Evidence of a Paucity of Genes That Interact with the Mitochondrion on the X in Mammals

Devin M. Drown*, Kevin M. Preuss, and Michael J. Wade
Department of Biology, Indiana University
*Corresponding author: E-mail: dmdrown@indiana.edu.
Accepted: July 16, 2012

Abstract
Mitochondria are essential organelles whose replication, development, and physiology are dependent upon coordinated interactions between gene combinations from both the mitochondrial and the nuclear genomes. The evolution of coadapted (CA) nuclear–mitochondrial gene combinations would be facilitated if such nuclear genes were located on the X-chromosome instead of on the autosome because of the increased probability of cotransmission. Here, we test the prediction of the CA hypothesis by investigating the chromosomal distribution of nuclear genes that interact with mitochondria. Using the online genome database BIOMART, we compared the density of genes that have a mitochondrion cellular component annotation across chromosomes in 16 vertebrates. We find a strong and highly significant genomic pattern against the CA hypothesis: nuclear genes interacting with the mitochondrion are significantly underrepresented on the X-chromosome in mammals but not in birds. We interpret our findings in terms of sexual conflict as a mechanism that may generate the observed pattern. Our finding extends single-gene theory for the evolution of sexually antagonistic genes to nuclear–mitochondrial gene combinations.

Key words: genomic conflict, gene transfer, sex chromosome, coadapted gene complex.

Introduction
Mitochondria are essential organelles whose replication, development, and physiology depend upon coordinated interactions between gene combinations from both the mitochondrial and the nuclear genomes. When genes from both genomes interact, their differing modes of inheritance, bi-parental versus maternal, can result in genomic conflict (Werren 2011). There is direct evidence of the genomic conflict generated by epistatic interactions between the mitochondrial genes and nuclear genes from studies of mitochondria involvement in sperm development (Wang 2004; Rajender et al. 2010; Paoli et al. 2011). Furthermore, introgression of foreign mitochondria into Drosophila populations has revealed sex-biased epistatic effects with X-linked nuclear genes (Rand et al. 2001, 2006; Montooth et al. 2010; Aw et al. 2011). When organismal fitness depends on the epistatic interactions between the nuclear and mitochondrial genomes (Rand et al. 2001, 2006; Dowling et al. 2007), the cotransmission of nuclear–mitochondrial gene combinations facilitates epistatic selection (Wade and Goodnight 2006; Brandvain and Wade 2009). Cotransmission is important to epistatic selection because it maintains gene combinations across generations. By increasing the heritability of gene combinations, cotransmission allows epistatic selection to act more effectively on trans-genomic interactions.

The question remains, as to how these genome interactions have shaped the geography of the genome. Theory predicts that the nuclear genome should evolve in a way to maintain nuclear–mitochondrial gene combinations (Wade and Goodnight 2006; Brandvain and Wade 2009). One means of increasing the heritability of nuclear–mitochondrial gene combinations involves changing the physical location of nuclear genes within the nuclear genome as opposed to changing location between the nuclear and mitochondrial genomes. The probability of cotransmission of X-linked genes and mitochondria is significantly higher than that of mitochondria and...
autosomal genes (0.67 versus 0.50 in a randomly mating population) (Rand et al. 2001). Therefore, the evolution of coadapted (CA) nuclear–mitochondrial gene combinations would be facilitated if nuclear genes interacting with the mitochondria were located on the X-chromosome instead on the autosomes because of the increased probability of cotransmission (Rand et al. 2001).

The CA nuclear–mitochondrial gene hypothesis posits that selection for beneficial nuclear–mitochondrial gene combinations and against poor ones is facilitated by the location of nuclear genes located on the X chromosome relative to the autosomes. A clear prediction of the CA hypothesis is that nuclear genes which interact with the mitochondrion (N-mt genes), should be overrepresented on the X relative to autosomes. Here, we tested the prediction of the CA by investigating the chromosomal distribution of nuclear genes that interact with the mitochondria.

**Methods**

**Data Collection**

Using the online genome database BIOMART (http://www.biomart.org) via their MartView interface (Smedley et al. 2009), we collected genomic information on the following organisms using the available ENSEMBL GENES 63 datasets: *Homo sapiens* (GRCH37.p3) (International Human Genome Sequencing Consortium 2004), *Pan troglodytes* (CHIMP2.1) (Chimpanzee Sequencing and Analysis Consortium 2005), *Gorilla gorilla* (gorGor3) (Scally et al. 2012), *Pongo pygmaeus* (PPYG2) (Locke et al. 2011), *Macaca mulatta* (MMUL_1.0) (Rhesus Macaque Genome Sequencing Analysis Consortium et al. 2007), *Callithrix jacchus* (callaC3) (GenBank Assembly ID GCA_000004665.1), *Mus musculus* (NCBI37) (Gregory et al. 2002), *Rattus norvegicus* (RGSC3.4) (Gibbs et al. 2004), *Oryctolagus cuniculus* (oryCun2.0) (Lindblad-Toh et al. 2011), *Canis familiaris* (CanFam_2.0) (Lindblad-Toh et al. 2005), *Bos taurus* (Btau_4.0) (Bovine Genome Sequencing Analysis Consortium et al. 2009), *Sus scrofa* (Sscrofa9) (GenBank Assembly ID GCA_000003025.4), *Equus caballus* (EquCab2) (Lindblad-Toh et al. 2011), *Monodelphis domestica* (monDom5) (Mikkelsen et al. 2007), *Gallus gallus* (WASHUC2) (Hillier et al. 2004), and *Taeiopygia guttata* (taeGut3.2.4) (Warren et al. 2010). We focused on mammalian datasets because they contain the greatest level of annotation, including chromosomal location. Also, we restricted our analysis to this single database so that we could make comparisons between genomes while minimizing data collection artifacts.

For each of the 16 genomes, we collected two gene sets: 1) complete list of all genes with annotated function; 2) including only those nuclear genes with mitochondrion annotation (N-mt genes) as specified by the Gene Ontology ID 0005739. From the database, we obtained for each gene: chromosomal location, ENSEMBLE gene name, and gene ID. Genes without a specific chromosomal location were excluded. We processed this output with scripts written in MATLAB (The Mathworks, Natick, MA) to calculate the counts of genes per chromosome in each genome for the two different datasets.

**Statistical Analysis**

We compared the density of N-mt genes on the X or Z chromosomes with that of the autosomes relative to the expectation based on overall gene density for 14 vertebrate genomes with male heterogamety and two vertebrate genomes with female heterogamety (the birds, *G. gallus* and *T. guttata*). We performed all of our statistical analysis in MATLAB. To determine if the set of N-mt genes was under- or overrepresented on each chromosome, we compared it with an expected count. For a particular genome, we used the following method to determine the expected number of N-mt genes per chromosome. We calculated the product of the total number of N-mt genes and the fraction of all annotated genes found on each chromosome. This expected gene count per chromosome assumes that the N-mt genes are distributed across the chromosomes in the same pattern as the total set of genes in the genome. The ratio of observed count to expected count will equal one when there is no under- or overrepresentation of N-mt genes on a particular chromosome. When the value is less (greater) than one, then there is an underrepresentation (overrepresentation) of N-mt genes on a chromosome.

We computed the comparisons as described above to limit any potential biases in how the data were collected. In order to create a bias in the density of this particular group of genes, some chromosomes would need to be over- or underannotated for N-mt genes in particular. We are unaware of any feature of the N-mt genes which make them more or less likely to be sequenced and subsequently recognized as genes relative to any other subset of genes. Our method is also robust to differences in annotation level between chromosomes within a genome. If for some reason there was a bias in the annotation level of autosomes compared with the sex chromosome, we would expect more genes to be annotated with the N-mt term as well. Our density of N-mt genes (rather than an absolute count) would not be affected.

We tested the significance of the underrepresentation on the sex chromosome by calculating a confidence interval of the mean over-/underrepresentation for the autosomes based on 10,000 replicate bootstrap samples and then determining if the sex chromosome representation fell below this confidence interval at a Šidák corrected alpha of 0.032, based on an individual alpha of 0.05 and 16 different comparisons (Šidák 1967).
Results
Using the online genome database BIOMART (http://www.biomart.org) via their MartView interface (Smedley et al. 2009), we examined the distribution of genes that have a mitochondrion cellular component annotation (Gene Ontology ID 0005739) across chromosomes in 16 vertebrates. We found a strong and significant underrepresentation of nuclear genes interacting with the mitochondrion on the sex chromosomes in all of the mammals (fig. 1) but in neither of the two birds. In mammals, although the sex chromosome did not always have the lowest density of N-mt genes relative to expectation, it was always in the lower end of the distribution of chromosomes. In both bird species, the sex chromosomes (Z) had the expected density of N-mt genes based on overall gene density.

Discussion
We have discovered a striking pattern in the distribution of genes among chromosomes: nuclear genes which interact with the mitochondrion are significantly underrepresented on the X-chromosome in mammals, but not in birds (fig. 1). These data are inconsistent with the predictions of the CA nuclear–mitochondrial gene hypothesis which proposes that due to the increased heritability of gene combinations, the N-mt genes should be overrepresented on the X-chromosome. We discuss other hypotheses that may help to explain the observed pattern of gene distributions.

We propose that the observed pattern may be explained by sexual conflict (SC). SC occurs when genes beneficial to the fitness of one sex are deleterious to the fitness of the other. The evolution of maternally transmitted organelles can lead to SC when mitochondrial mutations that interact epistatically with nuclear genes have favorable effects in females but are deleterious in males (Rice 1984; Rice et al. 2006; Vicoso and Charlesworth 2006). With epistatic selection, the effects of nuclear genes interacting with mitochondria manifest only on certain mitochondrial backgrounds and not on others (fig. 2A). Reciprocally, the effects of mitochondrial genes occur only in specific nuclear backgrounds. Importantly, changing the mitochondrial background alters the effects of the interacting nuclear genes without changing the nuclear gene sequence. Epistatic fitness interactions between nuclear genes and the mitochondrion can result in sexually antagonistic effects for nuclear genes.

With nuclear–mitochondrial epistasis, the force of genomic conflict (between the nucleus and the mitochondrion) is aligned with that of intralocus SC (between males and females). Because mitochondria have nearly exclusive maternal inheritance, selection can favor those mitochondrial genotypes that enhance female fitness, even if at the expense of male fitness (Partridge and Hurst 1998; Rand et al. 2006; Montooth et al. 2010; but see Wade and Brandvain 2009).

That is, mitochondrial mutations with epistatic effects favorable to females will spread owing to maternal inheritance (fig. 2B). Because the mitochondrial mutations are not selected on the basis of their effect in males, even mutations which carry unfavorable effects in males will spread as long as those mutations are favorable in females. As a result, a nuclear gene interacting with the mitochondria will become more deleterious to males over time.

There is direct evidence of sexually antagonistic epistatic interactions between the mitochondrial and nuclear genes from studies of the involvement of mitochondria in sperm development (Wang 2004; Rajender et al. 2010; Paoli et al. 2011). Furthermore, introgression of foreign mitochondria into Drosophila populations has revealed sex-biased epistatic effects for X-linked nuclear genes (Rand et al. 2001, 2006; Montooth et al. 2010; Aw et al. 2011).

Although SC based on segregating nuclear genes depends upon the nature of allelic dominance (Rice 1984), selection on the mitochondrial mutations is independent of the dominance relationships among alleles of its nuclear partners, requiring only positive fitness effects on females in some nuclear backgrounds. Furthermore, the SC process we are proposing does not depend on the simultaneous polymorphism of both nuclear and mitochondrial genes or the cotransmission of nuclear–mitochondrial gene combinations, which was a key to the alternative CA hypothesis. SC due to the spread of mitochondrial mutations with epistatic effects with nuclear loci can occur whether the nuclear partner of the mitochondrion is autosomal or X-linked.

As a mitochondrial mutation causing the SC spreads within the population, there is increasing selective pressure to resolve the SC. Several mechanisms for the nuclear resolution of the SC have the potential to alter the genomic distribution of N-mt genes. An evolutionary means for resolving such SC (among many) is via gene movement off of the X and onto the autosomes. This is a process involving several steps, including gene duplication, fixation and subsequent gene loss, and requiring significant chromosomal bias at one or all of these steps to account for the observed pattern (Wu and Yujun Xu 2003).

An alternative process is the creation of duplicate copies on the autosomes with sex-specific or sex-biased expression (Connnallon and Clark 2011; Gallach and Betrán 2011). Both proposed mechanisms involve the fixation of nuclear gene duplications which ameliorate the male detrimental effects caused by the mitochondrial mutation.

Theory and empirical evidence suggest that these duplicate copies accumulate on the autosomes (Connnallon and Clark 2011; Gallach and Betrán 2011). The bias in the location of the duplicate copies creates the overrepresentation on the autosomes relative to the X chromosome. Female beneficial (male deleterious) epistatic effects with mitochondrial mutations are not favored equally at the X and autosomal locations because X-linked loci spend two-thirds of their time in females and only one-third in males (Rice 1984). Autosomal loci, on
the other hand, spend half their time in males and such genes 
with male beneficial (female deleterious) epistatic effects 
would be expected to fix more easily on autosomes compared 
with the X chromosome. This autosomal-bias in selection ef-
ficiency may be particularly important for male traits involved 
in reproductive competition, where selection on male fertility 
can be several times stronger than opposing selection on 
females (Wade 1979; Shuster and Wade, 2003; Miller et al. 
2006). Over time, the difference in selection efficiency would 
lead to a greater cumulative deterioration of X-linked relative 
to autosomal nuclear genes as a result of mitochondrial mu-
tation pressure (fig. 2). Empirical support for this mechanism 
comes from Drosophila where Gallach et al. (2010) find that 
N-mt gene duplications accumulate additional copies on auto-
somes more readily than the X chromosome. Interestingly, 
these gene duplicates also show testis-specific expression that 
the authors suggest is evidence of SC resolution (Gallach et al. 2010).

Other hypotheses regarding the resolution of SC suggest 
that changes in the sex-specific expression of sexually antag-
onistic genes, possibly via cis-acting expression modifiers, re-
solve intralocus SC (Rice 1984; Rice and Chippindale 2001; 
Gibson et al. 2002; Connallon and Clark 2010; Stewart et al. 
2010). Our data are based on the genomic distribution of 
N-mt genes and not on sex differences in gene expression. 
Thus, although our data in mammals are consistent with reso-
lution of SC via gene movement or duplication, they do not 
exclude the possibility that SC antagonism may be resolved by 
alternative means. Our data suggest that mammalian N-mt 
genes may be useful for elucidating the mechanistic basis by 
which SC is resolved (Vibranovski et al. 2009; Gallach et al. 
2011; Han and Hahn 2012).

Fig. 1.—Underrepresentation of nuclear–mitochondrial genes on the sex chromosomes of mammals. Expected counts of genes are calculated 
based on the distribution of all annotated genes in the genome. The fraction of all genes present on a particular chromosome is then multiplied by 
the total number of N-mt genes in a particular genome to give the expected count. Ratios close to one show no over- or underrepresentation on a chromosome. 
Each chromosomal value for each genome is represented by a blue (autosomes) or red and pink (sex chromosome [X/Z]) circle. Lines connect the maximum 
and minimum autosomal ratios. Values in parenthesis after the genome name indicate the number of identified N-mt genes in the genome. Sex chromo-
somes that are significantly different from the autosomes are highlighted in red and nonsignificant values are pink (H. sapiens, P < 0.0001; P. troglodytes, 
P < 0.0001; G. gorilla, P < 0.0001; P. pygmaeus, P < 0.0001; M. mulatta, P < 0.0001; C. jacchus, P < 0.0001; M. musculus, P < 0.0001; R. norvegicus, P < 
0.0001; O. cuniculus, P < 0.0001; C. familiaris, P < 0.0001; B. taurus, P < 0.0001; S. scrofa, P = 0.0001; E. caballus, P < 0.0001; M. domestica, P < 0.0001; 
G. gallus, P = 0.7621; and T. guttata, P = 0.5281).
First, there is reduced mitochondrial mutation pressure in birds (Ogburn et al. 1998; Hickey 2008). Second, there is a near absence of nuclear gene retro-transposition in birds relative to mammals (Haas et al. 2001; Warren et al. 2010) but see (Ellegren 2011). Third, if the fitness effect is caused by epistatic interactions with the mitochondrion, then strict maternal inheritance of the mitochondria results in reduced cotransmission of the mitochondria and Z chromosomes compared to autosomes. Additionally, all of the cotransmission occurs from mother to sons at which point that mitochondrial-Z chromosome combination has a zero probability of being passed along to the next generation. The combination these factors suggests that our proposed process will be significantly weaker in birds, relative to mammals and that intralocus SC in birds may be resolved differently than it is in mammals.

We appreciate that our sample of genomes does not meet the criterion of phylogenetic independence. One could argue that we have a single contrast between two types of male heterogametic vertebrates (mammals and marsupials) and one type of female heterogametic vertebrate (birds).

Although other vertebrate genomes are available, none meet the criteria of having sex chromosomes (e.g., Danio rerio and Fugu rubripes) and the insect gene annotations in this database are not coincident with their annotations in other databases (e.g., drosophilids GO annotation in BIOMART relative to Flybase). It is well recognized that genomes are highly fluid, with frequent translocations among chromosomes, even though a recognizable amount of synteny is present (Hillier et al. 2004; Ross et al. 2005). Thus, despite its taxonomic limitations, our data suggest that the signal of a lack of N-mt genes on the X chromosomes of mammals is present and has been maintained across a diversity of species for tens of millions of years.

Our results significantly extend the single-gene theory for the evolution of sexually antagonistic genes to involve epistatic interactions between nuclear and mitochondrial genes. It also synergistically combines SC theory with results from genomic conflict theory, wherein, mitochondrial genes with deleterious effects on males spread if they are favorable to females owing to uniparental, maternal transmission. The high rate of such mitochondrial mutations and their subsequent adaptive spread leads to the continual accumulation of male deleterious fitness effects on X-linked genes (Rand et al. 2001). Our theory does not preclude patterns of sex-limited gene expression for sex-linked genes with direct (instead of epistatic) fitness effects. Our empirical finding adds to our knowledge of the forces affecting gene clustering and gene expression on sex chromosomes. Our results are consistent with a synergistic interaction between the forces of SC and genomic conflict affecting the distribution of N-mt genes on the sex chromosomes relative to the autosomes in vertebrates.

**Acknowledgments**

This research was improved by discussion with the working group at Indiana University and by comments from Kristi Montooth, Mira Han, and anonymous reviewers. This work was supported by National Institute of General Medical Sciences at the National Institutes of Health (grant numbers R01GM084238, R01GM065414 to M.J.W.). We thank the Marmoset Genome Sequencing Consortium for making their data publicly available and the Baylor College of Medicine—Human Genome Sequencing Center for providing the Marmoset Genome Project data. We thank the Swine Genome Sequencing Consortium for making their data publicly available and University of Illinois—Institute for Genomic Biology for providing the Sus scrofa data.

**Literature Cited**

Aw WC, Correa CC, Clancy DJ, Ballard JW. 2011. Mitochondrial DNA variants in Drosophila melanogaster are expressed at the level of the organismal phenotype. Mitochondrion 11:756–763.

Bovine Genome Sequencing Analysis Consortium, Eslak CG, Tellam RL, Worley KC. 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science 324:522–528.
to fixed incompatibilities between species of Drosophila. Evolution 64: 3364–3379.

Ogburn CE, et al. 1998. Cultured renal epithelial cells from birds and mice: enhanced resistance of avian cells to oxidative stress and DNA damage. J Gerontol Ser A Biol Sci Med Sci. 53:8287–8292.

Paoli D, et al. 2011. Mitochondrial membrane potential profile and its correlation with increasing sperm motility. Fertil Steril. 95:2315–2319.

Partridge L, Hurst LD. 1998. Sex and conflict. Science 281:2003–2008.

Rajender S, Rahul P, Mahdi AA. 2010. Mitochondria, spermatogenesis and male infertility. Mitochondrion 10:419–428.

Rand DM, Clark AG, Kann LM. 2001. Sexually antagonistic cytonuclear fitness interactions in Drosophila melanogaster. Genetics 159:173–187.

Rand DM, Fry A, Sheldahl L. 2006. Nuclear-mitochondrial epistasis and Drosophila aging: introgression of Drosophila simulans mtDNA modifies longevity in D-melanogaster nuclear backgrounds. Genetics 172:329–341.

Rhesus Macaque Genome Sequencing Analysis Consortium, et al. 2007. Evolutionary and biomedical insights from the rhesus macaque genome. Science 316:222–234.

Rice WR. 1984. Sex-chromosomes and the evolution of sexual dimorphism. Evolution 38:735–742.

Rice WR, Chippindale AK. 2001. Intersexual ontogenetic conflict. J Evol Biol. 14:685–693.

Rice WR, et al. 2006. Assessing sexual conflict in the Drosophila melano-gaster laboratory model system. Phil Trans R Soc B Biol Sci. 361:287–299.

Ross MT, et al. 2005. The DNA sequence of the human X chromosome. Nature 434:325–337.

Scally A, et al. 2012. Insights into hominin evolution from the gorilla genome sequence. Nature 483:169–175.

Shuster SM, Wade MJ. 2003. Mating systems and strategies. Princeton, NJ: Princeton University Press.

Sidák Z. 1967. Rectangular confidence region for the means of multi- variate normal distributions. J Am Stat Assoc. 62:626–633.

Smedley D, et al. 2009. BioMart—biological queries made easy. BMC Genomics 10:22.

Stewart AD, Pschedda A, Rice WR. 2010. Resolving intralocus sexual conflict: genetic mechanisms and time frame. J Heredity. 101:594–599.

Vibranovski MD, Zhang Y, Long MY. 2009. General gene movement off the X chromosome in the Drosophila genus. Genome Res. 19:897–903.

Vicoso B, Charlesworth B. 2006. Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet. 7:645–653.

Wade MJ. 1979. Sexual selection and variance in reproductive success. Am Nat. 114:742–747.

Wade MJ, Brandvain Y. 2009. Reversing mother’s curse: selection on male mitochondrial fitness effects. Evolution 63:1084–1089.

Wade MJ, Goodnight CJ. 2006. Cyto-nuclear epistasis: two-locus random genetic drift in hemphroditic and dioecious species. Evolution 60:643–659.

Wang PJ. 2004. X chromosomes, retrogenes and their role in male repro-duction. Trends Endocrinol Met. 15:79–83.

Warren WC, et al. 2010. The genome of a songbird. Nature 464:757–762.

Werren JH. 2011. Selfish genetic elements, genetic conflict, and evolution- ary innovation. Proc Natl Acad Sci U S A. 108:10863–10870.

Wu C-I, Yujun Xu E. 2003. Sexual antagonism and X inactivation—the SAXI hypothesis. Trends Genet. 19:243–247.