Title

Genomic scans reveal multiple mito-nuclear incompatibilities in population crosses of the copepod Tigriopus californicus

Journal

Evolution, 73(3)

ISSN

0014-3820

Authors

Lima, Thiago G
Burton, Ronald S
Willett, Christopher S

Publication Date

2019-03-01

DOI

10.1111/evo.13690

Peer reviewed

GENOMIC SCANS REVEAL MULTIPLE MITO-NUCLEAR INCOMPATIBILITIES IN POPULATION CROSSES OF THE COPEPOD *TIGRIOPUS CALIFORNICUS*

Thiago G. Lima[a,b,1], Christopher S. Willett[a] and Ronald S. Burton[b]

[a]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, 27599.

[b]Marine Biology Research Division, Scripps Institution of Oceanography, La Jolla, California, 92037.

[1]To whom correspondence should be addressed: Thiago G. Lima, MBRD, Scripps Institution of Oceanography, UCSD, 9500 Gilman Drive, La Jolla, CA, 92093; phone: (858) 534-7827; email: thiago.ghirello.lima@gmail.com

Running head: Nuclear-nuclear versus mito-nuclear incompatibilities

Keywords: Postzygotic reproductive isolation; hybrid inviability; *Tigriopus*; Pool-seq

Data archival location: Raw Illumina reads were deposited in NCBI SRA (XXXX); allele frequency data were deposited at Dryad (XXXXX)

Abstract

The evolution of intrinsic postzygotic isolation can be explained by the accumulation of Dobzhansky-Muller incompatibilities (DMI). Asymmetries in the levels of hybrid inviability and hybrid sterility are commonly observed between reciprocal crosses, a pattern that can result from the involvement of uniparentally inherited factors. The mitochondrial genome is one such factor that appears to participate in DMI in some crosses but the frequency of its involvement versus biparentally inherited factors is unclear. Here we assess the relative importance of incompatibilities between nuclear factors (nuclear-nuclear) versus those between mitochondrial and nuclear factors (mito-nuclear) in a species that lacks sex chromosomes. We used a Pool-seq approach to survey three crosses among genetically divergent populations of the copepod, *Tigriopus californicus*, for regions of the genome that are affected by hybrid inviability. Results from reciprocal crosses suggest that mito-nuclear incompatibilities are more common than nuclear-nuclear incompatibilities overall. These results suggest that in the presence of very high levels of nucleotide divergence between mtDNA haplotypes, mito-nuclear incompatibilities can be important for the evolution of intrinsic postzygotic isolation. This is particularly interesting considering this species lacks sex chromosomes, which have been shown to harbor a particularly high number of nuclear-nuclear DMI in several other species.

Keywords: Postzygotic reproductive isolation; hybrid inviability; *Tigriopus*; Pool-seq

Introduction

The formation of reproductive isolation through the evolution of hybrid incompatibilities (intrinsic postzygotic isolation), can often be attributed to the evolution of Dobzhansky-Muller incompatibilities (DMI [(Dobzhansky 1936; Muller 1942)]). One pattern that is observed in these crosses is that in reciprocal crosses asymmetries in hybrid inviability and hybrid sterility are commonly found. This pattern of asymmetry is called Darwin's corollary and is likely to result from DMI that have uniparentally inherited genetic elements comprising at least one partner in the interaction (Turelli and Moyle 2007). These uniparentally inherited factors can include things such as mitochondrial DNA (mtDNA), chloroplast DNA (cpDNA), sex chromosomes, and maternal transcripts. While sex chromosome often show up in many crosses as making key contributions to DMI (Tao et al. 2003; Masly and Presgraves 2007), in other crosses including those without sex chromosomes cytoplasmic factors such as mtDNA and cpDNA are increasingly showing up as important contributors (Burton et al. 2013). In animal taxa for which evidence for mito-nuclear interactions has been found it is not generally clear what are the relative contributions of mito-nuclear interactions versus nuclear/nuclear interactions to DMI leading to postzygotic reproductive isolation.

The accumulation of mito-nuclear incompatibilities is facilitated by a number of features shared by mtDNA in a variety of animal taxa, that might lead to it having outsized impacts despite its small size in comparison with the nuclear genome. One factor is that mtDNA generally has a higher rate of sequence evolution than the nuclear genome in most animal taxa (Willett 2012). In most cases mtDNA is maternally inherited and as such its haploid nature can expose DMI that would otherwise be masked in a diploid setting, (analogous to heteromorphic sex chromosomes). Additionally rapid evolution due to genomic conflicts can be particularly

72 pronounced given the different patterns of inheritance between nuclear and mtDNA genomes
73 (Gershoni et al. 2009; Chou and Leu 2015). The evolution of genomic conflicts could be
74 accelerated in many taxa by the higher mutation rate of mtDNA compared to nuclear DNA.
75 Combined, the faster rate of sequence evolution mito-interacting genes and the existence of
76 genomic conflicts in these genes could drive the evolution of compensatory changes in nDNA
77 within populations and could lead to DMI between populations or species where gene flow is
78 absent or low (Burton and Barreto 2012; Burton et al. 2013).

79 Strong support for the importance of mito-nuclear DMI has been found in the copepod
80 *Tigriopus californicus,* a species that lacks sex chromosomes. This copepod, which lives in high
81 intertidal pools on the west coast of North America, has polygenic sex determination with
82 several unlinked factors contributing to sex determination (Voordouw and Anholt 2002;
83 Alexander et al. 2014; 2015). *Tigriopus californicus* populations occupy rocky pools on
84 headlands that are often isolated from other headlands by long stretches of sandy beach. Gene
85 flow is highly restricted amongst populations (Burton 1997; Willett and Ladner 2009), and levels
86 of polymorphism within populations are very low (Willett 2012; Pereira et al. 2016). Reciprocal
87 crosses between divergent clades within this species show differences in patterns of reproductive
88 isolation depending on the direction of the cross (Ganz and Burton 1995; Peterson et al. 2013),
89 suggesting that mito-nuclear incompatibilities may be important for the formation of these
90 reproductive barriers. When populations with lower levels of divergence are crossed, first
91 generation hybrids (F_1) are usually equal in fitness, or even superior, to the parental populations.
92 while second generation hybrids (F_2) have, on average, lower fitness (Burton 1987; Edmands
93 1999; Willett 2008). When F_2 and F_3 hybrids are backcrossed to the maternal population, where

there is an increase in the proportion of the nDNA that matches mtDNA, hybrid fitness is rescued (Ellison and Burton 2008b).

Here, we were interested in determining the relative importance of nuclear-nuclear versus mito-nuclear DMI for hybrid breakdown in early stages of reproductive isolation between populations of *T. californicus*. We used a Pool-seq approach (Schlötterer et al. 2014) to sequence the genomes of pools of F_2 hybrids from three different pairs of reciprocal crosses, looking for deviations from expected allelic frequencies to determine regions of the genome that were affected by hybrid inviability. We show that mito-nuclear DMI are in general more common than nuclear-nuclear DMI, but that the relative contribution of different types of incompatibilities are unique in the different crosses.

Material and Methods

Population sampling, crossing design, DNA isolation and sequencing

Tigriopus californicus were collected from intertidal rocky pools at four sites in California, Abalone Cove (AB, 33°44' N, 118°22' W), Catalina Island (CAT, 33°27' N, 118°29' W), San Diego (SD, 32°44' N, 117°15' W), and Santa Cruz (SC, 36°57' N, 122°03' W). Animals were maintained in mass cultures in 400 mL beakers in seawater at 35 ppt and fed powdered commercial flake fish food as well as natural algae growth. Cultures were kept in incubators at 20°C with 12h light:dark cycle. Males and females used in crosses were sampled from culture beakers so that different crosses between the same populations included some of the genetic diversity of natural populations. Reciprocal crosses were setup between the AB populations and the SD, CAT and SC populations in 24 well culture plates, with a single pair of copepods in each well. F_2 hybrid breakdown in viability has been shown for the SD x AB and SC

x AB crosses (Burton 1987; Ellison and Burton 2008b), but no studies have been published using the CAT the population (although population crosses with similar levels genetic divergence typically show hybrid breakdown in this species (Edmands 1999)). Virgin females were obtained by separating females from clasped pairs (Burton 1985), and their non-mated status was confirmed by monitoring them in individual wells over a week, at which point males were added to each well. Twenty-four crosses between the parental populations were setup. F_1 hybrids from these crosses were separated into individual wells before they reached sexual maturity, to prevent siblings from mating with each other. F_1 x F_1 crosses were setup with a single pair per well again, and outcrossing was insured by crossing siblings from one cross to copepods from as many different crosses as possible, maximizing the number of combinations between the original 24 parental x parental crosses (within the same two population crosses). In both parental and F_1 x F_1 crosses, male fathers were removed from the cross as soon as nauplii were observed, while females were kept in the wells as they can produce multiple egg clutches from the single mating (Fig. 1a).

Crosses with the SD, CAT and SC female parents were setup and sequenced between 2012-2014. For each cross, 300 adult F_2 hybrids (150 males and females) were collected and pooled for DNA extraction. Crosses with the AB females were setup and sequenced from 2015-2016. For each of these crosses, two replicates of 100 males and 100 females were collected and pooled for DNA extraction. For the SDf x ABm and SCf x ABm crosses, DNA was isolated using the Qiagen DNeasy blood and tissue kit, with the suggested modification for extraction from insects (Qiagen). For all other crosses, DNA was isolated using a Phenol:Chloroform procedure (Sambrook and Russell 2006). Samples were sequenced as 100-bp paired-end (PE) libraries on the Illumina HiSeq 2000 for the SDf x ABm and SCf x ABm crosses, as 125-bp PE

libraries on the Illumina HiSeq 2500 for the CATf x ABm cross and as 100-bp PE libraries on the Illumina HiSeq 4000 for the ABf x SDm, ABf x CATm and ABf x SCm crosses. Results for the SDf x ABm cross have been published in Lima and Willett (2018). The difference in sequencing platforms used here should not affect or bias SNP determination or allele frequency estimation, as errors rates for these sequencing platforms are not significantly different.

Generation of consensus references for populations

Lima and Willett (2018) generated an AB reference genome sequence using the mapping file (BAM) available at (https://i5k.nal.usda.gov/Tigriopus_californicus; v1.0), where AB reads were mapped to the SD reference. We followed the same procedure to create consensus reference genomes for CAT and SC, by extracting the consensus sequences from the BAM files (CAT and SC reads mapped to the SD reference) using the Samtools and Bcftools pipeline (Li et al. 2009; Li 2011). We then compared the references between each pair of populations used in the three crosses, and made them equivalent by adding "N"s to any position where either the AB or the alternative populations also had an "N". This maintains the length of the references, but makes them comparable in terms of where reads can map, which is particularly important when the SD reference is considered (this population was *de novo* assembled, using diverse sequence stes, into a high-quality assembly [Barreto et al. 2018]). The purpose of creating these references is to allow the mapping of reads from both parental populations as well as hybrids, in order to identify SNPs that are fixed between populations in each cross. We take a very conservative approach that only considers regions of the genome where reads map with high alignment score, ignoring regions where divergence is too high and confidence in SNP calling may be low (see below for further details).

163

164 **Anchoring of scaffolds to chromosomes**

165 As noted in Lima and Willett (2018), a newer reference genome for the SD population
166 has recently become available (v2) (Barreto et al. 2018), where greater than 90% of the genome
167 is anchored to chromosomes. We used this reference to anchor and order the scaffolds from the
168 reference assembly used in the present study (v1) by BLASTing scaffolds from the v1 assembly
169 to the v2 assembly, and using these alignments to anchor and order the v1 scaffolds into the 12 *T.*
170 *californicus* chromosomes. This increased the percentage of the genome that is anchored to
171 chromosomes from ~30% to approximately 97% of the v1 reference length. In the process of
172 anchoring the v1 scaffolds it was determined that a few of these scaffolds were misassembled,
173 which in some cases can be observed as sharp changes in allele frequency in the hybrid datasets.
174 We removed these positions from the allele frequency plots. These misassembled regions can be
175 observed as small sections of the chromosome where allele frequency clearly deviates from the
176 trend in allele frequency change across the rest of the chromosome (Fig. S1a-3a).

177

178 **SNP database between parental populations**

179 Populations of *T. californicus* have been shown to be genetically stable and highly
180 segregated, with nearly no gene-flow between populations that are geographically very close
181 (Burton 1997; Willett and Ladner 2009; Pereira et al. 2016). Shared polymorphism in *T.*
182 *californicus* decreases exponentially with divergence, and even populations with ~ 1/3 the level
183 of divergence of the crosses presented here, share 0.6% or less of variable sites (Pereira et al.
184 2016). Within population polymorphism is also extremely low (Willett 2012; Pereira et al. 2016).
185 For the purpose of this study, we were interested in establishing a list of SNPs to be used as

186 markers across the genome (and not as a thorough survey of population differences). We aimed
187 to find SNPs that are fixed between the populations used in each cross. This means that
188 reciprocally fixed differences between populations are likely to reflect long term population
189 differences. We accomplished this by a) performing reciprocal mapping of reads of a population
190 to the reference sequence of another, b) considering only those position where all mapped reads
191 showed an alternative nucleotide to the reference ("fixed differences"), and c) comparing the
192 reciprocal mappings and keeping only SNP that were "fixed differences" in both mappings.
193 Reads were mapped reciprocally to the parental populations' reference of each cross,
194 using BWA with default parameters. Only reads that mapped with a MAPQ score > 20 were
195 kept, which excludes reads that map with low alignment score. We used PoPoolation2 (Kofler et
196 al. 2011b) to find positions across the genome where all reads had an alternative nucleotide to
197 that of the reference, considering only biallelic positions with coverage ≥ 10 (CAT mapped to
198 AB, 20X average coverage), ≥ 15X (SD mapped to AB, 35X average coverage) and ≥ 20X for
199 all other comparisons (SC mapped to AB = 50X; and 46X for AB mapped to the other three
200 populations). We chose different lower coverage cutoffs because the depth of coverage differed
201 for each population. Since we are only considering SNP positions with 100% of the reads having
202 an alternative nucleotide to that of the reference sequence, the difference in coverage between
203 populations should not have a significant effect on SNP selection. Populations with lower
204 coverages (CAT and SD), may include SNP positions that are not quite fixed between these
205 populations and AB. Since all three populations are compared to the reciprocal mapping of AB
206 reads, with a minimum coverage of 20X, many of these positions would be excluded.
207 Furthermore, SNP allele frequencies were averaged in sliding windows (see below), and

chromosome wide patterns were considered, making the effects of a posible small number of non-fixed SNP insignificant.

Hybrid read mapping, SNP identification, and allele frequency calculation

Illumina reads were trimmed for quality using PoPoolation (Kofler et al. 2011a) discarding bases with Phred quality scores lower than 25, and keeping reads of at least 50-bp after trimming. We followed the pipeline from Lima and Willett (2018) from mapping through allele frequency calculation and smoothing. This involved mapping reads from each cross to both of their parental genomes using BWA MEM with default parameters (Li and Durbin 2009), and keeping reads that mapped with MAPQ score > 20. Read counts for every variable position, as well as population specific allele counts were determined using PoPoolation 2. Only biallelic positions were considered where the minor allele had a minimum coverage of at least four. Allele frequencies were calculated as the AB allele frequency for each of the three pairs of crosses. Due to the large amount of noise in the allele frequency data (likely due to stochastic differences in coverage between SNPs, as well as the sampling of alleles from a pool), we averaged the allele frequency for a sliding window of 3000 consecutive SNPs, moving the window by 3000 SNPs each step (non-overlapping windows; Table 1). This sliding window size was chosen following Lima and Willett (2018), as it minimizes noise in allele frequency estimation, compared to smaller windows, without losing any signal. If the sliding window is done by position, averaging window sizes of 250Kbp, the pattern remains the same (Fig. S1b-S3b).

Identification of hybrid inviability patterns caused by nuclear-nuclear vs mito-nuclear incompatibilities

231 Statistical tests to determine if the allele frequency differs between the reciprocal crosses
232 in F_2 hybrids are conservative when using read counts and coverage, unless coverage is much
233 greater than 100X. That is because even if one of the homozygous genotypes is completely
234 lethal, the allele frequency would only change by 0.167 in either direction. Instead we
235 determined if the overall allele frequency distribution between reciprocal crosses was the same
236 using the Kolmogorov-Smirnov test (KS test). Since allele frequencies from SNPs in close
237 proximity would not be independent from each other, we compared the distributions of allele
238 frequencies after averaging allele frequencies in sliding windows of 2 Mbp. This decreased the
239 number of data points per chromosome to 6-8 allele frequency windows. In comparing allele
240 frequency distributions for the entire genome, or for each chromosome (see below) we expect
241 that if the pattern observed was caused by incompatibilities between nuclear factors (nuclear-
242 nuclear), the allele frequencies are expected to be skewed in the same direction in each reciprocal
243 cross. On the other hand, incompatibilities caused by problems between the nuclear genome and
244 mitochondrial genome are expected to lead to allele frequencies that are skewed towards the
245 population that matches the mitochondria. This may occur in only one direction of the cross or in
246 both directions, resulting in allele frequencies that are skewed in opposite directions in the
247 reciprocal crosses. A third pattern is also possible, where nuclear alleles from one population are
248 favored while having mitochondria from another population; we term this a "mismatch" pattern.
249 Again, this may happen in one direction of the cross, with the other direction showing no skew,
250 or in both directions with opposite allele frequencies (Fig. 1b-c).
251 To differentiate among these three potential patterns, we first looked for mito-nuclear
252 incompatibilities or mismatch patterns by determining the 10% and 90% allele frequency
253 quantiles for each chromosome for each cross, looking for chromosomes where the 10% quantile

of a cross did not overlap with the 90% quantile of its reciprocal cross. The pattern of skew in relation to the mtDNA-type in the cross would then determine if the deviation was consistent with a mismatch or mito-nuclear incompatibility pattern for each chromosome. To detect nuclear-nuclear incompatibilities, where the allele frequencies between reciprocal crosses are expected to show skews in the same direction, we used the method described in Lima and Willett (2018) which defined cutoffs for deviations from the expected allele frequency of 0.5 based on deviations seen in a naupliar dataset (Fig. 1c). The nauplii provide an estimate of experimental error in the estimation of these frequencies because little to no genotype specific selection occurs before nauplii hatch and no evidence of meiotic drive from F_1 parents has been describe in this species. Studies that compared nauplii to adults showed that nearly no deviations in expected Mendelian ratios of segregation occur prior to, or just after hatching, with nearly all effects of hybrid inviability taking place as nauplii develop post hatching (Willett and Ladner 2007; Pritchard et al. 2011; Foley et al. 2013; Willett et al. 2016; Lima and Willett 2018). We therefore used this SDf x ABm F2 naupliar dataset to calculate the 10% and 90% allele frequency quantiles for all chromosomes combined and found that these fell at ± 0.018 relative to the mean allele frequency (Fig. S4). We take a slightly more conservative cutoff and look for chromosomes where the 10% and 90% allele frequency quantiles were ± 0.02 away from 0.5 (see Table S1 for full results).

Divergence between populations

In order to determine the amount of genome-wide divergence between each of the pairs of populations used for crosses, we calculated the number of synonymous changes per synonymous sites (d*S*) for all genes annotated in the *T. californicus* genome (13,449 genes).

Alignments were done for each gene separately using PRANK (Löytynoja 2013), and positions where the quality of the alignment was low were removed with Gblocks (Castresana 2000), keeping codons intact. Alignments with length < 100bp were removed. Estimation of d*S* was done in PAML 4.8 (Yang 2007) in the program YN00, in pairwise comparisons between AB and the other populations. Only two genes showed dS > 1 across any of the combinations; removing or including these genes in the analysis did not affect the final dS average (individual values for each are in Table S2). While d*S* was calculated for pairwise comparisons, the alignment for all four populations was used, removing any position where any population had a gap or "N" in the sequence. This was done to minimize the effect that differences in assembly quality can cause when estimating d*S*. Values were averaged across all genes for each pairwise comparison (± standard error; Table 1). We also performed a sliding window averaging of *dS* using window sizes of 250kbp, as a way to determine if any chromosomal regions have particularly high levels of divergence (Fig. S1b-S3b).

Results

Descriptive analysis

The average depth of coverage ranged from 71.06 (SC x AB) to 229.20 (CAT x AB), which yielded between 2.1-2.4 million SNPs and 698-805 windows of 3000 SNP per cross, with mean window sizes ranging from 222,464-256,820 bp (Table 1). Divergences between AB and the other populations were calculated as the average genome-wide d*S*, a measure that approximates the level of genomic divergence between the populations. Values for d*S* show that divergences between AB-SD and AB-CAT are very similar, while AB-SC is slightly more divergent (Table 1).

300

301 **Patterns of hybrid inviability across different population crosses**

302 We were interested in determining the relative importance of nuclear-nuclear versus
303 mito-nuclear incompatibilities by comparing allele frequency changes caused by hybrid
304 inviability between reciprocal crosses. These two types of incompatibilities can be distinguished
305 by comparing reciprocal crosses, since the average composition of the nuclear genome of these
306 crosses should be the same, but the mtDNA would differ. As discussed previously, allele
307 frequency changes that are caused by nuclear-nuclear incompatibilities should affect both
308 reciprocals of the cross equally, while mito-nuclear incompatibilities would affect each direction
309 of the cross differently (though these incompatibilities need not be symmetric) (Fig. 1c). To test
310 this, we compared genome-wide allele frequency distributions between the reciprocal crosses
311 using a Kolmogorov-Smirnov test (KS test). The distributions of allele frequencies were
312 significantly different between the reciprocal crosses for all three crosses (KS test: SDxAB: D =
313 0.226, $N = 84$, $P = 0.027$; CATxAB: D = 0.398, $N = 83$ $P < 3.017^{-6}$; SCxAB: D = 0.298, $N = 84$
314 $P = 0.001$), indicating that hybrid inviability due to mito-nuclear interactions may have a
315 significant effect on the genome-wide allele frequency patterns (Fig. 2).

316

317 **Mito-nuclear incompatibilities**

318 When referring to each direction of a cross, we will use a two-letter acronym as follows:
319 DA (SDf x ABm, AD (ABf x SDm), CA (CATf x ABm), AC (ABf x CATm), SA (SCf x ABm)
320 and AS (ABm x SCf). Patterns consistent with mito-nuclear incompatibilities affect DA for
321 chromosomes 8 and 10, and AD for chromosomes 4 and 7 (Fig. 3a). CA is affected by mito-
322 nuclear incompatibilities for chromosomes 1, 2, 9 and 10, and AC for chromosome 1 and 3 (Fig.

3b). Chromosome 1 shows signs of a reciprocal mito-nuclear incompatibility between CA and AC. SA and AS do not show evidence of strong mito-nuclear incompatibilities, but for four chromosomes (c6, c8, c9 and c10) the allele frequencies between the reciprocal crosses are divergent and may indicate the presence of weaker mito-nuclear incompatibilities (Fig. 3c).

Nuclear-nuclear incompatibilities

When we consider nuclear-nuclear incompatibilities, DA and AD show strong allele frequency skews in chromosomes 1 and 3 both with an excess of AB alleles (Fig. 3a; Table S1 shows full results for all crosses). CA, AC show signs of nuclear-nuclear incompatibility in chromosome 6 (Fig. 3b), and SA and AS have skewed allele frequencies consistent with nuclear-nuclear incompatibilities for chromosome 5 (Fig. 3c). In both cases, the magnitude of the skew is different between the reciprocal crosses, indicating either variation in the expression of the incompatibility, or a more complicated combination of the nuclear-nuclear and mismatch patterns (Fig. 3).

Mismatch pattern

Surprisingly, a mismatch pattern is observed for at least one chromosome in all three crosses. A portion of chromosome 11 is skewed in DA with excess AB alleles, but no mismatch pattern is observed in AD (Fig. 3a). AC has the most extreme form of a mismatch pattern with chromosomes 6 and 7 showing strongly skewed allelic frequencies, which account for the majority of the skewed allelic frequency observed in this cross (16.3% of the SNP windows) (Fig. 3b). The reciprocal cross does not show a mismatch pattern. As mentioned above, allele frequencies are skewed for chromosome 6 for both directions of the cross, but are not nearly as

skewed as in AC. This may suggest a more complicated combination of the nuclear-nuclear and mismatch patterns. SA and AS show evidence of a reciprocal mismatch pattern for chromosome 7, while AS also has a mismatch pattern for chromosome 11 (Fig. 3c).

Mito-nuclear interactions in relation to mito-interacting nuclear genes

We plotted the chromosomal position of all annotated nuclear genes that code for proteins that are known to directly interact with mtDNA, mitochondrial encoded proteins or mitochondrial encoded RNAs (Table S3). These included proteins involved in oxidative phosphorylation (OXPHOS complexes 1, 3, 4 and 5), mitochondrial aminoacyl tRNA synthethases (mArs), as well as the three transcription associated genes: mitochondrial RNA polymerase (mtRPOL), mitochondrial transcription factor A (TFAM), and mitochondrial transcription factor 1 (TFB1) (Fig. 3). These are spread out across all 12 chromosomes, with at least 6 genes per chromosome, with the exception of chromosome 12 which only has one gene mapping to it (TFAM). Chromosome 10 has the most genes (29), both from OXPHOS complexes and mArs, and this is the only chromosome where all three crosses show evidence of mito-nuclear incompatibilities in at least one direction (DA, CA and a reciprocal pattern between SA and AS). Chromosome 7, which shows evidence of mito-nuclear incompatibilities in AD, and a mismatch pattern in AC, SA and AS, has 10 mito-interacting genes, two of these are the tightly linked mtRPOL and TFB1 genes. A cluster of mito-interacting genes on the left portion of chromosome 2 coincides with a mito-nuclear incompatibility pattern in CA, while a cluster on the right portion of chromosome 4 appear to coincide with a mismatch pattern in AC (Fig. 3; Table S3).

Discussion

The evolution of intrinsic postzygotic isolation, can occur through the accumulation of DMI between factors on different chromosomes (nuclear-nuclear), but may also occur between nuclear factors and the mitochondrial factors with which they interact (mito-nuclear [Burton and Barreto 2012]). In the present study, we show that in hybrids between populations of the copepod *Tigriopus californicus,* patterns of allele frequency deviation from Mendelian expectations, suggest mito-nuclear incompatibilities are more common than nuclear-nuclear incompatibilities (Table 1). There is, therefore, strong evidence that hybrid problems between interacting mitochondria and nuclear genes are particularly important for the evolution of intrinsic postzygotic isolation in this taxon. However, a direct causality for the role of mtDNA in this process remains to be demonstrated.

Given the design of our experiment, it is possible other factors besides inviability due to mito-nuclear DMI may have contributed to pattern of allele frequency observed. For example, meiotic drive in the F_1 hybrids could bias alleles that are observed in F_2 hybrids. However, this is unlikely to be the case, as all studies that have looked at genotypic or allele frequency ratios in nauplii found little to no deviations from expected Mendelian ratios (Willett and Berkowitz 2007; Pritchard et al. 2011; Foley et al. 2013; Lima and Willett 2018). This indicates that the deviations in allele frequency from the expected Mendelian ratio of 0.5, occur post hatching, as the nauplii develop (Willett et al. 2016). There is differential survival between male and female F_2 hybrids, with different loci showing deviations in each sex (Foley et al. 2013; Willett et al. 2016). However, in the present study we were interested in estimating the allele frequency patterns from equal number of males and females combined. Lastly, reciprocal crosses were setup at different times, which could confound allele frequency differences due to mito-nuclear

392 DMI, with temporal variation in allele frequencies within the parental populations. This is also
393 unlikely to have a significant effect on the patterns observed here. Populations of *T. californicus*
394 have been shown to have stable genetic divergence over several decades in some cases, and the
395 level of within population polymorphism is very low (Burton 1997; Willett and Ladner 2009;
396 Pereira et al. 2016). In addition, studies that have looked at the same cross through different
397 years, tend to find the same genomic regions being affected by hybrid inviability (Burton 1987;
398 Willett et al. 2016; Lima and Willett 2018). Small differences in allele frequency between
399 reciprocal crosses, however, may be due to temporal variation in the expression of specific DMI.
400 One of the interesting aspects of our results is the lack of commonality in allelic
401 frequency change across the crosses of different populations of *T. californicus*. Even though all
402 crosses involve the AB population, only chromosome 10 is affected in a similar manner across
403 all three crosses. Allelic frequencies along chromosome 10 suggest the presence of mito-nuclear
404 incompatibilities in both DA and CA, and a trend towards reciprocal mito-nuclear
405 incompatibility between SA and AS. All other genomic regions affected by hybrid inviability are
406 unique to each cross suggesting the nature of DMI as unique products of divergence between
407 allopatric populations, and possibly little parallelism among factors involved in DMI. *Tigriopus*
408 *californicus* populations are extremely isolated, with little to no gene flow even for populations
409 that are geographically very close (Burton 1997; Willett and Ladner 2009), and the level of
410 shared polymorphism decreases exponentially as divergence increases between populations
411 (Pereira et al. 2016). Therefore, even if selection pressures are similar across populations, the
412 lack of gene flow can lead to different solutions to local adaptation, as different mutations are
413 likely to appear in the different populations (Lima and Willett 2017). Furthermore, effective

414 populations sizes are very low in this species and genetic drift can lead to the fixation of different
415 alleles, regardless of selection (see below for further discussion on this).
416 Interestingly chromosome 10 has the highest number and density of mito-interacting
417 nuclear genes (with 17; all other chromosomes have between 1 [c12] and 10 [c8]) (Table S3). In
418 *T. californicus* recombination only occurs in males and large chromosomal blocks will be passed
419 on together to F_2 hybrids. Therefore, chromosomes, or regions of chromosomes, with high
420 density of mito-interacting nuclear genes may be more likely to lead to hybrid problems even if
421 each factor alone has only a small effect. It has been proposed that genomic architecture should
422 evolve to suppress recombination between nuclear encoded mitochondrial genes, and islands of
423 divergence associated with such genes have been observed in passerine birds (Sunnucks et al.
424 2017). Chromosomes 2 and 4 may support this idea, as the only skewed portion of the
425 chromosomes in CA and AC, respectively, coincide with the portion of the chromosome with
426 high density of mito-interacting nuclear genes. However, high density of mito-interacting genes
427 may not be necessary for the expression of mito-nuclear incompatibilities. Chromosome 7
428 suggests a mito-nuclear effect in AD, a mismatch pattern in AC, and reciprocal a mismatch
429 pattern in AS and SA. Two tightly linked genes responsible for transcription of the mitochondrial
430 genome are present in this chromosome (RNA polymerase [mtRPOL] and transcription factor B
431 [TFB1M]). Previous studies have shown that mismatched alleles between mtRPOL alleles and
432 the mitochondria population is associated with reduced hybrid fitness in some crosses (Ellison
433 and Burton 2006; 2008a; 2010), while in other crosses a mismatch pattern has been observed for
434 other markers (Ellison and Burton 2008a; Edmands et al. 2009).
435 In all three crosses, the distribution of allelic frequencies is significantly different
436 between the reciprocal crosses, suggesting the mitochondrial background has a significant effect

437 on the nuclear allelic frequency distribution. Surprisingly, however, this signal comes not only
438 from the effects of mito-nuclear incompatibilities, but from chromosomes affected by a
439 mismatch pattern. Pereira et al. (2016) found that effective population sizes are very small in this
440 species, with within population genetic diversity raging from 0.006-0.017 π_S (genetic diversity at
441 synonymous sites). Populations with small effective populations sizes should be strongly
442 affected by genetic drift, potentially leading to the fixation of slightly deleterious mutations
443 (Kimura 1968). Therefore, it is possible that populations of *T. californicus* have fixed slightly
444 deleterious mito-nuclear combinations and when given an alternative nuclear allele in hybrids,
445 this mismatch combination will increase fitness (a similar hypothesis has been proposed by
446 Edmands et al. 2009). These slightly deleterious mito-nuclear combinations (within each
447 population) might not cause inviability when in a completely parental population genome, but
448 when in a hybrid genome where other deleterious hybrid combinations are also affecting fitness,
449 their effect might be amplified.
450
451 **The importance of mito-nuclear incompatibilities for speciation**
452 The present results show mito-nuclear incompatibilities are more common than nuclear-
453 nuclear incompatibilities in *T. californicus* (across all crosses there are 4 cases where
454 chromosomes show evidence of nuclear-nuclear while 13 cases show evidence of mito-nuclear
455 incompatibilities [Fig. 3]) and may therefore be especially important in the formation of
456 reproductive isolation. This is also supported by the results of crosses between more divergent
457 clades of *T. californicus* where complete (or nearly so) F_1 sterility or inviability is observed. In
458 these crosses, the mitochondrial background appears to determine if F_1 hybrids are sterile or
459 inviable, with some crosses having sterile hybrids with one population's mitochondria, but

460 inviable hybrids with the other population's mitochondria (Ganz and Burton 1995; Peterson et al.
461 2013).

462 The importance of mito-nuclear incompatibilities in speciation is still debated, and much
463 of the work in the field has focused on the importance of sex chromosomes as contributors to the
464 evolution of intrinsic postzygotic barriers (Haldane 1922; Tao et al. 2003; Masly and Presgraves
465 2007). However, a large number of taxa possess modes of sex determination that do not involve
466 heteromorphic sex chromosomes (Bull 1983). In many of these cases mito-nuclear
467 incompatibilities appear to be more common than those between nuclear factors. This is
468 especially true if we also include chloroplast-nuclear incompatibilities in this category (cyto-
469 nuclear incompatibilities when considering DMI involving either cpDNA or mtDNA). Examples
470 of cyto-nuclear incompatibilities contributing to reproductive isolation are overwhelmingly from
471 plants (Fishman and Willis 2001; Sambatti et al. 2008; Rieseberg and Blackman 2010; Scopece
472 et al. 2010; Barnard-Kubow et al. 2016) (Fishman and Willis 2006; Sambatti et al. 2008;
473 Rieseberg and Blackman 2010; Scopece et al. 2010; Bernard-Kubow et al. 2016), but also in
474 yeast (Chou and Leu 2010; Chou et al. 2010), and *T. californicus* as possibly the best example in
475 animals (Ellison and Burton 2008b; 2010; Burton and Barreto 2012). Several groups of fish,
476 amphibians and reptiles have a range of sex determination mechanisms between closely related
477 taxa (Bull 1983; Hillis and Green 1990; Korpelainen 1990; Pokorná and Kratochvíl 2009), and
478 evidence of mito-nuclear incompatibilities has been observed in hybrids of these taxa with a
479 range of sex determination mechanisms (Bolnick et al. 2008; Gagnaire et al. 2013; Lee-Yaw et
480 al. 2014; Bar-Yaacov et al. 2015). These and other putative cases are reviewed in Sunnucks et al.
481 (2017) and Sloan et al. (2017).

There is no doubt that DMI involving sex chromosomes are important in taxa that bare them, exemplified by the large number of species that obey Haldane's Rule (where the heterogametic sex is affected disproportionately by intrinsic postzygotic isolation [Haldane 1922]), but even in some of these cases mito-nuclear incompatibilities can also contribute to decreased hybrid fitness, as has been observed in birds (McFarlane et al. 2016; Hill 2017; Morales et al. 2017; Lamb et al. 2018). One reason why the importance of mito-nuclear incompatibilities may have been underappreciated in these cases is because other reproductive barriers evolve very early in divergence for some of the most studied taxa. For example, complete hybrid male sterility is observed in several interspecific *Drosophila* crosses with *dS* ~0.05 (Turissini et al. 2018), which is the level of divergence of the crosses presented here. Another reason is that the likelihood of mito-nuclear incompatibilities evolving should be dependent on the rate of evolution of mtDNA (Burton and Barreto 2012). Therefore, mito-nuclear incompatibilities are more often observed in taxa with relatively high mtDNA substitution rates. In this sense, *Drosophila*, whose mtDNA has approximately 2x the substitution rate as those in the nuclear genome, should be less likely to evolve these types of incompatibilities than for example ungulates and primates, with substitution rate for mtDNA approximately 20-40 fold higher than for nDNA (Osada and Akashi 2011). For perspective, *T. californicus*' rate of mtDNA evolution has been estimated at 55 fold higher than that of nDNA (Willett 2012), and may therefore be particularly prone to evolve mito-nuclear DMI. A thorough assessment of the role of mtDNA substitution rates on mito-nuclear incompatibilities has yet to be completed.

503 **Literature cited**

504 Alexander, H. J., J. M. L. Richardson, and B. R. Anholt. 2014. Multigenerational response to
505 artificial selection for biased clutch sex ratios in *Tigriopus californicus* populations. J. Evol.
506 Biol. 27:1921–1929.

507 Alexander, H. J., J. M. L. Richardson, S. Edmands, and B. R. Anholt. 2015. Sex without sex
508 chromosomes: genetic architecture of multiple loci independently segregating to determine sex
509 ratios in the copepod *Tigriopus californicus*. J. Evol. Biol. 28:2196–2207.

510 Barreto, F. S., E. T. Watson, T. G. Lima, C. S. Willett, S. Edmands, W. Li, and R. S. Burton.
511 2018. Genomic signatures of mitonuclear coevolution across populations of *Tigriopus*
512 *californicus*. Nature Ecology & Evolution 2:1250.

513 Bar-Yaacov, D., Z. Hadjivasiliou, L. Levin, G. Barshad, R. Zarivach, A. Bouskila, and D.
514 Mishmar. 2015. Mitochondrial involvement in vertebrate speciation? The case of mito-nuclear
515 genetic divergence in chameleons. Genome Biology and Evolution 7:3322–3336.

516 Barnard-Kubow, K. B., N. So, and L. F. Galloway. 2016. Cytonuclear incompatibility
517 contributes to the early stages of speciation. Evolution 70:2752–2766.

518 Bolnick, D. I., M. Turelli, H. Lopez-Fernandez, P. C. Wainwright, and T. J. Near. 2008.
519 Accelerated mitochondrial evolution and "darwin's corollary": Asymmetric viability of
520 reciprocal F1 hybrids in centrarchid fishes. Genetics 178:1037–1048.

521 Bull, J. J. 1983. Evolution of sex determining mechanisms. The Benjamin/Cummings Publishing
522 Company, Inc.

523 Burton, R. S. 1987. Differentiation and integration of the genome in populations of the marine
524 copepod *Tigriopus californicus*. Evolution 41:504–513.

525 Burton, R. S. 1997. Genetic evidence for long term persistence of marine invertebrate
526 populations in an ephemeral environment. Evolution 51:993–998.

527 Burton, R. S. 1985. Mating system of the intertidal copepod *Tigriopus californicus*. Mar. Biol.
528 86:247–252.

529 Burton, R. S., and F. S. Barreto. 2012. A disproportionate role for mtDNA in Dobzhansky-
530 Muller incompatibilities? 21:4942–4957.

531 Burton, R. S., R. J. Pereira, and F. S. Barreto. 2013. Cytonuclear Genomic Interactions and
532 Hybrid Breakdown. Annu. Rev. Ecol. Evol. Syst. 44:281–302.

533 Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in
534 phylogenetic analysis. Molecular Biology and Evolution 17:540–552.

535 Chou, J.-Y., and J.-Y. Leu. 2010. Speciation through cytonuclear incompatibility: Insights from
536 yeast and implications for higher eukaryotes. Bioessays 32:401–411.

537 Chou, J.-Y., and J.-Y. Leu. 2015. The Red Queen in mitochondria: cyto-nuclear co-evolution,
538 hybrid breakdown and human disease. Front. Genet. 6:187.

539 Chou, J.-Y., Y.-S. Hung, K.-H. Lin, H.-Y. Lee, and J.-Y. Leu. 2010. Multiple Molecular
540 Mechanisms Cause Reproductive Isolation between Three Yeast Species. PLoS Biol
541 8:e1000432–13.

542 Coyne, J. A., and H. A. Orr. 1989. Two rules of speciation. Pp. 180–207 *in* D. Otte and J. A.
543 Endler, eds. Speciation and its consequences. Sunderland.

544 Dobzhansky, T. 1936. Studies on hybrid sterility. II. Localization of sterility factors in
545 *Drosophila pseudoobscura* hybrids. Genetics 21:113–135.

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

548 Edmands, S., S. L. Northrup, A. H. Evolution, 2009. 2009. Maladapted gene complexes within
549 populations of the intertidal copepod *Tigriopus californicus*? Evolution 63:2184–2192.

550 Ellison, C. K., and R. S. Burton. 2010. Cytonuclear conflict in interpopulation hybrids: the role
551 of RNA polymerase in mtDNA transcription and replication. J. Evol. Biol. 23:528–538.

552 Ellison, C. K., and R. S. Burton. 2006. Disruption of mitochondrial function in interpopulation
553 hybrids of *Tigriopus californicus*. Evolution 60:1382–1391.

554 Ellison, C. K., and R. S. Burton. 2008a. Genotype-dependent variation of mitochondrial
555 transcriptional profiles in interpopulation hybrids. Proc Natl Acad Sci USA 105:15831–15836.
556 National Academy of Sciences.

557 Ellison, C. K., and R. S. Burton. 2008b. Interpopulation hybrid breakdown maps to the
558 mitochondrial genome. Evolution 62:631–638.

559 Fishman, L., and J. H. Willis. 2001. Evidence for dobzhansky-muller incompatibilites
560 contributing to the sterility of hybrids between *Mimulus guttatus* and *M. nasutus*. Evolution
561 55:1932–1942.

562 Gagnaire, P.-A., E. Normandeau, S. A. Pavey, and L. Bernatchez. 2013. Mapping phenotypic,
563 expression and transmission ratio distortion QTL using RAD markers in the Lake Whitefish
564 (*Coregonus clupeaformis*). Mol. Ecol. 22:3036–3048.

565 Ganz, H. H., and R. S. Burton. 1995. Genetic differentiation and reproductive incompatibility
566 among Baja California populations of the copepod *Tigriopus californicus*. Mar. Biol. 123:821–
567 827. Springer-Verlag.

568 Gershoni, M., A. R. Templeton, and D. Mishmar. 2009. Mitochondrial bioenergetics as a major
569 motive force of speciation. Bioessays 31:642–650.

570 Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. Journal of genetics
571 12:101–109.

572 Hill, G. E. 2017. The mitonuclear compatibility species concept. The Auk 134:393–409.

573 Hillis, D. M., and D. M. Green. 1990. Evolutionary changes of heterogametic sex in the
574 phylogenetic history of amphibians. J. Evol. Biol. 3:49–64. Wiley/Blackwell (10.1111).

575 Kimura, M. 1968. Evolutionary rate at the molecular level. Nature 217:624–626.

576 Kofler, R., P. Orozco-terWengel, N. De Maio, R. V. Pandey, V. Nolte, A. Futschik, C. Kosiol,
577 and C. Schlötterer. 2011a. PoPoolation: a toolbox for population genetic analysis of next
578 generation sequencing data from pooled individuals. PLoS ONE 6:e15925–9.

579 Kofler, R., R. V. Pandey, and C. Schlotterer. 2011b. PoPoolation2: identifying differentiation
580 between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics
581 27:3435–3436.

582 Korpelainen, H. 1990. Sex ratios and conditions required for environmental sex determination in
583 animals. Biological Reviews 65:147–184.

584 Lamb, A. M., H. M. Gan, C. Greening, L. Joseph, Y. P. Lee, A. Morán-Ordóñez, P. Sunnucks,
585 and A. Pavlova. Climate‐driven mitochondrial selection: A test in Australian songbirds. Mol.
586 Ecol. 27: 898-918.

587 Lee-Yaw, J. A., C. G. C. Jacobs, and D. E. Irwin. 2014. Individual performance in relation to
588 cytonuclear discordance in a northern contact zone between long-toed salamander (*Ambystoma*
589 *macrodactylum*) lineages. Mol. Ecol. 23:4590–4602.

590 Levin, D. A. 2012. The long wait for hybrid sterility in flowering plants. New Phytologist
591 196:666–670.

592 Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping
593 and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–
594 2993.

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

597 Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R.
598 Durbin, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map
599 format and SAMtools. Bioinformatics 25:2078–2079.

600 Lima, T. G., and C. S. Willett. 2017. Locally adapted populations of a copepod can evolve
601 different gene expression patterns under the same environmental pressures. Ecol. and Evol.
602 7:4312-4325.
603

604 Lima, T. G., and C. S. Willett. 2018. Using Pool-seq to Search for Genomic Regions Affected by
605 Hybrid Inviability in the copepod *T. californicus*. J. Hered. 256:89.

606 Löytynoja, A. 2013. Phylogeny-aware alignment with PRANK. Pp. 155–170 *in* Multiple
607 Sequence Alignment Methods. Humana Press, Totowa, NJ.

608 Masly, J. P., and D. C. Presgraves. 2007. High-resolution genome-wide dissection of the two
609 rules of speciation in *Drosophila*. PLoS Biol 5:e243.

610 McFarlane, S. E., P. M. Sirkiä, M. Ålund, and A. Qvarnström. 2016. Hybrid dysfunction
611 expressed as elevated metabolic rate in male *Ficedula flycatchers*. PLoS ONE 11:e0161547–10.

612 Morales, H. E., P. Sunnucks, L. Joseph, and A. Pavlova. 2017. Perpendicular axes of
613 differentiation generated by mitochondrial introgression." Mol. Ecol. 26: 3241-3255.

614 Muller, H. J. 1942. Isolating mechanisms, evolution and temperature. Biol. Symp. 6:71–125.

615 Osada, N., and H. Akashi. 2011. Mitochondrial–nuclear interactions and accelerated
616 compensatory evolution: Evidence from the primate cytochrome c oxidase complex. Molecular
617 Biology and Evolution 29:337–346.

618 Pereira, R. J., F. S. Barreto, N. T. Pierce, M. Carneiro, and R. S. Burton. 2016. Transcriptome-
619 wide patterns of divergence during allopatric evolution. Mol. Ecol. 25:1478–1493.

620 Peterson, D. L., K. B. Kubow, M. J. Connolly, L. R. Kaplan, M. M. Wetkowski, W. Leong, B. C.
621 Phillips, and S. Edmands. 2013. Reproductive and phylogenetic divergence of tidepool copepod
622 populations across a narrow geographical boundary in Baja California. J. Biogeogr. 40:1664–
623 1675.

624 Phillips, B. C., and S. Edmands. 2012. Does the speciation clock tick more slowly in the absence
625 of heteromorphic sex chromosomes? 34:166–169.

626 Pokorná, M., and L. Kratochvíl. 2009. Phylogeny of sex-determining mechanisms in squamate
627 reptiles: are sex chromosomes an evolutionary trap? Zoological Journal of the Linnean Society
628 156:168–183. Wiley/Blackwell (10.1111).

629 Presgraves, D. C., and H. A. Orr. 1998. Haldane's Rule in taxa lacking a hemizygous X. Science
630 282:952–954.

631 Rieseberg, L. H., and B. K. Blackman. 2010. Speciation genes in plants. Annals of Botany
632 106:439–455.

633 Sambatti, J., D. Ortiz-Barrientos, E. J. Baack, and L. H. Rieseberg. 2008. Ecological selection
634 maintains cytonuclear incompatibilities in hybridizing sunflowers. Ecol Letters 11:1082–1091.
635 Wiley Online Library.

636 Sambrook, J., and D. W. Russell. 2006. Purification of nucleic acids by extraction with
637 Phenol:Chloroform. Cold Spring Harb Protoc 2006:pdb–prot4455.

638 Schlötterer, C., R. Tobler, R. Kofler, and V. Nolte. 2014. Sequencing pools of individuals —
639 mining genome-wide polymorphism data without big funding. Nat Rev Genet 15:749–763.
640 Nature Publishing Group.

641 Scopece, G., C. Lexer, A. Widmer, and S. C. 2010. Polymorphism of postmating reproductive
642 isolation within plant species. Taxon 59:1367–1374.

643 Sloan, D. B., J. C. Havird, and J. Sharbrough. 2017. The on-again, off-again relationship
644 between mitochondrial genomes and species boundaries. Mol. Ecol. 26:2212–2236.

645 Sunnucks, P., H. E. Morales, A. M. Lamb, A. Pavlova, and C. Greening. 2017. Integrative
646 approaches for studying mitochondrial and nuclear genome co-evolution in oxidative
647 phosphorylation. Front. Genet. 8:589–12.

648 Turelli, M. and L. C. Moyle. 2007. Asymmetric postmating isolation: Darwin's Corollary to
649 Haldane's Rule. Genetics 176: 1059-1088.

650 Turissini, D. A., J. A. McGirr, S. S. Patel, J. R. David, and D. R. Matute. 2017. The rate of
651 evolution of postmating-prezygotic reproductive isolation in *Drosophila*. Mol. Biol. and Evol.
652 35: 312-334.

653 Tao, Y., S. Chen, D. L. Hartl, and C. C. Laurie. 2003. Genetic dissection of hybrid
654 incompatibilities between *Drosophila simulans* and *D. mauritiana*. I. Differential accumulation
655 of hybrid male sterility effects on the X and autosomes. Genetics 164:1383–1398. Genetics.

656 Voordouw, M. J., and B. R. Anholt. 2002. Environmental sex determination in a splash pool
657 copepod. Biological Journal of the Linnean Society 76:511–520.

658 Willett, C. S. 2012. Quantifying the elevation of mitochondrial DNA evolutionary substitution
659 rates over nuclear rates in the intertidal copepod *Tigriopus californicus*. J Mol Evol 74:310–318.

660 Willett, C. S., and J. N. Berkowitz. 2007. Viability effects and not meoitic drive cause dramatic
661 departures from Mendelian inheritance for malic enzyme in hybrids of *Tigriopus californicus*
662 populations. J. Evol. Biol. 20: 1196-1205.

663 Willett, C. S. 2008. Significant Variation for Fitness Impacts of ETS Loci in Hybrids between
664 Populations of Tigriopus californicus. J Hered 99:56–65.

665 Willett, C. S., and J. T. Ladner. 2009. Investigations of fine-scale phylogeography in *Tigriopus*
666 *californicus* reveal historical patterns of population divergence. BMC Evol Biol 9:1–20.

667 Yang, Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Molecular Biology
668 and Evolution 24:1586–1591.

669

Tables

Table 1. Summary statistics for the three crosses of *T. californicus* populations. Windows are the average allele frequency for 3000 consecutive SNPs, in non-overlapping windows. Mean genome-wide allele frequency (± standard deviation) was calculated using the allele frequencies averaged across all windows for each cross. d*S* is the average rate of synonymous substitutions across all annotated genes in the *T. californicus* genome (± standard error). Allele frequency always refers to the AB allele.

Cross	Mean depth of coverage	Number of SNP	Number of windows	Base-pairs per window (± s.d.)	Mean genome-wide allele frequency (± s.d.)	d*S* (± SEM)	mt % nucleotide divergence
SD ♀ x AB ♂ (DA)	77	2,106,984	698	256,820 (± 61,282)	0.505 (± 0.056)	0.048 (± 1.99E-04)	20.8
AB ♀ x SD ♂ (**AD)**	98				0.521 (± 0.047)		
CAT ♀ x AB ♂ (**CA)**	229	2,433,118	805	222,464 (± 69,131)	0.478 (± 0.022)	0.048 (± 2.01E-04)	19.5
AB ♀ x CAT ♂ (**AC)**	118				0.489 (± 0.053)		
SC ♀ x AB ♂ (**SA)**	71	2,275,994	754	237,813 (± 75,634)	0.497 (± 0.030)	0.052 (± 4.07E-04)	20.7
AB ♀ x SC ♂ (**AS)**	107				0.487 (± 0.042)		

676 **Figure legends**

677 **Figure 1.** Expected allele frequencies for different potential incompatibility scenarios for *T.*
678 *californicus* hybrids with alternate mtDNA backgrounds. (A) Shows the reciprocal cross design
679 for crosses between two different populations with the copepod body color representing the
680 nuclear genome and the circle the mtDNA. These crosses result in two reciprocal cross
681 populations of F2 hybrids with variable nuclear genomes and alternate mtDNA-types. For
682 regions of the genome displaying mitonuclear coadaptation the expectation is that there will be a
683 higher allele frequency for the allele that matches the mtDNA-type with (B) showing a
684 hypothetical example of both a genomic region showing a match pattern and another region
685 showing a mismatch pattern. (C) Expected patterns of AB allele frequency in F2 hybrids
686 between two reciprocal crosses for each of six different outcomes consistent with the three
687 different scenarios of nuclear-nuclear, mito-nuclear, or mismatch incompatibilities. Bars depict
688 the range of possibilities for F2 AB allele frequencies for nuclear genes on the two different
689 mtDNA backgrounds in the two reciprocal crosses that are consistent with that outcome.
690

691 **Figure 2.** Allele frequency distributions for F_2 hybrids from three population crosses of *T.*
692 *californicus*. Allele frequencies are based on the allele frequency of the AB alleles (x-axis). Y-
693 axis is the count of allele frequency windows. Allele frequency windows are the average allele
694 frequency of 3000 consecutive SNPs. **a.** SD x AB; **b.** CAT x AB; **c.** SC x AB. Distributions in
695 red are for the direction of the cross with AB mitochondria and distributions in blue have the
696 mitochondria for the other populations.
697

698 **Figure 3.** Allele frequency plots for F_2 hybrids across 12 chromosomes from three population

crosses of *T. californicus*. Allele frequencies are based on the allele frequency of the AB alleles (y-axis). The x-axis indicates the relative position across each chromosome. Data points are the average allele frequency of 3000 consecutive SNPs. **a.** SD x AB; **b.** CAT x AB; **c.** SC x AB. Red dots indicate the direction of the cross with AB mitochondria and blue dots have the mitochondria for the other populations. Dark grey boxes indicate the level of variation observed in a null dataset. Light grey boxes indicate levels of allele frequency change considered strongly skewed. Colored diamonds below the allele frequency plots refer to the chromosomal positions of nuclear encoded genes that interact with mitochondria proteins. Green diamonds: OXPHOS genes; black diamonds: mitochondrial aminoacyl tRNA synthethases genes; magenta diamonds: transcription genes Differences between reciprocal crosses indicate the presence of mito-nuclear incompatibilities (shaded in red) or a mismatch pattern. Differences between each individual cross and the null allele frequency distribution indicates that cross deviates from the null expectation allele frequency, and suggest the presence of nuclear-nuclear incompatibilities (shaded in blue).

Figures

Figure 1

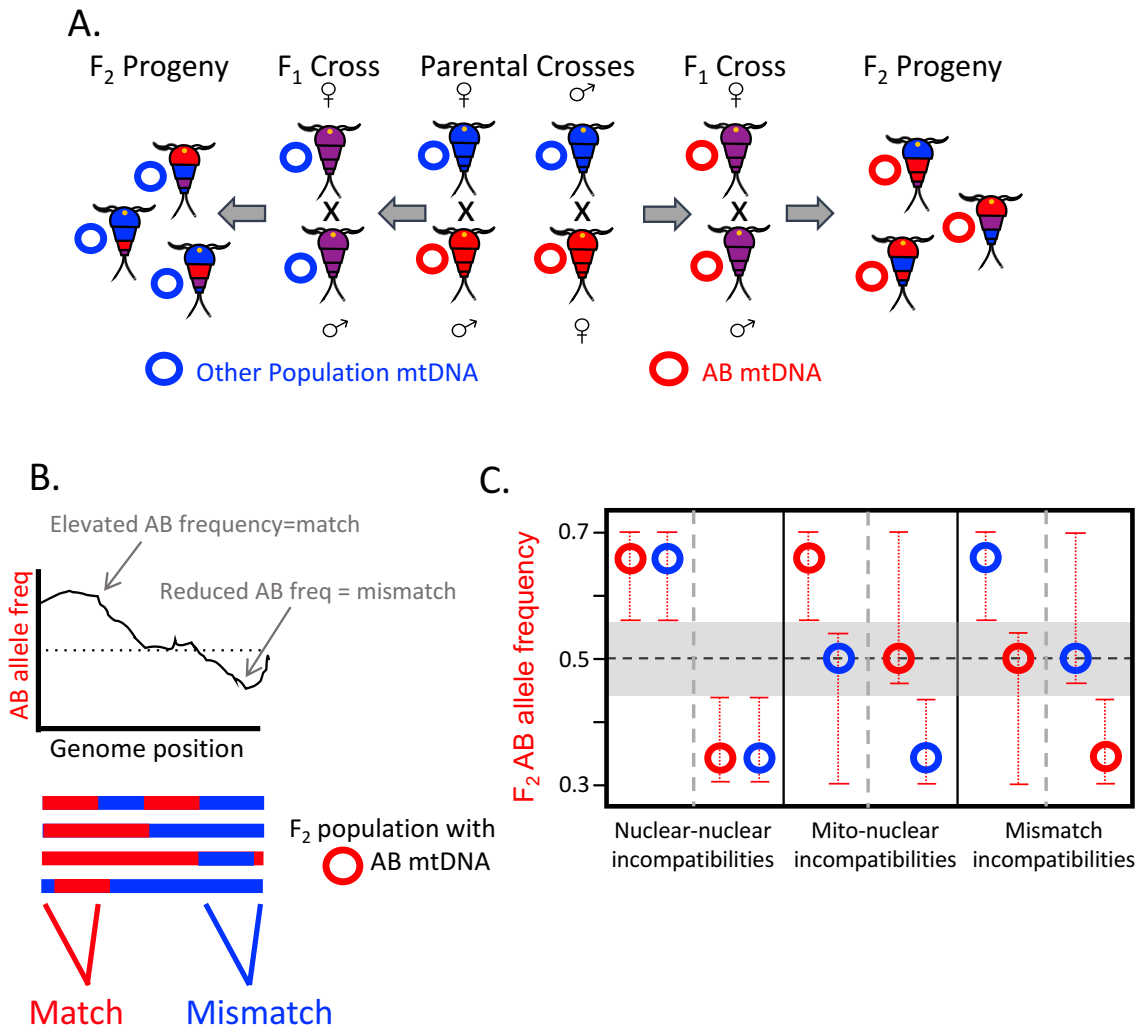


Figure 2

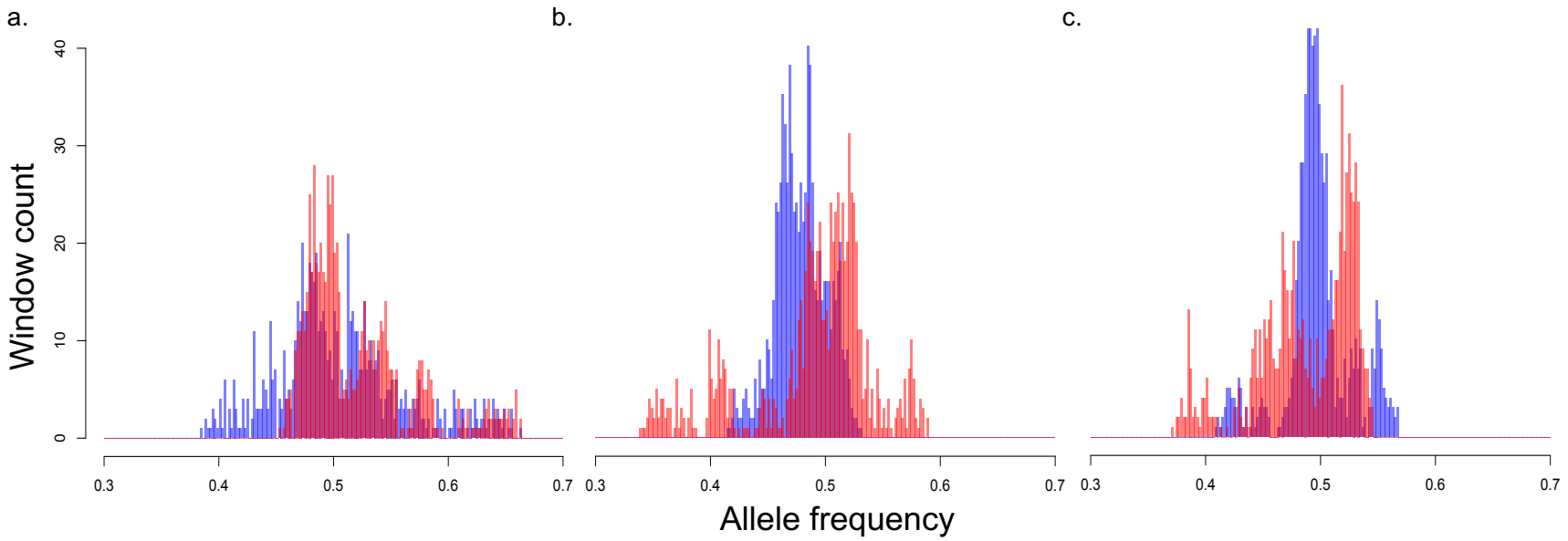
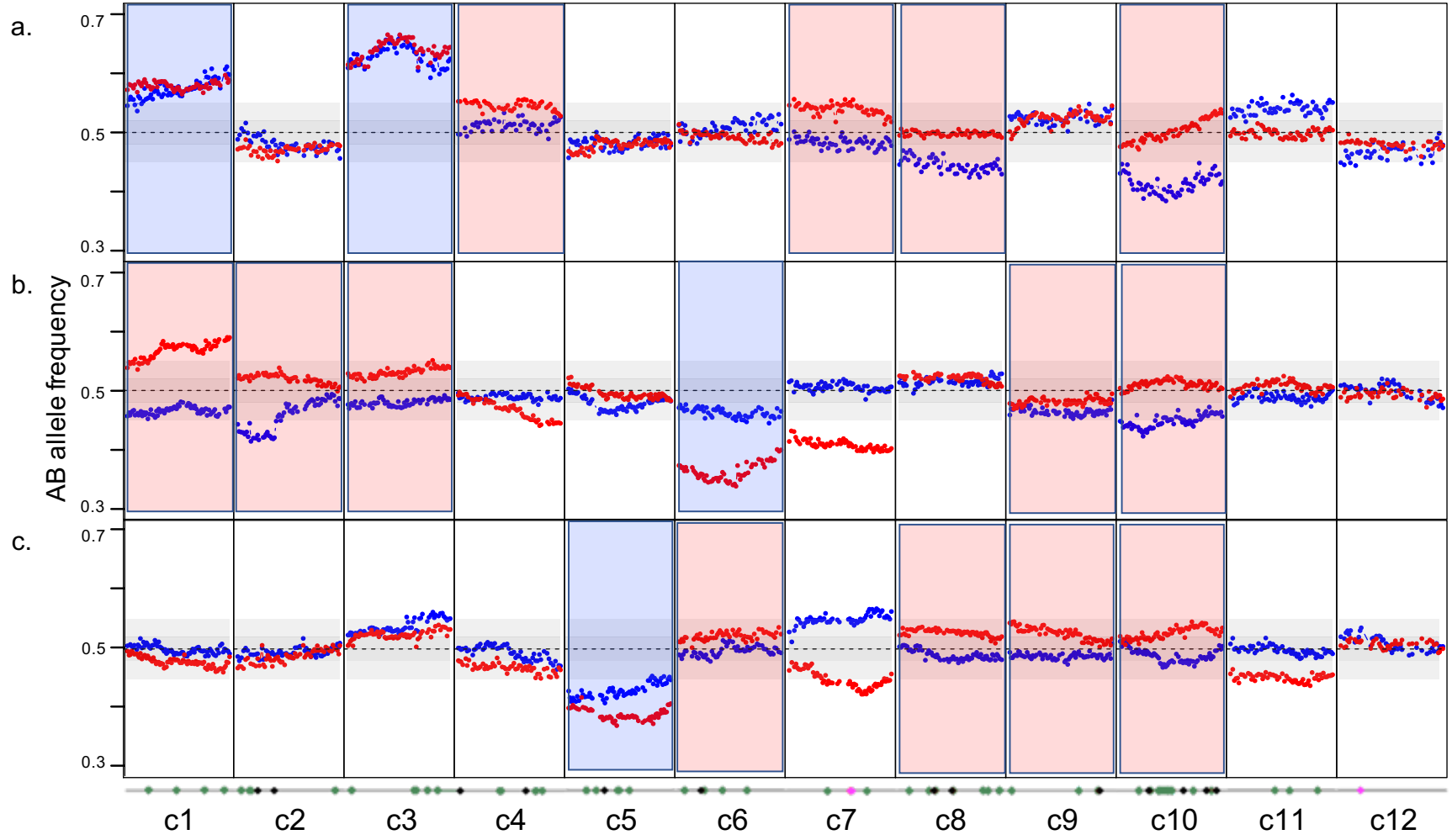


Figure 3


a.
b.
c.
AB allele frequency
0.7
0.5
0.3
c1
c2
c3
c4
c5
c6
c7
c8
c9
c10
c11
c12