, Supplementary Material online); a slower rate of genome sequence evolution might favor a slower rate of genome structure evolution (Shao et al. 2003; Xu et al. 2006). In addition, *trnI* lies in between *nad2* and *cox1* in *nad2-trnI-cox1-trnL2* minichromosome in the lice of horses and pigs. In the great ape lice, however, *trnI* is downstream *nad2* in *nad2* minichromosome and is upstream *cox1* in *cox1* minichromosome in the *Hoplopleura* rodent lice. The different locations of *trnI* between the great ape lice and the *Hoplopleura* rodent lice can be explained by two independent split events of *nad2-trnI-cox1-trnL2* minichromosome if it is ancestral to sucking lice (Jiang et al. 2013; Dong et al 2014a; Song et al. 2014). On the other hand, there is more evidence against the least fragmented mt genomes of the horse and pig lice to be the ancestral condition of the sucking lice. *First*, the least fragmented mt genomes are observed only in three species of one genus, *Haematopinus* (fig. 3*B*); other sucking lice have 11–20 types of minichromosomes. *Second*, the outgroup, the elephant louse, *H. elephantis* (suborder Rhynchophthirina), also has more minichromosomes than the lice of horses and pigs. The 33 mt genes found in the elephant louse are on 10 minichromosomes (Shao et al. 2015), in comparison to the full set of 37 mt genes on nine minichromosomes in the lice of horses and pigs (Jiang et al. 2013; Song et al. 2014). *Third*, *cox1* is on a minichromosome

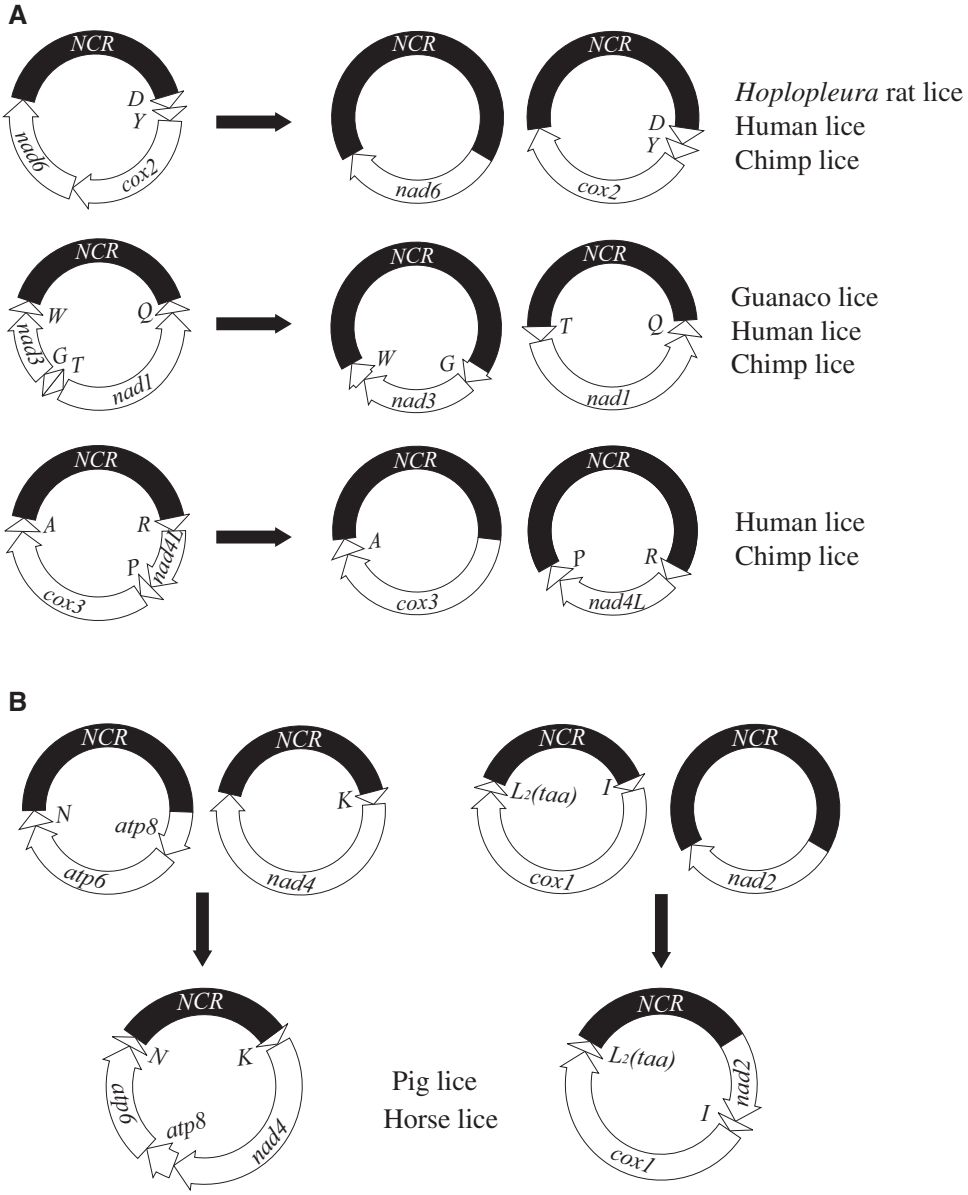


Fig. 6.—The ancestral mitochondrial minichromosomes of sucking lice that split in the lice of human, chimpanzee, rat and guanaco (*A*) and that merged in the lice of pig and horse (*B*). Gene name and transcription orientation are indicated in the coding region; non-coding regions (NCR) are in black. See fig. 1 legend for gene name abbreviation.

with *nad2*, and *nad4* is on another minichromosome with *atp8-atp6* in the lice of horses and pigs. However, in all other sucking lice and the outgroup, the elephant louse, *cox1* is on its own minichromosome without any other protein-coding or rRNA genes, so is *nad4* (fig. 3*B*).

The ancestral mt karyotype we established by parsimony principles for sucking lice is more fragmented than those of the lice of horses and pigs (Jiang et al. 2013; Song et al. 2014). The reduction of minichromosomes in the lice of horses and pigs can be accounted for by mergers of minichromosomes: two ancestral minichromosomes, *atp8-atp6-trnN* and *trnK-nad4*, merged into one minichromosome; so did another two ancestral minichromosomes, *trnI-cox1-trnL*$_2$ and *nad2* (fig. 6*B*). Probability-based methods, which weigh split and merger events differently, may or may not provide the same support for the merger hypothesis as the parsimony method. Shao and Barker (2011) reported evidence and a model for homologous recombination between minichromosomes; this

model can account for the minichromosome mergers we proposed for the lice of horses and pigs; the only extra requirement is the deletion of one of the two non-coding regions after mergers.

In conclusion, we sequenced the mt genome of the guanaco louse, *M. praelongiceps*, making mt genome data available for all of three clades of the sucking lice, that is, polyplacoid, microthoracoid and pediculoid. The mt genome of the guanaco louse consists of 11 minichromosomes; each minichromosome has 2–5 genes. The guanaco louse shares many features in mt karyotype with the rodent lice, more than with other sucking lice. Phylogenetic analyses of mt genome sequences, however, showed that the guanaco louse is more closely related to the lice of humans, chimpanzees, pigs and horses than to the rodent lice. The common features shared between the two major lineages of sucking lice (polyplacoid, and microthoracoid + pediculoid) in mt karyotype allowed us to establish a parsimony-based ancestral mt karyotype for the sucking lice. From this inferred ancestral mt karyotype, it is evident that numerous events of minichromosome split and merger have occurred in sucking lice since their MRCA, resulting in a complex and dynamic mt genome organization not yet seen in any other animal lineage.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We acknowledge funding support from the National Natural Science Foundation of China (No. 31420103902 to R.S. and H.L., No. 31401991 to H.L.), Australian Research Council (DP120100240 to R.S. and S.C.B.), Australia-China Science & Research Fund (ACSRF00980 to R.S.), and Beijing Natural Science Foundation (No. 6144027 to H.L.).

Literature Cited

Abascal F, Zardoya R, Telford MJ. 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 38:W7–W13.

Altschul SF, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.

Boore JL. 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27:1767–1780.

Cameron S, Johnson K, Whiting M. 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.

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.

Dong WG, et al. 2014a. Fragmented mitochondrial genomes are present in both major clades of the blood-sucking lice (suborder Anoplura): evidence from two *Hoplopleura* rodent lice (family Hoplopleuridae). BMC Genomics 15:751.

Dong WG, Song S, Jin DC, Guo XG, Shao R. 2014b. Fragmented mitochondrial genomes of the rat lice, *Polyplax asiatica* and *Polyplax spinulosa*: intra-genus variation in fragmentation pattern and a possible link between the extent of fragmentation and the length of life cycle. BMC Genomics 15:44.

Durden LA, Musser GG. 1994a. The mammalian hosts of the sucking lice (Anoplura) of the world: a host-parasite list. Bull Soc Vector Ecol 19:130–168.

Durden LA, Musser GG. 1994b. The sucking lice (Insecta, Anoplura) of the world: a taxonomic checklist with records of mammalian hosts and geographical distributions. Bull Am Mus Nat Hist 218:1–90.

Gibney VJ, Campbell JB, Boxler DJ, Clanton DC, Deutscher GH. 1985. Effects of various infestation levels of cattle lice (Mallophaga: tri-chodectidae and Anoplura: Haematopinidae) on feed efficiency and weight gains of beef heifers. J Econ Entomol 78:1304–1307.

Gish W, States DJ. 1993. Identification of protein coding regions by database similarity search. Nat Genet. 3:266–272.

Gordon JL, Byrne KP, Wolfe KH. 2009. Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern *Saccharomyces cerevisiae* genome. PLoS Genet. 5:e1000485.

Herd K, Barker SC, Shao R. 2015. The mitochondrial genome of the chimpanzee louse, *Pediculus schaeffi*: insights into the process of mitochondrial genome fragmentation in the blood-sucking lice of great apes. BMC Genomics 16:661.

Hornok S, et al. 2010. Survey on blood-sucking lice (Phthiraptera: Anoplura) of ruminants and pigs with molecular detection of *Anaplasma* and *Rickettsia* spp. Vet Parasitol. 174:355–358.

Jiang HW, Barker SC, Shao R. 2013. Substantial variation in the extent of mitochondrial genome fragmentation among blood-sucking lice of mammals. Genome Biol Evol. 5:1298–1308.

Jones BR, Rajaraman A, Tannier E, Chauve C. 2012. ANGES: reconstructing ANcestral GEnomeS maps. Bioinformatics 28:2388–2390.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30:772–780.

Kearse M, et al. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

Kim KC, Ludwig HW. 1978. The family classification of the Anoplura. Syst Entomol 3:249–284.

Kim KC. 1988. Evolutionary parallelism in Anoplura and eutherian mammals. In: Service MW, editor. Biosystematics of haematophagous insects. Oxford: Claredon Press. Vol. 37. p. 91–114.

Larkin MA, et al. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948.

Lartillot N, Rodrigue N, Stubbs D, Richer J. 2013. PhyloBayes-MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Syst Biol. 62:611–615.

Laslett D, Canback B. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24:172–175.

Lavrov DV. 2007. Key transitions in animal evolution: a mitochondrial DNA perspective. Integr Comp Biol. 47:734–743.

Lavrov D, Lang B. 2005. Transfer RNA gene recruitment in mitochondrial DNA. Trends Genet. 21:129–133.

Li H, et al. 2012. Comparative mitogenomic analysis of damsel bugs representing three tribes in the family Nabidae (Insecta: Hemiptera). PLoS ONE 7:e45925.

Li H, 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:e61685.
Light JE, Smith VS, Allen JM, Durden LA, Reed RL. 2010. Evolutionary history of mammalian sucking lice (Phthiraptera: Anoplura). BMC Evol Biol. 10:292.
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.
Neame E. 2009. Genome evolution: a manual way to the ancestral genome. Nat Rev Genet. 10:426.
Nelson WA, Shemanchuk JA, Haufe WO. 1970. *Haematopinus eurysternus*: blood of cattle infested with the short-nosed cattle louse. Exp Parasitol. 28:263–271.
Otter A, Twomey DF, Crawshaw TR, Bates P. 2003. Anaemia and mortality in calves infested with the long-nosed sucking louse (*Linognathus vituli*). Vet Rec 153:176–179.
Shao R, Barker SC. 2011. Chimeric mitochondrial minichromosomes of the human body louse, *Pediculus humanus*: Evidence for homologous and non-homologous recombination. Gene 473:36–43.
Shao R, Barker SC, Li H, Su Y. 2015. Fragmented mitochondrial genomes in two suborders of parasitic lice of eutherian mammals (Anoplura and Rhynchophthirina, Insecta). Sci Rep 5:17389.
Shao R, Campbell NJH, Barker SC. 2001. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). Mol Biol Evol. 18:858–865.
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, Zhu XQ, Barker SC, Herd K. 2012. Evolution of extensively fragmented mitochondrial genomes in the lice of humans. Genome Biol Evol. 4:1088–1101.
Smith VS, et al. 2011. Multiple lineages of lice pass through the K-Pg boundary. Biol Lett. 5:782–785.
Song S, Barker SC, Shao R. 2014. Variation in mitochondrial minichromosome composition between blood-sucking lice of the genus *Haematopinus* that infest horses and pigs. Parasit Vectors 7:144.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa andmixed models. Bioinformatics 22:2688–2690.
Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 56:564–577.
Wei DD, et al. 2012. The multipartite mitochondrial genome of *Liposcelis bostrychophila*: insights into the evolution of mitochondrial genomes in bilateral animals. PLoS ONE 7:e33973.
Xu W, Jameson D, Tang B, Higgs PG. 2006. The relationship between the rate of molecular evolution and the rate of genome rearrangement in animal mitochondrial genomes. J Mol Evol. 63:375–392.

Associate editor: Dennis Lavrov