Mitochondrial genomes of *Columbicola* feather lice are highly fragmented, indicating repeated evolution of minicircle-type genomes in parasitic lice

Andrew D. Sweet[1], Kevin P. Johnson[2] and Stephen L. Cameron[1]

[1] Department of Entomology, Purdue University, West Lafayette, IN, United States of America
[2] Illinois Natural History Survey, Prairie Research Institute, University of Illinois, Champaign, IL, United States of America

ABSTRACT

Most animals have a conserved mitochondrial genome structure composed of a single chromosome. However, some organisms have their mitochondrial genes separated on several smaller circular or linear chromosomes. Highly fragmented circular chromosomes ("minicircles") are especially prevalent in parasitic lice (Insecta: Phthiraptera), with 16 species known to have between nine and 20 mitochondrial minicircles per genome. All of these species belong to the same clade (mammalian lice), suggesting a single origin of drastic fragmentation. Nevertheless, other work indicates a lesser degree of fragmentation (2–3 chromosomes/genome) is present in some avian feather lice (Ischnocera: Philopteridae). In this study, we tested for minicircles in four species of the feather louse genus *Columbicola* (Philopteridae). Using whole genome shotgun sequence data, we applied three different bioinformatic approaches for assembling the *Columbicola* mitochondrial genome. We further confirmed these approaches by assembling the mitochondrial genome of *Pediculus humanus* from shotgun sequencing reads, a species known to have minicircles. *Columbicola* spp. genomes are highly fragmented into 15–17 minicircles between ∼1,100 and ∼3,100 bp in length, with 1–4 genes per minicircle. Subsequent annotation of the minicircles indicated that tRNA arrangements of minicircles varied substantially between species. These mitochondrial minicircles for species of *Columbicola* represent the first feather lice (Philopteridae) for which minicircles have been found in a full mitochondrial genome assembly. Combined with recent phylogenetic studies of parasitic lice, our results provide strong evidence that highly fragmented mitochondrial genomes, which are otherwise rare across the Tree of Life, evolved multiple times within parasitic lice.

INTRODUCTION

Mitochondria are crucial for the survival of most eukaryotic cells and organisms (although see *Karnkowska et al., 2016*). Because of this important function, the mitochondrial genome has a very stable structure across deep evolutionary scales. The overwhelming majority of metazoans have 37 mitochondrial genes on a single, ∼16,000 bp circular

Corresponding author
Andrew D. Sweet,
sweet19@purdue.edu

chromosome: 13 protein-coding genes (PCGs), two ribosomal RNAs (rRNAs), and 22 transfer RNAs (tRNAs) (*Wallace, 1982*; Wolstenhome, 1992; *Boore, 1992*; *Chen & Butow, 2005*). Significant deviations from this established genomic architecture can have dire consequences for the affected organism. For example, deletion within mitochondrial DNA in humans has been associated with several diseases, including Parkinson's and Kearns-Sayre syndrome (*Zeviani et al., 1988*; *Bender et al., 2006*). Likewise, mitochondrial genome rearrangement, mutation, and fragmentation from deletions has been linked to cell death, disease (e.g., cardiomyopathy), and aging in humans and *Drosophila* models (*Yoneda et al., 1995*; *Hayakawa et al., 1996*; *Kajander et al., 2000*; *Bender et al., 2006*; *Sánchez-Martínez et al., 2006*).

Although abnormal mitochondrial genome architecture in humans and model systems is often a pathological state, several other animals are known to possess uncharacteristic architecture under normal conditions. For example, fragmented mitochondrial genomes have seemingly arisen multiple times across the Tree of Life (*Smith & Keeling, 2015*). Parasitic nematodes in the genus *Globodera* have at least 6 circular chromosomes, each of which differs in gene content and order (*Gibson, Blok & Dowton, 2007*; *Gibson et al., 2007*). There is also evidence of recombination between the different chromosomes in *Globodera*. Fragmented mitochondrial genomes have also been identified in several insect genera including the booklouse *Liposcelis* with 2-3 chromosomes (*Wei et al., 2012*; *Shi et al., 2016*) and the thrips *Scirtothrips* with 2 chromosomes (*Dickey et al., 2015*). Mitochondrial genome fragments are not always circular, although cases of linear mitochondrial chromosomes are only known from outside of Bilateria (*Bendich, 1993*). For example, the unicellular ichyosporean *Amoebidium parasiticum* has a large (>200 kbp) mitochondrial genome consisting of several hundred linear chromosomes (Burger et al.,2003), the box jellyfish *Alatina moseri* has 8 linear chromosomes (*Smith et al., 2012*), and the Hydrozoa *Hydra magnipapillata* has two linear chromosomes (*Voigt, Erpenbeck & Worheide, 2008*).

The greatest known abundance of fragmented mitochondrial genomes within a single group is in the parasitic lice (Insecta: Phthiraptera), which include around 5,000 species. Many species of lice have a single mitochondrial chromosome, although in all species studied to date, the gene order is rearranged relative to the inferred ancestral insect mitochondrial genome (*Shao, Campbell & Barker, 2001*; *Price et al., 2003*; *Covacin et al., 2006*; *Cameron, Johnson & Whiting, 2007*; *Cameron et al., 2011*). However, 16 species of lice have, to date, been shown to have highly fragmented mitochondrial genomes with their genes divided across 9 to 20 "minicircles" depending on the species (*Jiang, Barker & Shao, 2013*; *Dong et al., 2014a*; *Dong et al., 2014b*; *Herd, Barker & Shao, 2015*; *Shao et al., 2015*; *Shao et al., 2017*). The first louse species shown to have minicircles was the human louse, *Pediculus humanus* (*Shao, Kirkness & Barker, 2009*; *Shao et al., 2012*). There are at least 20 mitochondrial chromosomes each between 3-4 kbp long in *P. humanus*. Each circle has a distinct coding region containing an average of two genes, but there is also a non-coding control region (CR) present in every circle, composed largely of motifs conserved between different minicircles. This region includes the D-loop which is involved in DNA replication and the initiation of transcription (*Clayton, 1982*; *Aquadro & Greenberg, 1983*). Other louse taxa with minicircles vary in their number of chromosomes (between nine and 20),

but the circles have a similar structure to those found in *P. humanus*: a few thousand base pairs in length, with variable coding regions and conserved CRs. The mechanism of mitochondrial genome fragmentation in lice is unknown, although *Cameron et al. (2011)* reported the gene for the nuclear-encoded mitochondrial single-stranded DNA-binding protein (mtSSB) is either missing or non-functional in *Pediculus*. Because mtSSB is involved in mitochondrial replication, the lack of this protein could inhibit replication of a normal, single 16kbp+ long chromosome, whereas replication of multiple, smaller chromosomes could proceed without mtSSB (*Pestryakov & Lavrik, 2008*; *Ruhanen et al. , 2010*).

Complete sets of mitochondrial minicircles in lice have only been found in mammalian sucking lice (Anoplura), elephant and warthog lice (Rhyncophthirina), and mammalian chewing lice of the family Trichodectidae. Together these lice form a highly-supported monophyletic group (*Johnson et al., 2018*), therefore implying mitochondrial genome fragmentation evolved once in parasitic lice (*Song et al., 2019*). *Cameron et al. (2011)* found evidence of mitochondrial genome fragmentation in three species of avian feather lice (Philopteridae), although these were only partial genomes, with only one chromosome sequenced and 6-19 genes recovered on a chromosome per louse species. These partial, fragmented mitochondrial genomes from avian lice differed substantially from those found in mammal lice. In particular, the avian louse genomes are typically larger, contain more protein-coding or ribosomal RNA genes (3–7 PCG/rRNA genes in avian lice, 0–2 in mammal lice), and lack a large (1,000 bp+) CR with conserved structural motifs. The fragmented mitochondrial genomes from avian lice reported by *Cameron et al. (2011)* more closely resembled those found in groups outside of lice (e.g., *Lipsoscelis*, *Scirtothrips*, and *Globodera*) than those found in mammalian lice. For this reason, *Cameron et al. (2011)* proposed definitions for the different mitochondrial genome fragmentation "types" found in avian versus mammalian lice. This distinction was emphasized by *Song et al. (2019)* in their proposed name Mitodivisia, the monophyletic clade composed of the three mammal-associated louse lineages, whose diagnostic feature was the minicircle form of extensive mitochondrial genome fragmentation. The extent of mitochondrial genome fragmentation in avian lice, and how similar their genome structures are to those found in mammalian lice, has yet to be evaluated.

Here, we test for mitochondrial minicircles in avian feather lice by focusing on *Columbicola*, a representative of one of the three main clades within Philopteridae (*Johnson et al., 2018*), and the only clade for which complete mitochondrial genomes have yet to be sequenced. Species of *Columbicola* parasitize pigeons and doves and are specialized to live between feather barbs in wing feathers of their hosts, although they eat downy feathers near the host's body (*Nelson & Murray, 1971*; *Marshall, 1981*; *Clayton et al., 1999*). These lice occur throughout the range of their host group, which includes every continent except Antarctica (*Price et al., 2003*). We have several motives for targeting this genus of feather lice. First, previous attempts to sequence and assemble the mitochondrial genome of *Columbicola* have been unsuccessful using traditional methods such as long PCR, suggesting the architecture of the genome is more complicated than a typical single chromosome (pers. obs.). Second, uncovering the mitochondrial genome structure of *Columbicola* would be highly informative for understanding mitochondrial evolution in parasitic lice.

The phylogenomic analysis of Johnson et al. (2018) recovered *Columbicola* on a relatively long branch representing one of three main clades in Philopteridae. However, *Columbicola* is not the earliest diverging lineage in the family; a clade including *Bothriometopus*, a genus with a full-sized, single-chromosome mitochondrial genome (Cameron, Johnson & Whiting, 2007) is the sister group to the rest of Philopteridae (Johnson et al., 2018). The remaining Philopteridae form a clade sister to that represented by *Columbicola*, and includes multiple genera documented to possess full-sized, single-chromosome mitochondrial genomes (*Campanulotes*: Covacin et al., 2006); *Ibidoecous* & *Coloceras*: (Cameron et al., 2011; *Falcolipeurus*: Song et al., 2019). Based on this phylogeny, *Columbicola* is therefore nested within clades of lice which possess single chromosome mitochondrial genomes. If *Columbicola* does not have a single chromosome, this would definitively support an independent origin of fragmentation (c.f. the fragmented genomes described by Cameron et al. (2011) which are partial and thus may not accurately represent genome structure in these species). Third, there is a well-supported and well-sampled phylogeny of the genus *Columbicola* from Boyd et al. (2017), which allows us to choose species that accurately represent the evolutionary diversity of the group. Finally, *Columbicola* is a model system in host-parasite ecology and evolution (Clayton & Johnson, 2003; Johnson et al., 2007; Bush & Malenke, 2008; Bush et al., 2010; Clayton, Bush & Johnson, 2016). Over two decades of experimental work on the genus provides us with the most abundant knowledge-base for any group of feather lice. Previous studies have related mitochondrial genome fragmentation to a variety of life-history traits shared by the louse taxa which possess them (e.g., life-cycle length or use of mammal hosts; Dong et al., 2014a; Dong et al., 2014b; Song et al., 2019). Investigating mitochondrial genome fragmentation in a feather louse genus with well-known ecology allows us to build on and contextualize these previous hypotheses.

In this study, we use existing and novel genome sequence data to assemble the mitochondrial genomes of four *Columbicola* species that represent major clades within the genus. We utilize three different bioinformatic methods for assembling mitochondrial genomes, and then annotate the resulting sequences to test for fragmentation and rearrangements within the genus. Finally, we confirm the reliability of our bioinformatic approach by assembling the fragmented mitochondrial genome of *Pediculus humanus*, for which mitochondrial genome structure has been previously verified with PCR and Southern blot methods.

MATERIALS & METHODS

Whole genome sequence data

We obtained whole genome shotgun Illumina sequencing reads from the NCBI SRA database for *Columbicola macrourae* ex. *Zenaida macroura* (SRR3161953), *C. passerinae* ex. *Columbina picui* (SRR3161931; hereafter "*C. passerinae* 1" following [43]), and *C. passerinae* ex. *Columbina cruziana* (SRR3161930; hereafter "*C. passerinae* 2"). We also generated new sequence data for a specimen of *C. columbae* from the feral rock pigeon (*Columba livia*). These species represent three of the four major clades identified within *Columbicola* by (Boyd et al., 2017) and thus are a good representation of this speciose genus

(90 described species; *Bush, Price & Clayton, 2009*). *Columbicola passerinae* is a known complex of cryptic species parasitizing New World ground-doves (*Johnson et al., 2007*). Although they are currently described as a single species based on morphology, individuals from *C. passerinae* form two well-defined clades structured according to mitochondrial (*cox1*) and genomic-level nuclear sequence data, biogeography, and host association (*Sweet & Johnson, 2016*; *Sweet & Johnson, 2018*). This complex is therefore a useful system for comparing mitochondrial architecture between two very closely related species.

The *Columbicola columbae* specimen was photographed as a voucher (available on the Illinois Data Bank) and total genomic DNA was extracted using a Qiagen QIAamp DNA Micro Kit (Qiagen, Valencia, CA, USA) with a 48-hour incubation period at 55 °C in proteinase K. The resulting DNA was quantified with a Qubit 3.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) using the standard protocol, fragmented with sonication for a mean insert size of 400 bp, and prepared for sequencing with a Hyper Library Preparation Kit (Kapa Biosystems, Wilmington, MA, USA). 150 bp paired-end reads were then generated in a single lane on an Illumina NovaSeq 6000 using a NovaSeq S4 reagent kit.

For both SRA and newly generated sequence data, we trimmed adapters and filtered reads following *Boyd et al. (2017)*. We removed duplicate read pairs with fastqSplitDups.py (https://github.com/McIntyre-Lab/mcscript), removed the first 5 nt from the 5′-end of de-duplicated reads to account for potential low-quality bases using Fastx_clipper v.0.0.14 (FASTX Toolkit, http://hannonlab.cshl.edu/fastx_toolkit/), removed bases with a phred quality score <28 from the 3′-end using a 1 nt window in Fastx_trimmer v.0.014, and after trimming removed reads <75 nt. We confirmed adapter and quality trimming by running the trimmed libraries through FastQC v.0.11.7 (Babraham Bioinformatics, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Mitochondrial genome assembly of *Columbicola*

We used three methods to assemble the mitochondrial genomes of *Columbicola*: two approaches that combine read-mapping and de novo assembly, and a standard de novo approach. The former two methods are also complementary for assembling more challenging sequences (e.g., repetitive non-coding regions). We initially used these approaches to assemble the mitochondrial genome of *C. passerinae* 2 and then searched for similar patterns in the other three *Columbicola* taxa. Our methodology is further outlined in Fig. 1.

First, we used a combination of aTRAM v.2.0 (*Allen et al., 2018*) and MITObim v.1.8 (Hahn, Bachmann, & Chevreux, 2013) to assemble contigs of mitochondrial sequences. aTRAM uses BLAST (*Altschul et al., 1990*) to identify reads from a user-specified target and then uses third-party software to assemble those particular reads de novo. We attempted to assemble all 13 mitochondrial protein coding genes in *C. passerinae* 2 with aTRAM, using the amino acid sequences of each gene from the published *Campanulotes compar* mitochondrial genome as target sequences (AY968672; *Covacin et al., 2006*). For each assembly, we ran aTRAM for three iterations, used ABySS (*Simpson et al., 2009*) for de novo assembly, and used 10% of the read library. Following *Sweet & Johnson (2018)*, we used a fraction of the library because mitochondrial reads are often much more prevalent
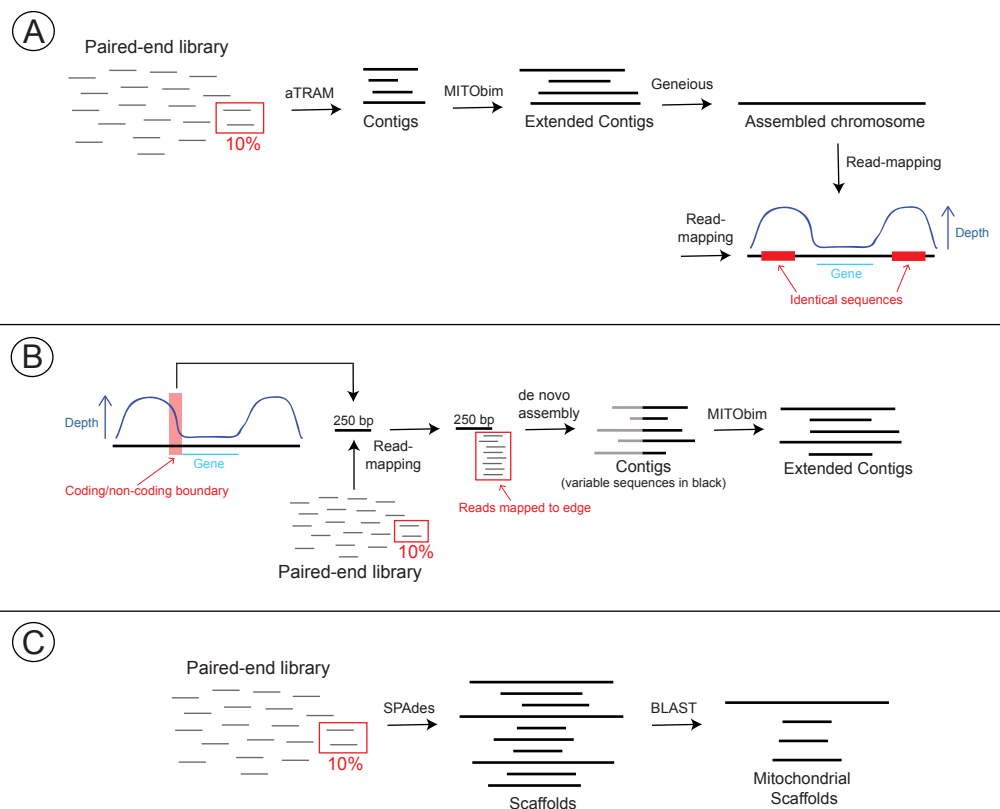
Figure 1 **Workflow detailing three approaches for assembling the mitochondrial genomes of**
***Columbicola* lice.** The first approach (A) uses de novo and read-mapping methods to assemble
mitochondrial contigs and identify peaks in read coverage. The second approach (B) utilizes coverage
information and MITOS annotations to identify the boundaries between coding and non-coding regions.
Reads are then mapped to this boundary region. Reads that map to the 5′ edge are extracted, trimmed
of conserved sequences, and used as starting references in MITObim. The third approach (C) is a more
standard de novo assembly using SPAdes and identifying assembled mitochondrial scaffolds with BLAST.

than nuclear reads in a shotgun sequencing library, and using only a portion of these
reads helps to avoid erroneous assemblies associated with numts, highly-repeated motifs,
or contaminants (including contaminants from index- swapping (*Carlsen et al., 2012*;
Schnell, Bohmann & Gilbert, 2015)). We then used the best- scoring contig from each gene
assembly as a starting reference in MITObim. MITObim is a pipeline that uses MIRA
v.4.0.2 (*Chevreux & Suhai, 1999*) to iteratively map reads against a starting reference
(e.g., a single gene) and then against subsequent contigs to potentially reconstruct whole
mitochondrial genomes. For each starting reference, we ran MITObim for a maximum of
100 iterations with the –quick option and utilizing paired-end information. Our MITObim
runs resulted in a single contig for each starting reference. We then used Geneious v.11.1.5
(Biomatters Ltd., Auckland, NZ) to assemble the MITObim contigs de novo, using
"Medium Sensitivity" and the default parameters.

Our second assembly method, an extension from our first approach, leverages the
conserved control regions (CR) among potential minicircles. Because the CRs have highly

Table 1 Read-mapping statistics for *Columbicola* and *Pediculus humanus* mitochondrial contigs. Contigs were assembled with MITObim using aTRAM-assembled starting references. Values are averaged across all contigs for each species. Statistics for each contig are listed in Table S3.

	Number of reads	Contig length (bp)	High coverage	Low coverage
C. *columbae*	50,501	7,100	3,034	747
C. *macrourae*	17,433	3,177	1,082	133
C. *passerinae* 1	12,155	2,847	821	112
C. *passerinae* 2	61,459	3,219	7,315	665
P. *humanus*	16,251	3,014	425	36

conserved sequence blocks among minicircles in mammal lice, we expected there to be many more reads that map to the non-coding CR in a given contig compared to those which map to genic regions. However, identifying large read-coverage differences across a contig can be used for more than just evidence for presence/absence of minicircles. We can use this coverage information for a single contig to identify the boundary between the non-coding control and genic regions, and then map reads to this "boundary region" to identify reads from multiple minicircles. *Song et al. (2019)* used a similar approach to assemble fragmented mitochondrial chromosomes in mammal lice. To do this, we identified potential CRs by aligning the MITObim contigs produced from the first assembly approach with MAFFT (*Katoh et al., 2002*) in Geneious (default parameters) and identifying conserved regions among the contigs (Fig. S1). We identified coding regions using the beta version of MITOS2 to annotate the MITObim contigs (*Bernt et al., 2013*). Using each annotated contig as a reference, we then mapped the paired-end reads used by MITObim for the assembly of that particular contig. Read-mapping was carried out in Geneious using "Medium-Low Sensitivity" and up to 5 iterations. As predicted, we identified coverage spikes upstream and downstream from the regions with annotated genes (Table 1, Fig. S2). For the *cox1* contig, we extracted a 250 bp region immediately upstream from the coding region, and mapped the entire 10% paired-end read library (i.e., not only reads that mapped to the contig in MITObim) to this region using Geneious (same parameters as above). We then identified reads that mapped to and extended beyond the 3′ end of the reference. Theoretically, these reads include sequences from all minicircles. We assembled the reads de novo in Geneious and identified contigs with mean coverage >20X. We then aligned these contigs with MAFFT, trimmed the conserved 5′ ends of each contig, and used the trimmed contigs (which are therefore putative coding portions of the minicircles) as starting references in MITObim (max 100 iterations, 10% read library).

Third, we used a standard de novo approach to assemble the *C. passerinae* 2 mitochondrial genome. We used SPAdes v.3.11.1 (*Bankevich et al., 2012*) to assemble the 10% read library using the –plasmid option to invoke plasmidSPAdes (*Antipov et al., 2016*). We assessed the assembly quality with the web version of QUAST (*Gurevich et al., 2013*). We then converted the resulting scaffolds to a BLAST-formatted database in Geneious, and identified scaffolds containing mitochondrial protein-coding genes using *tblastn* with amino acid sequences from *Campanulotes compar* as queries.

Assemblies with aTRAM, MITObim, and plasmidSPADES were run on a cluster with 24 cores, two 2.60 GHz Sky Lake processors, and 384 GB of memory maintained by Information Technology at Purdue (West Lafayette, IN, USA). De novo assembly, read mapping, and BLAST analyses in Geneious were run locally on a 2.9 GHz Intel i5 processor with 8 GB of memory.

To annotate the *C. passerinae* 2 mitochondrial genome, we initially used MITOS2 to annotate the assembled contigs. We then manually fine-tuned gene annotations (following the criteria in *Cameron, 2014b*) and identified the number of minicircles by focusing on annotated regions with conserved flanking regions. We identified conserved regions by using BLAST to find highly similar sequences (>80% identical, e-value $<1e^{-20}$) and searching for these sequences both within and between different contigs. We tested for evidence of circularization by identifying identical stretches of sequence >50 bp near the 3′ and 5′ ends of our contigs. We also confirmed circularization using AWA (*Machado et al., 2018*) with kmer length=31 and a minimum circular contig length of 1,000 bp (Table S1). After finalizing the annotations, we used Polymerase Chain Reactions (PCR) to verify our assembled minicircles (Article S1, Fig. S3, Table S2).

We also searched for minicircles in three other representatives of the *Columbicola* genus: *C. columbae*, *C. macrourae*, and *C. passerinae* 1. Evidence for minicircles in these taxa would indicate that the architecture is conserved across the genus. We used our first two approaches to assemble the mitochondrial genomes of these taxa: first generating longer contigs with aTRAM and MITObim and then assembling reads associated with CR/coding region boundaries to produce starting references for MITObim. We tested for a consistent architecture within *Columbicola* by using BLAST to compare our annotated *C. passerinae* 2 contigs to contigs from the other three species. We then compared the variation in gene-order and content between species and verified MITOS annotations following the process outlined in *Cameron (2014b)*.

Mitochondrial genome assembly of *Pediculus humanus*

As another confirmation of our computational approaches, we assembled the mitochondrial genome of *Pediculus humanus*, a species with known minicircles verified with both PCR and Southern Blotting (*Shao, Kirkness & Barker, 2009*). We downloaded the *P. humanus* A 2013 paired-end Illumina data from the GenBank SRA database (SRR5088471), cleaned the library, and assembled the minicircles using the same three approaches used for *Columbicola* (MITObim from aTRAM-assembled starting seeds, MITObim from CR/PCG edge starting seeds, and de novo with SPAdes). We used published gene sequences (translated PCGs and rRNA) from *P. humanus* (*Shao, Kirkness & Barker, 2009*) as targets for aTRAM, and mapped reads from the MITObim *cox1* assembly to identify a CR/protein-coding boundary. Because the *Pediculus* library was considerably larger than the *Columbicola* libraries (128,102,260 cleaned reads), we used 1% of the library (compared to 10% of the *Columbicola* libraries) in each assembly approach.

Table 2 Mitochondrial genome assembly information for four *Columbicola* species.

	Paired-end reads[a]	SRA accession	GenBank accessions	Edge contigs	Minicircles	AT%[c]
C. columbae	7,608,708	SRR8177102	MT113910–MT113924	19	15[b]	68.3
C. macrourae	7,617,990	SRR3161953	MT113902–MT113909 MT113925–MT113932	19	16	59.7
C. passerinae 1	3,557,924	SRR3161931	MT094283–MT094299	17	17	65.1
C. passerinae 2	4,553,898	SRR3161930	MT094266–MT094282	19	17	64.3

Notes.
[a]10% of total reads.
[b]Two protein-coding genes not assembled.
[c]Total AT% for both coding and non-coding regions across all minicircles.

RESULTS

Genome data

Paired-end library sizes varied among the four *Columbicola* genomes. There was an average of 7,598,374 reads from 10% of the adapter and quality-trimmed libraries. The number of reads ranged from 3,557,924 in *C. passerinae* 1 to 7,617,990 in *C. macrourae* (Table 2). We report statistics for 10% of the libraries because these were the reads used for assembling mitochondrial genomes.

Mitochondrial genome assembly of *Columbicola passerinae* 2

The software aTRAM produced assembled contigs for seven protein-coding genes in *C. passerinae* 2 using the *Campanulotes compar* amino acid bait sequences: *cox1* (1,530 bp), *cox2* (873 bp), *cox3* (801 bp), *cob* (1,089 bp), *nad1* (783 bp), *nad3* (477 bp), and *nad5* (945 bp). The reported contig lengths are from the best-scoring contigs that we subsequently used as starting references in MITObim. Resulting MITObim contigs had the following lengths: *cox1*: 3,246 bp (5 iterations); *cox2*: 2,665 bp (7 iterations); *cox3*: 3,597 bp (10 iterations); *cob*: 3,673 bp (11 iterations); *nad1*: 2,973 bp (7 iterations); *nad3*: 2,465 bp (7 iterations); *nad5*: 3,913 bp (10 iterations). An average of 61,459 reads mapped to each contig in the final MITObim iteration (Table S3). When we mapped these reads to the MITObim contigs in Geneious, all contigs showed elevated coverage near the 5′ and 3′ ends of the contigs and lower coverage toward the centers of the contigs. However, coverages toward the centers were still high. For example, coverage at the ends (positions 1–805 and 2,378–3,246) of the contig produced with the *cox1* reference was 12 times higher than coverage toward the center (positions 806–2,377) of the contig (Fig. S2, Table S3). The other contigs showed a similar pattern; among all contigs the coverage at the ends of each contig were on average 11 times higher than coverage near the centers of the contigs (Table 1, Table S3). When we annotated the contigs with MITOS, in every instance the lower-coverage regions corresponded to coding genes, whereas the regions with high-coverage peaks did not have any genes annotated with high confidence. We were unable to produce any additional contigs by assembling all the MITObim contigs together in Geneious. However, aligning each contig to itself with BLAST indicated >50 bp regions

of identical sequences in the high-coverage 5′ and 3′ regions, thus suggesting each contig is a complete circle (Fig. S2).

Although our first approach indicated there are several complete minicircles in *Columbicola*, we were not able to assemble every mitochondrial gene with this approach. One goal of our second approach was to assemble a set of contigs that contained all of the expected mitochondrial genes. The alignment of MITObim contigs from our first approach showed similar sequences near the ends of all contigs but highly variable sequences towards the middle of the contigs (Fig. S1). The variable regions also corresponded to annotated genes and with regions of lower coverage. Together, these results suggest the conserved, high-coverage regions are part of the conserved CRs in the minicircles. We identified 6,400 reads from the 10% read pool (6,226,688 reads) that mapped to the boundary of a CR and the 5′ end of *cox1*. De novo assembly of these reads resulted in 19 contigs with ≥20X coverage. Alignments of the 19 contigs showed conserved sequences near the 5′ ends and variable sequences towards the 3′ ends (Fig. S4). After trimming the conserved regions, the 19 contigs had an average length of 118 bp (min 31, max 139). Using the trimmed 19 contigs as starting references, MITObim produced contigs with an average length of 2,983 bp (min 1,814; max 4,213) from an average of 10 iterations per contig. BLAST identified >50 bp regions of identical sequences on the 5′ and 3′ ends of each MITObim contig.

The plasmidSPADES de novo assembly of the 10% read pool produced 2,024 contigs (N50: 12,832) and 2,020 scaffolds. We found BLAST hits for every *Campanulotes compar* protein-coding and ribosomal RNA gene among the scaffolds. However, *atp8*, *nad4l*, and *nad6* BLAST returns had high e-values >0.01 and low bit scores <25, indicating the homology of these sequences should be interpreted with caution. The longest scaffold with a high-scoring BLAST hit was 1,733 bp. These scaffolds were highly similar to corresponding contigs (i.e., containing the same protein-coding gene) assembled with MITObim (from both aTRAM and control edge starting references), with the only differences coming near or outside the gene boundaries identified by MITOS (Fig. S5). However, none of the de novo-assembled scaffolds covered the entire length of a MITObim contig. Relevant data for all three assembly approaches, including starting references and resulting contigs/scaffolds, are available on the Illinois Data Bank.

Annotation of the *Columbicola passerinae* 2 mitochondrial genome

An initial automated annotation of the control edge MITObim contigs with MITOS and subsequent manual annotations identified all 13 protein-coding genes, both rRNA genes, and 19 of the 22 tRNAs. These genes are separated on 17 separate mini-chromosomes. Although we recovered 19 contigs from MITObim, the assembly process can produce multiple contigs which possess identical coding regions, but whose CRs have sequence differences in portions of the CR outside of those conserved across all minicircles of *C. passerinae* 2. These 'duplicate' contigs were treated as assembly artefacts and removed from subsequent analyses. We determined the number of distinct chromosomes using a combination of the results described above, identifying conserved stretches of sequence and annotating the intervening regions (Fig. S6). Each protein-coding and rRNA gene is on its own circular chromosome, with two chromosomes containing only tRNA genes. The
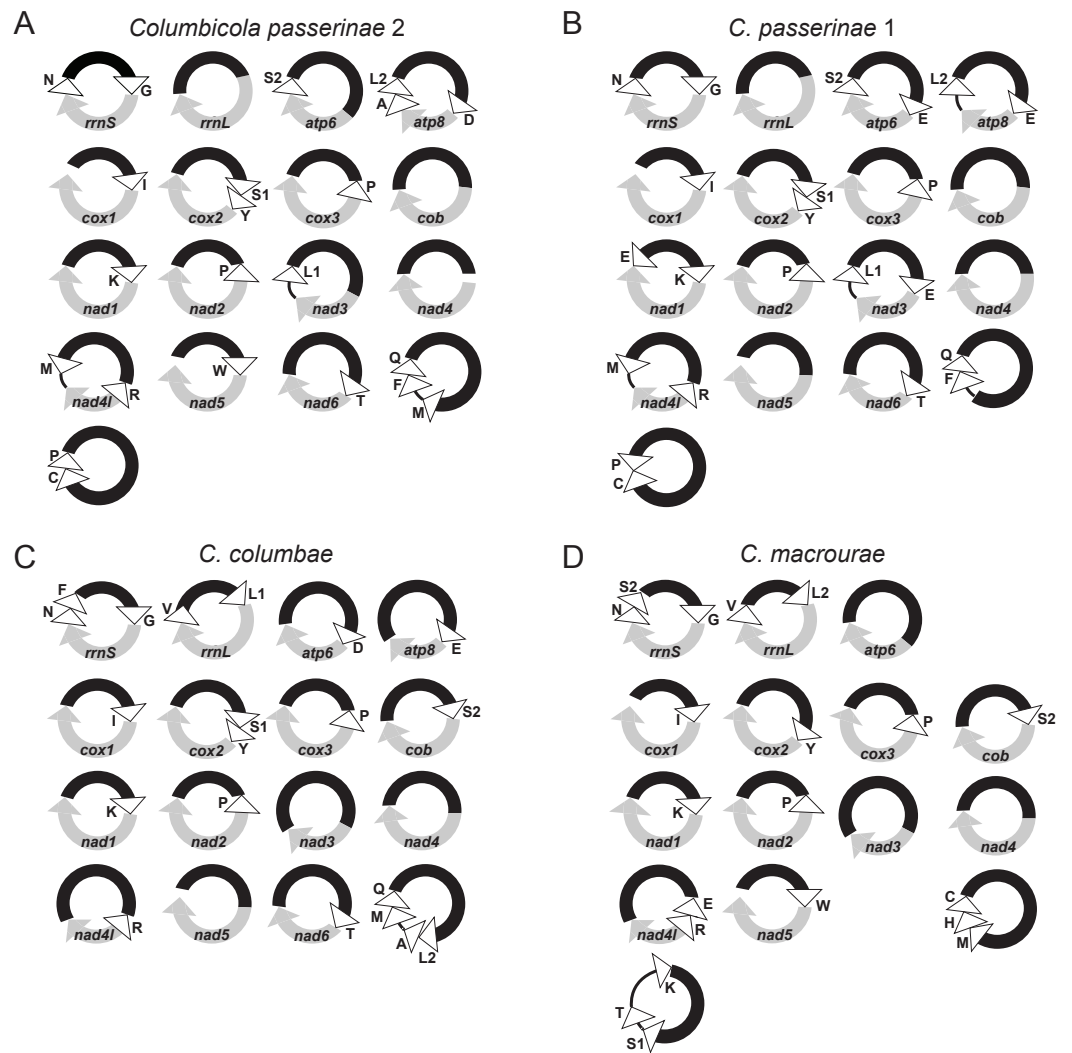
Figure 2 Mitochondrial genomes of four species of *Columbicola* lice: **(A)** *C. passerinae* **2, (B)** *C. passerinae* **1, (C)** *C. columbae,* **and (D)** *C. macourae.* Protein-coding and rRNA genes are indicated in gray, with arrows indicating 5′to 3′directionality in protein-coding genes. Annotated tRNAs are indicated with white triangles on the diagrams and are labeled with IUPAC amino acid letter abbreviations. Control regions are indicated with black segments. Partially assembled chromosomes have white space at the unassembled ends. Black lines indicate intergenic regions.

chromosomes range from 1271 bp to 2862 bp with two contigs (*cox1* and *nad4*) forming incomplete circles, as their assemblies lacked non-coding sequence in the 3′ direction of the coding gene (Fig. 2, Tables S3–S4). AWA and PCR results were consistent with the lengths of our assemblies and further confirmed circularity of most chromosomes (all but the chromosomes containing *cox1* and *nad4*) (Table S1, Fig. S3).

Assembly of other *Columbicola* species

Assemblies of the three other *Columbicola* species showed similar patterns to *C. passerinae* 2. Contigs from MITObim using the aTRAM-assembled references did not assemble together

in Geneious, but when aligned showed conserved regions of sequence with highly variable intervening regions. Although there was variation in the number of mapped reads and coverage values, for each species the conserved non-coding CRs in each contig had higher (4 to 9 times) coverage than the variable coding regions (Table 1). MITOS only identified genes in the variable, lower coverage regions. De novo assembly of reads that mapped to control/coding boundaries resulted in a similar number of contigs for most species (Table 2). MITOS was able to recover most genes from the resulting extended contigs from MITObim. BLAST searches identified similar regions flanking the annotated genes. As with *C. passerinae* 2, we used these flanking regions to identify individual minicircles. We recovered evidence for 15-17 minicircles in each species, with lengths of recovered complete chromosomes ranging from 1207-3118 bp for *C. passerinae* 1, 1262-3173 bp for *C. macourae,* and 1119-2657 bp for *C. columbae* (Fig. 2, Table S4).

Assembly of *Pediculus humanus*

Assemblies of the *P. humanus* mitochondrial chromosomes showed similar patterns to the *Columbicola* assemblies. aTRAM assembled 11/13 PCGs (all excluding *nad3* and *nad4l*). Using these assemblies as starting references in MITObim produced contigs with an average length of 3,014 bp. Read-mapping to these MITObim contigs showed large relative coverage difference between coding and non-coding regions (average of 12 times higher coverage in non-coding regions; Table 1 and Table S3). Using the contigs assembled from the CR/*cox1* edge reads, we recovered 17/18 chromosomes (we were unable to assemble the chromosome that contains *rrnS*) that showed high sequence similarity (>99% in coding regions) to the published *P. humanus* mitochondrial sequences from *Shao, Kirkness & Barker (2009)* (Fig. S7 ; contigs and alignments of published sequences are available on the Illinois Data Bank). There was lower similarity between the non-coding CRs, but most of these contigs still had >90% total similarity to the published sequences. Four contigs had particularly lower similarity (70–80%) with published sequences in the CRs, and five contigs were shorter than their comparable published chromosomes. We were able to assemble and identify 11 PCGs and both rRNA genes from *de novo* assembly. However, as with *Columbicola*, these contigs were shorter than contigs produced from the other two assembly approaches; most assemblies seemed to fail at the CRs (Fig. S8). There was also a chimeric contig containing portions of *nad5* and *rrnL*.

Comparative mitochondrial genome structure among *Columbicola*

Overall mitochondrial genome structure was consistent across *Columbicola* species, with each protein-coding and rRNA gene on a separate chromosome and two minicircles with only tRNAs. All four genomes showed AT compositional bias (average 67.2 AT% across all coding regions of the minicircles, ranging from 63.4–69.2 total AT% among the four species; Table 2, Table S4), which is consistent with some insect mitochondrial genomes and comparable to other louse mitochondrial genomes (*Cameron, Johnson & Whiting, 2007*; *Cameron, 2014a*). Additionally, protein coding genes aligned to the corresponding genes from the *C. passerinae* 2 assembly and had expected uncorrected distances between overlapping regions (13% between *C. passerinae* 1, and 22%–28% between the more
distantly related taxa). None of the completely assembled genes were significantly longer or shorter than homologous copies in other *Columbicola* species or other louse species, which suggests the assemblies were reliable. However, *nad5* was incompletely assembled in *C. passerinae* 1 and *C. columbicola*. Assembled contigs from *C. macrourae* did not include minicircles with either *atp8* or *nad6*; however, this is likely an assembly or annotation failure rather than evolutionary loss of these genes within this species. Both genes are comparatively short and have higher rates of amino-acid substitution than other louse mitochondrial genes creating challenges for BLAST. Additional aTRAM and BLAST searches using the *C. passerinae* 2 amino acid sequences of *atp8* and *nad6* as references likewise failed to identify these two genes in *C. macrourae*, further highlighting the difficulty of assembling or annotating them from short-read genomic data. Similarly, each *Columbicola* species assembly was missing between two (*C. columbae*) and five (*C. macrourae*) tRNA isotypes and had one or more duplicated tRNA isotypes. None of the canonical 22 tRNA isotypes were missing across all four *Columbicola* species, suggesting that missing tRNAs are likely present on other tRNA-only minicircles which failed to assemble or be adequately annotated.

Although the overall genome structure was consistent across the four *Columbicola* mitochondrial genomes, there was substantial variation in gene arrangement between *Columbicola* species. Of the 17 minicircles identified, only one (*trnI*-*cox1*) had a common gene arrangement across all four *Columbicola* species. Each of the other 14 minicircles containing a PCG or rRNA gene differed between at least two species due to the presence, absence, or translocation within the minicircle of accompanying tRNAs. These differences varied among comparisons: *C. passerinae* 2 and *C. passerinae* 1 shared 10 minicircles, *C. passerinae* 2 and *C. columbae* shared 7, *C. passerinae* 2 and *C. macrourae* shared 6, *C. passerinae* 1 and *C. columbae* shared 7, *C. passerinae* 1 and *C. macrourae* shared 4, and *C. macrourae* and *C. columbae* shared 9. The assembled and annotated chromosomes for all four *Columbicola* species are available on GenBank (Table 2). Relevant data generated from the various analyses, including assembled contigs and read-mapping references, are available on the Illinois Data Bank.

DISCUSSION

Fragmented mitochondrial genomes are a rare phenomenon in nature, with relatively few animals known to possess this genomic architecture. Extensive fragmentation of the mitochondrial genome into a large number of minicircles each with a small number of genes per chromosome is even rarer still. Amongst the insects for which complete/near-complete mitochondrial genomes have been sequenced, highly fragmented, minicircle-type genomes have only previously been found in a clade of parasitic mammal lice (Anoplura, Rhynchopthirina, and Trichodectidae), while less fragmented genomes consisting of 2–3 chromosomes have been found in free-living booklice of the genus *Liposcelis* (*Shi et al., 2016*), and a species of thrips (*Dickey et al., 2015*). In insects for which partial mitochondrial genomes have been sequenced, evidence for genome fragmentation has also been recovered for several genera of avian feather lice (*Anaticola, Philopterus,* and *Quadraceps*) (*Cameron et*

al., 2011). These exemplar chromosomes did not include the full suite of genes, although it is very likely the missing genes are present on other chromosomes that were not sequenced in these taxa. In the current study, we found strong evidence for highly-fragmented, minicircle-type genomes in another feather louse (Philopteridae) genus, *Columbicola*. We were able to assemble and annotate contigs which contained all of the protein-coding and rRNA genes in three species and 13 of the 15 genes in a fourth species (*C. macrourae*). We found no evidence for pseudogenes or for different copies of PCGs or rRNA genes on multiple minicircles, as has been found in the fragmented mitochondrial genome of the nematode *Globodera* (*Gibson, Blok & Dowton, 2007*; *Gibson, Blok & Dowton, 2007*). We were also able to recover 17 to 20 of the 22 tRNA genes in each species, and all 22 isotypes in at least one species. Other reported minicircular genomes from sucking lice are also missing between one to two tRNA genes per species, consistent with our findings here (*Song et al., 2019*). It is possible the missing genes are on one or more additional minicircles that we could not detect with our approach. Nevertheless, with at least 17 minicircles and an average of two genes per circle (and as few as one), the *Columbicola* mitochondrial genome appears to be much more fragmented than that of any other feather louse species reported by *Cameron et al. (2011)*, which had between 6 and 22 full genes on the recovered chromosomes. The mitochondrial genomes of *Columbicola* are also more fragmented than the mitochondrial genomes of most mammal associated lice (*Jiang, Barker & Shao, 2013*; *Dong et al., 2014a*; *Dong et al., 2014b*; *Herd, Barker & Shao, 2015*; *Shao et al., 2015*; *Shao et al., 2017*). *Columbicola* has a similar level of fragmentation to the human louse (*Pediculus humanus*), which has at least 20 minicircles and an average of two genes per circle (*Shao, Kirkness & Barker, 2009*; *Shao et al., 2012*). Like *Columbicola*, *Pediculus* also has several circles that only contain tRNA genes. This genome organization has not been found in other louse taxa with minicircles, in which all tRNAs that were detected are co-located with PCGs or rRNAs. However, undetected tRNA-only minicircles are a plausible explanation for putatively missing tRNA isotypes in these species.

Every recovered protein-coding gene in each species of *Columbicola* was on a separate minicircle. This is the first case of such an extreme level of fragmentation among PCGs in a metazoan mitochondrial genome. Even *Pediculus humanus*, which has a comparably fragmented genome in terms of the number of minicircles, has a circle containing both *atp6* and *atp8*. The genes *atp6* and *atp8* are consistently adjacent to one another in other fragmented or single louse mitochondrial genomes and have traditionally been understood as being co-transcribed genes (*Senior, 1988*; *Pelissier et al., 1995*). The two genes are separated in the mitochondrial genome of the ibis louse *Ibidoecus*, but this species' mitochondrial genome is not fragmented and has a canonical single mitochondrial chromosome (*Cameron et al., 2011*). In the case of *Columbicola* species, these genes appear to have separated onto different chromosomes. Once any fragmentation occurs, it is possible that there is no constraint on which parts of the genome become fragmented, because transcription already needs to be coordinated among different chromosomes.

There are considerable differences in tRNA gene arrangements among *Columbicola* species. Only one chromosome (*trnI-cox1*) was consistent across the four *Columbicola* species, and there are varying levels of shared arrangements between pairs of species.

Degrees of similarity in minicircle gene arrangements appears to mirror phylogenetic relatedness: *C. passerinae* 2 and *C. passerinae* 1 share the most minicircles (13), whereas *C. passerinae* 2 shares just 7 with *C. columbae* and 6 with *C. macrourae*. *Columbicola passerinae* 1 and 2 are very closely related, being cryptic species, whereas *C. columbae* and *C. macrouae* are in two different clades (*Boyd et al., 2017*). This consistency among taxa suggests mitochondrial gene arrangements could help clarify phylogenetic relationships within *Columbicola* or aid in species delimitation within species complexes (e.g., *C. passerinae*) (*Cameron et al., 2006*; *Cameron, 2014a*).

Arrangement comparisons are, however, complicated by tRNA duplications which are present in all four *Columbicola* species. For example, *C. passerinae* 1 has four, sequence identical copies of *trnE* which result in three minicircles differing in gene content from *C. passerinae* 2, suggesting recent duplication of this gene into what are otherwise conserved minicircles. Other tRNA duplications appear themselves to be part of conserved minicircle gene arrangements. For example, *trnP* is found in minicircles with *cox3* and *nad2* in all four species, and in a tRNA-only minicircle with *trnC* in both *C. passerinae 2* and *C. passerinae* 1. Within each species, all copies of *trnP* are sequence identical with the exception of *C. columbae* where the *nad2* associated copy has a large insertion between the anticodon and acceptor stems disrupting the variable loop. The presence of sequence-identical, duplicate copies of tRNAs across multiple minicircles, combined with the variation in tRNA arrangement between *Columbicola* species is consistent with recombination (or other gene conversion mechanisms) between minicircles as has been proposed elsewhere (*Dong et al., 2014a*; *Dong et al., 2014b*; *Shao, Kirkness & Barker, 2009*).

Overall, variation in which minicircles are present among different species of *Columbicola* is considerably higher than has been recorded for most other lice with minicircular mitochondrial genomes. Variation within a genus ranges from just 1 of 18 minicircles being different between two species of *Pediculus* (*Herd, Barker & Shao, 2015*) to 9 of 11 minicircles being different between species of *Polyplax* (Dong et al., 2015). Other than *Polyplax*, over two-thirds of minicircles are shared by all members of a genus of sucking louse (17/19 to 4/12) suggesting that minicircle arrangements are stable at the genus-level within mammal-associated lice (*Dong et al., 2014a*). As might be expected given this level of variation, few derived gene-boundaries are shared between *Columbicola* species and other lice, of which only *trnI-cox1* and *trnY-cox2* are likely of phylogenetic significance being shared by the common ancestor of a clade comprising all Ischnocera, Anoplura, and Rhychophthirina (c.f. *Shao et al., 2017*).

Assembling fragmented mitochondrial genomes with a fully bioinformatic approach

Assembling fragmented mitochondrial genomes can be a challenging and labor-intensive process. In this study, we were able to assemble the mitochondrial minicircles of lice using a fully bioinformatic approach from raw genomic Illumina reads. In particular, we used an approach that assembled short, variable contigs from reads that mapped to the edge of the control region (CR) and coding regions. These contigs served as starting references for assembling longer contigs, essentially acting as "bioinformatic PCR primers." By

identifying conserved non-coding CRs that flank adjacent coding regions, we could then detect separate circular chromosomes. Subsequent annotations recovered most of the 37 mitochondrial genes present in metazoans.

We used two other methods for assembling mitochondrial genomes: *de novo* and read-mapping extension. *De novo* assembly is fast (~12 min on five cores for SPAdes) and has the advantage of not requiring a starting reference. However, finding mitochondrial scaffolds is reliant on BLAST searches with a closely related target. In our case, we were not able to reliably identify every PCG or rRNA gene in each *Columbicola* species, presumably because the initial target was too distantly related (*Campanulotes*), the missing genes were particularly divergent (e.g., *nad6*), or the genes simply failed to assemble. It is also challenging to identify tRNA-only minicircles with a *de novo* approach, because tRNA genes are short (<100 bp) and can be highly divergent from even a closely-related target. Our *de novo* assemblies also had trouble assembling through the CR because of the high similarity of this region among minicircles, and as a result the *de novo* scaffolds are considerably shorter than contigs produced from other approaches (Fig. S5). Assemblies with aTRAM were also fast (assembled PCGs in ~3.5 min total) and using the assembled genes as starting seeds with read-mapping in MITObim allowed us to extend the contigs into the CR. However, we were limited by the starting seeds we could assemble with aTRAM, which as with standard *de novo* assembly tended to exclude faster-evolving PCGs and tRNA genes. Our approach based on the coding/non-coding boundary combines the benefits of *de novo* and read-mapping approaches, while also enabling us to assemble all mitochondrial minicircles (perhaps excluding small chromosomes that only contain tRNAs).

Our approach is an important progression that builds on previous work that primarily used wet laboratory methods or a combination of bioinformatics and laboratory methods to identify and confirm mitochondrial minicircles. The first minicircular mitochondrial genome from insects, that of *Pediculus humanus*, was assembled in 18 separate contigs from Sanger shotgun reads, with non-coding regions sequenced from clone libraries of PCR products generated by designing primers within the coding regions of assembled contigs (*Shao, Kirkness & Barker, 2009*). Circularization was also confirmed by restriction digest and Southern hybridization of purified mitochondrial DNA extractions from *Pediculus* (*Shao, Kirkness & Barker, 2009*). We were able to recover these same chromosomes from *Pediculus* using our bioinformatic approaches. Subsequently, a procedure was developed for sequencing minicircles by first sequencing several partial coding gene sequences, designing outward facing primers which would amplify the non-coding portion of several minicircles, then designing primers in conserved regions of the non-coding region to amplify all minicircles present in a given species. This mixed population of amplicons would then be sequenced using NGS methods to yield clean, minicircle contigs (*Shao et al., 2012*). Recently, *Song et al. (2019)* simplified this approach to minimize primer design and PCR steps by using initial partial coding sequences (generated by universal PCR primers) as seeds to directly assemble minicircles from genomic extracts. Similar to our approach, conserved CRs were identified for select mitochondrial genes (initially sequenced with Sanger) from whole genome assemblies and subsequently used to identify all contigs

derived from mitochondrial minicircles. Finally, *Song et al. (2019)* confirmed putative minicircle findings with PCR amplification and gel electrophoresis.

Approaches that rely on laboratory methods to either identify or confirm fragmented mitochondrial genomes are certainly useful and should continue as an option for these types of studies, particularly because there are some limitations to assembling these genomes with Next Generation Sequencing (NGS) data. For example, because it is repeated across multiple minicircles, it is challenging to assemble through long stretches of relatively conserved non-coding regions (e.g., CRs) with short-read data. The entire CR may not assemble for a particular minicircle, which prohibits confirmation of circularization. This was the case for a few of our assembled fragments (incomplete circles in Fig. 2). Due to their similarity, reads from the CRs could also misassemble among different minicircles, so the exact sequences of the CR should be interpreted with caution or reported with ambiguity codes (as we have done in this study). Additionally, a bioinformatic approach may fail to recover all genes present in the organism. This could particularly be an issue for tRNA-only minicircles or highly divergent PCGs such as *atp8* and *nad6*. However, laboratory-based methods can also fail to identify minicircles or genes, especially tRNAs. Additionally, the most rigorous wet-lab methods for confirming the presence of minicircles, such as Southern blotting, require large volumes of starting tissue, thus limiting their utility to species that are either physically large (unlike lice) or are maintained in culture, which is rarely feasible for the vast majority of louse species.

It should also be noted that bioinformatics methods do not appear to erroneously assemble fragmented mitochondrial genomes for taxa where they are not present. Many hundreds of insect mitochondrial genomes have been assembled from NGS datasets (e.g., *Linard et al., 2018*) using bioinformatics assembly procedures such as those presented here (e.g., *de novo* or MITObim baited assembly). Complete circularized mitochondrial genomes are routinely recovered, or when incomplete, assemblies typically only lack portions of the CR (e.g., *Gómez-Rodriguez et al., 2015*). In these assemblies there is no evidence for chimeras between incompletely assembled portions of genomes. Within insects, evidence for mitochondrial genome fragmentation derived from NGS sequencing and bioinformatics analysis is confined to taxa for which prior evidence for fragmentation had been collected (such as lice). Conversely, even within taxa such as lice where aberrant mitochondrial genomes appear to be common, NGS-only approaches have successfully assembled complete, canonical, single-chromosome mitochondrial genomes when such an architecture was present in a given taxon (e.g., for *Falcolipeurus* (*Song et al., 2019*).

Thus despite potential limitations, with an ever-increasing reliance on NGS data it is important to demonstrate that fragmented genomes can be assembled from sequence read libraries. In this regard there is capacity for the study of "aberrant" mitochondrial genome architectures to follow the path of "conventional" mitochondrial genome studies over the past decade where the expansion of NGS techniques has led to a sharply reduced reliance on PCR-centric methods (see review *Cameron (2014a)*). Avoiding PCR is especially relevant for rare specimens or for incorporating publicly available data (NCBI SRA), where it may not be possible to confirm genomic architecture with PCR due to a lack of DNA extracts. We assembled mitochondrial genome fragments from whole genome shotgun sequencing,
but our approach could also be used with reads generated from target capture methods (e.g., ultra-conserved elements (UCEs)). These genome reduction methods target nuclear regions, but mitochondrial reads are also captured at relatively high depth as "by-catch" in the process (*Amaral et al., 2015*; *Ströher et al., 2017*). Utilizing the ever-growing diversity of genomic data, most of which is publicly available through databases such as GenBank, allows for further exploration of mitochondrial genome architecture and arrangements on a broad taxonomic scale. The bioinformatic methods described here accurately reconstruct aberrant mitochondrial genomes whose structures have been verified by wet-lab procedures (i.e., *Pediculus*) and extend our capacity to search for fragmented mitochondrial genomes into taxa which cannot be verified by wet-lab methods (i.e., *Columbicola* and other feather-lice).

Evolutionary implications for highly fragmented mitochondrial genomes in *Columbicola*

Highly fragmented mitochondrial genomes in lice were first identified in human lice (*Pediculus humanus*) (*Shao, Kirkness & Barker, 2009*), and minicircle-type mitochondrial genomes have since been confirmed in several other genera of sucking lice (Anoplura (*Jiang, Barker & Shao, 2013*; *Dong et al., 2014a*; *Dong et al., 2014b*; *Herd, Barker & Shao, 2015*; *Shao et al., 2017*), elephant and warthog lice (Rhyncophthirina *Shao et al., 2015*), and chewing lice in the family Trichodectidae (*Cameron et al., 2011*; *Song et al., 2019*). Phylogenetic analyses have indicated that these three lineages of mammalian lice form a monophyletic group, which suggests highly fragmented mitochondrial genomes evolved once in the common ancestor of this group, as suggested by *Song et al. (2019)*. Given the rarity of this architecture in other groups of organisms and the implications of fragmentation for disease, it is very reasonable to conclude there have been few transitions from full to fragmented mitochondrial genomes. However, our study implies a second such transition to minicircle-type mitochondrial genomes in parasitic lice. The mitochondrial genomes of *Columbicola* species are the first recorded instance of a fragmented mitochondrial genome from within the family Philopteridae for which the complete genome has been sequenced. Other published complete mitochondrial genomes of Philopteridae include species of *Bothriometopus* (*Cameron, Johnson & Whiting, 2007*), *Ibidoecus* (*Cameron et al., 2011*), *Falcolipeurus* (*Song et al., 2019*), *Campanulotes* (*Covacin et al., 2006*), and *Coloceras* (*Cameron et al., 2011*). All five of these species have a single circular mitochondrial chromosome, although *Coloceras* has a full-sized chromosome and at least one additional heteroplasmic fragment missing approximately one third of the full-sized genome. Recent phylogenomic analysis shows that *Columbicola* is nested within these other three taxa (*Johnson et al., 2018*), so the minicircularized genomes of *Columbicola* likely represent an independent transition from a single chromosome (Fig. 3). It is possible that fragmentation occurred multiple times within the *Columbicola* genus, which could explain why we recovered fewer minicircles in *C. columbae.* However, given the similar levels of fragmentation and conserved gene order in some minicircles among the four *Columbicola* species, fragmentation is at most an infrequent phenomenon within the genus
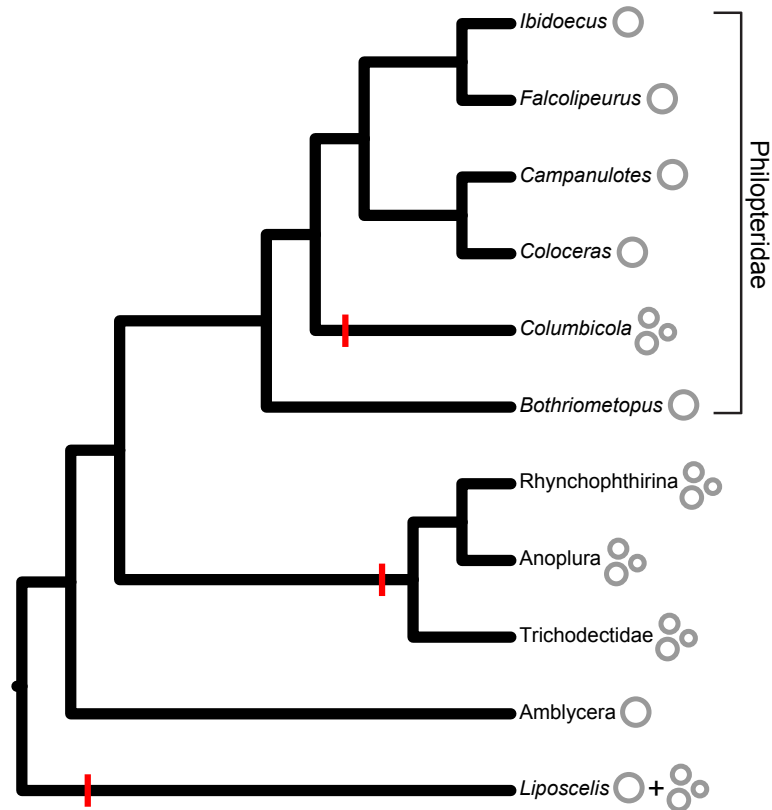
Figure 3 **Mitochondrial minicircles have evolved multiple times in parasitic lice and free-living relatives (*Liposcelis*).** The cladogram is adapted from *Johnson et al. (2018)* and features taxa with published full mitochondrial genomes, including *Columbicola* from the current study. Genome architecture (single or fragmented) is indicated next to the tip names. Presumed transitions from single to fragmented mitochondrial genomes are represented by red lines on the relevant branches.

and most parsimoniously represents a single independent origin. A more densely-sampled analysis is needed to confirm the history of fragmentation within the genus.

Cameron et al. (2011) identified partial fragmented mitochondrial genomes in three other genera of Philopteridae (*Anaticola*, *Quadraceps*, and *Philopterus*), so it is possible other transitions to fragmented genomes from single chromosomes have occurred in these feather lice. However, none of these taxa possess the minicircle-type genome architecture (i.e., few genes per chromosome and a relatively long CR with conserved motifs) reported here for *Columbicola*. The independent origin of mitochondrial minicircles in *Columbicola* and mammal-associated lice allows for phylogenetically independent examination of the life-history traits proposed to be related to this phenomenon. As *Columbicola* are feather-feeding and associated with birds, the proposals that mitochondrial minicircularization in lice is related to blood- or mammal feeding are unlikely (c.f. *Dong et al., 2014b*). Similarly, the generation-time (24 days) of *Columbicola*, which is the most fragmented mitochondrial genome yet sequenced, falls within the range of generation times (*Martin, 1934*) for other lice whose mitochondrial genomes are complete, or at least less fragmented, suggesting

that genome fragmentation is not causally related to short generation times. Multiple independent transitions from single to fragmented mitochondrial chromosomes in one group of organisms (parasitic lice) do, however, suggest that there are underlying potentially shared mechanisms for the formation of minicircles, such as the lack of a functional mtSSB gene, and/or similar selective pressures retained across relatively deep evolutionary time (~65 my). Future work can address these questions by building on the genome architectural framework established in this study.

CONCLUSIONS

The results from this study demonstrate that highly-fragmented, mitochondrial minicircles are present in species of the feather louse genus *Columbicola*. Based on a well-supported phylogeny of parasitic lice, mitochondrial minicircles in *Columbicola* also imply that mitochondrial genome fragmentation originated multiple times in lice, which is surprising given the reported detrimental effects of fragmentation and how infrequently the phenomenon has been observed across Metazoa. These results should encourage further investigation of mitochondrial architecture in insects and other animals. Additionally, we were able to assemble the complete *Columbicola* mitochondrial genome from NGS data using a fully bioinformatic approach, thus contributing to a methodological trend that has had increasing momentum in the last decade. Applying this approach to *P. humanus*, a louse species with known mitochondrial minicircles, showed the method to be reliable and consistent with previous laboratory-based methods for confirming the presence of minicircles. As NGS data becomes more commonplace in sequenced-based studies, it is important to continue developing approaches that can extract useful information, particularly for organisms with challenging or uncommon features.

ACKNOWLEDGEMENTS

Alvaro Hernandez and Christ Wright at the Roy J. Carver Biotechnology Center (University of Illinois, Champaign, IL, USA) were instrumental in facilitating genome library preparation and sequencing. We also thank Jyothi Thimmapuram at the Purdue University Bioinformatics Core (West Lafayette, IN, USA) for helping to set up and run software on the Purdue cluster.

Funding

This work was supported by the National Science Foundation DBI-1906262 to Andrew D. Sweet; National Science Foundation DEB-1239788 and DEB-1342604 to Kevin P. Johnson; and the Australian Research Council FT120100746 to Stephen L. Cameron. Publication of this article was funded in part by the Purdue University Libraries Open Access Publishing Fund. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
National Science Foundation: DBI-1906262.
National Science Foundation: DEB-1239788, DEB-1342604.
Australian Research Council: FT120100746.
Purdue University Libraries Open Access Publishing Fund.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Andrew D. Sweet conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Kevin P. Johnson conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Stephen L. Cameron conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

Raw genomic read data are available on the NCBI SRA: SRR8177102.

The assembled mitochondrial sequences are available at the Illinois Data Bank: Sweet, Andrew; Johnson, Kevin; Cameron, Stephen (2019): Data from: Mitochondrial genomes of Columbicola feather lice are highly fragmented, indicating repeated evolution of minicircle-type genomes in parasitic lice . University of Illinois at Urbana-Champaign. https://doi.org/10.13012/B2IDB-2211060_V2 and at GenBank: MT094266– MT094299 and MT113902–MT113932.

Data Availability

Raw genomic read data are available in Supplemental Information and at the NCBI SRA: SRR8177102.

The assembled mitochondrial sequences are available at the Illinois Data Bank: Sweet, Andrew; Johnson, Kevin; Cameron, Stephen (2019): Data from: Mitochondrial genomes of Columbicola feather lice are highly fragmented, indicating repeated evolution of minicircle-type genomes in parasitic lice . University of Illinois at Urbana-Champaign. https://doi.org/10.13012/B2IDB-2211060_V1.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.8759#supplemental-information.

REFERENCES

Allen JM, LaFrance R, Folk RA, Johnson KP, Guralnick RP. 2018. aTRAM 2.0: an improved, flexible locus assembler for NGS data. *Evolutionary Bioinformatics* **14**:117693431877454 DOI 10.1177/1176934318774546.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**:403–410 DOI 10.1016/S0022-2836(05)80360-2.

Amaral FR, Neves LG, Resende Jr MFR, Mobili F, Miyaki CY, Pellegrino KCM, Biondo C. 2015. Ultraconserved elements sequencing as a low-cost source of complete mitochondrial genomes and microsatellite markers in non-model amniotes. *PLOS ONE* **10**:e0138446 DOI 10.1371/journal.pone.0138446.

Antipov D, Hartwick N, Shen M, Raiko M, Lapidus A, Pevzner PA. 2016. plasmidSPAdes: assembling plasmids from whole genome sequencing data. *Bioinformatics* **32**:3380–3387 DOI 10.1093/bioinformatics/btw493.

Aquadro CF, Greenberg BD. 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* **103**:287–312.

Bankevich A, Nruk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* **19**:455–477 DOI 10.1089/cmb.2012.0021.

Bender A, Krishnan KJ, Morris CM, Taylor GA, Reeve AK, Perry RH, Jaros E, Hersheson JS, Betts J, Klopstock T, Taylor RW, Turnbull DM. 2006. High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and Parkinson disease. *Nature Genetics* **38**:515–517 DOI 10.1038/ng1769.

Bendich AJ. 1993. Reaching for the ring: the study of mitochondrial genome structure. *Current Genetics* **24**:279–290 DOI 10.1007/BF00336777.

Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, Putz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* **69**:313–319 DOI 10.1016/J.YMPEV.2012.08.023.

Boore JL. 1992. Animal mitochondrial DNA: structure and evolution. *International Review of Cytology* **141**:173–216 DOI 10.1016/S0074-7696(08)62066-5.

Boyd BM, Allen JM, Nguyen N, Sweet AD, Warnow T, Shaprio MD, Villa SM, Bush SE, Clayton DH, Johnson KP. 2017. Phylogenomics using target-restricted assembly resolves intra-generic relationships of parasitic lice (Phthiraptera: *Columbicola*). *Systematic Biology* **66**:896–911 DOI 10.1093/sysbio/syx027.

Bush SE, Kim D, Reed M, Clayton DH. 2010. Evolution of cryptic coloration in ectoparasites. *The American Naturalist* **176(4)**:529–535 DOI 10.1086/656269.

Bush SE, Malenke JR. 2008. Host defense mediates interspecific competition in parasites. *Journal of Animal Ecology* **77(3)**:558–564 DOI 10.1111/j.1365-2656.2007.01353.x.

Bush SE, Price RD, Clayton DH. 2009. Descriptions of eight new species of feather lice in the genus *Columbicola* (Phthiraptera: Philopteridae), with a comprehensive world checklist. *Journal of Parasitology* **95**:286–294 DOI 10.1645/GE-1799.1.

Cameron SL, Beckenbach AT, Dowton M, Whiting MF. 2006. Evidence from mitochondrial genomics on interordinal relationships in insects. *Arthropod Systematics and Phylogeny* **64**:27–34.

Cameron SL. 2014a. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annual Review of Entomology* **59**:95–117 DOI 10.1146/annurev-ento-011613-162007.

Cameron SL. 2014b. How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. *Systematic Entomology* **39**:400–411 DOI 10.1111/syen.12071.

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. *Journal of Molecular Evolution* **65**:589–604 DOI 10.1007/s00239-007-9042-8.

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 DOI 10.1146/annurev-ento-011613-162007.

Carlsen T, Aas AB, Lindner D, Vrålstad T, Schumacher T, Kauserud H. 2012. Don't make a mista(g)ke: is tag switching an overlooked source of error in amplicon pyrosequencing studies? *Fungal Ecology* **5**:747–749 DOI 10.1016/j.funeco.2012.06.003.

Chen XJ, Butow RA. 2005. The organization and inheritance of the mitochondrial genome. *Nature Reviews Genetics* **6**:815–825 DOI 10.1038/nrg1708.

Chevreux B, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information. *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics* **99**:45–56.

Clayton DA. 1982. Replication of animal mitochondrial DNA. *Cell* **28**:693–705 DOI 10.1016/0092-8674(82)90049-6.

Clayton DH, Bush SE, Johnson KP. 2016. *Coevolution of life on hosts: integrating ecology and history.* Chicago: The University of Chicago Press.

Clayton DH, Johnson KP. 2003. Linking coevolutionary history to ecological process: doves and lice. *Evolution* **57**:2335–2341 DOI 10.1111/j.0014-3820.2003.tb00245.x.

Clayton DH, Lee PLM, Tompkins DM, Brodie ED. 1999. Reciprocal natural selection on host-parasite phenotypes. *The American Naturalist* **154**:261–270 DOI 10.1086/303237.

Covacin C, Shao R, Cameron S, Barker SC. 2006. Extraordinary number of gene rearrangements in the mitochondrial genomes of lice (Phthiraptera: Insecta). *Insect Molecular Biology* **15**:63–68 DOI 10.1111/j.1365-2583.2005.00608.x.

Dickey AM, Kumar V, Morgan JK, Jara-Cavieres A, Shatters Jr RG, McKenzie CL, Osborne LS. 2015. A novel mitochondrial genome architecture in thrips (Insecta: Thysanoptera): extreme size asymmetry among chromosomes and possible recent control region duplication. *BMC Genomics* **16**:439 DOI 10.1186/s12864-015-1672-4.

Dong W-G, Song S, Guo X-G, Jin D-C, Yang Q, Barker SC, Shao R. 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 DOI 10.1186/1471-2164-15-751.

Dong W-G, Song S, Jin D-C, Guo X-G, 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 DOI 10.1186/1471-2164-15-44.

Gibson T, Blok VC, Dowton M. 2007. Sequence and characterization of six mitochondrial subgenomes from *Globodera rostochiensis*: multipartite structure is conserved among close nematode relatives. *Journal of Molecular Evolution* **65**:308–315 DOI 10.1007/s00239-007-9007-y.

Gibson T, Blok VC, Phillips MS, Hong G, Kumarasinghe D, Riley IT, Dowton M. 2007. The mitochondrial subgenomes of the nematode *Globodera pallida* are mosaics: evidence of recombination in an animal mitochondrial genome. *Journal of Molecular Evolution* **64**:463–471 DOI 10.1007/s00239-006-0187-7.

Gómez-Rodriguez C, Crampton-Platt A, Timmermans MJ, Baselga A, Vogler AP. 2015. Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages. *Methods in Ecology and Evolution* **6**:883–894 DOI 10.1111/2041-210X.12376.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**:1072–1075 DOI 10.1093/bioinformatics/btt086.

Hayakawa M, Katsumata K, Yoneda M, Tanaka M, Sugiyama S, Ozawa T. 1996. Age-related extensive fragmentation of mitochondrial DNA into minicircles. *Biochemical and Biophysical Research Communications* **226**:369–377 DOI 10.1006/bbrc.1996.1363.

Herd KE, 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 DOI 10.1186/s12864-015-1843-3.

Jiang H, Barker SC, Shao R. 2013. Substantial variation in the extent of mitochondrial genome fragmentation among blood-sucking lice of mammals. *Genome Biology and Evolution* **5**:1298–1308 DOI 10.1093/gbe/evt094.

Johnson KP, Nguyen N, Sweet AD, Boyd BM, Warnow T, Allen JM. 2018. Simultaneous radiation of bird and mammal lice following the K-Pg boundary. *Biology Letters* **14**:Article 20180141 DOI 10.1098/rsbl.2018.0141.

Johnson KP, Reed DL, Hammond Parker SL, Kim D, Clayton DH. 2007. Phylogenetic analysis of nuclear and mitochondrial genes supports species groups for *Columbicola* (Insecta: Phthiraptera). *Molecular Phylogenetics and Evolution* **45**:506–518 DOI 10.1016/j.ympev.2007.07.005.

Kajander OA, Rovio AT, Majamma K, Poulton J, Spelbrink JN, Holt IJ, Karhunen PJ, Jacobs HT. 2000. Human mtDNA sublimons resemble rearranged mitochondrial

genomes found in pathological states. *Human Molecular Genetics* **9**:2821–2835 DOI 10.1093/hmg/9.19.2821.

Karnkowska A, Vacek V, Zubacova Z, Treitl SC, Petrzelkova R, Eme L, Novak L, Zarsky V, Barlow L, Herman EK, Hroudova M, Dolezal P, Stairs CW, Roger AJ, Elias M, Dacks JB, Vlcek C, Hampl V. 2016. A eukaryote without a mitochondrial organelle. *Current Biology* **26**:1274–1284 DOI 10.1016/J.CUB.2016.03.053.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**:3059–3066 DOI 10.1093/nar/gkf436.

Linard B, Crampton-Platt A, Moriniere J, Timmermans MJTN, Anduja C, Arribas P, Miller KE, Lipecki J, Favreau E, Hunter A, Gomez-Rodriguez C, Barton C, Nie R, Gillett CPDT, Breeschoten T, Bocak L, Vogler AP. 2018. The contribution of mitochondrial metagenomics to large-scale data mining and phylogenetic analysis of Coleoptera. *Molecular Phylogenetics and Evolution* **128**:1–11 DOI 10.1016/j.ympev.2018.07.008.

Machado DJ, Janies D, Brouwer C, Grant T. 2018. A new strategy to infer circularity applied to four new complete frog mitogenomes. *Ecology and Evolution* **8**:4011–4018 DOI 10.1002/ece3.3918.

Marshall AG. 1981. *The ecology of ectoparasitic insects.* London: Academic Press.

Martin M. 1934. Life history and habits of the pigeon louse (*Columbicola columbae* [Linnaeus]). *Canadian Entomologist* **66**:6–16 DOI 10.4039/Ent666-1.

Nelson BC, Murray MD. 1971. The distribution of Mallophaga on the domestic pigeon (*Columba livia*). *International Journal for Parasitology* **1**:21–29 DOI 10.1016/0020-7519(71)90042-7.

Pelissier P, Camougrand N, Velours G, Guerin M. 1995. NCA3, a nuclear gene involved in the mitochondrial expression of subunits 6 and 8 of the Fo-F1 ATP synthase of *S. cerevisiae*. *Current Genetics* **27**:409–416 DOI 10.1007/BF00311209.

Pestryakov PE, Lavrik OI. 2008. Mechanisms of single-stranded DNA-binding protein functioning in cellular DNA metabolism. *Biochemistry* **73**:1388–1404 DOI 10.1134/S0006297908130026.

Price RD, Hellenthal RA, Palma RL, Johnson KP, Clayton DH. 2003. *The chewing lice: world checklist and biological overview.* Champaign: Illinois Natural History Survey.

Ruhanen H., et al 2010. Mitochondrial single-stranded DNA binding protein is required for maintenance of mitochondrial DNA and 7S DNA but is not required for mitochondrial nucleoid organisation. *Biochimica et Biophysica Acta—Bioenergetics* **1803**:931–939 DOI 10.1016/J.BBAMCR.2010.04.008.

Sánchez-Martínez A, Ningguang L, Clemente P, Adán C, Hernández-Sierra R, Ochoa P, Fernández-Moreno MA, Kaguni LS, Garesse R. 2006. Modeling human mitochondrial diseases in flies. *Biochim Biophys Acta, Bioenerg* **1757**:1190–1198 DOI 10.1016/j.bbabio.2006.05.008.

Schnell IB, Bohmann K, Gilbert MT. 2015. Tag jumps illuminated-reducing sequencing-to-sample misidentification in metabarcoding studies. *Molecular Ecology Resources* **15**:1289–1303 DOI 10.1111/1755-0998.12402.
Senior AE. 1988. ATP synthesis by oxidative phosphorylation. *Physiological Reviews* **68**:177–231 DOI 10.1152/physrev.1988.68.1.177.

Shao R, Barker SC, Li H, Song S, Poudel S, Su Y. 2015. Fragmented mitochondrial genomes in two suborders of parasitic lice of eutherian mammals (Anoplura and Rhynchophthirina, Insecta). *Scientific Reports* **5**:17389 DOI 10.1038/srep17389.

Shao R, Campbell NJH, Barker SC. 2001. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). *Molecular Biology and Evolution* **18**:858–865 DOI 10.1093/oxfordjournals.molbev.a003867.

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 Research* **19**:904–912 DOI 10.1101/gr.083188.108.

Shao R, Li H, Barker SC, Song S. 2017. The mitochondrial genome of the guanaco louse, *Microthoracius praelongiceps*: insights into the ancestral mitochondrial karyotype of sucking lice (Anoplura, Insecta). *Genome Biology and Evolution* **9**:431–445 DOI 10.1093/gbe/evx007.

Shao R, Zhu X-Q, Barker SC, Herd K. 2012. Evolution of extensively fragmented mitochondrial genomes in the lice of humans. *Genome Biology and Evolution* **4**:1088–1101 DOI 10.1093/gbe/evs088.

Shi Y, Chu Q, Wei D-D, Qui Y-J, Shang F, Dou W, Wang J-J. 2016. The mitochondrial genome of booklouse, *Liposcelis sculptilis* (Psocoptera: Liposcelididae) and the evolutionary timescale of *Liposcelis*. *Scientific Reports* **6**:30660 DOI 10.1038/srep30660.

Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. 2009. ABySS: a parallel assembler for short read sequence data. *Genome Research* **19**:1117–1123 DOI 10.1101/gr.089532.108.

Smith DR, Kayal E, Yanagihara AA, Collins AG, Pirro S, Keeling PJ. 2012. First complete mitochondrial genome sequence from a box jellyfish reveals a highly fragmented linear architecture and insights into telomere evolution. *Genome Biology and Evolution* **4**:52–58 DOI 10.1093/gbe/evr127.

Smith DR, Keeling PJ. 2015. Mitochondrial and plastid genome architecture: re-occurring themes, but significant differences at the extremes. *Proceedings of the National Academy of Sciences of the United States of America* **112**:10177–10184 DOI 10.1073/pnas.1422049112.

Song F, Li H, Lio G-H, Wang W, James P, Colwell DD, Tran A, Gong S, Cai W, Shao R. 2019. Mitochondrial genome fragmentation unites the parasitic lice of eutherian mammals. *Systematic Biology* **68**:430–440 DOI 10.1093/sysbio/syy062.

Ströher PR, Zarza E, Tsai WLE, McCormack JE, Feitosa RM, Pie MR. 2017. The mitochondrial genome of *Octostruma stenognatha* and its phylogenetic implications. *Insectes Sociaux* **64**:149–154 DOI 10.1007/s00040-016-0525-8.

Sweet AD, Johnson KP. 2016. Cophylogenetic analysis of New World ground-doves (Aves: Columbidae) and their parasitic wing lice (Insecta: Phthiraptera: *Columbicola*). *Molecular Phylogenetics and Evolution* **103**:122–132 DOI 10.1016/j.ympev.2016.07.018.

Sweet AD, Johnson KP. 2018. The role of parasite dispersal in shaping a host–parasite system at multiple evolutionary scales. *Molecular Ecology* **27**:5104–5119 DOI 10.1111/mec.14937.

Voigt O, Erpenbeck D, Worheide G. 2008. A fragmented metazoan organellar genome: the two mitochondrial chromosomes of *Hydra magnipapillata*. *BMC Genomics* **9**:350 DOI 10.1186/1471-2164-9-350.

Wallace DC. 1982. Structure and evolution of organelle genomes. *Microbiological Reviews* **46**:208–240 DOI 10.1128/MMBR.46.2.208-240.1982.

Wei D-D, Shao R, Yuan M-L, Dou W, Barker SC, Wang J-J. 2012. The multipartite mitochondrial genome of *Liposcelis bostrychophila*: insights into the evolution of mitochondrial genomes in bilateral animals. *PLOS ONE* **7**:e33973 DOI 10.1371/journal.pone.0033973.

Yoneda M, Katsumata K, Hayakawa M, Tanaka M, Ozawa T. 1995. Oxygen stress induces an apoptotic cell death associated with fragmentation of mitochondrial genome. *Biochemical and Biophysical Research Communications* **209**:723–729 DOI 10.1006/BBRC.1995.1559.

Zeviani M, Moraes CT, DiMauro S, Nakase H, Bonilla E, Schon EA, Rowland LP. 1988. Deletions of mitochondrial DNA in Kearns-Sayre syndrome. *Neurology* **38**:1339–1346 DOI 10.1212/WNL.38.9.1339.