American Genetic Association

Journal of Heredity, 2022, 171–183
https://doi.org/10.1093/jhered/esab068
Original Article
Advance Access publication November 1 2021

OXFORD

Original Article

Frequent Paternal Mitochondrial Inheritance and Rapid Haplotype Frequency Shifts in Copepod Hybrids

Jeeyun Lee and Christopher S. Willett

From the Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA (Lee and Willett).

Address correspondence to J. Lee at the address above, or e-mail: jeyn978@gmail.com.

Received June 1, 2021; First decision August 23, 2021; Accepted October 29, 2021.

Corresponding Editor: Jose Lopez

Abstract

Mitochondria are assumed to be maternally inherited in most animal species, and this foundational concept has fostered advances in phylogenetics, conservation, and population genetics. Like other animals, mitochondria were thought to be solely maternally inherited in the marine copepod *Tigriopus californicus,* which has served as a useful model for studying mitonuclear interactions, hybrid breakdown, and environmental tolerance. However, we present PCR, Sanger sequencing, and Illumina Nextera sequencing evidence that extensive paternal mitochondrial DNA (mtDNA) transmission is occurring in inter-population hybrids of *T. californicus*. PCR on four types of crosses between three populations (total sample size of 376 F1 individuals) with 20% genome-wide mitochondrial divergence showed 2% to 59% of F1 hybrids with both paternal and maternal mtDNA, where low and high paternal leakage values were found in different cross directions of the same population pairs. Sequencing methods further verified nucleotide similarities between F1 mtDNA and paternal mtDNA sequences. Interestingly, the paternal mtDNA in F1s from some crosses inherited haplotypes that were uncommon in the paternal population. Compared to some previous research on paternal leakage, we employed more rigorous methods to rule out contamination and false detection of paternal mtDNA due to non-functional nuclear mitochondrial DNA fragments. Our results raise the potential that other animal systems thought to only inherit maternal mitochondria may also have paternal leakage, which would then affect the interpretation of past and future population genetics or phylogenetic studies that rely on mitochondria as uniparental markers.

Keywords: marine invertebrate, mitochondria, mitonuclear incompatibility, nuclear mitochondrial DNA fragment (NUMT), paternal leakage

In most animal taxa, maternal mitochondrial inheritance is widely accepted as the norm, as deviations from this mode of inheritance could have significant evolutionary consequences. If paternal mitochondrial DNA (mtDNA) were also transmitted, there could be lethal genomic conflict due to mitochondrial competition (Hastings 1992; Hurst and Hamilton 1992), zygotes may receive sperm mtDNA that is damaged from deleterious effects associated with heavy respiration (Allen 1996), or there could be an increased propagation of deleterious mitochondrial mutations (Xu 2005). Additionally, several previous studies in humans and mice suggest that high frequency of

heteroplasmy—having multiple mitochondrial haplotypes within a cell—decreases fitness (Chomyn et al. 1992; Wallace 1999; Sharpley et al. 2012). To prevent these potential deleterious effects of paternal mtDNA, animals have evolved various mechanisms to exclude paternal mitochondria, such as ubiquitination or degradation of sperm mitochondria in primates (Sutovsky et al. 1999, 2000; Thompson et al. 2003), fruit flies (Reilly and Thomas Jr. 1980; DeLuca and O'Farrell 2012), and nematodes (Al Rawi et al. 2011; Sato and Sato 2011).

Despite mechanisms to ensure maternal inheritance of mtDNA, reports of paternal mtDNA in the offspring—termed paternal leakage—have been accumulating in hybrids of various animal species. A subset of examples includes paternal leakage in heterospecific or conspecific hybrids of fruit flies (Kondo et al. 1990; Matsuura et al. 1991; Sherengul et al. 2006; Dokianakis and Ladoukakis 2014), mice (Gyllensten et al. 1991; Shitara et al. 1998), periodical cicadas (Fontaine et al. 2007), ticks (Mastrantonio et al. 2019), and nematodes (Hoolahan et al. 2011; Ross et al. 2016). There are several bivalve species in which mitochondria are doubly inherited from both parents as a norm (Zouros et al. 1994; Gusman et al. 2016). While possible artifacts have not been completely ruled out in all of these former cases, evidence was consistent with at least some paternal leakage in experimental or natural (Matsuura et al. 1991; Mastrantonio et al. 2019) hybrid populations, along with continued maternal transmission of mtDNA. Some of these studies additionally tested and showed that leaked paternal mtDNA could proliferate to later developmental stages (Sherengul et al. 2006; Fontaine et al. 2007) and persist for multiple generations (Gyllensten et al. 1991). This phenomenon likely has significant consequences for how we understand mitochondria and its applications, because the assumption that mitochondria are only inherited maternally in animals has been the basis for using mtDNA as important molecular markers in phylogenetic inference (Avise et al. 1987), population genetics (Wilson et al. 1985), and conservation (Moritz 1994). Additionally, the presence of paternal mtDNA would impact estimates of genetic diversity and gene flow (Takahata and Maruyama 1981; Chapman et al. 1982; Hoolahan et al. 2011).

The mechanism of frequent paternal leakage and the effects of the phenomenon on hybrid fitness are presently unknown. It is speculated that paternal leakage is more commonly found in hybrids compared to pure populations because it is easier to experimentally detect paternal mtDNA that is genetically different from the maternal mtDNA (Kaneda et al. 1995; Shitara et al. 1998; Sutovsky et al. 2000), or because the mechanism to recognize and remove sperm mtDNA may not function as well in a hybrid embryo, which has genetically divergent sperm mtDNA from a different population (White et al. 2008). Literature on mitochondrial heteroplasmy, where multiple mitochondrial haplotypes are present in a cell, suggest that certain haplotypes are able to increase in abundance rapidly (Yoneda et al. 1992; Blok et al. 1997; White et al. 1999; Brandstätter et al. 2004), but whether this contributes to increased detection of paternal leakage in hybrids has not been explored. Whether specific paternal mtDNA haplotypes and frequent paternal leakage affect hybrid fitness are also unknown. In several species ranging from seed beetles to centrarchid fishes with highly coadapted mitochondrial and nuclear genomes (Burton et al. 2013), hybrids exhibit decreased fitness due to a breakdown in mito-nuclear coadaptation (Burton and Barreto 2012; Wolff et al. 2014). Paternal leakage may have complicated effects on hybrid fitness in such systems with mitonuclear coadaptation: inherited paternal mtDNA may be beneficial if they interact with the paternal nuclear DNA to restore coadpated functions, or alternatively be harmful if they compete with maternal mtDNA and cause genomic conflict (Hastings 1992; Hurst and Hamilton 1992).

Here we report substantial evidence for paternal mtDNA inheritance in inter-population hybrids of the marine copepod *T. californicus*. *Tigriopus californicus* is a rocky intertidal copepod with many genetically distinct populations along the western coast of North America (Burton 1997; Edmands 2001; Willett and Ladner 2009). This species has previously been thought to solely inherit maternal mitochondria (Burton and Lee 1994), and has been an important model to study the role of mtDNA in inter-population hybrid breakdown (Burton 1990; Edmands 1999; Ellison and Burton 2008; Burton and Barreto 2012; Healy and Burton 2020). MtDNA sequences have diverged in allopatry among populations of *T. californicus*, and several lines of evidence suggest that the mitochondrial and nuclear genomes within each population have coevolved (Rawson and Burton 2002; Barreto and Burton 2013; Barreto et al. 2018). When different populations mate, the hybrids may no longer have a complete set of co-adapted mitochondrial and nuclear products from each parent population. This results in suboptimal cellular functions or lowered fitness, called "mitonuclear incompatibility" (Ellison and Burton 2008; Barreto et al. 2015). Since mitochondrial genes greatly influence the fitness of *T. californicus* hybrids, the presence of paternal mtDNA presents novel evolutionary repercussions to consider, like potential impacts of paternal mtDNA on mitonuclear interaction and hybrid fitness.

In this study, we find evidence for paternal leakage in F1 hybrids of crosses between three genetically and geographically distant *T. californicus* populations in California: Abalone Cove (AB), Santa Cruz (SC), and San Diego (SD). Mitochondrial genome-wide divergence between AB and SC is 20.7%, and between AB and SD is 20.8% (Barreto et al. 2018). These values are much greater than the divergence values for coding regions in the nuclear genome, which are 2.54% between AB and SC, and 2.3% between AB and SD (Pereira et al. 2016). Although *T. californicus* has strikingly high levels of mitochondrial divergence even compared to other species within the subphylum Crustacea (Lefébure et al. 2006), it is still classified as a single species because there are few morphological differences across populations (Monk 1941), no premating isolation between populations (Ganz and Burton 1995; Palmer and Edmands 2000), and inter-population hybrids are viable and fertile (Burton 1990; Edmands 1999). We incorporated various types of PCR and sequencing methods to detect paternal leakage in these inter-population hybrids.

Given the techniques we used, one concerning artifact that could lead to experimental results masquerading as paternal leakage are nuclear mitochondrial DNA fragments (NUMTs). NUMTs are non-functional pieces of mtDNA that happen to get incorporated into the nuclear genome during a double-stranded break repair (Lopez et al. 1994; Bensasson et al. 2001). Each NUMT is typically under 1 kb (Richly and Leister 2004; Pamilo et al. 2007; Gaziev and Shaĭkhaev 2010), but there are species-specific exceptions such as NUMTs of lengths up to 5.5 kb in *Biomphalaria glabrata* gastropods, 15.5 kb in *Crassostrea gigas* bivalves (Sun and Yang 2016), and up to 12.6 kb in *Limulus polyphemus* crustaceans (Simpson et al. 2017). While it is plausible that genome size and the number of NUMTs are correlated (Bensasson et al. 2001), it is unclear whether taxonomic relationships are correlated to the size of NUMT fragments (Bensasson et al. 2001). For instance, unlike *B. glabrata*, another gastropod species *Aplysia californica* has a maximum NUMT size of 460 bp (Sun

and Yang 2016), and the only other study that investigated NUMT fragment size in a crustacean species reports 182 bp as the maximum size (Kowal et al. 2020). NUMTs pose problems for PCR-based methods in general because poorly designed PCR primers can amplify both true mtDNA and NUMTs and lead to erroneous interpretation (Zhang and Hewitt 1996; Bensasson et al. 2001; White et al. 2008). Previous studies of paternal leakage often used the PCR approach, but most only used singular or few methods of distinguishing true paternal mtDNA from NUMTs in the paternal nuclear genome (Sherengul et al. 2006; Fontaine et al. 2007; Hoolahan et al. 2011). One of the most comprehensive ways to exclude NUMTs is to first detect NUMTs via sequence analysis of the genome, avoid those regions for diagnostic PCRs and sequencing, ensure there are no premature stop codons characteristic of NUMTs in the sequences, and work with paternal mtDNA fragments in F1s that are larger than the detected NUMTs (Zhang and Hewitt 1996; Bensasson et al. 2001; Triant and DeWoody 2007; Song et al. 2008; White et al. 2008). Such rigorous procedure has been mainly implemented in relation to phylogenetics and DNA barcoding (Triant and DeWoody 2007; Song et al. 2008; Kim et al. 2013), but we incorporated it to minimize NUMTs in our study of paternal leakage. We used multiple molecular techniques to rule out artifacts and method-specific biases for detecting paternal mtDNA in F1 hybrids, quantified the proportion of F1 hybrid individuals with paternal leakage, and examined which paternal mitochondrial haplotypes are inherited by F1 hybrids.

Materials and Methods

Sample Collection

Tigriopus californicus samples were collected in 2013–2014 from Abalone Cove, California (AB: 33°44'16"N, 118°22'31"W), Santa Cruz, CA (SC: 36°56'58"N, 122°02'49"W), and San Diego, CA (SD: 32°45'N, 117°15'W). For all populations, copepods were sampled from multiple pools within an outcrop, where pools are genetically similar and mix over time (Burton and Swisher 1984; Willett and Ladner 2009). Collected populations were kept in 35 parts per thousand (ppt) artificial salt water (Instant Ocean, Spectrum Brands, VA), in an environmental chamber at 20°C and 12:12 hour light:dark cycle. Copepods were reared in these lab conditions for multiple generations before being used in experiments.

Inter-Population Cross Set Up

Tigriopus californicus females can be fertilized only after the final molt into adult copepods (Egloff 1966). Therefore, virgin females from pure populations were obtained either by raising small immature copepodids in individual wells of a 24-well culture plate until maturity or by separating a mating pair with an immature female copepodid—a *T. californicus* male mate guards a female until maturity, and upon mating releases the female (Burton 1985). A virgin female from one population and a male from another population were combined in each well of a new 24-well culture plate for one week, and subsequently the male parent was taken out for DNA extraction. The female parent was typically left to produce offspring until its death to increase the sample size of F1s, thus we did not generally collect DNA from females except to verify the presence of maternal mtDNA in a small number of individuals. DNA was extracted from F1 nauplii, copepodids, and adults, which are different developmental stages in chronological order. Four inter-population crosses were set up, with at least 48 replicates for each: SC female (f) x AB male (m), ABf x SCm, SDf x ABm, and ABf x SDm.

NUMT Identification

Potential regions of NUMTs were identified for populations AB and SD by using BLAST of the mitochondrial genome of a population (AB: NCBI GenBank accession DQ917373 from Burton et al. 2007; SD: NCBI GenBank accession DQ913891.2 from Burton et al. 2007) to the entire genome assembly of that population (SD: genome v.2.1, NCBI SRA SRX469409 from Barreto et al. 2018). The published genomes for *T. californicus* population AB (Barreto et al. 2018) were constructed based on mapping sequences from this population to the SD population's relatively high-quality genome and therefore do not contain unique sequences not found in the SD genome (potentially including unique NUMTs in AB). Thus, lower quality de novo assemblies were made from the AB population in attempts to uncover novel NUMTs in this population. The Illumina 400 bp sequences from population AB (NCBI SRA SRX2746703 from Barreto et al. 2018) were run on the computational cluster at UNC using the default parameters of SPAdes (v. 3.7.0; Bankevich et al. 2012). The resulting AB genome was highly fragmented and contained a number of bacterial contaminant scaffolds. To help remove some of these contaminating sequences, reads were mapped back to the new AB assembly using BWA-MEM (Li and Durbin 2009) and the coverage depth calculated for each scaffold using the idxstats command in SAMtools. For the resulting AB assembly, the modal coverage was 70 and scaffolds were trimmed as follows: for scaffolds >10kb those with coverage less than 40 were removed (most were checked by BLAST searches and confirmed to be bacterial); for scaffolds between 10 kb and 1 kb removed scaffolds with coverages less than 20; for scaffolds between 1 kb and 400 bp trimmed out those with coverage less than 10; and finally all scaffolds less than 400 bp size were removed (the AB de novo assembly will be available in the Dryad repository). This trimming resulted in a final assembly size of 182 Mb for AB, which is comparable in size to the initial assembly of the SD genome (NCBI SRA SRX469409 from Barreto et al. 2018) with bacterial contaminants removed of 181Mb and the improved published assembly size of 191 Mb.

For BLAST searches for NUMTs, default BLAST parameters were used, and all hit sequences returned from the BLAST search were designated as NUMTs if they had less than 100% Identity match to the mitochondrial genome. 100% identity matched hit sequences were discarded because they were likely the mtDNA itself included in the whole genome sequencing. This approach included hit fragments that may be too short to claim as mitochondrial fragments in nuclear DNA (NUMTs), but they were included in the complete list of NUMTs to be conservative. Because our preliminary results showed very few hybrids with inherited paternal SC mtDNA, we will not detail here the efforts to identify NUMTs in population SC.

BLAST results for population AB returned some hits with 100% identity match to SC mtDNA, indicating a very low level of SC mitochondrial contamination (<0.3% of total mtDNA reads) in the published AB raw Illumina sequences (NCBI SRA SRX2746703 from Barreto et al. 2018), and such contaminant hits were excluded from the list of NUMTs. These contaminated AB stocks in the lab were collected before 2012, and Illumina sequencing of hybrid populations with AB parents from subsequent collections of the AB population showed no evidence for SC mtDNA contamination. None of the AB copepod collections from before 2012 were used for the present study. We only used the AB sequences (NCBI SRA SRX2746703 from Barreto et al. 2018) after de-contaminating them for finding NUMTs as described above, obtaining the consensus sequence of population AB, and identifying polymorphisms at a few nucleotide sites, thus the interpretation of the data from this present study should be minimally impacted. Even if there were some unremoved

SC mtDNA contaminants, that would mean we used more stringent standards than necessary to avoid NUMTs and polymorphic sites within some PCR primers, and that we potentially didn't identify Sanger sequencing errors at a few nucleotide sites. To detail how existing SC mtDNA contaminants were removed for uses other than identifying NUMTs, reads in the published AB Illumina data (NCBI SRA SRX2746703 from Barreto et al. 2018) that mapped to anywhere within the entire SC mitochondrial genome (NCBI GenBank accession DQ917374 from Burton et al. 2007) were excluded in CLC Genomics Workbench (v.8.0.3). Mapping parameters were stringent and resulted in 1080 Illumina reads mapping to SC mtDNA (out of total about 79 million reads), where the reads had few nucleotide mismatches compared to the SC mtDNA.

Individual DNA Extraction and Standard PCR

DNA was extracted from whole bodies of single copepods by immersing them in 20 uL of lysis buffer consisting of Tris hydrochloride, potassium chloride, proteinase K, and Tween 20 (Hoelzel and Green 1992). The mixture was run in a thermocycler at 65°C for 1 hour then 99°C for 15 min. When lysing the copepods, special care was taken to exclude contamination of algae and debris—copepods were put in individual bubbles of freshly made artificial salt water for a few minutes, pipetted onto filter paper, then transferred into lysis buffer using a metal probe sterilized with 95% ethanol.

The resulting DNA was used in polymerase chain reaction (PCR) and agarose gel electrophoresis to detect the presence of inherited paternal mtDNA in hybrid individuals. PCR primers were designed to anneal within the *COB* gene (see Supplementary Figure S1) while avoiding areas with NUMTs that we detected. The forward primers contained two consecutive nucleotides at the 3' end that were unique to each population in the cross pair (and several other nucleotide mismatches unique to each population along the entire primer), thus were predicted to anneal only to one of the populations in the cross pair (see Supplementary Figure S1). Reverse primers were universal to all three populations (see Supplementary Table S3). Numerous control PCR reactions using DNA from pure population individuals of varying life stages and sex verified that the forward primer specific to one population did not amplify mtDNA of a different population (see Supplementary Figures S2 and S3). These control reactions also verified that there was no off-target amplification of the nuclear genome. PCR was done with one forward and one reverse primer at a time, and genotyping was done using the Promega Corporation PCR kit (catalog # M8295) or the New England Biolabs PCR kit (cat # M0267) with population-specific thermocycler conditions (see Supplementary Table S4). PCR products were run on 1% agarose gel and stained in ethidium bromide, and the population origin of the mtDNA was distinguishable based on band size—482 bp was from AB, 938 bp from SC, and 939 bp from SD (Figures 1 and 3). Only F1s with a distinct band that matched the size of the paternal band were scored as having paternal leakage. To be conservative, F1s with a faint and indistinct band that matched the paternal band size were excluded from the paternal leakage count. While this approach may lead to an underestimate of paternal leakage, it ensures that false positive detection of paternal mtDNA is not included the data.

Plasmid Purification of Copepod Pools and Long Range PCR

To additionally rule out the possibility of NUMTs, which are generally short and under 1 kb (Gaziev and Shaĭkhaev 2010), mtDNA was enriched, and long range PCR was done to detect long fragments of

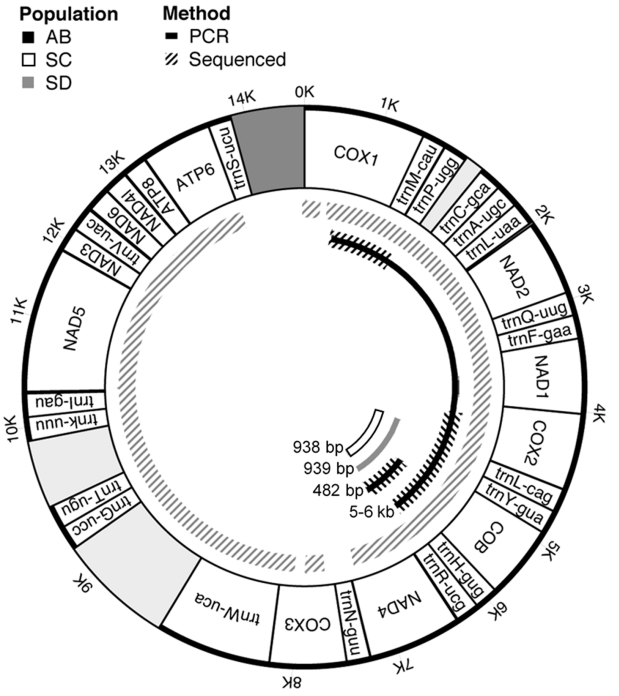

Figure 1. The locations of fragments resulting from PCR (standard PCR and long range PCR) and sequencing methods (Sanger sequencing and Nextera sequencing) along the entire *Tigriopus californicus* mtDNA (NCBI GenBank accession DQ917373 from Burton et al. 2007). Different experiments were done for mtDNA from the different *T. californicus* populations AB, SC, and SD. Inside of the donut-shaped gene map, the color and texture of the fragments indicate the population and experimental procedure (fragment sizes included if they were generated from PCR). Using SD as an example, sequencing was done along most of the mitochondrial genome, and PCR was done in the *COB* gene. Within the donut-shaped gene map itself, the dark grey region is the control region and the light grey regions are not annotated on NCBI GenBank.

paternal mtDNA in hybrids of one cross, SCf x ABm. Two pools of approximately 30 AB copepods and 30 SC copepods, each consisting of males, females with and without egg sacs were collected. Another pool of 40 F1s from the cross SCf x ABm, consisting of copepodids, males, and females was collected. MtDNA was enriched for each pool by using Qiagen's QIAprep Spin Miniprep Kit (catalog #27104), which aids in purifying plasmid DNA (Quispe-Tintaya et al. 2013).

The mtDNA-enriched DNA was subsequently used in long range PCR. Long range PCR primers were designed to span the mitochondrial *COB* gene to *COX1* gene (Figure 1), not anneal to NUMT regions, anneal specifically to the AB population and not the SC population by including two nucleotide mismatches at the 3' end (see Supplementary Table S3) and avoiding AB polymorphic nucleotide sites within the primers. PCR mixture was prepared using New England Biolabs' LongAmp® Taq DNA Polymerase (cat # M0323), and run in a thermocycler at 94°C for 30 s, cycled 27 times at 94°C for 30 s, 59°C for 15 s, and 65°C for 6 min 30 s, then final extension at 65°C for 10 min for SC and F1 DNA. For AB paternal population DNA, only 20 cycles were used because the band was too bright to identify the correct fragment size using 27 cycles. Long range PCR products were run on 1% agarose gel then stained in ethidium bromide, and if the sample contained paternal AB mtDNA, large 5–6 kb fragments were visible.

Sanger Sequencing of Standard PCR, Long Range PCR Products & Analysis

PCR products were sequenced using the paternal mtDNA primer to verify similarity to the paternal mtDNA sequence and to inspect if certain haplotypes are more likely to be transmitted to F1s. Specifically, standard PCR products from F1s of SCf x ABm, F1s of SDf x ABm, and the paternal population AB, and long range PCR products from F1s of SCf x ABm and the paternal population AB were purified and Sanger sequenced at Eton Bioscience Inc. using AB-specific primers. The maternal population samples from SD or SC did not produce any PCR bands using AB-specific primers, thus could not be sequenced. For sequencing of long range PCR products, an additional set of primers flanking the *COB* gene (see Supplementary Table S3) was used to obtain 2–3 kb of sequence.

TheSanger sequences were quality trimmed and mapped to the paternal population's common mitochondrial haplotype in Sequencher (v.5.2.4). To obtain the common AB haplotype, in CLC Genomics Workbench (v.8.0.3), the decontaminated AB Illumina reads (NCBI SRA SRX2746703 from Barreto et al. 2018) were aligned to the published *T. californicus* AB mitochondrial sequence (NCBI GenBank accession DQ917373 from Burton et al. 2007) and the extracted consensus sequence represented the most common haplotype in the general AB population. Of note, there were a few possible nucleotide errors in the published mtDNA sequence as there were a number of SNP variants between the published mtDNA and our consensus sequence of the AB Illumina data (see Supplementary Table S5). The control region was excluded from our consensus sequence because the coverage was unusually high with lots of repeats, impairing the ability to detect population specific mitochondrial reads. The resulting common AB haplotype was used as a reference to verify that haplotypes in AB male parents from this study were consistent with haplotypes in the general AB population. This was the case, as we identified only a few nucleotide mismatches between the most common AB haplotype and sequences of the male AB parent of crosses, and all nucleotide variants were present in the AB population Illumina reads.

Further analysis was done to test if all paternal mtDNA haplotypes in the pure AB and F1s of SCf x ABm and SDf x ABm contained any predicted stop codons in their coding regions, which is a common signature of NUMTs (Triant and DeWoody 2007; Song et al. 2008). First, NCBI blastx was done on the standard PCR and long range PCR sequences of AB to identify the reading frame and coding region. Then the translations of the leaked paternal AB haplotypes in the F1s were compared to that of translations from pure AB individuals in ExPASy (Artimo et al. 2012), to verify the presence/absence of stop codons within the coding region.

Illumina Nextera Sequencing & Analysis

A pool of 40 F1s (35 females, 5 males) from cross ABf x SDm was collected for plasmid DNA purification using Qiagen's QIAprep Spin Miniprep Kit (catalog # 27104). Library was prepped using Illumina's Nextera™ DNA Flex Library Prep kit and sequencing was done on the NovaSeq 6000.

The sequences were mapped to the common paternal haplotype in CLC Genomics Workbench (v.8.0.3). The common haplotype was obtained by aligning SD Illumina reads to the published *T. californicus* SD mitochondrial sequence and extracting the consensus sequence (Burton et al. 2007; Barreto et al. 2018). Again, the control region was excluded from the consensus sequence. The parameters used for aligning the Nextera reads to the most common AB and SD mtDNA haplotypes were stringent to ensure reads mapped to the correct population's mtDNA. The parameter settings were as follows: at least 98% of the total alignment length matched the reference sequence, and of that length at least 98% of the identity matched between the aligned read and the reference sequence. Additionally, the cost of mismatch between the read and reference sequence was set to 4 and non-specific matches were ignored; all other parameters were set to default values. This stringent approach may underestimate the true amount of paternal leakage, but provides stronger support for the presence of paternal mtDNA as it minimizes false positive detection of paternal leakage.

Contamination Tests

Several additional procedures were done to verify that paternal mtDNA PCR bands were not a result of laboratory contamination during the DNA lysis and PCR stage. Parts of steps in the individual DNA extraction and PCR procedure mentioned above were purposefully omitted to see if human error could have allowed contamination of paternal mtDNA and resulted in false detection of paternal leakage in the hybrid. A densely-populated AB petri dish was used to obtain DNA or to test for various potential sources of contamination since this was the paternal population for most crosses. In lysis buffer, we put clear water pipetted directly from the dish, water and algae pipetted from the dish (no living copepods were observed under microscope), and soaked a needle tool that was not wiped with ethanol after contacting an AB individual. Replicates of these samples were lysed, PCR amplified using AB-specific primers, and run on a gel along with positive control AB DNA.

Results

Frequent Paternal Leakage in Individual F1s and Long Paternal mtDNA Fragment in Pooled F1s

Surprisingly, standard PCR of individual F1 hybrids showed frequent paternal leakage for crosses SCf x ABm and SDf x ABm, with 31.1% and 58.5% of individuals showing leakage, respectively (Figure 2). For the reciprocal crosses ABf x SCm and ABf x SDm, there were fewer F1s with paternal leakage, with 2.0% and 11.7% respectively (Figure 2). The paternal bands in the F1s from both reciprocal

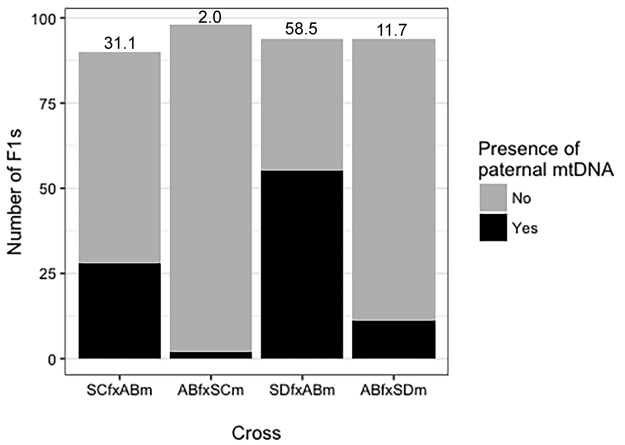

Figure 2. The number of *T. californicus* inter-population F1 offspring with and without paternal leakage based on standard PCR of the *COB* gene. The percentages of F1s with paternal leakage for each cross type are indicated at the top of the bars.

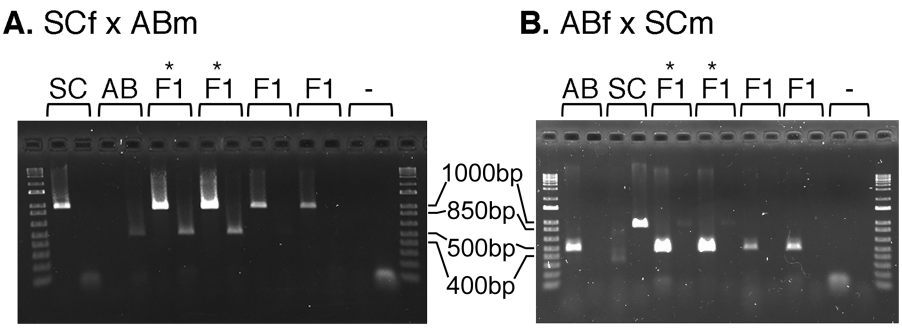

Figure 3. An example of an agarose gel image from standard PCR of the *COB* gene region using *Tigriopus californicus* individuals from the parental populations and F1s from several parent pairs. The first and last lanes contain the entire range of the Invitrogen™ 1kb Plus DNA ladder (cat # 10787018). The paired lanes show that the same individual's DNA was used twice for PCR: the first lane of the pair using maternal primers only and the second lane using paternal primers only. DNA from one individual from the maternal population, one individual from the paternal population, four F1 individuals, and negative controls consisting of the mastermix from each PCR condition without DNA are shown. F1s with an asterisk (*) above have paternal leakage as they show both maternal and paternal mtDNA bands. The expected size of SC mtDNA band is 938 bp and AB mtDNA band is 482 bp, and numbered arrows refer to the band sizes of the ladder. **A.** SCf x ABm parental population individuals and F1s, and **B.** ABf x SCm parental population individuals and F1s. The first two heteroplasmic F1s have faint but distinct paternal SC mtDNA bands.

crosses were relatively faint compared to the band in the paternal parent (Figure 3). Overall, the reciprocal crosses had fewer individuals showing paternal leakage. The continued maternal inheritance of mtDNA was verified by the robust PCR amplification of maternal bands in a subset of F1s from all crosses (see Supplementary Figure S4). Doing PCR with primers specific to the paternal population did not show any bands in individuals from the maternal population source (see Supplementary Figures S2 and S3). This suggests that potential haplotypes in the maternal population that might resemble the mitochondrial sequences of the paternal population were absent or at concentrations too low to be detected by PCR. This result also suggests there were no potential mis-primed sites elsewhere in the maternal genome. Additionally, the PCRs used to test for laboratory contamination showed no amplified bands on an agarose gel other than the positive control, demonstrating that contamination from the paternal population was unlikely in our samples (see Supplementary Figure S5).

Long range PCR on F1s of one representative cross type verified that long pieces of paternal mtDNA are present, minimizing the possibility of short NUMTs (Richly and Leister 2004; Pamilo et al. 2007; Gaziev and Shaĭkhaev 2010). Indeed, BLAST results of *T. californicus* showed that the average NUMT length was 223bp with a maximum of 2.4kb in population AB, and an average of 111bp and maximum of 403bp in SD (see Supplementary Table S1, Table S2). Long range PCR of the pool of SCf x ABm F1s resulted in a large 5-6kb fragment of AB paternal mtDNA in the agarose gel (Figure 1; see Supplementary Figure S6), suggesting the likely presence of paternal mtDNA in F1s as opposed to NUMTs. On the gel, no bands were visible for DNA isolated from the SC maternal population or the PCR negative controls, as expected since primers that only amplify AB mtDNA were used.

Sequencing of Leaked Paternal Fragments Reveals Haplotype Shifts in Some Crosses

Sanger sequencing of standard and long range PCR products verified that detected fragments in the hybrids are consistent with the paternal mtDNA sequence. Amplified mtDNA sequences in F1s from SCf x ABm and SDf x ABm were highly similar to the most common AB mitochondrial haplotype. For example, in the sequences of the standard PCR products, there were only four nucleotide mismatches out of about 400bp (*COB* region) in the F1s compared to AB (Table 1). This sequence divergence of 1% suggests the high likelihood that there were indeed AB mtDNA in hybrids, particularly when considering how the sequence divergence between the paternal and maternal population's *COB* region is around 20%.

The four nucleotide mismatches within the *COB* region showed evidence of substantial haplotype frequency shifts in the F1 generation (Table 1). At nucleotide positions 5622, 5636, and 5747, the nucleotides in all the F1s were different compared to their fathers in crosses involving both SC and SD females to AB males. At position 5801, a subset of the F1s had different nucleotides compared to the father. Except for the variant at position 5636, all variants in the F1s were present in the published Illumina sequence from pooled AB individuals (NCBI SRA SRX2746703 from Barreto et al. 2018), showing that those sites are heterozygous in the natural paternal population but F1s inherit the uncommon variant. There were no premature stop codons in all paternal haplotypes from AB and F1 individuals (Table 1).

Similarly, Sanger sequencing of long range PCR products showed that the sequence of the amplified band in the SCf x ABm F1s matched the majority of the paternal AB mtDNA sequence (Figure 4), but also included SNP variants that were uncommon in the AB population mtDNA haplotypes. Compared to the most common AB mtDNA haplotype, there were only 11 nucleotide mismatches in the putative paternal mtDNA sequence in pooled F1s, and one mismatch in the pooled ABs (Figure 4; Table 2) out of about 2,250bp. This further suggests the likely presence of paternal AB mtDNA in the F1s, and that some haplotypes are common in the F1 but at low frequency in the father individuals or in the general paternal population. Again, all variants in the F1s were found in the general AB population, except for position 5636, where T/C were present in the Sanger sequences while only A/C were present in the Illumina sequences of AB DNA (Table 1; Table 2). To verify that these variants are present in larger pools of AB individuals and not sequencing errors, we looked at published hybrid sequences of ABf x SDm F2s (Lima et al. 2019) and found each of the T/A/C variants present at position 5636. Of the 11 SNPs present in the sequenced region of the leaked paternal haplotypes in F1s, only one SNP was predicted to encode a premature stop codon (Table 2).

Nextera sequencing of ABf x SDm F1s showed some fragments that aligned to most of the length of the paternal mitochondrial

Table 1. Nucleotide variants in putative paternal AB mtDNA haplotypes in *COB* gene

Cross	**Position**[a]	5622	5636	5747	5801
	Nucleotide variants[b]	(T: 98.45% C: 1.55%)	(C: 99.96% A: 0.036%)	(T: 98.18% C: 1.78% A:0.04%)	(T: 99.72% A: 0.14% C: 0.07% G: 0.07%)
	Common AB haplotype[c]	**T**	C	T	T
SCf x ABm	ABm #1	T	C	T	T
	ABm #2	T	T[d]	T	T
	F1 #1 (father ABm #1)	C	C	C	T
	F1 #2 (father ABm #1)	C	C	C	T
	F1 #3 (father ABm #2)	C	C	C	A
SDf x ABm	ABm #3	T	T	T	T
	ABm #4	T	T	T	T
	F1 #4 (father ABm #3)	C	C	C	T
	F1 #5 (father ABm #3)	C	C	*NA*[e]	*NA*
	F1 #6 (father ABm #4)	C	C	C	A

All the paternal AB haplotype SNP variants in the Sanger sequences of standard PCR products in the *COB* coding region are shown. These sequences were generated from the actual AB male parent individuals and their offspring F1 individuals that have paternal leakage, from cross types SCf x ABm and SDf x ABm. The sequence lengths ranged from 450 to 471bp. Because Sanger sequencing does not detect low frequency alleles, only the most common haplotype or varying haplotypes that were present in similarly high proportions (i.e. heterozygous haplotypes) were obtained for each sequenced sample and reported here.

[a] The nucleotide positions are relative to the reference sequence, Common AB haplotype.

[b] The percentages of nucleotides at each position in the published AB genome containing Illumina reads of a pool of many AB individuals (NCBI SRA SRX2746703 from Barreto et al. 2018).

[c] The reference sequence was the most common AB haplotype, which was obtained by taking the consensus sequence from Illumina reads of many AB individuals (NCBI SRA SRX2746703 from Barreto et al. 2018).

[d] The T variant at position 5636 was not present in the Illumina reads used to obtain the reference sequence, but was present in other published *T. californicus* hybrid sequences (Lima et al. 2019).

[e] Missing nucleotides are indicated with '*NA*'.

sequence. Specifically, the Nextera sequences from ABf x SDm F1s covered 84% of the length of the consensus paternal SD mtDNA (Figure 1), albeit with low average coverage of 3.04 and a total of 296 reads. In contrast, when these F1 sequences were mapped to the consensus maternal AB mtDNA, 100% of the reference was covered, with average coverage of 33,448 reads and a total of 3,095,082 reads. Compared to the common haplotype in the general SD population, there were no SNP variants in the F1s' paternal mtDNA sequences. Standard PCR was done on the same DNA from ABf x SDm F1s used in Nextera sequencing but there was no amplification of SD paternal bands, consistent with very low levels of paternal mtDNA inheritance in this sample (and relatively low in this cross, see Figure 2). Although there is a small amount of paternal mtDNA in these F1s, the presence of fragments that match the majority of paternal mtDNA further decreases the possibility of NUMTs and suggests that the detection of paternal mtDNA is not method-specific.

Discussion

We found evidence consistent with a high prevalence of paternal leakage in the F1 hybrids of *T. californicus* inter-population crosses, with up to 59% of hybrid individuals from certain crosses having some level of paternal mitochondrial inheritance. We tried to verify the identity of the paternal mtDNA in the hybrids via contamination tests, two methods of PCR and several methods of sequencing. For some crosses, paternal haplotypes that were rare in the paternal population's mtDNAs were common in the F1s' mtDNA haplotypes.

The large extent of paternal leakage was unexpected in *T. californicus* hybrids, which were previously thought to only have maternal mitochondrial inheritance (Burton and Lee 1994). While most literature in other species report low copy numbers of paternal mtDNA or less than 10% of hybrids with paternal leakage, there are three studies that show comparable levels as our results for the SCf x ABm and SDf x ABm crosses. In periodical cicada hybrids, there was paternal leakage in 46% of offspring for a particular interspecific cross (Fontaine et al. 2007). There was 19–48% of paternal leakage in offspring of intraspecific *Drosophila* backcrosses, where the backcross consisted of F1 offspring crossed to a male from the paternal lineage (Sherengul et al. 2006). Additionally, 31–63% of offspring in interspecific *Drosophila* backcrosses showed paternal leakage (Sherengul et al. 2006). In 40% of hybrids from inter-population crosses of potato cyst nematodes, the paternal mtDNA replaced the maternal mtDNA (Hoolahan et al. 2011). Reports of frequent paternal leakage may increase if more directions of hybrid crosses are investigated. Several studies showed that reciprocal crosses of the same pair of populations could have contrasting levels of paternal leakage, particularly if one direction of the cross produces fewer or lower fitness offspring (Fontaine et al. 2007; Coleman-Hulbert 2010; Mastrantonio et al. 2019). In the present study, all crosses tested produced our target sample size of roughly 100 offspring (upon which the experiments were stopped) but the reciprocal crosses still showed drastically different prevalence of paternal leakage.

As paternal leakage is unexpected in animals, it is often met with fair criticism that artifacts could be inflating the frequency of paternal mtDNA (Luo et al. 2018; Balciuniene and Balciunas 2019; Lutz-Bonengel and Parson 2019). The previous studies with frequent paternal leakage mentioned above made some efforts to reduce contamination or NUMTs (although Hoolahan et al. (2011) were not explicit about how they did this), but we aimed to make a more thorough attempt to rule out artifacts. Sherengul et al. (2006) avoided NUMTs by amplifying PCR products over 1.8 kb, since there were fewer than ten known NUMTs totaling 500 bp in

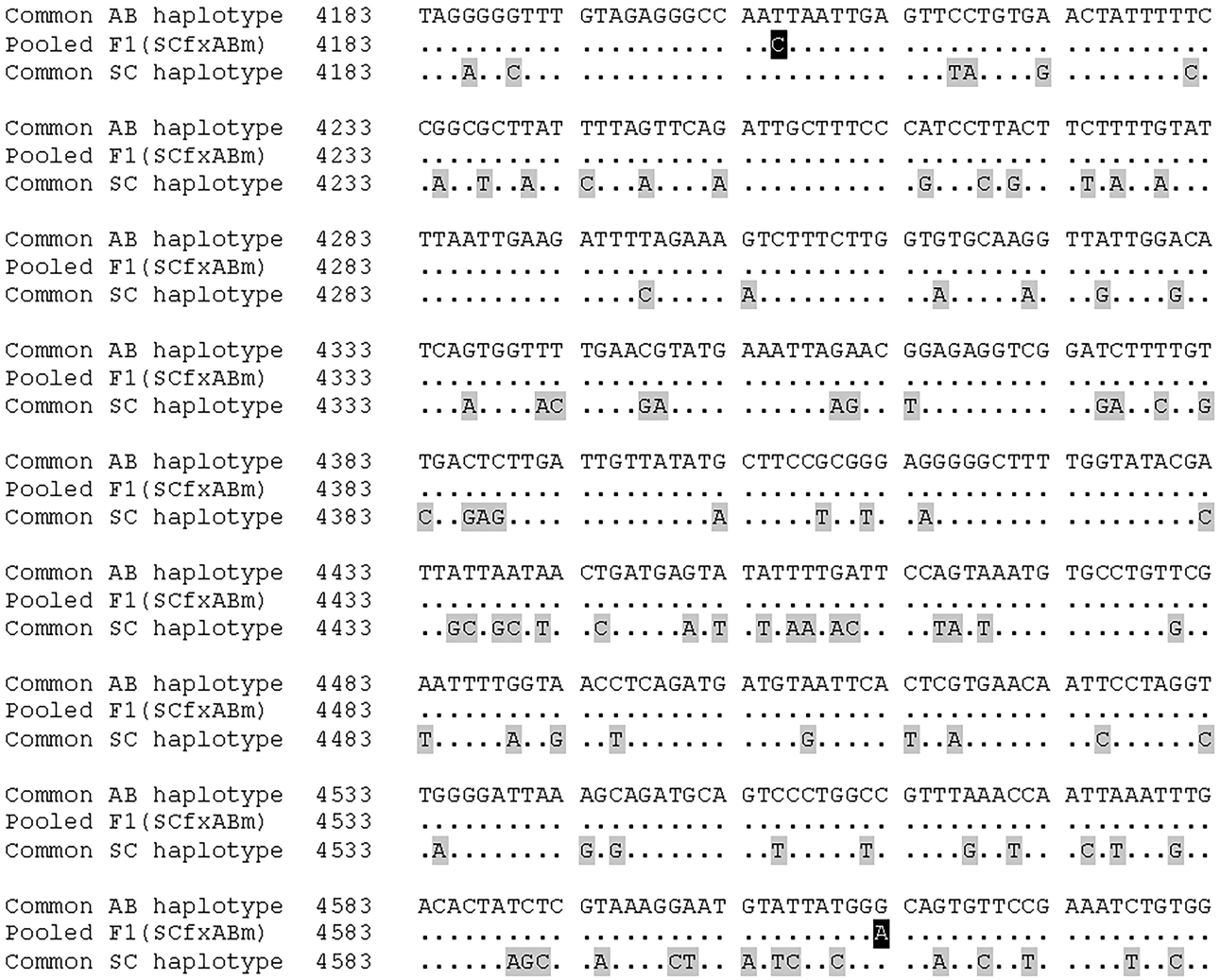

Figure 4. An example alignment of the most common AB haplotype, the most common SC haplotype, and a subset of the Sanger sequences of long range PCR fragments from plasmid-enriched and pooled SCf x ABm F1s in Table 2. Only information from nucleotide positions 4183 to 4633 are included as a representative subsample. SNPs are highlighted with black for the F1s, and grey for the SC sequences, and nucleotides that match the reference are shown with ".".

Drosophila melanogaster at the time (although subsequently greater numbers of NUMTs, some with length 1.5 kb, have been found in *D. melanogaster* subgroups) (Rogers and Griffiths-Jones 2012). We took a step further and made new comprehensive lists of NUMTs in relevant *T. californicus* populations, as NUMT discovery is regularly made and NUMTs can be species or population specific (Rogers and Griffiths-Jones 2012). Fontaine et al. (2007) tested for contamination, and minimized NUMTs since they found matching sequences of the *COX1* mitochondrial gene in the hybrids and paternal parent, and verified there were no stop codons in the hybrid paternal sequence. In addition to doing similar procedures to those in Fontaine et al. (2007), we tried to avoid NUMT regions at the primer design stage, and sequenced leaked paternal fragments longer than a single gene, since longer NUMTs are less common. To note, there were a number of gaps in the nuclear genome assemblies that we used, so some polymorphic NUMTs may not have been detected in the BLAST search, and there is a possibility we were PCR amplifying undetected polymorphic NUMT regions.

Other criticisms on reports of paternal leakage question the reliability of the PCR method to detect paternal mtDNA. The present study partially relied on the standard PCR approach for diagnosing paternal mtDNA, which has been widely used in previous literature to confirm paternal leakage (Sherengul et al. 2006, Fontaine et al. 2007, Dokianakis and Ladoukakis 2014, Ladoukakis and Zouros 2017). PCR methods do have several potential issues, for instance, the annealing efficiency of primers may vary across populations or even across haplotypes within a population, causing a bias in the frequency of detected paternal leakage. In the present study, the results from crosses SDf x ABm and SCf x ABm are still comparable to each other because the same AB population primers were used to diagnose paternal AB mtDNA. It may be that primers for populations SD and SC were worse at annealing compared to AB primers, leading to low amount of paternal leakage in crosses ABf x SDm and ABf x SCm. Spurious binding of primers was avoided by extensively testing all population primers on many pure population samples before using the primers for detection of paternal mtDNA in hybrids (see Supplementary Figures S2 and S3). As can happen with most other molecular methods, there could have been experimental errors while setting up the PCR reaction that led to paternal mtDNA not being detected in maternal population samples. However, the absence of paternal mtDNA bands in maternal population individuals was verified using many samples (see Supplementary Figures S2 and S3), varying

Table 2. Nucleotide variants in putative paternal AB mtDNA haplotypes in 5–6 kb fragments

	AB: 410-1352 [a] F1: 453-1049				AB: 4026-4892 F1: 4057-4994					AB: 4890-5785 F1: 4889-5504			
Position[b]	**825**	**865**	**971**	**1039**	**4205**	**4612**	**4784**	**4844**	**4849**	**5040**	**5084**	**5622**	**5636**
Nucleotide variants (%)[c]	T: 98.86 A: 1.11	T: 99.91 G: 0.093	A: 100	G: 99.97 – [d]: 0.034	T: 61.12 C: 38.85	G: 99.40 A: 0.60	A: 99.61 G: 0.36 C: 0.03	C: 98.53 T: 1.47	T: 98.33 C: 1.52	C: 99.80 T: 0.20	C: 98.64 T: 1.31	T: 98.45 C: 1.55	C: 99.96 A: 0.036
Common AB haplotype[e]	**T**	**T**	**A**	**G**	**T**	**G**	**A**	**C**	**T**	**C**	**C**	**T**	**C**
Pooled AB	T	T	A	G	T	G	A	C	T	C	C	T	T/C
Pooled F1 (SCf x ABm)	A	G	N[f]	C	C	A/G	G/A	T	C	T[g]/C	T	*NA*[h]	*NA*

All of the paternal AB haplotype nucleotide variants in the Sanger sequences of long range PCR products of length 5–6 kb that were generated from pooled ABs and pooled SCf x ABm F1s are shown. Because Sanger sequencing does not detect low frequency alleles, only the most common haplotype or varying haplotypes that were present in similarly high proportions (i.e. heterozygous haplotypes) were obtained for each sequenced sample and reported here.

[a] There were three sequenced fragments with nucleotide positions indicated in the first row, totaling 2140-2700 bp of sequence.

[b] The nucleotide positions are relative to the reference sequence, Common AB haplotype.

[c] The percentages of nucleotides at each position in the published AB genome containing Illumina reads of a pool of many AB individuals (NCBI SRA SRX2746703 from Barreto et al. 2018).

[d] The symbol '–' indicates there is a deletion in some AB Illumina reads aligned to the common AB haplotype.

[e] The reference sequence was the most common AB haplotype, which was obtained by taking the consensus sequence from Illumina reads of many AB individuals (NCBI SRA SRX2746703 from Barreto et al. 2018).

[f] 'N' corresponds to an unknown nucleotide.

[g] The T at position 5040 results in a premature stop codon.

[h] '*NA*' indicates missing bases because the F1 sequence ended before position 5622.

stocks of PCR reagents, several PCR machines, multiple PCR trials of positioning the samples in varying locations within the PCR machines, as well as taking other experimental precautions. Moreover, paternal mitochondrial sequences were not found in the Illumina sequences of the pool of maternal population individuals, consistent with the absence of paternal PCR bands.

While we cannot rule out the theoretical presence of extremely low frequency paternal-like maternal mtDNA that remained undetected in this study, we point out additional characteristics of the paternal-like mtDNA that we did detect which suggest these detected haplotypes likely did not originate from maternal populations. Firstly, our data in Table 1 shows two AB paternal-like haplotypes that are exactly the same with nucleotide variants at the same set of positions, in F1s from two different maternal population sources—SC and SD. The simplest way to explain this observation is that these shared AB-like mtDNA haplotypes came from the paternal AB parent rather than independently from the geographically distant SC and SD maternal parents. Secondly, Figure 4 shows that the detected AB-like haplotype in a pool of many SCf x ABm F1s is only 2 nucleotides different from the common paternal AB mtDNA haplotype, but 98 nucleotides different from the common maternal SC mtDNA haplotype within a subset of the Sanger sequenced long range PCR products mentioned in Table 2. It is difficult to explain how a typical SC maternal haplotype would acquire such high proportions of nucleotide substitutions in a manner that converges to the AB haplotype. Similarly in our Nextera sequencing data, the putative SD paternal mtDNA was an exact match of the common SD haplotype but about 20% divergent from AB maternal mtDNA haplotype.

Another interesting result from the present study was that a rare mitochondrial haplotype in the *T. californicus* male parents became the most common paternal haplotype in the F1s of some crosses. If paternal mtDNA were inherited through random sampling of various haplotypes, the most common haplotype in the father would be expected to also be predominant in the offspring. Any deviation in haplotype frequency between generations would indicate some other factor is playing a role in inheritance, such as genetic drift or natural selection. In *T. californicus*, it appears that higher frequency of paternal leakage could be correlated with the prevalence of the uncommon paternal haplotype. In our results, 31% of F1s of SCf x ABm had paternal leakage and had the rare paternal haplotype. Meanwhile, 12% of F1s of ABf x SDm had paternal leakage but low paternal mtDNA coverage in Nextera sequences, and had the most common SD paternal haplotype.

Previous literature on drastic frequency shifts in maternal mitochondrial haplotypes over a short period of time can give insight into our results. For example, a disease inducing point mutation in the human mitochondria was able to proliferate after one generation (Blok et al. 1997; White et al. 1999), and was even capable of becoming the most common haplotype (Blok et al. 1997). In two independent human families, a point mutation in the mother's mitochondrial control region became fixed in the child's mitochondrial genome (Brandstätter et al. 2004). It was speculated that one mitochondrion with the point mutation was preferentially amplified in the oocyte, or that subsets of mtDNA segregated and amplified (Blok et al. 1997; White et al. 1999; Brandstätter et al. 2004). Similar to how specific maternal haplotypes can increase in frequency, particular *T. californicus* paternal haplotypes could be preferentially amplified in the hybrid if they confer replication advantage—for instance a certain paternal haplotype will be amplified more if it is shorter than another haplotype due to a deletion (Yoneda et al. 1992). The premature stop codon in one leaked paternal haplotype in our study (Table 2) indicates some disruption of function, but we don't know if this stop codon results in large deletions that confer a replicative advantage. For future work, comparing the genetic structures of the various paternal haplotypes would give insight on how haplotype frequencies can shift dramatically.

Several studies provide clues about the mechanism of paternal mtDNA transmission in hybrids. In water frogs, a female hybrid with paternal mtDNA could pass those copies to the offspring (Radojičić et al. 2015), showing that paternal mtDNA can be passed through several generations. Paternal leakage was also reported in *Drosophila* and nematode backcross offspring (Sherengul et al. 2006; Ross et al. 2016). These studies suggest that either paternal mtDNA can be transmitted beyond the F1 stage, or that more opportunities to cross with the paternal lineage increase paternal leakage. Other researchers have speculated on how paternal mtDNA can enter the germline. Ladoukakis and Zouros (2017) suggest that an embryo could contain paternal mtDNA if the egg is genetically divergent from the sperm and fails to recognize and destruct sperm mitochondria during fertilization—this would explain the process of paternal leakage in hybrid individuals. Another insight can be gained by considering mussel species in the family Mytilidae, which, unlike other animal species, have one mitochondrial genome that is inherited from the mother for all offspring and one mitochondrial genome that is inherited from the father but is only found in male offspring (Zouros 2000). Recent studies found that the inherited paternal mtDNA have palindromic motifs in the control region that may increase transcription, have different open reading frames compared to maternal mtDNA, and have an extra cytochrome c oxidase subunit II gene (Passamonti et al. 2011; Zouros 2013). These genetic structures may help explain why transmission of paternal mtDNA along with maternal mtDNA is the norm in these mussels. In the *T. californicus* system, we do not know if there are such sequence differences in the leaked paternal mtDNA. In fact, the exact molecular mechanisms that cause paternal leakage in a typical animal system have not yet been directly tested, and the processes of paternal leakage are much less characterized compared to processes of uniparental maternal mtDNA inheritance.

It is largely unknown how extensive paternal leakage could affect hybrid fitness, but studies on how high amounts of heteroplasmy affect fitness can provide a clue. For example, induced heteroplasmic mice with initially equal proportion of two mitochondrial haplotypes had decreased cognition and physiological functions compared to homoplasmic mice (Sharpley et al. 2012). Harmful point mutations that cause MELAS syndrome in humans were selected against if they contributed more than 80% of the mitochondrial genome relative to the wild type mtDNA (Chomyn et al. 1992; Wallace 1999). In highly heteroplasmic *D. melanogaster* lines, mutant mitochondrial alleles were able to persist if there was another allele that complemented the wild type function (Ma et al. 2014). These studies suggest that high frequency of heteroplasmy can decrease or sometimes have no effect on fitness. It may be that paternal mtDNA in the hybrids behave like mutant mitochondrial haplotypes and thus affect fitness in similar ways. In our study, there is evidence of a potentially deleterious premature stop codon mutation in one leaked paternal haplotype, but it is unclear how this will impact fitness.

It remains unclear to what extent our results could influence interpretations of previous research in *T.s californicus*, but it is evident that the present study offers interesting directions for future research in this species. For instance, our results can be consistent with previous *T. californicus* literature if paternal leakage is frequent, but the leaked paternal mtDNA copy numbers within each hybrid individual are too low to have a substantial effect on fitness, or if initially

leaked paternal mtDNA do not get inherited to many subsequent generations. Preliminary research (Lee, unpublished) shows that leaked paternal mtDNA can persist at least up to the F2 generation, but we currently have no data on the copy numbers or concentration of the leaked paternal mtDNA within an individual *T. californicus*. Immediate future work in this species stemming from the present study could consist of quantifying paternal mtDNA copy numbers in hybrids of crosses found to have frequent paternal leakage, and investigating sequence differences of the uncommon versus common paternal haplotype in the F1 hybrids to gain an insight to the mechanism of paternal leakage. Beyond that, it would be intriguing to explore whether leaked paternal mtDNA could influence existing mitonuclear interactions in hybrids if the paternal mtDNA were to be artificially increased in concentration. While we stated earlier that paternal leakage is often reported in hybrids in previous literature and our results are consistent with that trend, it would be beneficial for future studies to also investigate if intrapopulation *T. californicus* crosses also exhibit paternal leakage to understand if this is a widespread phenomenon but has not been widely reported because it is difficult to detect.

Our results potentially challenge a variety of animal studies that utilize mitochondria as molecular markers (Hoolahan et al. 2011), and invite further studies on the prevalence and mechanisms of paternal leakage. As the process of paternal leakage is demystified, an interesting follow-up evolutionary question would be to test how extensive paternal leakage affects hybrid fitness. Our findings here will add to the growing evidence of paternal leakage in animals, and spark discussion on how paternal leakage could affect studies on molecular biology and evolution.

Supplementary Material

Supplementary data are available at *Journal of Heredity* online.

Funding

This work was supported by the National Science Foundation's Division of Integrative Organismal Systems (grant numbers NSF IOS-1155325 to C.S.W., NSF IOS-1555959 and IOS-2029156 to J. Kingsolver and C.S.W.); Kenan Trust Fund; and University of North Carolina at Chapel Hill's Graduate School Summer Research Fellowship.

Acknowledgments

We thank the "Illumina Workshop In A Box with Nextera Flex @ UNC with HTSF" workshop for Nextera sequencing supplies and services. We also thank Liana Kostak for help with rearing the hybrid crosses and collecting DNA, Jorge Santana, Sarah Slay, MaryAnn Bowyer for help with some of the research, members of Dr. Joel Kingsolver and Dr. Christopher Willett labs who provided feedback on the experimental design and manuscript, and Dr. Maria Servedio and Dr. Daniel Matute for input on the overall experimental design.

Data Availability

We will deposit the primary data for this manuscript as follows:

- Illumina Nextera sequences of ABf x SDm F1s will be uploaded to NCBI SRA.
- The Sanger sequences of long range PCR products spanning multiple mitochondrial genes, generated from pools of ABs and SCf x ABm F1s will be uploaded to NCBI GenBank. The Sanger sequences of long-range PCR products spanning multiple mtDNA genes from pooled samples and COB haplotypes from individual copepods will be uploaded to NCBI genbank.
- The AB genome assembly used for identifying NUMTs has been uploaded to Dryad (doi:10.5061/dryad.t76hdr828).

References

Al Rawi S, Louvet-Vallée S, Djeddi A, Sachse M, Culetto E, Hajjar C, Boyd L, Legouis R, Galy V. 2011. Postfertilization autophagy of sperm organelles prevents paternal mitochondrial DNA transmission. *Science*. 334:1144–1147.

Allen JF. 1996. Separate sexes and the mitochondrial theory of ageing. *J Theor Biol*. 180:135–140.

Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, *et al*. 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res*. 40:W597–W603.

Avise JC, Arnold J, Ball MR, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu Rev Ecol Evol Syst*. 18:489–522.

Balciuniene J, Balciunas D. 2019. A nuclear mtDNA Concatemer (Mega-NUMT) could mimic paternal inheritance of mitochondrial genome. *Front Genet*. 10:518.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, *et al*. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 19:455–477.

Barreto FS, Burton RS. 2013. Elevated oxidative damage is correlated with reduced fitness in interpopulation hybrids of a marine copepod. *Proc Biol Sci*. 280:20131521.

Barreto FS, Pereira RJ, Burton RS. 2015. Hybrid dysfunction and physiological compensation in gene expression. *Mol Biol Evol*. 32:613–622.

Barreto FS, Watson ET, Lima TG, Willett CS, Edmands S, Li W, Burton RS. 2018. Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. *Nat Ecol Evol*. 2:1250–1257.

Bensasson D, Zhang D, Hartl DL, Hewitt GM. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol Evol*. 16:314–321.

Blok RB, Gook DA, Thorburn DR, Dahl H-HM. 1997. Skewed segregation of the mtDNA nt 8993 (TrG) mutation in human oocytes. *Am J Hum Genet*. 60:1495–1501.

Brandstätter A, Niederstätter H, Parson W. 2004. Monitoring the inheritance of heteroplasmy by computer-assisted detection of mixed basecalls in the entire human mitochondrial DNA control region. *Int J Legal Med*. 118:47–54.

Burton RS. 1985. Mating system of the intertidal copepod *Tigriopus californicus*. *Marine Biology*. 86:247–252.

Burton RS. 1990. Hybrid breakdown in developmental time in the copepod *Tigriopus californicus*. *Evolution*. 44:1814–1822.

Burton RS. 1997. Genetic evidence for long term persistence of marine invertebrate populations in an ephemeral environment. *Evolution*. 51:993–998.

Burton RS, Barreto FS. 2012. A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities? *Mol Ecol*. 21:4942–4957.

Burton RS, Byrne RJ, Rawson PD. 2007. Three divergent mitochondrial genomes from California populations of the copepod *Tigriopus californicus*. *Gene*. 403:53–59.

Burton RS, Lee BN. 1994. Nuclear and mitochondrial gene genealogies and allozyme polymorphism across a major phylogeographic break in the copepod *Tigriopus californicus*. *Proc Natl Acad Sci U S A*. 91:5197–5201.

Burton RS, Pereira RJ, Barreto FS. 2013. Cytonuclear genomic interactions and hybrid breakdown. *Annu Rev Ecol Evol Syst*. 44:281–302.

Burton RS, Swisher SG. 1984. Population structure of the intertidal copepod *Tigriopus californicus* as revealed by field manipulation of allele frequencies. *Oecologia*. 65:108–111.

Chapman RW, Stephens JC, Lansman RA, Avise JC. 1982. Models of mitochondrial DNA transmission genetics and evolution in higher eucaryotes. *Genet Res*. 40:41–57.

Chomyn A, Martinuzzi A, Yoneda M, Daga A, Hurko O, Johns D, Lai ST, Nonaka I, Angelini C, Attardi G. 1992. MELAS mutation in mtDNA binding site for transcription termination factor causes defects in protein synthesis and in respiration but no change in levels of upstream and downstream mature transcripts. *Proc Natl Acad Sci U S A*. 89:4221–4225.

Coleman-Hulbert AL. 2010. *Mitochondrial inheritance and natural phenotypic variation among Caenorhabditis briggsae populations*. *Dissertations and Theses*. Paper 340. Portland, OR. doi:10.15760/etd.340

DeLuca SZ, O'Farrell PH. 2012. Barriers to male transmission of mitochondrial DNA in sperm development. *Dev Cell*. 22:660–668.

Dokianakis E, Ladoukakis ED. 2014. Different degree of paternal mtDNA leakage between male and female progeny in interspecific *Drosophila* crosses. *Ecol Evol*. 4:2633–2641.

Edmands S. 1999. Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. *Evolution*. 53:1757–1768.

Edmands S. 2001. Phylogeography of the intertidal copepod *Tigriopus californicus* reveals substantially reduced population differentiation at northern latitudes. *Mol Ecol*. 10:1743–1750.

Egloff DA. 1966. *Ecological aspects of sex ratio and reproduction in experimental and field populations of the marine copepod Tigriopus californicus* [*Thesis (Ph.D.*]. Stanford University.

Ellison CK, Burton RS. 2008. Interpopulation hybrid breakdown maps to the mitochondrial genome. *Evolution*. 62:631–638.

Fontaine KM, Cooley JR, Simon C. 2007. Evidence for paternal leakage in hybrid periodical cicadas (Hemiptera: Magicicada spp.). *PLoS One*. 2:e892.

Ganz HH, Burton RS. 1995. Genetic differentiation and reproductive incompatibility among Baja California populations of the copepod *Tigriopus californicus*. *Mar Biol*. 123:821–827.

Gaziev AI, Shaĭkhaev GO. 2010. Nuclear mitochondrial pseudogenes. *Mol Biol*. 44:405–417.

Gusman A, Lecomte S, Stewart DT, Passamonti M, Breton S. 2016. Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. *PeerJ*. 4:e2760.

Gyllensten U, Wharton D, Josefsson A, Wilson AC. 1991. Paternal inheritance of mitochondrial DNA in mice. *Nature*. 352:255–257.

Hastings IM. 1992. Population genetic aspects of deleterious cytoplasmic genomes and their effect on the evolution of sexual reproduction. *Genet Res*. 59:215–225.

Healy TM, Burton RS. 2020. Strong selective effects of mitochondrial DNA on the nuclear genome. *Proc Natl Acad Sci U S A*. 117:6616–6621.

Hoelzel AR, Green A. 1992. Analysis of population-level variation by sequencing PCR-amplified DNA. In: *Molecular Genetic Analysis of Populations: A Practical Approach*. p. 159–187.

Hoolahan AH, Blok VC, Gibson T, Dowton M. 2011. Paternal leakage of mitochondrial DNA in experimental crosses of populations of the potato cyst nematode *Globodera pallida*. *Genetica*. 139:1509–1519.

Hurst LD, Hamilton WD. 1992. Cytoplasmic fusion and the nature of sexes. *Philos Trans R Soc Lond B Biol Sci*. 247:189–194.

Kaneda H, Hayashi J, Takahama S, Taya C, Lindahl KF, Yonekawa H. 1995. Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis. *Proc Natl Acad Sci U S A*. 92:4542–4546.

Kim SJ, Lee KY, Ju SJ. 2013. Nuclear mitochondrial pseudogenes in *Austinograea alayseae* hydrothermal vent crabs (Crustacea: Bythograeidae): effects on DNA barcoding. *Mol Ecol Resour*. 13:781–787.

Kondo R, Satta Y, Matsuura ET, Ishiwa H, Takahata N, Chigusa SI. 1990. Incomplete maternal transmission of mitochondrial DNA in *Drosophila*. *Genetics*. 126:657–663.

Kowal K, Tkaczyk A, Pierzchała M, Bownik A, Ślaska B. 2020. Identification of mitochondrial dna (Numts) in the nuclear genome of daphnia magna. *Int J Mol Sci*. 21:8725.

Ladoukakis ED, Zouros E. 2017. Evolution and inheritance of animal mitochondrial DNA: rules and exceptions. *J Biol Res*. 24:2.

Lefébure T, Douady CJ, Gouy M, Gibert J. 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. *Mol Phylogenet Evol*. 40:435–447.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 25:1754–1760.

Lima TG, Burton RS, Willett CS. 2019. Genomic scans reveal multiple mitonuclear incompatibilities in population crosses of the copepod *Tigriopus californicus*. *Evolution*. 73:609–620.

Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien SJ. 1994. Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *J Mol Evol*. 39:174–190.

Luo S, Valencia CA, Zhang J, Lee NC, Slone J, Gui B, Wang X, Li Z, Dell S, Brown J, *et al*. 2018. Biparental inheritance of mitochondrial DNA in humans. *Proc Natl Acad Sci U S A*. 115:13039–13044.

Lutz-Bonengel S, Parson W. 2019. No further evidence for paternal leakage of mitochondrial DNA in humans yet. *Proc Natl Acad Sci U S A*. 116:1821–1822.

Ma H, Xu H, O'Farrell PH. 2014. Transmission of mitochondrial mutations and action of purifying selection in *Drosophila melanogaster*. *Nat Genet*. 46:393–397.

Mastrantonio V, Latrofa MS, Porretta D, Lia RP, Parisi A, Iatta R, Dantas-Torres F, Otranto D, Urbanelli S. 2019. Paternal leakage and mtDNA heteroplasmy in *Rhipicephalus spp*. ticks. *Sci Rep*. 9:1460.

Matsuura ET, Fukuda H, Chigusa SI. 1991. Mitochondrial DNA heteroplasmy maintained in natural populations of *Drosophila simulans* in Reunion. *Genetics Res*. 57:123–126.

Monk CR. 1941. Marine harpacticoid copepods from California. *Trans Am Microsc Soc*. 60:75–99.

Moritz C. 1994. Applications of mitochondrial DNA analysis in conservation: a critical review. *Mol Ecol*. 3:401–411.

Palmer CA, Edmands S. 2000. Mate choice in the face of both inbreeding and outbreeding depression in the intertidal copepod *Tigriopus californicus*. *Mar Biol*. 136:693–698.

Pamilo P, Viljakainen L, Vihavainen A. 2007. Exceptionally high density of NUMTs in the honeybee genome. *Mol Biol Evol*. 24:1340–1346.

Passamonti M, Ricci A, Milani L, Ghiselli F. 2011. Mitochondrial genomes and Doubly Uniparental Inheritance: new insights from *Musculista senhousia* sex-linked mitochondrial DNAs (*Bivalvia Mytilidae*). *BMC Genomics*. 12:442.

Pereira RJ, Barreto FS, Pierce NT, Carneiro M, Burton RS. 2016. Transcriptome-wide patterns of divergence during allopatric evolution. *Mol Ecol*. 25:1478–1493.

Quispe-Tintaya W, White RR, Popov VN, Vijg J, Maslov AY. 2013. Fast mitochondrial DNA isolation from mammalian cells for next-generation sequencing. *Biotechniques*. 55:133–136.

Radojičić JM, Krizmanić I, Kasapidis P, Zouros E. 2015. Extensive mitochondrial heteroplasmy in hybrid water frog (*Pelophylax spp*.) populations from Southeast Europe. *Ecol Evol*. 5:4529–4541.

Rawson PD, Burton RS. 2002. Functional coadaptation between cytochrome c and cytochrome c oxidase within allopatric populations of a marine copepod. *Proc Natl Acad Sci U S A*. 99:12955–12958.

Reilly JG, Thomas CA Jr. 1980. Length polymorphisms, restriction site variation, and maternal inheritance of mitochondrial DNA of *Drosophila melanogaster*. *Plasmid*. 3:109–115.

Richly E, Leister D. 2004. NUMTs in sequenced eukaryotic genomes. *Mol Biol Evol*. 21:1081–1084.

Rogers HH, Griffiths-Jones S. 2012. Mitochondrial pseudogenes in the nuclear genomes of *Drosophila*. *PLoS One*. 7:e32593.

Ross JA, Howe DK, Coleman-Hulbert A, Denver DR, Estes S. 2016. Paternal mitochondrial transmission in intra-species *Caenorhabditis briggsae* hybrids. *Mol Biol Evol*. 33:3158–3160.

Sato M, Sato K. 2011. Degradation of paternal mitochondria by fertilization-triggered autophagy in *C. elegans* embryos. *Science*. 334:1141–1144.

Sharpley MS, Marciniak C, Eckel-Mahan K, McManus M, Crimi M, Waymire K, Lin CS, Masubuchi S, Friend N, Koike M, *et al*. 2012.

Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. *Cell*. 151:333–343.

Sherengul W, Kondo R, Matsuura ET. 2006. Analysis of paternal transmission of mitochondrial DNA in *Drosophila*. *Genes Genet Syst*. 81:399–404.

Shitara H, Hayashi JI, Takahama S, Kaneda H, Yonekawa H. 1998. Maternal inheritance of mouse mtDNA in interspecific hybrids: segregation of the leaked paternal mtDNA followed by the prevention of subsequent paternal leakage. *Genetics*. 148:851–857.

Simpson SD, Ramsdell JS, Watson Iii WH, Chabot CC. 2017. The draft genome and transcriptome of the Atlantic horseshoe crab, *Limulus polyphemus*. *Int J Genomics*. 2017:7636513.

Song H, Buhay JE, Whiting MF, Crandall KA. 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proc Natl Acad Sci U S A*. 105:13486–13491.

Sun X, Yang A. 2016. Exceptionally large mitochondrial fragments to the nucleus in sequenced mollusk genomes. *Mitochondrial DNA A DNA Mapp Seq Anal*. 27:1409–1410.

Sutovsky P, Moreno RD, Ramalho-Santos J, Dominko T, Simerly C, Schatten G. 1999. Ubiquitin tag for sperm mitochondria. *Nature*. 402:371–372.

Sutovsky P, Moreno RD, Ramalho-Santos J, Dominko T, Simerly C, Schatten G. 2000. Ubiquitinated sperm mitochondria, selective proteolysis, and the regulation of mitochondrial inheritance in mammalian embryos. *Biol Reprod*. 63:582–590.

Takahata N, Maruyama T. 1981. A mathematical model of extranuclear genes and the genetic variability maintained in a finite population. *Genet Res*. 37:291–302.

Thompson WE, Ramalho-Santos J, Sutovsky P. 2003. Ubiquitination of prohibitin in mammalian sperm mitochondria: possible roles in the regulation of mitochondrial inheritance and sperm quality control. *Biol Reprod*. 69:254–260.

Triant DA, DeWoody JA. 2007. The occurrence, detection, and avoidance of mitochondrial DNA translocations in mammalian systematics and phylogeography. *J Mammal*. 88:908–920.

Wallace DC. 1999. Mitochondrial diseases in man and mouse. *Science*. 283:1482–1488.

White SL, Shanske S, McGill JJ, Mountain H, Geraghty MT, DiMauro S, Dahl HH, Thorburn DR. 1999. Mitochondrial DNA mutations at nucleotide 8993 show a lack of tissue- or age-related variation. *J Inherit Metab Dis*. 22:899–914.

White DJ, Wolff JN, Pierson M, Gemmell NJ. 2008. Revealing the hidden complexities of mtDNA inheritance. *Mol Ecol*. 17:4925–4942.

Willett CS, Ladner JT. 2009. Investigations of fine-scale phylogeography in *Tigriopus californicus* reveal historical patterns of population divergence. *BMC Evol Biol*. 9:139.

Wilson AC, Cann RL, Carr SM, George M, Gyllensten UB, Helm-Bychowski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, *et al*. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol J Linn Soc*. 26:375–400.

Wolff JN, Ladoukakis ED, Enríquez JA, Dowling DK. 2014. Mitonuclear interactions: evolutionary consequences over multiple biological scales. *Philos Trans R Soc Lond B Biol Sci*. 369:20130443.

Xu J. 2005. The inheritance of organelle genes and genomes: patterns and mechanisms. *Genome*. 48:951–958.

Yoneda M, Chomyn A, Martinuzzi A, Hurko O, Attardi G. 1992. Marked replicative advantage of human mtDNA carrying a point mutation that causes the MELAS encephalomyopathy. *Proc Natl Acad Sci U S A*. 89:11164–11168.

Zhang DX, Hewitt GM. 1996. Nuclear integrations: challenges for mitochondrial DNA markers. *Trends Ecol Evol*. 11:247–251.

Zouros E, Oberhauser Ball A, Saavedra C, Freeman KR. 1994. An unusual type of mitochondrial DNA inheritance in the blue mussel *Mytilus*. *Proc Natl Acad Sci U S A*. 91:7463–7467.

Zouros E. 2000. The exceptional mitochondrial DNA system of the mussel family *Mytilidae*. *Genes Genet Syst*. 75:313–318.

Zouros E. 2013. Biparental inheritance through uniparental transmission: the Doubly Uniparental Inheritance (DUI) of mitochondrial DNA. *Evol Biol*. 40:1–31.