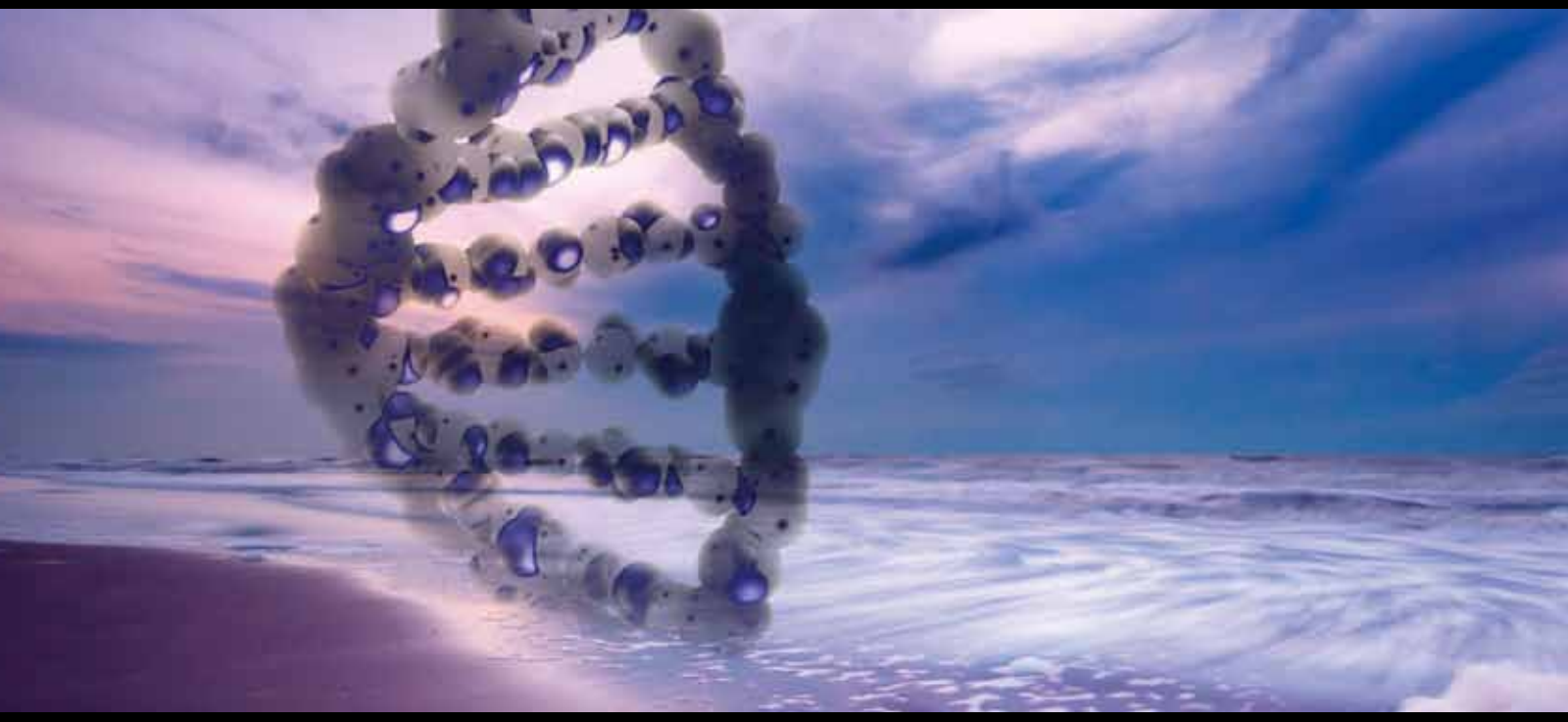
International Journal of Evolutionary Biology

The Evolution of Sex-Related Traits and Genes

Guest Editors: Alberto Civetta, José M. Eirín-López, Rob Kulathinal, and Jeremy L. Marshall



The Evolution of Sex-Related Traits and Genes

International Journal of Evolutionary Biology

The Evolution of Sex-Related Traits and Genes

Guest Editors: Alberto Civetta, José M. Eirín-López,
Rob Kulathinal, and Jeremy L. Marshall

Editorial Board

Contents

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 807218, 2 pages
doi:10.4061/2011/807218

Editorial

The Evolution of Sex-Related Traits and Genes

Alberto Civetta,[1] José M. Eirín-López,[2] Rob Kulathinal,[3] and Jeremy L. Marshall[4]

[1] *Department of Biology, University of Winnipeg, Winnipeg, MB, Canada R3B 2E9*
[2] *Department of Cell and Molecular Biology, University of A Coruña, E15071 Coruña, Spain*
[3] *Department of Biology, Temple University, Philadelphia, PA 19122, USA*
[4] *Department of Entomology, Kansas State University, Manhattan, KS 66506, USA*

Correspondence should be addressed to Alberto Civetta, a.civetta@uwinnipeg.ca

Received 28 June 2011; Accepted 28 June 2011

A growing number of studies are drawing attention to a consistent pattern of rapid evolution of sex-related traits and genes. Different hypotheses have been proposed to explain this rapid divergence. For traits and genes directly involved in reproduction, the pressures imposed by different sexes are likely to have driven rapid evolution, either by conflictive interactions, forms of female cryptic choice (female sex drive), or male differential use of traits and behavioral tactics (male sex drive). However, other factors such as subfunctionalization linked to gene duplications, birth and death processes, and pleiotropic effects, particularly those linked to immunity, can also drive fast change. The disruption of reproductive traits and gene interactions are of interest to those studying speciation because they can impede gene flow among diverging populations. Different studies have also shown that rapid changes and adaptive diversification between species can occur in reproductive developmental processes such as sex determination and gamete formation. Finally, high-throughput molecular technologies and bioinformatics approaches are allowing us to place patterns of gene evolution within the broad context of whole genome dynamics.

The ten papers published in this Issue tackle some of these fundamental questions on the evolution of sex-related traits and genes. Three opening review articles address problems relating to testing sexual selection at the molecular level, gene duplications and the evolution of sex bias patterns of gene expression, and the evolutionary origin of gametogenesis. A. Wong reviews recent studies in rodents and primates that have tested for associations between proxies of sexual selection and rates of evolution on male reproductive genes. He highlights how only a few studies have been able to establish significant associations. While this might cast some doubts on the role of sexual selection on the rapid evolution of male reproductive genes, he brings up some of the analytical problems linked to these prior studies and the need to incorporate multigene analysis and population genetics-based approaches. In their review, M. Gallach and collaborators address the problem faced by sexually reproducing species due to the need to generate two sexes with different patterns of tissue-specific gene expression and physiologies. They show that this conflict between sexes has led to a nonrandom distribution of sex-biased genes in different species and that gene duplication can help avoid sex-related genomic clashes leading to dosage compensation or meiotic sex chromosome inactivation as well as alleviate sexual conflict within species. Lastly, J. M. Eirín-López and J. Ausió provide an up-to-date insight into the evolutionary origins of the molecular mechanisms underlying sexual reproduction in metazoan animals. By bringing attention into the recent characterization of the DAZ family of reproductive proteins, the authors suggest that the wide conservation of a core reproductive machinery encoded by premeiotically expressed genes across Metazoa lends support to a common evolutionary origin for gametogenesis, rejecting the hypothesis that gametogenesis evolved multiple times in different lineages.

Seven research articles also address a variety of questions related to the evolution of sex-related traits and genes. Fox et al. reanalyzed data from seven human populations to show preferential paternal grandmothering behavior towards

granddaughters. They conclude that such behavior is likely driven by selection at selfish X-linked gene/s for alleles that increase the survival of grandmothers, because grandmothers who live longer provide longer care and benefits for their granddaughters. Thus, this work demonstrates that grandmaternal care accounts for the evolution of increased longevity in humans and is influenced by selection on the X-chromosome rather than by paternity uncertainty. Nelson and Neiman compared copulatory behavior in young and old asexual lineages of a freshwater snail. Contrary to an expected vestigialization of costly mating behaviors, the author found no differences in female copulation frequencies based on lineage age. Their results provide a rather cautionary warning regarding broad assumptions about the evolution of sex-related traits that are commonly linked to very complex aspects of the organism's biology. Two papers in this Issue deal with postcopulatory traits and genes. Greenspan and Clark tested differences in sperm competitive ability among strains of *Drosophila melanogaster* and found significant associations with polymorphisms residing on X-chromosome candidate genes. While these studies have traditionally been conducted by focusing on autosomal genes for seminal fluid proteins (SFP), their results suggests that such an approach has been biased and that we have been missing a large proportion of male reproductive variation linked to genes other than SFP. Sagga and Civetta surveyed divergence between species of Drosophila on noncompetitive postmating reproductive phenotypes and found evidence for a strong postmating prezygotic isolation (PPI) barrier among species of the virilis subgroup. They further established that PPI is due to reduced fertilization rates despite normal rates of sperm transfer and egg-laying. An inspection of sperm within female storage organs revealed low retention of heterospecific sperm leaving open the question of how female and female-male interactions might contribute to the exclusion of heterospecific sperm from storage. The paper by Jagadeeshan et al. also deals with the topic of species isolation and speciation in Drosophila. The authors analyzed 4,843 genes in 5 *D. melanogaster* subgroup species, seeking loci that exhibit accelerated rates of evolution associated with speciation. The finding that the majority of sex and reproduction-related genes are persistently subject to rapid evolution leads the authors to propose that selection, in combination with rapid demographic changes, makes a critical contribution to the elevated rates of evolution during speciation.

Finally, two papers look at larger macroevolutionary patterns of sex-bias evolution. Grassa and Kulathinal provide a functional and comparative genomic approach to understanding the role of sex-related genes in vertebrates. Using available ESTs from a range of tissues, fully assembled genomes, and well-curated gene models from five diverse vertebrate species, the authors characterize evolutionary rates in close to 5,000 reproductive and non-reproductive orthologs. Gonadal-specific genes that are lineage-specific show the greatest rate of evolutionary change. Surprisingly, an opposite trend emerges on functionally conserved orthologs. Kimball et al. catalog the presence and absence of sexually dimorphic characters involved in sexual signaling across the avian taxonomic clade, Phasianidae, which include pheasants and partridges. Six out of the nine dimorphic characters are visual cues while the remaining three relate to competitive ability. A variety of intersexual theories are described to explain the gains and losses of these sexually dimorphic characters in this lineage.

The evolution of sex-related traits and genes has clearly become a diverse and dynamic area of research and we hope that the articles included in this Issue will inspire, motivate, and challenge others to study the many questions that are emerging in this boundless field.

Acknowledgment

We would like to express our gratitude to all the authors and reviewers that contributed to the successful completion of this Issue.

Alberto Civetta
José M. Eirín-López
Rob Kulathinal
Jeremy L. Marshall

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 908735, 9 pages
doi:10.4061/2011/908735

Review Article

The Molecular Evolution of Animal Reproductive Tract Proteins: What Have We Learned from Mating-System Comparisons?

Alex Wong

Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, ON, Canada K1S 5B6

Correspondence should be addressed to Alex Wong, alex_wong@carleton.ca

Received 24 December 2010; Accepted 23 March 2011

Academic Editor: Alberto Civetta

Postcopulatory sexual selection is thought to drive the rapid evolution of reproductive tract genes in many animals. Recently, a number of studies have sought to test this hypothesis by examining the effects of mating system variation on the evolutionary rates of reproductive tract genes. Perhaps surprisingly, there is relatively little evidence that reproductive proteins evolve more rapidly in species subject to strong postcopulatory sexual selection. This emerging trend may suggest that other processes, such as host-pathogen interactions, are the main engines of rapid reproductive gene evolution. I suggest that such a conclusion is as yet unwarranted; instead, I propose that more rigorous analytical techniques, as well as multigene and population-based approaches, are required for a full understanding of the consequences of mating system variation for the evolution of reproductive tract genes.

1. Introduction

Over the last two decades, it has become widely appreciated that genes expressed in reproductive tissues, particularly in males, are among the most rapidly evolving genes in the genomes of many organisms. The phenomenon of rapid reproductive tract gene evolution is phylogenetically widespread, having been documented in vertebrates (e.g., [1–4]), invertebrates (e.g., [5, 6]), plants (e.g., [7, 8]), and fungi (e.g., [9]). As such, there is substantial interest in understanding the extent, causes, and consequences of the rapid evolution of reproductive genes.

A number of mechanisms have been proposed to explain the rapid evolution of reproductive tract genes, and the relative contribution of each of these mechanisms likely varies in different taxa. In plants, for example, high within-species diversity of interacting pollen and stigma proteins is vital for the avoidance of self-fertilization [10]. In other species, sexual conflict is thought to play a role in the rapid coevolution of interacting sperm and egg coat proteins, with selection on males favouring sperm that can rapidly penetrate the egg coat and selection on females to avoid polyspermy ([11], but see [12, 13]). Immune processes may also contribute to the rapid evolution of some reproductive proteins: in internally fertilizing animals in particular, copulation is likely to introduce potential pathogens into the female reproductive tract. An arms race between pathogens and hosts could, therefore, underlie selection on female- or male-derived antimicrobial proteins [13–16].

In this perspective, I will focus on the contribution of postcopulatory sexual selection (PCSS) to rapid reproductive protein evolution. Any form of competition amongst ejaculates or biased usage of sperm by females following the act of mating constitutes PCSS, and many such mechanisms have been described (e.g., [17, 18]). A particular form of PCSS, sperm competition, is particularly relevant in the context of the evolution of reproductive tract proteins, and it will be our chief concern in this paper. Sperm competition can occur in internally fertilizing animals if sperm from more than one male are simultaneously present in the reproductive tract of a single female [17, 19, 20]. In such a scenario, competition for fertilization opportunities exerts strong selection on traits that increase a male's paternity at the expense of his opponents. Moreover, females may gain a fitness advantage by biasing their sperm usage towards males of higher quality, more compatible males, males with "sexy" sperm, or males

that are more genetically distant (e.g., [21–24]). Such female choice may, in turn, drive sexual conflict, since such female adaptations will decrease the paternity share of some males, resulting in rapid evolution of both male and female traits.

PCSS, therefore, potentially acts on a wide variety of proteins that mediate male-male or male-female postcopulatory interactions. In animals, PCSS may act on sperm and egg surface proteins, as well as seminal fluid proteins and proteins present in the somatic portion of the female reproductive tract. Such proteins have been well studied from both the functional and evolutionary perspectives in a number of model organisms (e.g., [25, 26]). Sperm and egg coat proteins can play important roles in sperm-egg recognition and fusion, and thus may be subject to sexual conflict as described above. Moreover, seminal fluid proteins have been implicated in diverse processes in different animals, including sperm capacitation, sperm storage, and the control of postmating female behaviours, all of which could be subject to PCSS.

PCSS is often cited as a potential, and probable, cause of rapid reproductive protein evolution. Nonetheless, relatively few studies are able to distinguish the action of sexual selection from other processes, such as a host-pathogen arms race, that could also result in the rapid evolution of proteins expressed in reproductive tissues. Recently, a number of researchers have sought to clarify the effects of PCSS by comparing rates of reproductive tract protein evolution in taxa with different mating systems. Here, I critically review the methods used in these efforts, summarize their broad conclusions, and suggest several approaches for making further progress.

2. Rapid Evolution of Reproductive Tract Proteins

Before discussing comparative approaches to reproductive protein evolution, it will be useful to briefly review some of the methods and evidence behind the claim that reproductive tract proteins tend to evolve more rapidly than proteins expressed in other tissues. Most studies have focused on the DNA sequence evolution of protein-coding genes although rapid evolution and/or positive selection has also been documented at the level of protein diversity [27–29] and gene expression [30–32]. Here, I will focus on studies that compare DNA sequences between two or more species, although I will touch on some analyses of within-species sequence variation later on.

Typically, the rate of protein evolution is summarized by the parameter ω, which represents the ratio of dN (the rate of amino-acid changing nucleotide substitutions) to dS (the rate of silent nucleotide substitution). Under the assumption that silent substitutions do not affect fitness, the value of ω reflects the type of selection acting on a gene (or part of a gene). For most genes, average ω across the entire coding region is less than 1, indicating that most amino-acid changes are deleterious and are thus removed by purifying selection. An $\omega = 1$ indicates that amino-acid changes are neutral, while $\omega > 1$ is expected under diversifying (or positive) selection—that is, new amino acids tend to be favoured.

Over the past two decades, many studies have documented evidence for positive selection on selected reproductive tract genes in a variety of species (reviewed in [11, 26, 33–36]). More recently, several papers have shown that large sets of genes expressed in the male reproductive tract have a higher average ω than genes expressed in other tissues. In *Drosophila*, for example, Haerty and colleagues [5] showed that genes encoding testis- or seminal fluid-specific proteins evolve significantly more rapidly on average than do other classes of gene. Similarly, ω is higher on average for male reproductive tract-specific genes in comparison to other classes of genes in rodents [14, 37] and primates [3]. Interestingly, analogous patterns have generally not been found for genes expressed specifically in the female reproductive tract, with ω for egg, ovary, and/or somatic reproductive tract genes about the same as the genome-wide average in flies and primates [3, 5].

A simple elevation of ω amongst male reproductive tract genes does not necessarily imply the action of diversifying selection; however, this pattern could also be produced by a relaxation of constraint on reproductive tract genes. In order to infer the action of positive selection, it is important to show that $\omega > 1$. Averaged across an entire gene, this criterion is excessively strict, since positive selection likely acts on only a subset of sites. Thus, a number of methods have been developed that allow variation in ω within a gene, and that thereby allow the inference of positive selection (as $\omega > 1$) on a subset of codons. The most popular of these methods is Ziheng Yang's PAML (phylogenetic analysis by maximum likelihood), and a number of other methods build on the models developed by Yang and colleagues [38, 39]. The application of PAML to reproductive tract genes in *Drosophila* has confirmed that positive selection is at least partly responsible for their elevated ω [5]. In the case of the human/chimpanzee comparison, however, low sequence divergence substantially reduces power to detect positive selection [3], and unavailability of sequence data from multiple species has limited efforts to compare levels of positive selection between reproductive tract genes and nonreproductive tract genes in rodents [14]. More broadly, many studies have documented evidence for positive selection on reproductive tract genes in a wide variety of taxa (see [26, 33] for reviews). However, such studies often select rapidly diverging genes for subsequent PAML analysis (e.g., [37, 40]), and it is, therefore, not possible to contrast the relative impact of positive selection on reproductive-tract and nonreproductive tract genes in these cases.

It is important to note that the methods described in this section assume that rates of evolution are invariant across all lineages considered—that is, there is no variation in ω across the phylogeny of the species studied (Figure 1(a)). While a phylogeny-wide elevation of ω for reproductive tract genes is consistent with the hypothesis that PCSS acts on these genes, these analyses do not exclude other selective mechanisms. Other approaches, such as comparisons of ω between species with different mating systems, will help to clarify the effects of PCSS on rates of molecular evolution.

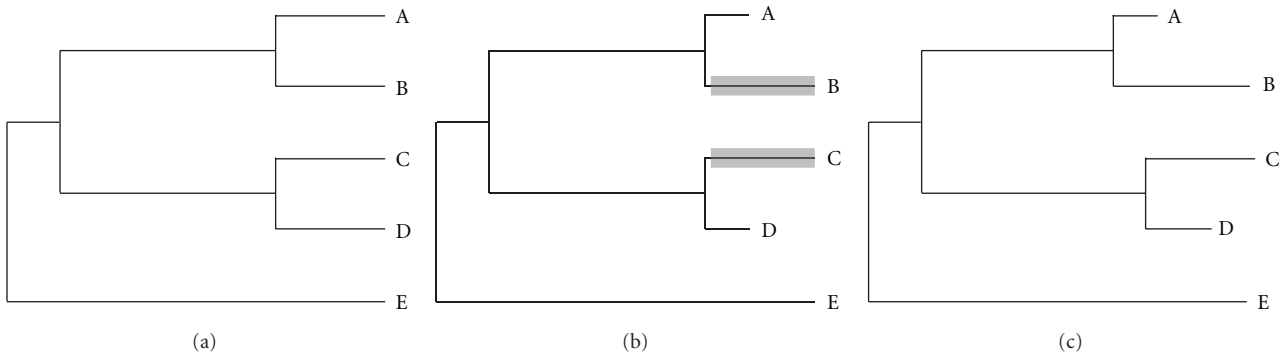


Figure 1: Three types of rate variation. In each panel, branch lengths are proportional to the rate of protein evolution. (a) Amongst the focal species A, B, C, and D, there is no variation in the rate of evolution. (b) Two rate classes. A and D share a rate, and B and C share a faster rate. In the current context, B and C represent polyandrous species, while A and D are monandrous. (c) Species-specific rates. Different rates of evolution are estimated for every lineage.

3. Mating System Variation Explains Morphological Variation in Reproductive Characters

Attempts to associate rates of molecular evolution with variation in the strength of PCSS take their cue from a long history of studies on the relationship between morphological characters and sexual selection. Starting in the late 1970s, a number of studies have examined the relationship between testis size and mating system, primarily in mammals (e.g., [41–49]) but also extending to other animals [50–53]. Many animal taxa show tremendous variation in testis size relative to body size, and it has been hypothesized that large testes are an adaptation to sperm competition. The logic is straightforward: for multiply mating species, such as chimpanzees, it should be in a male's interest to produce many sperm, either to have a numerical advantage in a single bout of sperm competition, or to avoid sperm depletion over the course of multiple matings. By contrast, for species in which females mate with only one male per estrous cycle, for example, gorillas, males need only produce sufficient sperm to ensure fertilization.

Early studies on the relationship between testis size and mating system used a nonphylogenetic approach [46]. Here, log(testes size) is regressed on log(body size), and residuals from this plot are used as a measure of relative testes size, with relatively large testes falling above the regression line and hence having positive residuals, and small testes falling below the regression line with negative residuals (Figure 2(a)). In the primate data presented in Figure 2, for example, it is evident that species in which females mate multiply (circles on Figure 2) have relatively large testes, whereas monandrous species (diamonds on Figure 2) tend to have small testes. However, as Felsenstein [54] pointed out, data from related species are not statistically independent such that regression can yield spurious relationships. Thus, later studies on the relationship between mating system and testis size have used explicitly phylogenetic methods and have generally shown that the association between multiple mating and large testes is, in fact, robust to phylogenetic effects [45, 47, 49, 50]. This relationship can be remarkably strong, with one study in rodents showing that variation in multiple paternity explains between 30% and 50% of variation in testes size [55].

The inferred intensity of sperm competition is associated with variation in a myriad of morphological traits beyond testis size. To give just a few examples, divergence in genital morphology is associated with the intensity of sperm competition in many animals [57], and in butterflies female, remating frequency is correlated with both testis size and sperm length [51]. In bats, testis size covaries negatively with brain size [58], probably due to tradeoffs in investment between these two energetically expensive organs. In addition, the presence and/or size of male accessory organs, which produce nonsperm components of the seminal fluid, are associated with sperm competition levels in gobies and in rodents [55, 59]. Similarly, the degree of seminal coagulation correlates with mating system in primates [60]. These results thus suggest that PCSS has evolutionary consequences for a wide range of phenotypic traits.

4. The Comparative Approach in Molecular Evolution

Recently, a number of studies have sought to test the hypothesis that rates of reproductive tract protein evolution, and/or levels of positive selection on these proteins, should increase with the intensity of PCSS or with one of its proxies, such as relative testes mass (see also [61] for an association between mating system and rates of evolution of immunity genes). These attempts represent an important step forward in trying to delineate the causes of the rapid evolution of reproductive tract proteins, rather than merely describe the phenomenon.

Two different methodological approaches have been used to test for a relationship between mating system and rates of molecular evolution. The first method uses categorical descriptions of the mating systems of the species of interest; the specific labels vary somewhat by taxon, but the important distinction is between species with a greater or lesser intensity of sperm competition. Species with low levels of sperm competition include those wherein females mate once

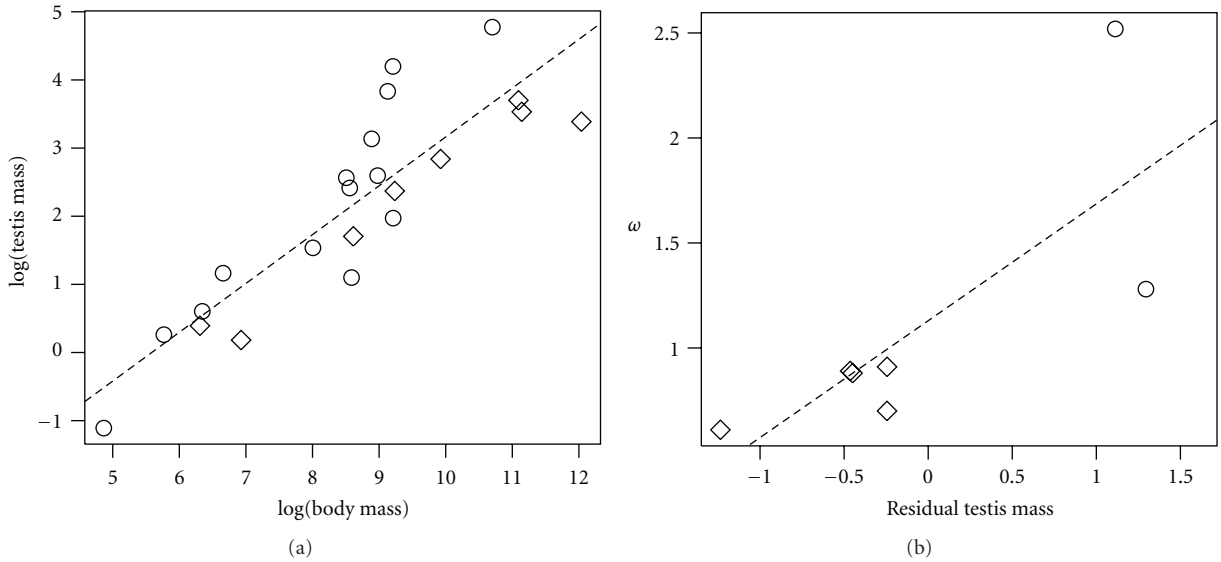


Figure 2: Consequences of sexual selection. (a) Regression of log(testis mass) on log(body mass). Polyandrous species (open circles) tend to have large testes for their body size, and hence fall above the regression line, while monandrous species (diamonds) tend to have relatively small testis. Data are from [47]. (b) Regression of ω on residual testis mass for *SEMG2*, showing that the rate of protein evolution increases with relative testis size. Estimates of ω are from [56], and testis mass residuals were calculated from panel (a).

per lifetime, species where females mate multiply but where sperm from different males do not overlap in time, and species where females mate with only one male per estrous cycle. Given a distinction between species with high or low sperm competition, the terminal branches of a phylogeny can be labelled by mating system (Figure 1(b)).

The phylogeny, labelled by mating system, can then be used to compare two models of sequence evolution [62]. Under the first model, which serves as the null hypothesis, positive selection (i.e., a subset of codons with $\omega > 1$) is not allowed for any lineage in the phylogeny. Under the second model, the alternative hypothesis, positive selection is allowed on a subset of codons in polyandrous lineages but not in monandrous lineages. Since the null hypothesis is a special case of the alternative, the two models can be compared via a likelihood ratio test. If the data fit the alternative hypothesis better than they do the null hypothesis, then this serves as evidence of an association between mating system and positive selection for the gene under study. This test is implemented in PAML, and corresponds to a comparison between models M2 and M0 described by Zhang and coworkers [62]. It should be noted that this test, as well as the lineage-invariant methods described above, will only detect selection consistently acting on the same codons and is thus likely to miss selection on different codons in different lineages.

This first approach to associating mating system and rates of molecular evolution, which I will refer to as the "discrete" comparative method, has been used by Ramm et al. [1] to study the evolution of seven rodent and two primate reproductive proteins, and by Finn and Civetta [63] to study the evolution of 13 sperm proteins in mammals. Using methods that do not allow variation in ω between lineages, Finn and Civetta found evidence for positive selection on 12 out of 13 genes encoding male expressed ADAM proteins, with positive selection on all 7 sperm surface ADAM proteins. In applying the discrete comparative method, however, only 1 of these 12 genes showed evidence for positive selection specifically along polyandrous lineages, with the 13th sperm-bound ADAM also showing evidence for lineage-specific selection in primates. Thus, the phylogeny-wide signal of positive selection does not appear to be solely, or even chiefly, due to selection in polyandrous species. Similarly, using lineage-invariant methods, Ramm et al. [1] found evidence for positive selection on 5 out of 7 rodent reproductive genes. Using the discrete comparative method, they found evidence for lineage-specific selection on only one of these genes, *Svs2*, which encodes a copulatory plug protein. Further comparative analyses conducted by Ramm et al. on 2 primate genes will be discussed below.

The second approach for testing for a relationship between mating system and rates of molecular evolution uses continuous measures of the intensity of sperm competition. The most often used metric of PCSS is residual testes mass: given the robustness of the finding that polyandrous species tend to have large testes, residual testes mass should be a good proxy for the intensity of sperm competition. Recently, Herlyn and Zischler [64] and others have proposed that sexual size dimorphism can also be used as a continuous proxy for sperm competition, with less dimorphism indicating stronger sperm competition. The reasoning behind this claim is that in highly dimorphic species, relatively large males will be able to enforce monandry. Finally, estimates of mating frequency—for example, through behavioural observation in primates (e.g., [56]), spermatophore counts in butterflies (e.g., [51]), or genetic analysis of offspring (e.g., [65–67])—can be used as more direct measures of the intensity of sperm competition.

In applying the continuous comparative method (as I will call it), ω is estimated separately for every branch of the phylogenetic tree under consideration (Figure 1(c)). Here, the ω estimate for each branch represents an average across all codons and so is unlikely to significantly exceed 1. Branch-specific estimates of ω are regressed on the continuous measure of sperm competition, and a significant positive relationship (negative in the case of size dimorphism) is interpreted as evidence for an effect of sperm competition on rates of molecular evolution.

Dorus et al. [56] presented the first use of the continuous comparative method on a reproductive protein, in which they studied the molecular evolution of the primate gene *SEMG2*, which encodes a copulatory plug protein. Lineage-invariant methods find robust evidence for positive selection on this gene, and Dorus et al. found a strong relationship between ω and residual testis size (data reanalyzed in Figure 2(b)), number of mates, and an ordinate rating of semen coagulation.

Several other studies have adopted a similar approach. Herlyn and Zischler [64] found a significantly negative relationship between branch specific ω and sexual size dimorphism in primates for the sperm ligand zonadhesin (*ZAN*), and Martin-Coello and colleagues [28] obtained a significant association between testis size and promoter divergence for the sperm-specific histone protamine 2. Hurle et al. [68], by contrast, were unable to detect a significant relationship between ω and mating system for any of six reproductive genes in primates (including *SEMG2*) despite finding evidence for positive selection on 5 of these genes using lineage-invariant methods. The case for an association between mating system and ω for *SEMG2* is nonetheless fairly strong: Ramm and colleagues [1] have found evidence for such a relationship using the discrete method for *SEMG2* although not for *SEMG1* (which encodes a second copulatory plug protein).

Table 1 summarizes the methods and results of studies investigating associations between mating system and rates of molecular evolution, using either the discrete or the continuous comparative methods. Notwithstanding heterogeneity in methods, sample sizes, and taxa, it is striking that very few reproductive genes show evidence for such an association: only 6 genes show evidence for an effect of mating system on rates of protein evolution even though 24 show evidence for positive selection using lineage-invariant methods. This unexpected lack of evidence for an association may indicate that processes other than PCSS underlie the rapid evolution of most reproductive tract proteins, contrary to popular wisdom. However, such a conclusion is likely premature, and I suggest in particular that methods ought to be improved in several ways.

5. Comparative Methods in Molecular Evolution: Methodological Issues

The statistical methods used to test for associations between mating system and rates of molecular evolution have not been entirely appropriate. The fundamental problem is that estimates of ω on the one hand and morphological or life-history characteristics on the other are different types of data. Phenotypic trait values, measured at the leaves of a phylogenetic tree, are instantaneous measures—that is, they indicate the state of the trait at each leaf *now*. Estimates of ω, by contrast, are typically an average for the entire branch. Due to this difference in data type, until very recently, there have not been any appropriate methods for detecting covariance between molecular evolutionary rates and phenotypic traits. Felsenstein's method of independent contrasts, for example, is appropriate for instantaneous data, since it explicitly models trait change along the branches of a phylogenetic tree. Attempts to use estimates of ω in an independent contrast framework, thereby conflate an average for an instantaneous measure. Even the discrete comparative method described above runs into problems here, since it assumes that the phenotypic state observed at a leaf has been constant for the entire branch, thus conflating an instantaneous measure for an average. It is not obvious to what extent these methodological concerns will lead to bad inferences, but ideally, we should aim to compare similar types of data in a single statistical framework.

Recently, several model-based methods have been developed for detecting associations between rates of substitution and morphological/life-history variation [69–71]. O'Connor and Mundy [69] and Mayrose and Otto [71] have both developed maximum-likelihood frameworks that simultaneously model molecular evolution and character state evolution. Both formulations use model comparisons to test for a coupling between rates of sequence evolution and character state evolution, in which the character of interest takes on a binary (1/0) value. Under the null model, no such association is present—that is, sequence evolution is independent of the phenotype under consideration. Under the alternative model, rates of molecular evolution vary systematically with the phenotypic character state. A likelihood ratio test can then be used to compare the two models, with a rejection of the null model suggesting an association between phenotypic variation and the rate of molecular evolution. Additional procedures are introduced to handle lineage-specific heterogeneity in rates of sequence evolution that is not associated with the trait of interest (an additional null model in the case of O'Connor and Mundy and parametric bootstrapping in the Mayrose and Otto method).

Currently, both of these methods implement nucleotide substitution models, and hence do not distinguish between synonymous and nonsynonymous substitutions (codon models are used for the latter purpose). As such, associations between molecular evolutionary rates and a phenotypic character are not necessarily specific to changes in protein sequence. Nonetheless, it should be possible to modify either method to use codon models.

Beyond the details of the model implementations, these two methods differ with respect to the formulation of the alternative model. Under the O'Connor and Mundy model, a subset of sites is allowed to evolve in association with the phenotypic trait of interest, with a background substitution rate that is independent of phenotype. By contrast, in the Mayrose and Otto model all sites are assumed to evolve in

Table 1: Associations between mating system variation and rates of molecular evolution.

Study	Taxon	No. genes studied	Method[1]	No. genes showing association	No. genes showing positive selection[2]
Dorus et al. 2004 [56]	Primates	1	C	1	1
Herlyn and Zischler 2007 [64]	Primates	1	C	1	1
Hurle et al. 2007 [68]	Primates	6	C	0	5
Ramm et al. 2008 [1]	Rodents	7	D	1	5
	Primates	2	D	1	2
Martin-Coello et al. 2009 [28]	Rodents	2	C	1	1
Finn and Civetta 2010 [63]	Primates	13	D	2	12
Total (Unique genes)[3]	—	29	—	6	24

[1] D = discrete comparative method; C = continuous comparative method.
[2] Inference of positive selection assuming no variation between lineages.
[3] Two genes (*SEMG1* and *SEMG2*) have been examined in multiple studies—*SEMG1* twice [1, 68] and *SEMG2* three times [1, 56, 68]. The "total (unique genes)" row reflects this replication and is therefore not a sum of the preceding rows.

a phenotype-dependent manner. Thus, the O'Connor and Mundy approach may be more suitable when selection drives the evolution of some, but not all, sites in a lineage specific manner, whereas the Mayrose and Otto method is best suited for detecting lineage-specific mutational effects that influence all sites.

Lartillot and Poujol [70] have developed a Bayesian method, CoEvol, for detecting associations between rates of molecular evolution and rates of change in phenotypic/life history traits. Unlike the O'Connor and Mundy [69] and Mayrose and Otto [71] methods, CoEvol considers continuous phenotypic characters, and implements a codon model of sequence evolution rather than a nucleotide model. CoEvol estimates a covariance matrix for the rates of change of dS, ω, and one or more morphological/life history characters. A high posterior probability associated with covariance between the rate of change in dS or ω on the one hand, and the rate of change of a phenotypic character on the other, serves as evidence for coupling between the molecular and phenotypic processes. Importantly, separate estimates of covariance for dS and ω should allow one to distinguish mutational effects (via covariance with dS) from selective effects (via covariance with ω).

The phylogenetic methods just described have each been applied to different datasets: Mayrose and Otto [71], for example, detected an effect of habitat salinity on the rate of molecular evolution in daphniids, and Lartillot and Poujol [70] detected negative associations between dS and mass and longevity in mammals. Only O'Connor and Mundy [69] have applied their method to reproductive proteins, specifically the primate seminal proteins *SEMG1*, *SEMG2*, and Zonadhesin (*ZAN*) as well as two genes not expected to be subject to sexual selection, *PI3* and *CYTB*. Using a binary mating system classification (multiply mating versus single mating), they found a significant association between substitution rate and phenotypic state for *SEMG2*, but not for the other four loci. Thus, their analyses are consistent with previous studies for *SEMG2*, but not for *ZAN* (recall that Herlyn and Zischler [64] reported a negative association between sexual size dimorphism and ω for the *ZAN* locus). The discrepancy in the case of *ZAN* could reflect differences in the way PCSS is measured (binary classification versus sexual size dimorphism), lack of power in the the O'Connor and Mundy method, or may suggest that the previously reported correlation between ω and dimorphism is a methodological artifact. Further studies will be required to distinguish these possibilities.

These new statistical approaches should prove to be powerful tools for investigating the consequences of mating system variation for substitution rates, since they will lend more confidence to inferences concerning associations (or lack thereof) between mating system and rates of molecular evolution. It will be particularly important to investigate the statistical properties of these methods (power, false positive rates under different conditions, etc.) in order to better understand when and how each should be used.

6. Inconsistency in the Targets of Selection

To date, most studies on the relationship between mating system and the molecular evolution of reproductive proteins have focused on one or a few genes. Thus, in order to detect an association, PCSS must act on the same gene in most lineages. While this may be true in some cases (e.g., *SEMG2*), it is not a foregone conclusion that the same genes will be subject to PCSS in different species. Indeed, the rapid turnover of reproductive tract genes in some species (e.g., *Drosophila*—[5, 72]), as well as observations of lineage specific positive selection on reproductive proteins [37, 63], suggests that there may be variation in the targets of selection between lineages.

If PCSS operates on different loci in different species, then it may be fruitful to compare average rates of evolution across many reproductive genes between taxa with different mating systems. Indeed, a handful of studies have done this and, in general, confirm the prediction that average rates of evolution are higher in taxa with higher levels of PCSS. For example, remating rates in the *repleta* group of the fruitfly genus *Drosophila* are substantially higher than remating rates in the *melanogaster* group (e.g., [73, 74]) such

that the intensity of PCSS is presumably higher for *repleta* group species. A number of recent studies have reported that average ω and rates of gene duplication are higher for both male and female reproductive tract genes in the *repleta* group than in the *melanogaster* group [75–78], as predicted if PCSS drives reproductive tract gene evolution. Similarly, I recently reported that rates of evolution are higher for testis-specific proteins in highly polyandrous chimpanzees than in humans, where historical levels of polyandry are thought to be more modest [3]. Work in additional taxa will be required to confirm these patterns (e.g., [79]), but these first studies suggest that we should not necessarily expect a strong signal of PCSS when looking at single genes alone.

In addition to this "many-gene" approach, molecular population genetic studies of candidate genes, rather than between-species comparisons, should provide complementary data on the effects of mating system on the evolution of genes encoding reproductive proteins. Here, candidate genes would be sequenced in multiple individuals from one or more species with strong PCSS, as well as in multiple individuals from one or more species with no or weak PCSS. A wide variety of statistical tests are available for inferring recent selection from such data [80, 81], and application of these tests would allow gene-by-gene tests of the prediction that positive selection should act in multiply mating species but not in monandrous species. The advantage of this approach is that selection need only have acted in a single species to detect a signal, whereas selection must be fairly consistent across a phylogeny to detect positive selection in between-species comparisons. Population genetic surveys in primates have proved informative with respect to semenogelin (*SEMG1* and *SEMG2*) evolution, for example. In chimpanzees, which are highly polyandrous, *SEMG1* shows strong evidence for recent positive selection [82] and an increase in protein length owing to a repeat expansion [83]. In gorillas, by contrast, both *SEMG1* and *SEMG2* carry multiple loss-of-function mutations, consistent with a loss of constraint due to the highly monandrous mating-system of this species [82, 83].

7. Summary and Conclusions

Despite widespread evidence for the rapid evolution of, and positive selection on, genes encoding reproductive tract proteins, comparative studies have had relatively little success in associating rates of protein evolution with the strength of PCSS. Only a handful of individual genes show robust evidence for accelerated evolution in polyandrous lineages in comparison to monandrous ones. This lack of evidence for a relationship between mating system and rates of molecular evolution may indicate that processes other than PCSS, such as immune interactions, drive the evolution of many reproductive tract proteins. I have suggested, however, that new analytical methods will add rigour to attempts to delineate the causes of rapid reproductive tract protein evolution. Moreover, if selection acts on different genes in different lineages, then a combination of multigene comparative studies, as well as population genetic studies, should prove useful.

Acknowledgments

The author thanks three anonymous reviewers, and the editors of this special issue for helpful comments on this paper. The author acknowledges funding support from NSERC, CIHR, and the Banting Fellowship program.

References

[1] S. A. Ramm, P. L. Oliver, C. P. Ponting, P. Stockley, and R. D. Emes, "Sexual selection and the adaptive evolution of mammalian ejaculate proteins," *Molecular Biology and Evolution*, vol. 25, no. 1, pp. 207–219, 2008.

[2] L. M. Turner and H. E. Hoekstra, "Adaptive evolution of fertilization proteins within a genus: variation in ZP2 and ZP3 in deer mice *(Peromyscus)*," *Molecular Biology and Evolution*, vol. 23, no. 9, pp. 1656–1669, 2006.

[3] A. Wong, "Testing the effects of mating system variation on rates of molecular evolution in primates," *Evolution*, vol. 64, no. 9, pp. 2779–2785, 2010.

[4] N. L. Clark and W. J. Swanson, "Pervasive adaptive evolution in primate seminal proteins," *PLoS Genetics*, vol. 1, no. 3, article e35, 2005.

[5] W. Haerty, S. Jagadeeshan, R. J. Kulathinal et al., "Evolution in the fast lane: rapidly evolving sex-related genes in *Drosophila*," *Genetics*, vol. 177, no. 3, pp. 1321–1335, 2007.

[6] N. L. Clark, J. Gasper, M. Sekino, S. A. Springer, C. F. Aquadro, and W. J. Swanson, "Coevolution of interacting fertilization proteins," *PLoS Genetics*, vol. 5, no. 7, Article ID e1000570, 2009.

[7] N. Takebayashi, P. B. Brewer, ED. Newbigin, and M. K. Uyenoyama, "Patterns of variation within self-incompatibility loci," *Molecular Biology and Evolution*, vol. 20, no. 11, pp. 1778–1794, 2003.

[8] K. Sato, T. Nishio, R. Kimura et al., "Coevolution of the S-locus genes SRK, SLG and SP11/SCR in *Brassica oleracea* and B. *rapa*," *Genetics*, vol. 162, no. 2, pp. 931–940, 2002.

[9] L. Wik, M. Karlsson, and H. Johannesson, "The evolutionary trajectory of the mating-type (mat) genes in *Neurospora* relates to reproductive behavior of taxa," *BMC Evolutionary Biology*, vol. 8, no. 1, p. 109, 2008.

[10] S. Takayama and A. Isogai, "Self-incompatibility in plants," *Annual Review of Plant Biology*, vol. 56, pp. 467–489, 2005.

[11] W. J. Swanson and V. D. Vacquier, "The rapid evolution of reproductive proteins," *Nature Reviews Genetics*, vol. 3, no. 2, pp. 137–144, 2002.

[12] J. D. Calkins, D. El-Hinn, and W. J. Swanson, "Adaptive evolution in an avian reproductive protein: ZP3," *Journal of Molecular Evolution*, vol. 65, no. 5, pp. 555–563, 2007.

[13] S. Berlin, L. Qu, and H. Ellegren, "Adaptive evolution of gamete-recognition proteins in birds," *Journal of Molecular Evolution*, vol. 67, no. 5, pp. 488–496, 2008.

[14] M. D. Dean, N. L. Clark, G. D. Findlay et al., "Proteomics and comparative genomic investigations reveal heterogeneity in evolutionary rate of male reproductive proteins in mice (*Mus domesticus*)," *Molecular Biology and Evolution*, vol. 26, no. 8, pp. 1733–1743, 2009.

[15] J. M. Good and M. W. Nachman, "Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis," *Molecular Biology and Evolution*, vol. 22, no. 4, pp. 1044–1052, 2005.

[16] M. K. N. Lawniczak, A. I. Barnes, J. R. Linklater, J. M. Boone, S. Wigby, and T. Chapman, "Mating and immunity in

invertebrates," *Trends in Ecology and Evolution*, vol. 22, no. 1, pp. 48–55, 2007.

[17] G. A. Parker, "Sperm competition and its evolutionary consequences in the insects," *Biological Reviews*, vol. 45, no. 4, pp. 525–567, 1970.

[18] W. G. Eberhard, *Female Control: Sexual Selection by Cryptic Female Choice*, Princeton University Press, Princeton, NJ, USA, 1996.

[19] G. A. Parker, "Sexual conflict over mating and fertilization: an overview," *Philosophical Transactions of the Royal Society of London: Series B*, vol. 361, no. 1466, pp. 235–259, 2006.

[20] T. R. Birkhead, "How stupid not to have thought of that: post-copulatory sexual selection," *Journal of Zoology*, vol. 281, no. 2, pp. 78–93, 2010.

[21] L. Keller and H. K. Reeve, "Why do females mate with multiple males—the sexually selected sperm hypothesis," *Advances in the Study of Behavior*, vol. 24, pp. 291–315, 1995.

[22] H. Kokko, M. D. Jennions, and R. Brooks, "Unifying and testing models of sexual selection," *Annual Review of Ecology, Evolution, and Systematics*, vol. 37, no. 1, pp. 43–66, 2006.

[23] M. Andersson and L. W. Simmons, "Sexual selection and mate choice," *Trends in Ecology and Evolution*, vol. 21, no. 6, pp. 296–302, 2006.

[24] L. W. Simmons, "The evolution of polyandry: sperm competition, sperm selection, and offspring viability," *Annual Review of Ecology, Evolution, and Systematics*, vol. 36, no. 1, pp. 125–146, 2005.

[25] F. W. Avila, L. K. Sirot, B. A. LaFlamme, C. D. Rubinstein, and M. F. Wolfner, "Insect seminal fluid proteins: identification and function," *Annual Review of Entomology*, vol. 56, pp. 21–40, 2011.

[26] T. M. Panhuis, N. L. Clark, and S. W. J., "Rapid evolution of reproductive proteins in abalone and *Drosophila*," *Philosophical Transactions of the Royal Society of London: Series B*, vol. 361, no. 1466, pp. 261–268, 2006.

[27] A. Civetta and R. S. Singh, "High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in *Drosophila melanogaster* and *Drosophila virilis* species," *Journal of Molecular Evolution*, vol. 41, no. 6, pp. 1085–1095, 1995.

[28] J. Martin-Coello, H. Dopazo, L. Arbiza, J. Ausió, E. R. S. Roldan, and M. Gomendio, "Sexual selection drives weak positive selection in protamine genes and high promoter divergence, enhancing sperm competitiveness," *Proceedings of the Royal Society B*, vol. 276, no. 1666, pp. 2427–2436, 2009.

[29] S. A. Ramm, L. McDonald, J. L. Hurst, R. J. Beynon, and P. Stockley, "Comparative proteomics reveals evidence for evolutionary diversification of rodent seminal fluid and its functional significance in sperm competition," *Molecular Biology and Evolution*, vol. 26, no. 1, pp. 189–198, 2009.

[30] C. Voolstra, D. Tautz, P. Farbrother, L. Eichinger, and B. Harr, "Contrasting evolution of expression differences in the testis between species and subspecies of the house mouse," *Genome Research*, vol. 17, no. 1, pp. 42–49, 2007.

[31] S. V. Nuzhdin, M. L. Wayne, K. L. Harmon, and L. M. McIntyre, "Common pattern of evolution of gene expression level and protein sequence in *Drosophila*," *Molecular Biology and Evolution*, vol. 21, no. 7, pp. 1308–1317, 2004.

[32] C. D. Meiklejohn, J. Parsch, J. M. Ranz, and D. L. Hartl, "Rapid evolution of male-biased gene expression in *Drosophila*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 17, pp. 9894–9899, 2003.

[33] N. L. Clark, J. E. Aagaard, and W. J. Swanson, "Evolution of reproductive proteins from animals and plants," *Reproduction*, vol. 131, no. 1, pp. 11–22, 2006.

[34] T. Chapman, "The soup in my fly: evolution, form and function of seminal fluid proteins," *PLoS Biology*, vol. 6, no. 7, article e179, 2008.

[35] A. Civetta, "Shall we dance or shall we fight? using DNA sequence data to untangle controversies surrounding sexual selection," *Genome*, vol. 46, no. 6, pp. 925–929, 2003.

[36] L. M. Turner and H. E. Hoekstra, "Causes and consequences of the evolution of reproductive proteins," *International Journal of Developmental Biology*, vol. 52, no. 5-6, pp. 769–780, 2008.

[37] L. M. Turner, E. B. Chuong, and H. E. Hoekstra, "Comparative analysis of testis protein evolution in rodents," *Genetics*, vol. 179, no. 4, pp. 2075–2089, 2008.

[38] N. Goldman and Z. Yang, "A codon-based model of nucleotide substitution for protein-coding DNA sequences," *Molecular Biology and Evolution*, vol. 11, no. 5, pp. 725–736, 1994.

[39] Z. Yang, R. Nielsen, N. Goldman, and A. M. K. Pedersen, "Codon-substitution models for heterogeneous selection pressure at amino acid sites," *Genetics*, vol. 155, no. 1, pp. 431–449, 2000.

[40] W. J. Swanson, A. Wong, M. F. Wolfner, and C. F. Aquadro, "Evolutionary expressed sequence tag analysis of *Drosophila* female reproductive tracts identifies genes subjected to positive selection," *Genetics*, vol. 168, no. 3, pp. 1457–1465, 2004.

[41] R. V. Short, "Sexual selection and its component parts, somatic and genital selection, as illustrated by man and the great apes," *Advances in the Study of Behavior*, vol. 9, pp. 131–158, 1979.

[42] R. W. Rose, C. M. Nevison, and A. F. Dixson, "Testes weight, body weight and mating systems in marsupials and monotremes," *Journal of Zoology*, vol. 243, no. 3, pp. 523–531, 1997.

[43] A. P. Møller, "Ejaculate quality, testes size and sperm competition in primates," *Journal of Human Evolution*, vol. 17, no. 5, pp. 479–488, 1988.

[44] G. Kenagy and S. Trombulak, "Size and function of mammalian testes in relation to body size," *Journal of Mammalogy*, vol. 67, no. 1, pp. 1–22, 1986.

[45] D. J. Hosken, "Testes mass in megachiropteran bats varies in accordance with sperm competition theory," *Behavioral Ecology and Sociobiology*, vol. 44, no. 3, pp. 169–177, 1998.

[46] A. H. Harcourt, P. H. Harvey, S. G. Larson, and R. V. Short, "Testis weight, body weight and breeding system in primates," *Nature*, vol. 293, no. 5827, pp. 55–57, 1981.

[47] A. H. Harcourt, A. Purvis, and L. Liles, "Sperm competition: mating system, not breeding season, affects testes size of primates," *Functional Ecology*, vol. 9, no. 3, pp. 468–476, 1995.

[48] G. Arnqvist, "Comparative evidence for the evolution of genitalia by sexual selection," *Nature*, vol. 393, no. 6687, pp. 784–786, 1998.

[49] G. Iossa, C. D. Soulsbury, P. J. Baker, and S. Harris, "Sperm competition and the evolution of testes size in terrestrial mammalian carnivores," *Functional Ecology*, vol. 22, no. 4, pp. 655–662, 2008.

[50] P. Stockley, M. J. G. Gage, G. A. Parker, and A. P. Møller, "Sperm competition in fishes: the evolution of testis size and ejaculate characteristics," *The American Naturalist*, vol. 149, no. 5, pp. 933–954, 1997.

[51] M. J. G. Gage, "Associations between body size, mating pattern, testis size and sperm lengths across butterflies," *Proceedings of the Royal Society B*, vol. 258, no. 1353, pp. 247–254, 1994.

[52] S. B. Emerson, "Testis size variation in frogs: testing the alternatives," *Behavioral Ecology and Sociobiology*, vol. 41, no. 4, pp. 227–235, 1997.

[53] S. Calhim and T. R. Birkhead, "Testes size in birds: quality versus quantity—assumptions, errors, and estimates," *Behavioral Ecology*, vol. 18, no. 1, pp. 271–275, 2007.

[54] J. Felsenstein, "Phylogenies and the comparative method," *The American Naturalist*, vol. 125, no. 1, pp. 1–15, 1985.

[55] S. A. Ramm, G. A. Parker, and P. Stockley, "Sperm competition and the evolution of male reproductive anatomy in rodents," *Proceedings of the Royal Society B*, vol. 272, no. 1566, pp. 949–955, 2005.

[56] S. Dorus, P. D. Evans, G. J. Wyckoff, S. S. Choi, and B. T. Lahn, "Rate of molecular evolution of the seminal protein gene SEMG2 correlates with levels of female promiscuity," *Nature Genetics*, vol. 36, no. 12, pp. 1326–1329, 2004.

[57] D. J. Hosken and P. Stockley, "Sexual selection and genital evolution," *Trends in Ecology and Evolution*, vol. 19, no. 2, pp. 87–93, 2004.

[58] S. Pitnick, K. E. Jones, and G. S. Wilkinson, "Mating system and brain size in bats," *Proceedings of the Royal Society B*, vol. 273, no. 1587, pp. 719–724, 2006.

[59] C. Mazzoldi, C. W. Petersen, and M. B. Rasotto, "The influence of mating system on seminal vesicle variability among gobies (Teleostei, Gobiidae)," *Journal of Zoological Systematics and Evolutionary Research*, vol. 43, no. 4, pp. 307–314, 2005.

[60] A. F. Dixson and M. J. Anderson, "Sexual selection, seminal coagulation and copulatory plug formation in primates," *Folia Primatologica*, vol. 73, no. 2–3, pp. 63–69, 2002.

[61] G. Wlasiuk and M. W. Nachman, "Promiscuity and the rate of molecular evolution at primate immunity genes," *Evolution*, vol. 64, no. 8, pp. 2204–2220, 2010.

[62] J. Zhang, R. Nielsen, and Z. Yang, "Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level," *Molecular Biology and Evolution*, vol. 22, no. 12, pp. 2472–2479, 2005.

[63] S. Finn and A. Civetta, "Sexual selection and the molecular evolution of ADAM proteins," *Journal of Molecular Evolution*, vol. 71, no. 3, pp. 231–240, 2010.

[64] H. Herlyn and H. Zischler, "Sequence evolution of the sperm ligand zonadhesin correlates negatively with body weight dimorphism in primates," *Evolution*, vol. 61, no. 2, pp. 289–298, 2007.

[65] A. Bretman and T. Tregenza, "Measuring polyandry in wild populations: a case study using promiscuous crickets," *Molecular Ecology*, vol. 14, no. 7, pp. 2169–2179, 2005.

[66] M. D. Dean, K. G. Ardlie, and M. W. Nachman, "The frequency of multiple paternity suggests that sperm competition is common in house mice *(Mus domesticus)*," *Molecular Ecology*, vol. 15, no. 13, pp. 4141–4151, 2006.

[67] S. C. Griffith, I. P. F. Owens, and K. A. Thuman, "Extra pair paternity in birds: a review of interspecific variation and adaptive function," *Molecular Ecology*, vol. 11, no. 11, pp. 2195–2212, 2002.

[68] B. Hurle, W. Swanson, and E. D. Green, "Comparative sequence analyses reveal rapid and divergent evolutionary changes of the WFDC locus in the primate lineage," *Genome Research*, vol. 17, no. 3, pp. 276–286, 2007.

[69] T. D. O'Connor and N. I. Mundy, "Genotype-phenotype associations: substitution models to detect evolutionary associations between phenotypic variables and genotypic evolutionary rate," *Bioinformatics*, vol. 25, no. 12, pp. i94–i100, 2009.

[70] N. Lartillot and R. Poujol, "A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters," *Molecular Biology and Evolution*, vol. 28, no. 1, pp. 729–744, 2010.

[71] I. Mayrose and S. P. Otto, "A likelihood method for detecting trait-dependent shifts in the rate of molecular evolution," *Molecular Biology and Evolution*, vol. 28, no. 1, pp. 759–770, 2010.

[72] A. K. Holloway and D. J. Begun, "Molecular evolution and population genetics of duplicated accessory gland protein genes in *Drosophila*," *Molecular Biology and Evolution*, vol. 21, no. 9, pp. 1625–1628, 2004.

[73] T. Markow, "Evolution of *Drosophila* mating systems," *Evolutionary Biology*, vol. 29, pp. 73–106, 1996.

[74] T. A. Markow and P. M. O'Grady, "Evolutionary genetics of reproductive behavior in *Drosophila*: connecting the dots," *Annual Review of Genetics*, vol. 39, pp. 263–291, 2005.

[75] E. S. Kelleher, W. J. Swanson, and T. A. Markow, "Gene duplication and adaptive evolution of digestive proteases in *Drosophila arizonae* female reproductive tracts," *PLoS Genetics*, vol. 3, no. 8, article e148, 2007.

[76] F. C. Almeida and R. Desalle, "Orthology, function and evolution of accessory gland proteins in the *Drosophila repleta* group," *Genetics*, vol. 181, no. 1, pp. 235–245, 2009.

[77] B. J. Wagstaff and D. J. Begun, "Adaptive evolution of recently duplicated accessory gland protein genes in desert *Drosophila*," *Genetics*, vol. 177, no. 2, pp. 1023–1030, 2007.

[78] B. J. Wagstaff and D. J. Begun, "Molecular population genetics of accessory gland protein genes and testis-expressed genes in *Drosophila mojavensis* and *D. arizonae*," *Genetics*, vol. 171, no. 3, pp. 1083–1101, 2005.

[79] J. R. Walters and R. G. Harrison, "Combined EST and proteomic analysis identifies rapidly evolving seminal fluid proteins in *Heliconius* butterflies," *Molecular Biology and Evolution*, vol. 27, no. 9, pp. 2000–2013, 2010.

[80] R. Nielsen, "Molecular signatures of natural selection," *Annual Review of Genetics*, vol. 39, pp. 197–218, 2005.

[81] J. D. Jensen, A. Wong, and C. F. Aquadro, "Approaches for identifying targets of positive selection," *Trends in Genetics*, vol. 23, no. 11, pp. 568–577, 2007.

[82] S. B. Kingan, M. Tatar, and D. M. Rand, "Reduced polymorphism in the chimpanzee semen coagulating protein, semenogelin I," *Journal of Molecular Evolution*, vol. 57, no. 2, pp. 159–169, 2003.

[83] M. I. Jensen-Seaman and W. H. Li, "Evolution of the hominoid semenogelin genes, the major proteins of ejaculated semen," *Journal of Molecular Evolution*, vol. 57, no. 3, pp. 261–270, 2003.

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 989438, 20 pages
doi:10.4061/2011/989438

Review Article

Gene Duplication and the Genome Distribution of Sex-Biased Genes

Miguel Gallach, Susana Domingues, and Esther Betrán

Department of Biology, University of Texas at Arlington, P.O. Box 19498, Arlington, TX 76019, USA

Correspondence should be addressed to Esther Betrán, betran@uta.edu

Received 29 December 2010; Revised 26 March 2011; Accepted 5 June 2011

Academic Editor: Rob Kulathinal

In species that have two sexes, a single genome encodes two morphs, as each sex can be thought of as a distinct morph. This means that the same set of genes are differentially expressed in the different sexes. Many questions emanate from this statement. What proportion of genes contributes to sexual dimorphism? How do they contribute to sexual dimorphism? How is sex-biased expression achieved? Which sex and what tissues contribute the most to sex-biased expression? Do sex-biased genes have the same evolutionary patterns as nonbiased genes? We review the current data on sex-biased expression in species with heteromorphic sex chromosomes and comment on the most important hypotheses suggested to explain the origin, evolution, and distribution patterns of sex-biased genes. In this perspective we emphasize how gene duplication serves as an important molecular mechanism to resolve genomic clashes and genetic conflicts by generating sex-biased genes, often sex-specific genes, and contributes greatly to the underlying genetic basis of sexual dimorphism.

1. Introduction

Sexual dimorphism occurs in species that produce differentiated sexes, most commonly, males and females. This implies that a single genome carries the information to generate two well-differentiated organisms characterized by sex-specific tissues, behaviors, and physiologies. As a consequence, the genome experiences different selective pressures when carried by males or females, and, therefore, for a single species, two optimal genomes will be selected [1]. This situation generates a conflict between sexes of the same species, known as sexual antagonism (SA), because the optimal genome for one of the sexes (i.e., the genome that confers maximum fitness to one of the sexes) is detrimental in the other sex, and *vice versa*. Understanding how sexual dimorphism occurs, what are the evolutionary forces driving sexual antagonism, and what mechanism can finally or temporarily resolve the conflict are important biological questions.

Currently, the existence of high-throughput technologies permits the study of gene expression on a genomic scale and the identification of genes that are differentially expressed between the sexes of a given species (i.e., sex-biased or sex-specific genes). These studies are helping evolutionary biologists understand how sexual dimorphism is attained. One of the most interesting results that came out of these studies has been the discovery that sex-biased genes are not randomly distributed in the genome of organisms with heteromorphic sex chromosomes (called XY or ZW chromosomes when males or females are the heterogametic sex, resp.; [2–6]). Heteromorphic chromosomes [7–10] are an exciting subject of study due to their unique biological features, such as hemizygosis in the sex carrying the heteromorphic sex chromosomes, dosage compensation (DC) and meiotic sex chromosome inactivation (MSCI). These biological features have been postulated to play a central role in the distribution of sex-biased genes as well as in sexual conflict.

The advent of the first genome sequences also revealed that gene duplication has played an important role in the distribution of sex-biased genes. Betrán et al. [11] analyzed the whole gene set of *D. melanogaster* and showed that retrogenes (i.e., genes originating from the retrotranscription of a parental mRNA molecule) are not randomly distributed throughout the genome. Instead, the authors found that there is a significant excess of retrogenes located on the autosomes that originated from X-linked parental copies (i.e., an excess of X to autosome duplication pattern), with any other pattern being significantly underrepresented. Importantly, an additional feature of these retrogenes is

that their expression tends to be male-specific while the respective parental genes have wide expression patterns [11]. Subsequent analyses have confirmed these patterns not only in other species [12–15], but also for DNA-mediated duplicates [14, 15], suggesting that the same evolutionary forces are probably favoring the relocation and sex-biased expression of duplicated genes. After these evidences, new empirical and theoretical studies are attempting to integrate gene duplication in the sexual conflict resolution.

In this perspective, we: (1) introduce methods for detecting sex-biased gene expression on a genomic scale and discuss the degree to which sexes and tissues contribute to gene expression differentiation, (2) describe analyses that have revealed the nonrandom distribution of sex-biased genes in distantly related species and discuss how much of these patterns can be explained by DC, MSCI and SA hypotheses, and (3) highlight the importance of gene duplication as a driver of the genome distribution of sex-biased genes and outline the recently proposed role of gene duplication as a means to resolve genomic clashes (e.g., escape from MSCI or DC; [12–16]), adaptive conflict, and intralocus sexually antagonistic conflict [17].

To favor a more friendly reading, we include two boxes with aside information that complement the text. In Box 1, we review the biological features of heteromorphic sex chromosomes, including hemizygosity of the X and Z chromosomes, DC, MSCI, and location of SA variation. In Box 2, we highlight examples of the multiple ways in which sex-biased genes can originate through gene duplication.

Box 1. The Biology of the Sex Chromosomes. In many organisms, sex is determined by the presence or absence of a pair of heteromorphic sex chromosomes. This pair of chromosomes can be XY or ZW when males or females are the heterogametic sex, respectively [7–10]. According to the theory of sex chromosome evolution, sex chromosomes were once a pair of equivalent autosomes that evolved into sex chromosomes after one of the homologs acquired a dominant factor for sex determination [18]. Two main processes are believed to drive sex chromosome morphological differentiation, although the speed and extent at which they occur can vary [10, 19]. First, the chromosome bearing the sex-determining gene degenerates as a consequence of the lack of recombination between the homologs [10, 20–23]. Different processes, such as Muller's ratchet, background selection, and selective sweeps, are probably involved in the degeneration of the Y/W chromosome, the importance of which has been discussed in several papers [10, 20–23]. Second, as the Y/W chromosome erodes, strong selective pressures will favor the generation of a compensatory regulatory mechanism of gene expression to balance functional aneuploidy in the heterogametic sex [24–26]. This mechanism is known as dosage compensation (DC; Figure 1), and it is attained in different ways in distant taxa.

For instance, in flies, DC results in an overexpression of the X chromosome in males; in worms, the expression of the two overexpressed X chromosomes in hermaphrodites is reduced by half, and there is overexpression of the single X chromosome of males; finally, female mammals inactivate one of the two X chromosomes and, like males, overexpress their unique active X chromosome (see [24–26, 28, 29]). These DC systems make the male to female and the X to Autosome expression ratios about the same [30, 31].

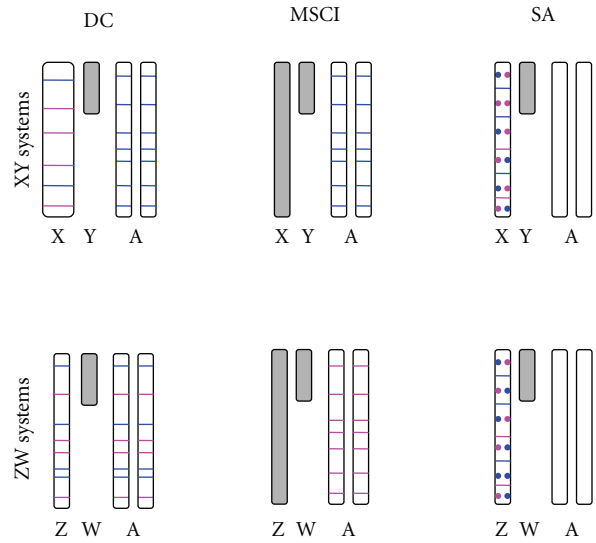


Figure 1: Predictions of the standard DC (Model II and Model V in Table 1), MSCI (Model I in Table 1) and SA (Model VII in Table 1) models on the location of sex-biased genes. In XY systems with DC, the X chromosome is hypertranscribed to equal the level of expression of both autosomal and X chromosome genes in females (not shown). Note that the X chromosome is wider (i.e., hypertranscribed) than the autosomes. Under this hypothesis, there will be physical constraints evolution of male-biased expression for highly expressed genes located on the X chromosome on the (blue lines), while this constraint does not exist for genes evolving female-biased expression (pink lines) or even for male-biased genes expressed at low levels. Consequently, most male-biased genes will be located on the autosomes. Because the ZW system do not have dosage compensation, male-biased genes will be detected on the Z chromosome, but a correction must be made, with the expectation then being that male- and female-biased genes will be equally distributed among the chromosomes. MSCI exists in both XY and ZW systems. Because most sex-biased genes are gonad specific, male- and female-biased genes are predicted to be preferentially located outside of the X and Z chromosomes, respectively; otherwise, they would be inactivated in the gonads. Sexual antagonism predicts that the X and Z chromosomes will carry most of the sexually antagonistic alleles (blue circles for genes beneficial to males and detrimental to females; pink circles for genes beneficial to females and detrimental to males), which has been confirmed in *Drosophila*. Under the classical model [27], resolution of this conflict would be achieved through the evolution of sex-biased and, optimally, sex-specific expression of the antagonistic gene. Therefore, the X and Z chromosomes are predicted to hold the most sex-biased genes. Gray is used to represent heterochromatin or inactivation of the X/Z chromosomes during MSCI.

The independent evolution of DC mechanisms in distantly related XX/XY organisms strongly suggest that the same evolutionary forces might favor the evolution of this compensatory mechanism, not only in XY but also in ZW systems. However, several recent experiments have ruled out the existence of global DC mechanisms acting on the Z chromosomes of birds and Lepidoptera [32–35]. In addition,

Prince et al. [36] recently described another interesting case: in the beetle *Tribolium castaneum*, the X chromosome in males is hypertranscribed such that the X to autosomes ratio is equal to 1, as expected when the compensation occurs in the heterogametic sex. However, the two X chromosomes in females are also hypertranscribed, resulting in an X to autosome ratio equal to 1.5 in this sex. The authors argue that these results probably reflect the absence of a general mechanism in females that inhibits or counterbalances the DC system working in males [36], as would be observed in Drosophila, mammals, or nematodes [30]. Altogether, these results indicate that females might be less vulnerable than males to the deleterious effects of the heterogamy [31] and that the concept of dosage compensation as universal necessity on heterogametic sexes must be reevaluated.

In addition to the well-known processes of Y chromosome degeneration and DC, it has been clear for some time that the X chromosome is inactivated in the male germline of eutherian [37–40] and metatherian mammals [41, 42] (i.e., MSCI). The occurrence of X chromosome inactivation in the male germline has also been observed in worm [5, 6, 43]. However, there has been some debate about the existence of MSCI in the Drosophila male germline [44–50]. A recent study of ZW inactivation in birds revealed that MSCI also occurs in heterogametic females [51]. Therefore, and in contrast to DC, MSCI is likely a more general phenomenon of sex chromosomes in the heterogametic sex (Figure 1). Many hypotheses have been proposed to explain the existence of MSCI. Recently, it has been revealed that meiotic silencing of unsynapsed chromatin (MSUC), a mechanism to silence selfish elements and unsynapsed (i.e., unpaired) DNA in order to prevent aberrant chromosome segregation, might be the basis of MSCI [39, 52–54].

An inevitable consequence of the Y/W chromosome degeneration is a state of hemizygosity in the heterogametic sex. This implies that recessive mutations linked to the X/Z will be exposed to selection in XY and ZW individuals. In addition, X/Z chromosomes will spend two thirds of their time in the homogametic sex and only one third in the heterogametic sex. These two circumstances have important consequences from a population and evolutionary point of view, especially for sexually antagonistic genes. In this perspective, we use the term sexually antagonistic gene or simply sexual antagonism (SA), to describe a gene where the two segregating alleles are selected in opposite directions in males and females (i.e., intralocus sexual conflict; [55]), unless otherwise stated. Theory predicts [27] that new SA alleles will increase in frequency and are more likely to remain polymorphic in the population when they are located on the X/Z chromosome than when they are on the autosomes. Hence, a dominant allele linked to the X that is beneficial to females but detrimental to males will spread through the population because the allele will spend two thirds of its time in females and only one third of its time in males. On the contrary, a recessive allele that is beneficial to males but detrimental to females will increase in frequency because it will be always affected by positive selection in hemizygous males but hidden from negative selection in heterozygous females [27]. The inverse rationale can be applied for ZW systems. Under such circumstances, modifiers of the expression have been proposed to evolve in order to reduce (and optimally inhibit) the expression of the SA gene in the harmed sex, resolving the sexual conflict and allowing the fixation of the antagonistic allele ([27]; Model VII in Table 1). From this, it is often assumed that the genomic location of sex-biased genes is a good indicator of the location of past SA genes [56, 57]. We say past because once the gene is sex biased, the antagonism is resolved.

In summary, two important predictions come from this model [27]: (1) because it is easier for SA alleles to increase in frequency when they are located on the X/Z chromosome, these chromosomes will carry most intralocus SA variation and (2) because SA is resolved through the evolution of sex-biased expression of the antagonistic loci, then most sex-biased genes will be also located on the X/Z chromosomes. However, there is now empirical evidence that sex-biased genes are not a good proxy for SA genes (see [89] and discussion in the text).

These biological features of sex chromosomes contribute to explain the location of genes with sex-biased expression and the genetic architecture of sexual dimorphism. However, as we will describe in the following sections, the evolution and distribution of sex-biased genes present several particularities that cannot be satisfactorily explained by any of the aforementioned biological phenomena unless gene duplication is introduced in the models (see several models in Table 1 and details in the text).

2. Sex-Biased Gene Expression

The development of high-throughput technologies has allowed the study of sex-biased gene expression on a genomic scale. Thanks to these technologies, we can measure the level of expression of all the genes in each sex and determine which genes are differentially expressed. Analyses of expressed sequence tags (ESTs), serial analysis of gene expression (SAGE), DNA microarrays, and, in recent years, massive parallel sequencing technologies are the methods most often used to quantify gene expression on a genomic scale [90]. Notably, most important advances in understanding genome-wide patterns of expression and their evolution derive from microarray technology. DNA microarrays are platforms containing series of microscopic spots with specific target sequences (i.e., oligonucleotides or cDNAs). Under stringent conditions, the probes (e.g., cDNAs or genomic DNA) are labeled with a fluorophore and will hybridize specifically with the target sequences such that the intensity of fluorescence emitted from a spot is proportional to the amount of the hybridizing probe [91]. Through this procedure, one can compare the relative expression of thousands of genes between tissues, individuals or strains (e.g., brain versus testis; larvae versus adults; males versus females; population A versus population B) in a single set of experiments. Although microarray technology has important limitations, many of which will fade away with the application of massive parallel sequencing technologies [90, 92], microarrays still have sufficient sensitivity as to detect small differences among samples [93].

Table 1: Models proposed to explain the location of sex-biased genes in genomes with heteromorphic sex chromosomes.

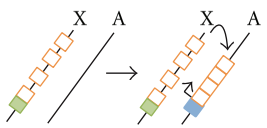
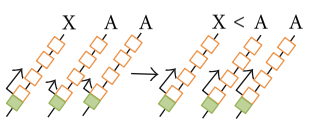
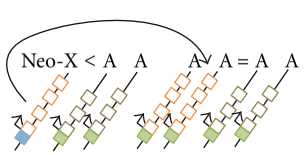
Model	Process	Features of the model					Supporting data	Conflicting data
		Gene duplication	Duplicate function	Parental gene loss	Sex-biased gene location	New gene selection regime		
(I) *Escape from MSCI:* genes are copied from X (or Z) to autosomes to replace the parental gene during meiotic sex chromosome inactivation in the heterogametic sex. Selective pressure should be strong on housekeeping genes that are needed in every cell, immediately after MSCI occurs, and new genes should keep parental function [11, 13, 58].		Yes	Germline	No	Autosomes	Purifying selection	(i) Pattern of movement off the X [11–13, 15, 58] (ii) Complementary expression of parental and new genes [16, 49] (iii) Parental and new genes have same function (e.g., [59]) (iv) Export recapitulates sex chromosome history [60] (v) There is a paucity of testis genes on the X [2, 61] (vi) In mammals, the distribution of male-biased genes is highly depending on cell type [62, 63].	(i) Genes copied from A to A also have male germline function [64] (ii) New genes often have a different function (e.g., [65, 66]) (iii) New genes often evolve under positive selection [67–70] (iv) New genes are lost [69] (v) Export is still on going [11–13, 15, 58] (vi) Paucity of somatic male-biased genes on the X [2, 61]
(II) *Dosage compensation (DC):* if there is a limit on how high you can express a gene and sex-biased expression is achieved increasing transcription [71], genes in the heteromorphic sex cannot achieve as high of an expression on the X (or Z) as on the autosomes and as a consequence an excess of highly expressed sex-biased genes on the autosomes might be observed [72].		No	NA	No	Autosomes	Purifying selection	(i) Highly expressed sex-biased genes on the autosomes have been observed [72]. (ii) Paucity of male-biased genes on the X [2, 61]	(i) Pattern of movement off the X [11–13, 15, 58] (ii) Genes copied from A to A also have male germline function [64]
(III)*Escape from DC:* in new sex chromosomes or chromosomal sections, genes might relocate to autosomes to attain dosage compensation. An efficient way is for the gene and regulatory region to copy to an autosome [14].		Yes	Sex biased	Yes	Autosomes	No change in selection regime	(i) Male-biased genes are relocating from neo-X chromosomes to autosomes [4, 14, 15].	

Table 1: Continued.

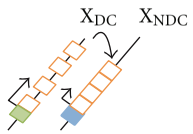
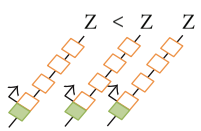
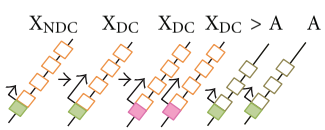
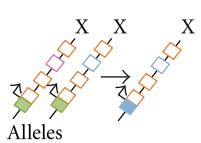
Model	Process	Features of the model					Supporting data	Conflicting data
		Gene duplication	Duplicate function	Parental gene loss	Sex-biased gene location	New gene selection regime		
(IV) *DC interference:* it has been proposed that dosage compensation might interfere (A) with male-biased expression (i.e., interfere with further upregulation in males of DC genes [73]) or (B) with the evolution of tissue-specific expression (e.g., testis-specific expression; [50])		Possibly	Sex biased/ Germline	No	NDC regions	NA	(i) Male-biased genes are locate in regions of the X that are not dosage compensated (i.e., NDC) [73].	(i) Genes copied from A to A also have male germline function [64] (ii) Export is still on going [11–13, 15, 58]
(V) *Absence of DC:* in the absence of dosage compensation, the heterogametic sex will have low expression of the genes on the X or Z chromosome [26, 28–30]. If you correct for this absence you expect not to observe a bias in the location of sexually dimorphic traits [74].		No	NA	NA	X (Z)	NA	(i) Absence of dosage compensation has been observed in birds and Lepidoptera [32–35].	
(VI) *DC effects in the homogametic sex:* the chromatin on the X chromosome might change in a way that the dosage compensated chromosome might now express higher than before in the homogametic sex. This might be interpreted as a lot of female biased genes on the X chromosome [36].		No	NA	NA	X	NA	(i) X has higher transcription than the autosomes [36]. (ii) X has different chromatin even in the homogametic sex [75].	(i) This does not apply to other lineages [30]
(VII) *Sexual antagonism (SA) resolved through the evolution of modifiers of expression:* sexual antagonistic variation (pink and blue alleles) will often appear on the X chromosome [27] and eventually modifiers of expression will evolve in one sex (e.g., females) to reduce expression of the gene leading to a dimorphic trait [27].		No	Sex biased	NA	X (or Z)	NA	(i) Changes in expression are often in cis-regulatory regions [71].	(i) There are a lot of genes with sex-biased expression generated through gene duplication [11–13, 15, 58, 64] or alternative splicing [76, 77]. (ii) Sex-biased genes evolve increasing expression [71]. (iii) SA seems to persist and not resolve [78].

Table 1: Continued.

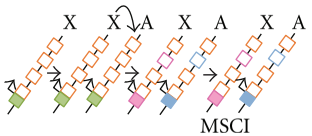
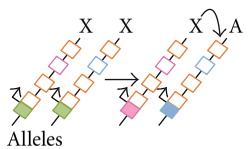
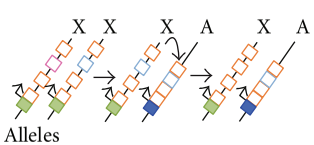
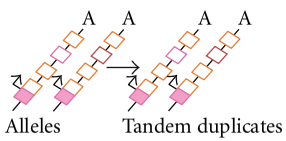
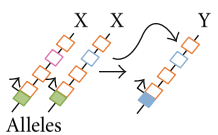
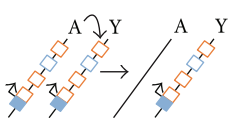
Model	Process	Features of the model					Supporting data	Conflicting data
		Gene duplication	Duplicate function	Parental gene loss	Sex-biased gene location	New gene selection regime		
(VIII) *SAXI:* this model proposes that sexual antagonism leads to MSCI. If a gene duplicates from X to A, the autosomal copy might masculinize and the X copy might feminize because it spends 2/3 of its time in females. And finally leading to MSCI [79].		Yes	Germline	No	X and autosomes	Specialization	(i) Pattern of movement off the X [11–13, 15, 58] (ii) Complementary expression of parental and new genes [16, 49] (iii) Export recapitulates sex chromosome history [60] (iv) There is a paucity of testis genes on the X [2, 61]	(i) Paucity of somatic male-biased genes on the X [2, 61, 70] (ii) Genes copied from A to A also have male germline function [64] (iii) Export is still on going [11–13, 15, 58] (iv) MSCI might be the outcome of the heterochromatization of unpaired chromatin [39, 53].
(IX) *SA resolution through gene duplication:* intralocus sexual antagonism has been proposed to be resolved through duplication and the evolution of a female- and a male-biased gene [80, 81].		Yes	Sex specific	No (Sex biased)	X and autosomes	Specialization	(i) There are a lot of genes with sex-biased expression generated through gene duplication [11–13, 15, 58, 64]	(i) Sex-biased genes generated through gene duplication are often tissue-specific genes [11–13, 15, 58, 64].
(X) *SA resolution through SA allele duplication:* intralocus sexual antagonism might be strong in housekeeping genes because of the pressures of sex-specific tissues. This SA will be resolved with the duplication of one of the SA alleles and the evolution of tissue-specific transcription [17, 64, 69].		Yes	Tissue specific (i.e., testis)	No	Autosomes	Positive selection (i.e., specialization or recurrent)	(i) Many testis-specific genes are generated through gene duplication [11–13, 15, 58, 64] from housekeeping genes from X but also from Autosomes. (ii) New genes often have a different function [65, 66] or evolve under positive selection [67–69] or are lost [69]. (iii) Export is still on going [11–13, 15, 58]. (iv) Since the parental gene still expresses in both sexes, there is still a source of SA explaining why SA persists [78].	(i) Parental and new genes have same function (e.g., [59])

Table 1: Continued.

Model	Process	Features of the model					Supporting data	Conflicting data
		Gene duplication	Duplicate function	Parental gene loss	Sex-biased gene location	New gene selection regime		
(XI) *Gene duplication of a sex-biased gene under diversifying selection:* the pressure of diversifying selection leads to balancing selection and gene duplication and the evolution of female- and male-biased genes from preexisting female- and male-biased genes [82, 83]		Yes	Sex specific	No (Sex biased)	X and autosomes	Positive selection (i.e., specialization or recurrent)	(i) Many sex-specific genes are generated through gene duplication from a preexisting sex-specific genes [82–85]	
(XII) *SA and duplication to the Y chromosome:* the Y chromosome will be a good location for genes with antagonistic effects between males and females [86, 87].		Yes	Sex specific often tissue specific	No	Y (W)	Positive selection (i.e., specialization or recurrent)	(i) Y chromosome gains genes through time [86, 88].	
(XIII) *Random gain and loss of genes on the Y chromosome:* neutral duplication and loss dynamics might lead to the gain of genes on the Y chromosome [88].		Yes	Sex specific often tissue specific	Yes	Y (W)	NA	(i) Y chromosome losses but also gains male-biased genes through time often through DNA-mediated duplication [88].	

Straight arrows indicate steps and level of expression if located in the transcription start site (TSS). Length of the arrow at TSS represents the level of transcription of the gene. Curved arrows indicate duplication or relocation. Cis-regulatory regions of genes are shown in filled boxes and exons in open boxes. Green and orange refer to transcription and function in both sexes. Pink refers to female transcription or specialization of an exon. Light blue refers to male transcription or specialization of an exon. Bright blue refers to tissue-specific transcription.

Microarray analyses have been successfully employed in many different studies [93], but only a handful of these studies have focused on detecting sex-biased expression, mostly in Drosophila [94]. A gene is considered to have sex-biased expression when its level of expression is significantly different between sexes. Obviously, the significance will depend on the experimental design as well as on the statistical test and the arbitrary cutoff applied. This is an important problem, as can be observed from the heterogeneous results obtained from different studies (Table 2), that reveals the necessity to standardize procedures in order to facilitate comparisons among studies [95]. As a general convention, a gene or transcript is considered sex biased when the normalized log signal quotient for the sexes differs by at least twofold [94]. This commonly accepted arbitrary cutoff [96] can be justified when the measurements for a significant proportion of genes are within twofold of the measurements from control hybridizations [2].

Most studies suggest that about half of Drosophila genes are sex biased, with most expression differences between sexes occurring in the gonads ([2, 61, 95–98]; see also Table 2). In addition, males often contribute more to the number of sex-biased genes and almost exclusively to the number of sex-specific genes [99, 100] as females express highly in ovaries many of the genes used for basic cellular functions. Recently, Prince et al. [36] analyzed more than 98% of the predicted coding transcripts in *T. castaneum* and found that 20% of them had differential expression between sexes, with 58% being female biased. However, at a 2-fold cutoff 75% of the sex-biased genes were male biased. On the other hand, Reinke et al. [5] found that about 29% of the elements tested in the nematode *Caenorhabditis elegans* (representing the 94% of the nematode's gene set) showed differentiated expression at twofold cutoff and that most of the expression divergence was generated by tissues in the gonads. However, in the case of *C. elegans*, hermaphrodites make a larger contribution to sex-biased genes than males. In the mosquito *Anopheles gambiae*, females also contribute more than males to the number of sex-biased genes [101]. It is likely that the differences in the number and types of sex-biased genes between mosquitoes and Drosophila reflect the important differences in behavioral and immune traits between the sexes in mosquitoes [101].

Sex-biased expression has been also studied in a few vertebrates, arriving at similar conclusions to those described for Drosophila. Microarray experiments in chicken and mice showed that in these organisms most of the sex-biased expression occurs in the gonads and that the contribution of males to the number of genes that show expression divergence between sexes is higher than that of females [102, 103]. Even in organisms without heteromorphic sex chromosomes or without sex chromosomes (i.e., Xenopus and Zebrafish), this holds true [104, 105].

In summary, males contribute more than females to the expression divergence between sexes in most analyzed species. In addition, further analyses have proven that male-biased expression contributes highly to interspecific gene expression divergence, indicating that male-biased genes also make a significant contribution to divergence among species in flies [3, 95, 97, 106] and mammals [107, 108]. Hence, the ratio of interspecific divergence in gene expression to intraspecific variation is significantly higher for male-biased genes than for female-biased and unbiased genes [106]. This result is consistent with positive selection operating over expression divergence in male-biased genes, while purifying or neutral selection would be operating on female-biased and unbiased genes. In agreement with this result, Conallon and Knowles [71] showed that the evolution of male-biased expression is mostly an active change; that is, male-biased expression is achieved by increasing the expression in males relative to females (for additional comments see [72, 109]).

The differences in the evolutionary patterns of male-biased compared to female-biased and unbiased genes are also observed at the protein sequence level. Male-biased genes have higher evolutionary rates (i.e., higher Ka/Ks ratios) than female-biased genes and unbiased genes that have been interpreted as adaptive evolution [84, 107, 110]. In addition, sex-biased expression and Ka/Ks ratios correlate with expression divergence in Drosophila [97, 111–114] and mammalian species [115], suggesting that a common selective force underlies all these processes [97]. This fast evolution at the sequence and expression level and the fact that most male-biased genes are testis-specific [2, 90, 95, 97, 98, 116] suggest that once expression testis-specific expression is achieved, genes may be released from pleiotropic constraints and, consequently, be able to evolve more quickly responding to the specific selective pressures of the particular tissue at the expression and sequence level than more widely expressed genes [116]. In agreement with this idea, a recent study revealed that the sex-specific expression in reproductive organs correlates with a high rate of gene evolution, not explained by a narrow pattern of expression [117].

Given that the origin of most gene expression divergence among species involves changes in male-biased gene expression and that this divergence might be conceivably driven by positive selection, it is crucial to understand the underlying evolutionary processes and forces driving these patterns. There are three different mechanisms that can lead to the evolution of sex-biased or sex-specific gene or isoform expression (Figure 2): (1) evolution of modifiers of the expression (i.e., *cis*- or *trans*-regulatory changes [118]), (2) generation of sex-specific transcriptional variants (with or without coding-exon duplication [76, 77]), and (3) generation of sex-specific genes by gene duplication [11–13, 15, 58, 64].

The evolution of modifier(s) of gene expression may turn a widely expressed gene into a sex-biased gene. The relative contribution of *cis*- and *trans*-regulatory changes to the evolution of expression divergence is a topic under study [119]. Analysis of allele expression in the F1 individuals resulting from crossing different lines of Drosophila allowed the estimation of the relative contribution of *cis*- and *trans*-regulatory changes to the expression divergence in this genus. Under this framework, many studies have shown that *cis*-effects explain more intraspecific [120, 121] and interspecific changes [122–124] than *trans*-effects do (but see also [125]). However, while in a few cases *cis*-effects are clearly implicated

Table 2: Summary of microarray studies data for sex-biased expression in nonvertebrate animals and male versus female contributions to sex-biased expression.

Species		Approach	Tissue	Cut-off	Sex-biased Percentage	Male/Female	References
Drosophila	*D. melanogaster*	cDNA Microarray (4,000 cDNAs)	Whole body	Twofold\ Fourfold	>50%\NA	Females\Males	Ranz et al. [3]
		FlyGEM Microarray	Gonads and whole body	Twofold	17%	Males (64%)	Parisi et al. [61]
		Affymetrix Microarray (18,800 transcripts)	Whole body	FDR ≤ 0.01\ Twofold	88%\25%	~Equal contribution from males and females	Ayroles et al. [98]
	D. pseudoobcura	Oligo Microarray	Whole body	Twofold	42%	Males (64%)	Jiang and Machado [95]
	Seven Drosophila	Oligo Microarray	Whole body	FDR ≤ 0.01	13% to 32%	Male mostly (5 out of 7 species)	Zhang et al. [97]
	D. simulans	cDNA Microarray (4,000 cDNAs)	Whole body	Twofold\ Fourfold	>50%\NA	Females\Males	Ranz et al. [3]
Tribolium castaneum		cDNA Microarray (98% of cDNAs)	Whole body	FDR ≤ 0.01	20%	Females (58%)	Prince et al. [36]
Caenorhabditis elegans		Microarray (94% of the gene set)	Whole bodies and mutants lacking germline	Twofold	29%	Hermaphrodites	Reinke et al. [5]
Anopheles gambiae		Affymetrix Microarray	Whole bodies	Fourfold	10%	Females (71%)	Hahn and Lanzaro [101]

in the regulation and expression divergence between sexes [118], its implication on a genomic scale is not known. *Cis*- and *trans*-regulatory changes towards sex-biased expression might be unlikely in single copy housekeeping genes, as these changes might be detrimental in depriving one of the sexes of a basic molecular function [17, 64].

The generation of sex-specific transcriptional variants may allow the acquisition of sex-specific protein isoforms (i.e., sex-specific alternatively spliced variants) without modifying the pattern of expression and/or function of the original isoform (i.e., the other alternatively spliced variant). This molecular level of sexual dimorphism has not been widely investigated. Research in Drosophila suggests that between 22% and 32% of the transcriptional variants are sex specific [126, 127]. A fraction of these variants are known to contribute to sex determination, but others are active only at the very end of the regulatory cascade, leading to a sexually dimorphic phenotype [126, 127]. Sex-specific splice variants are most abundant in the testis, they are also common in the head [127] and likely have important functions because they are highly conserved in Drosophila [127]. Sex-specific transcriptional variants have also been observed in mammals and often in the testis [77, 128].

Finally, gene duplication can also generate sex-specific genes. This is another way of generating sex-biased genes without modifying the expression pattern of the original (parental) gene [17, 64]. While particular examples of sex-specific duplicate genes have been accumulating for a long time (e.g., [58, 129]), including genes duplicated onto the Y chromosome [130], interest in the mechanisms behind this has increased since the first studies showing the nonrandom distribution of retrogenes and their tendency to be expressed specifically in males [11, 13]. Similar to the tissue-specific expression, gene duplication could also lead to the release from the pleiotropic constraints on the original function of the parental gene. Hence, gene duplication would allow for the evolution of new functions in one of the paralogs, while the other copy keeps its broad expression and original functions [17, 131]. In the following sections, we review the main studies in the last few years on the genomic location of sex-biased genes and the role of gene duplication in the evolution and genomic distribution of these genes.

3. Genomic Location of the Sex-Biased Genes

Many studies in XY systems have shown that X chromosomes in Drosophila [2–4], and nematodes [5, 6] have a significant underrepresentation of male-biased genes. In mammals, the distribution is highly dependent on cell type. For instance, there is an excess of male-biased premeiotic and postmeiotic genes on the X [62, 132] but the contrary is true for genes expressed during meiosis [63]. In addition, some X-linked genes seem to be overexpressed in brain in males and females [26]. In the particular case of Drosophila, chromosomal rearrangements have been critical in demonstrating that the compositional nature of the X chromosome is not just an

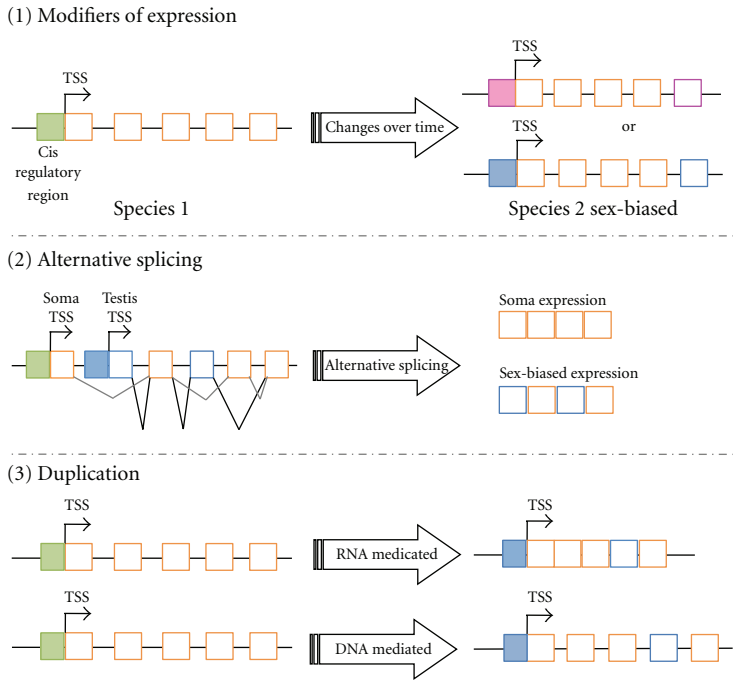


Figure 2: Sex-biased expression can be acquired in three different ways. (1) Changes in the cis-regulatory regions or in *trans*-acting factors can lead to a gene that is more highly expressed in females (pink) than in males or more highly expressed in males (blue) than in females. (2) A gene can have alternative transcripts through alternative splicing. One of those transcripts might be sex-biased in expression. This sex-biased expression can often, though not always, be detected by standard methods (e.g., microarray analyses) if it leads to higher expression in one sex, but it may require direct analysis of alternative transcripts. (3) Gene duplication can lead to the creation of sex-biased genes (e.g., a male-biased gene) through RNA- or DNA-mediated gene duplication. Gene duplication might create one or two sex-biased genes. For details of the proposed models, see Table 1. In the three instances, the exons might specialize. Gene expression patterns are shown with filled boxes and exons with open boxes. Green and orange refer to expression in both sexes. Pink refers to female-specific expression or specialization of an exon. Blue refers to male-specific expression or specialization of an exon.

historical peculiarity of this chromosome. Instead, evolutionary forces must be operating on the sex chromosomes such that newly formed X chromosomes develop the same composition as the original X chromosome to which it was fused [4, 133]. For instance, in the *Drosophila pseudoobscura* lineage, ancestral Muller elements A (the current XL arm, homologous to the X chromosome in *D. melanogaster*) and D (the current XR arm, homologous to the autosomal arm 3L in *D. melanogaster*) were fused, generating a new metacentric X chromosome about 10–18 million years ago [134, 135]. Interestingly, this originally autosomal arm underwent demasculinization during its evolution to a neo-X chromosome. There is currently an underrepresentation of male-biased genes on that chromosomal arm, something not observed in species where this chromosome is and autosome [4]. These observations strongly suggest that evolutionary forces might be favoring demasculinization of X chromosomes [2, 4]. It seems, however, that this trend might be only a feature of old male-biased genes since a trend for young male-biased genes to originate through gene duplication or "de novo" on the X chromosome has been observed [136–138]. These new male-biased genes are either translocated, lost or may lose the sex-biased expression as the trend reverses for older male-biased genes. For older male-biased genes, an excess is observed in the autosomes consistent with an overall excess of male-biased genes on autosomes in some species (i.e., Drosophila).

The patterns observed for the Z chromosomes, however, are not conclusive or consistent across species. This lack of consistency might reflect that the studies are often not done on whole organism but on particular tissues. For instance, analyses in chicken first suggested that male-specific genes expressed in brain are significantly overrepresented on the Z chromosome [139, 140]. In contrast, male-specific genes expressed in the testis appeared to be randomly distributed and female-specific genes expressed in brain and ovary were observed to be significantly underrepresented on the Z chromosome [139, 140]. Subsequent publications have contradicted these findings. More recently, Mořkovýský et al. [141] published another analysis showing that oocyte-genes are overrepresented on the Z chromosome. In addition, other studies suggest that testis-specific genes are overrepresented on the silkworm's Z chromosome, another organism with a ZW system [142]. Recently, Mank et al. analyzed the expression patterns genome-wide during chicken development and concluded that contrary to the prediction of current models [27], most sex-biased genes are located on autosomes [74].

Regardless of the difficult interpretation of the results obtained in organisms with ZW systems, disentangling the causes that shape the distribution of sex-biased genes in

genomes containing heteromorphic sex chromosomes is one of the most intensely active fields of research related to this topic. The actions of DC, MSCI, and SA on sex chromosomes have been suggested as a way to explain the nonrandom distribution of sex-biased genes.

DC systems hypertranscribe the X-linked genes in somatic male cells to raise the expression level to that of the autosomes and the two X chromosomes in females (reviewed in [30]). Because male-biased expression evolves in Drosophila by increasing the level of expression in males relative to females [71], evolving male-biased expression (i.e., higher expression in males) might be more difficult for highly expressed genes when located on the already hypertranscribed X chromosome than it would be for autosomal genes.

Vicoso and Charlesworth [72] tested this hypothesis by comparing the genomic locations of male-biased genes with different levels of expression and confirmed that highly expressed genes located on the X are less likely male biased than lowly expressed genes. Moreover, in agreement with this prediction, X-linked genes that are not dosage compensated (i.e., not bound by the dosage compensation complex) have a higher probability of showing male-biased expression than dosage compensated genes [73]. Therefore, under this hypothesis, the demasculinization of the X chromosome in Drosophila would be a consequence of a physical limitation on the ability of dosage compensated genes to increase their level of expression ([72]; Model II in Table 1). It has also been suggested that this effect is attributable to DC interfering with the evolution of male-biased expression (i.e., further upregulation of DC genes in males [73]; Model IV.A in Table 1). The most obvious limitations of the first model is that it only explains the X chromosome deficit for highly expressed male-biased genes, and a limitation of both is that most male-biased genes are testis-specific [2, 61], a tissue where DC is reportedly absent or at least not mediated by the dosage compensation complex [29, 45, 143]. Additionally, these models cannot explain the defeminization (i.e., the significant underrepresentation of female-biased genes) of the chicken Z chromosome, because, as we explained above, this organism has not evolved any dosage compensation mechanism.

An additional model (Model IV.B in Table 1) has just recently been proposed [50]. The authors suggest that the hyperacetylated state of the X chromosome in males due to DC reduces the capability of other regulatory mechanisms to repress tissue-specific X-linked genes in other tissues. According to this hypothesis, highly tissue-biased (testis-specific, as well as any other tissue-specific) genes will be selected to be located out of the X chromosome to assure its silenced state in other tissues (ovary or somatic tissues). This would explain the paucity of male-biased genes on the X as many of them are testis-specific genes. Interestingly, they also found that ovary is the only tissue where DC is not interfering with the evolution of tissue-specific expression. Further analyses are needed to understand the different effect of DC in male- and female-biased expression.

As described in Box 1, MSCI is likely a general property of heterogametic sexes. In nematodes [144], flies [47, 49], but see [50], mammals [37], and chicken [51], an early X or Z chromosome inactivation has been observed early in meiosis. During this stage of the spermatogenic/oogenic process, X/Z-linked genes are repressed, whereas autosomal genes are actively transcribed. The MSCI model postulates that it would be beneficial to duplicate an X/Z-linked gene to an autosome if it was required during MSCI. Because most sex-biased genes are gonad specific, MSCI could explain the demasculinization or defeminization of the X or Z chromosomes, respectively (Model I in Table 1). However, although this hypothesis could potentially explain a large fraction of the sex-biased genes, as most sex-biased genes are testis-biased [61], an important observation threatens this hypothesis. The rationale proposed by Parisi et al. [2] was that if MSCI were the sole cause of the X chromosome demasculinization, then we would expect a random distribution of somatic male-biased genes. Contrary to this expectation, demasculinization of the X chromosome is also observed in gonadectomized flies [2].

Finally, it has been both predicted [27] and observed [89, 145] that the X and Z carry the most intralocus SA polymorphisms. The resolution of this conflict following the classic model could explain the distribution of sex-biased genes [27], because the resolution is predicted to occur through the evolution of sex-biased expression, as introduced above. In particular, when the effects of a new antagonistic allele are mostly dominant or partially dominant, then female-beneficial (male-deleterious) alleles will increase in frequency on the X chromosome and these genes will evolve into female-biased genes. Similarly, male-biased genes will only have a good chance to evolve from SA genes in autosomes, explaining the relative demasculinization of the Drosophila X chromosome [2–4]. The inverse rationale can be applied in chicken to explain the defeminization of the Z chromosome ([139, 140] but see [141]). However, this prediction stands on the dominance effects of alleles, because when the antagonistic effects of a new allele are recessive, then male-beneficial (female-deleterious) alleles will increase in frequency on the X chromosome and will evolve into male-biased genes [27].

However, this hypothesis that might apply to some instances, also presents serious difficulties (Model VII in Table 1). First, X chromosomes hold the most intralocus SA variation (at least in Drosophila; [89, 145]), but, contrary to what the classical model predicts, the X and Z do not accumulate most sex-biased genes in Drosophila, nematodes and chicken [2, 4, 5, 56]. Second, contrary to predictions of the SA model, sex-biased expression seems to evolve frequently by increasing rather than decreasing the expression in one of the sexes [71]. Third, intralocus SA does not seem to be resolved but persists instead [78, 89]; probably because modifying the expression of certain genes (such as a single copy housekeeping gene) might not be as easy as proposed by the model when they are also needed to perform basic, not sexually differentiated, cellular tasks [17]. In those instances, the transient resolution of SA conflict might occur by gene duplication of the SA allele (see below; Model X), but, given that the broadly expressed gene will still be broadly expressed, SA can reappear (i.e., persist [17]).

In summary, the general conclusion from analyses performed in nematodes, flies and mammals is that there is a significant paucity of male-biased genes on the X chromosomes in these organisms [2–6, 63]. However, in the case of organisms with heterogametic females, the data are much more limited, and published results have been contradictory. Finally, although several hypotheses have been postulated, none of them can completely explain the distribution of sex-biased genes, and it is not known what amount of bias may be explained by each of the hypotheses. There might not be one model that fits all the data [14] but the role of gene duplication might be more important than previously acknowledged and could have multiple roles outside of that postulated in MSCI hypothesis (e.g., replacing the parental gene). We review the data supporting this in the next section and in Box 2 and remark how gene duplication likely has an unforeseen, multifaceted, and important role in the resolution of the X-related clashes, adaptive conflicts and sexual conflicts generated by the unique biology of this chromosome and sex-related selective pressures.

4. The Contribution of Gene Duplication to Sex-Biased Expression: Resolution of Gene Conflicts

Past analyses of retroposed copies of genes have revealed that gene duplicates escape the X chromosome and acquire male functions, often in the male germline [11–13, 16]. This pattern, often called the out-of-the-X pattern, is concomitant with sex chromosome formation in the human lineage [16, 60] and has been observed in other lineages, including mammals and flies, and in the parallel retroduplication of the same duplicated genes in closely related species [12, 16, 146], all of which suggest a shared underlying selective pressure. Unfortunately, analogous studies in chicken have not been fruitful due to the low number of retrogenes in the genome of this species, which is probably caused by the absence of the appropriate type of non-LTR retrotransposons [147].

To discern between mutation and selection as the driving force behind the out-of-the-X pattern observed in retrogenes, Emerson et al. [13] looked at retropseudogenes in the human genome. These authors found that retropseudogenes (i.e., nonfunctional duplicates which are supposed to evolve neutrally, and thus reflect mutational biases) do not show an out-of-the-X pattern, supporting the action of positive selection. However, this procedure cannot be applied in Drosophila, because its genome does not have enough retropseudogenes [12], and therefore, other evidence has to be gathered to uncover any mutational biases in this species. In one analysis, it was found that the out-of-the-X pattern is not explained by the insertional biases recently described for retrotransposable elements [148] that allegedly encode the machinery used for retrotransposition [149]. In *D. melanogaster*, transposable element (TE) insertions, and by extension retrogene insertions are affected by several factors, including recombination, genome compactness, gene expression, and the presence of coexpressed genes clusters [148]. Unlike retrogenes, the number of TE insertions on the X chromosome was higher than expected revealing that the pattern of retrogene distribution cannot be explained by the insertional biases of TEs and suggesting that selection could be a driving force. In another analysis, Vibranovski et al. [15] postulated that if the biased relocation pattern observed for retrogenes was simply a mutational process, we would expect a random distribution for relocated duplicates arising from other mutational mechanisms. Analyses of DNA-mediated duplicates instead showed relocation patterns similar to those observed for retrogenes, that is, an excess of X-to-autosome relocations and male-biased expression of the relocated genes [14, 15]. Studies of patterns of duplications and relocations (duplication with loss of the parental gene) that include not only retrogenes but also DNA-mediated duplications in Drosophila [4, 14, 15] have confirmed that retrogenes are duplicated on autosomes more often than expected. In addition, these studies have also uncovered that in neo-X chromosomes male-biased genes often get relocated to autosomes, mainly through DNA-mediated duplications [14].

In light of these results, it has been suggested that gene duplication is a mechanism that allows X-linked genes to overcome DC or MSCI and achieve male-biased expression (i.e., escape from genomic clashes, MSCI or DC; Table 1). Supporting these ideas, Bachtrog et al. [73] found that the probability that an X-linked gene will generate a relocated copy on an autosome is higher for compensated than noncompensated genes [73]. In addition, autosomal-linked retrogenes originated from X-linked parental genes are more often highly expressed in meiosis, while the respective parental genes seem to be inactivated [16, 49, 73]. Note that these two hypotheses state that escape from the X chromosome is a consequence of DC or MSCI.

Interestingly, there are some proposed SA models that involve gene duplication. For instance, the SAXI hypothesis (Model VIII in Table 1) suggests an alternative and conceptually different explanation [79]. This hypothesis states that rather than causing relocation of genes from the X chromosome, MSCI is a consequence of the continuous duplication of SA genes from the X to the autosomes. This process would end with the accumulation on the X chromosome of female beneficial antagonistic genes that are detrimental in spermatogenesis. Consequently, MSCI evolved to avoid the expression of these X-linked genes that would be detrimental in male gonads. In other words, SA is the evolutionary force that drives demasculinization of the X chromosome and germline X inactivation. Wu and Xu's paper not only suggest an explanation for the out-of-the-X pattern but also states that gene duplication is a way of resolving sexual antagonism, an idea that was suggested earlier ([80, 81]; Model IX in Table 1). These duplicative models have been recently developed further to try to explain the location of sex-biased genes ([150]; Model IX in Table 1). In these models, the resolution of the intralocus sexual antagonism by gene duplication has been proposed to occur through the creation of a male-specific and a female-specific duplicate gene (Model IX in Table 1).

All the aforementioned hypotheses, although satisfactory in some instances, cannot cope with the total complexity of this phenomenon (Table 1). For instance, DC cannot

explain the complementary expression between parental and testis-specific duplicated genes or the way relocated genes evolve [64–69, 85]; MSCI cannot explain why relocated male-biased genes, which supposedly replace the parental function in the germline, can have a different function, evolve under positive selection or eventually become lost [65–69]; the SAXI hypothesis cannot explain why the export of genes out of the X chromosome is still ongoing and why SA in the X chromosome persists. In addition, recent results present additional challenges to these models [64] by demonstrating that the relocation patterns may vary depending on the type of genes that are being analyzed and that a big fraction of testis-specific genes have been duplicated genes from one autosome to another. Some recent models of the resolution of the intralocus sexual antagonism by gene duplication and the generation of a male-specific and a female-specific gene predict that accumulation of female-biased genes on the X chromosomes and male-biased genes on the autosomes is independent of the dominance effects and can account for the observed autosomal locations of male-biased genes [150]. However, these models do not completely fit the data because most sex-specific genes are male-specific genes derived from broadly expressed genes or sex-specific genes duplicated from other sex-specific genes [4, 11–13, 15, 16, 64, 82–85, 94, 97, 151–153].

Recently, a new duplicative model has been suggested [17, 64]. Gallach et al. [64] studied the duplication patterns of Drosophila mitochondrial genes encoded in the nucleus and found an extreme rate of duplicate relocation for both, RNA- and DNA-mediated duplicates. Interestingly, 83% of the relocated genes evolve at higher rates than their respective parental genes and are expressed only in the testis. Importantly, they also found a significant excess of autosome to autosome relocations [64]. The authors concluded that a particular SA model (Model X in Table 1) was the most satisfactorily explanation for their results and suggested that gene duplication might be an important mechanism for resolving intralocus SA because of the high potential for relocated genes to develop male-specific expression [154–156]. The high potential of relocated genes to develop male-specific expression was proposed to be the result of both insertional biases close to germline genes that facilitate male-specific expression together with selection [64].

Based on these observations, Gallach and Betrán [17] have suggested a new duplicative model for the resolution of intralocus SA (Model X in Table 1). According to this model, relocation and testis-specific expression of the antagonistic allele will result in the parental gene keeping its original function and expression patterns (i.e., wide expression in males and females) with the antagonistic allele fixed in the population as a new gene with testis-specific functions. This model has the power to help interpret multiple observations. It can explain why the X chromosome holds the most sexually antagonistic alleles but is not the place where the conflict is resolved; it can explain why the gene export persists as well as the significant excess of X to autosome duplications; it can also explain why an excess of autosome to autosome duplications is observed for certain sets of genes [64]. In agreement with this hypothesis, there are data that indirectly indicate that male-biased genes originate more often through gene duplication than unbiased or female-biased genes [84, 94, 97, 151].

The contribution of gene duplication to sex-biased expression might also come from the duplication of a preexisting sex-biased gene and the driving force could be, in this case, not intralocus sexual conflict but interlocus sexual conflict [55]. It is now clear that a large fraction of sex-biased genes may result from gene duplication of preexisting sex-biased genes [82, 84, 85, 152, 153]. This has occurred multiple times in some gene families (e.g., female reproductive proteases or accessory gland proteins [83–85]). These are sex-specific genes that interact with the other sex and positive and diversifying selection have been proposed to act on those genes as a consequence of strong sexual selection [82, 83]. This process is different from the model above that proposes that the resolution of intralocus sexual antagonism occurs by gene duplication, because the parental genes are already sex specific. Therefore, the model that would fit these data (Model XI in Table 1) is a model that combines positive selection and balancing selection maintaining segregating variation in sex-biased genes with gene duplication; that is, there is balancing selection or adaptive conflict due to diversifying selection and the resolution of this genetic conflict by gene duplication [157]. In the particular cases of the female reproductive proteases and accessory gland proteins, the recurrent action of positive and diversifying selection might lead to a particular dynamic process that involves an increase in gene number (i.e., creation of a gene family) but also eventual pseudogenization of some genes [83–85].

Interestingly, in these last duplicative models (Model X and Model XI in Table 1), there is observed (or assumed) preexisting allelic variation in the parental gene resulting from strong diversifying or sexually antagonistic selection that promotes the gene duplication ([157]; see also model XII in Table 1) and the same selective pressures might persist in the form of positive selection or specialization of the duplicated gene [17]. Finally, the high turnover with respect to gene loss of male-biased genes [84, 97] likely leads to an underestimation of the impact of gene duplication as a means to generate sex-biased genes through these models.

The extent to which each model in Table 1 can explain the genomic location of sex-biased genes is lineage-dependent and some of these models deserve further investigation in many lineages. However, it is clear, as reviewed in this section and highlighted by the examples in Box 2, that gene duplication has an important role in the origin of sex-biased genes. The role of gene duplication could potentially be very large if we consider that the rate of new gene per generation is known to be very high, as revealed by new genomic data [158]. Consistent with this view, recent data from Drosophila suggest that the adult male transcriptome contains more genes of recent origin than the adult female transcriptome indicating that new genes have a role as sex-biased genes in Drosophila [159]. The role of gene duplication in creating sexually dimorphic genes is likely to hold even in organisms without sex chromosomes (e.g., Zebrafish) as the adult transcriptome appears to be young in this species as well

[159]. Note that only a small fraction of the models in Table 1 will apply to genomes without sex chromosomes.

Box 2. Examples of the Generation of Sex-Biased and Sex-Specific Genes through Gene Duplication. Gene duplication plays a major role in evolution, because it can increase the dose of a gene, partition the function of a gene or create a new function [160, 161]. Functional data have accumulated revealing that gene duplication patterns contribute greatly to sexual dimorphism in genomes with heteromorphic sex chromosomes, as introduced in the text. In this Box, we highlight some relevant examples of gene duplications, with or without relocation (Figure 3).

Gene Duplications onto the Y Chromosome. Duplication of genes to the Y has been observed in mammals, flies, and even in plants [86, 88, 162, 163]. These studies suggest that relocation onto the Y chromosome is one way this chromosome avoids being completely eroded (Models XII and XIII in Table 1). The most spectacular example of this observed to date occurs in Drosophila, where at least thirteen protein-coding genes have been originated through DNA-mediated duplication from the autosomes to the Y chromosome [88, 164–167].

Because the Y chromosome is a nonrecombining chromosome, it is expected to degenerate over time. Nevertheless, it was found that gene duplications and their architecture within the Y chromosome play a major role in offsetting this process in some species. Examples of this can be found in humans and chimpanzees, where duplicated genes on the Y chromosome have been found in a head-to-head orientation in long palindromes [168, 169]. Further analyses of the chimpanzee male-specific region of the Y chromosome (MSY) revealed that the chimpanzee Y chromosome contains twice as many palindromes as the human MSY but also has lost some palindromes present in the MSY region in humans. This result suggests an extraordinary turnover of genes on the chimpanzee Y chromosome. Other Y or W chromosomes have also been found to contain multiple gene copies disposed in a palindrome-like structure that have been demonstrated to be undergoing concerted evolution [170–172]. Hence, gene duplication is a way of not only generating sex-specific genes but preserving the functions of these genes in nonrecombining chromosomes.

Gene Duplications That Escape the X Chromosome. Convergent acquisition of X to autosome retrogenes strongly suggests the existence of selective pressure acting on the fixation of these retrogenes. The retrogene *Utp14*, a gene involved in pre-rRNA processing and ribosome assembly, is a convincing example of convergence, as the recurrent emergence of this retrogenes from the same parental gene has been found in four distinct mammalian lineages [146, 173]. Potrzebowski and colleagues [16] also found two examples of recurrent emergence of retrogenes that have moved from the X-to-autosomes in mammals (*Pgk* and *Centrin*). Other studies have also revealed recurrent retroduplications from the X to autosome of in Drosophila [12]. For instance, there

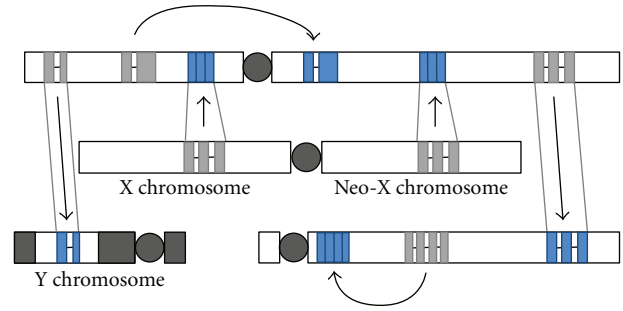


Figure 3: Gene duplication to the Y, out of the X and neo-X chromosomes and between autosomes, often leads (with or without the loss of the parental gene) to sex-biased expression (often male tissue-specific expression; blue). Sex-biased expression (in blue) appears to evolve often as long as the duplication involves relocation. Both RNA-mediated and DNA-mediated gene duplication appear to contribute to these effects.

appears to have convergent evolution of the *Dntf-2* and *Ran* retrogenes in three independent lineages: *D. melanogaster*, *D. ananassae*, and *D. grimshawi* [69].

Some Gene Duplications Enter the X Chromosome. In mammals, in addition to the pattern of gene duplicates leaving the X chromosome, there is also an excess of retrogene duplicates entering the X chromosome in mammals [13]. Because retropseudogenes also show this trend, this bias is partly explained by mutational biases. However, because the trend for retropseudogenes is not as strong as it is for new retrogenes, selection was postulated to have shaped these biases [13] and with the selective pressures involved only recently being proposed [174]. A fraction of these new genes are expressed in spermatogonia or postmeiotically and support the action of selection in favor of X-linked spermatogenesis genes that are expressed before or after MSCI. For instance, *USP26* is a X-linked gene that is expressed in the spermatogonia and is evolving under positive selection [175]. Other new genes entering the X might be involved in the maternal-fetal conflict in placenta. For instance, a retrocopy of *Fth1*, a gene involved in iron metabolism, entered the X chromosome and evolved placental-specific functions.

Autosome to Autosome Gene Duplications. Gallach et al. [64] recently analyzed all duplicated mitochondrial genes in the nucleus of *D. melanogaster*. The authors showed that both RNA- and DNA-mediated mitochondrial gene duplications exhibit an unexpectedly high rate of relocation, and found an excess of X to autosome and also of autosome to autosome duplications. However, and in agreement with previous analyses, they found a significant deficiency of autosome to X duplications. They called this the avoidance-of-the-X pattern [64]. In addition, they found that the relocated genes (including those that originated from an autosomal parental gene) tend to have testes-specific expression and functions related to energy production-related functions. These genes evolve faster than their parental genes [64], and some of them (*CG18418* and *CG6255*) have been identified as having been

under positive selection [110], indicating that they have been selected for different functions in the male germline. This study suggest that although out of the X is the main pattern observed in genomes, other patterns, such as the avoidance-of-the-X pattern, may appear when a particular set of genes are analyzed independently. These observations are not explained by any of the models in Table 1, except model X, that is, SA resolution through SA allele duplication and the evolution of tissue-specific expression in the new gene.

Tandem Duplication. As introduced in the previous section and in the model XI in Table 1, there are sex-specific gene families that experience ongoing diversifying selection, leading to segregating variation, gene duplication and adaptive evolution. Examples of these are accessory gland proteins and female reproductive proteases [83, 152]. These genes that are often duplicated in tandem, maintain the patterns of expression of the parental genes (male or female tissue-specific expression), and exhibit recurrent pseudogenization, copy number variation, and gene conversion, revealing the strength and the changing direction of selection. Gene duplication is an important mechanism for diversifying preexisting sex-specific functions.

5. Conclusions

The advent of high-throughput technologies has allowed evolutionary biologists to analyze sexual dimorphisms in expression at the genomic level. Microarrays have been primarily used for these analyses, although next generation sequencing technologies (i.e., RNA-seq) are now being successfully used and will probably allow more direct, homogeneous, and conclusive results for many types of analyses. Despite the associated technical problems, important conclusions regarding the origin and evolution of sex-biased and sex-specific expression can be extracted from microarray analyses. A significant fraction of genes exhibit biased expression in one sex, and male bias is more common than female bias. In addition, male-biased genes tend to be expressed in a tissue-specific manner, generating most of the sex-biased expression in testis. In addition, male-biased genes evolve in a significantly different fashion from female-biased and unbiased genes: male-biased genes normally evolve increased levels of expression in this sex, the ratio interspecific divergence to intraspecific divergence is high compared with female-biased, and unbiased genes and male-biased genes have higher evolutionary rates than other genes. These observations provide evidence that male-biased genes are an important source of evolutionary change, are under strong selective pressures, and contribute greatly to sexual dimorphism.

Sex-biased genes are not randomly distributed across genomes. Instead, demasculinization of the X chromosome (i.e., a significant absence of male-biased genes in this chromosome in most cell types) is observed in many species and conflicting, inconclusive results, have been published regarding the Z chromosome. Three main hypotheses (and their derivatives) have been suggested to explain this: dosage compensation (DC), meiotic sex chromosome inactivation, (MSCI) and sexual antagonism (SA). Each one of these hypotheses may apply to a big fraction of the sex-biased genes in a given lineage, but, in another lineage, the particular hypothesis might have low predictive power. Many studies in the last few years have emphasized that gene duplication is likely an important mechanism for the evolution of sex-biased gene expression in situations where the parental genes are constrained by genomic clashes, DC or MSCI, adaptive conflict, or sexual conflict (SA). Recent study of trends in gene relocation and the evolutionary features of the new genes suggest that gene duplication might be an important mechanism for the resolution of intralocus sexual conflict. In particular, we emphasize models that may deal with many of the unexplained observations reported for DC, MSCI, and the classical model of SA resolution. Upcoming of new genome sequences from organisms with XY and ZW systems, better studies of gene gains and losses, increased availability of more precise and unbiased expression data, and the functional analyses of the relevant genes will permit the explanatory power of each hypothesis to be measured in every lineage.

Indeed, the same forces that are driving fast evolution of sex-related proteins, such as sexual selection and sexual antagonism, might also promote the creation of new sex-biased genes through gene duplication and help explain the young age of these genes and their rapid turnover.

Acknowledgments

The authors would like to thank Jeff Demuth for providing feedback and two anonymous reviewers for numerous comments on this work. This research was supported by UTA startup funds and Grant nos. GM071813 and ARRA GM071813-0551 from National Institutes of Health, USA, to E. Betrán.

References

[1] B Oliver, "Sex-biased genomic expression," in *Evolutionary Genomics and Proteomics*, M. Pagel and A. Pomiankowski, Eds., pp. 193–206, Sinauer Associates, Sunderland, Mass, USA, 2008.

[2] M. Parisi, R. Nuttall, D. Naiman et al., "Paucity of genes on the Drosophila X chromosome showing male-biased expression," *Science*, vol. 299, no. 5607, pp. 697–700, 2003.

[3] J. M. Ranz, C. I. Castillo-Davis, C. D. Meiklejohn, and D. L. Hartl, "Sex-dependent gene expression and evolution of the Drosophila transcriptome," *Science*, vol. 300, no. 5626, pp. 1742–1745, 2003.

[4] D. Sturgill, Y. Zhang, M. Parisi, and B. Oliver, "Demasculinization of X chromosomes in the Drosophila genus," *Nature*, vol. 450, no. 7167, pp. 238–241, 2007.

[5] V. Reinke, I. S. Gil, S. Ward, and K. Kazmer, "Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans," *Development*, vol. 131, no. 2, pp. 311–323, 2004.

[6] V. Reinke, H. E. Smith, J. Nance et al., "A global profile of germline gene expression in C. elegans," *Molecular Cell*, vol. 6, no. 3, pp. 605–616, 2000.

[7] D. Charlesworth, "Plant sex determination and sex chromosomes," *Heredity*, vol. 88, no. 2, pp. 94–101, 2002.

[8] D. Charlesworth, "Plant sex chromosomes," *Genome Dynamics*, vol. 4, pp. 83–94, 2008.

[9] M. Schartl, "Sex chromosome evolution in non-mammalian vertebrates," *Current Opinion in Genetics and Development*, vol. 14, no. 6, pp. 634–641, 2004.

[10] J. A. Marshall Graves, "Weird animal genomes and the evolution of vertebrate sex and sex chromosomes," *Annual Review of Genetics*, vol. 42, pp. 565–586, 2008.

[11] E. Betrán, K. Thornton, and M. Long, "Retroposed new genes out of the X in Drosophila," *Genome Research*, vol. 12, no. 12, pp. 1854–1859, 2002.

[12] Y. Bai, C. Casola, C. Feschotte, and E. Betrán, "Comparative genomics reveals a constant rate of origination and convergent acquisition of functional retrogenes in Drosophila," *Genome Biology*, vol. 8, no. 1, article R11, 2007.

[13] J. J. Emerson, H. Kaessmann, E. Betrán, and M. Long, "Extensive gene traffic on the mammalian X chromosome," *Science*, vol. 303, no. 5657, pp. 537–540, 2004.

[14] R. P. Meisel, M. V. Han, and M. W. Hahn, "A complex suite of forces drives gene traffic from Drosophila X chromosomes," *Genome Biology and Evolution*, vol. 1, pp. 176–188, 2009.

[15] M. D. Vibranovski, Y. Zhang, and M. Long, "General gene movement off the X chromosome in the Drosophila genus," *Genome Research*, vol. 19, no. 5, pp. 897–903, 2009.

[16] L. Potrzebowski, N. Vinckenbosch, A. C. Marques, F. Chalmel, B. Jégou, and H. Kaessmann, "Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes," *PLoS Biology*, vol. 6, no. 4, p. e80, 2008.

[17] M. Gallach and E. Betrán, "Intralocus sexual conflict resolved through gene duplication," *Trends in Ecology & Evolution*, vol. 26, pp. 222–228, 2011.

[18] I. Marn, M. L. Siegal, and B. S. Baker, "The evolution of dosage-compensation mechanisms," *BioEssays*, vol. 22, no. 12, pp. 1106–1114, 2000.

[19] R. Bergero and D. Charlesworth, "The evolution of restricted recombination in sex chromosomes," *Trends in Ecology and Evolution*, vol. 24, no. 2, pp. 94–102, 2009.

[20] W. R. Rice, "Evolution of the Y sex chromosome in animals," *BioScience*, vol. 46, no. 5, pp. 331–343, 1996.

[21] B. T. Lahn and D. C. Page, "Four evolutionary strata on the human X chromosome," *Science*, vol. 286, no. 5441, pp. 964–967, 1999.

[22] D. Bachtrog, "The temporal dynamics of processes underlying Y chromosome degeneration," *Genetics*, vol. 179, no. 3, pp. 1513–1525, 2008.

[23] B. Charlesworth, "The evolution of sex chromosomes," *Science*, vol. 251, no. 4997, pp. 1030–1033, 1991.

[24] B. S. Baker, M. Gorman, and I. Marin, "Dosage compensation in Drosophila," *Annual Review of Genetics*, vol. 28, pp. 491–521, 1994.

[25] T. W. Cline and B. J. Meyer, "Vive la différence: males vs Females in flies vs worms," *Annual Review of Genetics*, vol. 30, pp. 637–702, 1996.

[26] D. K. Nguyen and C. M. Disteche, "Dosage compensation of the active X chromosome in mammals," *Nature Genetics*, vol. 38, no. 1, pp. 47–53, 2006.

[27] W. R. Rice, "Sex chromosomes and the evolution of sexual dimorphism," *Evolution*, vol. 38, no. 4, pp. 735–742, 1984.

[28] J. C. Lucchesi, W. G. Kelly, and B. Panning, "Chromatin remodeling in dosage compensation," *Annual Review of Genetics*, vol. 39, pp. 615–651, 2005.

[29] V. Gupta, M. Parisi, D. Sturgill et al., "Global analysis of X-chromosome dosage compensation," *Journal of Biology*, vol. 5, article 3, 2006.

[30] T. Straub and P. B. Becker, "Dosage compensation: the beginning and end of generalization," *Nature Reviews Genetics*, vol. 8, no. 1, pp. 47–57, 2007.

[31] J. E. Mank, "The W, X, Y and Z of sex-chromosome dosage compensation," *Trends in Genetics*, vol. 25, no. 5, pp. 226–233, 2009.

[32] H. Ellegren, L. Hultin-Rosenberg, B. Brunstrom, L. Dencker, K. Kultima, and B. Scholz, "Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes," *BMC Biology*, vol. 5, article 40, 2007.

[33] Y. Itoh, E. Melamed, X. Yang et al., "Dosage compensation is less effective in birds than in mammals," *Journal of Biology*, vol. 6, no. 1, article 2, 2007.

[34] J. E. Mank and H. Ellegren, "All dosage compensation is local: gene-by-gene regulation of sex-biased expression on the chicken Z chromosome," *Heredity*, vol. 102, no. 3, pp. 312–320, 2009.

[35] X. Zha, Q. Xia, J. Duan, C. Wang, N. He, and Z. Xiang, "Dosage analysis of Z chromosome genes using microarray in silkworm, Bombyx mori," *Insect Biochemistry and Molecular Biology*, vol. 39, no. 5-6, pp. 315–321, 2009.

[36] E. G. Prince, D. Kirkland, and J. P. Demuth, "Hyperexpression of the X chromosome in both sexes results in extensive female bias of X-linked genes in the flour beetle," *Genome Biology and Evolution*, vol. 2, no. 1, pp. 336–346, 2010.

[37] C. Richler, H. Soreq, and J. Wahrman, "X inactivation in mammalian testis is correlated with inactive X-specific transcription," *Nature genetics*, vol. 2, no. 3, pp. 192–195, 1992.

[38] C. Richler, G. Ast, R. Goitein, J. Wahrman, R. Sperling, and J. Sperling, "Splicing components are excluded from the transcriptionally inactive XY body in male meiotic nuclei," *Molecular Biology of the Cell*, vol. 5, no. 12, pp. 1341–1352, 1994.

[39] J. M. A. Turner, "Meiotic sex chromosome inactivation," *Development*, vol. 134, no. 10, pp. 1823–1831, 2007.

[40] J. R. McCarrey, C. Watson, J. Atencio et al., "X-chromosome inactivation during spermatogenesis is regulated by an Xist/Tsix-independent mechanism in the mouse," *Genesis*, vol. 34, no. 4, pp. 257–266, 2002.

[41] J. L. Hornecker, P. B. Samollow, E. S. Robinson, J. L. Vandeberg, and J. R. McCarrey, "Meiotic sex chromosome inactivation in the marsupial Monodelphis domestica," *Genesis*, vol. 45, no. 11, pp. 696–708, 2007.

[42] S. H. Namekawa, J. L. VandeBerg, J. R. McCarrey, and J. T. Lee, "Sex chromosome silencing in the marsupial male germ line," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 23, pp. 9730–9735, 2007.

[43] M. Reuben and R. Lin, "Germline X chromosomes exhibit contrasting patterns of histone H3 methylation in caenorhabditis elegans," *Developmental Biology*, vol. 245, no. 1, pp. 71–82, 2002.

[44] E. Lifschytz and D. L. Lindsley, "The role of X-chromosome inactivation during spermatogenesis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 69, no. 1, pp. 182–186, 1972.

[45] L. Rastelli and M. I. Kuroda, "An analysis of maleless and histone H4 acetylation in Drosophila melanogaster spermatogenesis," *Mechanisms of Development*, vol. 71, no. 1-2, pp. 107–117, 1998.

[46] D. W. Rogers, M. Carr, and A. Pomiankowski, "Male genes: X-pelled or X-cluded?" *BioEssays*, vol. 25, no. 8, pp. 739–741, 2003.

[47] W. Hense, J. F. Baines, and J. Parsch, "X chromosome inactivation during Drosophila spermatogenesis," *PLoS Biology*, vol. 5, no. 10, p. e273, 2007.

[48] H. D. Hoyle, J. A. Hutchens, F. R. Turner, and E. C. Raff, "Regulation of β-tubulin function and expression in Drosophila spermatogenesis," *Developmental Genetics*, vol. 16, no. 2, pp. 148–170, 1995.

[49] M. D. Vibranovski, H. F. Lopes, T. L. Karr, and M. Long, "Stage-specific expression profiling of Drosophila spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes," *PLoS Genetics*, vol. 5, no. 11, Article ID e1000731, 2009.

[50] L. M. Mikhaylova and D. I. Nurminsky, "Lack of global meiotic sex chromosome inactivation, and paucity of tissue-specific gene expression on the Drosophila X chromosome," *BMC Biology*, vol. 9, p. 29, 2011.

[51] S. Schoenmakers, E. Wassenaar, J. W. Hoogerbrugge, J. S. E. Laven, J. A. Grootegoed, and W. M. Baarends, "Female meiotic sex chromosome inactivation in chicken," *PLoS Genetics*, vol. 5, no. 5, Article ID e1000466, 2009.

[52] J. Schimenti, "Synapsis or silence," *Nature Genetics*, vol. 37, no. 1, pp. 11–13, 2005.

[53] J. M. Turner, S. K. Mahadevaiah, O. Fernandez-Capetillo et al., "Silencing of unsynapsed meiotic chromosomes in the mouse," *Nature Genetics*, vol. 37, no. 1, pp. 41–47, 2005.

[54] J. T. Lee, "Sex chromosome inactivation: the importance of pairing," *Current Biology*, vol. 15, no. 7, pp. R249–R252, 2005.

[55] G. S. van Doorn, "Intralocus sexual conflict," *Annals of the New York Academy of Sciences*, vol. 1168, pp. 52–71, 2009.

[56] J. E. Mank, "Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome," *American Naturalist*, vol. 173, no. 2, pp. 141–150, 2009.

[57] J. E. Mank and H. Ellegren, "Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds," *Evolution*, vol. 63, no. 6, pp. 1464–1472, 2009.

[58] J. R. McCarrey and K. Thomas, "Human testis-specific PGK gene lacks introns and possesses characteristics of a processed gene," *Nature*, vol. 326, no. 6112, pp. 501–505, 1987.

[59] L. Zhong and J. M. Belote, "The testis-specific proteasome subunit Prosalpha6T of D. melanogaster is required for individualization and nuclear maturation during spermatogenesis," *Development*, vol. 134, no. 19, pp. 3517–3525, 2007.

[60] A. McLysaght, "Evolutionary steps of sex chromosomes are reflected in retrogenes," *Trends in Genetics*, vol. 24, no. 10, pp. 478–481, 2008.

[61] M. Parisi, R. Nuttall, P. Edwards et al., "A survey of ovary-, testis-, and soma-biased gene expression in Drosophila melanogaster adults," *Genome Biology*, vol. 5, no. 6, p. R40, 2004.

[62] J. L. Mueller, S. K. Mahadevaiah, P. J. Park, P. E. Warburton, D. C. Page, and J. M. A. Turner, "The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression," *Nature Genetics*, vol. 40, no. 6, pp. 794–799, 2008.

[63] P. P. Khil, N. A. Smirnova, P. J. Romanienko, and R. D. Camerini-Otero, "The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation," *Nature Genetics*, vol. 36, no. 6, pp. 642–646, 2004.

[64] M. Gallach, C. Chandrasekaran, and E. Betrán, "Analyses of nuclearly-encoded mitochondrial genes suggest gene duplication as a mechanism for resolving intralocus sexual antagonistic conflict in Drosophila," *Genome Biology and Evolution*, vol. 2, pp. 835–850, 2010.

[65] R. Kalamegham, D. Sturgill, E. Siegfried, and B. Oliver, "Drosophila mojoless, a retroposed GSK-3, has functionally diverged to acquire an essential role in male fertility," *Molecular Biology and Evolution*, vol. 24, no. 3, pp. 732–742, 2007.

[66] P. V. Danshina, C. B. Geyer, Q. Dai et al., "Phosphoglycerate kinase 2 (PGK2) is essential for sperm function and male fertility in mice," *Biology of Reproduction*, vol. 82, no. 1, pp. 136–145, 2010.

[67] E. Betrán, Y. Bai, and M. Motiwale, "Fast protein evolution and germ line expression of a Drosophila parental gene and its young retroposed paralog," *Molecular Biology and Evolution*, vol. 23, no. 11, pp. 2191–2202, 2006.

[68] E. Betrán and M. Long, "*Dntf-2r*: a young Drosophila retroposed gene with specific male expression under positive Darwinian selection," *Genetics*, vol. 164, no. 3, pp. 977–988, 2003.

[69] C. Tracy, J. Rio, M. Motiwale, S. M. Christensen, and E. Betrán, "Convergently recruited nuclear transport retrogenes are male biased in expression and evolving under positive selection in Drosophila," *Genetics*, vol. 184, no. 4, pp. 1067–1076, 2010.

[70] J. E. Quezada-Diaz, T. Muliyil, J. Río, and E. Betrán, "Drcd-1 related: a positively selected spermatogenesis retrogene in Drosophila," *Genetica*, vol. 138, no. 9, pp. 925–937, 2010.

[71] T. Connallon and L. L. Knowles, "Intergenomic conflict revealed by patterns of sex-biased gene expression," *Trends in Genetics*, vol. 21, no. 9, pp. 495–499, 2005.

[72] B. Vicoso and B. Charlesworth, "The deficit of male-biased genes on the D. melanogaster X chromosome is expression-dependent: a consequence of dosage compensation?" *Journal of Molecular Evolution*, vol. 68, no. 5, pp. 576–583, 2009.

[73] D. Bachtrog, N. R. T. Toda, and S. Lockton, "Dosage compensation and demasculinization of X chromosomes in Drosophila," *Current Biology*, vol. 20, no. 16, pp. 1476–1481, 2010.

[74] J. E. Mank, K. Nam, B. Brunström, and H. Ellegren, "Ontogenetic complexity of sexual dimorphism and sex-specific selection," *Molecular Biology and Evolution*, vol. 27, no. 7, pp. 1570–1578, 2010.

[75] Y. Zhang and B. Oliver, "An evolutionary consequence of dosage compensation on Drosophila melanogaster female X-chromatin structure?" *BMC Genomics*, vol. 11, no. 1, article 6, 2010.

[76] S. A. Vemuganti, T. A. Bell, C. O. Scarlett, C. E. Parker, F. de Villena, and D. A. O'Brien, "Three male germline-specific aldolase A isozymes are generated by alternative splicing and retrotransposition," *Developmental Biology*, vol. 309, no. 1, pp. 18–31, 2007.

[77] K. C. Kleene, "A possible meiotic function of the peculiar patterns of gene expression in mammalian spermatogenic cells," *Mechanisms of Development*, vol. 106, no. 1-2, pp. 3–23, 2001.

[78] A. Pischedda and A. K. Chippindale, "Intralocus sexual conflict diminishes the benefits of sexual selection," *PLoS Biology*, vol. 4, no. 11, p. e356, 2006.

[79] C.-I. Wu and E. Y. Xu, "Sexual antagonism and X inactivation—the SAXI hypothesis," *Trends in Genetics*, vol. 19, no. 5, pp. 243–247, 2003.

[80] R. Bonduriansky and S. F. Chenoweth, "Intralocus sexual conflict," *Trends in Ecology and Evolution*, vol. 24, no. 5, pp. 280–288, 2009.

[81] W. R. Rice and A. K. Chippindale, "The evolution of hybrid infertility: perpetual coevolution between gender-specific and sexually antagonistic genes," *Genetica*, vol. 116, no. 2-3, pp. 179–188, 2002.

[82] E. S. Kelleher, N. L. Clark, and T. A. Markow, "Diversity enhancing selection acts on a female reproductive protease family in four sub-species of Drosophila mojavensis," *Genetics*, vol. 187, pp. 865–876, 2011.

[83] B. J. Wagstaff and D. J. Begun, "Adaptive evolution of recently duplicated accessory gland protein genes in desert Drosophila," *Genetics*, vol. 177, no. 2, pp. 1023–1030, 2007.

[84] W. Haerty, S. Jagadeeshan, R. J. Kulathinal et al., "Evolution in the fast lane: rapidly evolving sex-related genes in Drosophila," *Genetics*, vol. 177, no. 3, pp. 1321–1335, 2007.

[85] L. M. Mikhaylova, K. Nguyen, and D. I. Nurminsky, "Analysis of the Drosophila melanogaster testes transcriptome reveals coordinate regulation of paralogous genes," *Genetics*, vol. 179, no. 1, pp. 305–315, 2008.

[86] B. T. Lahn, N. M. Pearson, and K. Jegalian, "The human Y chromosome, in the light of evolution," *Nature Reviews Genetics*, vol. 2, no. 3, pp. 207–216, 2001.

[87] R. A. Fisher, "The evolution of dominance," *Biological Reviews*, vol. 6, pp. 345–368, 1931.

[88] L. B. Koerich, X. Wang, A. G. Clark, and A. B. Carvalho, "Low conservation of gene content in the Drosophila Y chromosome," *Nature*, vol. 456, no. 7224, pp. 949–951, 2008.

[89] P. Innocenti and E. H. Morrow, "The sexually antagonistic genes of Drosophila melanogaster," *PLoS Biology*, vol. 8, no. 3, Article ID e1000335, 2010.

[90] H. Ellegren and J. Parsch, "The evolution of sex-biased genes and sex-biased gene expression," *Nature Reviews Genetics*, vol. 8, no. 9, pp. 689–698, 2007.

[91] B. Lemos, C. R. Landry, P. Fontanillas et al., "Evolution of genomic expression," in *Evolutionary Genomics and Proteomics*, M. Pagel and A. Pomiankowski, Eds., pp. 81–118, Sinauer Associates, Sunderland, Mass, USA, 2008.

[92] P. A. t Hoen, Y. Ariyurek, H. H. Thygesen et al., "Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms," *Nucleic Acids Research*, vol. 36, no. 21, article e141, 2008.

[93] A. Whitehead and D. L. Crawford, "Variation within and among species in gene expression: raw material for evolution," *Molecular Ecology*, vol. 15, no. 5, pp. 1197–1211, 2006.

[94] F. Gnad and J. Parsch, "Sebida: a database for the functional and evolutionary analysis of genes with sex-biased expression," *Bioinformatics*, vol. 22, no. 20, pp. 2577–2579, 2006.

[95] Z. F. Jiang and C. A. Machado, "Evolution of sex-dependent gene expression in three recently diverged species of Drosophila," *Genetics*, vol. 183, no. 3, pp. 1175–1185, 2009.

[96] W. Jin, R. M. Riley, R. D. Wolfinger, K. P. White, G. Passador-Gurgell, and G. Gibson, "The contributions of sex, genotype and age to transcriptional variance in Drosophila melanogaster," *Nature Genetics*, vol. 29, no. 4, pp. 389–395, 2001.

[97] Y. Zhang, D. Sturgill, M. Parisi, S. Kumar, and B. Oliver, "Constraint and turnover in sex-biased gene expression in the genus Drosophila," *Nature*, vol. 450, no. 7167, pp. 233–237, 2007.

[98] J. F. Ayroles, M. A. Carbone, E. A. Stone et al., "Systems genetics of complex traits in Drosophila melanogaster," *Nature Genetics*, vol. 41, no. 3, pp. 299–307, 2009.

[99] B. R. Graveley, A. N. Brooks, J. W. Carlson et al., "The developmental transcriptome of Drosophila melanogaster," *Nature*, vol. 471, pp. 473–479, 2010.

[100] B. Daines, H. Wang, L. Wang et al., "The Drosophila melanogaster transcriptome by paired-end RNA sequencing," *Genome Research*, vol. 21, pp. 315–324, 2011.

[101] M. W. Hahn and G. C. Lanzaro, "Female-biased gene expression in the malaria mosquito Anopheles gambiae," *Current Biology*, vol. 15, no. 6, pp. R192–R193, 2005.

[102] J. E. Mank, L. Hultin-Rosenberg, M. T. Webster, and H. Ellegren, "The unique genomic properties of sex-biased genes: insights from avian microarray data," *BMC Genomics*, vol. 9, article 148, 2008.

[103] X. Yang, E. E. Schadt, S. Wang et al., "Tissue-specific expression and regulation of sexually dimorphic genes in mice," *Genome Research*, vol. 16, no. 8, pp. 995–1004, 2006.

[104] J. H. Malone, D. L. Hawkins, and P. Michalak, "Sex-biased gene expression in a ZW sex determination system," *Journal of Molecular Evolution*, vol. 63, no. 4, pp. 427–436, 2006.

[105] C. M. Small, G. E. Carney, Q. Mo, M. Vannucci, and A. G. Jones, "A microarray analysis of sex- and gonad-biased gene expression in the zebrafish: evidence for masculinization of the transcriptome," *BMC Genomics*, vol. 10, article 579, 2009.

[106] C. D. Meiklejohn, J. Parsch, J. M. Ranz, and D. L. Hartl, "Rapid evolution of male-biased gene expression in Drosophila," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 17, pp. 9894–9899, 2003.

[107] P. Khaitovich, I. Hellmann, W. Enard et al., "Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees," *Science*, vol. 309, no. 5742, pp. 1850–1854, 2005.

[108] C. Voolstra, D. Tautz, P. Farbrother, L. Eichinger, and B. Harr, "Contrasting evolution of expression differences in the testis between species and subspecies of the house mouse," *Genome Research*, vol. 17, no. 1, pp. 42–49, 2007.

[109] T. Connallon and A. G. Clark, "Sex linkage, sex-specific selection, and the role of recombination in the evolution of sexually dimorphic gene expression," *Evolution*, vol. 64, no. 12, pp. 3417–3442, 2010.

[110] M. Pröschel, Z. Zhang, and J. Parsch, "Widespread adaptive evolution of Drosophila genes with sex-biased expression," *Genetics*, vol. 174, no. 2, pp. 893–900, 2006.

[111] J. M. Good, C. A. Hayden, and T. J. Wheeler, "Adaptive protein evolution and regulatory divergence in Drosophila," *Molecular Biology and Evolution*, vol. 23, no. 6, pp. 1101–1103, 2006.

[112] S. V. Nuzhdin, M. L. Wayne, K. L. Harmon, and L. M. McIntyre, "Common pattern of evolution of gene expression level and protein sequence in Drosophila," *Molecular Biology and Evolution*, vol. 21, no. 7, pp. 1308–1317, 2004.

[113] B. Lemos, B. R. Bettencourt, C. D. Meiklejohn, and D. L. Hartl, "Evolution of proteins and gene expression levels are coupled in Drosophila and are independently associated with mRNA abundance, protein length, and number of protein-protein interactions," *Molecular Biology and Evolution*, vol. 22, no. 5, pp. 1345–1354, 2005.

[114] C. G. Artieri, W. Haerty, and R. S. Singh, "Association between levels of coding sequence divergence and gene misregulation in Drosophila male hybrids," *Journal of Molecular Evolution*, vol. 65, no. 6, pp. 697–704, 2007.

[115] P. Khaitovich, S. Paabo, and G. Weiss, "Toward a neutral evolutionary model of gene expression," *Genetics*, vol. 170, no. 2, pp. 929–939, 2005.

[116] J. E. Mank, L. Hultin-Rosenberg, M. Zwahlen, and H. Ellegren, "Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression," *American Naturalist*, vol. 171, no. 1, pp. 35–43, 2008.

[117] R. P. Meisel, "Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein coding sequence evolution," *Molecular Biology and Evolution*, vol. 28, no. 6, pp. 1893–1900, 2011.

[118] T. M. Williams and S. B. Carroll, "Genetic and molecular insights into the development and evolution of sexual dimorphism," *Nature Reviews Genetics*, vol. 10, no. 11, pp. 797–804, 2009.

[119] J. M. Ranz and C. A. Machado, "Uncovering evolutionary patterns of gene expression using microarrays," *Trends in Ecology and Evolution*, vol. 21, no. 1, pp. 29–37, 2006.

[120] K. A. Hughes, J. F. Ayroles, M. M. Reedy et al., "Segregating variation in the transcriptome: cis regulation and additivity of effects," *Genetics*, vol. 173, no. 3, pp. 1347–1355, 2006.

[121] A. Genissel, L. M. McIntyre, M. L. Wayne, and S. V. Nuzhdin, "Cis and trans regulatory effects contribute to natural variation in transcriptome of Drosophila melanogaster," *Molecular Biology and Evolution*, vol. 25, no. 1, pp. 101–110, 2008.

[122] P. J. Wittkopp, B. K. Haerum, and A. G. Clark, "Regulatory changes underlying expression differences within and between Drosophila species," *Nature Genetics*, vol. 40, no. 3, pp. 346–350, 2008.

[123] P. J. Wittkopp, B. K. Haerum, and A. G. Clark, "Evolutionary changes in cis and trans gene regulation," *Nature*, vol. 430, no. 6995, pp. 85–88, 2004.

[124] R. M. Graze, L. M. McIntyre, B. J. Main, M. L. Wayne, and S. V. Nuzhdin, "Regulatory divergence in Drosophila melanogaster and D. simulans, a genomewide analysis of allele-specific expression," *Genetics*, vol. 183, no. 2, pp. 547–561, 2009.

[125] C. J. McManus, J. D. Coolon, M. O. Duff, J. Eipper-Mains, B. R. Graveley, and P. J. Wittkopp, "Regulatory divergence in Drosophila revealed by mRNA-seq," *Genome Research*, vol. 20, no. 6, pp. 816–825, 2010.

[126] L. M. McIntyre, L. M. Bono, A. Genissel et al., "Sex-specific expression of alternative transcripts in Drosophila," *Genome Biology*, vol. 7, no. 8, article R79, 2006.

[127] M. Telonis-Scott, A. Kopp, M. L. Wayne, S. V. Nuzhdin, and L. M. McIntyre, "Sex-specific splicing in Drosophila: widespread occurrence, tissue specificity and evolutionary conservation," *Genetics*, vol. 181, no. 2, pp. 421–434, 2009.

[128] R. Blekhman, J. C. Marioni, P. Zumbo, M. Stephens, and Y. Gilad, "Sex-specific and lineage-specific alternative splicing in primates," *Genome Research*, vol. 20, no. 2, pp. 180–189, 2010.

[129] M. Long and C. H. Langley, "Natural selection and the origin of jingwei, a chimeric processed functional gene in Drosophila," *Science*, vol. 259, no. 5104, pp. 91–95, 1993.

[130] R. Saxena, L. G. Brown, T. Hawkins et al., "The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned," *Nature Genetics*, vol. 14, no. 3, pp. 292–299, 1996.

[131] H. Innan and F. Kondrashov, "The evolution of gene duplications: classifying and distinguishing between models," *Nature Reviews Genetics*, vol. 11, no. 2, pp. 97–108, 2010.

[132] P. J. Wang, J. R. McCarrey, F. Yang, and D. C. Page, "An abundance of X-linked genes expressed in spermatogonia," *Nature Genetics*, vol. 27, no. 4, pp. 422–426, 2001.

[133] M. Gallach, V. Arnau, and I. Marin, "Global patterns of sequence evolution in Drosophila," *BMC Genomics*, vol. 8, p. 408, 2007.

[134] S. Ramos-Onsins and M. Aguade, "Molecular evolution of the Cecropin multigene family in Drosophila: functional genes vs. pseudogenes," *Genetics*, vol. 150, no. 1, pp. 157–171, 1998.

[135] K. Tamura, S. Subramanian, and S. Kumar, "Temporal patterns of fruit fly (Drosophila) evolution revealed by mutation clocks," *Molecular Biology and Evolution*, vol. 21, no. 1, pp. 36–44, 2004.

[136] Y. E. Zhang, M. D. Vibranovski, P. Landback, G. A. B. Marais, and M. Long, "Chromosomal redistribution of male-biased genes in mammalian evolution with two bursts of gene gain on the X chromosome," *PLoS Biology*, vol. 8, no. 10, Article ID e1000494, 2010.

[137] Y. E. Zhang, M. D. Vibranovski, B. H. Krinsky, and M. Long, "Age-dependent chromosomal distribution of male-biased genes in Drosophila," *Genome Research*, vol. 20, no. 11, pp. 1526–1533, 2010.

[138] M. T. Levine, C. D. Jones, A. D. Kern, H. A. Lindfors, and D. J. Begun, "Novel genes derived from noncoding DNA in Drosophila melanogaster are frequently X-linked and exhibit testis-biased expression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 26, pp. 9935–9939, 2006.

[139] V. B. Kaiser and H. Ellegren, "Nonrandom distribution of genes with sex-biased expression in the chicken genome," *Evolution*, vol. 60, no. 9, pp. 1945–1951, 2006.

[140] R. Storchova and P. Divina, "Nonrandom representation of sex-biased genes on chicken Z chromosome," *Journal of Molecular Evolution*, vol. 63, no. 5, pp. 676–681, 2006.

[141] L. Mořkovýky, R. Storchova, J. Plachy et al., "The chicken Z chromosome is enriched for genes with preferential expression in ovarian somatic cells," *Journal of Molecular Evolution*, vol. 70, no. 2, pp. 129–136, 2009.

[142] K. P. Arunkumar, K. Mita, and J. Nagaraju, "The silkworm Z chromosome is enriched in testis-specific genes," *Genetics*, vol. 182, no. 2, pp. 493–501, 2009.

[143] B. Oliver, "Genetic control of germline sexual dimorphism in Drosophila," *International Review of Cytology*, vol. 219, pp. 1–60, 2002.

[144] W. G. Kelly, C. E. Schaner, A. F. Dernburg et al., "X-chromosome silencing in the germline of C. elegans," *Development*, vol. 129, no. 2, pp. 479–492, 2002.

[145] J. R. Gibson, A. K. Chippindale, and W. R. Rice, "The X chromosome is a hot spot for sexually antagonistic fitness variation," *Proceedings of the Royal Society B*, vol. 269, no. 1490, pp. 499–505, 2002.

[146] J. Bradley, A. Baltus, H. Skaletsky, M. Royce-Tolland, K. Dewar, and D. C. Page, "An X-to-autosome retrogene is required for spermatogenesis in mice," *Nature Genetics*, vol. 36, no. 8, pp. 872–876, 2004.

[147] L. W. Hillier, W. Miller, E. Birney et al., "Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution," *Nature*, vol. 432, pp. 695–716, 2004.

[148] P. Fontanillas, D. L. Hartl, and M. Reuter, "Genome organization and gene expression shape the transposable element distribution in the Drosophila melanogaster euchromatin," *PLoS Genetics*, vol. 3, no. 11, pp. 2256–2267, 2007.

[149] C. Esnault, J. Maestre, and T. Heidmann, "Human LINE retrotransposons generate processed pseudogenes," *Nature Genetics*, vol. 24, no. 4, pp. 363–367, 2000.

[150] T. Connallon and A. G. Clark, "The resolution of sexual antagonism by gene duplication," *Genetics*, vol. 187, pp. 919–937, 2011.

[151] M. N. Arbeitman, E. E. Furlong, F. Imam et al., "Gene expression during the life cycle of Drosophila melanogaster," *Science*, vol. 297, no. 5590, pp. 2270–2275, 2002.

[152] E. S. Kelleher and T. A. Markow, "Duplication, selection and gene conversion in a Drosophila mojavensis female reproductive protein family," *Genetics*, vol. 181, no. 4, pp. 1451–1465, 2009.

[153] E. S. Kelleher, W. J. Swanson, and T. A. Markow, "Gene duplication and adaptive evolution of digestive proteases in Drosophila arizonae female reproductive tracts," *PLoS Genetics*, vol. 3, no. 8, p. e148, 2007.

[154] N. Vinckenbosch, I. Dupanloup, and H. Kaessmann, "Evolutionary fate of retroposed gene copies in the human genome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 9, pp. 3220–3225, 2006.

[155] B. Loppin, D. Lepetit, S. Dorus, P. Couble, and T. L. Karr, "Origin and neofunctionalization of a Drosophila paternal effect gene essential for zygote viability," *Current Biology*, vol. 15, no. 2, pp. 87–93, 2005.

[156] Y. Bai, C. Casola, and E. Betrán, "Evolutionary origin of regulatory regions of retrogenes in Drosophila," *BMC Genomics*, vol. 9, article 241, 2008.

[157] S. R. Proulx and P. C. Phillips, "Allelic divergence precedes and promotes gene duplication," *Evolution*, vol. 60, no. 5, pp. 881–892, 2006.

[158] F. Zhang, W. Gu, M. E. Hurles, and J. R. Lupski, "Copy number variation in human health, disease, and evolution," *Annual Review of Genomics and Human Genetics*, vol. 10, pp. 451–481, 2009.

[159] T. Domazet-Loso and D. Tautz, "A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns," *Nature*, vol. 468, no. 7325, pp. 815–819, 2010.

[160] M. Long, E. Betrán, K. Thornton, and W. Wang, "The origin of new genes: glimpses from the young and old," *Nature Reviews Genetics*, vol. 4, no. 11, pp. 865–875, 2003.

[161] S. Ohno, *Evolution by Gene Duplication*, Springer, New York, NY, USA, 1970.

[162] S. Matsunaga, E. Isono, E. Kejnovsky et al., "Duplicative transfer of a MADS box gene to a plant Y chromosome," *Molecular Biology and Evolution*, vol. 20, no. 7, pp. 1062–1069, 2003.

[163] H. Skaletsky, T. Kuroda-Kawaguchl, P. J. Minx et al., "The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes," *Nature*, vol. 423, no. 6942, pp. 825–837, 2003.

[164] A. B. Carvalho, B. P. Lazzaro, and A. G. Clark, "Y chromosomal fertility factors kl-2 and kl-3 of Drosophila melanogaster encode dynein heavy chain polypeptides," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 24, pp. 13239–13244, 2000.

[165] A. B. Carvalho, B. A. Dobo, M. D. Vibranovski, and A. G. Clark, "Identification of five new genes on the Y chromosome of Drosophila melanogaster," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 23, pp. 13225–13230, 2001.

[166] M. D. Vibranovski, L. B. Koerich, and A. B. Carvalho, "Two new Y-linked genes in Drosophila melanogaster," *Genetics*, vol. 179, no. 4, pp. 2325–2327, 2008.

[167] F. J. Krsticevic, H. L. Santos, S. Januario, C. G. Schrago, and A. B. Carvalho, "Functional copies of the Mst77F gene on the Y chromosome of Drosophila melanogaster," *Genetics*, vol. 184, no. 1, pp. 295–307, 2010.

[168] J. F. Hughes, H. Skaletsky, T. Pyntikova et al., "Chimpanzee and human y chromosomes are remarkably divergent in structure and gene content," *Nature*, vol. 463, no. 7280, pp. 536–539, 2010.

[169] S. Rozen, H. Skaletsky, J. D. Marszalek et al., "Abundant gene conversion between arms of palindromes in human and ape Y chromosomes," *Nature*, vol. 423, no. 6942, pp. 873–876, 2003.

[170] A. Geraldes, T. Rambo, R. A. Wing, N. Ferrand, and M. W. Nachman, "Extensive gene conversion drives the concerted evolution of paralogous copies of the SRY gene in European rabbits," *Molecular Biology and Evolution*, vol. 27, no. 11, pp. 2437–2440, 2010.

[171] E. L. C. Verkaar, C. Zijlstra, E. M. van 't Veld, K. Boutaga, D. C. J. van Boxtel, and J. A. Lenstra, "Organization and concerted evolution of the ampliconic Y-chromosomal TSPY genes from cattle," *Genomics*, vol. 84, no. 3, pp. 468–474, 2004.

[172] N. Backstrom, H. Ceplitis, S. Berlin, and H. Ellegren, "Gene conversion drives the evolution of HINTW, an ampliconic gene on the female-specific avian W chromosome," *Molecular Biology and Evolution*, vol. 22, no. 10, pp. 1992–1999, 2005.

[173] J. Rohozinski and C. E. Bishop, "The mouse juvenile spermatogonial depletion (jsd) phenotype is due to a mutation in the X-derived retrogene, mUtp14b," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 32, pp. 11695–11700, 2004.

[174] L. Potrzebowski, N. Vinckenbosch, and H. Kaessmann, "The emergence of new genes on the young therian X," *Trends in Genetics*, vol. 26, no. 1, pp. 1–4, 2010.

[175] R. Nielsen, C. Bustamante, A. G. Clark et al., "A scan for positively selected genes in the genomes of humans and chimpanzees," *PLoS Biology*, vol. 3, no. 6, p. e170, 2005.

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 972457, 7 pages
doi:10.4061/2011/972457

Review Article

Boule and the Evolutionary Origin of Metazoan Gametogenesis: A Grandpa's Tale

José M. Eirín-López[1] and Juan Ausió[2]

[1] *CHROMEVOL-XENOMAR Group, Departamento de Biología Celular y Molecular, Facultade de Ciencias, Universidade da Coruña, Campus de A Zapateira s/n, E15071 A Coruña, Spain*
[2] *Department of Biochemistry and Microbiology, University of Victoria, V8W 3P6 Victoria, BC, Canada V8W 3P6*

Correspondence should be addressed to José M. Eirín-López, jeirin@udc.es

Received 16 January 2011; Revised 18 April 2011; Accepted 9 May 2011

Academic Editor: Rob Kulathinal

The evolution of sex remains a hotly debated topic in evolutionary biology. In particular, studying the origins of the molecular mechanisms underlying sexual reproduction and gametogenesis (its fundamental component) in multicellular eukaryotes has been difficult due to the rapid divergence of many reproductive proteins, pleiotropy, and by the fact that only a very small number of reproductive proteins specifically involved in reproduction are conserved across lineages. Consequently, during the last decade, many efforts have been put into answering the following question: did gametogenesis evolve independently in different animal lineages or does it share a common evolutionary origin in a single ancestral prototype? Among the various approaches carried out in order to solve this question, the characterization of the evolution of the DAZ gene family holds much promise because these genes encode reproductive proteins that are conserved across a wide range of animal phyla. Within this family, *BOULE* is of special interest because it represents the most ancestral member of this gene family (the "grandfather" of *DAZ*). Furthermore, *BOULE* has attracted most of the attention since it represents an ancient male gametogenic factor with an essential reproductive-exclusive requirement in urbilaterians, constituting a core component of the reproductive prototype. Within this context, the aim of the present work is to provide an up-to-date insight into the studies that lead to the characterization of the DAZ family members and the implications in helping decipher the evolutionary origin of gametogenesis in metazoan animals.

1. Preliminary Considerations on the Evolution of Sexual Reproduction

The appearance of sexual reproduction constituted an important breakthrough with critical genetic, cellular, physiological, and evolutionary implications. This is mainly due to three reasons [1, 2]. Firstly, it provided a mechanism for DNA crossing-over and recombination during meiosis [3] leading to the generation of genetically diverse gametes [4, 5]. Second, it permitted the differentiation of a germinal cell lineage in multi-cellular eukaryotes, responsible for the generation of haploid gametes through a sequential process known as gametogenesis [6]. Third, sexual reproduction involved the differentiation of two sexes in which male- and female-specific gametes are generated by means of a sequential process involving sex determination, mitotic proliferation, meiosis, and gamete differentiation [7, 8].

The evolution of sexual reproduction has represented an important milestone in evolutionary biology due to its relevance to the genetic diversification of individuals within a species and its consequences for speciation. Many aspects of the unique cell division process of meiosis associated with sexual reproduction are well known to have been highly conserved across eukaryotes (i.e., the key components of the meiotic machinery for chromosomal synapses and recombination), sharing a common evolutionary origin [9–12]. However, the complexity of sexual reproduction goes beyond meiosis ultimately leading to the differentiation of sexually dimorphic male sperm and female eggs through a process known as gametogenesis. In metazoan animals, sex-specific gametes are produced, displaying distinct differentiation patterns that result in different size, motility and gamete morphologies. These are sex-specific or sex-biased, with no conserved male- or female-specific gametogenic

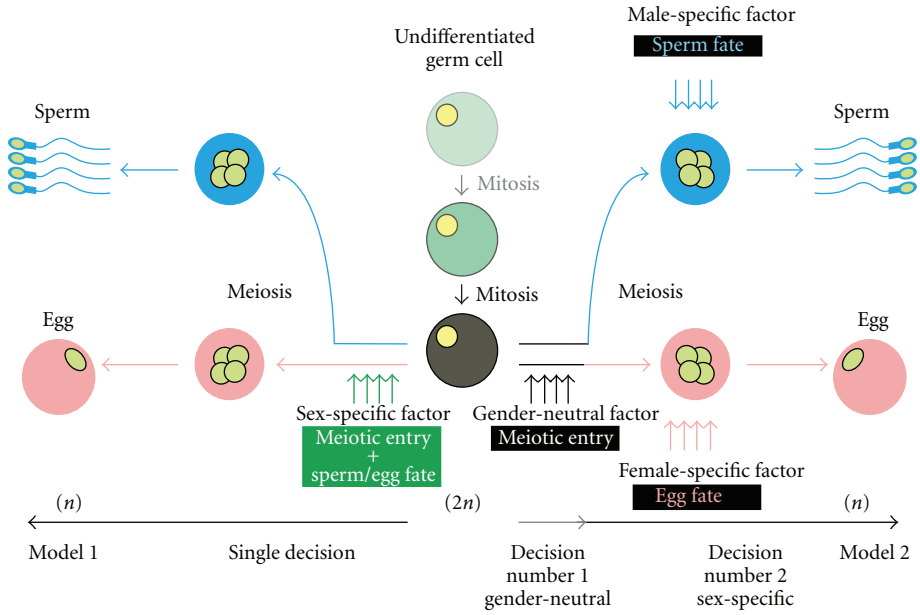


Figure 1: Schematic representation of the germ cell fate decisions following two alternative models (adapted from [13]) involving either a single sex-specific decision (model 1) or two decisions (gender neutral and sex specific, model 2). Uncommitted, male-committed, and female-committed germ cells are indicated in grey, blue, and pink, respectively.

factors [13, 14]. Furthermore, genes for sex-specific traits usually diverge very rapidly in males and females, and, with very few exceptions, only a very small number of reproductive proteins specifically involved in reproduction (broadly defined as those that act after copulation and that mediate gamete usage, storage, signal transduction, and fertilization) are conserved across lineages [15, 16]. Such dichotomy is in sharp contrast with the requirement of meiosis by both males and females, raising an important question as to what extent the features of sexual reproduction can be conserved. In other words, did gametogenesis evolve independently multiple times in different animal lineages (achieving a similar functional goal through convergent evolution) or can its origin be traced back to a single prototype sharing a common evolutionary origin (ensued by a rapid divergence of most of components of the reproductive machinery)? In this latter instance, it would be reasonable to expect a prevalence of a few core components from the ancient prototype [7].

Among the different studies attempting to address such an interesting issue (see [6, 16, 17] for review), those carried out by E. Y. Xu and collaborators during the last decade stand out due to the identification of a very conserved family of reproductive proteins (the deleted in azoospermia (DAZ) gene family) across a wide range of animal phyla. Particularly, the gene *BOULE* is of critical interest within this family since it seems to be invariably conserved both in protostome and deuterostome lineages during metazoan evolution. Furthermore, molecular evolutionary analyses indicate that *BOULE* escapes from the process of rapid adaptive evolution that is typically operating in many reproductive genes expressed postmeiotically [18] as well as in many premeiotic and meiotic genes [19–21]. Hence, *BOULE* has been the object of detailed studies, constituting a potentially conserved male-gametogenic requirement that provides support for a common evolutionary origin for spermatogenesis in metazoan animals [7].

2. Gametogenesis Involves Tough Decisions

During the early stages of gametogenesis (premeiotic gametogenesis), animal germ cells proliferate through mitotic divisions while kept on an undifferentiated state mediated by RNA-binding proteins [13]. During the process, the first critical decision that germ cells must take involves the arrest of mitotic divisions in order to enter into the meiosis stage that will lead to the subsequent formation of the haploid sex-specific gametes. To this end, germ cells must be informed (via somatic cell signaling) of whether they are in a male or in a female body in order to subsequently differentiate into male-specific spermatozoa or female-specific eggs, with this constituting a second critical decision. The timing of the first of these decisions (mitosis/meiosis) is often sex-specific and closely related with the second decision (sperm/egg). The molecular mechanisms underlying such critical choices during early (premeiotic) gametogenesis have been a subject of debate during the last ten years, leading to the proposal of two main models that differ in the steps involved in how the two basic decisions are made (Figure 1). The first model (model 1) involves a single sex-specific regulator responsible for both the entrance in meiosis and the determination of sperm or egg cell differentiation. A second model calls for the presence of a gender-neutral regulator responsible for the switch from mitosis to meiosis, as well as a second sex-specific regulator involved in the sperm/egg decision [22].

Recent reports have provided evidence favoring model 1 over model 2, arguing that RNA-binding proteins are responsible for keeping germ cells on an undifferentiated

state during mitotic proliferation [23–25], while maintaining the reversibility of such process that would allow for a transient regulation in a yet unknown fashion. Indeed, RNA-interference experiments on key regulators carried out in nematodes have revealed that sexual identity is labile, with adult females generating eggs that are amenable to switch and follow a spermatogenic development [26] and vice-versa [27]. Notably, the decision about the sexual fate of germ cells appears to be adopted at around the same time they exit mitosis to enter meiosis, underscoring the connection between the two processes [28]. On the other hand, while certain sex-specific decisions occur around the time when germline cells undergo meiosis, other decisions happen far earlier [29]. Although it is still premature to determine the universality of such mechanism, the model in which a conserved factor directs the main decisions within the core mechanism of germ cell development, may provide a very interesting insight into the evolution of the gametogenesis and sex reproduction processes.

Once the choices for mitosis/meiosis and male/female have been taken, germ cells undergo a postmeiotic cell differentiation leading to the formation of sex-specific gametes. Many of the structures and proteins appearing during these late stages seem to be related to the establishment of reproductive barriers that have critical implications for the process of speciation [30]. Therefore, it is not surprising that in many instances they evolve rapidly. However, it is important to note that although the alteration of any of these genes often leads to intraspecific infertility, none of the trends associated with them appear to be critical for proper zygote formation nor pose a roadblock for the differentiation process. This suggests that while genes involved in gamete specialization are very important for different aspects of intra-and interspecific fertilization [31] they are not the core determinants of gametogenesis.

3. The DAZ Gene Family Encodes Widely Conserved Gametogenic Factors

Although the major steps of dimorphic (male/female) gametogenesis in animals are very similar, the identification of male- and female-specific gametogenic factors common to all lineages of animals has shown to be anything but obvious [13, 14]. Reproductive proteins are often involved in other general cellular processes (besides reproduction itself), and, thus, the relevant question pertains as to whether such conservation is due to their crucial role in gametogenesis or whether it is the result of the pleoitropic functions they perform in their reproductive and somatic functions [32–34]. The key to solve this paradox probably lies in the identification of remnants of an ancient gametogenic core, lending support to a common origin of sexually dimorphic traits among animals, or in other words, the identification of conserved male- or female-specific gametogenic proteins across large evolutionary ranges. According to Xu and collaborators [7, 8], the components of such gametogenic core in the reproductive prototype should fulfill four requirements, including: (a) presence in most of major lineages of multicellular animals with sexual reproduction, (b) evolutionary origin at the same time as sexual reproduction, (c) conservation of sequence expression and function in different phyla; and most importantly (d) being selectively involved in gametogenesis in one of the two sexes and being excluded from any other pleiotropic process outside reproduction that could condition their evolutionary conservation.

Among all reproductive-associated genes known to date, the Deleted in AZoospermia (DAZ) gene family has probably been the group of reproductive factors fulfilling the above criteria that have attracted more interest from researchers during the last ten years. In humans, the DAZ gene family encompasses three genes referred to as *BOULE*, *DAZ*-like (*DAZL*), and *DAZ* encoding translational regulators with common features, including the presence of a RNA-binding domain with signature (ribonucleoprotein, RNP) RNP-1/RNP-2 motifs, as well as DAZ repeats rich in N, Y, and Q residues ([7, 8, 35], Figure 2(a)). The *BOULE* gene maps to chromosome 2 in humans and to chromosome 3 in mouse, at a region where a male sterile mutation is located which is syntenic to human chromosome 2 [8]. Human *BOULE* shares identical RNP-1 and RNP-2 motifs with fly and nematode *boule* [36, 37]), acting as a meiotic regulator expressed during late stages of male-specific meiosis (Figure 2(b)). Although defects in BOULE cause meiotic arrest predominantly in males, interference with meiosis in females has been also indicated in the nematode *boule* homolog known as *daz-1* [37]. In contrast, mutations on the *DAZL* gene (located at chromosome 3 in humans) interfere with both male and female germ cell development [8], with an expression pattern that begins early in development and continues through the meiotic divisions of gametogenesis (Figure 2(b)). Finally, the *DAZ* gene emerged on the Y chromosome in humans showing structural (protein 95% similar) and functional similarity to *DAZL* (Figure 2(b)). However, in contraposition to *DAZL*, *DAZ* expression is restricted to males [8, 35], albeit is not essential for the completion of spermatogenesis [38, 39], and its deletion is linked to infertility [7, 8].

The evolutionary conservation observed in the DAZ family members across different animal lineages, together with the lack of evidence of positive Darwinian selection acting on their members, have sparked the interest in this family as potential remnants of an ancient gametogenic core providing support to a common origin for gametogenesis in metazoans. However, and most importantly, the relevance of the study of the DAZ gene family within this context is further supported by the fact that all DAZ members are restricted to germ cells, eliminating the potential masking effects of pleiotropy in the study of the evolution of reproductive mechanisms.

4. The Conservation of *Boule* Supports a Common Evolutionary Origin of Gametogenesis from an Ancestral Prototype in Metazoans

Molecular evolutionary and phylogenetic analyses indicate that the *DAZ* gene family consists of two subfamilies

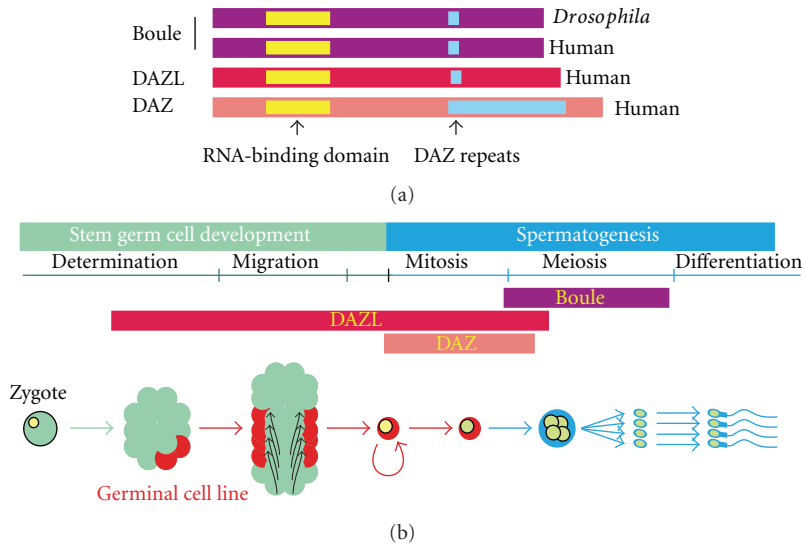


Figure 2: Molecular structure and expression patterns of DAZ gene family members during germ cell development and differentiation [8]. (a) diagram of *BOULE*, *DAZL*, and *DAZ* genes displaying the RNA-binding domains (yellow) and the DAZ repeats (blue). (b) expression patterns of *BOULE*, *DAZL*, and *DAZ* genes represented by horizontal lines during gametogenesis. *BOULE* is an ancient meiotic regulator conserved in all metazoans, giving rise to a gene family required for novel vertebrate germ cell functions. While *BOULE* encompasses a meiotic expression, *DAZL* and *DAZ* evolved novel premeiotic functions related to stem germ cell proliferation and differentiation later on during evolution.

(*DAZL* and *BOULE*) involved in different stages of germ cell development ([8], Figure 3). *BOULE* represents the ancestral single copy gene founder of the *DAZ* family, (the "grandfather" of *DAZ*) giving rise to *DAZL* (the "father" of *DAZ*) through gene duplication events most likely in the ancestral lineage of bony fishes before the emergence of tetrapods [7, 8]. A gene duplication process seems to be also responsible for the emergence of *DAZ* from *DAZL*, appearing in the Y chromosome of primates after the divergence between New World monkeys and Old World monkeys (Figure 3), approximately 30–40 MYA [8]. Y-linked *DAZ* went through two more gene duplication events as recently as 55,000 years ago, giving rise to a cluster of four *DAZ* genes [40–42]. While it seems clear that both *BOULE* and *DAZL* are subject to purifying selection, the selective process guiding *DAZ* evolution has been controversial as purifying selection as well as positive Darwinian selection of *DAZ* have both been described depending on the lineage of primates analyzed [43]. Perhaps, as suggested by Xu and colleagues, *DAZ* has yet to evolve a function essential for the completion of spermatogenesis, maybe by probing different evolutionary possibilities of nature [8].

Among all *DAZ* members, only *Boule* homologs have been identified in protostomes (fly and nematode [36, 37]) and deuterostomes (humans [7, 8, 35]), displaying a high degree of conservation across a large evolutionary time. Additionally, *Boule* exhibits functional similarity in both lineages, as revealed by the ability of a human *Boule* transgene to partially rescue the *boule* function in *Drosophila* [44]. Consequently, *Boule* has attracted a great deal of interest as one strong candidate to represent a conserved reproductive factor in the core machinery of sexual reproduction from metazoans. Indeed, recent studies suggest that *Boule* is restricted to animals, with homologs across cnidarians and bilaterian species encompassing a high degree of conservation both at the RNA-recognition motif (RRM) as well as in their genomic structure (intron-exon boundaries). These data suggest that *Boule* was already present 600 MYA in urbilaterians as well as in eumetazoans, evolving under a strong purifying selection process, whose intensity is not even relieved by the presence of a partially redundant *Dazl* function [7, 45].

In order to completely fulfill the requirements defined in the previous section for a general gametogenic factor, it remained to be demonstrated that the functional constraints acting on *Boule* were restricted to one sex. In the case of protostomes, the analysis of *Boule* expression revealed a different requirement for male (*Drosophila*) and female (nematode) reproduction [36, 37]. In deuterostomes, the expression of a *Boule* homolog was also reported both in testes and ovaries of the fish medaka [46]. The answer to this apparent contradiction came from recent studies carried out in mouse, unveiling the presence of 2 different *Boule* transcripts as a result of alternative splicing that result in a major male-specific type and a secondary type expressed in males but also in early embryonic male and female gonads [7]. Thus, while urbilaterian *Boule* was important for gametogenesis ancestrally, experimental data seems to suggest that the predominant expression of *Boule* in animals was restricted to testes. Indeed, and although exceptions to the male-specific function of *Boule* also happened during evolution in a lineage-specific manner such as in nematodes, the functional relevance of *Boule* in male gametogenesis has been further assessed using mice mutants producing

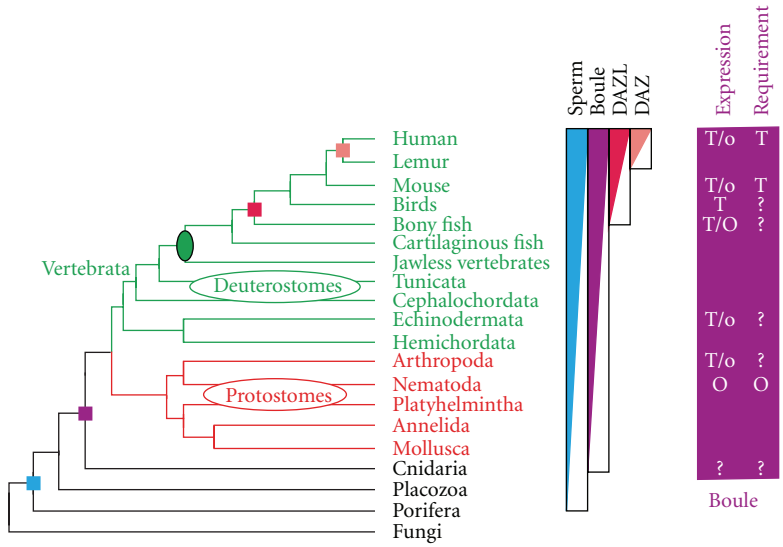


Figure 3: Evolutionary distribution of motile sperm and members of the *DAZ* family among major lineages of animals adapted from [7, 8]. Motile sperm is found in all major phyla of metazoan animals (evolutionary origin indicated by a blue box). The ancient gene *Boule* in the common ancestor of bilateria is indicated by a purple box in the tree topology, likely originated during the evolution of eumetazoans. Its function was spermatogenesis specific based on the predominance of a testis-biased expression in diverse bilaterian lineages as well as in the conservation of male reproductive function in the fly and in mice. *DAZL* arose from the ancestral *BOULE* through gene duplication events, most likely in the ancestral lineage of bony fishes after the divergence from cartilaginous fishes (indicated by a red box) but before the emergence of tetrapods, and is lacking in protostomes. *DAZ* arose from *DAZL* in the primate lineage (pink box) becoming integrated in the Y chromosome later on during primate evolution, after the divergence between New World monkeys and Old World monkeys, approximately 30–40 MYA. Testis expression/functional requirement (T) or ovary expression/functional requirement (O) for *Boule* in different metazoans is indicated in the right-hand side of the tree in purple background. Low levels of *Boule* expression in the ovary are indicated by the lowercase letter "o", in order to distinguish them from abundant expression referred to as "O".

a truncated Boule protein. The resulting male homozygous phenotype matches exactly the mutant *boule* phenotype in *Drosophila*, supporting the role of *Boule* as a reproductive factor widely conserved (structurally and functionally) across eumetazoan animals [7].

5. Concluding Remarks and Future Perspectives

The evolution of sex has constituted a very important subject in evolutionary biology since the very beginning of this discipline, and the study of the evolution of the molecular mechanisms underlying sexual reproduction has proven to be quite challenging due to the rapid divergence of an important part of the reproductive proteins and to the masking effect of pleiotropy (somatic functions in addition to reproductive functions). Within this complex scenario, the characterization of conserved gametogenic factors encompassing reproductive-exclusive roles within the DAZ family has been critical in order to help decipher the evolutionary origin of gametogenesis in metazoan animals. *BOULE* represents the most ancestral *DAZ* member, encompassing an essential reproductive-exclusive requirement in urbilaterians and a high degree of conservation across metazoan animals resulting from purifying selection. This provides support to its role as an ancient male gametogenic factor whose function has been conserved over 600 MY of evolution. Interestingly, the members of the DAZ family share a common molecular nature (RNA-binding proteins) with the molecular signals mediating the entrance of germ cells into meiosis and the differentiation of sperm/eggs. The functional prevalence of *BOULE* throughout metazoan evolution, together with the increasing support favoring a single decision model (responsible for both the entrance to meiosis and the determination of sperm or egg cell differentiation), seem to suggest that male- and female-specific gametogenesis evolved from a common somatic ancestral prototype. This likely took place early in metazoan evolution, instead of arising independently in different lineages.

Although the key role of *Boule* constituting a core component of the metazoan reproductive prototype seems to be well established, further studies will be necessary in order to clarify certain aspects pertaining its evolutionary origin, including detailed analyses of *Boule* expression in additional protostomes and deuterostomes, its characterization in outgroup of bilaterians, as well as the characterization of its subcellular expression. Furthermore, and as previously suggested by Xu and collaborators, the search for a female-specific core component of the reproductive prototype that performs a role as a gametogenic factor (similar to *Boule* in males) will be of an outmost importance [7, 8]. From a more general perspective, the studies reviewed in the present work raise several new questions regarding the specific mechanisms involved in gametogenesis and their evolution. For instance, how do novel regulators of gametogenesis such as the (inhibitor of growth) ING2 protein, a potential tumor suppressor involved in the regulation

of human spermatogenesis through p53- and chromatin-mediated mechanisms [47], fit into the picture?, are these regulators evolutionarily conserved or are they circumscribed to certain mammalian lineages?, if so, when and how were they recruited into spermatogenesis?. The answers to such questions will help to increase our knowledge on the origin and the functional evolution of sexual reproduction, probably one of the most important processes in Biology.

Nonstandard Abbreviations

DAZ: Deleted in azoospermia
DAZL: *DAZ*-like
MY: Million years
RRM: RNA-recognition motif.

Acknowledgments

The authers would like to thank Ana Meira for critical reading of the preliminary version of the manuscript and Rob Kulathinal and two anonymous reviewers for numerous comments on this work. This work was supported by a contract within the Ramon y Cajal Subprogramme from the Spanish Government-MICINN, a Travel Grant from the Xunta de Galicia (INCITE 2006–2010, European Social Fund) and a Research Grant from the Xunta de Galicia (10-PXIB-103-077-PR) awarded to J.M.E.-L. Support was also provided by a Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC-OGP-0046399-02) (to J.A.).

References

[1] J. F. Crow, "Advantages of sexual reproduction," *Developmental Genetics*, vol. 15, no. 3, pp. 205–213, 1994.
[2] A. D. Peters and S. P. Otto, "Liberating genetic variance through sex," *BioEssays*, vol. 25, no. 6, pp. 533–537, 2003.
[3] T. Allers and M. Lichten, "Differential timing and control of noncrossover and crossover recombination during meiosis," *Cell*, vol. 106, no. 1, pp. 47–57, 2001.
[4] S. C. Barrett, "The evolution of plant sexual diversity," *Nature Reviews Genetics*, vol. 3, no. 4, pp. 274–284, 2002.
[5] G. S. Roeder, "Meiotic chromosomes: it takes two to tango," *Genes & Development*, vol. 11, no. 20, pp. 2600–2621, 1997.
[6] H. Ellegren and J. Parsch, "The evolution of sex-biased genes and sex-biased gene expression," *Nature Reviews Genetics*, vol. 8, no. 9, pp. 689–698, 2007.
[7] C. Shah, M. J. Vangompel, V. Naeem et al., "Widespread presence of human BOULE homologs among animals and conservation of their ancient reproductive function," *PLoS Genetics*, vol. 6, no. 7, Article ID e1001022, 2010.
[8] E. Y. Xu, F. L. Moore, and R. A. Pera, "A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in metazoans," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 13, pp. 7414–7419, 2001.
[9] M. A. Handel and J. C. Schimenti, "Genetics of mammalian meiosis: regulation, dynamics and impact on fertility," *Nature Reviews Genetics*, vol. 11, no. 2, pp. 124–136, 2010.
[10] R. Kumar, H. M. Bourbon, and B. de Massy, "Functional conservation of Mei4 for meiotic DNA double-strand break formation from yeasts to mice," *Genes & Development*, vol. 24, no. 12, pp. 1266–1280, 2010.
[11] M. A. Ramesh, S. B. Malik, and J. M. Logsdon, "A phylogenomic inventory of meiotic genes: evidence for sex in Giardia and an early eukaryotic origin of meiosis," *Current Biology*, vol. 15, no. 2, pp. 185–191, 2005.
[12] A. M. Villeneuve and K. J. Hillers, "Whence meiosis?" *Cell*, vol. 106, no. 6, pp. 647–650, 2001.
[13] J. Kimble and D. C. Page, "The mysteries of sexual identity: the germ cell's perspective," *Science*, vol. 316, no. 5823, pp. 400–401, 2007.
[14] M. M. Matzuk and D. J. Lamb, "Genetic dissection of mammalian fertility pathways," *Nature Cell Biology*, vol. 4, pp. S41–S49, 2002.
[15] W. Haerty, S. Jagadeeshan, R. J. Kulathinal et al., "Evolution in the fast lane: rapidly evolving sex-related genes in Drosophila," *Genetics*, vol. 177, no. 3, pp. 1321–1335, 2007.
[16] W. J. Swanson and V. D. Vacquier, "The rapid evolution of reproductive proteins," *Nature Reviews Genetics*, vol. 3, no. 2, pp. 137–144, 2002.
[17] G. J. Wyckoff, W. Wang, and C. I. Wu, "Rapid evolution of male reproductive genes inn the descent of man," *Nature*, vol. 403, no. 6767, pp. 304–309, 2000.
[18] J. M. Good and M. W. Nachman, "Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis," *Molecular Biology and Evolution*, vol. 22, no. 4, pp. 1044–1052, 2005.
[19] J. A. Anderson, W. D. Gilliland, and C. H. Langley, "Molecular population genetics and evolution of Drosophila meiosis genes," *Genetics*, vol. 181, no. 1, pp. 177–185, 2009.
[20] A. Civetta, S. A. Rajakumar, B. Brouwers, and J. P. Bacik, "Rapid evolution and gene-specific patterns of selection for three genes of spermatogenesis in Drosophila," *Molecular Biology and Evolution*, vol. 23, no. 3, pp. 655–662, 2006.
[21] D. J. Obbard, F. M. Jiggins, D. L. Halligan, and T. J. Little, "Natural selection drives extremely rapid evolution in antiviral RNAi genes," *Current Biology*, vol. 16, no. 6, pp. 580–585, 2006.
[22] A. E. Baltus, D. B. Menke, Y. C. Hu et al., "In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication," *Nature Genetics*, vol. 38, no. 12, pp. 1430–1434, 2006.
[23] J. R. Huynh and D. St Johnston, "The role of BicD, Egl, Orb and the microtubules in the restriction of meiosis to the Drosophila oocyte," *Development*, vol. 127, no. 13, pp. 2785–2794, 2000.
[24] R. Mendez and J. D. Richter, "Translational control by CPEB: a means to the end," *Nature Reviews Molecular Cell Biology*, vol. 2, no. 7, pp. 521–529, 2001.
[25] B. E. Thompson, D. S. Bernstein, J. L. Bachorik, A. G. Petcherski, M. Wickens, and J. Kimble, "Dose-dependent control of proliferation and sperm specification by FOG-1/CPEB," *Development*, vol. 132, no. 15, pp. 3471–3481, 2005.
[26] M. Otori, T. Karashima, and M. Yamamoto, "The Caenorhabditis elegans homologue of deleted in azoospermia is involved in the sperm/oocyte switch," *Molecular Biology of the Cell*, vol. 17, no. 7, pp. 3147–3155, 2006.
[27] P. J. Chen, A. Singal, J. Kimble, and R. E. Ellis, "A novel member of the tob family of proteins controls sexual fate in Caenorhabditis elegans germ cells," *Developmental Biology*, vol. 217, no. 1, pp. 77–90, 2000.
[28] M. K. Barton and J. Kimble, "fog-1, a regulatory gene required for specification of spermatogenesis in the germ line of Caenorhabditis elegans," *Genetics*, vol. 125, no. 1, pp. 29–39, 1990.

[29] A. Casper and M. Van Doren, "The control of sexual identity in the Drosophila germline," *Development*, vol. 133, no. 15, pp. 2783–2791, 2006.

[30] S. R. Palumbi, "Speciation and the evolution of gamete recognition genes: pattern and process," *Heredity*, vol. 102, no. 1, pp. 66–76, 2009.

[31] J. D. Lewis, Y. Song, M. E. de Jong, S. M. Bagha, and J. Ausió, "A walk though vertebrate and invertebrate protamines," *Chromosoma*, vol. 111, no. 8, pp. 473–482, 2003.

[32] D. S. Chu, H. Liu, P. Nix et al., "Sperm chromatin proteomics identifies evolutionarily conserved fertility factors," *Nature*, vol. 443, no. 7107, pp. 101–105, 2006.

[33] S. Dorus, S. A. Busby, U. Gerike, J. Shabanowitz, D. F. Hunt, and T. L. Karr, "Genomic and functional evolution of the Drosophila melanogaster sperm proteome," *Nature Genetics*, vol. 38, no. 12, pp. 1440–1445, 2006.

[34] K. McGuigan, L. Rowe, and M. W. Blows, "Pleiotropy, apparent stabilizing selection and uncovering fitness optima," *Trends in Ecology & Evolution*, vol. 26, pp. 22–29, 2011.

[35] K. Kee, V. T. Angeles, M. Flores, HA. N. Nguyen, and R. A. Reijo Pera, "Human DAZL, DAZ and BOULE genes modulate primordial germ-cell and haploid gamete formation," *Nature*, vol. 462, no. 7270, pp. 222–225, 2009.

[36] C. G. Eberhart, J. Z. Maines, and S. A. Wasserman, "Meiotic cell cycle requirement for a fly homologue of human deleted in Azoospermia," *Nature*, vol. 381, no. 6585, pp. 783–785, 1996.

[37] T. Karashima, A. Sugimoto, and M. Yamamoto, "Caenorhabditis elegans homologue of the human azoospermia factor DAZ is required for oogenesis but not for spermatogenesis," *Development*, vol. 127, no. 5, pp. 1069–1079, 2000.

[38] R. Reijo, R. K. Alagappan, P. Patrizio, and D. C. Page, "Severe oligozoospermia resulting from deletions of azoospermia factor gene on Y chromosome," *The Lancet*, vol. 347, no. 9011, pp. 1290–1293, 1996.

[39] R. Reijo, T. Y. Lee, P. Salo et al., "Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene," *Nature Genetics*, vol. 10, no. 4, pp. 383–393, 1995.

[40] C. Carani, J. Gromoll, M. H. Brinkworth, M. Simoni, G. F. Weinbauer, and E. Nieschlag, "cynDAZLA: a cynomolgus monkey homologue of the human autosomal DAZ gene," *Molecular Human Reproduction*, vol. 3, no. 6, pp. 479–483, 1997.

[41] J. Gromoll, G. F. Weinbauer, H. Skaletsky et al., "The old world monkey DAZ (Deleted in AZoospermia) gene yields insights into the evolution of the DAZ gene cluster on the human Y chromosome," *Human Molecular Genetics*, vol. 8, no. 11, pp. 2017–2024, 1999.

[42] R. Saxena, J. W. de Vries, S. Repping et al., "Four DAZ genes in two clusters found in the AZFc region of the human Y chromosome," *Genomics*, vol. 67, no. 3, pp. 256–267, 2000.

[43] J. P. Bielawski and Z. Yang, "Positive and negative selection in the DAZ gene family," *Molecular Biology and Evolution*, vol. 18, no. 4, pp. 523–529, 2001.

[44] E. Y. Xu, D. F. Lee, A. Klebes, P. J. Turek, T. B. Kornberg, and R. A. Reijo Pera, "Human BOULE gene rescues meiotic defects in infertile flies," *Human Molecular Genetics*, vol. 12, no. 2, pp. 169–175, 2003.

[45] D. W. Houston, J. Zhang, J. Z. Maines, S. A. Wasserman, and M. L. King, "A Xenopus DAZ-like gene encodes an RNA component of germ plasm and is a functional homologue of Drosophila boule," *Development*, vol. 125, no. 2, pp. 171–180, 1998.

[46] H. Xu, Z. Li, M. Li, L. Wang, and Y. Hong, "Boule is present in fish and bisexually expressed in adult and embryonic germ cells of medaka," *PLoS One*, vol. 4, no. 6, Article ID e6097, 2009.

[47] M. Saito, K. Kumamoto, A. I. Robles et al., "Targeted disruption of Ing2 results in defective spermatogenesis and development of soft-tissue sarcomas," *PLoS One*, vol. 5, no. 11, Article ID e15541, 2010.

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 165919, 9 pages
doi:10.4061/2011/165919

Research Article

The Selfish Grandma Gene: The Roles of the X-Chromosome and Paternity Uncertainty in the Evolution of Grandmothering Behavior and Longevity

Molly Fox,[1] Johannes Johow,[2] and Leslie A. Knapp[1]

[1] *Department of Biological Anthropology, University of Cambridge, Pembroke Street, Cambridge CB2 3QY, UK*
[2] *Zentrum für Philosophie und Grundlagen der Wissenschaft, Universität Giessen, Otto-Behaghel-Straße 10 C, 35394 Giessen, Germany*

Correspondence should be addressed to Leslie A. Knapp, lak26@cam.ac.uk

Received 6 October 2010; Revised 7 December 2010; Accepted 18 January 2011

Academic Editor: Alberto Civetta

When considering inclusive fitness, it is expected that individuals will provide more care towards those with whom they are more closely related. Thus, if a selfish X-linked genetic element influenced care giving, we would expect care giving to vary with X-relatedness. Recent studies have shown that X-chromosome inheritance patterns may influence selection of traits affecting behavior and life-history. Sexually antagonistic (SA) zygotic drive could encourage individuals to help those with whom they are more likely to share genetic material at the expense of other relatives. We reanalyze previously reported data in light of this new idea. We also evaluate the effects of paternity uncertainty on SA-zygotic drive. Our evidence suggests that human paternal discrepancy is relatively low. Using published models, we find the effects of paternal discrepancy do not override opportunity for selection based on X-relatedness. Based on these results, longevity and grandmothering behaviors, including favoritism, may be more heavily influenced by selection on the X-chromosome than by paternity uncertainty.

1. Introduction

Care giving between family members may be influenced by genes in ways that encourage people to treat relatives differently according to their degrees of relatedness [1, 2]. The importance of genetics in care giving behaviors within families is somewhat intuitive: one would expect a woman to care more for her son than for her nephew, and more for her sister than for her cousin. In other words, it is expected that people vary the amount of care they provide proportionally to their genetic relatedness with family members. It follows that a gene which encourages such a care giving pattern may also be adaptive, as those who carry it help others who are most likely to carry it.

The idea that differential relatedness encourages preferential behaviors is not new. Many publications have reported evidence supporting kin selection and several recent studies have explored the ways in which adopted children may be treated differently than biological children, how step-parents may invest less in step children than in biological children, and how the extent of paternal care may vary based on likelihood of paternity [3–7].

Along these same lines, grandmothering behavior has been implicated in our species' unique post-menopausal longevity. The advantages that grandmothers bestow upon certain grandchildren may create opportunity for the selection of selfish genetic elements that increase longevity. Beyond this, it has been suggested that sexually antagonistic zygotic drive (SA-zygotic drive) may contribute to the behavioral pattern of some grandmothers helping granddaughters at the expense of grandsons [8].

Recent research has shown how inheritance patterns of the X-chromosome may create opportunity for selection of traits affecting human behavior and life history. Here, we reanalyze previously published data in light of the SA-zygotic drive argument. We also re-evaluate data related to prehistoric rates of paternal discrepancy and consider how discrepancy would affect SA-zygotic drive. We present

models that examine how paternity uncertainty and X-linked selfish mutations may influence selection. We find that even the highest estimated rates of paternity uncertainty do not override models for selection on grandmothers based on X-chromosome relatedness. Therefore, the differential genetic relatedness between family members may explain the ways in which women treat their grandchildren, as well as the longevity of our species.

2. Grandmothering Behavior

2.1. X-Linked Grandmother Hypothesis. The grandmother hypothesis, originally formulated to account for menopause itself, has since often been utilized in discussions of postmenopausal longevity [9–11]. This view holds that postmenopausal longevity evolved in our species because women with genetic elements coding for increased lifespan experienced increased inclusive fitness, as they were able to increase their daughters' fertility and the survivorship of their grandchildren [9, 11]. Fox et al. [12] proposed an X-linked grandmother hypothesis, based on the fact that there is variation in X-chromosome sharing between grandmothers and grandchildren depending on the sex of the grandchild and whether the grandmother is from the matriline or patriline. This differential genetic relatedness creates differential incentives for grandmothers to invest in grandchildren. In Fox et al.'s analysis of seven populations, the variation in grandmothers' effect on grandchild likelihood of mortality correlated with their X-relatedness [12].

2.2. X-Linked Granddaughter Favoritism Hypothesis. The differential X-relatedness between grandmothers and grandchildren creates opportunity for genes that affect behaviors associated with grandparenting to cluster on the X-chromosome. When paternal grandmothers (PGMs) invest in granddaughters, there is a better return on that investment for the X-chromosome than for the autosomes, so X-linked alleles for grandparenting will be more strongly selected than autosomal alleles [13, 14].

One pattern of grandparenting behavior observed in Fox et al.'s [12] meta-analysis is that of PGMs decreasing survivorship of grandsons. This phenomenon can be viewed in light of selfish genetic elements on the X-chromosome. SA-zygotic drive refers to selfish genetic material on the X or Y chromosomes that helps offspring who carry it and harms offspring who do not carry it [15]. Rice et al.'s [8] mathematical model reveals the circumstances under which natural selection would cause X-linked mutations that affect grandparenting behavior to persist. This can be thought of as an "X-Linked Granddaughter Favoritism Hypothesis." For a selfish X-linked mutation, the only relatedness that affects selection is X-chromosome relatedness. X-relatedness varies by line of descent and sex of grandchild, so an X-linked mutation in a woman has a 50% chance of being transmitted to her son's daughter, 0% chance of being transmitted to her son's son, and a 25% chance of being transmitted to her daughter's child of either sex. Using these values in Rice et al.'s [8] mathematical model shows that a dominant X-linked mutation causing all grandparents to help granddaughters at the expense of grandsons would increase in frequency as long as the magnitude of the cost to grandsons is no more than twice the benefit to granddaughters. What if the X-linked mutation were only expressed in certain grandparents (Table 1)? An X-linked mutation that causes only females (i.e., grandmothers and not grandfathers) to help granddaughters at the expense of grandsons would increase in frequency as long as the expense to grandsons is no more than three times the benefit to granddaughters. An X-linked mutation that is only expressed in PGMs would increase in frequency as long as there was a benefit to granddaughters, no matter what the effect on grandsons. This means that if an X-linked mutation arose which only affected how women treat their sons' children (in other words, the way paternal grandmothers treat their grandchildren) in terms of helping granddaughters at the expense of grandsons, there would be no hindrance to that mutation reaching fixation in the population. Overall, there are many opportunities for mutations to accumulate on the X-chromosome that cause granddaughters to be favored at the expense of grandsons. Although selection for this phenotype occurs only in PGMs, Rice et al.'s [8] model indicates that it can evolve in other grandparents as a correlated effect. Table 1 shows the predictions of the X-Linked Granddaughter Favoritism Hypothesis and the circumstances under which this phenotype would accumulate.

Table 1: The circumstances under which an X-linked gene coding for favoritism of granddaughters would persist in a population (based on Rice et al. [8] mathematical model).

When X-linked mutation helping granddaughters at expense of grandsons is expressed in:	It would increase in frequency provided the expense to grandsons is no more than (values below) times the benefit to granddaughters
All grandparents (dominant expression)	2
All grandparents (additive expression)	1.5
Grandmothers only (dom or add expression)	3
Paternal grandmothers only (dom or add expression)	no limit

The present study analyzes the data from Fox et al.'s [12] meta-analysis of seven geographically and temporally varied populations [16], in light of Rice et al.'s [8] discussion of SA-zygotic drive. Rice et al.'s [8] model suggests that granddaughters should be favored at the expense of grandsons. The predictions (Table 1) in order of increasing effect strength are that granddaughters are helped at the expense of grandsons by (1) and (2) All grandparents, (3) Grandmothers, and (4) Paternal grandmothers. The third prediction, that all grandmothers might favor granddaughters at the expense of grandsons, is not supported by the data, as the maternal grandmother (MGM) never exhibits this trend. However, in six of the seven populations PGMs have the predicted

Table 2: Data from Fox et al. [12] analyzed according to predictions based on SA-zygotic drive model (Rice et al. [8]). PGM: paternal grandmother; MGM: maternal grandmother; SA: sexually antagonistic; GD: granddaughter; GS: grandson. Check mark indicates that the population data in Fox et al. does conform to the Rice et al. prediction, and a dash indicates that it does not.

Population	PGM helps GD and harms GS	MGM helps GD and harms GS
Germany	✓	—
England	✓	—
Ethiopia	✓	—
Canada	✓	—
Japan	✓	—
Gambia	—	—
Malawi	✓	—

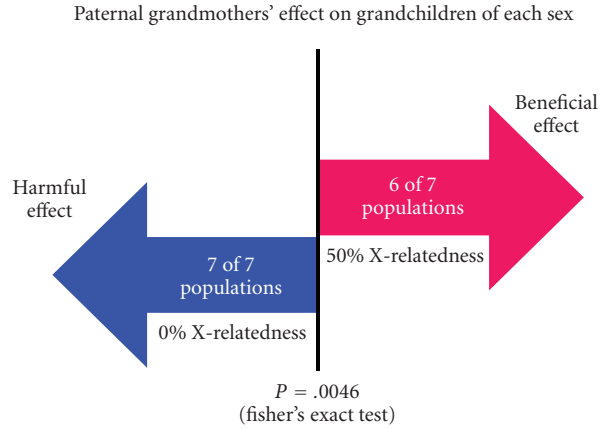


Figure 1: Analysis of the PGMs effect on grandchildren using data from Fox et al. [12]. Red (top) represents granddaughters, and blue (bottom) represents grandsons.

effect of helping granddaughters and harming grandsons, providing some support for the fourth prediction (Table 2).

Considering Fox et al.'s [12] results, the PGM had a harmful effect on grandsons in all seven populations, and a helping effect on granddaughters in six of the seven populations (Figure 1). This trend was statistically significant: Fisher's exact test $P = .0046$. These results are consistent with the X-Linked Granddaughter Favoritism Hypothesis, which suggests that selfish genetic material on the X-chromosome coding for helping granddaughters at the expense of grandsons should be most strongly favored as it is only expressed in PGMs. This PGM-grandson harming behavior, first noticed by Jamison et al. [17] who described the effect in their own data as "startling to say the least," is consistent with the presence of X-linked mutations encoding sexually antagonistic phenotypes.

The aforementioned studies found evidence of PGMs favoring granddaughters, consistent with the fourth prediction of the X-Linked Granddaughter Favoritism Hypothesis (Table 1). But based on the limitations of the statistics and the number of study populations, this may not be the most sensitive method that could be employed to test the hypothesis. When each of the 28 effects measured in Fox et al.'s meta-analysis [12] are considered individually, only five were statistically significant, although the directionalities of the effects were highly significant (Figure 1). The conclusions of Jamison et al. [17] and Fox et al. [12], therefore, provide a limited amount of evidence for the fourth prediction of the X-Linked Granddaughter Favoritism Hypothesis. Further research is needed to verify a PGM-specific trend. Also, it is important to note that these studies only examine grandchild mortality rates, not behavior, health, or any other measure of favoritism. These studies were conducted not to analyze behaviors, but rather, as an opportunity to evaluate evidence related to the grandmother hypothesis. Therefore, if the specific predictions of Rice et al. [8] are to be tested rigorously, perhaps we should look at behavior, rather than mortality rates.

Evidence supporting favoritism of granddaughters via SA-zygotic drive comes from questionnaire studies in which grandparents and grandchildren are asked to evaluate their relationships with each other. Euler and Weitzel [18] found that grandparents provided more care to granddaughters than to grandsons. Participants were asked to rank amount of care on a scale from 1 to 7, and mean granddaughter care was 4.45 and grandson care was 4.23. These results support the first prediction of the X-Linked Granddaughter Favoritism Hypothesis (Table 1). Adding their own data to that of Euler and Weitzel, Chrastil et al. [14] found that granddaughters were favored over grandsons by both MGMs ($P < .0001$) and PGMs ($P = .003$). This favoritism of granddaughters over grandsons provides further support for the third prediction of the hypothesis (Table 1).

3. Longevity

3.1. Sexually Antagonistic Zygotic Drive and Grandmother Longevity. The X-Linked Granddaughter Favoritism Hypothesis can account for SA-zygotic drive causing some, or even all, grandparents (via side-effect of selection on PGM) to carry X-linked traits that induce favoritism of granddaughters at the expense of grandsons. By the same logic, SA-zygotic drive may cause perpetuation of an X-linked longevity gene.

If, as suggested by the evidence presented above, some grandmothers favor granddaughters, then those girls with grandmothers who live longer would have the greatest advantage, as they would experience the benefits of that favoritism longer. This effect may be tempered by costs associated with having a grandmother, which may increase as she ages. Additionally, the benefits of a grandmother may only benefit young grandchildren. Further research should explore these and other limits of grandmother benefits. Nonetheless, if a grandmother has X-linked genetic elements causing her to live longer to at least a certain extent, her granddaughters may disproportionately survive. The result might be that the X-linked genetic elements will increase in frequency in the population.

There may be natural selection for selfish X-linked alleles that help one sex of grandchild at the expense of the other. If presence of a PGM (i.e., surviving) for more years helps girls and harms boys, then there is opportunity for natural selection of X-linked alleles that increase longevity. Using Rice et al.'s [8] formula, in which relatedness (R) refers to X-relatedness because this hypothesis considers only X-linked traits, a selfish genetic element will be favored as long as the following condition is true:

$$R_{\text{Helped}} \times B_{\text{Helped}} > R_{\text{Harmed}} \times C_{\text{Harmed}}. \quad (1)$$

As described in Rice et al. [8], R_{Helped} is the relatedness to the individual helped, B_{Helped} is the benefit to the individual helped, R_{Harmed} is the relatedness to the individual harmed, and C_{Harmed} is the cost to the individual harmed. Therefore, if an allele encoding greater longevity is X-linked, it will increase in frequency as long as one of the conditions listed in Table 1 is met.

(1) The magnitude of grandparent longevity's harming effect on grandsons is no more than twice the magnitude of the helping effect on granddaughters.

(2) If an X-linked longevity allele is only expressed in females (i.e., grandmothers), it will increase in frequency as long as the magnitude of grandmother longevity's harming effect on grandsons is no more than three times its helping effect on granddaughters.

(3) If an X-linked longevity allele is only expressed in PGMs (in other words, only affects the way a woman treats her sons' children), then it will be favored without constraint.

In sum, SA-zygotic drive could contribute to our species' unique phenomenon of postmenopausal longevity, as a consequence of X-linked selfish genetic elements being favored in certain grandparents.

3.2. Grandmother Alloparenting and Longevity. Many proponents of the grandmother hypothesis have suggested that postmenopausal longevity has evolved in our species because grandmothers can bolster their inclusive fitness by reducing the weaning age of their grandchildren and thereby diminish the interbirth interval of their daughters and/or daughters-in-law and enhance the survivorship of their grandchildren especially as toddlers. Grandmothers may be in a unique position to increase their number of descendants and the likelihood of those descendants' survival without compromising their own fertility.

A recent study by Kachel and coworkers [19] set out to quantify whether grandmothering could actually be a strong enough selective force to account for the perpetuation of longevity. The authors ran three mathematical simulations to test if grandmothering could increase inclusive fitness enough to influence the evolution of human longevity and/or age at weaning and survival of grandchildren. While their results claimed to prove that grandmothering cannot account for longevity, in fact their results do not conflict with the new X-Linked Grandmother Hypothesis [8, 12]. This is because Kachel et al.'s [19] study only included maternal grandchildren. Their model did not consider the effects of the paternal line and assumed that grandmothers did not provide care for their sons' children. Their results contradict studies which suggest that maternal grandmothering accounts for our species' longevity [16, 20, 21], and they cite paternal discrepancy as the reason that only maternal grandmothers are relevant to the adaptive circumstances leading to postmenopausal longevity.

If, however, SA-zygotic drive is responsible for the evolution of grandmothering and longevity alleles, the asymmetry in genetic relatedness along the paternal line is an important consideration, despite potential problems of paternity uncertainty. The previous section of this article suggests that longevity could be a result of selection purely on the PGM, and recent work by Fox et al. [12] and Rice et al. [8] suggests that PGMs' care for granddaughters could be the key to selection for grandmother care (Tables 1 and 2). Thus, Kachel et al.'s [19] conclusion that maternal grandmothering cannot account for the selection of genetic factors affecting longevity is not in conflict with the possibility that the PGMs behavior drives selection for longevity. Further research should investigate the specific behaviors of grandmothers, and the particular ways in which granddaughters are helped and grandsons are harmed. Nevertheless, paternal relatives play an important role in the X-Linked Granddaughter Favoritism Hypothesis.

4. Paternity Uncertainty

Paternal discrepancy refers to cases in which a man raises a child as his own when unbeknownst to him, he is not the biological father. If this were often the case, there would be little incentive not only for men to invest in paternal care, but also for patrilineal kin to invest in caring for his children at all. With respect to the X-Linked Granddaughter Favoritism Hypothesis, high rates of paternal discrepancy would result in little selective pressure for women to engage in caretaking behaviors towards their sons' children.

Many previous studies of the grandmother hypothesis do not distinguish between MGMs and PGMs [22, 23], and those that do distinguish between matrilineal and patrilineal relatedness tend to frame their analysis around paternity uncertainty [16, 20, 21]. Prominent researchers have claimed that selection for grandmothering behaviors and postmenopausal longevity is a result of selection exclusively on the MGM. Some studies, such as the aforementioned paper by Kachel et al. [19], have even left PGMs out of their analysis entirely under the assumption that paternity uncertainty renders PGMs role immaterial in the evolution of human longevity. As described above, PGMs are integral to the bases of all X-linked grandmother hypotheses [8, 12, 14]. Therefore, two questions cannot be ignored: how prevalent has paternal discrepancy been throughout our species existence, and how prevalent would it have to be to refute X-linked theories of longevity selection?

We suggest that paternal discrepancy may not have been much different during pre-history than it is today, based

on studies of the Y-chromosome as well as anthropological information from modern hunter-gatherers (see Supplementary Material I available online at doi:10.4061/2011/165919). Based on an extensive literature review (see below), we suggest that this rate is 1.3–3.7%. We can reanalyze the likelihood of selfish X-linked genes accumulating using Rice et al.'s inequality equations [8] by taking into account paternal discrepancy. We find that the thresholds for the accumulation of X-linked mutations causing certain grandparents to favor granddaughters at the expense of grandsons are altered only slightly. The thresholds are reported below.

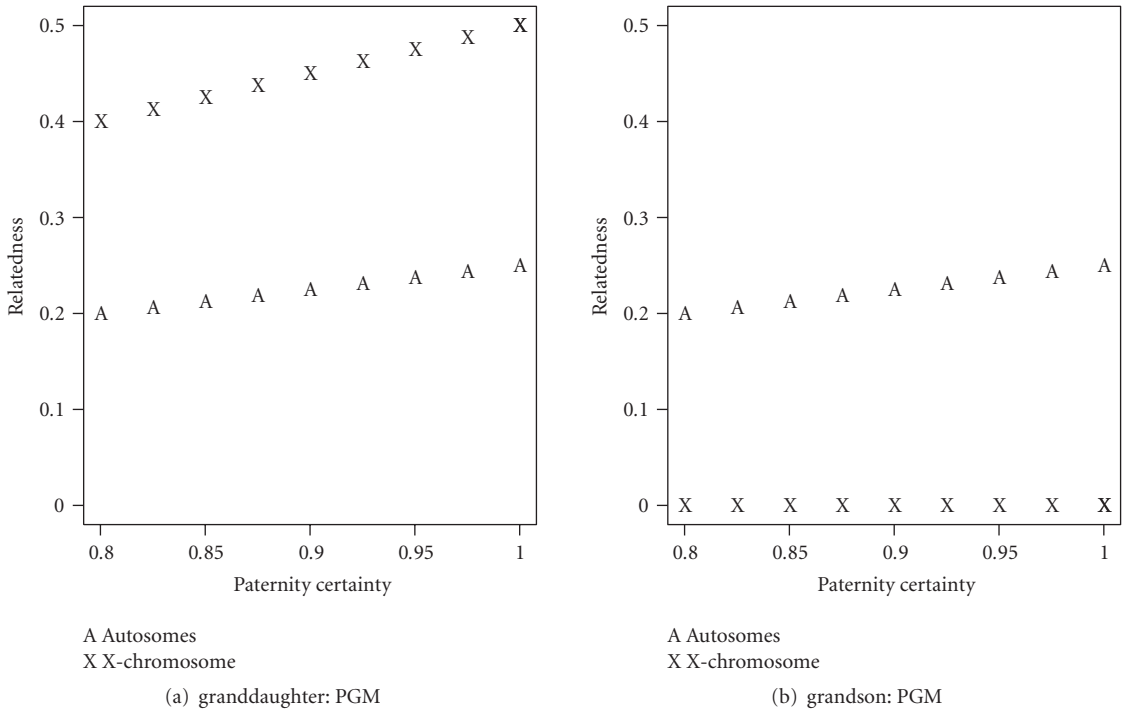
4.1. Prevalence of Paternal Discrepancy. Paternal discrepancy is often cited in academic literature as an unsubstantiated 10% in the modern human populations (e.g., [28–30]), but there is evidence that the actual rates are far lower. Bellis et al. [27] and Simmons et al. [25] performed meta-analyses on geographically varied samples of 20,871 people from 17 populations, and 16,523 people from 12 populations, respectively. All of these people underwent biological tests for purposes other than discovering paternity; therefore, the studies avoided bias towards discrepancy. Bellis et al. found that median paternal discrepancy was 3.7%, and Simmons et al. [25] reported the rate was 1.3%.

Two other studies analyzed the Y-chromosome to measure paternal discrepancy in ancient populations. Sykes and Irven [26] found a highly significant association between British men based on surnames and Y-chromosome haplotype, tracing back to a common paternal ancestor 700 years ago. Based on their data, Sykes and Irven [26] calculated a paternal discrepancy rate of 1.3%. A similar study analyzed the Y-chromosome similarities among modern "Cohanim" Jews, the supposed descendants of the biblical Moses [24]. Skorecki et al. [24] found that within this population, whose lineage dates back to 3,300 years ago, there is no evidence of paternal discrepancy from non-Cohanim Jews to complicate patterns of Y-chromosome inheritance. The authors show that paternity certainty is close to 100% with high probability. Although it is possible that extramarital paternity may have occurred with a man sharing the same surname, and thus discrepancy would not be detected, these estimates of paternal discrepancy are not only low but are also consistent with results published by Simmons et al. [25–27].

4.2. The Effect of Paternity Uncertainty on Selection for X-Linked Longevity Trait. Paternity uncertainty would surely change the likelihood of a PGM sharing an allele with her grandchild. Therefore, we have added paternal discrepancy into previously published calculations regarding the accumulation of X-linked mutations for grandmothering behavior and longevity. With this, we can show a range of PGM-grandchild relatedness given a generous variety of paternal discrepancy conditions. We use Rice et al's [8] equations to calculate the effect magnitudes for which an X-linked granddaughter favoritism trait would increase in frequency.

While the varying relatedness between maternal and paternal grandmothers with granddaughters and grandsons has been reported before (e.g., [12, 14]), these predicted relatedness values can be re-evaluated by considering rates of paternal discrepancy. Paternal discrepancy changes some aspects of the X-chromosome and autosomal genetic relatedness between (a) PGMs and granddaughters, and (b) PGM and grandsons (Figure 2; see Supplementary Material Table S7 for mathematical methods available online at doi:10.4061/2011/165919). Previous authors have suggested that paternity uncertainty may result in PGMs being statistically unlikely to share genes with their grandchildren and, as a consequence, selection for grandmothering traits act only on MGMs. The best estimates of both current and ancient paternal discrepancy (see above and [24–27]) range from 1.3–3.7%, although literature and textbooks often claim an unfounded 10%. To consider the widest range of possible values, we have modeled PGM-grandchild relatedness with paternal discrepancy ranging from 0% to 20% (Figure 2). These graphs show that although paternal discrepancy has some impact upon genetic relatedness, the comparisons between grandmother-grandchild pairs remain largely the same. The X-relatedness between a PGM and grandson is always 0%, and so hypotheses related to behaviors associated with this relationship, based on sharing no X-linked genes, still hold no matter what the amount of paternity uncertainty. The X-relatedness between a PGM and granddaughter is 50% given total paternity certainty. Even when paternity uncertainty is as high as 20% (i.e., there is a 20% chance that the PGM's son is not the biological father of the granddaughter), the X-relatedness between the PGM and granddaughter is 40%. This is because Hamiltonian relatedness refers to the statistical likelihood that two individuals share a given gene, rather than the percent of genetic material two individuals share [1, 2]. Compared to the PGM-grandson relatedness of 0% and MGM-grandchild X-relatedness of 25% (which are all relationships unaffected by paternity uncertainty), PGM-granddaughter X-relatedness of 40% is still significantly higher than all other grandmother-grandchild X-relatedness. Even given an unlikely 20% rate of paternal discrepancy, the 40% chance of sharing X-linked genetic material between a PGM and granddaughter is still much higher than with a son's son (0%) and between a MGM and granddaughter (25%). Thus, there remains the same expected favoritism as Fox et al. [12] suggested (see Table 1 in Fox et al. [12]).

Rice et al. [8] calculated the circumstances under which an X-linked allele causing favoritism of granddaughters over grandsons would accumulate (Table 1). Using their inequality equations (see Table 2 of Rice et al. [8]), we have calculated new values to describe the circumstances under which the hypothetical X-linked granddaughter favoritism allele would increase in frequency, given varying degrees of paternal discrepancy (Table 3). Following Rice et al. [8], we calculate the likelihood that an X-linked mutation, which causes grandparents to help their granddaughters at the expense of their grandsons, would accumulate as long as the detriment to grandsons is not more than a calculable magnitude greater than the benefit to granddaughters. Given dominant allele expression and complete paternity certainty,



(a) granddaughter: PGM

(b) grandson: PGM

Figure 2: Hamiltonian r value for autosomal and X-relatedness between paternal grandmothers (PGM) and grandchildren; that is, the likelihood that any given autosomal or X-linked gene in the PGM will be present in her grandchild. These values are expressed in terms of a range of paternity certainty. (a) Shows the relatedness between a PGM and her granddaughter, and (b) shows the relatedness between a PGM and her grandson. For example, when paternal certainty is 100%, a granddaughter has a 50% chance of carrying any given X-linked allele of her PGMs', and a 25% chance of carrying any given autosomal allele of her PGMs'. See Supplementary Material Table S7 for mathematical methods available online at doi:10.4061/2011/165919.

Table 3: The circumstances under which an X-linked gene coding for favoritism of granddaughters would accumulate in a population using Rice et al. [8] model, given five rates of paternal discrepancy: 0% (as reported in Rice et al. [8], and consistent with Skorecki et al. [24]); 1.3% (estimate based on Simmons et al. [25] and Sykes and Irven [26]); 3.7% (estimate based on Bellis et al. [27]); 10% (popular unfounded figure included here to show range of possibility).

Rate of paternal discrepancy →	0%	1.30%	3.70%	10%
When X-linked mutation helping granddaughters at expense of grandsons is expressed in:	Its frequency would increase provided the expense to grandsons is no more than (values below) times the benefit to granddaughters			
All grandparents (dominant expression)	2	1.993457	1.98167	1.947368
All grandparents (additive expression)	1.5	1.331876	1.329235	1.321429
Grandmothers only (dom or add expression)	3	2.97400	2.9280	2.8000
Paternal grandmothers only (dom or add expression)	No limit	No limit	No limit	No limit

the threshold for selection is grandson harm at twice the expense of granddaughter help. Using the three rates of paternal discrepancy from the literature review above (0%; 1.3%; 3.7% [24–27]) and also the popular figure of 10%, this threshold remains above 1.9. In other words, even given the highest estimated rate of prehistoric paternal discrepancy (10%), a dominant X-linked mutation that causes grandparents to help granddaughters at the expense of grandsons would accumulate as long as the expense to grandsons were no more than 1.95 times the benefit to granddaughters. The figures for additive expression, sex-specific, and lineage-specific expression are given in Table 3. The paternity uncertainty induced changes in threshold appear to be minor enough that the possibility of SA-zygotic drive towards granddaughter favoritism and longevity is not compromised.

The most comprehensive analysis possible measures the circumstances under which an X-linked mutation would increase in frequency, given a rate of paternal discrepancy ranging from 0 (total certainty; all fathers identify their children accurately) to 1 (total discrepancy; all fathers identify their children inaccurately) (Figure 3). Although

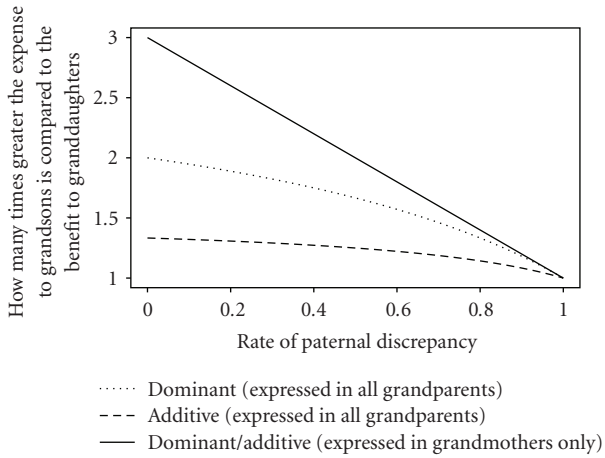


Figure 3: The threshold for an X-linked mutation causing grandparents (or grandmothers) to favor their granddaughters at the expense of their grandsons to accumulate. These curves represent the thresholds for which such a mutation would increase in frequency. The threshold can be described as the maximum number of times greater the expense of this mutation would be to grandsons, compared to the benefit of the mutation to granddaughters. These values were calculated using the mathematical model from Rice et al. [8, supplement]. We suggest that paternal discrepancy among our species would be approximately 1.3–3.7%. However, this graph shows a range of paternal discrepancy from 0% (all paternity is identified correctly) to 100% (all paternity is identified inaccurately).

total discrepancy is implausible, it is useful to visualize a curve that depicts the threshold for accumulation changes for the hypothetical X-linked mutation. A more specific analysis focuses on the curve where the threshold changes for allele frequency increase with rates of paternal discrepancy ranging from 0% to 10% (Figure 4). This segment of the larger threshold curve (Figure 3) displays the rate at which paternal discrepancy affects the threshold of the proportion benefit to granddaughters versus detriment to grandsons. This is the segment we consider to be the most likely range of paternal discrepancy rates among modern and ancient human populations. While increasing paternal discrepancy creates a stricter criterion for allele frequency increase (less detriment to grandsons compared to benefit to granddaughters), this effect is not strong (Table 3; Figure 4).

5. Conclusion

The asymmetrical inheritance pattern of the X-chromosome may influence selection among traits related to behavior and life history. The variation in X-relatedness between grandmothers and grandchildren, based on sex and lineage, may create opportunity for selection of genes that affect grandmothering strategy and longevity. Here, we have reanalyzed data from seven previously-studied populations, in light of Rice et al.'s [8] suggestion of SA-zygotic drive. The analysis explores the circumstances under which an X-linked mutation would persist, causing grandmothers to behave preferentially towards granddaughters at the expense of grandsons. The results show that six of the seven populations conform to a prediction of this hypothesis: that PGMs have a beneficial effect on granddaughters and a harmful effect on grandsons. Further research should explore how consistent this trend is between populations, and should see if this trend exists in modern industrialized populations. Additionally, future research should explore the behavioral mechanisms involved in this pattern.

Preferential grandmothering behavior may be present in other species as well. Johnstone and Cant [31] recently reported that whales represent another clade in which postmenopausal longevity is consistently observed. Among certain whales, as a female gets older, her genetic relatedness to the members of her local group increases. This suggests that it is increasingly advantageous for her to care for individuals in her social group because she is increasingly likely to be closely related to them. The benefits of this strategy may contribute to longevity in whales. Also, some whale species are known to favor sons over their daughters, and this may directly affect fitness of individuals. Further research into preferential behaviors within families should extend to other species, for the purposes of understanding our species in the context of others.

The extent of care giving behaviors among the paternal line in our own species is often analyzed in terms of degree of paternity certainty. Many assumptions are made regarding the prevalence and importance of paternity uncertainty in the evolution of grandmothering behaviors and longevity. A review of the relevant literature ranging from cultural anthropology to genetics suggests that paternal discrepancy may be 1.3–3.7%, and there is evidence that rates today are similar to rates in prehistoric times, although more research needs to be done to confirm this. By evaluating a wide range of rates of paternal discrepancy, our models (adapted from Rice et al. [8]) suggest that the thresholds for selection of X-linked grandmothering traits are not dramatically influenced by paternal discrepancy, even when the rates are extremely high. Thus, there is opportunity for selection based on asymmetrical genetic relatedness, such as differential inheritance of sex chromosomes.

SA-zygotic drive may contribute to the evolution of human longevity. If the benefits of having a living grandmother are sufficiently advantageous for certain individuals, then this could lead to selection for longevity on a larger scale. Further research should probe the mechanisms by which paternal grandmothers have a beneficial effect on granddaughters and a detrimental effect on grandsons, in light of incentives for longevity. Additionally, as our understanding of functional genetics increases, finding X-linked traits influencing longevity and care giving would provide support for the hypotheses described herein.

Researchers should also further investigate the magnitude of the effects grandmothers have on different grandchildren. Finally, although attention has primarily focused on grandmothers and the X-chromosome, we think the roles of grandfathers and the Y-chromosome should also be explored in light of SA-zygotic drive.

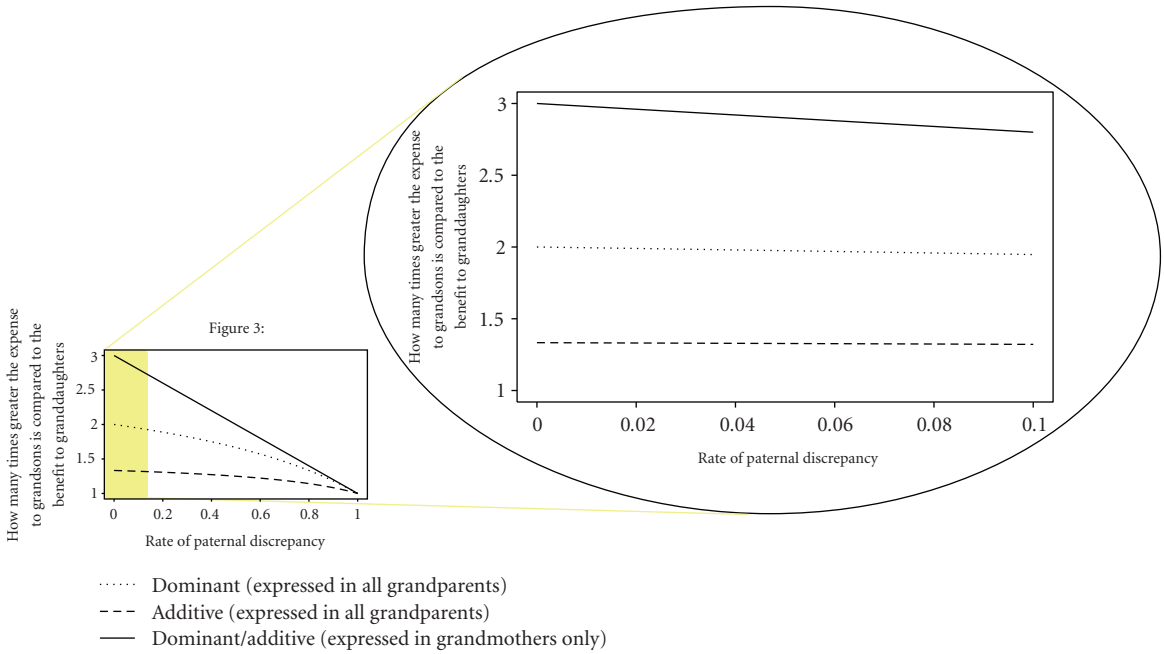


Figure 4: The threshold for an X-linked mutation causing grandparents (or grandmothers) to favor their granddaughters at the expense of their grandsons to accumulate. The threshold can be described as the maximum number of times greater the expense of this mutation would be to grandsons, compared to the benefit of the mutation to granddaughters. These values were calculated using the mathematical model from Rice et al. [8, supplement]. The range of paternal discrepancy is 0% (all paternity is identified accurately) and 10% (1 in 10 instances paternity is identified inaccurately). This range was chosen because previous studies suggest that our species' normal rates of paternal discrepancy may range from 1.3% to 3.7%, although many sources claim an unsubstantiated rate of 10%. Therefore, the range in this graph is meant to be inclusive and show a more sensitive scale of invasion threshold than Figure 3.

Acknowledgments

The authors thank Philipp Braun for his assistance with statistical analyses, and they also thank The Gates Cambridge Trust.

References

[1] W. D. Hamilton, "The genetical evolution of social behaviour. I," *Journal of Theoretical Biology*, vol. 7, no. 1, pp. 1–16, 1964.

[2] W. D. Hamilton, "The genetical evolution of social behaviour. II," *Journal of Theoretical Biology*, vol. 7, no. 1, pp. 17–52, 1964.

[3] K. G. Anderson, "The life histories of American stepfathers in evolutionary perspective," *Human Nature*, vol. 11, no. 4, pp. 307–333, 2000.

[4] K. G. Anderson, H. Kaplan, and J. Lancaster, "Paternal care by genetic fathers and stepfathers—I: reports from Albuquerque men," *Evolution and Human Behavior*, vol. 20, no. 6, pp. 405–431, 1999.

[5] J. Santrock, R. Warshak, C. Lindbergh, and L. Meadows, "Children's and parents' observed social behavior in stepfather families," *Child Development*, vol. 53, pp. 472–480, 1982.

[6] L. D. Borders, L. K. Black, and B. K. Pasley, "Are adopted children and their parents at greater risk for negative outcomes?" *Family Relations*, vol. 47, no. 3, pp. 237–241, 1998.

[7] N. Barber, *Why Parents Matter: Parental Investment and Child Outcomes*, Bergin & Garvey, London, UK, 2000.

[8] W. R. Rice, S. Gavrilets, and U. Friberg, "The evolution of sex-specific grandparental harm," *Proceedings of the Royal Society B*, vol. 277, no. 1694, pp. 2727–2735, 2010.

[9] K. Hawkes, J. F. O'Connell, and N. G. Blurton Jones, "Hadza women's time allocation, offspring provisioning, and the evolution of long postmenopausal life spans," *Current Anthropology*, vol. 38, no. 4, pp. 551–577, 1997.

[10] K. Hawkes, J. F. O'Connell, N. G. Blurton Jones, H. Alvarez, and E. L. Charnov, "The grandmother hypothesis and human evolution," in *Adaptation and Human Behavior: An Anthropological Perspective*, pp. 231–252, 2000.

[11] K. Hawkes, "Grandmothers and the evolution of human longevity," *American Journal of Human Biology*, vol. 15, no. 3, pp. 380–400, 2003.

[12] M. Fox, R. Sear, J. Beise, G. Ragsdale, E. Voland, and L. A. Knapp, "Grandma plays favourites: X-chromosome relatedness and sex-specific childhood mortality," *Proceedings of the Royal Society B: Biological Sciences*, vol. 277, no. 1681, pp. 567–573, 2010.

[13] J. A. Wilder, "Do grandmothers who play favorites sow seeds of genomic conflict?" *BioEssays*, vol. 32, no. 6, pp. 457–460, 2010.

[14] E. R. Chrastil, W. M. Getz, H. A. Euler, and P. T. Starks, "Paternity uncertainty overrides sex chromosome selection for preferential grandparenting," *Evolution and Human Behavior*, vol. 27, no. 3, pp. 206–223, 2006.

[15] W. R. Rice, S. Gavrilets, and U. Friberg, "Sexually antagonistic "zygotic drive" of the sex chromosomes," *PLoS Genetics*, vol. 4, no. 12, Article ID e1000313, 2008.

[16] E. Voland and J. Beise, "Opposite effects of maternal and paternal grandmothers on infant survival in historical Krummhörn," *Behavioral Ecology and Sociobiology*, vol. 52, no. 6, pp. 435–443, 2002.

[17] C. S. Jamison, L. L. Cornell, P. L. Jamison, and H. Nakazato, "Are all grandmothers equal? A review and a preliminary test of the "grandmother hypothesis" in Tokugawa Japan," *American Journal of Physical Anthropology*, vol. 119, no. 1, pp. 67–76, 2002.

[18] H. A. Euler and B. Weitzel, "Discriminative grandparental solicitude as reproductive strategy," *Human Nature*, vol. 7, no. 1, pp. 39–59, 1996.

[19] A. F. Kachel, L. S. Premo, and J. J. Hublin, "Grandmothering and naturalselection," *Proceedings of the Royal Society B: Biological Sciences*. In press.

[20] G. Ragsdale, "Grandmothering in Cambridgeshire, 1770–1861," *Human Nature*, vol. 15, no. 3, pp. 301–317, 2004.

[21] K. Hawkes, J. F. O'Connell, N. G. Blurton Jones, H. Alvarez, and E. L. Charnov, "Grandmothering, menopause, and the evolution of human life histories," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 3, pp. 1336–1339, 1998.

[22] K. Hill and A. M. Hurtado, "The evolution of premature reproductive senescence and menopause in human females—an evaluation of the "grandmother hypothesis"," *Human Nature*, vol. 2, no. 4, pp. 313–350, 1991.

[23] M. Lahdenperä, V. Lummaa, S. Helle, M. Tremblay, and A. F. Russell, "Fitness benefits of prolonged post-reproductive lifespan in women," *Nature*, vol. 428, no. 6979, pp. 178–181, 2004.

[24] K. Skorecki, S. Selig, S. Blazer et al., "Y chromosomes of Jewish priests," *Nature*, vol. 385, no. 6611, pp. 32–36, 1997.

[25] L. W. Simmons, R. C. Firman, G. Rhodes, and M. Peters, "Human sperm competition: testis size, sperm production and rates of extrapair copulations," *Animal Behaviour*, vol. 68, no. 2, pp. 297–302, 2004.

[26] B. Sykes and C. Irven, "Surnames the Y chromosome," *American Journal of Human Genetics*, vol. 66, no. 4, pp. 1417–1419, 2000.

[27] M. A. Bellis, K. Hughes, S. Hughes, and J. R. Ashton, "Measuring paternal discrepancy and its public health consequences," *Journal of Epidemiology and Community Health*, vol. 59, no. 9, pp. 749–754, 2005.

[28] A. M. Johnson, C. H. Mercer, B. Erens et al., "Sexual behaviour in Britain: partnerships, practices, and HIV risk behaviours," *Lancet*, vol. 358, no. 9296, pp. 1835–1842, 2001.

[29] S. Platek and T. K. Shackelford, *Female Infidelity and Paternal Uncertainty*, 2006.

[30] S. Macintyre and A. Sooman, "Non-paternity and prenatal genetic screening," *Lancet*, vol. 338, no. 8771, pp. 869–871, 1991.

[31] R. A. Johnstone and M. A. Cant, "The evolution of menopause in cetaceans and humans: the role of demography," *Proceedings of the Royal Society B: Biological Sciences*, vol. 277, no. 1701, pp. 3765–3771, 2010.

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 439046, 7 pages
doi:10.4061/2011/439046

Research Article

Persistent Copulation in Asexual Female *Potamopyrgus antipodarum*: Evidence for Male Control with Size-Based Preferences

Amanda E. Nelson and Maurine Neiman

Department of Biology, University of Iowa, 143 Biology Building, Iowa City, IA 52242-1324, USA

Correspondence should be addressed to Amanda E. Nelson, amanda-nelson@uiowa.edu

Received 30 September 2010; Revised 12 January 2011; Accepted 3 February 2011

Academic Editor: Jeremy Marshall

Transitions from sexual to asexual reproduction provide a useful context for investigating the evolutionary loss of nonfunctional traits. It is often assumed that useless behaviors or structures will degrade, but this process is poorly understood. *Potamopyrgus antipodarum* is an ancestrally sexual New Zealand freshwater snail characterized by numerous independent transitions to asexual all-female lineages. The availability of multiple independently-derived asexual lineages of various time since derivation from sexual ancestors means that the *P. antipodarum* system is well-suited for the study of trait loss related to mating behavior and copulation. Here, we asked whether mating behavior in asexual female *P. antipodarum* degrades with increasing asexual lineage age. While copulation frequency did not differ in females from old versus young asexual lineages, *post hoc* analyses indicated that it was instead positively associated with mean lineage female size. We observed that female *P. antipodarum* take a passive physical role in copulatory interactions, indicating that female behavior may not be a useful variable for detection of sex-related vestigialization in this system. Instead, males seem to be in proximate control of copulation frequencies, meaning that male mating behavior may be a primary determinant of the expression of mating behavior in asexual female *P. antipodarum*.

1. Introduction

Vestigialization, the evolutionary loss of formerly useful and presently nonfunctional morphological, molecular, and behavioral traits, is an important evolutionary process that remains poorly understood [1–3]. The multiple independent transitions from sexual to asexual reproduction across eukaryotic phylogeny [4] and the likelihood that mating is costly [5–14] provide a powerful context in which to pose questions regarding rates of vestigialization. For example, we might expect that parthenogenetic lineages that do not need to copulate or be fertilized by another individual to reproduce may experience relaxed selection for the maintenance of sexual structures and behaviors, setting the stage for trait loss [1, 5, 15–17]. Selection may even actively favor the loss of traits related to mating behavior and copulation in asexual lineages if these traits are costly [1–3]. Empirical examples of decay of mating behavior associated with transitions to asexuality include the documentation of increasing reluctance to mate in female *Drosophila mercatorum* following several generations of artificial selection for parthenogenetic reproduction [18] and the observation of decay of female mating behavior in natural strains of the parasitoid wasp *Apoanagyrus divericornis* subject to infection-induced thelytokous parthenogenesis [19].

Potamopyrgus antipodarum, a prosobranch snail native to New Zealand freshwater lakes and streams, is an excellent model system in which to examine rates of behavioral vestigialization. First, many *P. antipodarum* populations contain both obligately sexual, dioecious individuals and obligately parthenogenetic females [20, 21], such that asexual females often coexist with males in their native populations. Second, asexual *P. antipodarum* lineages are of multiple, independent origins from sexual *P. antipodarum* [21, 22] and vary in age from less than 70,000 years since derivation from sexual *P. antipodarum* ("young") to over 500,000 years

Table 1: Copulation frequency, shell length, and lineage age in asexual cultures.

Culture	Mean copulation frequency ± 1 SD	Mean shell length ± 1 SD	Lineage age	*N*
Taylor[a,b]	0.13 ± 0.13	4.27 ± 0.33	YOUNG	14
Mapourika[a]	0.13 ± 0.18	4.32 ± 0.31	OLD	13
Denmark A*	0.15 ± 0.10	4.30 ± 0.24	OLD	5
Duluth*	0.16 ± 0.37	4.70 ± 0.22	OLD	4
Okareka[a,b]	0.18 ± 0.11	4.37 ± 0.58	OLD	15
Sarah[a]	0.19 ± 0.15	4.54 ± 0.50	YOUNG	14
Poerua	0.24 ± 0.15	4.53 ± 0.21	YOUNG	14
Waikaremoana Lab**	0.25 ± 0.15	4.65 ± 0.47	OLD**	19
Evelyn	0.31 ± 0.16	4.67 ± 0.32	YOUNG	13
Waikaremoana Field**	0.32 ± 0.15	4.65 ± 0.42	OLD**	19
Tarawera[c]	0.35 ± 0.15	5.09 ± 0.82	OLD	14
Taupo[c]	0.39 ± 0.21	5.60 ± 0.61	OLD	15
All females	0.25 ± 0.17	4.67 ± 0.59	—	159

Culture of origin (rank ordered by increasing copulation frequency), mean copulation frequency ± SD, mean shell length ± SD, lineage age, and number of individual asexual females used from each culture. "Young" indicates lineages derived from sexual ancestors <70,000 years ago, while "old" indicates lineages derived >500,000 years ago [23]. *Denotes the only two population names that do not refer to lakes of origin in New Zealand; the ancestor of the Denmark A lineage was collected in Denmark in the early 1990s, and the ancestor of the Duluth lineage was collected in 2007 in Minnesota, USA from Lake Superior. **Waikaremoana populations consist of multiple lineages of old asexuals. [a]Denotes cultures that differed significantly from Taupo in mean copulation frequency via Bonferroni-corrected chi-square comparisons. [b]Denotes significant differences in copulation frequency from Tarawera. [c]Taupo differed significantly in mean length from all cultures except Tarawera (*t*-tests). All differences are significant at $P < .05$.

since derivation ("old") [23], allowing for the comparison of behavioral consequences of the absence of sex on differential evolutionary time scales. Previous studies showed no difference in copulation frequency of male *P. antipodarum* with sexual versus asexual females from young lineages when males were housed simultaneously with both types of female [24] and no apparent reproductive benefits of copulation to asexual females [25].

Here, our goal was to generate new insights into the persistence of mating behavior in asexual *P. antipodarum* through comparison of copulatory behavior in multiple young and old asexual lineages. We predicted that if mating is sufficiently costly, mating behavior will be degraded in old lineages relative to their younger counterparts. We found that there are no differences in asexual female copulation frequencies based on lineage age. Instead, *post hoc* analysis shows that asexual female copulation frequency increases with female body size and suggests that male *P. antipodarum* are in physical control of copulatory behavior. Male control of copulation would mean that copulation frequencies do not represent physical actions on the part of females and implies that copulation frequencies may not reflect female evolutionary change or lack thereof.

2. Materials and Methods

2.1. Experimental Methods. We haphazardly selected 170 sexually mature (>3 mm in length [26, 27]) asexual female *P. antipodarum* from twelve lab-raised lineages or field-collected cultures that consisted of either young (<70,000 years since derivation from sexual ancestors) or old (>500,000 years since derivation from sexual ancestors) asexual clades (Table 1). Time since derivation from sexual ancestors was previously established for source populations through phylogenetic analysis of cytochrome b mtDNA haplotypes [23]. Lab-raised and field-collected subsets of snails of common origin will hereafter be referred to as "cultures." To control for prior exposure to males, field-collected females were selected from samples with <5% males, indicating that males—and thus sexuals—are rare in source populations [28, 29]. We haphazardly collected sexually mature males (possessing a visible penis [30]) from a culture consisting of three inbred sexual lineages derived from a single New Zealand lake (Lake Alexandrina) and three cultures from New Zealand lakes with high relative frequencies of sexual *P. antipodarum* (Lakes Kaniere, Rotoroa, and Wairarapa). We used 16–56 males from each of these cultures according to male availability. We also included 15 males (individual identity unknown) drawn from a mixture of inbred Lake Alexandrina cultures and a culture from Lake Kaniere. We recorded the culture of origin for all *P. antipodarum* and used nail polish to make a culture-specific mark on each individual.

Each asexual female was housed separately in a 0.95 L plastic cylindrical cup containing ~700 mL of water. We added an ~1 cm^3 piece of chalk to each cup as a source of dietary calcium. Each snail was fed 1 g/L *Spirulina* (a common lab food for *P. antipodarum*) 3 × weekly. Cups were maintained at 16°C on a 12 : 12 hr light:dark cycle. Females that died were replaced for the first week only, resulting in 159 females at experiment end. One male was placed in each cup and redistributed to a different cup ("rotated") twice weekly, so that male-female pairs from all possible culture combinations were observed and to control for variation in male mating behavior across cultures. Five males of mixed

Alexandrina/Kaniere origin were distributed randomly after all other males had been placed in containers. The four-week experiment included eight rotations (two per week), each constituting a single "copulation opportunity" for a distinct male-female pair, resulting in a total of 1272 pairwise rotations (or "opportunities"). In other words, each distinct snail pair had a 2.5 day "opportunity" to copulate.

2.2. Observation of Courtship and Copulatory Behavior. Specific behaviors observed repeatedly and involving physical contacts between males and females were recorded (see Supplemental Table 1 in Supplementary Material available online at doi: 10.4061/2011/439046) five days a week for four weeks at approximately 830, 1100, 1300, and 1500 hours. Each observation lasted 15–20 minutes, which was the time required to assess the behaviors for each pair. While we recorded four distinct behaviors, analyses focused solely on copulation as this behavior was most explicitly defined and the only behavior observed with sufficient frequency to be amenable to meaningful statistical interpretation (supplemental Table 2). Copulation in *P. antipodarum* has been previously specified as the assumption of an aperture-to-aperture position maintained over ~20–90 minutes [25]. All 416 observed aperture to aperture contacts occurred with the male situated in a dextralateral position on the female's shell, consistent with Fretter's [31] observations of specific mating positions in other prosobranch taxa. This distinct position was classified as copulation. Further, each observation of the copulatory position was examined using 10x magnification in order to confirm that body-to-body contact was occurring between male-female pairs.

The other three behaviors included "resistance," defined as a female mounted by a male twisting so that her body made an abrupt radial motion. The same behavior was observed when males housed without females were mounted by other males. The behavior we termed "riding" was defined as a male mounted on a female shell away from her aperture and remaining stationary while the female foraged. "Checking" was defined as the male actively, often (relatively) rapidly circling the female's shell.

2.3. Statistical Analysis. Copulation data were discrete and left skewed, such that nonparametric analyses were required. We analyzed copulation frequencies in terms of the proportion of copulation opportunities (as defined above) in which copulation was observed within cultures. We performed 2×2 chi-square tests to determine whether there were differences in the proportion of opportunities in which members of young versus old asexual cultures copulated. We also performed 2×2 chi-square tests to detect differences in copulation duration and repetition, respectively, between young and old asexual cultures, but these results are not included as they were qualitatively the same as results using the proportion of opportunities in which copulation was observed. We also used chi-square tests within old cultures (2×8), within young cultures (2×4), among all cultures combined (2×12), and pairwise between all combinations of cultures (66 2×2 tests, Bonferroni-corrected threshold at $P = .0008$) in order to compare variation between old and young cultures to overall variation in copulation frequency. All chi-square comparisons were performed by hand.

2.4. Post Hoc Analysis. We observed marked variation across asexual female cultures in both copulation frequency and size. Since female *P. antipodarum* possess notable heritable variation in shell size [32], and because males from other snail taxa prefer to copulate with larger females [33, 34], we then asked whether across-culture size differences affected copulation frequencies. We quantified size in each female at the end of the experiment by measuring shell length (longest dorsal distance in mm from aperture to apex of shell) and shell width (mm across widest portion of shell). We also used these data to quantify "obesity" (width : length), following Winterbourn [35]. We used two-tailed Spearman's rho coefficients in SPSS to detect correlations (across all pairwise variable combinations) between mean culture copulation frequency and mean culture length, width, and obesity. Because mean length and width were positively correlated with mean copulation frequency and mean obesity was not (see results), we used stepwise linear regression to determine whether length and/or width were significant predictors of copulation frequency. Next, we used Bonferroni-corrected pairwise Student's t-tests to quantify culture-level differences in female length (66 t-tests for each parameter, Bonferroni threshold at $P = .0008$).

478 out of 511 observed physical contacts between males and asexual females consisted of apparent physical activity only on the part of the male (see Supplemental Tables 1-2), suggesting that male *P. antipodarum* may control copulation frequency. Chi-square tests were used to detect differences in copulation frequency across all four male cultures (2×4 test) and between pairs of male cultures (6 2×2 tests, Bonferroni threshold at $P = .008$). Data from the mixed Alexandrina/Kaniere origin males were not included in these analyses because specific male origins were unknown. Among both males and females, we observed marked differences between cultures with the highest and lowest overall mean copulation frequencies. These differences provided an opportunity to compare the effects of male versus female culture on copulation frequencies. In order to do this, we used the highest and lowest relative mean copulation frequencies for pairwise comparison of male-female culture combinations in a Scheirer-Ray-Hare extension of a Kruskal-Wallis test. If males control copulation frequency, we expected to find a marked effect of male culture and little or no effect of female culture on observed pairwise copulation frequencies.

3. Results

We observed 416 copulations involving 159 females over the course of 314 copulation opportunities (as defined above) resulting in copulation. On average, females copulated 2.62 ± 2.11 SD times. This value is similar in magnitude to observations reported in Neiman and Lively (2005) in which mean number of copulations for sexual and asexual

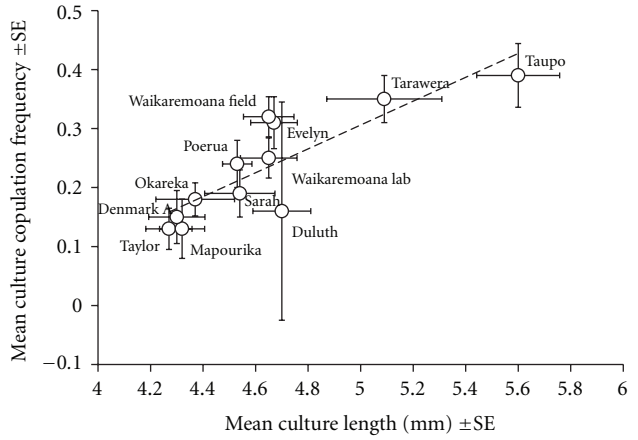


Figure 1: Plot of mean asexual female culture copulation frequency ± SE versus mean culture shell length ± SE. Mean copulation frequency increases with mean shell length for asexual females.

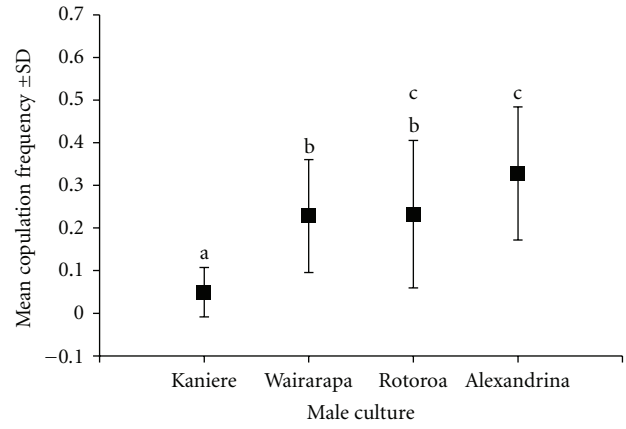


Figure 2: Differences in mean copulation frequency (±1 standard deviation) for all four male cultures. Letters in common above error bars indicate $P > .05$ in a Bonferroni-corrected chi-square test for differences in copulation frequency. Bonferroni-adjusted P values for significant differences are as follows: P(a/b) = .0012–.00012, P(b/c) = .009, and P(a/c) = .00006.

females were 3.49 ± 0.47 SD and 2.58 ± 0.40 SD, respectively [24]. 78.61% of copulations lasted 20 minutes to 1.5 hours, within the range reported in Neiman and Lively (2005). 14.66% of copulations lasted from two to four hours, 4.33% lasted between four and six hours, and 1.92% lasted six or more hours. Our observations of copulation durations substantially longer than those in the previous study may be due to larger sample sizes revealing greater variation in this behavior. Another nonmutually exclusive explanation may be differences in experimental design (male-female pairs in our study versus multiple males housed with multiple females in Neiman and Lively (2005)) [24].

There were no significant differences in the proportion of opportunities in which females from old versus young cultures copulated ($\chi^2 = 3.15$, $P = .0759$, df = 1; Table 1). In contrast, there was marked across-lineage variance in copulation frequency for old asexual female cultures ($\chi^2 = 32.50$, $P = .0003$, df = 7), young asexual female cultures ($\chi^2 = 10.84$, $P = .0126$, df = 3), and across all asexual female cultures ($\chi^2 = 47.33$, $P < .00001$, df = 11).

Copulation frequency among all asexual female cultures was positively correlated with both length ($\rho = 0.824$, $P = .001$, $n = 12$; Figure 1) and width ($\rho = 0.639$, $P = .025$, $n = 12$). Length and width were not correlated ($\rho = 0.481$, $P = .114$, $n = 12$). Obesity was correlated only with width ($\rho = 0.931$, $P < .001$, $n = 12$), not with copulation frequency ($\rho = 0.351$, $P = .263$, $n = 12$). While length significantly and positively affected copulation frequency in a stepwise linear regression ($\beta = 0.732$, $P = .039$, $n = 12$), no further analysis was conducted with width since it was not associated with copulation frequency ($\beta = 0.088$, $P = .778$, $n = 12$). Bonferroni-corrected t-tests showed that one female culture, Taupo, was significantly longer than all other female cultures except Tarawera (Table 1), such that the longest cultures also exhibited the highest mean copulation frequencies relative to other cultures. The mean length of females from old cultures (4.75 mm ± 0.67 SD) was similar to that of young cultures (4.50 mm ± 0.38 SD) and the longest cultures differed from the shortest cultures irrespective of lineage age (Table 1). Across all cultures, shell length was a significant predictor of copulation frequency, but only between the low and high ends of the distribution of female shell lengths (Table 1; Figure 1). Females from Mapourika, Taylor, and Okareka represented the lowest mean copulation frequency and females from Taupo, Tarawera, and Waikaremoana Field represented the highest mean copulation frequency cultures. These were the cultures we used for Scheirer-Ray-Hare analysis. Denmark A and Duluth cultures were excluded from this analysis due to small sample sizes (Table 1).

There were significant differences in copulation frequency among male cultures ($\chi^2 = 42.91$, $P < .00001$, df = 3). Alexandrina and Kaniere males differed most in copulation frequency (Alexandrina having the highest overall mean copulation frequency and Kaniere the lowest, Figure 2). These two male cultures were included in the Scheirer-Ray-Hare test. This test revealed an effect of male culture on mean pairwise copulation frequency at high and low extremes of copulation frequency (SS = 75, H = 5.98, $P = .015$, df = 1) but showed no effect of female culture (SS = 12.12, H = 0.97, $P = .325$, df = 1) and no interaction between male and female culture (SS = 1.38, H = 0.11, $P = .740$, df = 1).

4. Discussion

There was no apparent effect of lineage age on copulation frequency in female asexual *P. antipodarum*. Instead, differences in copulation frequency seem primarily to be a function of differences in female shell length (Figure 1). Further analyses suggest that male rather than female copulatory tendencies might largely determine copulation frequencies and that female size differences might provide a basis for male mating preferences.

In the absence of a lineage age effect, we find no evidence of vestigialization of reproductive behaviors. In fact, the significant variation in copulation frequency within age classes shows that there is considerably more variation within than between old and young cultures. Nonmutually exclusive explanations for the apparent lack of a role for asexual lineage age in determining female copulatory behavior include insufficient time for neutral mutation and/or natural selection to degrade the genetic basis of useless or costly female reproductive behaviors, that mating behavior is maintained by selection, or that apparently old asexual lineages have an undiscovered recent sexual ancestor (reviewed in [36]). It is also possible that maintaining mechanisms of mate choice, such as the avoidance of undesirable mates, may be costly to females. Some have even speculated that receptivity to mating may actually increase in asexual females depending on which sex-related genes (i.e., genes associated with mating versus those associated with avoidance of mating) undergo relaxed selection [37].

It is also possible that asexual female *P. antipodarum* gain nonreproductive benefits from copulation, as in several invertebrate taxa. For instance, seminal fluid includes components that increase female longevity in the field cricket, *Gryllus lineaticeps* [38], and the bedbug, *Cimex lectilarius* [39]. Multiply-mated female *Callobruchus maculatus* (seed beetles) show higher tolerance of starvation conditions [10, 40], but under normal conditions this benefit does not overcome a cost of harassment by males [10]. Traits that have lost a given function may be maintained by selection if a novel function is gained that increases fitness [1]. Under a model of antagonistic sexual coevolution [41], it is expected that internal female reproductive structures, such as sperm storage organs, are more likely to be modified by natural or sexual selection than external female genitalia [42]. Dillon [43] observed that structures originally associated with sperm storage are maintained in asexual *P. antipodarum* and assumed that these structures retain their original function. However, it has been suggested that one such structure, the bursa copulatrix, is modified in asexual female *P. antipodarum* to function for digestion of waste materials [44]. If the bursa copulatrix serves a digestive function in asexual female *P. antipodarum* and still receives sperm, it could provide the basis for a nutritive benefit of copulation. However, in taxa where putative nutritive benefits of ejaculate components have been extensively studied, the proximate mechanism by which such benefits are conferred remains unclear (e.g., [10]).

Post hoc comparisons indicating that male rather than female culture-level tendencies affect mean copulation frequencies points to male control as one possible determinant of relative differences in asexual female copulation frequencies. Further, copulation in *P. antipodarum* occurs via male mounting of the female shell, a mechanism which is thought to confer greater control to the snail on top [45]. Regardless of the evolutionary implications of persistent copulatory behavior in asexual female *P. antipodarum*, male control of copulation implies male control of copulation frequencies [46]. This would mean that copulatory behavior may not be a useful metric by which to detect sex-related trait changes in females. However, the proximate physical mechanisms of mating observed herein may belie an important role for females; chemical signaling via mucus may play an important role in prosobranch snail mating [31]. In other words, female *P. antipodarum* may control copulation frequencies via biochemical or physiological traits not measured in this experiment.

The finding that longer females engage in more copulation has important implications for copulatory control. Female *P. antipodarum* show a positive relationship between size and fecundity [47, 48]. This, along with the observed effect of length on copulation frequency in *P. antipodarum* females, suggests that male mating behavior may be influenced by length as a signal for fecundity. That length most affects copulation frequency at the high and low extremes of size is not without precedent. For example, in guppies (*Poecilia reticulata*), male mating preference based on size is markedly stronger as the relative size difference between females increases [49]. Male preference for larger mates in *Drosophila melanogaster* reduces variance in female relative fitness because larger, otherwise more fecund females suffer fitness costs from increased harassment by males relative to their smaller counterparts [12].

It is a standard assumption that females are more likely to be the "choosy" gender [50], yet male preferences, though little studied, may dominate copulation dynamics [12, 51]. If male *P. antipodarum* are both choosy and in proximate control of copulation, why do they continue to copulate with asexual females? One possibility is that males basing mating preference on female size cannot discriminate between sexual and asexual female *P. antipodarum*, which do not consistently differ in size [32]. This phenomenon has been documented in hybridogenetic water frogs, in which indiscriminant male mating behavior is more likely to persist when sexual and asexual (hybridogenetic) females are of similar size [51]. Alternately, males may be doing the "best of a bad job" by balancing the cost of discrimination versus the cost of futile mating [52].

One caveat of our study was that in only using asexual females from all-asexual source populations, we were unable to ask whether sympatry with sexual males and/or females may drive differences in asexual female copulation frequencies. We did, however, use males from source populations that vary widely in their proportions of asexual females. Interestingly, the male culture that copulated most with asexual females (Alexandrina) came from a source population with a high relative proportion of asexual females (50%). Conversely, the male culture with the lowest copulation frequency with asexual females (Kaniere) was from a nearly all-sexual source population. Likewise, the two male cultures with intermediate copulation frequencies (Wairarapa and Rotoroa) came from source populations with ~10% asexuals (Neiman, unpublished). That males with greater prior exposure to asexual females copulated with these females more often is counter to the prediction that males are more likely to discriminate when the likelihood of pairing with an undesirable female is higher [52].

The persistence of copulation between male and asexual female *P. antipodarum* appears on the surface to be

maladaptive for both. Our results indicate that broad assumptions about circumstances favoring vestigialization of characters may belie very complex organismal biology. *Potamopyrgus antipodarum* as a system reflects the dynamic behavioral and morphological factors that may influence reproductive biology. Far more detailed observation and analysis is needed in order to disentangle the evolutionary forces at work in this unique and interesting system.

Acknowledgments

Thanks are due to the Guest Editor, Dr. Jeremy Marshall, and two anonymous reviewers for constructive and insightful comments on our paper. The authors thank members of the Neiman laboratory for logistic support and comments on an earlier version of the paper. Thanks are due to Christina Schiltz for edits, moral support, and helping to set up this experiment. Thanks are also due to Claire Tucci for help with our experimental procedure, and Dr. Stephen Hendrix for editorial advice. Thanks also are due to Dr. Jonathan Poulton, the University of Iowa Honors Program, and the University of Iowa, Department of Biology.

References

[1] D. C. Lahti, N. A. Johnson, B. C. Ajie et al., "Relaxed selection in the wild," *Trends in Ecology and Evolution*, vol. 24, no. 9, pp. 487–496, 2009.

[2] D. W. Fong, T. C. Kane, and D. C. Culver, "Vestigialization and loss of nonfunctional characters," *Annual Review of Ecology and Systematics*, vol. 26, no. 1, pp. 249–268, 1995.

[3] M. E. Dorken, K. J. Neville, and C. G. Eckert, "Evolutionary vestigialization of sex in a clonal plant: selection versus neutral mutation in geographically peripheral populations," *Proceedings of the Royal Society B*, vol. 271, no. 1555, pp. 2375–2380, 2004.

[4] M. Neiman, S. Meirmans, and P. G. Meirmans, "What can asexual lineage age tell us about the maintenance of sex?" *Annals of the New York Academy of Sciences*, vol. 1168, pp. 185–200, 2009.

[5] M. Daly, "The cost of mating," *The American Naturalist*, vol. 112, no. 986, pp. 771–774, 1978.

[6] A. E. Magurran and B. H. Seghers, "A cost of sexual harassment in the guppy, *Poecilia reticulata*," *Proceedings of the Royal Society B*, vol. 258, no. 1351, pp. 89–92, 1994.

[7] S. F. Chenoweth, D. Petfield, P. Doughty, and M. W. Blows, "Male choice generates stabilizing sexual selection on a female fecundity correlate," *Journal of Evolutionary Biology*, vol. 20, no. 5, pp. 1745–1750, 2007.

[8] D. J. Hosken and T. A. R. Price, "Genital evolution: the traumas of sex," *Current Biology*, vol. 19, no. 13, pp. R519–R521, 2009.

[9] P. W. Bateman, J. W. H. Ferguson, and C. A. Yetman, "Courtship and copulation, but not ejaculates, reduce the longevity of female field crickets (*Gryllus bimaculatus*)," *Journal of Zoology*, vol. 268, no. 4, pp. 341–346, 2006.

[10] M. den Hollander and D. T. Gwynne, "Female fitness consequences of male harassment and copulation in seed beetles, *Callosobruchus maculatus*," *Animal Behaviour*, vol. 78, no. 5, pp. 1061–1070, 2009.

[11] L. Gay, P. E. Eady, R. Vasudev, D. J. Hosken, and T. Tregenza, "Costly sexual harassment in a beetle," *Physiological Entomology*, vol. 34, no. 1, pp. 86–92, 2009.

[12] T. A.F. Long, A. Pischedda, A. D. Stewart, and W. R. Rice, "A cost of sexual attractiveness to high-fitness females," *PLoS Biology*, vol. 7, no. 12, Article ID e1000254, 2009.

[13] I. Schlupp, R. M. Knab, and M. J. Ryan, "Sexual harassment as a cost for molly females: bigger males cost less," *Behaviour*, vol. 138, no. 2, pp. 277–286, 2001.

[14] K. U. Heubel and M. Plath, "Influence of male harassment and female competition on female feeding behaviour in a sexual-asexual mating complex of mollies (*Poecilia mexicana, P. formosa*)," *Behavioral Ecology and Sociobiology*, vol. 62, no. 11, pp. 1689–1699, 2008.

[15] H. J. Muller, "The Darwinian and modern conceptions of natural selection," *Proceedings of the American Philosophical Society*, vol. 93, no. 6, pp. 459–470, 1949.

[16] H. L. Carson, L. S. Chang, and T. W. Lyttle, "Decay of female sexual behavior under parthogenesis," *Science*, vol. 218, no. 4567, pp. 68–70, 1982.

[17] M. Neiman, "Physiological dependence on copulation in parthenogenetic females can reduce the cost of sex," *Animal Behaviour*, vol. 67, no. 5, pp. 811–822, 2004.

[18] H. Ikeda and H. L. Carson, "Selection for mating reluctance in females of a diploid parthenogenetic strain of *Drosophila mercatorum*," *Genetics*, vol. 75, no. 3, pp. 541–555, 1973.

[19] J. W. A. M. Pijls, H. J. van Steenbergen, and J. J. M. van Alphen, "Asexuality cured: the relations and differences between sexual and asexual *Apoanagyrus diversicornis*," *Heredity*, vol. 76, no. 5, pp. 506–513, 1996.

[20] C. Wallace, "Parthenogenesis, sex and chromosomes in *Potamopyrgus*," *Journal of Molluscan Studies*, vol. 58, no. 2, pp. 93–107, 1992.

[21] M. F. Dybdahl and C. M. Lively, "Diverse, endemic and polyphyletic clones in mixed populations of a freshwater snail (*Potamopyrgus antipodarum*)," *Journal of Evolutionary Biology*, vol. 8, no. 3, pp. 385–398, 1995.

[22] M. Neiman and C. M. Lively, "Pleistocene glaciation is implicated in the phylogeographical structure of *Potamopyrgus antipodarum*, a New Zealand snail," *Molecular Ecology*, vol. 13, no. 10, pp. 3085–3098, 2004.

[23] M. Neiman, J. Jokela, and C. M. Lively, "Variation in asexual lineage age in *Potamopyrgus antipodarum*, a New Zealand snail," *Evolution*, vol. 59, no. 9, pp. 1945–1952, 2005.

[24] M. Neiman and C. M. Lively, "Male New Zealand mud snails (*Potamopyrgus antipodarum*) persist in copulating with asexual and parasitically castrated females," *American Midland Naturalist*, vol. 154, no. 1, pp. 88–96, 2005.

[25] M. Neiman, "Embryo production in a parthenogenetic snail (*Potamopyrgus antipodarum*) is negatively affected by the presence of other parthenogenetic females," *Invertebrate Biology*, vol. 125, no. 1, pp. 45–50, 2006.

[26] D. C. Richards and D. C. Shinn, "Intraspecific competition and development of size structure in the invasive snail *Potamopyrgus antipodarum* (Gray, 1853)," *American Malacological Bulletin*, vol. 19, no. 1-2, pp. 33–37, 2004.

[27] M. F. Dybdahl and S. L. Kane, "Adaptation vs. phenotypic plasticity in the success of a clonal invader," *Ecology*, vol. 86, no. 6, pp. 1592–1601, 2005.

[28] J. A. Fox, M. F. Dybdahl, J. Jokela, and C. M. Lively, "Genetic structure of coexisting sexual and clonal subpopulations in a freshwater snail (*Potamopyrgus antipodarum*)," *Evolution*, vol. 50, no. 4, pp. 1541–1548, 1996.

[29] C. M. Lively and J. Jokela, "Temporal and spatial distributions of parasites and sex in a freshwater snail," *Evolutionary Ecology Research*, vol. 4, no. 2, pp. 219–226, 2002.

[30] M. Haase, "The radiation of hydrobiid gastropods in New Zealand: a revision including the description of new species based on morphology and mtDNA sequence information," *Systematics and Biodiversity*, vol. 6, no. 1, pp. 99–159, 2008.

[31] V. Fretter, "Prosobranchs," in *The Mollusca, Volume 7: Reproduction*, A. S. Tompa, N. H. Verdonk, and J. A. M. van den Biggelaar, Eds., Academic Press, Orlando, Fla, USA, 1984.

[32] J. Jokela, C. M. Lively, M. F. Dybdahl, and J. A. Fox, "Evidence for a cost of sex in the freshwater snail *Potamopyrgus antipodarum*," *Ecology*, vol. 78, no. 2, pp. 452–460, 1997.

[33] O. V. Ambrogio and J. A. Pechenik, "Do sex-changing male snails use mate choice to get a jump on their "size advantage"?" *Marine Biology*, vol. 156, no. 10, pp. 2173–2180, 2009.

[34] J. M. Koene, K. Montagne-Wajer, and A. Maat, "Aspects of body size and mate choice in the simultaneously hermaphroditic pond snail *Lymnaea stagnalis*," *Animal Biology*, vol. 57, no. 2, pp. 247–259, 2007.

[35] M. Winterbourn, "The New Zealand species of *Potamopyrgus* (Gastropoda: Hydrobiidae)," *Malacologia*, vol. 10, pp. 283–321, 1970.

[36] A. M. Schurko, M. Neiman, and J. M. Logsdon, "Signs of sex: what we know and how we know it," *Trends in Ecology and Evolution*, vol. 24, no. 4, pp. 208–217, 2009.

[37] K. Kraaijeveld, P. Franco, B. M. Reumer, and J. J. M. van Alphen, "Effects of parthenogenesis and geographic isolation on female sexual traits in a parasitoid wasp," *Evolution*, vol. 63, no. 12, pp. 3085–3096, 2009.

[38] W. E. Wagner, R. J. Kelley, K. R. Tucker, and C. J. Harper, "Females receive a life-span benefit from male ejaculates in a field cricket," *Evolution*, vol. 55, no. 5, pp. 994–1001, 2001.

[39] K. Reinhardt, R. A. Naylor, and M. T. Siva-Jothy, "Ejaculate components delay reproductive senescence while elevating female reproductive rate in an insect," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 51, pp. 21743–21747, 2009.

[40] C. W. Fox, "Multiple mating, lifetime fecundity and female mortality of the bruchid beetle, *Callosobruchus maculatus* (Coleoptera: Bruchidae)," *Functional Ecology*, vol. 7, no. 2, pp. 203–208, 1993.

[41] M. Andersson, *Sexual Selection*, Princeton University Press, Princeton, NJ, USA, 1994.

[42] D. Joly and M. Schiffer, "Coevolution of male and female reproductive structures in *Drosophila*," *Genetica*, vol. 138, no. 1, pp. 105–118, 2009.

[43] R. T. Dillon, *The Ecology of Freshwater Molluscs*, Cambridge University Press, Cambridge, UK, 2000.

[44] V. Fretter and A. Graham, *British Prosobranch Molluscs: Their Functional Anatomy and Ecology*, Ray Society, London, UK, 1962.

[45] T. Asami, R. H. Cowie, and K. Ohbayashi, "Evolution of mirror images by sexually asymmetric mating behavior in hermaphroditic snails," *The American Naturalist*, vol. 152, no. 2, pp. 225–236, 1998.

[46] M. Ridley, "The control and frequency of mating in insects," *Functional Ecology*, vol. 4, no. 1, pp. 75–84, 1990.

[47] E. Gaino, F. Scoccia, T. Lancioni, and A. Ludovisi, "The invader mudsnail *Potamopyrgus antipodarum* in the Tiber River basin (Central Italy)," *Italian Journal of Zoology*, vol. 75, no. 3, pp. 253–261, 2008.

[48] T. M. Tibbets, A. C. Krist, R. O. Hall Jr., and L. A. Riley, "Phosphorus-mediated changes in life history traits of the invasive New Zealand mudsnail (*Potamopyrgus antipodarum*)," *Oecologia*, vol. 163, no. 3, pp. 549–559, 2010.

[49] L. D. Dosen and R. Montgomerie, "Female size influences mate preferences of male guppies," *Ethology*, vol. 110, no. 3, pp. 245–255, 2004.

[50] M. Daly and M. Wilson, *Sex, Evolution, and Behavior: Adaptations for Reproduction*, Duxbury Press, Belmont, Calif, USA, 1978.

[51] D. S. Schmeller, R. O'Hara, and H. Kokko, "Male adaptive stupidity: male mating pattern in hybridogenetic frogs," *Evolutionary Ecology Research*, vol. 7, no. 7, pp. 1039–1050, 2005.

[52] T. J. Kawecki, "Unisexual/bisexual breeding complexes in Poeciliidae: why do males copulate with unisexual females?" *Evolution*, vol. 42, no. 5, pp. 1018–1023, 1988.

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 485460, 11 pages
doi:10.4061/2011/485460

Research Article

Male-Female Interactions and the Evolution of Postmating Prezygotic Reproductive Isolation among Species of the Virilis Subgroup

Nada Sagga and Alberto Civetta

Department of Biology, University of Winnipeg, 515 Portage Avenue, Winnipeg, MB, Canada R3B 2E9

Correspondence should be addressed to Alberto Civetta, a.civetta@uwinnipeg.ca

Received 25 October 2010; Accepted 3 February 2011

Academic Editor: Jeremy Marshall

Reproductive isolation reduces breeding between species. Traditionally, prezygotic and postzygotic barriers to reproduction have been broadly studied, but in recent years, attention has been brought to the existence of barriers that act after copulation but before fertilization. Here, we show that when *D. virilis* females from different geographic locations mate with *D. novamexicana* males, egg laying is normal, but fertilization rates are severely reduced, despite normal rates of sperm transfer. This reduction in fertilization is probably due to lower retention of heterospecific sperm in female storage organs one-to-two days after copulation. An inspection of egg hatchability in crosses between females and males from other virilis subgroup species reveals that isolation due to poor egg hatchability likely evolved during the diversification of *D. virilis/D. lummei* from species of the novamexicana-americana clade. Interestingly, the number of eggs laid by *D. virilis females* in heterospecific crosses was not different from the numbers of eggs laid in conspecific crosses, suggesting that females exert some form of cryptic control over the heterospecific ejaculate and that future studies should focus on how female and female-sperm interactions contribute to the loss or active exclusion of heterospecific sperm from storage.

1. Introduction

Species can be reproductively isolated by a variety of mechanisms that as a whole reduce gene flow between them. These mechanisms can be broadly classified into prezygotic and postzygotic isolating barriers [1]. A large number of studies have drawn attention to isolating barriers in which gene flow is reduced after mating but prior to zygote formation [2–10]. While some of these barriers are competitive, others are a consequence of male-female incompatibilities and are noncompetitive, disrupting the sperm's capacity to reach and/or fertilize an egg. In Drosophila, noncompetitive isolating mechanisms have been well described. Sperm transfer from *D. arizonae* males to *D. mojavensis* females generates an insemination reaction in which a large mass forms in the uterus that obstructs ovulation and ultimately fertilization [9]. In crosses between *D. simulans* females and *D. sechellia* males, fewer sperm are transferred, and egg-laying and hatchability are reduced [5]. Sperm are also depleted more quickly in crosses between *D. santomea* and *D. yakuba* [10].

The virilis subgroup consists of five species: *D. virilis*, *D. lummei*, *D. novamexicana*, *D. americana americana*, and *D. americana texana*. There is evidence for both premating and postmating isolation among these five species. *Drosophila virilis* females have the highest crossability with all heterospecific males of the subgroup, whereas *D. virilis* males show the strongest courtship discrimination against all heterospecific females with the result that very low numbers of hybrids are produced [11]. A recent study has shown that when *D. novamexicana* females are exposed to *D. virilis* males for up to two weeks, only 14% of females produce progeny. Moreover, *D. virilis* males are able to recognize heterospecific females at the first stage of courtship (tapping), indicating strong premating isolation [12]. For the reciprocal cross between *D. virilis* females and *D. novamexicana* males, there is evidence of strong postzygotic isolation with 7% hybrid

male fertility [13]. *Drosophila virilis* females mated with *D. lummei* males produce 95% fertile hybrid males, but 47% of the hybridizations die before becoming pupae, and 25% fail to emerge from their case [14]. When *D. virilis* females mate with *D. a. texana* males, one-third of the hybrid male progeny is sterile [13, 15].

The occurrence of postmating prezygotic isolation (PPI) barriers among species of the virilis subgroup has also been documented. For example, early studies showed that *D. virilis* females mated to heterospecific males produced very few offspring. These studies suggested, although qualitatively, that the low production of progeny was due to sperm immobility (mortality) in the seminal receptacle of the females' reproductive tract for crosses between *D. virilis* females and *D. a. americana* or *D. a. texana* males but not for crosses involving *D. novamexicana* males [17, page 489]. A recent study has shown a low rate of egg hatchability when *D. virilis* females are mated to *D. a. americana* males with most of the laid eggs not being fertilized [16].

Drosophila virilis females mated to *D. novamexicana* males have previously shown to hatch a low proportion of eggs after mating, but the reasons for such low egg hatchability remain unclear [17]. Here, we explore mechanisms of reproductive isolation between *D. virilis* females and heterospecific males by mainly, but not exclusively, focusing on crosses with *D. novamexicana* males. We first test strains of different geographic origin to confirm that prior reports of low hatchability for this cross are not strain dependant. Secondly, we use outbred populations to rule out the possibility that laboratory inbreeding might create or exacerbate some of the isolation phenotypes detected. Thirdly, we show that females hatch low proportion of eggs after a heterospecific cross but lay similar numbers of eggs in heterospecific and conspecific crosses. Fourthly, within the virilis subgroup, we show that the only cross for which isolation between species due to low hatching of eggs has not evolved is that between *D. virilis* females and *D. lummei* males. Finally, using outbred *D. virilis* females and outbred *D. novamexicana* males, we show that the low hatchability of eggs is not a consequence of embryo mortality but due to problems associated with female storage of heterospecific sperm.

2. Materials and Methods

2.1. Drosophila Species and Maintenance. Five geographically diverse strains of *D. virilis* (Argentina 15010-1051.49, California 15010-1051.00, Japan 15010-1051.09, Mexico 15010-1051.48, and Russia 15010-1051.52), two *D. novamexicana* strains (New Mexico 1301.08 and Utah 1301.08), a strain of *D. lummei* (Japan 1011.08), and a strain of *D. americana texana* (wild type 1041.16) were obtained from the San Diego Drosophila Stock Center. The number of eggs laid and the proportion of eggs hatched can be affected by inbreeding in laboratory strains of Drosophila, so we also established outbred populations of *D. virilis* and *D. novamexicana* by mixing equal numbers of individuals from the different strains in a population cage for five generations. Flies were reared in round-bottom bottles (64 × 130 mm) containing standard cornmeal-yeast-agar-molasses (CYAM) medium. Bottles were kept in a 12 : 12 light-dark cycle and at 18°C –20°C. For stock maintenance, flies were allowed to freely mate and laid progeny in fresh media, the adults were discarded after 18 days, and a new generation of newly emerged adults were transferred to fresh media.

2.2. Setting Up Crosses for Experimental Testing. Crosses were performed between *D. virilis* females and *D. novamexicana* males as well as *D. lummei* and *D. a. texana* males. Reciprocal crosses among species were also tested with the exception, due to strong premating isolation [12], of crosses using *D. virilis* males. Conspecific crosses were used as controls. Bottles from each species stock were emptied and inspected daily for new adult emergences. Newly emerged flies were lightly anesthetized using CO_2 gas. Virgin females and males were separated by sex and placed in cylindrical vials (28.5 × 95 mm) containing CYAM medium. Males and females were aged for 10 days before use to ensure sexual maturity [18] and crosses were set up with one male and one female. Counts were obtained of eggs laid by females and the proportion of eggs hatched. For crosses between outbred *D. virilis* and *D. novamexicana*, we also estimated the proportion of fertilized eggs, and sperm fate was tracked within females.

2.3. Egg Hatchability. Each conspecific and heterospecific pair was placed in an egg-laying chamber made using a polystyrene Petri dish (60 × 15 mm) containing fresh CYAM medium attached to a 100 mL graduated polypropylene beaker. Every 24 hours, flies were slightly anesthetized using CO_2, the Petri dish was removed, and a new dish with fresh CYAM medium was attached to the chamber. The replacement of dishes was continued for a total of five consecutive days. Eggs laid in each dish were daily counted, and 48 hours later hatched eggs were scored. Unhatched eggs can be recognized as a white compact shape, while hatched eggs appear as an empty outer chorion membrane due to larvae emergence.

2.4. Fertilization of the Eggs. Single male and female heterospecific and conspecific crosses, using *D. virilis* and *D. novamexicana* outbred strains, were set up in vials. Courtship behavior was observed until copulation occurred and for no longer than six hours. After copulation, females were transferred individually into egg-laying chambers. Females were transferred daily for five days to fresh egg-laying chambers, and the number of eggs laid was counted. 48 hours after eggs were laid, hatched eggs were scored. Unhatched eggs were collected from the media with a wooden handle dissecting pin and placed on a clean microscope slide. A drop of 1X PBS was added to the eggs. Eggs were manually dechorionated by removing the dorsal appendage and gently pressing at the posterior pole using minute pins (0.1 mm diameter). The inner vitelline membrane was removed by immersing the dechorionated eggs in a small tube containing a 1 : 1 solution of heptane and 90% methanol. The eggs

dropped to the layer between heptane and methanol and slowly descended to the bottom of the tube when their waxy layer was lost [19]. Eggs without vitelline membrane are almost transparent and easily damaged. Intact eggs were collected by pouring the heptane-methanol solution on a small piece of dark cloth. Within one minute, the solution evaporated and the eggs were visible on the cloth surface. A couple of drops of 1X PBS were added on the eggs using a glass Pasteur pipette. The eggs were then gently picked using an insect pin, placed on a clean microscope slide containing 300 nM DAPI (Molecular Probes, D3571) and incubated in a dark room for 30 minutes. After incubation, the eggs were examined under a Nikon Eclipse (E400) fluorescence microscope for evidence of nuclear division (fertilization).

2.5. Tracking of Sperm. A single male and female were set up as described in the previous section. At different intervals after mating (0, 24 and 48 hours), inseminated females were transferred into vials. The vials were flash frozen by submersion in liquid nitrogen. Females were then transferred to small tubes and stored in a freezer at −70°C. Flies were placed in a drop of 1X PBS and each female reproductive tract was separated from the rest of the body. Sperm storage organs, uterus, pair of spermatheca, and seminal receptacle, were furthered separated, and each was placed on a fresh drop of 1X PBS on a single clean microscope slide. These slides were dried in an oven set to 60°C for 5 minutes, fixed in 3 : 1 methanol-glacial acetic acid for 5 minutes, and washed three times in 1X PBS [5]. Dissected tissues were stained using 300 nM of DAPI by incubating in a dark room for 30 minutes. Slides were examined under a Nikon Eclipse (E400) fluorescence microscope, and the presence of sperm was determined.

2.6. Data Analysis. The number of eggs, laid, the proportion of hatched eggs and the proportion of eggs fertilized were compared among different crosses using a one-way analysis of variance (ANOVA). When significant differences were found among crosses, a posteriori Tukey test was performed to test which cross averages were significantly different from one another. All statistical tests were conducted in SPSS (version 12.0). Comparisons of number of females with sperm in storage were done using a 2×3 Chi-Square test.

3. Results

The average number of eggs laid by females mated to heterospecific males was not significantly different than the number laid by *D. virilis* females mated to conspecific males (Figure 1). Only crosses involving two *D. virilis* strains, Russia ($F_{2,101} = 28.60$; $P < .001$) and Japan ($F_{2,91} = 22.23$; $P < .001$), showed a significantly lower number of eggs laid than *D. novamexicana* females mated to conspecific males (Figure 1). In contrast, the proportion of eggs hatched from the heterospecific cross was consistently and significantly lower than the proportion of hatches from both of the conspecific crosses (Figure 2). Depending on the *D. virilis* strain used, the proportion of unhatched eggs varied between 0.84 and 0.96 (Figure 2). When using flies from the outbred populations, we found that the number of eggs laid was not different among crosses ($F_{2,77} = 2.83$; $P = .065$) (Figure 1) but that 0.94 of the eggs laid by *D. virilis* females mated with *D. novamexicana* males did not hatch (Figure 2).

Unhatched eggs could result from fertilized eggs that fail to develop or unfertilized eggs. We DAPI stained unhatched eggs to test for nuclear division. We counted the number of eggs hatched as fertilized and tested unhatched eggs laid from both conspecific and heterospecific crosses using the outbred populations. We found significant differences in the proportion of fertilized eggs ($F_{2,68} = 173.42$; $P < .001$) due to a significantly lower proportion of only five percent of eggs fertilized by *D. novamexicana* males that mated with *D. virilis* females (Tukey post hoc test: $P < .001$) (Figure 3). These results indicate that the vast majority of unhatched eggs in heterospecific crosses between *D. virilis* females and *D. novamexicana* males are the result of some form of PPI.

PPI can result from problems in sperm transfer during copulation, problems with sperm storage, or the inability of sperm to fertilize heterospecific eggs. We used the *D. virilis* and *D. novamexicana* outbred populations to test whether sperm transfer and/or storage was affected in the heterospecific cross. We did this by observing copulations and dissecting females immediately after mating (0 hours) and at two intervals of 24 and 48 hours after mating. Immediately after mating, large numbers of sperm transferred to the females were found to be located in the uterus in both conspecific and heterospecific crosses. Because of the large numbers and the fact that the sperm head of *D. novamexicana* and *D. virilis* males is a needle-like structure that sometimes only faintly stains with DAPI, we tested differences between crosses by scoring numbers of females with or without sperm in storage. There were no significant differences in numbers of females with sperm in storage immediately after mating (0 hours), with all females having large amounts of sperm in their uterus. At 24 hours after mating, we observed a slight decline in numbers of females with sperm in storage for the heterospecific crosses. We only found significant differences between intra- and interspecific crosses for both the spermatheca and the seminal receptacle at 48 hours after mating, with a higher number of females mated to heterospecific males having no sperm in storage (spermatheca: $\chi^2 = 11.31$, $P = .004$; seminal receptacle: $\chi^2 = 25.23$, $P < .001$) (Figure 4). The most striking difference was that only 1 out of 21 *D. virilis* females mated with *D. novamexicana* males had few sperm cells in the seminal receptacle at 48 hours after mating. Overall, the heterospecific cross shows a different pattern of sperm storage within the female reproductive tract than the conspecific crosses (Figure 4).

Using males from other species of the virilis subgroup, we show that the number of eggs produced by *D. virilis* females is significantly higher than the number of eggs laid by *D. a. texana* ($F_{2,48} = 10.77$; $P < .001$) and *D. lummei* ($F_{2,49} = 3.74$; $P < .031$) (Figure 5(a)). Interestingly, the number of eggs laid by *D. virilis* females mated to heterospecific males is not significantly different than the number of eggs laid by *D. virilis* females mated with conspecific males

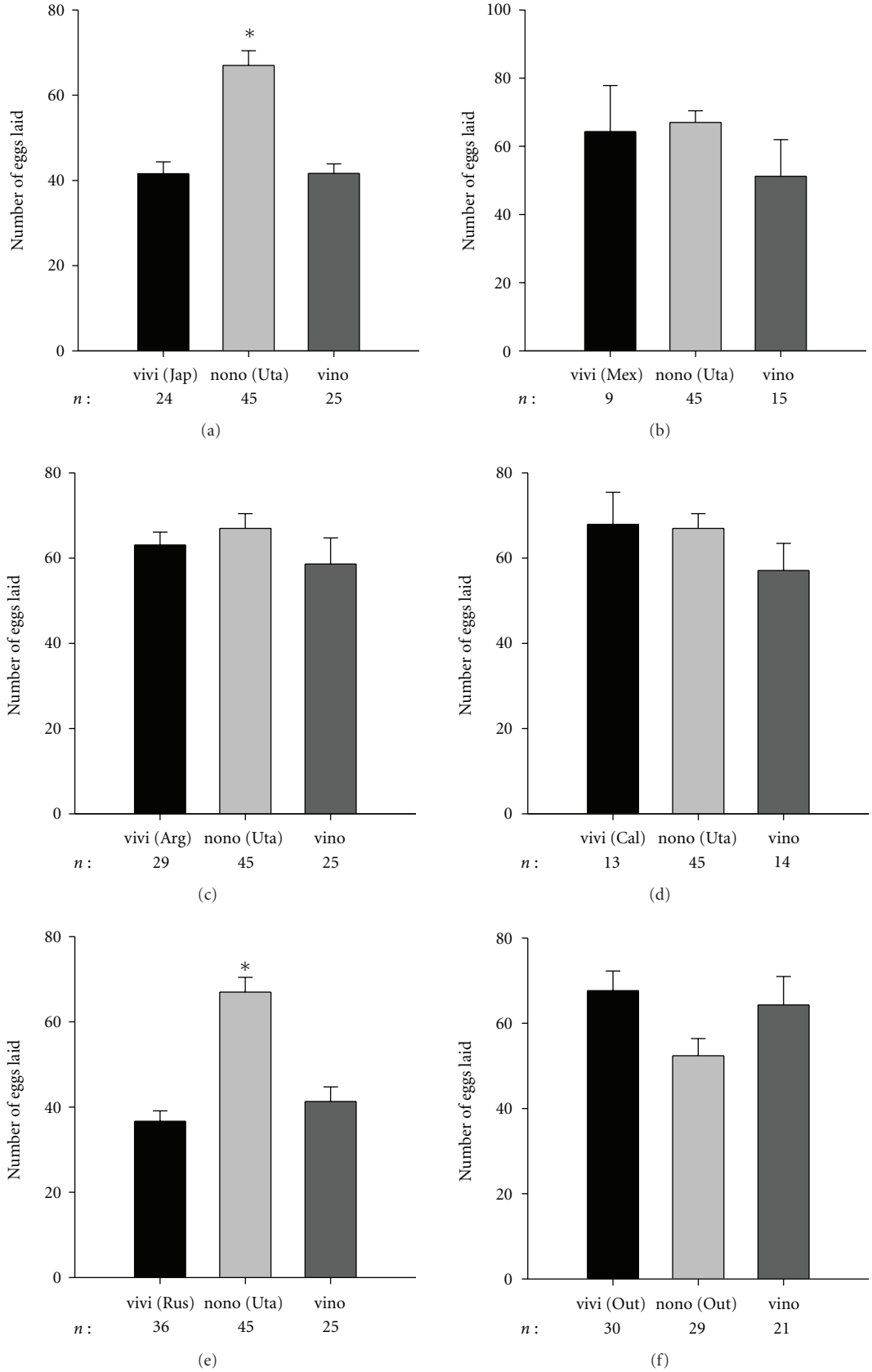


Figure 1: Average number of laid eggs and associated standard error. *n*: the number of females tested. A conspecific cross is denoted as vivi or nono, for *D. virilis* and *D. novamexicana*, respectively, with the origin of the strain used shown in between parenthesis. Results for the heterospecific cross are labeled as vino. Asterisks above columns are used to label statistically different averages.

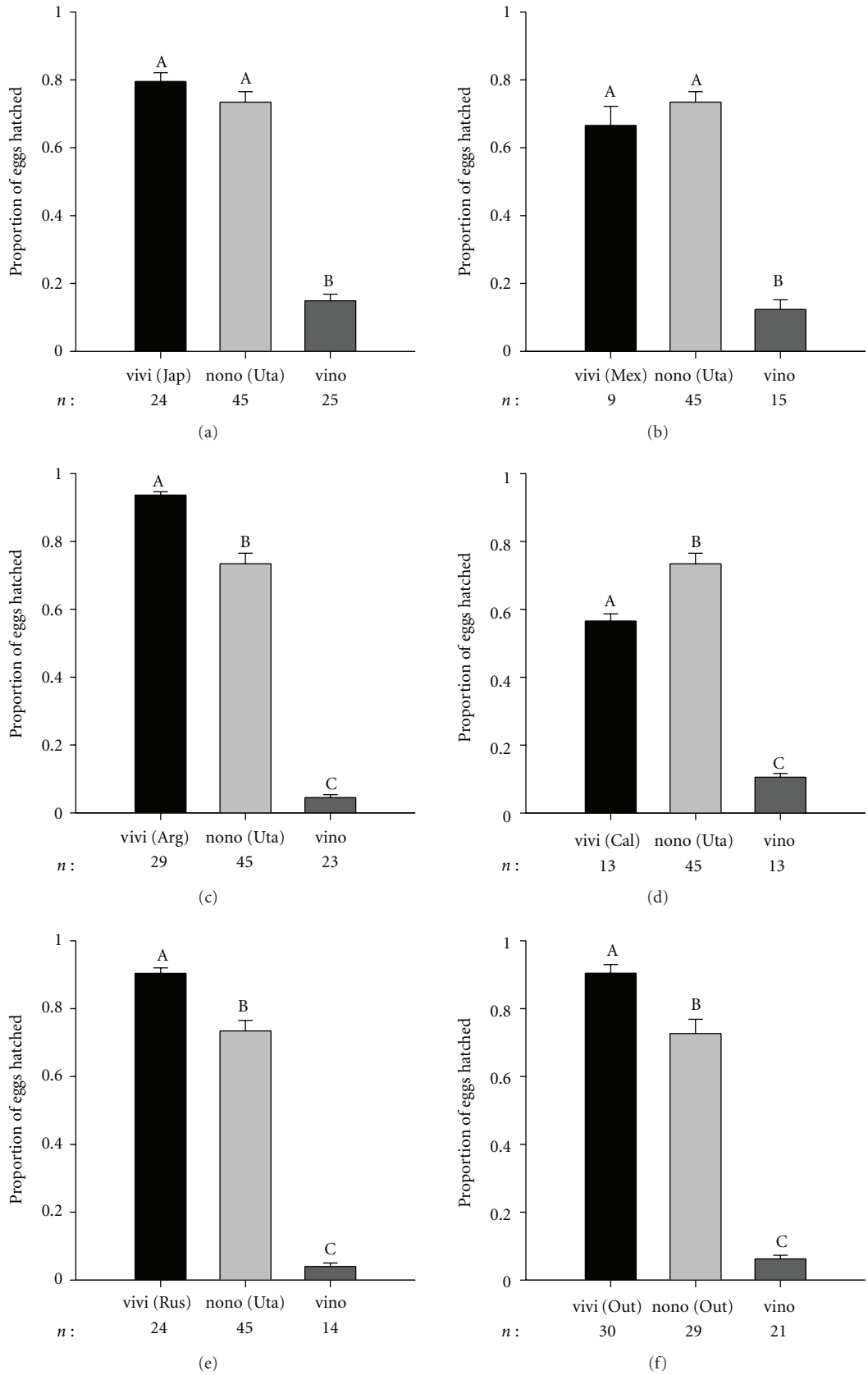


FIGURE 2: Average proportion of hatched eggs and associated standard error. Labels are as in Figure 1 except that shared letters above columns indicate that averages are not statistically different.

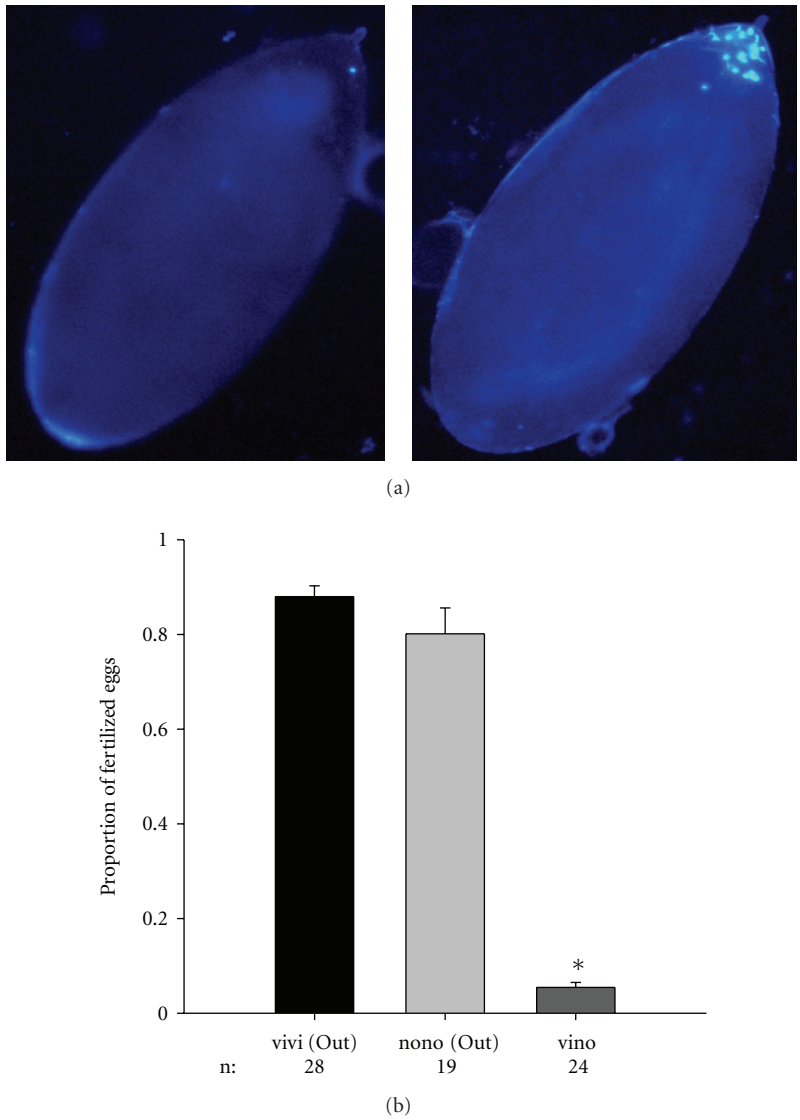


Figure 3: DAPI staining of eggs showing no nuclear division in an unhatched egg from a heterospecific cross (left panel) and a cluster of dividing nuclei at 2 hours after egg laying by a female mated to a conspecific male (a). Average proportion of fertilized eggs and associated standard error (b). Labels are as in Figure 1.

(Tukey post hoc test: $P = 1.00$; $P = .567$; $P = .709$; $P = .944$ for *D. virilis* Japan × *D. novamexicana*, *D. virilis* Russia × *D. novamexicana*, *D. virilis* × *D. a. texana*, and *D. virilis* × *D. lummei* resp.) (Figures 1 and 5(a)). However, egg hatchability is always lower in heterospecific crosses than both conspecific crosses, with the exception of *D. virilis* females × *D. lummei* males (Figures 2 and 5(b)).

4. Discussion

Studies of reproductive isolation between species of the virilis subgroup have shown extensive variation in the strength of the isolating barrier depending on the geographic origin of the strains tested (reviewed in [11]). Therefore, it is important to show that PPI between *D. virilis* females and *D. novamexicana* males is independent of the geographic origin of the strains used and thus fixed rather than polymorphic. *Drosophila virilis* females of different geographic origin crossed with *D. novamexicana* males hatched between 4% and 16% of eggs, in line with a prior single strain assay that found only 15% of eggs hatched within two days after mating [17, see Table 103]. The uniformity of PPI among geographically distinct *D. virilis* strains suggests that PPI with *D. novamexicana* is a byproduct of divergent evolution rather than reinforcement though ecological differences between the two species may have reduced any possible interactions.

Contrary to what has been found among other species of Drosophila (e.g., [10]), *D. virilis* females mated to heterospecific males produce numbers of eggs that are not different from the numbers produced by *D. virilis* mated to conspecific males. Numerous studies have shown the influence of accessory gland proteins (ACPs) on female oviposition in

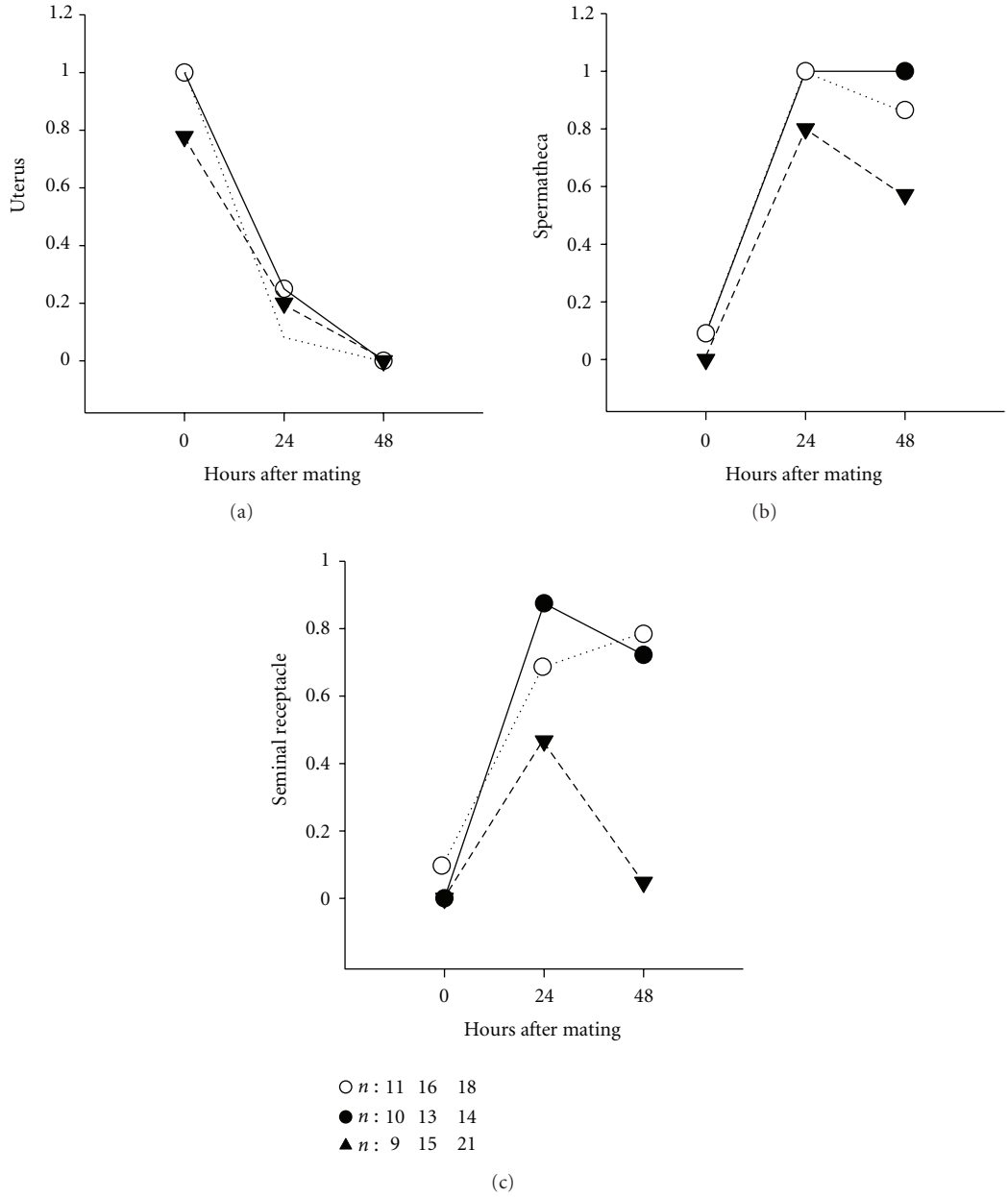


Figure 4: Proportion of females with sperm found in different storage organs immediately after mating (0 hours), at 24 hours and 48 hours after mating. White circles are *D. novamexicana* × *D. novamexicana*, black circles are *D. virilis* × *D. virilis*, and triangles are *D. virilis* × *D. novamexicana*. *n*: the number of females tested.

inseminated *D. melanogaster* females (reviewed in [20]). For example, both ovulin (ACP26Aa) and the sex peptide (ACP70A) have been implicated in stimulating egg production and ovulation after mating [21–24]. Males can influence female egg laying, but the egg-laying effect triggered by the ejaculate is also mediated by females' molecular counterparts (reviewed in [20]). Therefore, it is equally likely that either males of the virilis subgroup has not diverged enough in ACPs, and other ejaculate proteins content or that *D. virilis* females have retained the ability to recognize a wide variety of male-derived egg-laying stimulating signals from the ejaculate. The first possibility seems unlikely given the fast rate of evolution of ACPs in Drosophila. ACPs are known to become neofunctionalized, lose their function or simply be lost even in comparisons among closely related species of Drosophila [25–27]. However, despite the fact that current genome information (http://flybase.org/) shows no orthologs of known *D. melanogaster* ACPs in *D. virilis*, we cannot fully rule out ACPs, as we have no knowledge of divergence among these proteins within the virilis subgroup. Our results demonstrate that in every heterospecific cross involving *D. virilis* females, egg laying is determined by the female of the species. This supports the hypothesis that *D. virilis* females have retained the molecular ability to recognize

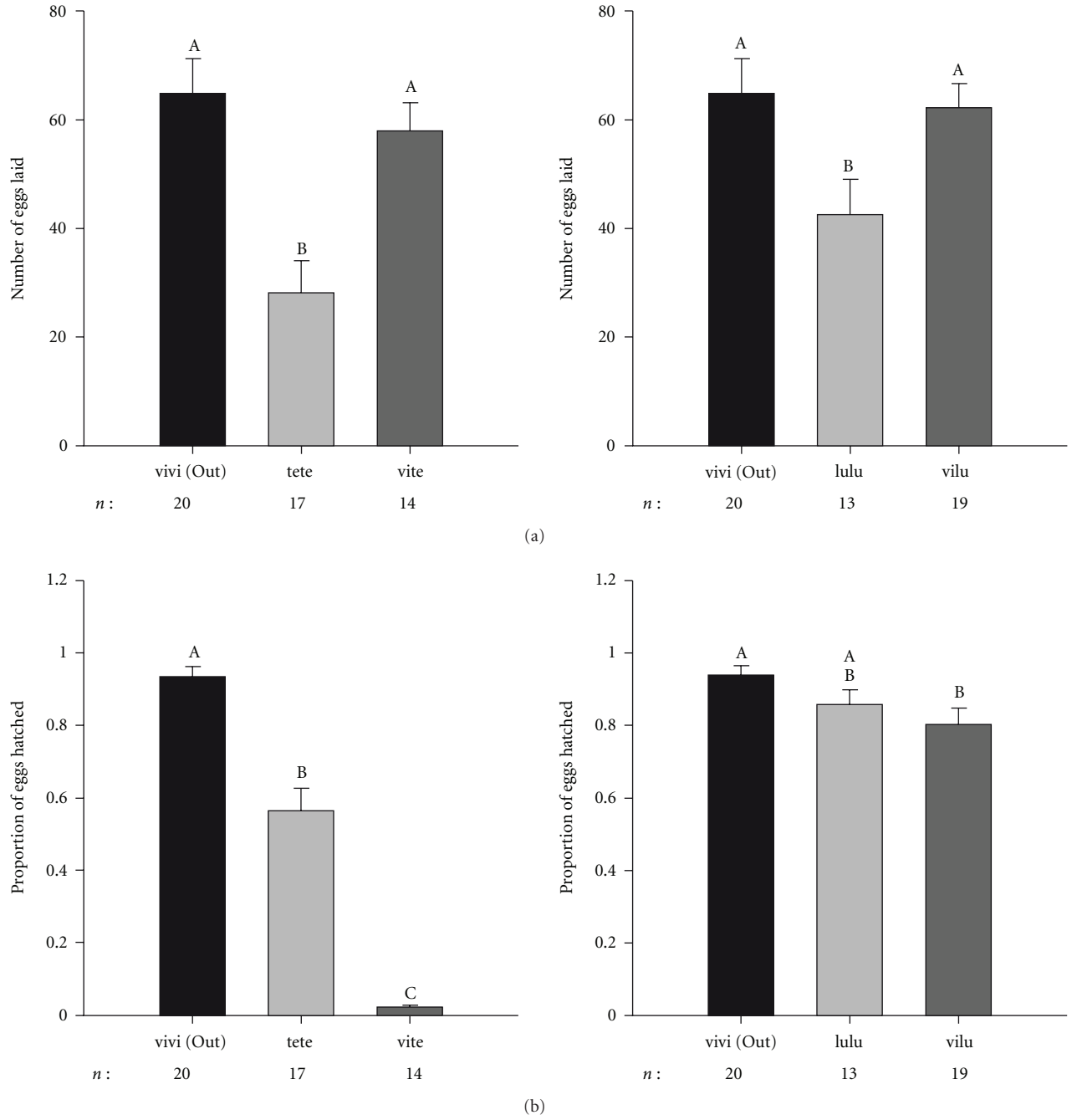


Figure 5: Average number of laid eggs and associated standard error for crosses involving *D. virilis* females and either *D. a. texana* (te) or *D. lummei* (lu) males (a). Average proportion of hatched eggs and associated standard error for crosses involving *D. virilis* females and either *D. a. texana* (te) or *D. lummei* (lu) males (b). Other labels are as in Figures 1 and 2.

a wide variety of egg-stimulating signals within the ejaculate. However, not all females within the species subgroup have retained this form of cryptic control. Preliminary data suggests that in crosses involving *D. lummei*, females respond by laying significantly higher number of eggs than in conspecific crosses (supplementary Figure 1S and supplementary Tables 1S and 2S in supplementary material available online at doi: 10.4061/2011/485460). This is a pattern more likely expected when males can overtake female resources to their advantage. Overall, it appears that complex male-female interactions likely drive the evolution of egg laying phenotypes among species within the virilis subgroup.

Heterospecific crosses, with the exception of the cross between *D. virilis* females and *D. lummei* males, fail to fertilize enough eggs. Detailed studies done in *D. melanogaster* have characterized male ejaculate components that are known to influence the ability of sperm to properly store within female storage organs [28–33]. We also know that

female secretions from the spemathecae and parovaria are required by sperm to fertilize the eggs [34–36], but it is unclear whether they affect sperm viability, sperm retention, or release from female storage. Our data from crosses between *D. virilis* females and *D. novamexicana* males shows that sperm is normally transferred but rapidly lost and such depletion of sperm could contribute to low egg fertilization. We do not know whether females actively dump or simply lose sperm during egg laying. Given the time frame at which sperm are being lost, and based solely on data from studies using *D. melanogaster* [37, 38], it seems unlikely that sperm is ejected by females but rather lost during the egg-laying process.

The production of similar numbers of eggs by females mated to heterospecific males when compared to conspecific crosses, and the fact that most eggs fail to hatch is likely costly as it represents an energetic burden for females to lay a large number of unfertilized eggs. However, rapid loss of sperm from the female storage organs after mating might explain why there has not been strong selection against the high egg-laying phenotype. A female mating to a heterospecific male could actively dump or simply lose sperm from storage and become available and receptive to another mate. There is, in fact, evidence that *D. melanogaster* females with less sperm in storage are more likely to remate [39] and that female remating, in at least another Drosophila species, increases when a male is sperm depleted [40].

The PPI barrier among species within the subgroup is strong (Figures 2, 5 and 1S). Depending on the heterospecific cross performed, between approximately 75% and 98% of eggs failed to hatch. The cross between *D. virilis* females and *D. lummei* males was the only one for which we did not detect any form of PPI (see supplementary Table 2S). *Drosophila virilis* and *D. lummei* are the basal palearctic species to the subgroup, with chromosomal changes arisen during dispersal of the most recently derived species of *D. a. americana*, *D. a. texana*, and *D. novamexicana* (the novamexicana-americana clade) into North America [11]. Based on molecular data, *D. virilis* and *D. lummei* consistently appear as ancestral to species of the novamexicana-americana clade, with *D. virilis* being the oldest species followed by *D. lummei* [41, 42]. Moreover, estimated times of divergence show *D. virilis* arose about 6.5 Mya, closely followed by *D. lummei* (approximately 5.5 Mya) and a much later diversification of the novamexicana-americana clade around 1 Mya [41]. The available phylogenetic information therefore supports *D. virilis* and *D. lummei* as ancestral species and more closely related to each other than to the other three species in the subgroup. Our results support that PPI due to low egg hatchability most likely evolved during the diversification of the novamexicana-americana clade from the two basal species of the virilis subgroup. Interestingly, PPI also appears to be a strong form of isolation among species within the most recently derived clade (novamexicana-americana).

Finally, we know that *D. virilis* males show strong premating isolation, but *D. virilis* females readily mate with heterospecifics [11, 12]. Coyne and Orr [43, 44] combined information on phylogenetic divergence and strength of premating and postzygotic isolation in the genus Drosophila to conclude that premating isolating barriers evolve earlier than other forms of isolation between diverging populations. It is then puzzling why *D. virilis* females do not show strong premating isolation with other species of the virilis subgroup. One possibility is that ordering isolating barriers by time of divergence (as in [43, 44]) is not fully informative of their actual contribution to isolation as one cannot assume total independence among isolation mechanisms. Premating isolation might, therefore, not necessarily be the first barrier to hybridization. The other possibility might be that premating behavioural isolation between males of the ancestral species (*D. virilis*) and derived female species resulted from the evolution of polymorphism in receptors of derived male species to detect both short ancestral (*D. virilis*) and long derived species female cuticular hydrocarbons [45]. If so, the monomorphic male receptors in *D. virilis* males might not be able to detect heterospecific females as suitable mates [12].

In summary, our results demonstrate that *D. virilis* females exert control over egg-laying rates after mating and that PPI due to problems with egg hatching evolved as a strong reproductive isolating barrier among some of the species of the virilis subgroup. Our data from crosses with *D. novamexicana* males and data from earlier studies using *D. a. americana* and *D. a. texana* males indicate that PPI is due to different problems faced by the heterospecific sperm in female storage. It is yet unclear to what extent interspecies divergence of proteins of the male ejaculate and/or female reproductive tract secretions might contribute to heterospecific sperm problems in the *D. virilis* female reproductive tract.

Acknowledgments

The authors would like to thank Jens Franck and two anonymous referees for their comments and suggestions. This work was funded by an NSERC Discovery Grant to A. Civetta. N. Sagga was fully funded by a graduate scholarship from the Government of Saudi Arabia.

References

[1] J. A. Coyne and H. A. Orr, *Speciation*, Sinauer Associates, Sunderland, Mass, USA, 2004.

[2] P. G. Gregory and D. J. Howard, "A post-insemination barrier to fertilization isolates two closely related ground crickets," *Evolution*, vol. 48, no. 3, pp. 705–710, 1994.

[3] C. S. C. Price, "Conspecific sperm precedence in *Drosophila*," *Nature*, vol. 388, no. 6643, pp. 663–666, 1997.

[4] D. J. Howard, P. G. Gregory, J. Chu, and M. L. Cain, "Conspecific sperm precedence is an effective barrier to hybridization between closely related species," *Evolution*, vol. 52, no. 2, pp. 511–516, 1998.

[5] C. S.C. Price, C. H. Kim, C. J. Gronlund, and J. A. Coyne, "Cryptic reproductive isolation in the *Drosophila simulans* species complex," *Evolution*, vol. 55, no. 1, pp. 81–92, 2001.

[6] L. W. Simmons, *Sperm Competition and Its Evolutionary Consequences in the Insects*, Princeton University Press, Princeton, NJ, USA, 2001.

[7] N. Sellier, J. M. Brun, M. M. Richard, F. Batellier, V. Dupuy, and J. P. Brillard, "Comparison of fertility and embryo mortality following artificial insemination of common duck females (*Anas Platyrhynchos*) with semen from common or Muscovy (*Cairina Moschata*) drakes," *Theriogenology*, vol. 64, no. 2, pp. 429–439, 2005.

[8] T. R. Birkhead and J. P. Brillard, "Reproductive isolation in birds: postcopulatory prezygotic barriers," *Trends in Ecology and Evolution*, vol. 22, no. 5, pp. 266–272, 2007.

[9] E. S. Kelleher and T. A. Markow, "Reproductive tract interactions contribute to isolation in *Drosophila*," *Fly*, vol. 1, no. 1, pp. 33–37, 2007.

[10] D. R. Matute and J. A. Coyne, "Intrinsic reproductive isolation between two sister species of Drosophila," *Evolution*, vol. 64, no. 4, pp. 903–920, 2010.

[11] L. H. Throckmorton, "The virilis species group," in *The Genetics and Biology of Drosophila*, M. Ashburner, H. L. Carson, and J. N. Thompson Jr., Eds., vol. 3b, pp. 227–296, Academic Press, London, UK, 1982.

[12] D. Nickel and A. Civetta, "An X chromosome effect responsible for asymmetric reproductive isolation between male *Drosophila virilis* and heterospecific females," *Genome*, vol. 52, no. 1, pp. 49–56, 2009.

[13] H. A. Orr and J. A. Coyne, "The genetics of postzygotic isolation in the *Drosophila virilis* group," *Genetics*, vol. 121, no. 3, pp. 527–537, 1989.

[14] J. Lumme and E. Heikkinen, "Viability of first and second generation hybrids of *Drosophila virilis* and *Drosophila lummei*," *Heredity*, vol. 65, pp. 435–447, 1990.

[15] K. Lamnissou, M. Loukas, and E. Zouros, "Incompatibilities between Y chromosome and autosomes are responsible for male hybrid sterility in crosses between *Drosophila virilis* and *Drosophila texana*," *Heredity*, vol. 76, p. 6, 1996.

[16] A. L. Sweigart, "The genetics of postmating, prezygotic reproductive isolation between *Drosophila virilis* and *D. americana*," *Genetics*, vol. 184, no. 2, pp. 401–410, 2010.

[17] J. T. Patterson and W. S. Stone, *Evolution in the Genus Drosophila*, Macmillan and Co., New York, NY, USA, 1952.

[18] T. A. Markow and P. M. O'Grady, "Drosophila biology in the genomic age," *Genetics*, vol. 177, no. 3, pp. 1269–1276, 2007.

[19] R. M. Warn and A. Warn, "Microtubule arrays present during the syncytial and cellular blastoderm stages of the early *Drosophilia embryo*," *Experimental Cell Research*, vol. 163, no. 1, pp. 201–210, 1986.

[20] M. F. Wolfner, "Battle and ballet: molecular interactions between the sexes in drosophila," *Journal of Heredity*, vol. 100, no. 4, pp. 399–410, 2009.

[21] M. Soller, M. Bownes, and E. Kubli, "Mating and sex peptide stimulate the accumulation of yolk in oocytes of Droosophila melanogaster," *European Journal of Biochemistry*, vol. 243, no. 3, pp. 732–738, 1997.

[22] M. Soller, M. Bownes, and E. Kubli, "Control of oocyte maturation in sexually mature Drosophila females," *Developmental Biology*, vol. 208, no. 2, pp. 337–351, 1999.

[23] S. A. Monsma and M. F. Wolfner, "Structure and expression of a Drosophila male accessory gland gene whose product resembles a peptide pheromone precursor," *Genes & Development*, vol. 2, no. 9, pp. 1063–1073, 1988.

[24] Y. Heifetz, O. Lung, E. A. Frongillo, and M. F. Wolfner, "The Drosophila seminal fluid protein Acp26Aa stimulates release of oocytes by the ovary," *Current Biology*, vol. 10, no. 2, pp. 99–102, 2000.

[25] D. J. Begun and H. A. Lindfors, "Rapid evolution of genomic Acp complement in the *melanogaster* subgroup of *Drosophila*," *Molecular Biology and Evolution*, vol. 22, no. 10, pp. 2010–2021, 2005.

[26] D. J. Begun, H. A. Lindfors, M. E. Thompson, and A. K. Holloway, "Recently evolved genes identified from *Drosophila yakuba* and *D. erecta* accessory gland expressed sequence tags," *Genetics*, vol. 172, no. 3, pp. 1675–1681, 2006.

[27] W. Haerty, S. Jagadeeshan, R. J. Kulathinal et al., "Evolution in the fast lane: rapidly evolving sex-related genes in Drosophila," *Genetics*, vol. 177, no. 3, pp. 1321–1335, 2007.

[28] A. Wong, S. N. Albright, J. D. Giebel et al., "A role for Acp29AB, a predicted seminal fluid lectin, in female sperm storage in *Drosophila melanogaster*," *Genetics*, vol. 180, no. 2, pp. 921–931, 2008.

[29] D. M. Neubaum and M. F. Wolfner, "Mated Drosophila melanogaster females require a seminal fluid protein, Acp36DE, to store sperm efficiently," *Genetics*, vol. 153, no. 2, pp. 845–857, 1999.

[30] T. Chapman, D. M. Neubaum, M. F. Wolfner, and L. Partridge, "The role of male accessory gland protein Acp36DE in sperm competition in *Drosophila melanogaster*," *Proceedings of the Royal Society Biological Sciences*, vol. 267, no. 1448, pp. 1097–1105, 2000.

[31] M. J. Bertram, D. M. Neubaum, and M. F. Wolfner, "Localization of the Drosophila male accessory gland protein Acp36DE in the mated female suggests a role in sperm storage," *Insect Biochemistry and Molecular Biology*, vol. 26, no. 8-9, pp. 971–980, 1996.

[32] F. W. Avila and M. F. Wolfner, "Acp36DE is required for uterine conformational changes in mated Drosophila females," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 37, pp. 15796–15800, 2009.

[33] J. L. Mueller, J. R. Linklater, K. R. Ram, T. Chapman, and M. F. Wolfner, "Targeted gene deletion and phenotypic analysis of the Drosophila melanogaster seminal fluid protease inhibitor Acp62F," *Genetics*, vol. 178, no. 3, pp. 1605–1614, 2008.

[34] R. C. Anderson, "A study of the factors affecting fertility of lozenge females of *Drosophila melanogaster*," *Genetics*, vol. 30, pp. 280–296, 1945.

[35] A. K. Allen and A. C. Spradling, "The Sf1-related nuclear hormone receptor Hr39 regulates Drosophila female reproductive tract development and function," *Development*, vol. 135, no. 2, pp. 311–321, 2008.

[36] A. Prokupek, F. Hoffmann, S. I. Eyun, E. Moriyama, M. Zhou, and L. Harshman, "An evolutionary expressed sequence tag analysis of drosophila spermatheca genes," *Evolution*, vol. 62, no. 11, pp. 2936–2947, 2008.

[37] R. R. Snook and D. J. Hosken, "Sperm death and dumping in Drosophila," *Nature*, vol. 428, no. 6986, pp. 939–941, 2004.

[38] M. K. Manier, J. M. Belote, K. S. Berben, D. Novikov, W. T. Stuart, and S. Pitnick, "Resolving mechanisms of competitive fertilization success in *Drosophila melanogaster*," *Science*, vol. 328, no. 5976, pp. 354–357, 2010.

[39] M. E. A. Newport and M. H. Gromko, "The effect of experimental design on female receptivity to remating and its impact on reproductive success in *Drosophila melanogaster*," *Evolution*, vol. 38, pp. 1261–1272, 1984.

[40] N. Luck and D. Joly, "Sexual selection and mating advantages in the giant sperm species, *Drosophila bifurca*," *Journal of Insect Science*, vol. 5, p. 10, 2005.

[41] G. S. Spicer and C. D. Bell, "Molecular phylogeny of the *Drosophila virilis* species group (Diptera: Drosophilidae) inferred from mitochondrial 12S and 16S ribosomal RNA genes," *Annals of the Entomological Society of America*, vol. 95, no. 2, pp. 156–161, 2002.

[42] B. C. Caletka and B. F. McAllister, "A genealogical view of chromosomal evolution and species delimitation in the *Drosophila virilis* species subgroup," *Molecular Phylogenetics and Evolution*, vol. 33, no. 3, pp. 664–670, 2004.

[43] J. A. Coyne and H. A. Orr, "Patterns of speciation," *Evolution*, vol. 43, pp. 362–381, 1989.

[44] J. A. Coyne and H. A. Orr, "Patterns of speciation in *Drosophila*," *Evolution*, vol. 52, pp. 295–303, 1997.

[45] R. J. Bartelt, M. T. Armold, A. M. Schaner, and L. L. Jackson, "Comparative analysis of cuticular hydrocarbons in the *Drosophila virilis* species group," *Comparative Biochemistry and Physiology—Part B: Biochemistry and*, vol. 83, no. 4, pp. 731–742, 1986.

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 214280, 10 pages
doi:10.4061/2011/214280

Research Article

Associations between Variation in X Chromosome Male Reproductive Genes and Sperm Competitive Ability in *Drosophila melanogaster*

Leah Greenspan[1,2] and Andrew G. Clark[1]

[1] *Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853-2801, USA*
[2] *Department of Developmental Biology, Memorial Sloan-Kettering Cancer Center, New York, NY 10065-6094, USA*

Correspondence should be addressed to Andrew G. Clark, ac347@cornell.edu

Received 11 January 2011; Accepted 7 March 2011

Academic Editor: Jeremy Marshall

Variation in reproductive success has long been thought to be mediated in part by genes encoding seminal proteins. Here we explore the effect on male reproductive phenotypes of X-linked polymorphisms, a chromosome that is depauperate in genes encoding seminal proteins. Using 57 X chromosome substitution lines, sperm competition was tested both when the males from the wild-extracted line were the first to mate ("defense" crosses), followed by a tester male, and when extracted-line males were the second to mate, after a tester male ("offfense" crosses). We scored the proportion of progeny sired by each male, the fecundity, the remating rate and refractoriness to remating, and tested the significance of variation among lines. Eleven candidate genes were chosen based on previous studies, and portions of these genes were sequenced in all 57 lines. A total of 131 polymorphisms were tested for associations with the reproductive phenotypes using linear models. Nine polymorphisms in 4 genes were found to show significant associations (at a 5% FDR). Overall, it appears that the X chromosomes harbor abundant variation in sperm competition, especially considering the paucity of seminal protein genes. This suggests that much of the male reproductive variation lies outside of genes that encode seminal proteins.

1. Introduction

In nature, *Drosophila melanogaster* females often mate with multiple partners [1–3]. Due to the female fly's ability to store sperm, having multiple partners provides an evolutionary advantage for female flies allowing them to choose the strongest and most fit sperm to fertilize their eggs. Females' tendency to remate has provided an opportunity for selection to operate on differential sperm success, and this opportunity is thought to have resulted in robust sperm competition among male fruit flies. In order for male flies to sire as many progeny as possible, their ejaculate must either provide a means for their sperm to outcompete other sperm in order to fertilize as many eggs as possible, or it may contain substances that discourage the female fly from remating [4]. Although recent studies have shown that males vary in their ability to gain fertilizations under competitive conditions, the precise mechanism is unknown. Factors that influence the outcome of sperm competition and postcopulatory sexual selection include differences in sperm delivery and storage, seminal fluid composition, female egg laying rate, and female remating latency [4, 5].

Sexual antagonism occurs when there is an evolutionary advantage of a trait in one sex that reduces the fitness of the opposite sex. Since *D. melanogaster* females have the capability of storing sperm for use in fertilizations over a period of many days, they have the opportunity to produce progeny from multiple fathers. In addition to the fact that there may be differential use of the stored sperm from different males, there is an opportunity for postcopulatory sexual selection. Seminal proteins have developed in male flies that manipulate the female's reproductive behavior, including her tendency to remate and the rate at which she lays eggs [4]. These proteins increase the fitness of the male

fly by increasing the probability of paternity; however, some of these proteins have been found to be toxic to females, resulting in a conflict, whereby what is good for the male appears to have fitness-lowering consequences to the female [4]. To compensate for this decrease in fitness, female fruit flies have developed a defense mechanism against the toxic seminal fluid. This sets up the opportunity for an arms race between the male and female. One consequence of this is that accessory gland protein-encoding genes (Acps) are among the fastest evolving genes in Drosophila [6].

Different strains of *D. melanogaster* vary in the degree of toxicity of the male seminal fluid and also in the degree of effectiveness of the defense system in females. Many experiments have been performed exploring the specific genes that are involved in sperm competition. In Fiumera et al. [7], a survey of 95 second chromosome extraction lines identified wide variability in all the reproductive phenotypes measured, including the extent to which the first-mating and second-mating males successfully sire progeny in double-mating contests with a standard marker male. SNPs in Acp-encoding genes and other candidate genes were genotyped across the lines, and nine significant associations were identified between polymorphisms in the male reproductive genes and sperm competition. Pleiotropy was seen between two genes: a polymorphism in the *Acp33A* gene correlates with both P1′ and P2′ (arcsine square root transformation of P1 and P2) and a polymorphism in the *CG17331* gene correlates with both elevated P2′ and reduced refractoriness. A similar study of variation on the third chromosome of *D. melanogaster* [8] yielded four genes related to sperm competition that presented evidence of pleiotropic effects: *CG6168*, *CG14560*, *Acp62F*, and *Esterase-6*. These studies revealed extraordinary magnitudes of difference in fitness among genotypes, and even the single-gene associations reveal effect sizes on fitness of greater than 10%. This poses a challenge to understand how such a high level of intergenotype fitness variation can be maintained in the population. Fiumera et al. [8] identified some cases where the pleiotropic effects of the single nucleotide variants were antagonistic, that is, the allele that was favored by sperm competition might be disfavored in female-induced fecundity. While this is a possible mechanism for maintaining polymorphism, so far, the evidence does not support the idea that antagonistic pleiotropy is sufficiently prevalent to explain standing levels of polymorphism.

To complete a genome-wide survey of polymorphisms that mediate differences in sperm competitive ability, we sought to determine whether the X chromosome contains any genes that also can affect sperm competition. The X chromosome has additional features distinct from the autosomes that impact the tendency for variation in sperm competitive fitness to impact polymorphism. The different doses of the X chromosome in XY males and XX females lead to a dosage-compensation mechanism in males that is based on elevated transcription of X-chromosome genes [9]. However, in Drosophila, dosage compensation does not occur in the germ cells, leading to a lower level of expression of certain X chromosome genes [10]. Another reason to study the X chromosome is its hemizygous state in males. Due to males' hemizygous sex chromosomes, there is no masking of recessive effects. Slightly deleterious mutations are more efficiently selected against when they are on the X chromosome, and so, their frequency under mutation-selection balance will be lower. But the lack of masking also promotes the accumulation of mutations on the X chromosome that present favorable attributes in males. An additional reason to study the X chromosome is that only one or two of the structural genes for accessory gland proteins (Acps) are found on the X chromosome. Therefore, if there is variation in sperm competition in X substitution lines, either it is not mediated by Acps, or there are additional Acps or Acp-modifying enzyme genes to be discovered on the X. Thus, the goal of this study is to investigate the variation among sperm competition of *D. melanogaster* in genetically identical flies that differ only in their X chromosomes and to determine which polymorphisms are causing this variation through sequencing and association tests.

Genes that might be involved in the process of sperm competition have been found by a number of routes. Parisi et al. [10] used an Incyte microarray to assay the expression of 14,142 transcripts that were predicted to have sex-differential expression profiles. From the total number of transcripts identified on the X chromosome, only those that had a high testis-to-ovary expression ratio were considered as possible candidate genes. Other potential candidate genes were the 27 genes identified by Drnevich et al. [11] that showed different expression levels due to within-population variation of male reproductive success under competitive conditions (MCRS). Significant differences were detected between high MCRS and low MCRS genotypes using a combination of traditional quantitative genetics and oligonucleotide microarrays. Genes on the X chromosome, whose P values indicated significant association between expression variation and male fertility were selected to be analyzed for polymorphisms affecting sperm competition. Another source of potential candidate genes was the result of efforts to apply mass spectrometry analysis to the sperm proteome [12–14]. There are currently 1108 proteins in the *D. melanogaster* sperm proteome that were identified using whole-sperm mass spectrometry. Of the proteins identified in these studies, those residing on the X chromosome were considered as potential candidate genes for association testing. Many proteins that were identified by mass spectrometry had unknown function, and thus could possibly play a role in sexual antagonism in flies. The data collected from Haerty et al. [6] provided other potential candidate genes. Haerty et al. applied phylogenic analysis by maximum likelihood and examined 2505 sex and reproduction-related genes (SRR genes) across the twelve species of Drosophila, whose genome sequence was available. For the purposes of this study, genes found on the X chromosome that were undergoing positive selection were considered as possible candidate genes for association testing. To refine the list, we ran all candidate genes through FlyAtlas [15] to determine their expression levels in the testes and the male accessory glands compared to the whole body. Only those genes that showed high expression in either the testes or male accessory glands were examined further.

2. Materials and Methods

2.1. Drosophila Stocks. Balancer stocks were used to construct 57 X chromosome substitution lines, whose X chromosomes were derived from a wild population in Ithaca, NY, by Dr. Erin Hill-Burns, and details of the line construction can be found in Hill-Burns and Clark [16]. These lines, which were used in all phases of this study, are coisogenic for the second, third, and fourth chromosomes, but each has a unique homozygous X chromosome. The experimental flies had the red wild-type eye color. In the sperm competition assays, tester flies were a standard *cn bw* lab strain, which is phenotypically white eyed. Lines were raised at 24°C on 12 h light/dark cycle using standard agar-dextrose-yeast medium.

2.2. Scoring Sperm Competition Phenotypes. The methods for scoring sperm competition phenotypes follow closely the methods of Fiumera et al. [8] and Baker et al. [9]. *cn bw* females were mated to two different males in succession in order to determine the extent of sperm competition among the lines of flies. In the "defense" crosses, the first male to mate were males from the X-replacement lines, and the second males were *cn bw*. P1 is the sperm competition index for these crosses, and it scores the proportion of progeny sired by the first (X-replacement line) males. In the "offense" crosses, the first males to mate were *cn bw*, and the second males were from the X-replacement lines. P2 is the sperm competition index for these crosses, and it scores the proportion of progeny sired by the second (X-replacement line) males. We also scored the fecundity of the females in the defense crosses (when the experimental male is first to mate) and offense crosses (when the experimental male is second to mate), where fecundity is defined as the total count of offspring produced by each female across all three vials. The remating rate is the proportion of females that had mated to a *cn bw* male and subsequently remate with an experimental male. The final statistic is refractoriness to remating, which expresses the proportion of females that mate once with an experimental male and then fail to remate.

In the offense experiment, virgin *cn bw* females were first mated to virgin *cn bw* males in a mass mating on the morning of day one. In the afternoon of day one, individual females were transferred, without anesthesia, to vial one, and the males were discarded. In the afternoon of day three, two virgin experimental males were added to vial one. The next day, the females were transferred to vial two. On day seven, the female was transferred to vial three and allowed to lay eggs until day 11 when she was discarded. The progeny were then scored for eye color to determine paternity at 15, 17, and 19 days after the initiation of egg laying for vials one, two, and three, respectively. For the defense experiment, the experimental male was the first male to mate. On the first day, the *cn bw* females were individually mated with the experimental X substitution line males. On day three, the females were transferred to vial two, and the males were discarded. On day four, two *cn bw* males were added to vial two. By day six, the females were transferred to vial three, and the males were discarded. The females were discarded on day nine, and the progeny were scored for eye color at 15, 17, and 19 days after the initiation of egg laying for vials one, two, and three, respectively. The sources of variation were tested for statistical significance across both the offense and defense experiments. General linear models were used to assess the significance of among-line variability in the face of block and replicate variation, using the R statistics package lme4 (http://www.r-project.org/). Permutation tests based on the χ^2 statistics were used to test for significant associations with SNP genotypes.

2.3. Sequencing of Candidate Genes. A total of eleven male reproductive candidate genes on the X chromosome (*CG15208, CG15200, CG4349, CG15035, CG12470, CG9806, CG1559, CG9156, CG17450, CG32819, and CG32820*) were chosen for DNA sequencing to test for association between any sequence polymorphisms and the degree of sperm competition. Genes were chosen based on (1) microarray data from Parisi et al. [10] examining global gene expression in *D. melanogaster*, (2) quantitative genetics and oligonucleotide microarrays from Drnevich et al. [11] that investigated male reproductive success under competitive conditions, (3) whole-mass sperm spectrometry from Dorus et al. [12] and Findlay et al. [13] that identified proteins of the *D. melanogaster* sperm proteome, and (4) phylogenic analysis by maximum likelihood from Haerty et al. [6] that examined sex and reproduction-related genes across all twelve species of Drosophila. All candidate genes were run through the database FlyAtlas [15] to determine their expression levels in the testes and male accessory glands (MAGs) compared to the whole fly. Nine of these genes demonstrated high expression in the testes, while two genes showed high expression in the male accessory glands (Table 1).

Primers were designed for each gene using Primer 3 [17]. PCR products ranged from about 700–1000 base pairs long. Coding regions of each gene were flanked by these primers and amplified using PCR. The PCR product was then sequenced using Applied Biosystems Automated 3730 DNA Analyzer with Big Dye Terminator chemistry and AmpliTaq-FS DNA Polymerase. Reference sequences for each gene were obtained using the FlyBase database [18], and the contigs for each primer set were assembled and aligned according to these references. Sequencher 4.6 (Gene Codes Corporation, 2007) was used to detect polymorphic nucleotides and indels within the aligned contigs.

2.4. Statistical Methods. Statistical tests for the fecundity of the offense and defense experiment, P1′, P2′, remating rate, and refractoriness were performed using the R statistical package version 2.01. General linear models were used to assess the significance of heterogeneity among the different lines for each statistical test. For the tests of sperm competition, the model was

$$Y_{ijk} = \mu + L_i + B(L)_{ij} + \varepsilon_{ijk}, \quad (1)$$

where Y_{ijk} is the response variable (P1′, P2′, or fecundity), μ is the grand mean, L_i is the effect of the ith line, $B(L)_{ij}$ is the block effect (the design was a balanced nested design), and ε_{ijk} is the error term, quantified by the among replicate

Table 1: Candidate male reproductive genes on the X chromosome. Criteria for selecting genes for DNA sequencing and association testing are presented. Testes/whole body and male accessory gland (MAG)/whole body columns present ratios of expression from FlyAtlas data.

Gene	Testes/whole body	MAG/whole body	Predicted function[e]
CG15208[a,c]	14.2	0.1	Unknown
CG15200[a,b]	10.4	0	Unknown
CG4349[a,d]	9.5	0.3	Iron ion binding (*Fer3HCH*)
CG15035[a]	15.9	0.1	Protein phosphatase
CG12470[a,d]	7.6	0	Unknown
CG9806[c]	0.5	92.5	Membrane alanyl aminopeptidase
CG1559[c]	0.7	1.00	DNA/RNA helicase (*Upf1*)
CG9156[c,d]	10.9	0	Protein phosphatase 1 (*Pp1-13c*)
CG17450[d]	14.4	0	Microtubule associated
CG32819[d]	14.4	0	Microtubule associated
CG32820[d]	14.4	0	Microtubule associated

Gene candidacy determined from: [a]Parisi et al. (2003) [10], [b]Drnevich et al. (2004) [11], [c]Haerty et al. (2007) [6], [d]Dorus et al. (2006) [12], [e]Flybase (http://www.flybase.org).

variance. The linear model was fitted with the lme4 package of the statistics program R (http://cran.r-project.org/), and homogeneity of variances, approximate fit of residuals to normality, and absence of outliers were all assessed with R utilities. Tests of heterogeneity in remating rate and refractoriness were essentially tests of equality of proportions under binomial sampling, and were tested by Chi-square. The association between each SNP and the reproductive phenotypes was tested with linear models, testing the null hypothesis that the phenotypes of the lines bearing the alternative two alleles for each SNP were sampled from the same distribution. Because many tests were performed for each phenotype, we report the results of these tests as the false-discovery rates or q-values [19].

3. Results

3.1. Reproductive Phenotype Variation. All 57 X-chromosome lines were successfully scored in both the offense and the defense experiments. Twenty replicate double matings were attempted for each line; however, not all lines yielded all twenty replicates with doubly mated females. Overall 57,919 progeny were scored from the 57 lines for the defense experiment, and 92,467 progeny were scored from the 57 lines for the offense experiment. For the defense experiment in which the X experimental male was presented first, a total of 1130 females successfully mated, with 531 double mating and 599 mating only once. For the offense experiment in which the X experimental male was presented second, a total of 1172 females successfully mated, with 645 double mating and 527 mating only once.

Six statistics were recorded and analyzed for each line. The first two statistics were P1 and P2, which are the proportion of offspring fathered by the experimental X chromosome males when he is either the first or second male to mate with the female. The terms P1′ and P2′ refer to the arcsine square root transformed values of P1 and P2. This transformation is done to stabilize variances, which otherwise become inherently small when values approach zero or one. The next two statistics were the fecundity of the offense and defense experiments, which is the total count of offspring produced by each doubly mated female. The "remating rate" is defined as the proportion of females that remate when the experimental male is second to mate. Finally, "refractoriness" expresses the proportion of females that do not remate after having mated once with an X-line experimental male.

The results of the linear model fits revealed significant genetic variation among the X replacement lines in P1′, P2′, and fecundity of the offense experiment (Table 2). There was no significant variation in fecundity for the defense experiment across the 57 X substitution lines ($P = .5265$). In this case, it appears that genotype does not affect female fecundity across the lines when the X experimental males are first to mate. Fecundity showed a good fit to the normal distribution, and linear models confirmed that the fecundities of females across the defense crosses were homogeneous.

Results for the fecundity counts of the offense experiment differed slightly from those from the defense experiment. For one, fecundity of the offense experiment showed significant variation among the 57 *D. melanogaster* lines ($P = 1.98 \times 10^{-8}$). Thus, in this case, genotype does seem to affect the difference in fecundity seen across lines.

Considering sperm competitive metrics, the P1′ statistic shows highly heterogeneous variation among the lines ($P = 1.38 \times 10^{-6}$) indicating that genotype does have a significant effect on the proportion of offspring fathered by the experimental X chromosome males when he is the first to mate. Figure 1 shows the distribution of the P1 (and P2) line means, showing a mode for the line mean of P1 around 0.4. Figure 2 shows a boxplot of the P1 statistics stratified by line, giving a visual confirmation of the linear model that the lines are highly heterogeneous.

The P2′ statistic was found to be highly significant across the 57 X substitution lines ($P = 2.2 \times 10^{-16}$), thus indicating that genotype has an effect on the variation seen for the proportion of progeny sired by the experimental X

Table 2: Reproductive phenotype statistics. P1′ is the proportion of progeny sired by the first male in defense crosses (arcsine transformed). P2′ is the proportion of progeny sired by the second male in offense crosses (arcsine transformed). Summary of the reproductive phenotypes statistics using linear models, a total of 57,919 progeny were counted for the defense experiment, and 92,467 progeny were counted for the offense experiment.

Phenotype	Mean across lines	Range of line means	D.F.	Test statistic	P value
Fecundity-D	80.69	52–103.583	56	$F = 0.98$	.5265
P1′	0.8666	0.53088–1.5708	56	$F = 2.25$	1.38×10^{-6}
Fecundity-O	110.79	76.125–153.75	56	$F = 2.53$	1.98×10^{-8}
P2′	1.01	0.27237–1.49932	56	$F = 7.12$	2.20×10^{-16}

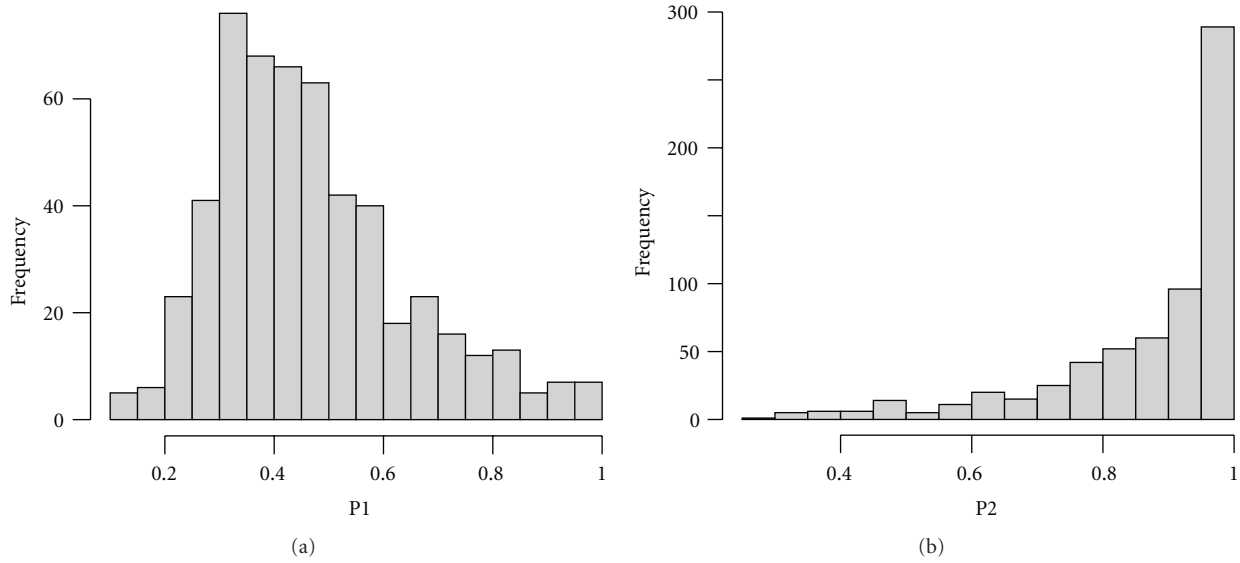


Figure 1: *The distribution of P1 and P2 across 57 X substitution lines.* P1 is the defense test, where the X-line experimental male was the first to mate (followed by a *cn bw* control male), and the P2 assay uses the reverse order. P2 tests the ability of the X experimental males to displace resident sperm. In both cases, arcsine transformation was applied before statistical testing.

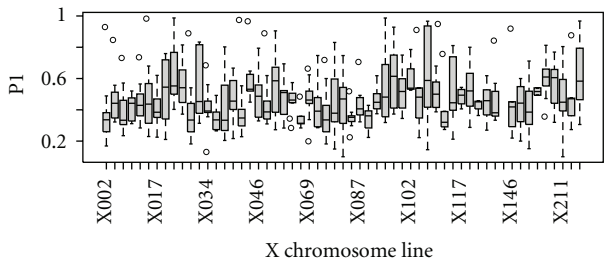


Figure 2: *Boxplot of the P1 scores of the 57 X replacement lines.* The proportion of offspring sired by the X experimental male out of 57 naturally derived X chromosomes of *D. melanogaster* when the respective X-line male is the first to mate (followed by the *cn bw* control male). P1 scores are highly heterogeneous across lines ($P = 1.38 \times 10^{-6}$).

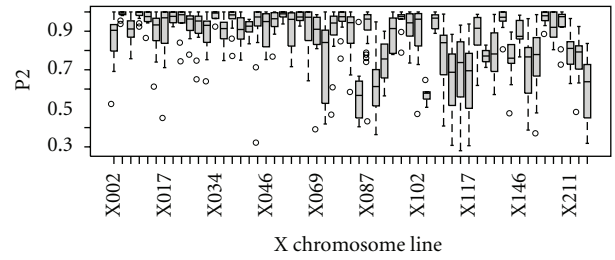


Figure 3: *Boxplot of the P2 scores of the 57 X replacement lines.* The proportion of offspring sired by the X experimental male out of 57 naturally derived X chromosomes of *D. melanogaster* when the respective X-line male is the second to mate (after the *cn bw* control male). P2 scores are highly heterogeneous across lines ($P = 2.2 \times 10^{-16}$).

chromosome male when he is the second to mate. The right panel of Figure 1 displays the distribution of the P2, indicating a highly skewed distribution (statistical analysis was performed on arcsine transformed data). Figure 3 presents a boxplot showing the heterogeneity across lines in the P2 statistic.

Next, we examined the variation among lines in the mating success rate of second males (in the offense crosses), and the ability of the first-mating males to induce refractoriness in their female mates (in the defense crosses). A Chi-square test found that remating rate showed statistically significant heterogeneity across lines ($P < .001$).

Table 3: Remating rate and refractoriness. Summary of the remating rate and refractoriness using Chi-square analysis.

Reproductive phenotype	Mean value across lines	Range of line means	D.F.	Chi-square statistic	P value
Remating rate	0.55	0.2105–0.8000	56	94.809	<.001
Refractoriness	0.53	0.0000–0.7500	56	97.072	<.001

This indicates that the different male genotypes among the lines varied in their ability to induce already mated females to remate. Refractoriness, on the other hand, is the proportion of females that do not remate after having mated with an X experimental male. This phenotype also was statistically significantly heterogeneous ($P < .001$) indicating that there is heterogeneity among lines in the ability of males to induce refractoriness among females. Table 3 shows the average remating and refractoriness rates as well as their significance values.

Correlation tests were performed to determine if certain reproductive phenotypes were significantly correlated with each other (Table 4). There was a significant correlation seen between fecundity of the defense experiment and fecundity of the offense experiment ($P = .0098$). These two parameters showed a positive linear correlation indicating that females who had a higher number of offspring when they mated with the X experimental males first tended to also have a higher number of offspring when they mated with the X experimental males second. Additionally, the fecundity of the offense experiment demonstrated a statistically significant correlation with the transformed P2 data ($P = 7.58 \times 10^{-7}$). The strong positive correlation implies that females that had a higher number of progeny when they mated with the X experimental second males also had a higher proportion of progeny from that second male. Furthermore, fecundity of the defense experiment also showed a significant correlation P2′ ($P = .0134$). This correlation suggests that the higher number of offspring sired in the defense experiment, the larger the proportion of offspring sired from the X experimental males in the offense experiment. A significant correlation was also seen between the P1′ and refractoriness ($P < .0001$). This strong negative correlation suggests that males with strong defense induce a higher portion of females to be refractory to remating. Another statistically significant correlation was found between the fecundity of the offense experiment and the remating rate ($P < .001$). This correlation indicates that as the fraction of females that tend to remate increases, the number of progeny sired in the offense experiment increases. A final correlation observed among the reproductive phenotypes was between P2′ and remating rate ($P < .0001$ resp.). This strong positive correlation suggests that males that have high P2, and thus strong sperm competition offense, were much more successful at inducing already mated females to remate. None of the other correlations were significant.

Table 4: Correlations Between Reproductive Phenotypes. Correlation coefficients and P values for correlations among line means of reproductive phenotypes.

Variables	Correlation coefficient	P value
Fec-O versus Fec-D	0.3395	.0098
P1′ versus P2′	0.0411	.7614
Fec-O versus P2′	0.6013	7.58×10^{-7}
Fec-D versus P2′	0.3258	.0134
P1′ versus refractoriness	−0.483	<.0001
Fec-O versus remating	0.434	<.001
P2′ versus remating	0.695	<.0001

3.2. Identifying SNPs and Association Tests. A total of 131 SNPs and indels were found within the 11 candidate genes using Sequencher 4.6. Three of the 11 genes (*CG15200*, *CG4349*, and *CG15035*) were fully sequenced, while the rest had only portions of the genes sequenced. Association tests were run using linear models in the R statistical software to determine if there were any significant associations between the SNPs and indels, and the reproductive phenotypes.

Nine significant associations were found between the polymorphisms and the reproductive phenotypes ($q < 0.05$). Of the nine associations, two were located in the 3′UTR region, two were located in intronic regions, and five were located in the coding regions. The variants in the coding regions consisted of one nonsynonymous mutation, one synonymous mutation, and two frameshifting deletions. One frameshift deletion was found in gene *CG9156* position 402, associated with both fecundity of the defense experiment and the P1′ statistic. Table 5 displays all the significant polymorphism-phenotype combinations.

Two polymorphisms and a deletion were found to associate significantly with fecundity of the offense experiment. Two genes, *CG15200* and *CG15208*, each had one polymorphism, and there was a deletion allele in gene *CG17450*. Alleles of these genes associated with either an increase or decrease in the total amount of progeny sired when the X experimental male was the second to mate. Since high fecundity is related to a higher fitness, these alleles can be considered either advantageous or disadvantageous. The polymorphisms *CG15200* SNP 471 and *CG17450* deletion 487-488 showed a significant decrease in fecundity ($q = 0.008$ and $q = 0.048$ resp.), while *CG15208* SNP 1295 showed a significant increase in fecundity of the offense experiment ($q = 0.012$). These three SNPs explained 11.5%, 19.3%, and 12.8% of the interline variance in fecundity in the offense crosses.

Two polymorphisms were found to be associated with changes in P2′. Because P2 measures the paternity success (it is the proportion of offspring sired by the second-mating male), it is a direct proxy measure for reproductive fitness. The first associated polymorphism was found in gene *CG15208* at position 1019, while the second polymorphism

Table 5: Significant Associations between reproductive phenotypes and DNA sequence polymorphisms found in each gene. q-value is derived from the linear model testing homogeneity of the phenotype across lines stratified by the targeted SNP.

Gene name	Association	Index number	q-value	Location in gene	Mutation	Amino acid change
CG15200	Fecundity-O	471	0.0079	3′ UTR		
CG15208	P2′	1019	0.0200	Coding	Nonsyn	Serine → Proline
CG15208	Fecundity-O	1295	0.0121	3′′UTR		
CG17450	P2′	210	0.0482	Coding	Synon	
CG17450	Fecundity-O	487-488	0.0478	Intron		
CG17450	Fecundity-D	788–795	0.0146	Intron		
CG9156	P1′	159-160	0.0328	Coding	Frameshift	
CG9156	Fecundity-D	402	0.0182	Coding	Frameshift	
CG9156	P1′	402	0.0032	Coding	Frameshift	

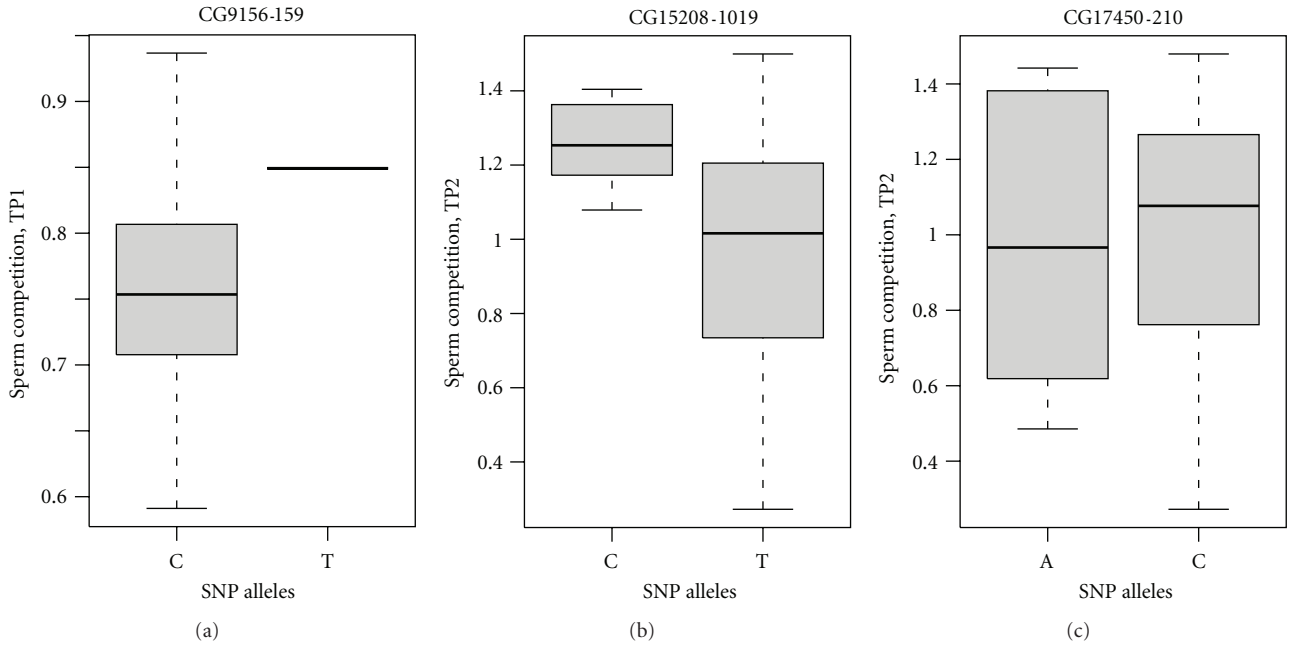


Figure 4: *SNP allelic associations with P1 and P2.* (a): associations between a deletion found on the gene *CG9156* and the arcsine square transformation of the proportion of progeny sired by the X experimental male when he was the first to mate. Gene *CG9156* deletion 159-160 had 21 lines with a deletion in both positions, 4 lines with a deletion in just the second position, and 14 lines with the reference sequence AA. This deletion showed a statistically significant decrease in the P1′ statistic ($q = 0.033$). (b): for *CG15208* position 1019, there were 7 lines with the rare allele C and 48 lines with the common allele T. Those lines containing the rare allele showed a significant increase in P2′ ($q = 0.020$). (c): for *CG17450* position 210, 4 lines had the rare allele G, while 41 lines contained the common allele C. This rare allele also demonstrated a statistically significant increase in P2′ ($q = 0.048$).

was found in gene *CG17450* at position 210. Both polymorphisms were low frequency alleles that showed an increase in P2′. Since both fecundity of the offense and fecundity of the defense experiment were significantly correlated with P2′, we can conclude that these low frequency alleles may be advantageous, since in the lab environment, they lead to an increase in fitness. This increase in fitness is a result of the increase in the number of offspring from the X experimental male when he is second to mate. Even though the female already mated, she has a tendency to use more of the second male's sperm if he possessed the novel allele. The right two panels of Figure 4 illustrate the difference in P2′ across genotypes of the associated SNPs. These two SNPs explained 15.4% and 19.8% of the interline variance in P2.

Two mutations were found to associate with fecundity in the defense experiment. One polymorphism was found in gene *CG17450*, and the SNP was immediately flanked by a seven basepair deletion. This mutation demonstrated an increase in the total number of progeny sired in the defense experiment, and it explained 13.4% of the interline variance in fecundity defense. Since an increase in fecundity is related to an increase in fitness, these rare alleles and deletion can be considered advantageous due to the increased reproductive fitness they confer. Gene *CG9156* position 402 contained

a deletion that was significantly associated with the fecundity of the defense experiment. This SNP explained 3.96% of the interline variance in fecundity in the defense crosses.

Two deletions were found to associate with P1′, the sperm competition parameter for first-mating males. One deletion was found in gene *CG9156*, at position 159. The leftmost panel of Figure 4 depicts the association between this polymorphism and P1′. The deletion allele appears to be associated with a decrease in P1′, and it explained 17.8% of the interline variance in P1′. No other associations with this deletion allele were detected, so the reduced P1′ alone suggests that it has a net negative effect on fitness. Gene CG9156 at position 402 also contained a deletion that was significantly associated with the P1′ statistic, and it explained 3.9% of the interline variance in fecundity.

4. Discussion

The overall magnitude of variation seen among these 57 X substitution lines is roughly consistent with what was found in similar studies of the second and third chromosomes [7, 8]. However, these results are still surprising, given the paucity of seminal protein genes found on the X chromosome [20]. Many of the attributes of male reproductive fitness varied significantly across this set of X-chromosome extracted lines, including fecundity of the offense experiment, P1′, P2′, remating rate, and refractoriness to remating. Thus male, genotype seems to have a significant effect on all these reproductive phenotypes. Although, fecundity of the defense experiment did not show significant variation across lines, two mutations were found to have a significant association with the male-induced differences in fecundity.

The different reproductive phenotypes that were scored in our collection of X-chromosome replacement lines showed a strong pattern of pairwise correlations. The pattern of correlations overall shows exactly the opposite of what we would call antagonistic pleiotropy. The alleles that were found to be best for one reproductive phenotype tended to be best for all. In such a situation, one would expect the alleles with this best-at-everything phenotype to become fixed in the population, begging the question of what maintains this high level of genetic variation in these important fitness traits.

DNA sequence analysis of the 57 X substitution lines using simple statistical tests allowed us to determine if there were polymorphisms among the candidate genes that were associated with variation in aspects of reproductive fitness. A total of 131 mutations including single nucleotide polymorphisms, insertions, and deletions were found among the 11 candidate genes that were analyzed. Association tests were performed between these mutations and the varying reproductive phenotypes seen in the X experimental males. Although association tests with these polymorphisms do not prove causation of the reproductive phenotypes, they do give insight into evolutionary changes that may be occurring, and thus prove important to study.

Of the 131 mutations identified, nine were found to significantly associate with specific phenotypes. These mutations were found among four genes which contain naturally occurring polymorphisms that are associated with differences in mating. *CG15208* showed two significant associations with reproductive phenotypes, but there is as yet no identified known or predicted function. The first was at position 1019 in the coding region. This polymorphism caused a nonsynonymous mutation changing a serine into a proline, and was found to have an association with the proportion of progeny sired by the X experimental male when he was the second to mate. The other polymorphism found in this gene was at position 1295 and located in the 3′UTR region. There was found to be a statistically significant association between this polymorphism and the fecundity of the offense experiment. *CG15208* was originally chosen to be analyzed due to its high testis-to-ovary expression ratio, as determined by Parisi et al. [10], and its high level of expression in the testes compared to the whole body (14.4), determined by FlyAtlas [15]. Furthermore, this gene is one of the sex and reproduction-related genes that was found to be undergoing positive selection [6]. Thus, this gene has been shown to be important in the variation of reproductive phenotypes. However, more investigation is necessary to understand the mechanisms behind these polymorphisms that lead to phenotypic variation.

CG15200 does not have a known function but was chosen as a candidate gene because of its high testis-to-ovary expression ratio determined by Parisi et al. [10] and the significant difference seen in male reproductive success under competitive conditions [11]. This gene was also shown to have a higher expression level in the testes than in the whole body (10.4). One significant polymorphism found in this gene was at position 471 in the 3′UTR region, where many variants influence the transcript abundance of a gene. This polymorphism was shown to be significantly associated with fecundity in the offense experiment.

In *CG17450*, several mutations were found that associated with three different reproductive phenotypes, thus making this gene the most pleiotropic of all the genes analyzed. *CG17450* is known to be associated with microtubule binding and microtubule cytoskeleton organization and biogenesis, but at present, we do not have any understanding of the mechanism whereby polymorphisms in this gene should mediate fitness differences. However, this gene was chosen as a good candidate gene, because it was shown to be part of the Drosophila sperm proteome [12–14] as well as to have a high expression level in the testes compared to the whole body (14.4). The first polymorphism identified in this gene was at position 210 in the coding region. This SNP caused a synonymous mutation but was shown to affect the proportion of progeny sired by the X experimental male when he was the second to mate. The second mutation seen in this gene was a deletion in the intronic region of the gene at position 487-488. This deletion seemed to relate to a decrease in the fecundity of the offense experiment. The final mutation identified in this gene was an SNP followed by a deletion at positions 788–795. This mutation was also found in the intronic region of the gene but was associated with an increase in the fecundity of the defense experiment.

The fourth gene that showed segregating variation that was significantly associated with reproductive phenotypes

Table 6: Primers used to amplify the desired region of each of the 11 candidate genes.

Gene name	Forward primer (F)	Reverse primer (R)	PCR product
CG15208	CGCGATATGATCGAACAGTG	TTGCAAAATGCGAATAGCTG	787 bp
CG15200	TCCTTCAAAGTCAATGCTGAGA	TTGCAGTGAAGCACAAAACA	785 bp
CG4349	CGACCCAGTTGCCTTTTAGA	AAACAATTGATCGGCCAAAA	852 bp
CG15035	GATGTGACACCCCTTCGATT	CACTTTACCAGATCGCACCA	839 bp
CG15035	GCAGCTGTACGGCGTGTAT	TGACCACTTTGTTTTGGAGAGA	850 bp
CG12470	GATCACGTTGCTTTGCTGAA	ACGAAGAAGAGCCGGGTAAT	749 bp
CG9806	TATTTCGGGAGAGCAGTCGT	TGATGGAGAAGGCCAATTTC	821 bp
CG1559	GCAGATGGACGGTTTGCTAT	TGTTTGCGTGATTTTCCTTG	972 bp
CG9156	TCCATGAAACACGAACTGGA	GTCTGGAGAATGGGACTGGA	709 bp
CG17450	AGAAGAAGCCACCGAGGACT	AAAGTGGCCTACTCCGATTG	823 bp
CG17450	ATATCGCATCAGCCATGTCA	CCGAGAGGTTGCTGAAAGAC	945 bp
CG32819	GCCAGGCTACCGTTTAATCA	GAATTTGGACTGGGTGTTGG	987 bp
CG32820	GGCCCACATGCTAATCAACT	ATGACATGGCTGATGCGATA	913 bp

was *CG9156*. This gene is known to be a protein phosphatase and is important for protein amino acid dephosphorylation. It was chosen as a good candidate gene to analyze because it was shown to be a sex and reproduction-related gene that was undergoing positive selection [6], and it is clearly expressed as part of the sperm proteome [12]. This gene also demonstrates an 11-fold increase in expression in the testes compared to the whole body. *CG9156* was found to have two deletions that caused frameshift mutations that affected different reproductive phenotypes. The first deletion is at positions 159-160 and is located in the coding region. This deletion is associated with a decrease in the proportion of progeny sired by the X experimental male when he is the first to mate. The second deletion in this gene is also located in the coding region and is found at position 402. This mutation was found to associate with the fecundity of the defense experiment and also with the proportion of offspring sired by the X experimental male when he is the first to mate.

The other seven genes analyzed did not show significant associations between their mutations and variation in the reproductive phenotypes. However, all of these genes contain important functions that if disabled could potentially alter reproductive phenotypes. *CG4349* is the ferritin 3 heavy chain (gene symbol *Fer3HCH*) and is known to be involved in ferrous and ferric iron binding, while *CG15035* is inferred to be a protein phosphatase. *CG12470* has no known function but is found in the sperm proteome [12] and is shown to have a high testis to ovary expression ratio [10]. Gene *CG9806* is involved in alanyl aminopeptidase activity and gene *CG1559* is needed for DNA/RNA helicase activity as well as nucleic acid binding. The last two genes *CG32819* and *CG32820* are both associated with microtubule function. Although these genes did not contain polymorphisms that significantly associated with reproductive phenotypes, they may still have an important effect on the phenotypes. Most of the genes were not sequenced in their entirety, and it is possible that significant effects of sequence variants within large introns were missed.

Due to a lack of accessory gland proteins on the X chromosome, it is tempting to assume that differences in the X chromosome would not cause differences in sperm-competition phenotypes. However, this study has identified several mutations on the X chromosome that are associated with either fecundity, remating tendencies, or measures of sperm competition. Perhaps it is the hemizygous state of male flies that is allowing recessive polymorphisms that would normally be masked to be seen or the differences in dosage compensation that occurs in the X chromosome that is causing these changes. It is also possible that we have not yet been exhaustive in our discovery of accessory gland proteins and that there are yet to be discovered Acps on the X. But, we can be confident that the candidate genes that had significant SNP associations are not Acps, because of their absence of expression in the male accessory gland. However, these genes could affect Acps as modifying enzyme genes that affect development of function of accessory glands. Whatever the cause, the data indicates that genotypic variation on the X chromosome does affect reproductive phenotypes including sperm competition. Genes *CG15208*, *CG15200*, *CG17450*, and *CG9156* are good candidates for further investigation seen as they all contain mutations that are significantly associated with different reproductive phenotypes. Finally, we note that the significant effects that were found in this study were obtained in a particular context, and while we have confidence that the genes involved have effects in that context, changes in many experimental attributes, including the genotypes of other competing flies in the assays, may alter the sperm competitiveness rankings in unpredictable ways [21].

Appendix

The forward and reverse primers utilized can be found listed in Table 6.

Acknowledgments

The authors thank Xiaoyun Wang, Sean Hackett, Anna Beavis, and Scott Pilla for technical assistance. This work was

supported in part by NIH Grant no. RO1 HD059060 to A. G. Clark and Mariana Wolfner and NSF DEB-0743125 to A. G. Clark and Anthony Fiumera.

References

[1] R. C. Griffiths, S. W. McKechnie, and J. A. McKenzie, "Multiple mating and sperm displacement in a natural population of *Drosophila melanogaster*," *Theoretical and Applied Genetics*, vol. 62, no. 1, pp. 89–96, 1982.

[2] L. G. Harshman and A. G. Clark, "Inference of sperm competition from broods of field-caught *Drosophila*," *Evolution*, vol. 52, no. 5, pp. 1334–1341, 1998.

[3] M. Imhof, B. Harr, G. Brem, and C. Schlötterer, "Multiple mating in wild *Drosophila melanogaster* revisited by microsatellite analysis," *Molecular Ecology*, vol. 7, no. 7, pp. 915–917, 1998.

[4] M. F. Wolfner, "The gifts that keep on giving: physiological functions and evolutionary dynamics of male seminal proteins in *Drosophila*," *Heredity*, vol. 88, no. 2, pp. 85–93, 2002.

[5] F. W. Avila, L. K. Sirot, B. A. Laflamme, C. D. Rubinstein, and M. F. Wolfner, "Insect seminal fluid proteins: identification and function," *Annual Review of Entomology*, vol. 56, pp. 21–40, 2011.

[6] W. Haerty, S. Jagadeeshan, R. J. Kulathinal et al., "Evolution in the fast lane: rapidly evolving sex-related genes in *Drosophila*," *Genetics*, vol. 177, no. 3, pp. 1321–1335, 2007.

[7] A. C. Fiumera, B. L. Dumont, and A. G. Clark, "Sperm competitive ability in *Drosophila melanogaster* associated with variation in male reproductive proteins," *Genetics*, vol. 169, no. 1, pp. 243–257, 2005.

[8] A. C. Fiumera, B. L. Dumont, and A. G. Clark, "Associations between sperm competition and natural variation in male reproductive genes on the third chromosome of *Drosophila melanogaster*," *Genetics*, vol. 176, no. 2, pp. 1245–1260, 2007.

[9] B. S. Baker, M. Gorman, and I. Marin, "Dosage compensation in *Drosophila*," *Annual Review of Genetics*, vol. 28, pp. 491–521, 1994.

[10] M. Parisi, R. Nuttall, D. Naiman et al., "Paucity of genes on the *Drosophila* X chromosome showing male-biased expression," *Science*, vol. 299, no. 5607, pp. 697–700, 2003.

[11] J. M. Drnevich, M. M. Reedy, E. A. Ruedi, S. Rodriguez-Zas, and K. A. Hughes, "Quantitative evolutionary genomics: differential gene expression and male reproductive success in *Drosophila melanogaster*," *Proceedings of the Royal Society B*, vol. 271, no. 1554, pp. 2267–2273, 2004.

[12] S. Dorus, S. A. Busby, U. Gerike, J. Shabanowitz, D. F. Hunt, and T. L. Karr, "Genomic and functional evolution of the *Drosophila melanogaster* sperm proteome," *Nature Genetics*, vol. 38, no. 12, pp. 1440–1445, 2006.

[13] G. D. Findlay, M. J. MacCoss, and W. J. Swanson, "Proteomic discovery of previously unannotated, rapidly evolving seminal fluid genes in *Drosophila*," *Genome Research*, vol. 19, no. 5, pp. 886–895, 2009.

[14] E. R. Wasbrough, S. Dorus, S. Hester et al., "The *Drosophila melanogaster* sperm proteome-II (DmSP-II)," *Journal of Proteomics*, vol. 73, no. 11, pp. 2171–2185, 2010.

[15] V. R. Chintapalli, J. Wang, and J. A. T. Dow, "Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease," *Nature Genetics*, vol. 39, no. 6, pp. 715–720, 2007.

[16] E. M. Hill-Burns and A. G. Clark, "X-linked variation in immune response in *Drosophila melanogaster*," *Genetics*, vol. 183, no. 4, pp. 1477–1491, 2009.

[17] S. Rozen and H. Skaletsky, "Primer3 on the WWW for general users and for biologist programmers," in *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, S. Krawetz and S. Misener, Eds., pp. 365–386, Humana Press, Totowa, NJ, USA, 2000.

[18] M. A. Crosby, J. L. Goodman, V. B. Strelets, P. Zhang, and W. M. Gelbart, "FlyBase: genomes by the dozen," *Nucleic Acids Research*, vol. 35, supplement 1, pp. D486–D491, 2007.

[19] J. D. Storey, "A direct approach to false discovery rates," *Journal of the Royal Statistical Society. Series B*, vol. 64, no. 3, pp. 479–498, 2002.

[20] K. R. Ram and M. F. Wolfner, "Seminal influences: *Drosophila* Acps and the molecular interplay between males and females during reproduction," *Integrative and Comparative Biology*, vol. 47, no. 3, pp. 427–445, 2007.

[21] C. Y. Chow, M. F. Wolfner, and A. G. Clark, "The genetic basis for male x female interactions underlying variation in reproductive phenotypes of *Drosophila*," *Genetics*, vol. 186, no. 4, pp. 1355–1365, 2010.

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 595121, 11 pages
doi:10.4061/2011/595121

Research Article

Is Speciation Accompanied by Rapid Evolution? Insights from Comparing Reproductive and Nonreproductive Transcriptomes in *Drosophila*

Santosh Jagadeeshan,[1,2] Wilfried Haerty,[1] and Rama S. Singh[1]

[1] *Department of Biology, McMaster University, Hamilton, ON, Canada L8S 4KI*
[2] *Smithsonian Tropical Research Institute, P. O. Box 0834-03092, Balboa, Ancón, Panama*

Correspondence should be addressed to Rama S. Singh, singh@mcmaster.ca

Received 10 January 2011; Revised 4 April 2011; Accepted 19 May 2011

Academic Editor: Jose M. Eirin-Lopez

The tempo and mode of evolutionary change during speciation have remained contentious until recently. While much of the evidence claiming speciation is an abrupt and rapid process comes from fossil data, recent molecular phylogenetics show that the background of gradual evolution is often broken by accelerated rates of molecular evolution during speciation. However, what kinds of genes affect or are affected by speciation remains unexplored. Our analysis of 4843 protein-coding genes in five species of the *Drosophila melanogaster* subgroup shows that while ~70% of genes follow clock-like evolution, between 17–19.67% of loci show signatures of accelerated rates of evolution in recently formed species. These genes show 2-3-fold higher rates of substitution in recently diverged species compared to older species. This fraction of loci affects a diverse range of functions. Only a small proportion of reproductive genes experience speciation-related accelerated changes but many sex-and -reproduction related genes show an interesting pattern of persistent rapid evolution suggesting that sex-and-reproduction related genes are under constant selective pressures. The identification of loci associated with accelerated evolution allows us to address the mechanisms of rapid evolution and speciation, which in our study appears to be a combination of both selection and rapid demographical changes.

1. Introduction

The tempo and mode of evolutionary change during speciation have remained a contentious issue for more than five decades. Evidence for abrupt and rapid changes during speciation came from fossil evidence but lacked mechanistic explanations for such a process [1–7]. Since the mid-1970s, molecular phylogenetic studies began to associate increased genetic changes with speciation events suggesting that rates of molecular evolution might be altered during speciation [8–11]. This trend has been recently confirmed by molecular phylogenies using large numbers of genes and a wide variety of taxa, explicitly showing that speciation is accompanied by accelerated rates of molecular evolution [12–15].

What remains unknown is the mode of change during speciation, that is, the kind of genetic changes associated with speciation. More specifically, the numbers and kinds of protein coding genes that change during speciation remain unexplored. Accelerated evolution is only observed in a fraction of genes analyzed by Pagel et al. [13] as well as in other genome-wide estimates of molecular divergences [16–18]. A more systematic search to identify protein-coding genes that experience accelerated evolution during speciation will allow us to directly address the mode of evolutionary change during speciation, that is, divergences, in what kinds of genes, affect or are affected by speciation.

Taxa-wide evidence of sex- and reproduction-related genes evolving rapidly in sibling species [19, 20] has raised the possibility that sex-related genes might preferentially experience elevated rates of evolution during the speciation process, or may even drive speciation by causing reproductive isolation among diverging populations. This framework is supported by the fact that almost all candidate "speciation" genes identified, so far, are mainly sex related (with the exception of a few genes with other functions), evolve rapidly between closely related species, often show signatures of

adaptive evolution and have been invoked in the rapid evolution of hybrid sterility in different organisms [21–32]. After controlling for incomplete lineage sorting in the melanogaster subgroup, we put this framework to test by analyzing 4843 protein coding genes in the Drosophila transcriptome for signatures of speciation-related accelerated changes.

Divergence trends between recently formed species and old species provide a proxy to detecting signatures of speciation-related accelerated changes. Molecular divergences are expected to be proportional to the duration of a species' existence; newly formed species will have accumulated less molecular divergence compared to older species. Signatures of speciation-related accelerated evolution will be manifested as relatively higher rates of molecular evolution in newly formed species compared to species with longer postspeciation divergence times. Accordingly we asked the following questions: (1) are rates of molecular evolution in protein-coding genes affected by speciation events? That is, do genes show unexpectedly elevated rates of evolution in newly formed species relative to older species? (2) do sex-related genes preferentially exhibit accelerated speciation-related changes relative to non-sex related genes? And (3) do genes with accelerated rates of evolution show evidence of positive selection?

2. Methods

2.1. Rationale. If evolutionary rates of protein coding genes are not affected by speciation and evolve at a constant rate, we should be able to find a correlation between the length of time that species have diverged and the proportion of molecular divergence between these species. Species diverged for longer lengths of time would have accumulated proportionally higher amounts of genetic changes relative to species diverged for shorter periods of time. Conversely, if rates of molecular evolution were indeed affected by speciation, then this correlation will be broken and newer species might show relatively higher evolutionary rates compared to older species. Our analyses therefore primarily exploit the nature of molecular changes *since* divergence of species pairs (recent versus older) to infer speciation-related accelerated changes.

2.2. The Phyletic System. We used a phyletic system comprised of three pairs of closely related species from the melanogaster subgroup *D. simulans* and *D. sechellia* diverged about 0.3–0.6 Mya [33], *D. melanogaster* and *D. simulans* diverged about 4.3–6.5 Mya, and *D. yakuba* and *D. erecta* diverged about 8.1–12.7 Mya [34]. Divergence times for *D. simulans-D. melanogaster* and *D. yakuba-D. erecta* were recently reestimated by using over 2977 nuclear genes and by implementing a novel genomic-mutation distance approach correcting for codon bias [34] whereas the divergence times for the *D. simulans-D. sechellia* split were estimated using a small number of genes [33]. All of these divergence dates broadly concur with other independent estimates using mitochondrial loci (reviewed in Powell 1997), and there is also some general concordance with what little paleontological evidence is available for Drosophila (see [34]). We also used *D. pseudoobscura* and *D. persimilis* from the *D. obscura* group diverged for about 0.85 Mya [34]. These two species belong to the *Obscura* group that diverged 55 Mya from the melanogaster group [34], which brings about problems of saturation in d_S and gene expression differences [35]. Nevertheless, this species pair diverged for roughly the same amount of time as *D. simulans-D. sechellia*, a comparison of rates of protein evolution between the two species pairs will be useful to determine if acceleration in rates of evolution is commonly found among newly derived species.

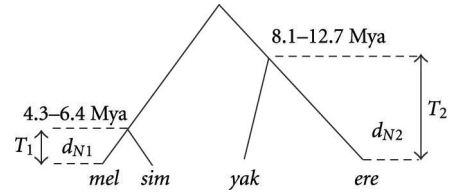


Figure 1: Schematic representation of rational implemented in our methodology. d_{N1}: proportion of nonsynonymous substitutions in the more recently diverged species pair (diverged for time T_1). d_{N2}: proportion of nonsynonymous substitutions in species that have diverged for longer periods of time (T_2).

2.3. Estimating Differences in Rates of d_N in relation to Speciation. According to the divergence times recently reported by Tamura et al. [34], *D. erecta* and *D. yakuba* have diverged for an estimated length of time that is ~2-3 times greater than the divergence time between *D. melanogaster* and *D. simulans*. Given the neutrally expected linear relationship between genetic divergence and time, we should therefore expect ~2-3 times greater divergence in genes between *D. yakuba* and *D. erecta* relative to *D. melanogaster* and *D. simulans*. We calculated the expected rate of molecular divergence (rate of nonsynonymous mutations per nonsynonymous sites d_N and rate of synonymous mutations per synonymous sites d_S) using the relationship given below. Since our interest ultimately was to determine differences in rates of protein evolution between the different species pairs we focused more on d_N. The ratio of nonsynonymous divergence (d_N) between two-species pairs, a recently diverged species pair 1 and an older species pair 2, must be proportional to the ratio of their divergence times (T), such that:

$$\log_2\left(\frac{T_1}{T_2}\right) = \log_2\left(\frac{d_{N1}}{d_{N2}}\right), \quad (1)$$

where, for a given species pair, T_1 = divergence time of the more recently diverged pair and T_2 = divergence time of the species with longer post-speciation divergence time. In our data, for example, T_1 = 4.3–6.5 Mya (mel-sim) and T_2 = 8.1–12.7 Mya (yak-ere), or, T_1 = 0.4–0.6 Mya (sim-sec) and T_2 = 8.1–12.7 Mya (yak-ere) (see Figure 1). We similarly calculated the ratio of synonymous divergence to species divergence time in all species pairs. Because of existing uncertainties in the divergence time estimates, particularly for the newer species [36, 37], we worked with

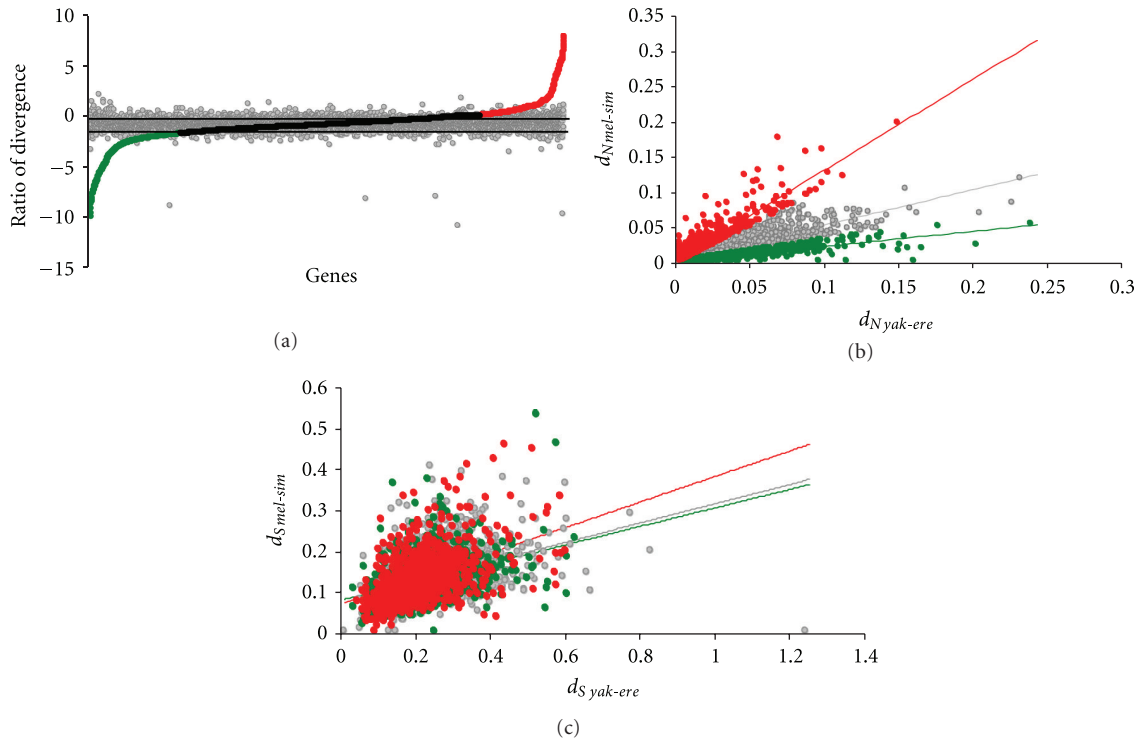


Figure 2: Distribution of ratio of divergence between *D. melanogaster-D. simulans* versus *D. yakuba-D. erecta*. (a) Distribution of molecular-divergence $\log_2$ (d_{N1}/d_{N2}) to divergence-time ratio $\log_2$ (T_1/T_2), between *D. melanogaster-D. simulans* and *D. yakuba-D. erecta*. Colored circles represent d_N ratios, and gray circles represent d_S ratios. The upper limit of the clockwise category is given by the $\log_2$ ratio of divergence times (6.5/8.1) and the lower limit by the $\log_2$ ratio 4.3/12.7. (b) Relationship of d_N between *D. melanogaster-D. simulans* and *D. yakuba-D. erecta*. The regression lines for each rate categories are also plotted. (c) Relationship of d_S between *D. melanogaster-D. simulans* versus *D. yakuba-D. erecta*; color codes: red: accelerated, black: clock-like, green: slow-rate categories, and grey: d_S ratios.

the estimated *range of divergence times* for each of the species pairs by incorporating the published upper and lower limits of divergence time estimates from Tamura et al. [34]. For instance, in comparisons between *D. melanogaster-D. simulans* versus *D. yakuba-D. erecta* pairs, the upper divergence-time limit represents a scenario where *D. melanogaster* and *D. simulans* diverged 4.3 Mya and *D. yakuba* and *D. erecta* diverged 12.7 Mya. Similarly, the lower divergence time limit represents a scenario where *D. melanogaster* and *D. simulans* diverged 6.5 Mya, and *D. yakuba* and *D. erecta* diverged 8.1 Mya. When $\log_2$ (d_{N1}/d_{N2}) is plotted against $\log_2$ (T_1/T_2), genes evolving in a clock-like manner (i.e., $d_{N1} = [d_{N2} * (T_1/T_2)]$) would fall within these divergence time-ranges. Genes that fall above the divergence time-ranges indicate accelerated evolution in the newer species and those falling below the lower limit are considered to evolve slowly in the newly formed species or to have much higher rates of evolution in the older species lineage.

Comparisons using the *D. simulans-D. sechellia* species pair grossly overestimated the number of genes in the accelerated rate category (94.89% of the genes fall under the accelerated category Figure 3(a)). This is an unlikely scenario given that the range of d_S falls well above the divergence times plotted in this graph. We believe that this is most likely due to a gross underestimation of the divergence time between *D. simulans* and *D. sechellia*. The phylogenetic relationships in the *simulans* triad (*D. simulans, D. mauritiana,* and *D. sechellia*) have been contentious and unresolved [35] and the sole source of recent divergence times estimates using molecular data comes from Kliman et al.'s study [33], which used a small number of genes. We took a very generalized approach to reevaluate the divergence time for the *D. simulans-D. sechellia* split using d_S estimated from our dataset (Figure 2) in order to replot Figure 3(a). Relative to nonsynonymous divergence, we expect the ratio of synonymous divergence-to-species divergence times for most loci to fall within the clock-like category for most species comparisons (as it did in the previous comparison). We therefore applied the d_S boundaries conforming to the clock-like category from *D. melanogaster-D. simulans* versus *D. yakuba-D. erecta* comparisons (Figure 2) to the *D. simulans-D. sechellia* species pair. We were able to arrive at a generalized estimate of 1.16–3.05 Mya for the *D. simulans-D. sechellia* split which appears more likely (Kumar 2007, pers comm.). We therefore employed a slightly more stringent approach by applying 95% confidence intervals to the existing divergence boundaries (using d_S ratios), which consequently incorporated previous outliers (Table 1).

2.4. Sequences and Rate Analyses. All sequences (*D. melanogaster, D. simulans, D. sechellia, D. yakuba, D. erecta, D. pseudoobscura*, and *D. persimilis*) were obtained from the

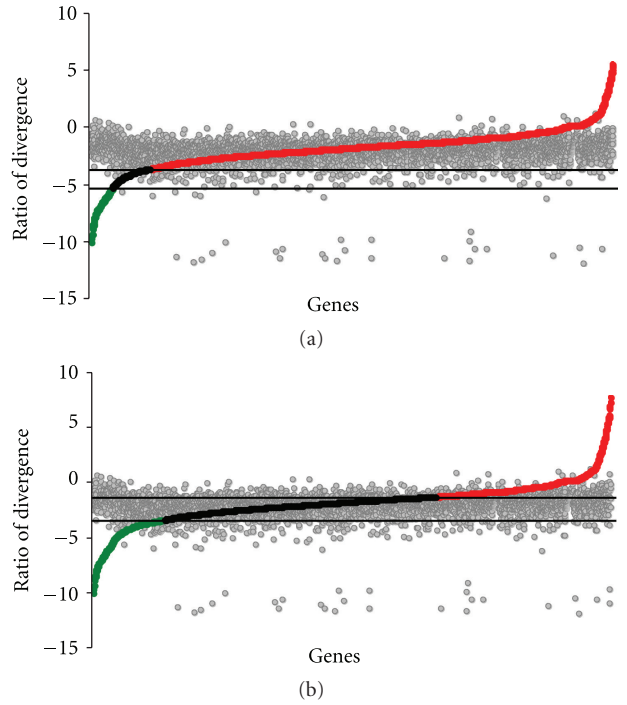


Figure 3: Distribution of ratio of divergence between *D. simulans-D. sechellia and D. yakuba-D. erecta*. (a) Distribution of ratio of amino acid replacement substitutions in separate clades $\log_2 (d_{N1}/d_{N2})$ to divergence time (according to Kliman et al. [33]) ratio $\log_2 (T_1/T_2)$, *D. simulans-D. sechellia* and *D. yakuba-D. erecta*. (b) Re-evaluation of divergence times between these species. d_N ratios are plotted in colors, and d_S ratios are plotted in grey. Color codes: red circles accelerated evolution, black: clock-like, green: slow evolution categories, grey: d_S ratios.

recently sequenced genomes available on FlyBase (*Drosophila* 12 Genomes consortium 2007, http://www.flybase.org/). Sequences were aligned according to the corresponding protein alignment using CLUSTALW ver. 1.8 [38]. In order to remove any potential bias due to incomplete lineage sorting effects [39, 40], for each gene we compared the likelihood of trees differing in the placement of *D. yakuba* and *D. erecta* using PAML and restricted our analysis to genes for which the best tree involved *D. yakuba* and *D. erecta* as sister species. Pairwise estimates of d_N and d_S were determined using the program codeml in PAML [41]. Estimates of d_N, d_S, and ω along each lineage using branch site models and outputs of models M7 (neutral) versus M8 (positive selection) were also determined using PAML [42] and retrieved from a recent genome-wide analysis [43]. We were able to compute estimates of divergence to time ratio (see above) for 4843 orthologs between *D. melanogaster-D. simulans* versus and *D. yakuba-D. erecta*, and 4327 orthologs between *D. simulans-D. sechellia* versus *D. yakuba - D. erecta* comparisons as well as for 3988 genes from *D. pseudoobscura-D. persimilis* versus all melanogaster species pairs.

2.5. Classification of Genes according to Site of Expression. We classified genes according to their tissues of expression (testis, ovary, and head) by using the NCBI EST database (NCBI, http://www.ncbi.nlm.nih.gov/UniGene/). Genes that could not be classified into any tissue category were referred as unspecified. Based on EST data, we were able to classify tissue of expression for 3040 out of 4843 genes for comparisons involving *D. melanogaster-D. simulans* versus *D. yakuba-D. erecta*, 2686 genes for comparisons involving *D. simulans-D. sechellia* versus *D. yakuba-D. erecta* (Table S1).

3. Results

3.1. Accelerated Rates of Molecular Evolution in Newly Formed Species. Molecular divergence estimates, with respect to species divergence times, showed that protein coding genes fell into three distinct rate categories: (1) *accelerated evolution in younger species:* genes showing higher rates of molecular divergence in newly formed species relative to species diverged for longer lengths of time. (2) *clock-like evolution*: genes showing molecular divergence that corresponds to species divergence times, and (3) *slow evolution in younger species*: genes showing lower rates of molecular divergence in newly formed species in comparison to species with longer divergence times (Figure 2). Table 1 summarizes the fraction of genes that fall under each rate category for every species pair compared.

Most genes in our dataset (61–74%) fell into the clock-like rate category in all comparisons, indicating that evolutionary trajectories of most genes are unaffected by speciation events and their rates remained constant (Table 1). A small but discernable fraction of genes (17–19%) showed signatures of speciation-related accelerated evolution. Nonsynonymous divergences in these genes were 2-3-fold higher in newly formed species compared to older species (Table 1, Figures 2 and 3). A plot of d_N and d_S estimates between *D. melanogaster-D. simulans* versus *D. yakuba-D. erecta* shows that the distribution of d_N estimates is quite distinct between the accelerated, clock-like, and slow-rate categories while the distribution of d_S estimates is not (Figures 2(b) and 2(c)). This quite clearly indicates that elevated proportions of d_N is not always accompanied by correspondingly elevated proportions of d_S in genes with evidence of accelerated evolution, which is generally a sign of selection driven changes. We also obtained similar results using protein divergence estimates instead of the ratios of nucleotide divergence data See files in Supplementery Material available online at doi: 10.4061/2011/595121 . About 8–19% of protein coding genes that fell into the slow rate category showed extremely low levels of nonsynonymous divergence in newly formed species compared to older species (e.g., Avg $d_{N\,mel\text{-}sim} = 0.0057 \pm 0.007$, Avg $d_{N\,yak\text{-}ere} = 0.2276 \pm 0.081$, Figures 2, 4, and 5). This may be indicative of genes that have remained conserved in the evolution of the recently diverged species but that have diverged substantially with time in the older lineages.

These results quite unambiguously indicate two important trends: firstly, despite variances associated with divergence estimates, the 2-3-fold higher nonsynonymous divergence in newly formed species compared to species with

Table 1: Number and (percent) of genes falling under accelerated, clock-like, and slow-rate categories in recently diverged species when compared to older species pairs. (a) Analysis using the published divergence boundaries according to Kliman et al. [33] for *D. simulans* and *D. sechellia*, and Tamura et al. [34] for *D. melanogaster-D. simulans* and *D. yakuba-D. erecta*. (b) Analysis using the revised divergence time estimates for *D. simulans* and *D. sechellia*. (c) Analysis applying the 95% confidence interval boundaries on revised divergence estimates.

	Rate Categories		
	Slow	Clock-like	Accelerated
(a)			
mel-sim versus *yak-ere*			
d_N	1096 (22.63)	2251 (46.48)	1496 (30.89)
d_S	348 (7.19)	3628 (74.91)	867 (17.90)
sim-sec versus *yak-ere*			
d_N	173 (3.99)	314 (7.26)	3840 (88.75)
d_S	37 (0.86)	184 (4.25)	4106 (94.89)
(b)			
sim-sec versus *yak-ere*			
d_N	594 (13.73)	2275 (52.58)	1458 (33.69)
d_S	311 (7.19)	3241 (74.90)	775 (17.91)
(c)			
mel-sim versus *yak-ere*			
d_N	963 (19.89)	2996 (61.86)	884 (18.25)
d_S	242 (5.00)	4359 (90.00)	242 (5.00)
sim-sec versus *yak-ere*			
d_N	469 (10.84)	3007 (69.49)	851 (19.67)
d_S	216 (5.00)	3895 (90.00)	216 (5.00)
pse-per versus *sim-sec*			
d_N	327 (8.20)	2962 (74.27)	699 (17.53)
d_S	199 (5)	3590 (90)	199 (5)

much longer divergence times clearly represents increases in rates of protein evolution either during or immediately after speciation. Secondly, accelerated rates of molecular evolution are most apparent in nonsynonymous divergence and not in synonymous divergence (see Table 1, Figures 2 and 4), a broadly accepted sign of selection driven changes.

3.2. Higher Representation of Sex-Related Genes in the Accelerated and Clock-Like Rate Categories. Identifying the range and, particularly, the kinds of loci in each rate category (accelerated, clock-like and slow) provides a starting point to broadly address the effects of demographic factors and selection during speciation. Demographic factors (drift, bottlenecks etc.) would affect a wide variety of loci whereas selection driven divergence would only affect specific functional classes of genes, such as sex-related genes which are expected to drive reproductive isolation in diverging populations [45, 46]. Due to lack of functional information for a large number of *Drosophila* genes, we used tissue of expression as a general and presumable indication of function (testis and ovary as reproductive tissues versus head, a presumably nonreproductive tissue). We tested the null hypothesis that genes expressed in each tissue-type are equally distributed within the accelerated, clock-like and slow rate categories. Because the expression data was determined using *D. melanogaster* data, only species from the *D. melanogaster* subgroup were analyzed. Given the long divergence time between the *Obscura* group and the melanogaster group (55 Mya), it is likely that patterns of gene expression may be radically different in *D. pseudoobscura* and *D. persimilis*.

In the *D. melanogaster-D. simulans* versus *D. yakuba-D. erecta* comparison, a significantly higher proportion of testis-specific genes occupy the accelerated and clock-like rate categories compared to the slow-rate category (χ^2 testis specific = 8.08 and 6.78, df = 1, and $P = 0.0134$ and 0.0277, respectively, and a Bonferroni correction was applied, Table 2, [44]). Genes expressed in all three tissues—testis, ovary and head mostly fall in the accelerated category compared to both slow and clock-like rate categories (χ^2 testis, ovary-and-head, = 5.96 df = 1, $P = 0.0439$, $\chi^2 = 19.99$, df = 1, $P = 2.34 \times 10^{-5}$, Bonferroni corrections were applied [44], Table 2). No significant differences were observed in the proportions of all other genes classes between rate categories.

In the *D. simulans-D. sechellia/D. yakuba-D.* erecta comparison, a significantly small proportion of testis-specific genes fall into the accelerated category compared to clock-like and slow rate categories ($\chi^2 = 10.04$ and 13.91, df = 1, $P = 4.59 \times 10^{-3}$ and 5.75×10^{-4}, a Bonferroni correction was applied, Table 2).

3.3. "Persistence" of Rapid Evolution in Testis-Specific Genes over Time. That most testis-specific genes fall under the clock like rate category and only a small proportion fall in the

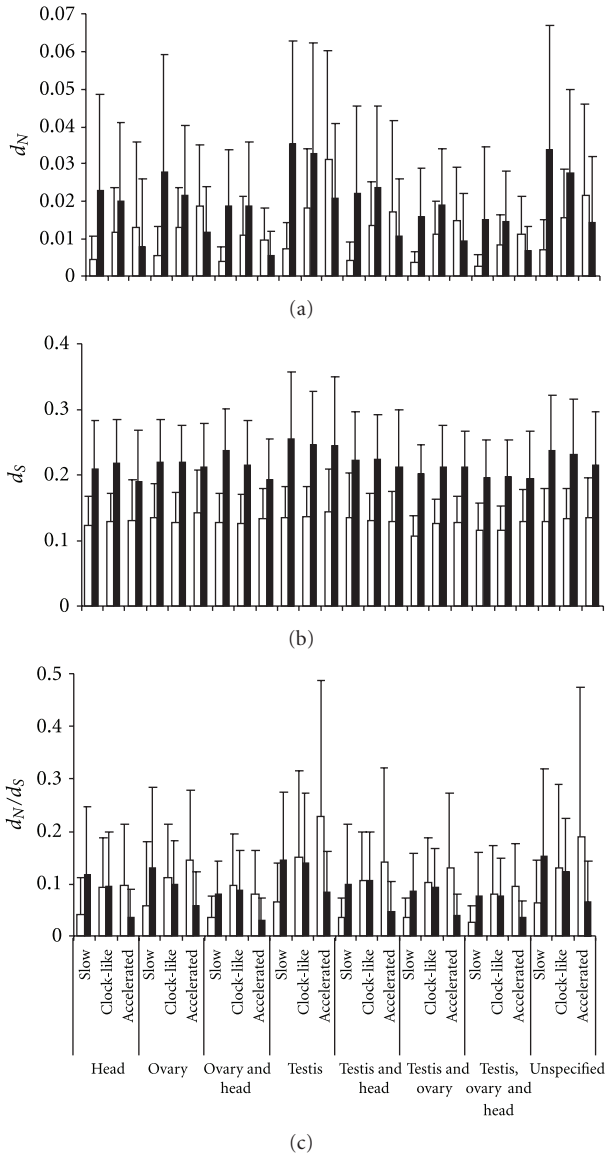


Figure 4: Relationship between rate category and tissue of expression. Distribution of (a) d_N, (b) d_S, and (c) d_N/d_S between rate categories (accelerated, neutral, and slow) and between tissue classifications. White bars: *D. melanogaster-D. simulans*, and black bars: *D. yakuba-D. erecta*. The error bars represent the standard deviation.

accelerated rate category would appear to contradict earlier evidence that testis specific genes are in general evolving rapidly [20, 47, 48]. We hypothesized that a "persistence" of rapid rates of evolution in testis specific genes (rapid rates of evolution over time) will explain why testis-specific genes largely occupy the clock-like rate category. Pagel et al. [13] also found such persisting elevated rates of evolution in many lineages. Persistence of rapid rates of evolution can be verified if testis specific genes in both younger and older lineages have high substitution rates in both the accelerated and clock-like rate categories compared to other genes. To test this, we first looked for a global difference in d_N and d_S between tissue categories as well as between rate categories (Figure 5). We also performed more detailed analyses of differences in d_N and d_S between each tissue category using a Tuckey HSD test. Invariably, testis-specific genes show significantly higher d_N and d_N/d_S compared to genes in all other tissues in both younger and older species pairs ($P < 0.05$ for both, Figures 4(a), 4(b), and 4(c)) (additional file 3). This implies that testis genes in the clock-like category are evolving faster than other genes but at a constant rate, whereas those in the accelerated rate category have experienced elevated rates of evolution in response to stronger selection during or after speciation. This is an important trend as it reveals a tempo of molecular evolution in the different classes of genes.

3.4. Genes Evolving under Positive Selection in the Accelerated, Clock-Like, and Slow Rate Categories. We compared the proportion of genes showing evidence of positive selection ($\omega > 1.0$) among the different rate categories by implementing site models, branch-site and branch models in PAML [42]. Applying models M7 versus M8, none of the rate categories showed overrepresentation of genes evolving positively (slow versus accelerated, $\chi^2 = 0.88$, df = 1, $P > 0.05$, accelerated, slow versus clock-like $\chi^2 = 0.93$ and 2.46, df = 1, $P > 0.05$ and 0.35 after Bonferroni corrections, Table 3). Using the branch-site model however, we observed a significant over-representation of genes showing positive selection in *D. simulans* in the accelerated rate category compared to the clock-like and slow rate categories ($\chi^2 = 20.82$ and 28.6, df = 1, $P = 1.51 \times 10^{-5}$ and 2.67×10^{-7}, respectively after Bonferroni correction). Branch model tests also show a greater proportion of genes with foreground $\omega >$ background ω in the accelerated rate category in the *D. simulans* branch and the branch leading to the *D. melanogaster* clade (Table 3). A list of genes detected to be evolving under positive selection in each lineage can be found in additional file 4. Several genes involved in sensory stimuli, immune response, gametogenesis (spermatogenesis), transcription regulation, and hybrid incompatibilities (including *Hmr*, [21] are amongst the genes that show relatively large ω estimates. Population genomic study of six *D. simulans* strains compared to *D. melanogaster* [49] found significant evidence of directional selection in genes affecting reproduction or spermatogenesis. Among the 1270 genes that show evidence of adaptive evolution from Singh [46], 505 are found in our comparison of *D. melanogaster-D. simulans/D. yakuba-D. erecta*. Among the 505 genes, 360 fall in the clock-like category, 108 in the accelerated category and 37 fall in the slow rate category. We detect a significant enrichment of genes under positive selection among the clock-like category $\chi^2 = 7.25$, df = 1, $P = 7.09 \times 10^{-3}$) while we found a significant paucity among the slow rate category ($\chi^2 = 40.05$, df = 1, $P = 2.48 \times 10^{-10}$). No significant effect is observed for the genes classified as rapidly evolving in the younger species ($\chi^2 = 2.72$, df = 1, $P = 0.099$).

3.5. Effect of Local Recombination Rates. Gene evolution may be influenced by their chromosomal location [50, 51] as well

Table 2: Number of genes present in each rate category when comparing *D. melanogaster-D. simulans* to *D. yakuba-D. erecta* and *D. simulans-D. sechellia* and to *D. yakuba-D. erecta*. Under brackets are the proportion of genes within tissue and the proportion within the rate category. Enrichment within a category was tested using χ^2, test and the Bonferroni corrections were applied [44].

Tissue	Accelerated	Clock-like	Slow	Total
D. melanogaster-D. simulans versus *D. yakuba-D. erecta*				
H	129 (17.18); (14.59)	445 (60.59); (14.85)	177 (22.63); (18.38)	**751 (15.51)**
O	47 (20.26); (5.32)	147 (63.79); (4.91)	38 (15.95); (9.95)	**232 (4.79)**
OH	35 (15.70); (3.96)	138 (63.68); (4.61)	50 (20.62); (5.19)	**223 (4.6)**
T	142 (20.67); (16.06)	440 (65.50); (14.69)	105 (14.70); (10.90)	**687 (14.19)**
TH	114 (20.36); (12.90)	357 (64.46); (11.92)	89 (15.72); (9.24)	**560 (11.56)**
TO	32 (20.65); (3.62)	99 (64.52); (3.30);	24 (14.84); (2.49)	**155 (3.2)**
TOH	116 (21.17); (13.12)	231 (55.97); (7.71)	80 (17.57); (8.31)	**427 (8.82)**
un	269 (14.88); (30.43)	1139 (64.66); (38.02)	400 (20.79); (41.54)	**1808 (37.33)**
Total	**884 (18.25)**	**2996 (61.86)**	**963 (19.89)**	**4843**
D. simulans-D. sechellia versus D. yakuba-D. erecta				
H	149 (21.94); (17.51)	458 (67.45); (15.23)	72 (10.6); (15.35)	**679 (15.69)**
O	32 (16.24); (3.76)	146 (74.11); (4.86)	19 (9.64); (4.05)	**197 (4.55)**
OH	48 (32.65); (5.64)	130 (65.99); (4.32)	19 (9.64); (4.05)	**197 (4.55)**
T	101 (16.13); (11.87)	427 (68.21); (14.20)	98 (15.65); (20.90)	**626 (14.47)**
TH	101 (19.65); (11.87)	363 (70.62); (12.07)	50 (9.73); (10.67)	**514 (11.88)**
TO	26 (19.55); (3.06)	99 (74.44); (3.29)	8 (6.02); (1.71)	**133 (3.07)**
TOH	77 (22.65); (9.05)	217 (63.82); (7.22)	46 (13.53); (9.81)	**340 (7.86)**
un	317 (19.32); (37.25)	1167 (71.11); (38.81)	157 (9.77); (33.48)	**1641 (37.92)**
Total	**851 (19.67)**	**3007 (69.49)**	**469 (10.84)**	**4327**

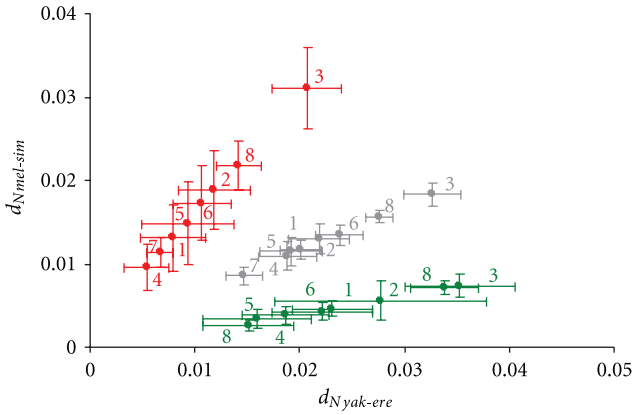


Figure 5: Relationship between evolutionary rates and tissue of expression. Relationship between $d_{N\ D.melanogaster\text{-}D.\ simulans}$ and $d_{N\ D.\ yakuba\text{-}D.\ erecta}$ for gene classifications in each rate category: red: accelerated, grey: clock-like, green: slow. Tissue classifications: (1) head-specific; (2) ovary-specific; (3) testis-specific; (4) ovary and head; (5) ovary and testis; (6) testis and head; (7) Testis; ovary and head; (8) unknown. Error bars represent 95% confidence intervals.

as by local recombination rates [43]. Using recombination rates in *D. melanogaster* computed by Hey and Kliman [52], we found significantly higher average recombination rates for genes in the accelerated rate category compared to the clock-like and slow categories in the *D. melanogaster-D. simulans/D. yakuba-D. erecta* comparison (Kruskal Wallis rank sum test, $P < 2.2 \times 10^{-16}$ and $P = 0.0234$, respectively, a Bonferroni correction was applied). There was no significant difference between the clock-like and slow rate categories (Kruskal Wallis rank sum test, $P = 0.323$, a Bonferroni correction was applied). The *D. simulans-D. sechellia/D. yakuba-D. erecta* comparison showed no such effect of recombination in all rate categories (Kruskal Wallis rank sum test, $P = 0.67$, 0.30 and for clock-like versus accelerated, accelerated versus slow, clock-like versus slow, respectively after Bonferroni correction).

4. Discussion

4.1. What Kinds of Genes Change during Speciation?. Despite recent evidence linking accelerated rates of molecular evolution to speciation events [12, 13], the kinds of protein coding genes that might experience accelerated rates of evolution during or immediately after speciation require investigation [15]. Are rates of evolution in all genes likely to be altered during or even after speciation? If not, what genes do show speciation associated changes? Answers to these questions will be crucial to understanding the mechanism(s) driving speciation and the molecular evolutionary consequences of speciation.

This study analyzes the ratio of molecular evolutionary rates to the ratio of species divergence times, therefore clock-like evolution should not be taken as a sign of gradual evolution; it only implies a constant rate of evolution. This is exemplified by the large representation of testis-specific genes in the clock-like rate category, which are in fact

evolving much faster than other genes in the same category. Furthermore, *Drosophila* is one of those groups that lack detailed fossil records and incomplete taxon sampling is an obvious but unavoidable problem in this, or any study of this nature. Apart from extinct taxa, undiscovered and undescribed taxa are an additional problem. We are also lacking the genome sequence of *D. mauritiana*, a member of the *D. simulans-D. sechellia-D. mauritiana* triad as well as the genome of *D. santomea* which recently split from *D. yakuba* [53]. These factors are bound to affect true assessments of speciation-related changes in rates of molecular evolution. However, within the currently accepted phylogenetic network of the melanogaster subgroup, our study provides an approach to identify genes that change during speciation and out results report the fraction and identity of genes in the *Drosophila* transcriptome that show signatures of accelerated speciation-related changes that can be further investigated.

4.2. Evidence for Accelerated Evolution. In our study, evidence of speciation-related accelerated evolution is detected in a small but discernable fraction of protein coding genes. This corresponds to what Pagel et al. [13] found across taxa. But what is striking is that the acceleration is most apparent in nonsynonymous divergence. That d_S in most genes fall under the clock-like category as opposed to d_N is not entirely surprising, as we expect d_N changes to be more sensitive to selection and may therefore show higher variance. Nevertheless, the 2-3 fold higher estimates of d_N in the recently diverged species compared to species with much longer divergence times is a strict deviation from clock-like evolution despite the variance that may be involved. More importantly, our survey reveals that the more recently diverged species (*D. simulans-D. sechellia*) have a slightly greater proportion of genes that show accelerated rates of evolution (Table 1), further strengthening the case for a causal link between acceleration in rates of molecular evolution and speciation.

4.3. Evidence for Persistent Rapid Evolution in Sex Genes. Sexually reproducing organisms are influenced both by natural and sexual selection. Widespread rapid evolution of sex-related genes in *Drosophila* genomes [18] imply that regardless of how speciation occurred, sex and reproduction related genes that are causally involved in establishing reproductive isolation would be constantly under strong selection [46]. Therefore we would not expect to find an over-representation of sex-related genes in the accelerated rate category alone. This is specifically illustrated by the persistence of rapid rates of evolution in testis specific genes in our study. In addition, testis-specific genes and genes expressed in all three tissues are over-represented in the accelerated and clock-like rate category in the *D. melanogaster-D. simulans* pair (Table 2). But in the more recently diverged *D. simulans-D. sechellia* pair, only testis-specific genes are overrepresented in the clock-like and slow rate categories (Table 2). These results support a scenario where sex-related genes are under constant but higher selective pressure.

Table 3: Comparison of the proportion of genes showing evidence of positive selection under the site model, branch-site model, and branch model between slow ($n = 963$), clock-like ($n = 2996$), and accelerated ($n = 884$) categories. *P* values for the different tests were corrected using a Bonferroni correction prior to the comparisons. *mel: D. melanogaster, sim: D. simulans, sec: D. sechellia, simsecmel*: branch leading to the *D. melanogaster clade, yak: D. yakuba, ere: D. erecta, and yakere*: branch leading to the *D. yakuba* clade.

Model	Slow	Clock-like	Accelerated
Site model			
M7/M8	6	40	9
Branch-site model			
mel	0	8	0
sim	5	51	39
sec	10	29	9
simsecmel	3	6	3
yakere	3	4	0
Branch model			
mel	0	6	6
sim	0	3	24
sec	1	13	3
simsecmel	1	59	80
yak	20	20	0
ere	36	26	0
yakere	68	83	0

4.4. Factors Driving Accelerated Evolution during Speciation. We find no evidence of widespread positive selection in the accelerated rate category. The higher proportion of nonsynonymous divergence we observe in the accelerated rate category cannot be ruled out due to relaxed constraint or accumulation of slightly deleterious mutations [16, 54, 55]. The only tenable links to acceleration that we find appear to be influenced by local recombination rates (higher in the *D. melanogaster-D. simulans* comparison but lower in the *D. simulans-D. sechellia* comparison) and a relatively small proportion of genes with lineage specific positive selection (Table 3). Variation in recombination rates between species can be an important force driving divergence, as new allele combinations can be produced at different rates within species [56]. Rate of recombination also appears to be positively correlated with levels of polymorphism and high polymorphism would be expected to correlate with levels of divergence [57]. Nonetheless, sex-related genes involved in gametogenesis, hybrid incompatibilities as well as genes involved in metabolism and sensory functions do show evidence of positive selection. This implies a loci-specific role of sex genes and sexual selection in speciation along with some ecological specialization. Our data therefore indicates that acceleration in rates of evolution is not purely a consequence of strong selection but most likely a combination of factors, such as demographic processes as well as some form of selection (a few genes in the accelerated category are evolving positively, Table 3). These data are interesting particularly in *D. sechellia*, which is a relatively young species and one that is likely to have undergone founder speciation (see [33]).

According to the founder effect model of speciation [3], speciation occurs as a consequence of major demographic processes in which bottlenecks play an important role. In such a case, a large number of loci involved in a wide range of functions would be affected (but see [58] for a more recent critical analysis of bottlenecks and speciation). Our results as well as those of Pagel et al. [13] do not support the notion of a genetic revolution as a consequence of speciation; the range of loci affected during speciation is rather limited. However, loci affecting a wide range of functions show accelerated rates of evolution (Figures 1, 2, and 3, Tables 1 and 2). This supports a scenario where, as founding populations adapt to new ecological niches the evolution of modified or new behaviors affecting sex and reproduction as well as other ecological adaptations can occur rapidly [59–62]. *D. sechellia* has evolved an intricate ecological relationship with its plant host *Morinda citrifolia*, on which *D. sechellia* females oviposit their eggs [63]. Compounds in the pulp of the immature *M. citrifolia* fruit are toxic to other species of the melanogaster subgroup but not to *D. sechellia* larve that feed and grow on it [63]. A large proportion of head-expressed genes show accelerated evolution in the *D. simulans-D. sechella* pair, which might reflect rapid behavioral and sensory modifications that occurred during the host-plant specialization between *D. sechellia* and *M. citrifolia*. Genes involved in sensory perception (*Gr2a, Gr21a, Gr43a, Obp50a, Or67d, king-tubby, CG32683*), sensory organ development (*amos, Brd, mib1, Oseg1, Poxn*), detoxification (*kraken*), metabolism, and oogenesis (*cup, retn, kel, spir, Tm1*) found to be involved in host specialization by Dworkin and Jones [64] are among those that show evidence of accelerated speciation-related rates of evolution in our study.

The *D. sechellia* data demonstrates the importance of founder-effect and subsequent ecological divergence in driving speciation and as a consequence, causing elevated rates of evolution of relevant protein coding genes associated with the speciation event. Such cases of speciation driven by founder-effect and subsequent resource specialization are not uncommon in Drosophila [45, 59, 65]. *D. sechellia, D. mauritiana* and *D. santomea* (a recently described sibling species to *D. yakuba*) are all insular species within the melanogaster subgroup. Among these, *D. sechellia and D. santomea* show resource specialization having evolved special ecological relationships of utilizing specific host plants either for food or oviposition [53, 63–65], while their parent lineages do not. Such species-specific ecological adaptations are likely to influence accelerated evolution in the newly formed species but not necessarily in the parent lineages. Our study support this reasoning; we find ∼400 genes in the accelerated rate category in the *D. simulans* versus *D. sechellia* comparison; all of which fall in the clock-like rate category in the *D. melanogaster* versus *D. simulans* comparison (see additional file 5). Many of these genes, as mentioned above, have been found to be involved in host-plant specialization [64]. Further detailed study of such genes between *D. sechellia* and *D. simulans*, as well as the inclusion of *D. santomea* genome (when completed) in a similar study should shed some light on our claim. However, it should be noted that while *D. sechellia* is isolated from *D. simulans, D. santomea* and its sibling species *D. yakuba* are sympatric on the Island of San Tome and there is substantial divergence of sexual traits between the two species [53, 66–69]. We might therefore expect to see higher proportions of sex-related genes as well as genes involved in ecological specializations to show signatures of accelerated evolution between *D. santomea* and *D. yakuba*, much like what we observed between the sympatric *D. melanogaster* and *D. simulans* in our study.

5. Conclusion

Signatures of speciation-related accelerated rates of evolution are detected in all newly evolved species analyzed in this study. This is not observed as a widespread genomic signature but is restricted to a fraction of the genome that affects widely different functions. The range and kinds of loci in this rapidly evolving fraction identified in this study complements the recently reported punctuational effect observed across a variety of taxa [13, 15] and improves our focus of studying the genetics of speciation. Population genetic studies of these 'candidate' genes can establish the nature of selective forces driving their elevated rates of evolution. In our case, we find a demography-selection driven effect. Testis specific genes and sex-related genes show persistently high rates of evolution, indicating that sexual selection is a constant pressure in diverging and established populations. In the *D. simulans-D. sechellia* pair, our data reinforce previous evidence that founder effect and ecological specialization played important roles in their speciation process. Our results support the growing appreciation that speciation is often driven by a combination of demographic fluctuations, ecological adaptation, and sexual selection and can seldom be attributed to one single important factor [45, 70].

Acknowledgments

The authors are grateful to the AAA group for making the genomes data available to them. Richard Morton, Alberto Civetta, Sudhir Kumar, Harilaos Lessios, and two anonymous reviewers provided useful and critical comments that greatly improved this manuscript. The authors would also like to thank Ben Evans, Jonathan Stone, and Osamu Miura for very useful discussions, comments, and criticisms on previous versions of this manuscript. This study was supported by a National Science and Engineering Research Grant to R.S.S. and a Smithsonian Institute Molecular Evolution Postdoctoral Fellowship to S. Jagadeeshan. S. Jagadeeshan and W. Haerty contributed equally to this work.

References

[1] N. Eldredge and S. J. Gould, "Punctuated equilibria: an alternative to phyletic gradualism," in *Models in Paleobiology*, Freeman, San Francisco, Calif, USA, 1972.

[2] S. J. Gould and N. Eldredge, "Punctuated equilibrium comes of age," *Nature*, vol. 366, no. 6452, pp. 223–227, 1993.

[3] T. G. Barraclough, A. P. Vogler, and P. Harvey, "Revealing the factors that promote speciation," *Philosophical Transactions of the Royal Society B*, vol. 353, no. 1366, pp. 241–249, 1998.

[4] J. B. C. Jackson and A. H. Cheetham, "Tempo and mode of speciation in the sea," *Trends in Ecology and Evolution*, vol. 14, no. 2, pp. 72–77, 1999.

[5] T. G. Barraclough and V. Savolainen, "Evolutionary rates and species diversity in flowering plants," *Evolution*, vol. 55, no. 4, pp. 677–683, 2001.

[6] M. J. Benton and P. N. Pearson, "Speciation in the fossil record," *Trends in Ecology and Evolution*, vol. 16, no. 7, pp. 405–411, 2001.

[7] G. Hunt, "The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 47, pp. 18404–18408, 2007.

[8] J. C. Avise and F. J. Ayala, "Genetic change and rates of cladogenesis," *Genetics*, vol. 81, no. 4, pp. 757–773, 1975.

[9] J. C. Avise and F. J. Ayala, "Genetic differentiation in speciose versus depauperate phylads: evidence from the California minnows," *Evolution*, vol. 30, no. 1, pp. 46–58, 1976.

[10] J. C. Avise, "Is evolution gradual or rectangular? Evidence from living fishes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 74, no. 11, pp. 5083–5087, 1977.

[11] D. P. Mindell, J. S. Sites, and D. Graur, "Speciation evolution: a phylogenetic test with allozymes in Sceloporus (Reptilia)," *Cladistics*, vol. 5, no. 1, pp. 49–61, 1989.

[12] A. J. Webster, R. J. H. Payne, and M. Pagel, "Molecular phylogenies link rates of evolution and speciation," *Science*, vol. 301, no. 5632, p. 478, 2003.

[13] M. Pagel, C. Venditti, and A. Meade, "Large punctuational contribution of speciation to evolutionary divergence at the molecular level," *Science*, vol. 314, no. 5796, pp. 119–121, 2006.

[14] I. J. Leitch, J. M. Beaulieu, K. Cheung, L. Hanson, M. A. Lysak, and M. F. Fay, "Punctuated genome size evolution in Liliaceae," *Journal of Evolutionary Biology*, vol. 20, no. 6, pp. 2296–2308, 2007.

[15] C. Venditti and M. Pagel, "Speciation as an active force in promoting genetic evolution," *Trends in Ecology and Evolution*, vol. 25, no. 1, pp. 14–20, 2010.

[16] L. Arbiza, J. Dopazo, and H. Dopazo, "Positive selection, relaxation, and acceleration in the evolution of the human and chimp genome," *PLoS Computational Biology*, vol. 2, no. 4, Article ID e38, 2006.

[17] J. Savard, D. Tautz, and M. J. Lercher, "Genome-wide acceleration of protein evolution in flies (Diptera)," *BMC Evolutionary Biology*, vol. 6, Article ID 7, 2006.

[18] A. G. Clark, M. B. Eisen, D. R. Smith et al., "Evolution of genes and genomes on the *Drosophila* phylogeny," *Nature*, vol. 450, no. 7167, pp. 203–218, 2007.

[19] T. M. Panhuis, N. L. Clark, and W. J. Swanson, "Rapid evolution of reproductive proteins in abalone and *Drosophila*," *Philosophical Transactions of the Royal Society B*, vol. 361, no. 1466, pp. 261–268, 2006.

[20] W. Haerty, S. Jagadeeshan, R. J. Kulathinal et al., "Evolution in the fast lane: rapidly evolving sex-related genes in *Drosophila*," *Genetics*, vol. 177, no. 3, pp. 1321–1335, 2007.

[21] D. A. Barbash, P. Awadalla, and A. M. Tarone, "Functional divergence caused by ancient positive selection of a *Drosophila* hybrid incompatibility locus," *PLoS Biology*, vol. 2, no. 6, Article ID E142, 2004.

[22] N. J. Brideau, H. A. Flores, J. Wang, S. Maheshwari, XU. Wang, and D. A. Barbash, "Two Dobzhansky-Muller Genes interact to cause hybrid lethality in *Drosophila*," *Science*, vol. 314, no. 5803, pp. 1292–1295, 2006.

[23] J. P. Masly, C. D. Jones, M. A. F. Noor, J. Locke, and H. A. Orr, "Gene transposition as a cause of hybrid sterility in *Drosophila*," *Science*, vol. 313, no. 5792, pp. 1448–1450, 2006.

[24] D. C. Presgraves, L. Balagopalan, S. M. Abmayr, and H. A. Orr, "Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*," *Nature*, vol. 423, no. 6941, pp. 715–719, 2003.

[25] C. T. Ting, S. C. Tsaur, M. L. Wu, and C. I. Wu, "A rapidly evolving homeobox at the site of a hybrid sterility gene," *Science*, vol. 282, no. 5393, pp. 1501–1504, 1998.

[26] J. Wittbrodt, D. Adam, B. Malitschek et al., "Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu locus in *Xiphophorus*," *Nature*, vol. 341, no. 6241, pp. 415–421, 1989.

[27] K. S. Zigler, M. A. McCartney, D. R. Levitan, and H. A. Lessios, "Sea urchin bindin divergence predicts gamete compatibility," *Evolution*, vol. 59, no. 11, pp. 2399–2404, 2005.

[28] H. A. Orr, "The genetic basis of reproductive isolation: insights from *Drosophila*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, supplement 1, pp. 6522–6526, 2005.

[29] W. J. Swanson and V. D. Vacquier, "The rapid evolution of reproductive proteins," *Nature Reviews Genetics*, vol. 3, no. 2, pp. 137–144, 2002.

[30] E. C. Metz and S. R. Palumbi, "Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin," *Molecular Biology and Evolution*, vol. 13, no. 2, pp. 397–406, 1996.

[31] K. S. Zigler and H. A. Lessios, "Evolution of bindin in the pantropical sea urchin Tripneustes: comparisons to bindin of other genera," *Molecular Biology and Evolution*, vol. 20, no. 2, pp. 220–231, 2003.

[32] C. G. Artieri, W. Haerty, and R. S. Singh, "Association between levels of coding sequence divergence and gene misregulation in *Drosophila* male hybrids," *Journal of Molecular Evolution*, vol. 65, no. 6, pp. 697–704, 2007.

[33] R. M. Kliman, P. Andolfatto, J. A. Coyne et al., "The population genetics of the origin and divergence of the *Drosophila simulans* complex species," *Genetics*, vol. 156, no. 4, pp. 1913–1931, 2000.

[34] K. Tamura, S. Subramanian, and S. Kumar, "Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks," *Molecular Biology and Evolution*, vol. 21, no. 1, pp. 36–44, 2004.

[35] Y. Zhang, D. Sturgill, M. Parisi, S. Kumar, and B. Oliver, "Constraint and turnover in sex-biased gene expression in the genus *Drosophila*," *Nature*, vol. 450, no. 7167, pp. 233–237, 2007.

[36] S. Kumar, A. Filipski, V. Swarna, A. Walker, and S. B. Hedges, "Placing confidence limits on the molecular age of the human-chimpanzee divergence," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 52, pp. 18842–18847, 2005.

[37] F. Rodríguez-Trelles, R. Tarrío, and F. J. Ayala, "A methodological bias toward overestimation of molecular evolutionary time scales," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 12, pp. 8112–8115, 2002.

[38] J. D. Thompson, D. G. Higgins, and T. J. Gibson, "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap

penalties and weight matrix choice," *Nucleic Acids Research*, vol. 22, no. 22, pp. 4673–4680, 1994.

[39] D. A. Pollard, V. N. Iyer, A. M. Moses, and M. B. Eisen, "Widespread discordance of gene trees with species tree in *Drosophila*: evidence for incomplete lineage sorting," *PLoS Genetics*, vol. 2, no. 10, p. e173, 2006.

[40] A. Wong, J. D. Jensen, J. E. Pool, and C. F. Aquadro, "Phylogenetic incongruence in the *Drosophila melanogaster* species group," *Molecular Phylogenetics and Evolution*, vol. 43, no. 3, pp. 1138–1150, 2007.

[41] Z. Yang, "PAML: a program package for phylogenetic analysis by maximum likelihood," *Computer Applications in the Biosciences*, vol. 13, no. 5, pp. 555–556, 1997.

[42] Z. Yang and R. Nielsent, "Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages," *Molecular Biology and Evolution*, vol. 19, no. 6, pp. 908–917, 2002.

[43] A. M. Larracuente, T. B. Sackton, A. J. Greenberg et al., "Evolution of protein-coding genes in *Drosophila*," *Trends in Genetics*, vol. 24, no. 3, pp. 114–123, 2008.

[44] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, 1969.

[45] J. A. Coyne and A. H. Orr, *Speciation*, Sinauer Associates, Sunderland, Mass, USA, 2004.

[46] R. S. Singh, "Toward a unified theory of speciation," in *Evolutionary Genetics from Molecules to Morphology*, R. S. Singh and C. Krimbas, Eds., pp. 570–608, 2000.

[47] A. Civetta and R. S. Singh, "High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in *Drosophila melanogaster* and *Drosophila virilis* group species," *Journal of Molecular Evolution*, vol. 41, no. 6, pp. 1085–1095, 1995.

[48] S. Jagadeeshan and R. S. Singh, "Rapidly evolving genes of *Drosophila*: differing levels of selective pressure in testis, ovary, and head tissues between sibling species," *Molecular Biology and Evolution*, vol. 22, no. 9, pp. 1793–1801, 2005.

[49] D. J. Begun, A. K. Holloway, K. Stevens et al., "Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*," *PLoS Biology*, vol. 5, no. 11, p. e310, 2007.

[50] J. F. Baines, S. A. Sawyer, D. L. Hartl, and J. Parsch, "Effects of X-linkage and sex-biased gene expression on the rate of adaptive protein evolution in *Drosophila*," *Molecular Biology and Evolution*, vol. 25, no. 8, pp. 1639–1650, 2008.

[51] N. D. Singh, A. M. Larracuente, and A. G. Clark, "Contrasting the efficacy of selection on the X and autosomes in drosophila," *Molecular Biology and Evolution*, vol. 25, no. 2, pp. 454–467, 2008.

[52] J. Hey and R. M. Kliman, "Interactions between natural selection, recombination and gene density in the genes of *Drosophila*," *Genetics*, vol. 160, no. 2, pp. 595–608, 2002.

[53] D. Lachaise, M. Harry, M. Solignac, F. Lemeunier, V. Benassi, and M. L. Cariou, "Evolutionary novelties in islands: *Drosophila santomea*, a new melanogaster sister species from São Tomé," *Proceedings of the Royal Society B*, vol. 267, no. 1452, pp. 1487–1495, 2000.

[54] J. Charlesworth and A. Eyre-Walker, "The other side of the nearly neutral theory, evidence of slightly advantageous back-mutations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 43, pp. 16992–16997, 2007.

[55] S. Y. W. Ho, M. J. Phillips, A. Cooper, and A. J. Drummond, "Time dependency of molecular rate estimates and systematic overestimation of recent divergence times," *Molecular Biology and Evolution*, vol. 22, no. 7, pp. 1561–1568, 2005.

[56] R. K. Butlin, "Recombination and speciation," *Molecular Ecology*, vol. 14, no. 9, pp. 2621–2635, 2005.

[57] R. J. Kulathinal, S. M. Bennett, C. L. Fitzpatrick, and M. A. F. Noor, "Fine-scale mapping of recombination rate in *Drosophila* refines its correlation to diversity and divergence," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 29, pp. 10051–10056, 2008.

[58] M. Nei, "Bottlenecks, genetic polymorphism and speciation," *Genetics*, vol. 170, no. 1, pp. 1–4, 2005.

[59] H. L. Carson, "The genetic system, the deme, and the origin of species," *Annual Review of Genetics*, vol. 21, pp. 405–423, 1987.

[60] J. B. Losos, K. I. Warheit, and T. W. Schoener, "Adaptive differentiation following experimental island colonization in Anolis lizards," *Nature*, vol. 387, no. 6628, pp. 70–73, 1997.

[61] J. B. Losos, T. W. Schoener, R. B. Langerhans, and D. A. Spiller, "Rapid temporal reversal in predator-driven natural selection," *Science*, vol. 314, no. 5802, p. 1111, 2006.

[62] J. N. Thompson, "Rapid evolution as an ecological process," *Trends in Ecology and Evolution*, vol. 13, no. 8, pp. 327–329, 1998.

[63] S. R'Kha, P. Capy, and J. R. David, "Host-plant specialization in the *Drosophila melanogaster* species complex: a physiological, behavioral, and genetical analysis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 5, pp. 1835–1839, 1991.

[64] I. Dworkin and C. D. Jones, "Genetic changes accompanying the evolution of host specialization in Drosophila sechellia," *Genetics*, vol. 181, no. 2, pp. 721–736, 2009.

[65] M. L. Cariou, J. F. Silvain, V. Daubin, J. L. Da Lage, and D. Lachaise, "Divergence between *Drosophila santomea* and allopatric or sympatric populations of *D. yakuba* using paralogous amylase genes and migration scenarios along the Cameroon volcanic line," *Molecular Ecology*, vol. 10, no. 3, pp. 649–660, 2001.

[66] A. J. Moehring, A. Llopart, S. Elwyn, J. A. Coyne, and T. F. C. Mackay, "The genetic basis of prezygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to mating preference," *Genetics*, vol. 173, no. 1, pp. 215–223, 2006.

[67] A. J. Moehring, A. Llopart, S. Elwyn, J. A. Coyne, and T. F. C. Mackay, "The genetic basis of postzygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to hybrid male sterility," *Genetics*, vol. 173, no. 1, pp. 225–233, 2006.

[68] J. A. Coyne, S. Y. Kim, A. S. Chang, D. Lachaise, and S. Elwyn, "Sexual isolation between two sibling species with overlapping ranges: *Drosophila santomea* and *Drosophila yakuba*," *Evolution*, vol. 56, no. 12, pp. 2424–2434, 2002.

[69] J. E. Blyth, D. Lachaise, and M. G. Ritchie, "Divergence in multiple courtship song traits between *Drosophila santomea* and *D. yakuba*," *Ethology*, vol. 114, no. 7, pp. 728–736, 2008.

[70] M. G. Ritchie, "Evolution: feathers, females, and fathers," *Science*, vol. 318, no. 5847, pp. 54–55, 2007.

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 274975, 9 pages
doi:10.4061/2011/274975

Research Article

Elevated Evolutionary Rates among Functionally Diverged Reproductive Genes across Deep Vertebrate Lineages

Christopher J. Grassa[1] and Rob J. Kulathinal[2]

[1] *Department of Botany, University of British Columbia, 6270 University Boulevard, Vancouver, BC, Canada V6T 1Z4*
[2] *Department of Biology, Temple University, 1900 North 12th Street, Philadelphia, PA 19122, USA*

Correspondence should be addressed to Rob J. Kulathinal, robkulathinal@temple.edu

Received 1 February 2011; Revised 17 May 2011; Accepted 23 May 2011

Academic Editor: Jose M. Eirin-Lopez

Among closely related taxa, proteins involved in reproduction generally evolve more rapidly than other proteins. Here, we apply a functional and comparative genomics approach to compare functional divergence across a deep phylogenetic array of egg-laying and live-bearing vertebrate taxa. We aligned and annotated a set of 4,986 1 : 1 : 1 : 1 : 1 orthologs in *Anolis carolinensis* (green lizard), *Danio rerio* (zebrafish), *Xenopus tropicalis* (frog), *Gallus gallus* (chicken), and *Mus musculus* (mouse) according to function using ESTs from available reproductive (including testis and ovary) and non-reproductive tissues as well as Gene Ontology. For each species lineage, genes were further classified as tissue-specific (found in a single tissue) or tissue-expressed (found in multiple tissues). Within independent vertebrate lineages, we generally find that gonadal-specific genes evolve at a faster rate than gonadal-expressed genes and significantly faster than non-reproductive genes. Among the gonadal set, testis genes are generally more diverged than ovary genes. Surprisingly, an opposite but nonsignificant pattern is found among the subset of orthologs that remained functionally conserved across *all* five lineages. These contrasting evolutionary patterns found between functionally diverged and functionally conserved reproductive orthologs provide evidence for pervasive and potentially cryptic lineage-specific selective processes on ancestral reproductive systems in vertebrates.

1. Introduction

Over the past 550 million years, evolutionary processes have generated a diverse array of vertebrate species. Taxa that include fishes, birds, reptiles, and mammals evolved unique suites of adaptations allowing them to prosper in the most extreme sea, air, and land environments. Vertebrate diversity spans morphological innovations, developmental programming, cellular responses, as well as behaviors and life histories, and such differences become increasingly evident when taxa are compared across deep phylogenies. Studying the evolutionary patterns of functional change across this subphylum provides an opportunity to understand the evolutionary processes that have been important throughout vertebrate evolution. Yet, to date, clear *common* functional signatures that are in rapid flux across all vertebrate taxa have not been identified indicating the historical presence of a variety of niche- and lineage-dependent selective processes.

While functional evolutionary signals are not apparent across diverse phylogenetic lineages, when more closely related species such as sister species or multiple species within a single genus are compared, reproductive traits consistently reveal high diversity among species. This reproductive signature has been known for centuries, beginning with Linnaeus' binomial classification system [1]. Charles Darwin, in his 1871 treatise on sexual selection, also catalogued highly differentiated secondary sexual organs between closely related bird and mammal species [2]. Over a century later, William Eberhard described the diversity of morphological differences found in male secondary sexual traits, including vertebrate genitalia [3]. Both Darwin and Eberhard explain this higher male variance as the result of female mate choice or male-male competition on sexually selected traits within populations. The last three decades have amassed more vertebrate examples including cichlids [4], frogs [5], and primates [6] indicating that selection on reproductive

traits may be a common underlying evolutionary process in vertebrates.

Studying rates of morphological character change demonstrates how certain functional classes evolve relative to others and provides a lens into evolutionary processes of the past. While this framework works well on closely related species, signatures diminish when applied to distantly related taxa due to the presence of lineage-specific rates of development, selective constraints, and genetic architectural differences [7]. In addition, there are many processes in which the selected phenotype may be hidden or cryptic to human observers. Such phenotypes often occur at the molecular level and include immune response [8], gametic interactions [9], and pheromonal exchange [10]. To systematically understand the relative roles of different functional classes in the evolutionary history of vertebrates, and hence the role of certain selective processes, it would be instructive to employ a common and unbiased framework on a representative sample of taxa.

With the availability of annotated genomic sequences across an ever-expanding number of taxa in addition to associated functional data (e.g., ESTs, GO) that can link genes to function, an operational framework is emerging that compares rates of functional change across varying degrees of phylogenetic relatedness [11, 12]. By applying this functional and comparative genomics approach, we now can use normalized information from sequences to infer how functional categories of genes have changed in the past. Combining the two domains of time and function can provide valuable information about the history of these lineages, in particular, how certain selective forces act upon certain reproductive processes such as gamete recognition, oogenesis, spermatogenesis, and adult behavior.

In this paper, we quantify the rates of change among reproductive and non-reproductive genes in five distantly related vertebrate lineages. We functionally categorize ~5,000 orthologs using available testis, ovary, and non-reproductive EST libraries in each species and find that individual vertebrate lineages generally follow a pattern of greater divergence in genes solely expressed in the gonads compared to genes expressed in non-reproductive tissue. In most cases, the testis appears to be driving gonadal divergence. However, an opposing pattern emerges when we compare evolutionary rates among the much smaller subset of tissue-expressed genes that have remained functionally conserved across vertebrates ($dN_{testis} < dN_{ovary} < dN_{non\text{-}reproductive}$). Using this framework, we are beginning to unmask a pattern of rapid and cryptic molecular evolution on lineage-specific reproductive features that are part of conserved developmental processes, thus, providing a common underlying genetic basis of functional evolutionary change in the vertebrate subphylum.

2. Materials and Methods

2.1. Orthology and Estimates of Divergence. Protein coding genes from *A. carolinensis, G. gallus, D. rerio, M. musculus,* and *X. tropicalis* were used in this analysis. Orthologs for each species pair were obtained from BioMart (http://uswest.ensembl.org/biomart/index.html). Orthologs were filtered so that only transitive sets of 1 : 1 : 1 : 1 : 1 orthologs remained, producing 4,986 sets of 5-species orthologs. We excluded all paralogous relationships (including 10,122 1 : 1 : 2 : 1 : 1 relationships, where "2" denotes paralogous sequences from the zebrafish lineage) in order to maintain a relatively ambiguous ortholog set. The protein coding CDS and amino acid sequence of each gene's longest transcript were also obtained from BioMart: in the case of transcript length ties, the transcript with the lower incremental Ensemble ID number was used. Multiple sequence alignments for each orthologous set of proteins were generated using MUSCLE (version 3.8; [13]) and then back-translated using corresponding CDS and a custom Perl script (available from CJG on request). All 1 : 1 : 1 : 1 : 1 alignments in addition to their associated functional assignments will be made available via *lizardbase* (http://www.lizardbase.org/) as an active link to current *A. carolinensis* annotations in *lizardbase*'s genome browser, JBrowse, and *lizardbase*'s Resources Page. All alignments will also be made available on the Resources page in *lizardbase.*

A protein distance matrix was calculated for each protein alignment using the Jones-Taylor-Thornton (JTT) model in the prodist program from the Phylip suite of phylogenetic programs (version 3.69; [14]). Consensus phylogenetic trees were generated using concatenated sequences from both CDS and its associated protein sequences (See in Supplementary Material available online at doi:10.4061/2011/274975 Supplementary Figure 1). For a given gene from each species, the mean of its four orthologous protein distances was used as one of two estimates of sequence divergence. A matrix of nonsynonymous substitutions per nonsynonymous site, dN, was also estimated for each codon alignment using Nei and Gojobori's method [15] using the SNAP Perl program [16], and its mean dN across four orthologs was used as an estimate of sequence divergence.

2.2. Functional Annotation Using EST Libraries and Gene Ontologies. ESTs from each of the five species were filtered as "normal adult" tissue from NCBI's dbEST (downloaded in October 2009) and assigned to species-specific tissue libraries (see Supplementary Table 1) based on either organ or tissue fields in the Genbank record. EST sequences were locally indexed and aligned to genes from the same species using a standalone version of blastn (version 2.22; [17]). EST-to-gene alignments of at least 100 nucleotides, 90% identity, and an *E*-value of $e - 20$ were used as alignment criteria. For each of the five species lineages, genes with at least three ESTs (i.e., hits) meeting the above alignment criteria were assigned to seven non mutually exclusive functional classes: (1) genes with hits in only the testis were classified as testis-specific; (2) genes with hits in the testis and another tissue(s) were classified as testis-expressed; (3) genes with hits in only the ovary were classified as ovary-specific; (4) genes with hits in the ovary and another tissue(s) were classified as ovary-expressed; (5) genes with hits in only the testis and/or ovary were classified as gonadal-specific; (6) genes with hits in the testis and/or ovary, in addition to non-reproductive tissue(s), were classified as gonadal-expressed;

Table 1: Functional classification of reproductive (testis, ovary, gonadal) and non-reproductive orthologs in vertebrate species. Genes were assigned to at least one of seven functional categories (see Section 2.2 for explanation). Non-reproductive genes are found in neither the testis nor ovary EST libraries but are present in other tissues.

Functional classification	*A. carolinensis*	*D. rerio*	*X. tropicalis*	*G. gallus*	*M. musculus*
Testis-specific	55	49	129	12	43
Testis-expressed	613	2511	2825	1011	1659
Ovary-specific	87	21	13	16	0
Ovary-expressed	889	2422	1367	2033	350
Gonadal-specific	243	126	187	53	45
Gonadal-expressed	1109	3142	3109	2447	1848
Non-reproductive	243	537	613	940	2159
Total annotated genes (out of 4986)	**1352**	**3679**	**3722**	**3387**	**4007**

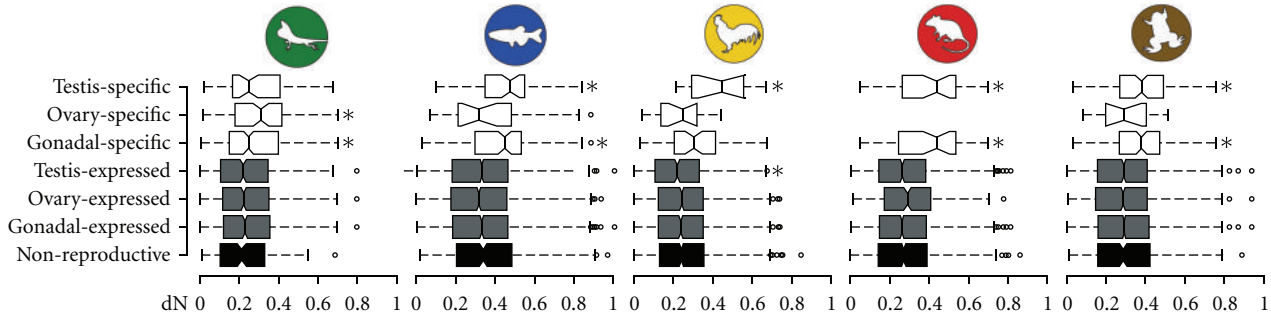


Figure 1: Protein divergence versus functional class across vertebrate lineages. Boxplots show the distribution of dN, nonsynonymous substitutions per nonsynonymous site in seven functional classes for each of the five species, *A. carolinensis*, *D. rerio*, *G. gallus*, *M. musculus*, and *X. tropicalis*. The three tissue-specific classes are found on the top (nonshaded), tissue-expressed classes are below in grey, and the non-reproductive functional class is indicated on the bottom, in black. Asterisks on the right-hand side of a boxplot signifies a highly significant ($P < 0.001$) difference in mean, as given by the Wilcoxon rank sum test, when compared to the non-reproductive class. No ovary-specific genes were identified in *M. musculus*.

(7) genes with hits from an assortment of non-reproductive tissues (see Supplementary Table 1) that were neither testis nor ovary were classified as non-reproductive. Thus, for each of the five species, genes with sufficient EST coverage fell into at least one functional class (Table 1). The difference between the mean dN of each reproductive class and the mean dN of the non-reproductive class was tested using an unpaired two-sample two-sided Wilcoxon rank sum test ([18]; Figure 1).

We also compared evolutionary rates in *functionally conserved* genes, that is, those orthologs that do not change functional class across all five lineages, according to our EST annotations. Interestingly, we were not able to identify a single gonadal-specific gene, but were able to identify functionally conserved subsets of testis-expressed ($n = 95$), ovary-expressed ($n = 16$), and non-reproductive ($n = 3$) orthologs. Figures 2(a)–2(g) provides Venn diagrams for all species combinations in each functional class. Figure 3 compares dN across four (nonzero) functional classes.

To complement the functional annotations generated by ESTs, we linked the 10% most diverged orthologs to the GO categories, Biological Process (BP) and Cellular Component (CC) in each species. GeneMerge [19] was used to test for statistically significant over-represented functional terms. A "word cloud" that relates the frequency of each GO term to its font size was generated for four of the five species (Figure 4, Supplementary Figure 2). *X. tropicalis* was excluded from this analysis due to its sparse GO term set.

3. Results and Discussion

In this study, we chose five distantly related vertebrate species that fit the following criteria: (1) the presence of a well-assembled and freely available genome sequence, (2) the existence of well-curated gene models, (3) the availability of appreciable numbers of testis, ovary, and non-reproductive ESTs at dbEST, and (4) the condition that all five species, together, represent divergent clades thus presenting a deep vertebrate phylogeny with a diverse breadth of functional differences. After filtering out alignments that were of poor quality or had ambiguous orthologous relationships, a consensus tree-based off-concatenated CDS sequences from 4,986 orthologs was generated using the five vertebrate species. The tree's topology was well supported in 100% of 1000 bootstrap replicates (Supplementary Figure 1). A concatenated protein tree-demonstrated the same topology and support (not shown) and mirrored published vertebrate phylogenies (e.g., [20]). We note that these ~5,000 orthologs represent a relatively "well-behaved" and conserved gene set

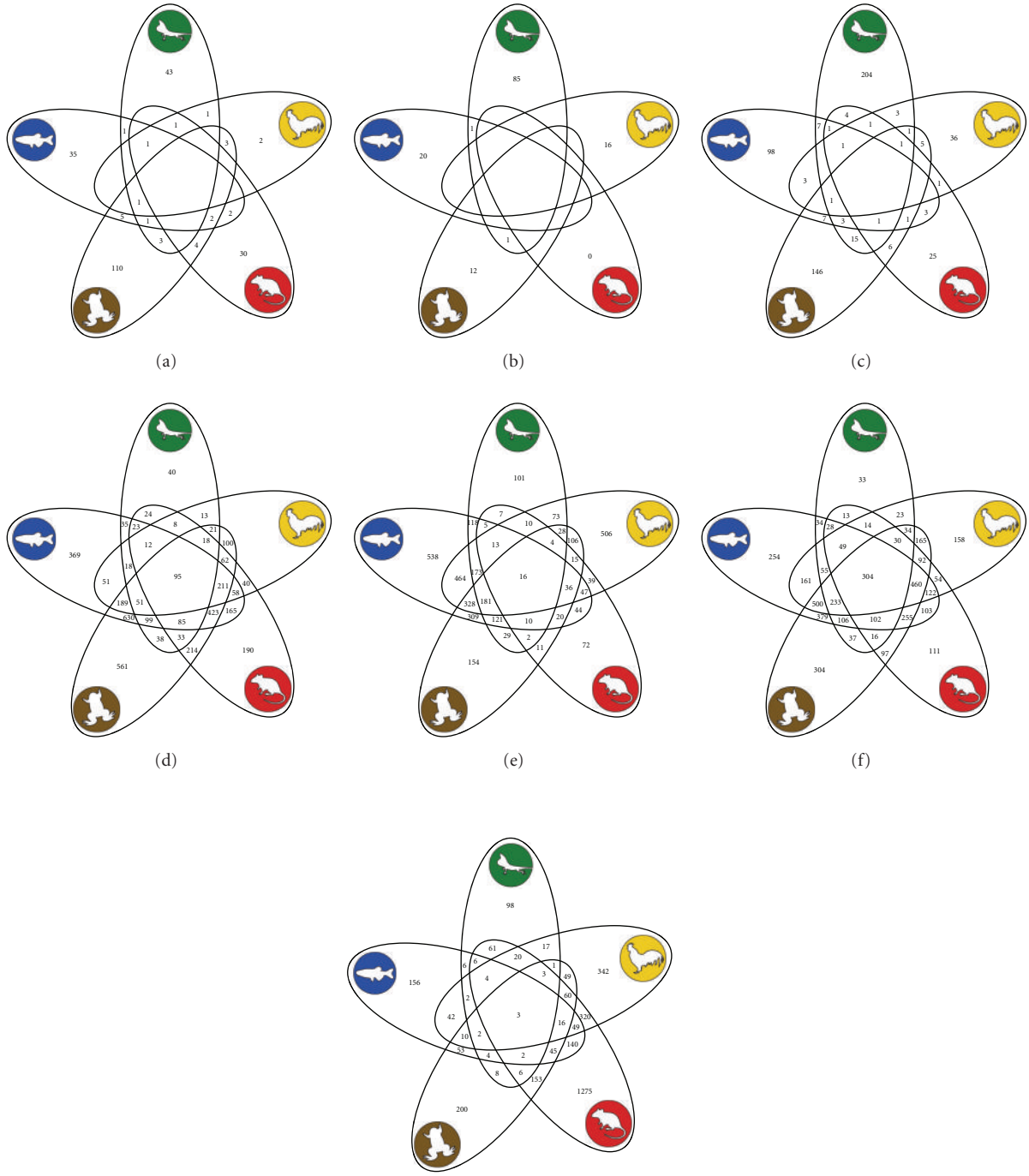


Figure 2: Venn diagrams of common functionally conserved genes across all five vertebrate species. For each of the seven functional classes, the number of genes found in all combination of species intersections and exclusions are listed. (a) testis-specific, (b) ovary-specific, (c) gonadal-specific, (d) testis-expressed, (e) ovary-expressed, (f) gonadal-expressed, (g) non-reproductive.

that do not possess paralogs in any of the five lineages. This study focuses on 1 : 1 : 1 : 1 : 1 orthologs and ignores complications arising from neo-/subfunctionalization caused by gene duplication events [21, 22], particularly those found in the zebrafish lineage after an ancient duplication event [23].

We used the extensive EST libraries publically available for each species in order to categorize genes into functional classes. Our objective was to generate a standardized sample of genes in each of the reproductive and non-reproductive functional classes, for each species. Historically, EST libraries

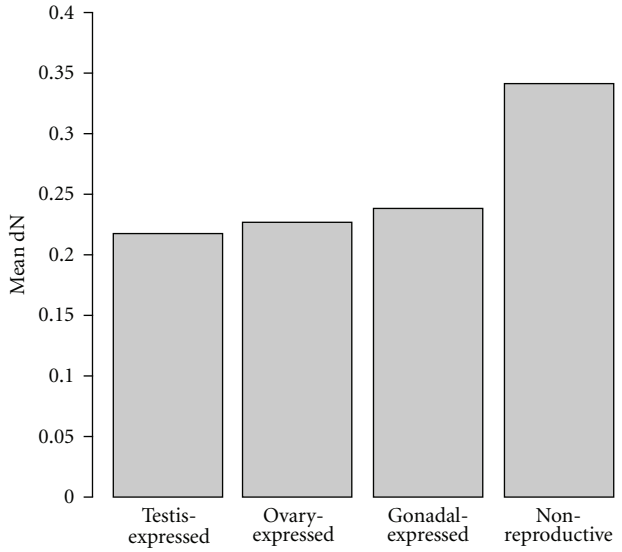


Figure 3: dN among functionally conserved classes across all five vertebrate species. Only four of the seven functional classes contained genes that were found in the same functional class across zebrafish, Anolis, Xenopus, chicken, and mouse. Functional classes were not significantly different from each other.

were originally developed to assist in the genome annotation process (e.g., [24]). The quantity, quality, specificity, and tissue-diversity of EST libraries vary considerably across species (see Supplementary Table 1) and are largely a function of each research community's priorities and preferences for each of the five sequenced genomes. Since our principal objective is to compare reproductive versus. non-reproductive levels of molecular divergence in vertebrates, we sought to generate pooled gene samples derived from testis and ovary (i.e., reproductive) and non-reproductive tissue (any adult tissue that does not contain a sex-specific organ or tissue). In addition, genes from tissue-specific (or tissue-limited) classes were differentiated from "tissue-expressed" genes that are expressed more ubiquitously. This approach enables us to compare functional gene classes using relatively large sample sizes and ample statistical power.

A total of seven functional classes were assigned to genes in each of the five species (see Section 2). Table 1 summarizes the number of genes that are contained in each functional class for each species. It is important to note that the proportion of reproductive (e.g., testis, ovary) to non-reproductive genes in each species is not necessarily indicative of the total fraction of reproductive genes found in each genome but, again, reflects each community's specialized interests in generating certain libraries. In addition, overall EST library coverage can be different by an order of magnitude. For example, at last count, the mouse has nearly 5 million ESTs deposited in dbEST, while the green lizard has only 150,000 ESTs. The broader EST coverage in mouse may explain why our screen failed to identify any ovary-specific genes in this taxon. In contrast, since the anoles EST set includes only three non-reproductive tissues at a lower coverage than other species, this may also explain the relatively high number of ovary-specific genes in this species. With such large differences in EST coverage in each of the five species, it is important to understand the limits of these analyses.

Overall, our results provide evidence of a general pattern of rapid reproductive change over deep vertebrate lineages. Each of the five vertebrates demonstrate significantly higher protein divergence in gonadal genes compared to non-reproductive genes (Figure 1). Rapidly evolving testis genes appear to be driving much of the pattern of higher gonadal-specific gene divergence in these lineages: four of the five taxa—zebrafish, Xenopus, chicken, and mouse—all share significantly higher testis-specific divergence. Interestingly, these three taxa include two of the more basal taxa, Xenopus and zebrafish (Supplementary Figure 1), supporting that this pattern spans broad phylogenetic groups across the vertebrate subphylum. In green lizards, we observe a contrasting pattern of gonadal divergence as ovary-specific genes appear to be driving the significantly higher divergence of gonadal genes (but see caveat above). Thus, while we see a general pattern of significantly higher divergence among reproduction-specific genes across all vertebrate lineages, there may be large differences in the subset of reproductive genes that are diverging.

In Drosophila, we also see a similar pattern of rapidly evolving gonadal genes from EST libraries. Reproductive genes from the testis and ovary and non-reproductive genes from the brain have been used to characterize sexually dimorphic expression patterns [25–27] as well as to compare the evolution of reproductive genes relative to non-reproductive genes [28–30]. A recent study using 12 genomes in Drosophila and an extensive EST set from *D. melanogaster* also found that rates of evolution among testis-expressed genes are significantly higher than genes expressed in the ovary or head [12]. A number of studies in mammals have also demonstrated a similar pattern of higher divergence rates in male reproductive genes [11, 31–33].

This higher divergence of reproductive genes, and in particular, male-specific proteins, supports the hypothesis that sexual selection may be an important driver of evolutionary change and extends sexual selection theory to the level of molecules such as those found in gametogenesis and fertilization [34–36]. The strength of this molecular signature indicates the pervasive and cryptic nature of this process: much of this pattern would remain hidden without a comparative and systematic treatment of genome-wide sequence data. We also note that reproductive proteins, particularly those regulating sperm development, are of particular interest to researchers studying mechanisms of reproductive isolation because hybrid male sterility may be the product of the rapid evolution of male reproductive genes: spermatogenesis appears to be a selected target of hybrid male fertility breakdown [37–41]. In addition, there is mounting evidence that positive selection drives the evolution of genes controlling key transitions during both spermatogenesis and oogenesis [42, 43].

Other functional classes of testis-associated genes have also been found in Drosophila. Genes encoding proteins secreted by male accessory glands (Acps), the ejaculatory

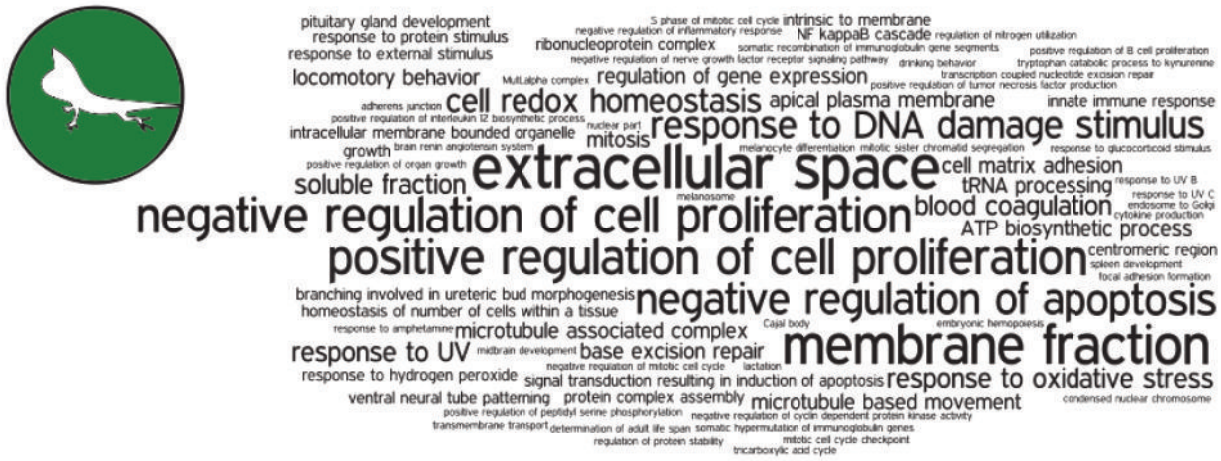


Figure 4: Word-size frequency distribution of Gene Ontology (GO) terms for the most diverged orthologs in *A. carolinensis*. Associated GO terms for the top 10% diverged ortholog subset are displayed according to size, based on the frequency of that term. GO terms from Biological Process (BP) and Cellular Component (CC) were used. Similar GO-based word-size frequencies based on 10% most diverged orthologs from *M. musculus, G. gallus, and D. rerio* are found in Supplementary Figure 2.

duct, and the ejaculatory bulb, as well as many components of *D. melanogaster* seminal fluid, are known to be rapidly evolving. These proteins are transferred from the male to the female along with sperm during mating and mediate a series of postmating events [44–46]. Furthermore, there is ample evidence of adaptive evolution at several loci that encode *D. melanogaster* seminal fluid proteins [47–52]. Whether a similar signal among secretory reproductive classes is found in vertebrate lineages is an intriguing question.

While a clear pattern of rapid testis-specific divergence emerges from our lineage-specific annotations, we then asked whether the same evolutionary pattern holds across genes that have maintained a similar function across all five vertebrate species. In other words, what are the relative rates of evolutionary change across *functionally conserved* classes? Surprisingly, the numbers of genes per class were drastically reduced to the point that only four classes—testis-expressed, ovary-expressed, gonadal-expressed, and non-reproductive genes—share genes in common across all five species (Figure 2). Furthermore, a decreasing but nonsignificant trend of evolutionary rates was found among these four functional classes: $dN_{\text{non-reproductive}} > dN_{\text{gonad-expressed}} > dN_{\text{ovary-expressed}} > dN_{\text{testis-expressed}}$. Overall, this functionally conserved group describes a subset of the data with a contrasting evolutionary pattern, thereby demonstrating that testis-specific genes are affected by a variety of evolutionary forces. In a recent study, Dean et al. [33] performed a genomic and proteomic study on six tissue types from the male reproductive tract of mouse (excluding testis) and found that one tract, the seminal vesicle, had significantly higher rates of divergence while the other five tracts showed significantly lower rates of divergence when compared to other proteins. Our results demonstrate a similar high variance of evolutionary rates within the testis.

A. carolinensis was the outlier of the five vertebrate taxa with a significantly higher divergence among ovary-specific genes. Ovaries have also been shown to be sites of rapid divergence in *D. melanogaster* as part of a molecular coevolutionary process between sperm and egg. A number of rapidly evolving genes have been found expressed in the female reproductive tract and potentially secreted [53] or induced in the female reproductive tract by mating [54–56]. Further characterization of the green anole genome, in addition to other Lepidosauria genomes and genomic resources that will soon be available, will allow us to address whether female lizards are indeed driving sexual selective processes and whether this is a common lineage-specific process among squamate reptiles or simply an artifact of EST functional annotation.

While aligning orthologs to ESTs offers a powerful approach for functional annotation, it is important to procure a more granular understanding of process, function and localization. Therefore, we took the 10% most diverged orthologs and associated each species' corresponding gene to its Gene Ontology (GO), namely, Biological Process (BP) and Cellular Component (CC). A word cloud in which the font size is a function of the frequency of statistically over-represented functional phrases in diverged orthologs is shown for *A. carolinensis* in Figure 4. GO-associated word frequencies for fish, mouse, and chicken are found in Supplementary Figure 2. The density of each word cloud for each species reflects the amount of curation effort in Gene Ontology within these species' communities. As expected, we don't see much overlap between the GO and EST approaches to functional annotation. Reproductive function is a poorly annotated ontological class harboring a level of characterization that will not substantially improve until

more geneticists and molecular biologists study reproductive loci in greater detail.

4. Concluding Remarks

Patterns parsed from extant genomes can inform us about the underlying evolutionary processes that have acted upon lineages in the past. As a functional class, reproductive-specific genes are more rapidly evolving than other functional gene classes, and it appears that testis genes are driving this pattern of divergence in the majority of vertebrate lineages. This work sets the stage for a more nuanced analysis of divergence leveraged against function across diverse taxa. With more genomes and ESTs generated, greater effort can be afforded to better estimate the probability that a gene is a member of a particular functional class, even when the number of ESTs and libraries are quite different between species. Newer data types such as RNAseq will certainly help solve the sampling bias problem with better coverage and more tissues sampled. Future studies that include paralogous sequences to evaluate birth/death processes and *de novo* gene functionalization models (including incorporating the large number of paralogs from zebrafish) in the context of functional class will also be useful in addressing the role of reproductive genes in vertebrate evolution.

It is remarkable that across very distant phylogenetic lineages, we detect the same evolutionary patterns found among closely related species: high lineage-specific reproductive diversity and, in particular, a high variance in male reproductive characters. These parallel patterns support the contention that sexual selection on both morphological and molecular characters may be an important, common, and pervasive feature of vertebrate evolution.

Abbreviations

GO: Gene ontology
EST: Expressed sequence tag
CDS: Coding sequence
BP: Biological process
CC: Cellular component
dN: Nonsynonymous substitutions per nonsynonymous site.

Acknowledgments

The authors would like to thank Dr. Tonia Hsieh (Temple University) for guidance and support during all stages of this project. They also thank Dr. Ed Braun (University of Florida) for technical advice and the Beaty Biodiversity Research Centre at the University of British Columbia for use of their SciBorg cluster. C. J. Grassa dedicates his contributions to Thomas and Sarah Grassa.

References

[1] C. Linnaeus, *Systema Naturae*, Nieuwkoop, The Netherlands, 1735, reprinted in 1964 by B. de Graaf.

[2] C. R. Darwin, *The Descent of Man, and Selection in Relation to Sex*, John Murray, London, UK, 1871.

[3] W. G. Eberhard, *Sexual Selection and Animal Genitalia*, Harvard University Press, Cambridge, Mass, USA, 1985.

[4] W. Salzburger and A. Meyer, "The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics," *Naturwissenschaften*, vol. 91, no. 6, pp. 277–290, 2004.

[5] K. E. Boul, W. C. Funk, C. R. Darst, D. C. Cannatella, and M. J. Ryan, "Sexual selection drives speciation in an Amazonian frog," *Proceedings of the Royal Society B*, vol. 274, no. 1608, pp. 399–406, 2007.

[6] A. H. Harcourt and J. Gardiner, "Sexual selection and genital anatomy of male primates," *Proceedings of the Royal Society B*, vol. 255, no. 1342, pp. 47–53, 1994.

[7] J. Roux and M. Robinson-Rechavi, "Developmental constraints on vertebrate genome evolution," *PLoS Genetics*, vol. 4, no. 12, Article ID e1000311, 2008.

[8] B. Korber, M. LaBute, and K. Yusim, "Immunoinformatics comes of age," *PLoS Computational Biology*, vol. 2, no. 6, article e71, 2006.

[9] W. J. Swanson and V. D. Vacquier, "The rapid evolution of reproductive proteins," *Nature Reviews Genetics*, vol. 3, no. 2, pp. 137–144, 2002.

[10] T. R. Shirangi, H. D. Dufour, T. M. Williams, and S. B. Carroll, "Rapid evolution of sex pheromone-producing enzyme expression in Drosophila," *PLoS Biology*, vol. 7, no. 8, Article ID e1000168, 2009.

[11] C. I. Castillo-Davis, F. A. Kondrashov, D. L. Hartl, and R. J. Kulathinal, "The functional genomic distribution of protein divergence in two animal phyla: coevolution, genomic conflict, and constraint," *Genome Research*, vol. 14, no. 5, pp. 802–811, 2004.

[12] W. Haerty, S. Jagadeeshan, R. J. Kulathinal et al., "Evolution in the fast lane: rapidly evolving sex-related genes in Drosophila," *Genetics*, vol. 177, no. 3, pp. 1321–1335, 2007.

[13] R. C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Research*, vol. 32, no. 5, pp. 1792–1797, 2004.

[14] J. Felsenstein, *PHYLIP (Phylogeny Inference Package) Version 3.6*, Department of Genome Sciences, University of Washington, Seattle, Wash, USA, 2005.

[15] M. Nei and T. Gojobori, "Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions," *Molecular Biology and Evolution*, vol. 3, no. 5, pp. 418–426, 1986.

[16] B. Korber, "HIV signature and sequence variation analysis," in *Computational Analysis of HIV Molecular Sequences*, A. G. Rodrigo and G. H. Learn, Eds., chapter 4, pp. 55–72, Kluwer Academic, Dodrecht, The Netherlands, 2000.

[17] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990.

[18] D. F. Bauer, "Constructing confidence sets using rank statistics," *Journal of the American Statistical Association*, vol. 67, pp. 687–690, 1972.

[19] C. I. Castillo-Davis and D. L. Hartl, "GeneMerge—post-genomic analysis, data mining, and hypothesis testing," *Bioinformatics*, vol. 19, no. 7, pp. 891–892, 2003.

[20] S. B. Hedges, "Molecular evidence for the origin of birds," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 7, pp. 2621–2624, 1994.

[21] M. Lynch and A. Force, "The probability of duplicate gene preservation by subfunctionalization," *Genetics*, vol. 154, no. 1, pp. 459–473, 2000.

[22] M. Lynch and J. S. Conery, "The evolutionary fate and consequences of duplicate genes," *Science*, vol. 290, no. 5494, pp. 1151–1155, 2000.

[23] J. S. Taylor, Y. Van de Peer, I. Braasch, and A. Meyer, "Comparative genomics provides evidence for an ancient genome duplication event in fish," *Philosophical Transactions of the Royal Society B*, vol. 356, no. 1414, pp. 1661–1679, 2001.

[24] M. G. Reese, G. Hartzell, N. L. Harris, U. Ohler, J. F. Abril, and S. E. Lewis, "Genome annotation assessment in *Drosophila melanogaster*," *Genome Research*, vol. 10, no. 4, pp. 483–501, 2000.

[25] J. Andrews, G. G. Bouffard, C. Cheadle, J. Lü, K. G. Becker, and B. Oliver, "Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis," *Genome Research*, vol. 10, no. 12, pp. 2030–2043, 2000.

[26] M. Parisi, R. Nuttall, P. Edwards et al., "A survey of ovary-, testis-, and soma-biased gene expression in Drosophila melanogaster adults," *Genome Biology*, vol. 5, no. 6, article R40, 2004.

[27] R. S. Singh and R. J. Kulathinal, "Male sex drive and the masculinization of the genome," *BioEssays*, vol. 27, no. 5, pp. 518–525, 2005.

[28] M. B. Coulthart and R. S. Singh, "High level of divergence of male-reproductive-tract proteins, between *Drosophila melanogaster* and its sibling species, *D. simulans*," *Molecular Biology and Evolution*, vol. 5, no. 2, pp. 182–191, 1988.

[29] A. Civetta and R. S. Singh, "High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in Drosophila melanogaster and Drosophila virilis group species," *Journal of Molecular Evolution*, vol. 41, no. 6, pp. 1085–1095, 1995.

[30] S. Jagadeeshan and R. S. Singh, "Rapidly evolving genes of Drosophila: differing levels of selective pressure in testis, ovary, and head tissues between sibling species," *Molecular Biology and Evolution*, vol. 22, no. 9, pp. 1793–1801, 2005.

[31] J. M. Good and M. W. Nachman, "Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis," *Molecular Biology and Evolution*, vol. 22, no. 4, pp. 1044–1052, 2005.

[32] L. M. Turner, E. B. Chuong, and H. E. Hoekstra, "Comparative analysis of testis protein evolution in rodents," *Genetics*, vol. 179, no. 4, pp. 2075–2089, 2008.

[33] M. D. Dean, N. L. Clark, G. D. Findlay et al., "Proteomics and comparative genomic investigations reveal heterogeneity in evolutionary rate of male reproductive proteins in mice (Mus domesticus)," *Molecular Biology and Evolution*, vol. 26, no. 8, pp. 1733–1743, 2009.

[34] A. Civetta and R. S. Singh, "Sex-related genes, directional sexual selection, and speciation," *Molecular Biology and Evolution*, vol. 15, no. 7, pp. 901–909, 1998.

[35] R. S. Singh and R. J. Kulathinal, "Sex gene pool evolution and speciation: a new paradigm," *Genes and Genetic Systems*, vol. 75, no. 3, pp. 119–130, 2000.

[36] R. J. Kulathinal and R. S. Singh, "The nature of genetic variation in sex and reproduction-related genes among sibling species of the *Drosophila melanogaster* complex," *Genetica*, vol. 120, no. 1–3, pp. 245–252, 2004.

[37] R. J. Kulathinal and R. S. Singh, "Cytological characterization of premeiotic versus postmeiotic defects producing hybrid male sterility among sibling species of the *Drosophila melanogaster* complex," *Evolution*, vol. 52, no. 4, pp. 1067–1079, 1998.

[38] P. Michalak and M. A. F. Noor, "Genome-wide patterns of expression in Drosophila pure species and hybrid males," *Molecular Biology and Evolution*, vol. 20, no. 7, pp. 1070–1076, 2003.

[39] W. Haerty and R. S. Singh, "Gene regulation divergence is a major contributor to the evolution of Dobzhansky-Muller incompatibilities between species of Drosophila," *Molecular Biology and Evolution*, vol. 23, no. 9, pp. 1707–1714, 2006.

[40] R. J. Kulathinal and R. S. Singh, "The molecular basis of speciation: from patterns to processes, rules to mechanisms," *Journal of Genetics*, vol. 87, no. 4, pp. 327–338, 2008.

[41] D. C. Presgraves, L. Balagopalan, S. M. Abmayr, and H. A. Orr, "Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila," *Nature*, vol. 423, no. 6941, pp. 715–719, 2003.

[42] A. Civetta, S. A. Rajakumar, B. Brouwers, and J. P. Bacik, "Rapid evolution and gene-specific patterns of selection for three genes of spermatogenesis in Drosophila," *Molecular Biology and Evolution*, vol. 23, no. 3, pp. 655–662, 2006.

[43] V. L. Bauer DuMont, H. A. Flores, M. H. Wright, and C. F. Aquadro, "Recurrent positive selection at *Bgcn*, a key determinant of germ line differentiation, does not appear to be driven by simple coevolution with its partner protein bam," *Molecular Biology and Evolution*, vol. 24, no. 1, pp. 182–191, 2007.

[44] T. Chapman and S. J. Davies, "Functions and analysis of the seminal fluid proteins of male *Drosophila melanogaster* fruit flies," *Peptides*, vol. 25, no. 9, pp. 1477–1490, 2004.

[45] A. Wong and M. F. Wolfner, "Sexual behavior: a seminal peptide stimulates appetites," *Current Biology*, vol. 16, no. 7, pp. R256–R257, 2006.

[46] K. R. Ram and M. F. Wolfner, "Sustained post-mating response in *Drosophila melanogaster* requires multiple seminal fluid proteins," *PLoS Genetics*, vol. 3, no. 12, article e23, 2007.

[47] M. Aguadé, N. Miyashita, and C. H. Langley, "Polymorphism and divergence in the Mst26A male accessory gland gene region in Drosophila," *Genetics*, vol. 132, no. 3, pp. 755–770, 1992.

[48] S. C. Tsaur, C. T. Ting, and C. I. Wu, "Positive selection driving the evolution of a gene of male reproduction, Acp26Aa, of Drosophila: II. Divergence versus polymorphism," *Molecular Biology and Evolution*, vol. 15, no. 8, pp. 1040–1046, 1998.

[49] M. Aguadé, "Positive selection drives the evolution of the Acp29AB accessory gland protein in Drosophila," *Genetics*, vol. 152, no. 2, pp. 543–551, 1999.

[50] D. J. Begun, P. Whitley, B. L. Todd, H. M. Waldrip-Dail, and A. G. Clark, "Molecular population genetics of male accessory gland proteins in Drosophila," *Genetics*, vol. 156, no. 4, pp. 1879–1888, 2000.

[51] W. J. Swanson, A. G. Clark, H. M. Waldrip-Dail, M. F. Wolfner, and C. F. Aquadro, "Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in Drosophila," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 13, pp. 7375–7379, 2001.

[52] J. L. Mueller, K. Ravi Ram, L. A. McGraw et al., "Cross-species comparison of Drosophila male accessory gland protein genes," *Genetics*, vol. 171, no. 1, pp. 131–143, 2005.

[53] T. M. Panhuis and W. J. Swanson, "Molecular evolution and population genetic analysis of candidate female reproductive genes in Drosophila," *Genetics*, vol. 173, no. 4, pp. 2039–2047, 2006.

[54] M. K. N. Lawniczak and D. J. Begun, "A genome-wide analysis of courting and mating responses in *Drosophila melanogaster* females," *Genome*, vol. 47, no. 5, pp. 900–910, 2004.

[55] M. K. N. Lawniczak and D. J. Begun, "Molecular population genetics of female-expressed mating-induced serine proteases in *Drosophila melanogaster*," *Molecular Biology and Evolution*, vol. 24, no. 9, pp. 1944–1951, 2007.

[56] L. A. McGraw, G. Gibson, A. G. Clark, and M. F. Wolfner, "Genes regulated by mating, sperm, or seminal proteins in mated female *Drosophila melanogaster*," *Current Biology*, vol. 14, no. 16, pp. 1509–1514, 2004.

SAGE-Hindawi Access to Research
International Journal of Evolutionary Biology
Volume 2011, Article ID 423938, 16 pages
doi:10.4061/2011/423938

Research Article

A Macroevolutionary Perspective on Multiple Sexual Traits in the Phasianidae (Galliformes)

Rebecca T. Kimball, Colette M. St. Mary, and Edward L. Braun

Department of Biology, University of Florida, P.O. Box 118525, Gainesville, FL 32611, USA

Correspondence should be addressed to Rebecca T. Kimball, rkimball@zoo.ufl.edu

Received 2 October 2010; Accepted 26 February 2011

Academic Editor: Rob Kulathinal

Traits involved in sexual signaling are ubiquitous among animals. Although a single trait appears sufficient to convey information, many sexually dimorphic species exhibit multiple sexual signals, which may be costly to signalers and receivers. Given that one signal may be enough, there are many microevolutionary hypotheses to explain the evolution of multiple signals. Here we extend these hypotheses to a macroevolutionary scale and compare those predictions to the patterns of gains and losses of sexual dimorphism in pheasants and partridges. Among nine dimorphic characters, including six intersexual signals and three indicators of competitive ability, all exhibited both gains and losses of dimorphism within the group. Although theories of intersexual selection emphasize gain and elaboration, those six characters exhibited greater rates of loss than gain; in contrast, the competitive traits showed a slight bias towards gains. The available models, when examined in a macroevolutionary framework, did not yield unique predictions, making it difficult to distinguish among them. Even with this limitation, when the predictions of these alternative models were compared with the heterogeneous patterns of evolution of dimorphism in phasianids, it is clear that many different selective processes have been involved in the evolution of sexual signals in this group.

1. Introduction

The evolution of sexually dimorphic traits, such as armaments, bright coloration, or exaggerated traits has long been of interest to evolutionary biologists, as such traits are often thought to reduce fitness through natural selection (e.g., [1]). However, in spite of the potential negative consequences of these traits, many taxa exhibit not just one, but often multiple, sexually dimorphic traits. In nearly all systems examined, there is evidence that these multiple, sexually dimorphic signals are used in both intrasexual (competition, typically male-male competition) and intersexual (mate choice, typically female mate choice) interactions. Traits traditionally thought to have evolved via competition include those that directly affect the physical competitive ability of males, such as larger body size, antlers, or other armaments, while traits typically associated with mate choice include signals such as coloration or specialized ornaments such as modified feathers that are unlikely to have arisen via natural selection [1]. There is a long-standing expectation that taxa may have multiple sexually dimorphic traits as a result of the separate action of competition and mate choice (e.g., [1]). If intrasexual and intersexual processes are acting independently, we would expect the evolution of dimorphism in a trait classically associated with competition to be unlinked (over evolutionary time and in a given taxa) with the evolution of dimorphism in a signaling trait that might be associated with mate choice.

Although the evolution of a sexually dimorphic trait via competition and a second trait via mate choice may be expected within the same taxon, how can we explain the presence of multiple sexually dimorphic traits that are likely to have evolved through the same selective pressure? In the case of competition, where each individual trait may confer a net competitive advantage (e.g., each type of armament increases male competitive ability in an additive fashion), it would be predicted that multiple traits should accumulate over evolutionary time. In contrast, assessing multiple signals used for intersexual interactions is likely to be costly for female receivers, making it difficult to understand why

females would bear the cost of assessing multiple signaling traits in mate choice. Thus, the majority of theories to explain multiple sexual signals has developed in the context of mate choice (reviewed in [2, 3]).

Because signals can be costly to assess, produce, and maintain, the development of theory to explain the gain and maintenance of multiple traits used in intersexual signaling has focused on the identification of plausible mechanisms to explain why multiple signals, rather than a single signal, are used by a receiver. Candolin [2] reviewed seven classes of explanations including (1) different messages are conveyed by the different signals, (2) that there are back-up or redundant signals, (3) some signals are used primarily in species recognition, (4) some signals are unreliable or Fisherian cues and thus not informative, (5) two signals are received more efficiently than one (receiver psychology), (6) different sensory environments favor different signals, and (7) intersexual conflict/antagonistic coevolution results in a dynamic in which older signals become uninformative antes and newer signals arise. In his more recent review, Bro-Jørgensen [3] also summarizes this theory, albeit slightly differently. In addition to the explanations discussed by Candolin [2], Bro-Jørgensen [3] expands the list to include the rare male effect (i.e., the hypothesis that an unusual signal is always advantageous) and proposes a novel explanation for the advantage of multiple sexual signals: the idea that fluctuating signaling environments favor different signals at different times. Bro-Jørgenson [3] argues that since dynamic social and physical signaling environments are the rule, rather than the exception, this dynamic hypothesis may be a more general explanation for multiple signaling systems than what has been proposed previously.

As can be seen, there is a wealth of plausible hypotheses available to explain the use of multiple sexual signals (Table 1). A critical task at present is to evaluate whether or not there is any dominant mechanism at work across taxa, or whether many different hypotheses will be necessary to understand the ubiquity of multiple sexual signals. The existing hypotheses are framed in a strongly microevolutionary context. As a result, we have detailed predictions about the expression of multiple traits and their use and costs within species (or populations) [2], but much less clarity about the expectations at the macroevolutionary (above species) scale regarding patterns of trait diversification, loss, or the association of traits in lineages where multiple sexual signals occur. However, looking across lineages over evolutionary time should clarify whether there might be one or a few mechanisms, or whether multiple hypotheses are needed to explain the broad patterns seen across taxa.

In an effort to relate this existing (mostly microevolutionary) theory to patterns quantified at the macroevolutionary scale, we have reviewed the major classes of hypotheses available to explain the existence of multiple sexual signals (Table 1) and extended these (when necessary) to make predictions at the macroevolutionary scale. Although competition is expected to lead to the accumulation of traits associated with physical combat over evolutionary time (as described above), we also recognize that intrasexual competition can lead to the evolution of signals that indicate competitive advantage (e.g., [14, 15]). Just as the assessment of multiple signals is likely to be costly for females choosing mates, the assessment of multiple signals indicating competitive advantages in intrasexual interactions are also likely to be costly. Thus, we argue that many of the hypotheses put forward to explain multiple cues in intersexual communication (mate choice) should also apply to the signals used in intrasexual interactions as well (e.g., [16]).

In extending the predictions of the existing theory to the macroevolutionary scale, we made the simplifying assumptions that signals of the same mode (e.g., coloration) are more likely to be redundant in the information they convey, whereas the presence of multiple signals of different modes (e.g., color patterns versus specialized structures) are more likely to communicate multiple different messages or to act in concert. These assumptions can be logically argued and are supported by data (e.g., [17, 18] but see [19]).

The upper portion of Table 1 focuses on the models that deal with signaling in general (applicable to both mate choice and competition), summarizes the major classes of hypotheses to explain multiple sexual signals, and provides our macroevolutionary expectations for sexual dimorphism and the pattern of trait accumulation and loss over time from each of models. In the lower portion of Table 1, we discuss hypotheses that are specific just to mate choice. Specifically, this includes rare male advantages in mate choice [9], antagonistic coevolution in sexual signaling [10, 11], and Fisher Runaway processes [4, 12, 13]. These models do not make specific predictions about the types of traits that should evolve but have been framed in a dynamic context and thus make more explicit macroevolutionary predictions than the majority of hypotheses.

To evaluate support for the macroevolutionary predictions of the models, we used a comparative framework to examine the distribution of sexual dimorphism in multiple traits in the family Phasianidae (chickens, pheasants, and allies). Although this is a very broad-brush approach, it has the potential to highlight general patterns and trends. Thus, this can be viewed as the first step for more detailed studies in specific clades that further refine our understanding of the evolution of multiple sexual traits.

2. Methods

2.1. Study System. The avian family Phasianidae (within the order Galliformes) contains some of the best-studied avian species, including the domestic chicken (*Gallus gallus*) and turkey (*Meleagris gallopavo*), both of which now have complete genomes available [21, 22]. Species within the family exhibit a wide range of variation in size and external morphology [23]. Species range greatly in size, varying from relatively small (~35 g in *Coturnix chinensis*) to quite large (up to 5 kg in the turkey). While many species have little to no sexual size dimorphism, in others males may be more than 30% larger than females. Some species also exhibit one to several spurs on the tarsi, though these are absent in many species [24]. In some cases, spurs are found in both sexes, while in others they are restricted to males or males have

Table 1: Major hypotheses to explain multiple sexual signals and their predictions extended to the macroevolutionary scale. Terminology largely follows that of Candolin [2] and Bro-Jørgensen [3]. Macroevolutionary predictions inferred from the logic of the hypothesis when not specifically specified in the literature with the assumption that signals of a similar mode are more likely to be redundant in the information they convey, whereas multimodal signals are more likely to communicate multiple messages or to act in concert. Patterns are evaluated for traits involved in mate choice, signaling traits (MC), traits involved in mating competition (Com), and across both sets of characters (Total). Since we were interested in falsification of the hypotheses, we compared our results with the patterns that would contradict (or refute) the hypothesis. Y (yes) indicates that our data refute the hypothesis, N (no) indicates our results cannot refute the hypothesis, and "—" means the data cannot be used to address the hypothesis.

Hypothesis	Description	Extending the hypothesis to the macroevolutionary scale	Macroevolutionary patterns consistent with hypothesis	Macroevolutionary patterns that refute the hypothesis	For each prediction, does our data refute the hypothesis?		
Sexual Signaling (models that can be applied to signals used in inter- and intrasexual signaling)					MC	Com	Total
Multiple messages [4, 5]	Each signal conveys unique information	Selection on each signal is independent, thus the evolutionary trajectories should be as well. Further, we do not expect consistent changes in the number of messages over evolutionary time.	(i) No net accumulation of signals within the same mode of signal	(i) Consistent increase/decrease in number of traits used in signaling	(i) Y	(i) N	(i) Y
			(ii) Independent patterns of loss and gain in different modes	(ii) Increases in traits within a mode	(ii) N	(ii) —	(ii) N
			(iii) Increase in the number of signaling modes	(iii) Correlated trait evolution	(iii) Y	(iii) —	(iii) Y
Redundant/backup signals [4, 5]	All signals convey the same information, the combination reduces errors in communication.	Signals will accumulate over evolutionary time.	(i) Net gain in signaling traits	(i) Net loss of signals	(i) Y	(i) N	(i) Y
			(ii) Net increase in signaling traits within a mode	(ii) Increases in signaling mode that are not accompanied by increases within a mode	(ii) N	(ii) N	(ii) N
Interacting signals: *in this category are lumped receiver psychology hypotheses, alerting signals, and emergent messages*, [6–8]	The combination of signals (i.e., their interaction) enhances the communication of a single message.	Signals will accumulate over evolutionary time. However, once an effective complement of signals has evolved no further dynamics are predicted.	(i) Net gain in signaling traits	(i) Net loss of signals	(i) Y	(i) N	(i) Y
			(ii) Net increase in signaling modes	(ii) Replacement of signaling traits	(ii) N	(ii) N	(ii) N
			(iii) Correlated increases in trait modes	(iii) Net increases in signaling traits within a mode.	(iii) N	(iii) —	(iii) N

Table 1: Continued.

Hypothesis	Description	Extending the hypothesis to the macroevolutionary scale	Macroevolutionary patterns consistent with hypothesis	Macroevolutionary patterns that refute the hypothesis	For each prediction, does our data refute the hypothesis?		
Multiple/dynamic sensory environments [3]	Because organisms move frequently among different signaling environments (physical and social), alternative signals are required to convey either the same or different messages.	Conveying even a single message across environments favors signaling in more modes; if individuals experience multiple environments, these might arise in a correlated manner; multiple messages increase expected diversity of signals and modes. Loss and gain of signals is expected with environmental change.	(i) Correlated accumulation of signaling traits, both within and among modes (ii) Net increases in signaling traits within a lineage followed by stasis (iii) Greater signaling trait diversity in taxa with wider geographic distributions (iv) Ongoing gain and loss of traits (i.e., a dynamic equilibrium)	(i) Net gain or loss of signals (ii) No environmental correlates of trait evolution	(i) Y (ii) —	(i) N (ii) —	(i) Y (ii) —
Signaling in mate choice (models that specific incorporate intersexual dynamics)					MC		Total
Rare male effects [9]	Females prefer novel/rare males.	There should be steady replacement of signaling traits over evolutionary time leading to a dynamic multitrait equilibrium. Signaling mode should have no effect on this process.	(i) No net accumulation of signaling traits or modes. (ii) Replacement of particular signals while maintaining the same level of dimorphism	(i) Net trait or mode accumulation or loss (ii) Net diversification within a mode (iii) Correlated trait evolution	(i) Y (ii) N (iii) Y		(i) N (ii) N (iii) Y

Table 1: Continued.

Hypothesis	Description	Extending the hypothesis to the macroevolutionary scale	Macroevolutionary patterns consistent with hypothesis	Macroevolutionary patterns that refute the hypothesis	For each prediction, does our data refute the hypothesis?	
Sexually antagonistic coevolution [10, 11]	Informative signals will degrade in their information content, due to the conflicts between males and females	Informative signals will arise and remain as antes after their information content has been degraded. Thus they remain alongside newly evolved, informative signals	(i) Net accumulation of traits	(i) Trait loss	(i) Y	(i) Y
			(ii) No differences in accumulation rates according to mode	(ii) No net trait accumulation	(ii) N	(ii) N
				(iii) Correlated trait evolution	(iii) Y	(iii) Y
				(iv) Differences in trait accumulation within and among modes	(iv) Y	(iv) Y
Fisher Runaway [4, 12, 13]	Traits arise via a Fisher Runaway process accumulate in lineages	Traits arise and become fixed in species via a Fisher process and therefore tend to accumulate in lineages.	(i) Net accumulation of traits	(i) Trait loss	(i) Y	(i) Y
			(ii) No differences in accumulation rates according to mode	(ii) No net trait accumulation	(ii) Y	(ii) Y
				(iii) Correlated trait evolution	(iii) Y	(iii) Y
				(iv) Differences in trait accumulation within and among modes	(iv) Y	(iv) Y

more spurs than females (e.g., females typically have one, while males may have up to three).

Plumage is also quite variable, ranging from species that are cryptically colored with no specialized feathers or extravagant coloration in either sex, to others in which males are strikingly colored, exhibit high levels of iridescence, or have specially modified feathers [23]. Specialized feather types are found in many species, such as greatly elongated or modified tails as well as head modifications such as crests or ear tufts; in some cases these can be found in both sexes, while in others they are either restricted to males or are more exaggerated in males. Additionally, some taxa have very uniquely modified feathers, typically found only in males, such as the highly modified tail coverts of peafowl (*Pavo* spp.), the elongated hackle and saddle feathers of most junglefowl species (*Gallus* spp.), or the elongated secondary wing feathers in the argus pheasant (*Argusianus argus*).

Some species in the Phasianidae (and in galliformes overall) also have regions on the head and neck that lack plumage, instead showing the "flesh" underneath [25]. These fleshy traits can vary from simply bare regions around the eye to the complete absence of feathers on the head and neck of the turkey and specialized structures such as the comb and wattle of junglefowl. In some phasianid species, these fleshy traits have the ability to be "erected" in which the trait can greatly change its size and appearance in a short time, and then be retracted (often being barely visible in the nonerect state). Taxa that are able to rapidly erect these fleshy structures form a clade [25, 26], suggesting that the physiological basis for trait erection may have evolved a single time. Fleshy traits are typically found in both sexes when they are present in a species. However, the traits can still be dimorphic, since males often exhibit larger or more brightly colored fleshy regions than females. Moreover, the ability to erect these traits is found only in males.

In part due to the large number of ornamental or exaggerated traits, the role of both competition and mate choice has been extensively studied in several species in this family. Results of these studies have varied and highlight the role of sexual selection in this group. For example, in red junglefowl (the ancestor to domestic chickens, also *Gallus gallus*), the size of the male comb appears to be most consistently used by females in mating decisions (e.g., [27]), and along with body size, correlates with dominance [28]. Manipulation of social structure indicates that females prefer males with large combs, even when these are not the dominant male [29]. In the Indian peafowl (*P. cristatus*), females prefer males with more ocelli (eye-like spots) in the modified tail coverts that form the train (e.g., [30, 31]), though this may not be consistent across populations [32]. Both train and tarsus length appeared to be involved in competition in the Indian peafowl [31]. In wild turkey, both competition and mate choice favor males with long snoods [33, 34], the fleshy protrusion above the beak. Studies in the ring-necked pheasant (*Phasianus colchicus*) and various species of grouse have been less clear, but implicate various aspects of morphology and display behaviors in mate choice (e.g., [35–38]).

Early studies on relationships within the Phasianidae showed little consensus (reviewed by [20, 39]). The family exhibits limited osteological variation (e.g., [40, 41]), and the group appears to have undergone a relatively rapid radiation [42], both of which probably contribute to the difficulties associated with resolving phylogenetic relationships. More recent studies using both whole mitochondrial sequences, sequences from multiple nuclear loci, as well as insertions of transposable elements, have resulted in a better-resolved and more stable phylogeny [25, 43–49]. Specifically, these studies agree on several key points. First, the grouse and turkeys, traditionally separated into separate subfamilies or families, nest within the Phasianidae. Second, the traditional grouping of pheasants (dimorphic, highly ornamented, large bodied) and the partridges and old world quail (cryptic, little ornamentation, small bodied) into separate clades within the family is erroneous. Instead, the characteristics that have traditionally been used to define pheasants and to differentiate them from partridges have been suggested to have arisen independently in multiple lineages [42], though this hypothesis has not been tested explicitly.

The suggestion that highly ornamented and dimorphic "pheasants" repeatedly evolved from more monomorphic "partridge" ancestors suggests that studying the Phasianidae may provide insight into the forces that have led to complex patterns of sexual dimorphism and the evolution of multiple male secondary sexual traits. Recently, it has been shown that the evolution of sexual size dimorphism within the Phasianidae may be related to mating system (particularly lekking; [50]), though other types of traits that are potentially involved in sexual selection (e.g., plumage, fleshy traits, and spurs) were not considered. Herein we use a comparative approach to examine the evolution of dimorphism in morphological traits that have been implicated in sexual selection, mapping the gain and loss of dimorphism across the family to generalize about the evolution of multiple sexual signals.

2.2. Phylogenetic Estimation. To obtain a taxon-rich phylogeny, we used a supermatrix analysis (a large-scale data matrix) rather than a supertree analysis (meta-analysis of previous work) [51, 52]. To ensure all gene partitions were represented by a good range of taxa in the supermatrix, we restricted our analyses to the six data partitions analyzed by Kimball and Braun [25], which included two mitochondrial coding regions (cytochrome *b* [CYB] and NADH-ubiquinone oxidoreductase chain 2 [ND2]) and four nuclear introns (ovomucoid intron G [OVMintG], β-fibrinogen intron 7 [FGBint7], pterin-4-α-carbinolamine dehydratase intron 3 [PCBD1int3], and rhodopsin intron 1 of [RHOint1]). Kimball and Braun [25] obtained a complete data matrix comprising all of these gene regions for 44 galliformes, representing all galliform families but placing an emphasis on the Phasianidae. For the present study, we extended the taxonomic coverage by retrieving all sequences homologous to the six focal loci in the NCBI database using BLASTN [53] and filtering the sequences to retain a single representative per species. Although the present study focused on Phasianidae, all nonphasianid galliform species that were available for these partitions were included as

outgroups. CYB and ND2 sequences were equal in length (with the exception of modest variation in the CYB stop codon (see [54]), making them straightforward to align. The nuclear intron sequences were added to the Kimball and Braun [25] alignment and then optimized by eye. We were able to add data to all partitions in the Kimball and Braun [25] matrix. This resulted in a data matrix in which 51.8% of the cells were filled; all taxa (170 species) were represented by some mitochondrial data and 99 taxa (70 of which were phasianids) also included nuclear intron data.

Maximum likelihood (ML) phylogenetic analyses were performed using the GTRMIX method implemented in RAxML 7.0.4 [55], which performs searches using the GTRCAT [56] method to accommodate among-sites rate heterogeneity followed by a final optimization using the GTR+Γ model. ML searches were conducted using 25 distinct randomized maximum parsimony (MP) starting trees. RAxML analyses were conducted both with and without partitioning the data. Partitioned analyses were conducted using 10 data partitions, one for each nuclear intron and three for each of the mitochondrial coding regions (one partition for each of the codon positions). Bootstrap analyses used the GTRCAT method and 500 replicates.

To accommodate uncertainty in phylogenetic reconstruction when examining patterns of character change, we used MrBayes 3.1 [57, 58] to generate a set of trees that accommodates the uncertainty in both topology and branch lengths. For these analyses, the data were partitioned as described above for RAxML. For each partition, we used the AIC to select the best fitting model that is implemented in MrBayes. We ran four chains (three of which were heated) for 7.5×10^7 generations and discarded the first 1×10^7 generations. We sampled every 1×10^4 generations, resulting in a set of 6500 trees that were used for trait reconstruction (see below).

2.3. Trait Coding and Analyses. The focus of this study was to examine gains and losses in sexual dimorphism, so all traits were coded as binary characters, where 0 = monomorphic and 1 = dimorphic. Thus, a monomorphic score (0) could mean a trait was absent in both sexes or present and similar in both sexes; a dimorphic score (1) could mean either that the trait was present in both sexes but larger or modified in one sex or present in one sex and absent in the other. Most information used to code traits was obtained from Madge and McGowan [23], though this source was supplemented with information in Johnsgard [59–61]. Although the Phasianidae were the focus of this study, we also coded data for the five species of guineafowl (Numididae) and six available New World quail (Odontophoridae) represented in the sequence dataset to polarize traits.

A total of nine binary characters were scored for dimorphism (Table 2). Five of these corresponded to plumage differences: (1) plumage color, (2) tail length, (3) head ornaments (e.g., crests and ear tufts), (4) other plumage differences (e.g., iridescence, unusual plumage features not covered in other traits), and (5) whether the male plumage was highly elaborated compared to females. The remaining four traits included (6) fleshy traits, (7) spurs, (8) wing size, and (9) overall body size. Body size is difficult to measure, and we examined several possibilities. First, we looked at a qualitative difference based on descriptions in Madge and McGowan [23]. Second, we took weight measurements for males and females from Madge and McGowan [23], although some of these were based on single individuals and/or captives whose weights may not be typical. Third, we considered the weights from Lislevand et al. [50], which included values that were based on three or more individuals, primarily measured during the breeding season. For the two weight measures, we considered species dimorphic if there was at least 10% difference between the sexes [20, 50]. Since the patterns using weight data were similar to those using qualitative measurements (including in the relative differences between loss and gains), but were for many fewer species, we only report those results using the first approach.

Several additional characters were generated from this initial set, as defined in Table 2. We also scored whether species were dimorphic in any plumage trait (10) and dimorphic in any competitive trait (11). To highlight species that exhibited strong dimorphism in either plumage or competitive traits, species were scored for two additional binary characters: (12) high plumage dimorphism and (13) high competitive dimorphism. To look at overall patterns, we also made several composite traits, including totaldimorphism (character 14; sum of the state of characters 1–9), total signal dimorphism (character 15; sum of characters 1–6), total competitive dimorphism (character 16; sum of characters 7–9), and total plumage dimorphism (character 17; sum of characters 1–5).

To assess overall rates of change and relative rates of gains and losses, we used maximum likelihood reconstruction in BayesTraits [62] using the 6500 phylogenetic trees from our Bayesian analysis (above). From this, we estimated the transition rate and the bias (the ratio of the gain rate to the loss rate) for each binary character. Although the ML estimates obtained using this procedure do not represent posterior distributions in the Bayesian sense, this approach captures the impact of phylogenetic uncertainty on our estimates and it has the advantage of being independent of priors. We report the median estimate as well as the top 2.5 percentile and the bottom 2.5 percentile of the distribution (95% range) for both the rate and bias. To determine whether there was a significant difference between gains and losses in the binary characters (1–13), we used Mesquite [63] and the ML tree from RAxML to test whether a two-rate model (gains $\neq$ losses; AsymmMk) was significantly better than a one-rate model (gains = losses) with a likelihood ratio test, assuming that twice the difference between models in their log likelihood is χ^2 distributed with 1 degree of freedom.

To understand the overall patterns in the ancestor, we used the output from BayesTraits [62] to assess the probability of being dimorphic for characters 1 to 9. We did this at three possible ancestral states: (1) the ancestor of the outgroup taxa (Numididae and Odontophoridae) and the Phasianidae, (2) the ancestor to the Phasianidae, and (3) the

Table 2: Characters scored as monomorphic or dimorphic.

	Character	Description
1	Plumage-Color	Dimorphic scored for obvious differences between the sexes in coloration (subtle differences, such as slightly brighter in males or more extensive spotting in one sex were scored as monomorphic)
2	Plumage-Tail Length	Dimorphic scored when tail (or tail coverts) was much longer or modified in males relative to female tail
3	Plumage-Head Ornaments	Dimorphic scored if ear tufts, crests, or capes either present in males or larger in males
4	Plumage-Other	Dimorphic scored if male plumage exhibits iridescence or modified feathers not covered in other categories
5	Plumage-Elaboration	Dimorphic scored if the male plumage pattern or color is dramatic and very obvious relative to female plumage
6	Fleshy Traits	Dimorphic scored if one sex larger or obviously brighter than the other; since the goal was to explore sexual selection on these traits, the appearance of the trait when displayed was scored
7	Spurs	Dimorphic scored if males had more spurs than females, or in a few cases male spurs noticeably longer than female spurs
8	Wing Length	Dimorphic if at least a 5% difference between the sexes (following Höglund [20]). Used average wing length if given; if only a range was given, then the midpoint of the range was used for calculations)
9	Body Size	Dimorphic if species noted as having larger males or smaller females; scored as monomorphic if differences were noted as slight (weight was not used as many measures in the literature taken from captive individuals and/or dead specimens whose weight may not be representative)
10	Any Plumage	Scored as one if dimorphic in any of the five plumage traits
11	Any Competitive	Scored as one if dimorphic in either spurs, wing length, or body size
12	High Plumage	Scored as one if the species exhibited dimorphism in at least three of the five plumage characters
13	High Competitive	Scored as one if the species exhibited dimorphism in at least two of the three competitive characters
14	Total	Sum of all characters scored as dimorphic in a species
15	Total Signal	Sum of the plumage and fleshy trait dimorphism, both of which are used in visual displays
16	Total Competitive	Sum of spur, wing, and body size characters scored as dimorphic
17	Total Plumage	Sum of the five plumage characters scored as dimorphic

ancestor to the core Phasianidae (excluding the Arborophilinae from the remaining phasianids, Figure 1). For each trait, we summed the probability of being dimorphic for all nine characters (character 14), total signal dimorphism (character 15), total competitive (character 16), and total plumage (character 17) across all trees. We then calculated the median value and the 95% range.

To determine whether changes in one type of trait (e.g., plumage) were associated with transitions in another (e.g., fleshy traits), we looked for correlations using correlated changes [65] as implemented in Mesquite 2.72 [63]. For each test, we used 10 likelihood iterations and ran 100 simulations to establish the distribution. We ran this test with several combinations of traits. First, we examined whether there was a relationship between having dimorphism in any plumage trait (10) with any competitive trait (11). We also tested all pairwise combinations of high plumage dimorphism (12), high competitive dimorphism (13), and fleshy (6) characters to examine whether there was an association between being highly dimorphic in two different types of traits.

3. Results and Discussion

3.1. Phylogeny. The combined alignment of 170 taxa (127 phasianids) was 6296 bp in length, of which 2184 bp were mitochondrial and 4112 bp were nuclear. The nuclear intron data included some regions that were difficult to align or present only in a small number of taxa (e.g., the mountain quail (*Oreortyx pictus*) had an autapomorphic insertion that corresponded to a 579-bp segment of an endogenous retrovirus (ERV)). The nuclear intron data also included two microinversions. None of the rare genomic changes (the ERV insertion and microinversions) united controversial groups: two were autapomorphic and one united the genus *Gallus*. After the difficult-to-align regions and rare genomic changes were excluded from analyses, there were 2828 bp of nuclear intron data available for the phylogenetic analyses. For the data that were analyzed, there were 1298 variable mitochondrial sites (1159 were parsimony-informative) and there were 1877 variable nuclear intron sites (1451 were parsimony-informative).

Our analyses of the supermatrix yielded trees (Figure 1) that were more similar to other recent analyses of galliform phylogeny based upon sequence data (e.g., [25, 26, 43, 44, 46, 49, 66, 67]) than to the available galliform supertrees [68, 69]. Points of agreement between the supermatrix trees and other studies include the strong support for dividing Galliformes into five major clades: Megapodidae (megapodes), Cracidae (chachalacas, guans, and curassows), Numididae (guineafowl), Odontophoridae (New World quail), and Phasianidae (pheasants and partridges). Relationships among these groups (especially the position of the New World quail) are variable in supertrees

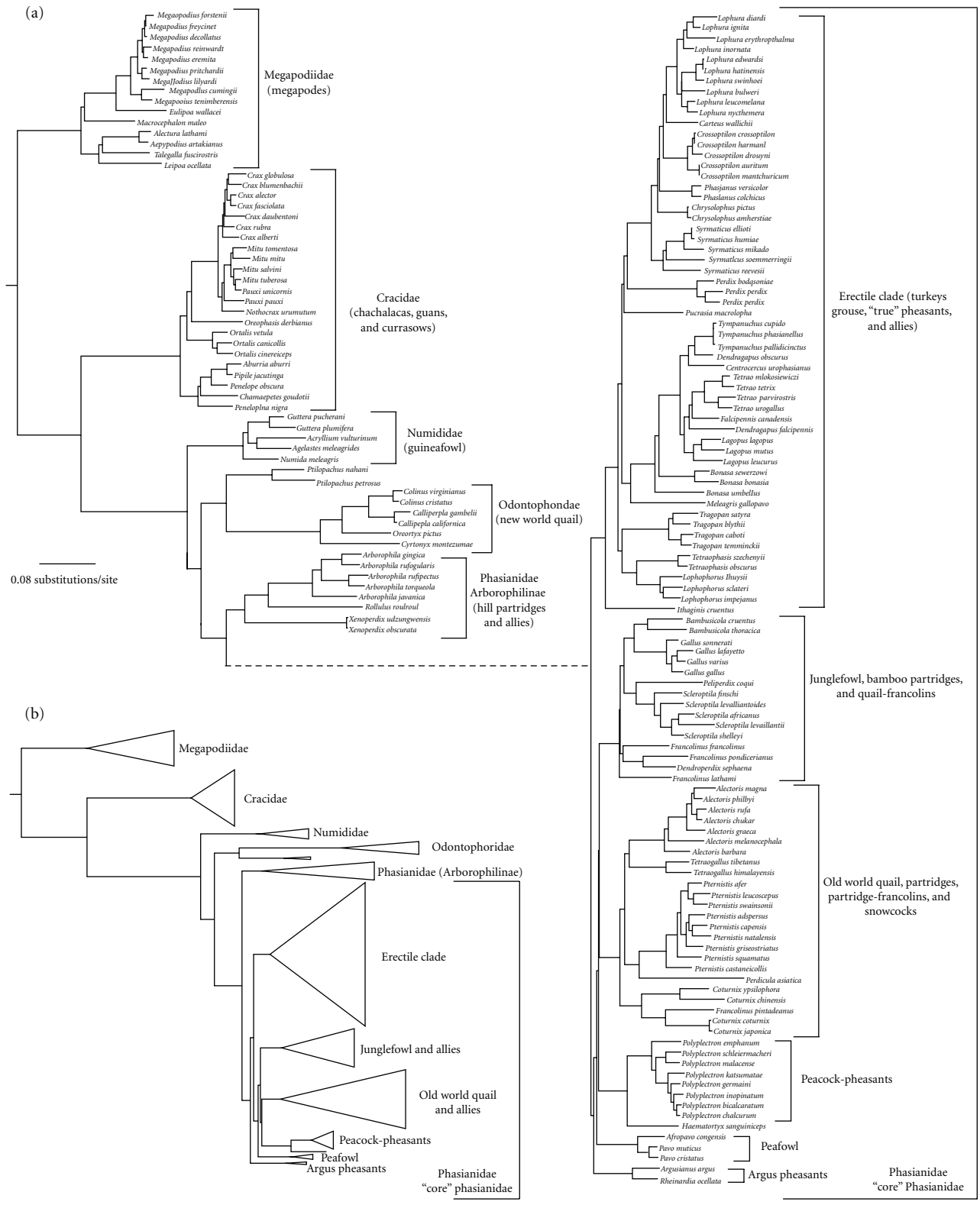


Figure 1: Supermatrix phylogeny of Galliformes. Estimates of phylogeny for galliformes were obtained by unpartitioned (a) and partitioned (b) ML analyses. The root of Galliformes was placed between Megapodidae and other galliformes, a position consistent with large-scale studies that included both outgroups and members of all five families [43, 64]. Stars indicate that the adjacent branch had strong (≥95%) bootstrap support, whereas dots nodes indicate that the adjacent branch had moderate (≥70%) bootstrap support. The topology and degree of support for the partitioned Bayesian MCMC analysis were virtually identical to that of the partitioned ML analysis (if the tendency for Bayesian posterior probabilities to exceed ML bootstrap values is considered).

Table 3: Patterns of evolutionary change using values estimated from BayesTraits. *indicates traits where there was a significantly better fit to a two-rate model than a one-rate model.

		Median rate	95% Range for rate	Median gains/loss	95% Range for gains/loss
1	Plumage-Color	1.65	1.32–2.11	0.78	0.45–1.00
2	Plumage-Tail Length	1.34	1.04–1.57	0.26	0.17–0.42
3	Plumage-Head Ornaments	3.48	2.73–4.36	0.25	0.21–0.33*
4	Plumage-Other	1.51	1.20–1.88	0.17	0.11–0.23*
5	Plumage-Elaboration	2.41	1.87–3.49	0.22	0.17–0.31*
6	Fleshy Traits	0.88	0.57–1.08	0.57	0.26–1.38
7	Spurs	0.39	0.32–0.48	0.84	0.32–1.77
8	Wing Length	18.40	3.68–891.1	1.21	0.77–1.26
9	Body Size	1.67	1.13–2.36	1.55	0.83–2.43
10	Any Plumage	1.54	1.10–1.79	0.92	0.41–1.21
11	Any Competitive	0.92	0.79–1.18	7.58	4.75–9.57
12	High Plumage	1.71	1.36–2.07	0.14	0.10–0.24*
13	High Competitive	2.08	1.78–2.90	1.25	0.93–1.81

[68, 69] but relationships among these clades in our tree are consistent with studies based upon independent gene regions [25, 44, 46, 67], and the combination of both morphological and molecular data [43].

The deepest divergence within the Phasianidae was between two well-supported clades, the Arborophilinae [43] and a large "core phasianid" clade. Arborophilinae includes hill partridges (*Arborphilia* spp.), crested wood-partridge (*Rollulus roulroul*), and recently discovered African forest partridges (*Xenoperdix* spp.), consistent with results from total evidence studies [43]. The core phasianid clade included the strongly supported "erectile clade," which includes a number of taxa that are able to rapidly erect fleshy traits [25, 26]. The erectile clade has also been found in other recent studies [43, 67]. The majority of the remaining core phasianids form a clade in both analyses presented here, with the exception of the argus pheasants (*Argusianus argus* and *Rheinardia ocellata*) which form the sister group of the remaining core phasianids in the partitioned analysis. This agreement with prior estimates of galliform phylogeny (e.g., [25, 43, 49, 67]) indicates the trees are sufficiently accurate to allow rigorous examination of the evolution of sexual dimorphism in this order.

Despite the congruence between our supermatrix trees and prior estimates of galliform phylogeny, the differences between trees obtained with and without partitioning (Figure 1) raise the question of whether the core phasianids can be divided into two large clades (i.e., the erectile clade and a second clade that comprises the remaining core phasianids) or three clades (the erectile clade, the argus pheasants, and the remaining core phasianids). Recent phylogenetic studies of galliformes (e.g., [25, 26, 43, 47–49, 67]) have reached contrasting conclusions regarding this question, but there is no consistency among studies. Moreover, the relevant branches are short, and support for the conflicting relationships is limited both here and in previous studies. Given this region of conflict, we mapped traits on both trees (Figure 1), though only the results from the unpartitioned tree are shown (using the partitioned tree yielded similar results).

3.2. Character Evolution. As can be seen when looking at total dimorphism (character 14), species that are highly dimorphic (e.g., dimorphic in seven or more characters) can be found throughout the Phasianidae (Figure 2). While some clades are comprised of species that are generally highly dimorphic, and others of species that exhibit little dimorphism, there are several cases in which taxa with relatively high levels of dimorphism are sister to taxa that have very low levels of dimorphism. As expected given this level of variation throughout the family, all nine dimorphism characters exhibit both gains and losses, and many are evolving at relatively high rates (Table 3). The lowest rate of change occurs in spur dimorphism, with the other traits typically exhibiting rates of change that are between 2- to 10-fold greater than for spur dimorphism (although wing length shows a much greater relative rate of change, topological uncertainty had a relatively large impact upon the estimates on this character, and the 95% range is so broad, it is difficult to compare this to the other characters).

Although all characters exhibited gains and losses, the relative rate of gains and losses differed among characters (Table 3). Signal traits (plumage and fleshy traits) exhibited a greater rate of loss than gain (gain/loss values <1), while most competitive characters showed a greater rate of gains than losses (the exception is spurs, where the median value is near one, indicating the rate of gains is similar to that of losses). However, in most cases, the difference between the rate of gains and losses was not substantial, and the estimated 95% range of gains to losses across the Bayesian trees frequently included one. Although the 95% ranges that we report reflect the impact of topological uncertainty upon the ML estimate (not the confidence interval on the ML estimate given a fixed topology), the results based upon the 95% range were typically consistent with the results of the likelihood ratio test when it was used to select the best fitting model. All traits with a 95% range that included one did not show a significantly better fit to a two-rate model (gains and losses occur at different rates) relative to a one-rate model (equal gain and loss rates). In contrast, many traits in which the 95%

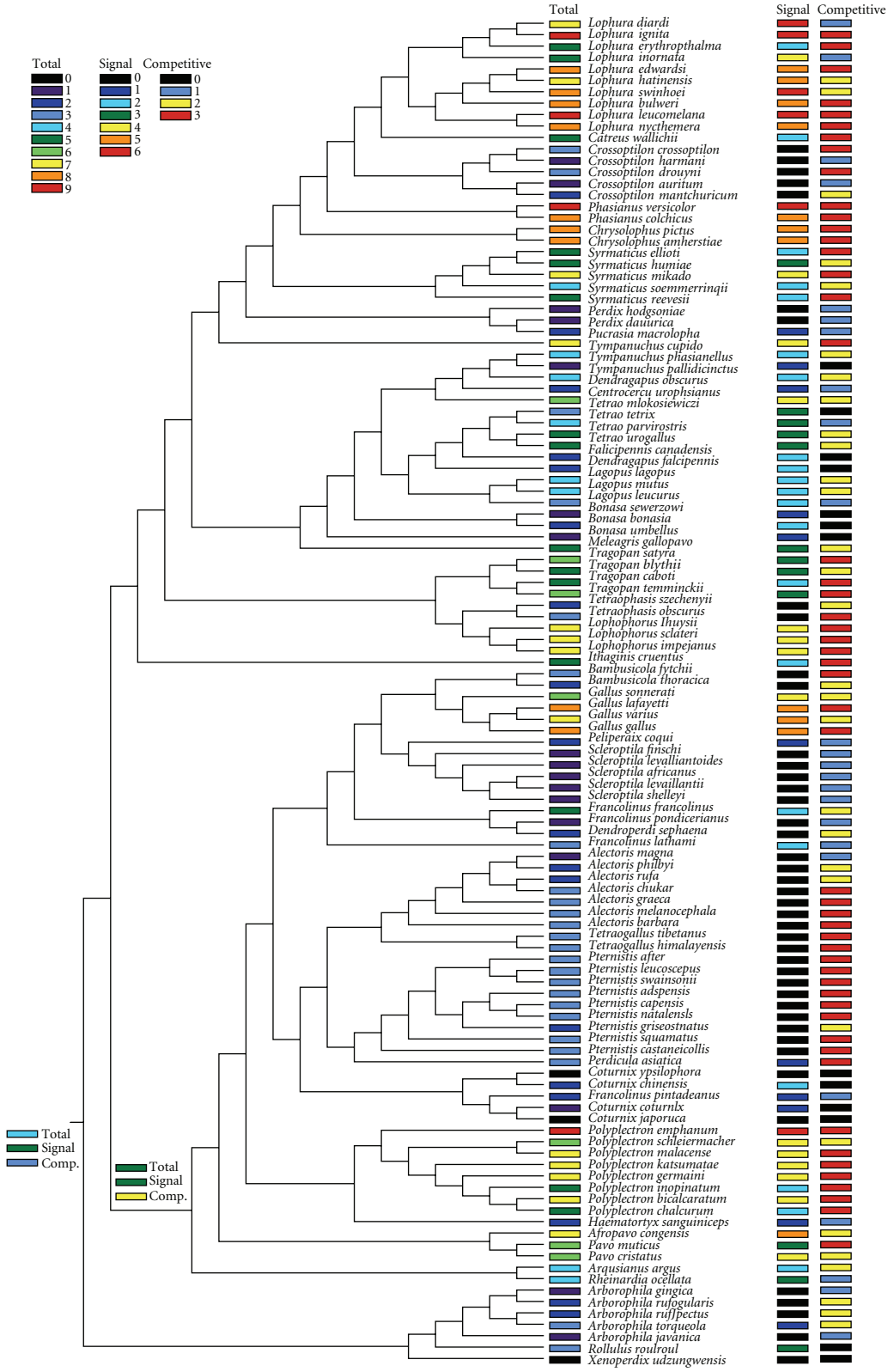


Figure 2: Patterns of evolution for sexual dimorphism in the Phasianidae. Values for total dimorphism (character 14), total signal dimorphism (character 15), and total competitive dimorphism (character 16) in extant taxa are presented to the right of the unpartitioned ML topology. The estimated ancestral states for these dimorphism characters are presented for the Phasianidae and the core Phasianidae. Ancestral state estimates reflect the median of the set of ML estimates obtained using the trees sampled from the Bayesian MCMC analysis.

range did not include one showed a significantly better fit to a two-rate model (Table 3; note that the likelihood ratio test for character 2 approached a significant improvement with a two-rate model; 2Δ log likelihood = 3.70, $2\Delta_{crit}$ = 3.84, df = 1, P = .054).

Directional selection for continued elaboration of sexually dimorphic traits is assumed in many models of sexual selection (e.g., [1]), suggesting that loss of dimorphism should be rare. However, more recent studies have recognized that loss of dimorphism can occur [66, 70]. Our study further emphasizes this point, as the rate of losses exceeded gains in most of the nine dimorphism characters (Table 3). Moreover, even for characters where the rate of gains exceeded that of losses (i.e., characters 8 and 9), the rate for gains was less than twice that of losses. In contrast, the rate of losses was three to four times greater than that of gains for the several of the other characters (i.e., characters 2 through 5).

This pattern of loss suggests that gains in dimorphism likely occurred early in the evolution of the Phasianidae (or earlier), making it more likely that subsequent transitions were losses rather than gains. The summed probability of being dimorphic for all nine characters was low in the common ancestor of Numididae, Odontophoridae, and Phasianidae (Table 4), but increases in the common ancestor of the Phasianidae and is even larger in the ancestor to the core phasianids (Table 4, Figure 2). Similar patterns are seen when summing the signal and competitive traits separately (Table 4, Figure 2), suggesting this early gain of dimorphism involved both types of traits. These results suggest that the characteristics that have traditionally been used to define pheasants and differentiate them from partridges may not have arisen independently in multiple lineages, as suggested by Kimball et al. [42]. Instead, the common ancestor of the core phasianids may have been a dimorphic and "pheasant-like," with the partridge lineages arising multiple times due to the loss of dimorphism.

The only specific traits for which there is a low (<0.05) probability that the ancestor to all phasianids was dimorphic were fleshy traits (character 6) and spurs (character 7). At the base of the core phasianids, only dimorphism in fleshy traits appears very unlikely, while the state for spurs is more equivocal (probability of dimorphism is 0.39). The absence of dimorphism in fleshy traits is consistent with hypotheses that fleshy traits evolved through natural selection as a mechanism of thermoregulation in both sexes of Galliformes (and thus are present but monomorphic in many basal lineages), and only in some of the Phasianidae were these co-opted for sexual selection [25, 71] leading to the evolution of dimorphism for this trait.

To assess whether the pattern of transitions to and from a monomorphic state are similar to the patterns inferred using individual characters, we scored the presence or absence of any plumage dimorphism (character 10). Patterns of gains and losses of dimorphism in any plumage character exhibit roughly equal rates of gains and losses, in contrast to the pattern seen with the individual plumage characters that generally show a higher rate of losses (Table 3). The common ancestor to the core phasianids has a relatively high summed probability of dimorphism in the five individual plumage characters (Table 4), making it likely the ancestor that had already gained multiple plumage characters. Under this scenario, any plumage character can only transition to monomorphism if all of the individual dimorphic characters are loss, whereas gain only requires a change in one trait. Thus, loss of dimorphism in any plumage is expected to be much less common for this character than in the individual characters, as we observe.

A very different pattern is seen when birds are scored for the presence of dimorphism in any competitive trait (character 11), where gains appear much more common than losses (Table 3). Although the rate of gain is greater than that of loss, a one-rate model (equal gains and losses) could not be rejected for this character (2Δ log likelihood = 0.2; $2\Delta_{crit}$ = 3.84, df = 1, P = .66). This suggests that the likelihood surface for this character is very flat, such that very different gain to loss ratios should result in only minor differences in likelihood values. Since an equal rate of gains to losses could not be rejected it remains possible that the gain to loss ratio is actually much lower than estimates obtained using the trees sampled by the Bayesian analysis (Table 3), making it difficult to interpret patterns of changes in this character at this time.

Some species exhibit more extreme dimorphism, being dimorphic for multiple traits. We examined transitions to and from high degrees of dimorphism in same type of trait. Species that exhibited high plumage dimorphism relative to other species (character 12) showed a strong bias toward loss of dimorphism (Table 3). This pattern would be expected given the greater rate of loss for the individual plumage characters and the relatively high probability of plumage dimorphism in multiple traits in the ancestor to the phasianids (Table 4). Similarly, as expected from the patterns observed in individual competitive traits, species that exhibited high competitive dimorphism (character 13) showed the opposite pattern, with gains exceeding losses in extreme dimorphism (Table 3).

Strong sexual selection could select for increased dimorphism in two types of traits (i.e., one signal and one competitive trait). An examination of the overall level of dimorphism in signal and competitive characters indicates that there are clades that are highly dimorphic for both types of traits, clades that exhibit little dimorphism in both types of traits, or clades that are highly dimorphic in one, but not both, types of traits (Figure 2). Thus, there is not a strong, consistent pattern that emerges at this broad level. However, there do appear to be some correlation among transitions in the characters (Table 5) suggesting that the transitions among character states may not be independent of other characters. For example, gains in the any plumage (character 10) and any competitive (character 11) are correlated with fleshy traits (character 6), while being highly dimorphic for plumage (character 12) appears to lead to more gains in both fleshy traits (character 6) and being highly dimorphic for competitive traits (character 13).

In summary, the patterns we observed are complex and suggest that no single dominant explanation will be sufficient to explain our observations. Both monomorphism

Table 4: Reconstructed dimorphism at ancestral nodes. Ancestral values are the summed dimorphism at each node, followed by the 95% range in parentheses.

	Maximum possible	Ancestor of outgroups	Ancestor of Phasianidae	Ancestor of core phasianids
Total dimorphism (14)	9	2.8 (1.7–3.4)	4.0 (2.4–5.3)	5.4 (4.6–7.1)
Signal dimorphism (15)	6	1.8 (0.4–2.3)	2.7 (1.5–3.6)	3.3 (2.6–4.4)
Competitive dimorphism (16)	3	1.0 (0.7–1.9)	1.3 (0.8–1.8)	2.1 (1.6–2.9)
Plumage dimorphism (17)	5	1.8 (0.4–2.2)	2.7 (1.5–3.6)	3.3 (2.6–4.4)

Table 5: Correlated patterns of change between characters (character number indicated in parentheses). *indicates $P < .05$, **indicates $P < .01$.

		Δ log likelihood	Highest transition
Any plumage (10)	Any competitive (11)	9.47**	Any plum. 0 → 1, if Any comp. = 0
Fleshy traits (6)	Any plumage (10)	11.62**	Any plum. 0 → 1, if Fleshy = 1
Fleshy traits (6)	Any competitive (11)	5.31*	Any comp. 0 → 1, if Fleshy = 1
High Plumage (12)	High competitive (13)	5.39**	High comp. 0 → 1, if High plum. = 1
Fleshy traits (6)	High plumage (12)	4.97**	Fleshy 0 → 1, if High plum. = 1
Fleshy traits (6)	High competitive (13)	3.77	No relationship

and dimorphism evolved multiple times for all characters (Figure 2 and Table 3). When all traits were considered together there was no clear trend towards gain or loss, although there were some weaker patterns. Consistent with the hypothesis that gains in dimorphism in traits that are directly involved in competition are expected, competitive characters exhibited a larger number of gains than signaling characters. Surprisingly, much of the gain in signal character dimorphism appears in the ancestor to modern phasianids (Figure 2) and losses predominate for these characters within the group (Table 3). This is in contrast both to expectation (e.g., [1]) and our previous hypothesis for the Phasianidae [42].

3.2.1. Observed versus Predicted Patterns of Character Evolution. In Table 1, we summarize the patterns we identified that are either inconsistent with our predictions given specific models (and would therefore falsify the model) or consistent with our predictions. It is immediately obvious by reference to Table 1 that there was no single model that was consistently supported that also could not be refuted by our results. Falsification of many models was due to our observed pattern of loss of dimorphism. The overall pattern of trait loss from an early and highly dimorphic ancestor or set of ancestors is one that has not been strongly considered in the field of sexual selection (despite other studies that have highlighted the importance of loss of sexually dimorphic traits [66, 70]). Since the sexual selection literature has driven the development of hypotheses to explain the presence of multiple sexual signals, loss of signaling traits has not been explicitly considered in these models as well. Thus, these results highlight the need for a more balanced treatment of trait loss and gain in this area. Furthermore, the diversity of observed patterns of gain and loss of dimorphism among lineages within the Phasianidae suggests that the processes affecting the evolution of multiple signaling systems in any single lineage are either temporally variable or diverse. Thus, there appears to be no single, predominant mechanism to explain the evolution of multiple sexual signals.

That said, our test of the multiple/dynamic sensory environments hypothesis [3], essentially the idea that fluctuating signaling environments may favor different signals at different times, is particularly weak. This hypothesis has the potential to be more general than the other hypotheses tested, due to its flexibility and the wide array of patterns it can explain as a result. Within the hypothesis are nested the possibilities that different aspects of a mate or competitor are important in different contexts (i.e., the multiple messages hypothesis) and that different environments may demand different signals to communicate the same message (i.e., receiver psychology). Thus, it is also more comprehensive than other proposed models, as well as being more dynamic. However, adequate tests of this model demand a phylogenetic approach that considers environmental variables alongside signaling traits. Without those comparisons, we cannot seriously evaluate whether this more general model could explain the type of patterns we observed, and such analyses should be conducted in the future.

In contrast, those models for which the phylogenetic predictions are best specified, that is, the models that specifically incorporate interspecific dynamics (bottom portion of Table 1), have little support across the Phasianidae. Thus, none of these models appears to be sufficiently general so as to explain the distribution of multiple sexual signals within this family. This is true, whether we focus exclusively on those characters that are most likely to be the result of mate choice or also include characters that are associated with competitive ability (which females often consider in mate choice; reviewed in [14]).

The remaining hypotheses that we examined are both supported and refuted across the Phasianidae. Thus, even within this one family, the diversity of factors contributing to the evolution and maintenance of multiple sexual signals across taxa appears to be greater than postulated by any single model. However, it is also not possible to exclude any

of these models as potential explanations for the patterns in at least some lineages within the Phasianidae. The difficulty in excluding any hypothesis may reflect one of two factors. First, it may be that all of these processes contribute to the evolution of multiple signaling traits in this family to at least some degree, though the relative importance of each may vary across taxa. Alternatively, this may also reflect that the macroevolutionary predictions of each hypothesis are currently too imprecise to exclude some models. Additional refinement of the predictions each model makes over evolutionary time might allow the exclusion of some of these in the future.

Taken together, this comparison of data with models to explain the evolution of multiple sexual signals suggests we need to focus on (1) a more explicit consideration of trait loss in the context of sexual selection and the evolution of multiple sexual signals and (2) formally extending the existing models on the evolution of multiple sexual signals to the macroevolutionary scale and thus, strengthening our predictions. Our observed pattern of diversification and losses of multitrait dimorphism in the Phasianidae highlights the complexity of testing among the existing models, particularly as the models are currently described. Additionally, it is clear we need to incorporate environmental characters into these types of analysis to fully test existing models. Finally, to better relate theoretical predictions about the functionality of signaling and competitive traits with their resulting phylogenetic distribution we also need to better understand the costs of different signal types to both signalers and receivers (e.g., [72]), the effectiveness of different types of signals across environments, and the nature of interactions amongst signals.

3.3. Broader Implications and Future Directions. Sexual selection is hypothesized to have led to rapid evolutionary change [1]. If this is true, there are several expectations that should be apparent in a study such as this. First, there may be gains of multiple sexual signals along a single branch in a phylogenetic tree, rather than a slow stepwise gain of such traits at the time that the strength of sexual selection increased. Second, a shift to increasing sexual selection in a lineage may drive rapid radiations and lead to short internodes shortly after that period. Both of these patterns are consistent with our observations in the Phasianidae. The base of the core phasianids contains may short branches (e.g., Figure 1), and it is at this time that it appears that there was a gain in multiple sexually dimorphic traits. Understanding whether this is a general phenomenon or a pattern specific to the Phasianidae will require looking for similar associations between rapid evolutionary radiation and the gain of multiple sexually dimorphic traits in other groups. Additionally, the potential to evolve sexual dimorphism in suites of traits in a short time period may depend on the genetic basis of such traits, which is likely to be complex. To understand the specific genetic changes responsible for sexually dimorphic characters is likely to require combining analyses like these with experimentation in specific species, but should lead to a greater knowledge of the processes that may lead to multiple sexual signals.

Many models for the evolution of sexual signals have been phrased, either explicitly or implicitly, in a microevolutionary framework (see Table 1). This has led to tests using behavioral assays within species focused on assessing the strength and nature of sexual selection. These types of studies are clearly important, but another area that will be important to explore in the future is the development of explicit mathematical models that can be used to examine patterns of sexual signal evolution in a macroevolutionary framework. These can then be used to gain a greater understanding of the evolution of sexual dimorphism in the Phasianidae and other groups of organisms, potentially revealing any unifying themes in the evolution of sexual signals over relatively large time scales.

Acknowledgments

This paper was greatly improved by helpful suggestions from an anonymous reviewer and Rob Kulathinal. This research was facilitated by funding from the National Science Foundation (DEB-0228682) to R. T. Kimball, E. L. Braun, and D. W. Steadman.

References

[1] M. Andersson, *Sexual Selection*, Princeton University Press, Princeton, NJ, USA, 1994.

[2] U. Candolin, "The use of multiple cues in mate choice," *Biological Reviews of the Cambridge Philosophical Society*, vol. 78, no. 4, pp. 575–595, 2003.

[3] J. Bro-Jørgensen, "Dynamics of multiple signalling systems: animal communication in a world in flux," *Trends in Ecology and Evolution*, vol. 25, no. 5, pp. 292–300, 2010.

[4] A. P. Møller and A. Pomiankowski, "Why have birds got multiple sexual ornaments?" *Behavioral Ecology and Sociobiology*, vol. 32, no. 3, pp. 167–176, 1993.

[5] R. A. Johnstone, "The tactics of mutual mate choice and competitive search," *Behavioral Ecology and Sociobiology*, vol. 40, no. 1, pp. 51–59, 1997.

[6] T. Guilford and M. S. Dawkins, "Receiver psychology and the design of animal signals," *Trends in Neurosciences*, vol. 16, no. 11, pp. 430–436, 1993.

[7] T. Guilford and M. S. Dawkins, "Receiver psychology and the evolution of animal signals," *Animal Behaviour*, vol. 42, no. 1, pp. 1–14, 1991.

[8] C. Rowe, "Receiver psychology and the evolution of multicomponent signals," *Animal Behaviour*, vol. 58, no. 5, pp. 921–931, 1999.

[9] H. Kokko, M. D. Jennions, and A. Houde, "Evolution of frequency-dependent mate choice: keeping up with fashion trends," *Proceedings of the Royal Society B*, vol. 274, no. 1615, pp. 1317–1324, 2007.

[10] B. Holland and W. R. Rice, "Perspective: chase-away sexual selection: antagonistic seduction versus resistance," *Evolution*, vol. 52, no. 1, pp. 1–7, 1998.

[11] G. S. Van Doorn and F. J. Weissing, "Sexual conflict and the evolution of female preferences for indicators of male quality," *American Naturalist*, vol. 168, no. 6, pp. 742–757, 2006.

[12] A. Pomiankowski and Y. Iwasa, "Evolution of multiple sexual preferences by Fisher's runaway process of sexual selection," *Proceedings of the Royal Society B*, vol. 253, no. 1337, pp. 173–181, 1993.

[13] R. O. Prum, "Phylogenetic tests of alternative intersexual selection mechanisms: Trait macroevolution in a polygynous clade (Aves: Pipridae)," *American Naturalist*, vol. 149, no. 4, pp. 668–692, 1997.

[14] A. Berglund, A. Bisazza, and A. Pilastro, "Armaments and ornaments: an evolutionary explanation of traits of dual utility," *Biological Journal of the Linnean Society*, vol. 58, no. 4, pp. 385–399, 1996.

[15] R. O. Vasconcelos, J. M. Simões, V. C. Almada, P. J. Fonseca, and M. C. P. Amorim, "Vocal behavior during territorial intrusions in the lusitanian toadfish: boatwhistles also function as territorial 'keep-out' signals," *Ethology*, vol. 116, no. 2, pp. 155–165, 2010.

[16] A. S. Chaine and B. E. Lyon, "Intrasexual selection on multiple plumage ornaments in the lark bunting," *Animal Behaviour*, vol. 76, no. 3, pp. 657–667, 2008.

[17] C. Rowe, "Sound improves visual discrimination learning in avian predators," *Proceedings of the Royal Society B*, vol. 269, no. 1498, pp. 1353–1357, 2002.

[18] J. A. Pfaff, L. Zanette, S. A. MacDougall-Shackleton, and E. A. MacDougall-Shackleton, "Song repertoire size varies with HVC volume and is indicative of male quality in song sparrows (Melospiza melodia)," *Proceedings of the Royal Society B*, vol. 274, no. 1621, pp. 2035–2040, 2007.

[19] J. T. Thompson, A. N. Bissell, and E. P. Martins, "Inhibitory interactions between multimodal behavioural responses may influence the evolution of complex signals," *Animal Behaviour*, vol. 76, no. 1, pp. 113–121, 2008.

[20] J. Hoglund, "Size and plumage dimorphism in lek-breeding birds: a comparative analysis," *American Naturalist*, vol. 134, no. 1, pp. 72–87, 1989.

[21] International Chicken Genome Sequencing Consortium, "Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution," *Nature*, vol. 432, no. 7018, pp. 695–716, 2004.

[22] R. A. Dalloul, J. A. Long, A. V. Zimin et al., "Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): genome assembly and analysis," *PLoS Biology*, vol. 8, no. 9, Article ID e1000475, 2010.

[23] S. Madge and P. McGowan, *Pheasants, Partridges and Grouse*, Princeton University Press, Princeton, NJ, USA, 2002.

[24] G. W. H. Davison, "Avian spurs," *Journal of Zoology*, vol. 206, pp. 353–366, 1985.

[25] R. T. Kimball and E. L. Braun, "A multigene phylogeny of Galliformes supports a single origin of erectile ability in non-feathered facial traits," *Journal of Avian Biology*, vol. 39, no. 4, pp. 438–445, 2008.

[26] R. T. Kimball, E. L. Braun, J. D. Ligon, E. Randi, and V. Lucchini, "Using molecular phylogenetics to interpret evolutionary changes in morphology and behavior in the Phasianidae," *Acta Zoologica Sinica*, vol. 52, supplement, pp. 362–365, 2006.

[27] J. D. Ligon, R. Kimball, and M. Merola-Zwartjes, "Mate choice by female red junglefowl: the issues of multiple ornaments and fluctuating asymmetry," *Animal Behaviour*, vol. 55, no. 1, pp. 41–50, 1998.

[28] J. D. Ligon, R. Thornhill, M. Zuk, and K. Johnson, "Male-male competition, ornamentation and the role of testosterone in sexual selection in red jungle fowl," *Animal Behaviour*, vol. 40, no. 2, pp. 367–373, 1990.

[29] T. S. Johnsen, M. Zuk, and E. A. Fessler, "Social dominance, male behaviour and mating in mixed-sex flocks of red jungle fowl," *Behaviour*, vol. 138, no. 1, pp. 1–18, 2001.

[30] M. Petrie, H. Tim, and S. Carolyn, "Peahens prefer peacocks with elaborate trains," *Animal Behaviour*, vol. 41, no. 2, pp. 323–331, 1991.

[31] A. Loyau, M. S. Jalme, and G. Sorci, "Intra- and intersexual selection for multiple traits in the peacock (*Pavo cristatus*)," *Ethology*, vol. 111, no. 9, pp. 810–820, 2005.

[32] M. Takahashi, H. Arita, M. Hiraiwa-Hasegawa, and T. Hasegawa, "Peahens do not prefer peacocks with more elaborate trains," *Animal Behaviour*, vol. 75, no. 4, pp. 1209–1219, 2008.

[33] R. Buchholz, "Male dominance and variation in fleshy head ornamentation in wild turkeys," *Journal of Avian Biology*, vol. 28, no. 3, pp. 223–230, 1997.

[34] R. Buchholz, "Female choice, parasite load and male ornamentation in wild turkeys," *Animal Behaviour*, vol. 50, no. 4, pp. 929–943, 1995.

[35] C. Mateos, "Sexual selection in the ring-necked pheasant: a review," *Ethology Ecology and Evolution*, vol. 10, no. 4, pp. 313–332, 1998.

[36] L. M. Brodsky, "Ornament size influences mating success in male rock ptarmigan," *Animal Behaviour*, vol. 36, no. 3, pp. 662–667, 1988.

[37] K. Holder and R. Montgomeri, "Context and consequences of comb displays by male rock ptarmigan," *Animal Behaviour*, vol. 45, no. 3, pp. 457–470, 1993.

[38] P. T. Rintamäki, J. Höglund, E. Karvonen et al., "Combs and sexual selection in black grouse (*Tetrao tetrix*)," *Behavioral Ecology*, vol. 11, no. 5, pp. 465–471, 2000.

[39] C. G. Sibley and J. E. Ahlquist, *Phylogeny and Classification of Birds: A Study in Molecular Evolution*, Yale University Press, New Haven, Conn, USA, 1990.

[40] P. R. Lowe, "Some preliminary notes on the anatomy and systematic position of *Afropavo congensis* Chapin," in *Proceedings of the 9th International Ornithological Congress*, J. Delacour, Ed., pp. 219–230, Rouen, France, 1938.

[41] J. Delacour, *The Pheasants of the World*, Saiga, Surrey, UK, 1977.

[42] R. T. Kimball, E. L. Braun, P. W. Zwartjes, T. M. Crowe, and J. D. Ligon, "A molecular phylogeny of the pheasants and partridges suggests that these lineages are not monophyletic," *Molecular Phylogenetics and Evolution*, vol. 11, no. 1, pp. 38–54, 1999.

[43] T. M. Crowe, R. C. K. Bowie, P. Bloomer et al., "Phylogenetics, biogeography and classification of, and character evolution in, gamebirds (Aves: Galliformes): effects of character exclusion, data partitioning and missing data," *Cladistics*, vol. 22, no. 6, pp. 495–532, 2006.

[44] M. H. Armstrong, E. L. Braun, and R. T. Kimball, "Phylogenetic utility of avian ovomucoid intron G: a comparison of nuclear and mitochondrial phylogenies in Galliformes," *Auk*, vol. 118, no. 3, pp. 799–804, 2001.

[45] E. J. Smith, L. Shi, and Z. Tu, "*Gallus gallus* aggrecan gene-based phylogenetic analysis of selected avian taxonomic groups," *Genetica*, vol. 124, no. 1, pp. 23–32, 2005.

[46] W. A. Cox, R. T. Kimball, and E. L. Braun, "Phylogenetic position of the New World quail (Odontophoridae): eight nuclear loci and three mitochondrial regions contradict morphology and the Sibley-Ahlquist tapestry," *Auk*, vol. 124, no. 1, pp. 71–84, 2007.

[47] V. B. Kaiser, M. van Tuinen, and H. Ellegren, "Insertion events of CR1 retrotransposable elements elucidate the phylogenetic branching order in galliform birds," *Molecular Biology and Evolution*, vol. 24, no. 1, pp. 338–347, 2007.

[48] J. O. Kriegs, A. Matzke, G. Churakov et al., "Waves of genomic hitchhikers shed light on the evolution of gamebirds (Aves: Galliformes)," *BMC Evolutionary Biology*, vol. 7, article no. 190, 2007.

[49] Y. Y. Shen, L. Liang, Y. B. Sun et al., "A mitogenomic perspective on the ancient, rapid radiation in the Galliformes with an emphasis on the Phasianidae," *BMC Evolutionary Biology*, vol. 10, no. 1, article no. 132, 2010.

[50] T. Lislevand, J. Figuerola, and T. Székely, "Evolution of sexual size dimorphism in grouse and allies (Aves: Phasianidae) in relation to mating competition, fecundity demands and resource division," *Journal of Evolutionary Biology*, vol. 22, no. 9, pp. 1895–1905, 2009.

[51] H. N. Bryant, "The cladistic of matrix representation with parsimony analysis," in *Phylogenetic Supertrees. Combining Information to Reveal the Tree of Life Computational Biology*, O. R. P. Binnida-Emonds, Ed., vol. 3, pp. 353–368, Kluwer Academic Publishers, Dordrecth, The Netherlands, 2004.

[52] J. Gatesy and M. S. Springer, "A critique of matrix representation with parsimony supertrees," in *Phylogenetic Supertrees. Combining Information to Reveal the Tree of Life Computational Biology*, O. R. P. Binnida-Emonds, Ed., vol. 3, pp. 369–388, Kluwer Academic Publishers, Dordrecth, The Netherlands, 2004.

[53] C. Camacho, G. Coulouris, V. Avagyan et al., "BLAST+: architecture and applications," *BMC Bioinformatics*, vol. 10, article no. 421, 2009.

[54] E. Randi, V. Lucchini, A. Hennache, R. T. Kimball, E. L. Braun, and J. D. Ligon, "Evolution of the mitochondrial DNA control region and cytochrome *b* genes and the inference of phylogenetic relationships in the avian genus *Lophura* (Galliformes)," *Molecular Phylogenetics and Evolution*, vol. 19, no. 2, pp. 187–201, 2001.

[55] A. Stamatakis, "RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models," *Bioinformatics*, vol. 22, no. 21, pp. 2688–2690, 2006.

[56] A. Stamatakis, "Phylogenetic models of rate heterogeneity: a high performance computing perspective," in *Proceedings of 20th IEEE/ACM International Parallel and Distributed Processing Symposium (IPDPS '06)*, Rhodos, Greece, April 2006.

[57] J. P. Huelsenbeck and F. Ronquist, "MRBAYES: Bayesian inference of phylogenetic trees," *Bioinformatics*, vol. 17, no. 8, pp. 754–755, 2001.

[58] F. Ronquist and J. P. Huelsenbeck, "MrBayes 3: Bayesian phylogenetic inference under mixed models," *Bioinformatics*, vol. 19, no. 12, pp. 1572–1574, 2003.

[59] P. A. Johnsgard, *Grouse and Quails of North America*, University of Nebraska Press, Lincoln, Neb, USA, 1973.

[60] P. A. Johnsgard, *The Quails, Partridges, and Francolins of the World*, Oxford University Press, Oxford, UK, 1988.

[61] P. A. Johnsgard, *The Pheasants of the World*, Oxford University Press, Oxford, UK, 1986.

[62] M. Pagel, A. Meade, and D. Barker, "Bayesian estimation of ancestral character states on phylogenies," *Systematic Biology*, vol. 53, no. 5, pp. 673–684, 2004.

[63] "Mesquite: A modular system for evolutionary analysis Version Version 2.72," http://mesquiteproject.org.

[64] S. J. Hackett, R. T. Kimball, S. Reddy et al., "A phylogenomic study of birds reveals their evolutionary history," *Science*, vol. 320, no. 5884, pp. 1763–1768, 2008.

[65] M. Pagel, "Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters," *Proceedings of the Royal Society B*, vol. 255, no. 1342, pp. 37–45, 1994.

[66] R. T. Kimball, E. L. Braun, D. J. Ligon, V. Lucchini, and E. Randi, "A molecular phylogeny of the peacock-pheasants (Galliformes: *Polyplectron* spp.) indicates loss and reduction of ornamental traits and display behaviours," *Biological Journal of the Linnean Society*, vol. 73, no. 2, pp. 187–198, 2001.

[67] A. J. Bonilla, E. L. Braun, and R. T. Kimball, "Comparative molecular evolution and phylogenetic utility of 3′-UTRs and introns in Galliformes," *Molecular Phylogenetics and Evolution*, vol. 56, no. 2, pp. 536–542, 2010.

[68] N. Kolm, R. W. Stein, A. Ø. Mooers, J. J. Verspoor, and E. J. A. Cunningham, "Can sexual selection drive female life histories? A comparative study on galliform birds," *Journal of Evolutionary Biology*, vol. 20, no. 2, pp. 627–638, 2007.

[69] S. H. Eo, O. R. P. Bininda-Emonds, and J. P. Carroll, "A phylogenetic supertree of the fowls (Galloanserae, Aves)," *Zoologica Scripta*, vol. 38, no. 5, pp. 465–481, 2009.

[70] J. J. Wiens, "Widespread loss of sexually selected traits: How the peacock lost its spots," *Trends in Ecology and Evolution*, vol. 16, no. 9, pp. 517–523, 2001.

[71] R. Buchholz, *Adaptive Functions of Fleshy Ornamentation in Wild Turkeys and Related Birds*, Dept. of Zoology, University of Florida, Gainesville, Fla, USA, 1994.

[72] R. S. Wilson, C. H. Condon, G. David, S. FitzGibbon, A. C. Niehaus, and K. Pratt, "Females prefer athletes, males fear the disadvantaged: different signals used in female choice and male competition have varied consequences," *Proceedings of the Royal Society B*, vol. 277, no. 1689, pp. 1923–1928, 2010.