Volff J-N (ed): Gene and Protein Evolution.
Genome Dyn. Basel, Karger, 2007, vol 3, pp 101–118

Evolution of Gene Function on the X Chromosome Versus the Autosomes

N.D. Singh, D.A. Petrov

371 Serra Mall, Stanford, Calif., USA

Abstract

Sex chromosomes have arisen from autosomes many times over the course of evolution. This process generates chromosomal heteromorphy between the sexes, which has important implications for the evolution of coding and noncoding sequences on the sex chromosomes versus the autosomes. The formation of sex chromosomes from autosomes involves a reduction in gene dosage, which can modify properties of selection pressure on sex-linked genes. This transition also generates differences in the effective population size and dominance characteristics of novel mutations on the sex chromosome versus the autosomes. All of these changes may affect both patterns of in situ gene evolution and the rates of interchromosomal gene duplication and movement. Here we present a synopsis of the current understanding of the origin of sex chromosomes, theoretical context for differences in rates and patterns of molecular evolution on the X chromosome versus the autosomes, as well as a summary of empirical molecular evolutionary data from *Drosophila* and mammalian genomes.

Origin of Sex Chromosomes

Sex chromosomes are thought to be derived from an ancient pair of autosomes, have arisen independently several times over the course of evolution (for review see [1]), and are one of the mechanisms by which sex is determined. Sex chromosomes are generally morphologically and genetically distinct, though in some cases of recently formed sex chromosomes, the homologous neo-sex chromosomes have not had time to differentiate significantly and are thus quite similar. The heterogametic sex can be either male or female; in the case of female heterogamety the sex chromosomes are denoted Z and W while in the case of male heterogamety they are referred to as X and Y. Both types of sex chromosome evolution have arisen in diverse lineages, though the XX/XY system of

male heterogamety appears to be more common, and is found in plants, mammals, insects, amphibians, reptiles, and several groups of teleost fish, among others [2, 3]. Birds and butterflies represent the best-studied examples of the ZZ/ZW system, though this gonosomal system is also found in plants, insects, amphibians, reptiles, as well as fish [2, 3]. For the sake of simplicity, the remainder of this discussion will be in the context of male heterogametic systems.

In spite of the independent origins of these sex chromosomes in diverse taxa, there are striking similarities among them. Notably, lack of recombination in the heterogametic sex and the erosion of the Y chromosome appear to be hallmarks of sex chromosome evolution [1–4]. This degeneration of the Y typically involves gene loss as well as chromatin state transitions, and as a result, Y chromosomes are typically smaller than their homologous counterparts and are largely heterochromatic [2]. That there are common features among evolutionarily distinct sex chromosomes may be suggestive of a general framework for sex chromosome evolution.

The current model (for review see [1]) for the evolution of sex chromosomes presumes that the sex chromosomes originate from a homologous pair of freely recombining autosomes. The first step in the transition between the autosomal and sex-linked state is the restriction of recombination between the proto-X and the proto-Y, though explanations for why such suppression evolves remain elusive [2]. Most likely, if alleles contributing to sex-determination arise, recombination between these genes will be deleterious, as it would generate maladapted sexual phenotypes [4]. Following this crucial step, natural selection may continue to favor tight linkage of the genes involved in sex determination. Selection may also promote the linkage of these sex-determining loci with sexually antagonistic genes, which are advantageous in one sex while deleterious in the other, as genes with sexually antagonistic consequences that cosegregate with the appropriate sex-determining locus will enjoy a selective advantage over those that segregate independently. This process will differentiate the sex chromosomes further, which will in turn generate selective pressure for the suppression of recombination over an increasingly large window [4]; inversions and rearrangements, which have been implicated in the evolution of the human sex chromosomes [5, 6], may play a role at this stage.

Once the X and the Y are differentiated and recombination between them has reached sufficiently low levels, the Y chromosome begins to erode genetically. There are five main theories for this degeneration, and it is not yet clear which process is primarily governing this genetic erosion of the Y chromosome (for review see [7, 8]). One possibility is that the Y chromosome degenerates due to the effects of Muller's Ratchet [9, 10], or the continued and stochastic loss of the chromosome class containing the fewest deleterious mutations in non-recombining, finite populations. A background selection model [11] similarly

predicts a gradual decline in fitness of the Y chromosome; background selection reduces the effective population size of the non-recombining Y chromosome by removing deleterious alleles and neutral variants on this allelic background from the population, which facilitates the fixation of mildly deleterious mutants due to lowered effective population size.

Another hypothesis is 'weak-selection Hill-Robertson effect' [12], which posits that fixation probabilities of alleles are altered by fixation probabilities of nearby, linked mutations; this general model can contribute to the degeneration of the Y chromosome but the timescales required under this model may be too great to be of biological significance [8]. Genetic hitchhiking [13] may also play a role in Y chromosome degeneration [14], because fixation of novel adaptive alleles also results in fixation of all linked deleterious mutations. As the Y chromosome is unable to recombine, repeated bouts of positive selection will lead to an accumulation of deleterious Y-linked alleles. However, this model as well requires timescales that are too great to be appropriate for sex chromosome evolution. Orr's and Kim's "ruby in the rubbish" model [15] is also based on the inability of the non-recombining Y chromosome to evolve adaptively. In this model, beneficial Y-linked mutations rarely reach fixation, creating a large fitness difference between the X and Y chromosomes, which generates selective pressure to decrease the expression of these deleterious Y-linked genes. Thus, there are several non-mutually exclusive hypotheses for the processes underlying the evolution of the Y chromosome, and further work is needed to elucidate the relative roles of these forces in Y chromosome evolution.

In this review, we will focus on the implications of the degeneration of the Y chromosome for the evolution of coding sequences on the X chromosome, particularly with respect to differences in molecular evolution between the X and the autosomes. We will present an introduction to the theoretical and conceptual frameworks for understanding the evolution of X-linked sequences, as well as summaries of available experimental evidence to date.

Theoretical Predictions for Molecular Evolution of X-linked and Autosomal Genes

Three major differences between the X chromosome and the autosomes stem from the evolution of sex chromosomes from autosomes, each of which has important implications for the evolution of coding and noncoding sequences on these chromosome sets. First, as a consequence of the presence of only a single copy of the X chromosome in males, the X and the autosomes may differ with respect to effective population size. Assuming equal numbers of breeding males and breeding females, the effective population size of the

X chromosome should be 3/4 that of the autosomes, since there are only three X chromosomes for every four autosomes segregating in the population. However, other factors will serve to modulate the effective population size of the X chromosome relative to that of the autosomes such as sex-specific life history traits or differences in breeding success between males and females. In fact, if the effective number of breeding females far exceeds the effective number of breeding males, the effective population size of the X chromosome can equal or even exceed that of the autosomes [16, 17]. Effective population size is an important molecular evolutionary parameter, as increases in effective population size can increase the efficacy of natural selection on weakly adaptive or mildly deleterious mutations.

The effective haploidy of the X chromosome in males also affects the visibility of novel mutations to natural selection. If new mutations are at least partially recessive, the selective effects of novel autosomal variants can be masked by the homologous allele in heterozygous individuals. In contrast, X-linked alleles are immediately visible to selection in males, as males are hemizygous for the X chromosome. Although the X chromosome spends only 1/3 of its evolutionary history in males, the increased exposure of X-linked alleles to selection in these hemizygous individuals enhances the efficacy of both positive and negative selection [18], which can categorically alter rates of molecular evolution between the X and the autosomes.

Finally, that the sex chromosomes are formed from a pair of ancestral autosomes generates a dosage problem for X-linked genes. Indeed, the degradation of the Y chromosome involves gene loss [4, 8], and as a result, X-linked genes generally lack a functional counterpart on the Y and are thus present at one half the copy number that they were in the ancestral, autosomal state. Such a reduction in gene dosage is likely to be deleterious for many genes and a number of dosage compensation mechanisms have evolved in several lineages to remediate the effects of this dosage problem. *Drosophila* and mammals provide the most well-characterized mechanisms of dosage compensation (for review see [19, 20]); in *Drosophila*, dosage compensation is mediated by transcriptional upregulation of X-linked genes in males, whereas in mammals one copy of the X chromosome is transcriptionally downregulated in females.

Though much dosage compensation is achieved through transcriptional regulation, the reduction in gene dosage of X-linked genes may also alter the strength of selection on other mutations that could also serve to partially remediate the dosage problem. For instance, duplication or transposition events from the sex chromosome to the autosomes in some cases may be selectively favored to mitigate the reduction in gene dosage [21–23]. In addition, selection may favor increased codon bias for X-linked genes, which may increase levels of active protein [24] and thus partially compensate for the dosage problem. Even

if dosage equilibration is reached between the sexes, selective pressures may still affect the X chromosome and the autosomes differently.

The chromosomal heteromorphy which results from the evolution of sex chromosomes from autosomes can thus modify properties of selection pressure, effective population size and dominance characteristics of mutations in X-linked genes. These parameters will interact to affect rates and patterns of evolution on the X chromosome relative to the autosomes, and the relative rates of evolution of X-linked versus autosomal genes under various parameter combinations can be examined. In particular, the relative rates of evolution between the chromosome sets will depend on two properties of new mutations: dominance characteristics and sojourn times.

The coefficient of dominance of a new mutation is a key determinant of the relative rate of spread of that new allele on the X chromosome versus the autosomes. Assuming equal numbers of breeding males and breeding females, rates of evolution on the X chromosome will exceed those on the autosomes if new mutations are on average at least partially recessive [18, 25], for both small and large coefficients of selection [26]. In contrast, rates of substitution of mildly deleterious alleles on the autosomes will exceed those on the X [18]. For codominant mutations, rates of evolution should be comparable between the X and the autosomes and rates of fixation of at least partially dominant mutations on the autosomes will exceed those of the X chromosome, for mutations of both small and large selective effects [18, 26].

Sojourn time of novel mutants differs between the X and the autosomes, and this may also affect rates of evolution for non-neutral mutations between the X and the autosomes. In general, the sojourn time for new beneficial mutations will be shorter if the mutations are X-linked rather than autosomal [25, 27], though the magnitude of the difference in rates of evolution between the X and the autosomes does depend on both the relative numbers of breeding males and females as well as the coefficient of selection [25]. This inequality in transit time holds across all dominance coefficients [25, 26], and results from the greater variance in fitness in the haploid versus the diploid state. Given that changes in allele frequency are a function of variance in fitness [28], that the X chromosome spends 1/3 of its evolutionary history in the haploid state suggests that the change in frequency of a novel beneficial allele will be greater if it is X-linked.

It should be noted that all of the above theoretical predictions are predicated on selection acting on novel mutational variants. If natural selection instead predominantly operates on standing variation, with formerly deleterious mutations becoming advantageous, these predictions no longer hold. Instead, rates of adaptive evolution are slower for X-linked alleles than for autosomal alleles [18], regardless of the dominance coefficient of these alleles [29].

Theoretical predictions are also sensitive to the strength and direction of selection acting on mutant alleles in the two sexes. If selection is not presupposed to be acting equally in the two sexes and rather, if there are opposing selective pressures in males versus females, as is the case for sexually antagonistic genes, then X-linked mutations that benefit males can spread through a population under a less restrictive range of parameter values than is required for autosomal invasion [30]. The opposite appears to be true when the mutations favor females at the cost of males, at least for partially recessive mutations [18].

Patterns of In Situ Evolution of X-linked and Autosomal Genes

The above theoretical considerations suggest that under certain conditions, X-linked loci will have higher rates of adaptive evolution than autosomal loci. For this to be the case, selection must be operating on novel allelic variants, and the selective effects of these mutations must be on average at least partially recessive and be equal in the two sexes [18]. In addition, fixation of alleles under positive selection should predominate over fixation of slightly deleterious alleles. Empirical evidence in support of this 'faster-X' hypothesis is thus taken as confirmation that such assumptions are biologically reasonable.

Early testing for faster-X evolution yielded largely contradictory results. A recent comparison of rates of protein evolution in 254 coding sequences from *D. melanogaster* and *D. simulans* showed comparable rates of evolution for X-linked and autosomal genes [31]. In contrast, comparative genomic data from pairs of orthologous genes from these species as well as *D. pseudoobscura* and *D. miranda* suggest that rates of evolution on the X chromosome exceed those on the autosomes [32], and that rates of adaptive evolution between pairs of X-linked gene duplicates appear to be higher than those for duplicate gene pairs residing on the autosomes [33]. The discrepancy between these two major findings may be a consequence of experimental design; faster-X evolution predicts an increased rate of adaptive evolution for an X-linked gene relative to the rate of evolution that a gene would experience if it were autosomal. As a result, paired comparisons, either between orthologous sequences from different species or paralogous sequences within species, may be more appropriate for testing the faster-X model.

With several *Drosophila* genomes fully sequenced, investigating faster-X in this system is now possible at a large scale. Using the full genomes of *D. melanogaster*, *D. pseudoobscura*, *D. yakuba* as well as large-scale sequence data from *D. miranda*, Thornton and colleagues revisited this question at the genomic scale [34]. Using either only whole genome data from the three Drosophilids or a smaller dataset of 202 coding sequences from all four species

did not qualitatively or quantitatively affect the results; rates of protein evolution are not significantly different between the X and the autosomes. The authors suggest that this lack of support for faster-X evolution indicates that either new mutations are not on average partially recessive or that adaptive evolution originates from mutation-selection equilibrium [34].

Results from mammalian genomes are more consistent, with several studies offering evidence in support of a faster-X model of evolution. A whole genome comparison of the human and chimpanzee genomes reveals rates of protein evolution (estimated as Ka/Ks, or the ratio of nonsynonymous substitutions per nonsynonymous site to synonymous substitutions per synonymous site) of X-linked genes exceeding rates of evolution on the autosomes [35]. In addition, a scan for positively selected genes in these genomes revealed an enrichment of X-linked genes, suggesting that X-linked genes have an increased tendency for adaptive evolution relative to autosomal genes, which is also consistent with the faster-X model [36]. Similarly, inferring selective events using linkage disequilibrium among single nucleotide polymorphisms results in a 2-fold enrichment of putatively selected loci on the X chromosome in humans [37]. These genomic phenomena are also recapitulated within smaller functional categories of genes; sex-linked mammalian sperm proteins, for instance, evolve more rapidly than autosomal sperm proteins [38, 39], and X-linked testis-expressed genes have higher rates of evolution (normalized to account for local mutation rate) than those on the autosomes [40], as do X-linked testis-expressed homeobox genes [41]. While these data do suggest that X-linked genes may indeed evolve more rapidly than autosomal genes in mammalian genomes, there is a possibility that ascertainment bias also plays a role in generating these patterns; an overrepresentation of X-linked genes in rapidly evolving proteins, for instance, could reflect differences in gene complements between the X and the autosomes rather than higher rates of adaptive evolution for X-linked genes.

In addition to examining patterns of interspecific divergence of X-linked versus autosomal genes, patterns of variability within species at sex-linked and autosomal loci can also shed light on the forces contributing to the evolution of coding and noncoding sequences on these chromosome sets. Adaptive evolution and purifying selection will both contribute to levels of intraspecific variation via the effects of genetic hitchhiking and background selection, respectively; differences in the relative contributions of these evolutionary processes in X-linked versus autosomal genes may manifest as differences in standing levels of variation. Importantly, background selection and hitchhiking models make distinct predictions regarding the relative levels of diversity of X-linked and autosomal genes. Background selection models predict higher levels of neutral variation on the X chromosome [11, 27, 42, 43]. The reduction

of variation at neutral sites due to the removal of deleterious alleles and linked neutral variants by purifying selection is most pronounced when deleterious alleles reach high frequencies. Given that purifying selection is more effective on the X chromosome due to the hemizygosity of the X chromosome in males, deleterious mutants are maintained at lower population frequencies if they are X-linked than they would be if they were autosomal. In essence, the X chromosome has a larger effective number of deleterious-mutation free chromosomes, and as a consequence, increased levels of standing neutral variation, relative to the autosomes.

In contrast to the background selection model, genetic hitchhiking may lead to lower levels of polymorphism on the X chromosome versus the autosomes. Because the sojourn time of new adaptive mutations is shorter for X-linked versus autosomal genes [25, 27], there may be fewer recombinational opportunities during a selective sweep of an X-linked gene than there would be in an autosomal gene. In addition, if new beneficial mutations are at least partially recessive on average, X-linked genes simply evolve more rapidly from adaptive evolution, which would result in an increase in the number of selective sweeps of X-linked alleles over autosomal alleles per unit time. Theoretical results suggest that such a hitchhiking model will yield lower diversity on the X chromosome than on the autosomes if new beneficial mutations are partially recessive in systems such as *Drosophila* in which there is no recombination in males, and under a broader range of dominance coefficients if there is recombination in males, as is the case in humans [26].

As these models make different predictions with respect to expected levels of neutral sequence variation on the X and the autosomes, comparing X-linked and autosomal polymorphism can provide insight into the relative roles of background selection versus hitchhiking models. There are numerous studies of levels of molecular polymorphism in *D. melanogaster* and *D. simulans* (for review see [44]). Within *D. simulans*, levels of diversity are consistently lower at X-linked loci [42, 45–47]. In *D. melanogaster*, differences in sequence variation between the X and the autosomes seem heavily dependent on population. For ancestral African populations, X-linked diversity levels are consistently higher than expected under the assumption of equal numbers of breeding males and breeding females, while X-linked diversity appears to be depressed in derived populations of this species [45, 48, 49].

Studies of variation in other taxa have revealed similar patterns. In humans, the densities of single nucleotide polymorphisms and microsatellite markers are considerably lower on the X chromosome relative to the autosomes [50, 51], and noncoding sequence diversity as well as microsatellite variability also seem to be reduced on the human X chromosome [51–53]. The density of SSLP markers as well as polymorphism at these loci in mouse are also depressed for X-linked

loci relative to autosomal loci [54]; this X-specific deficit in polymorphic markers is also found in rat [55]. In chicken as well as two flycatcher species, polymorphism data from intronic sequences also support a reduction in diversity of Z-linked alleles relative to autosomal levels [56, 57].

On balance, it appears as though positive selection does play a role in the evolution of the sex chromosomes. While specifically testing for faster-X in *Drosophila* has generated inconsistent results, patterns of X chromosome evolution in mammals appear wholly consistent with the faster-X model. The discrepancies within the *Drosophila* studies may in fact suggest that positive selection is comparatively rare in this system (although see [58]), or may result from the breakdown of one or more of the assumptions in the faster-X model. Intraspecific patterns of sequence variability are more consistent overall among taxa, with a general trend towards reduced polymorphism of the X (or Z) chromosomes. Such a reduction is consistent with a model of genetic hitchhiking or reduced effective population size, and may thus implicate positive selection in the evolution of the X chromosome.

It should be noted that purifying selection, or selection against deleterious alleles, may be more efficient on the X chromosome as well, in accordance with theoretical predictions [18]. Evidence in support of this model stems largely from studies of the evolution of codon bias on the X chromosome and the autosomes. Codon bias refers to the unequal usage of synonymous codons in protein coding sequences, and is thought to be maintained by the balance among mutation, random genetic drift, and selection on translational efficiency/accuracy [59–62].

Codon bias of X-linked genes appears to be higher than codon bias of genes on the autosomes in *Drosophila* [63–65] and *C. elegans* [65]. This increase in codon bias on the X chromosome in these two systems is not mediated by other known correlates of codon bias such as recombination rate, protein length, or level of gene expression. In addition, the X-specific elevation in codon bias does not result from the identities or functions of the genes residing on this chromosome, as comparisons of codon bias in pairs of X-linked and autosomal duplicate genes in *D. melanogaster* and *C. elegans*, as well as pairs of orthologous genes involved in an X-autosome translocation in *D. melanogaster* and *D. pseudoobscura* also support the increase in codon bias on the X chromosome [65]. Thus, the increase in codon bias on the X chromosome in *Drosophila* and *C. elegans* appears to be due entirely to X-linkage, and is thus consistent with an increased efficacy of purifying selection on the X chromosome.

Importantly, while many of the above observations are indeed consistent with the increased efficacy of both directional and purifying selection on the X chromosome, which is suggestive of an in situ evolutionary model, it is also possible that the observed differences between rates and patterns of molecular

evolution are due to other forces shaping the properties of the resident genes on these different chromosomes. External forces may be of particular importance in light of observed differences between the X and the autosomes with respect to gene content and gene movement, which are the focus of the remainder of this discussion.

Gene Complements of the X Chromosome Versus the Autosomes

In addition to predicting differences in rates of molecular evolution between the X and the autosomes, theoretical models further suggest that differences in gene content may evolve between the X chromosome and the autosomes. In particular, genes with different selective effects in males versus females may accumulate differently on the X versus the autosomes, thus shaping the complements of the genes residing on these chromosome sets. More precisely, genes that benefit one sex at the cost of the other, known as sexually antagonistic genes, can accumulate faster on the X chromosome than on the autosomes under certain conditions [30]. For both partially dominant and partially recessive alleles, sexually antagonistic alleles that are beneficial in males though detrimental to females enjoy higher fixation rates if they are X-linked, though the requisite conditions differ somewhat between these classes of dominance coefficients [30]. Mutations in X-linked genes that benefit females at the cost of males can also increase in frequency more rapidly than comparable autosomal mutations, particularly if these mutations are dominant [30]. Thus, rates of accumulation of sexually antagonistic alleles benefiting females or males on the X chromosome can exceed those rates on the autosomes under a variety of conditions, which may play a role in shaping the gene content of the X versus the autosomes.

In *Drosophila*, there appears to be a relative dearth of male-biased genes and an enrichment of female-biased genes on the X chromosome. The genomic distribution of secreted accessory gland proteins, for example, which are heavily implicated in male reproduction, is shifted significantly away from the X chromosome [21]. In addition, genes with male-biased patterns of germline and somatic gene expression are comparatively rare on the *Drosophila* X [66, 67], and genes with female-biased germline expression are enriched on the X relative to expectation [67].

Similar patterns have been documented in *C. elegans*. In addition to genes with germline expression, male-biased germline expressed genes are less likely to be found on the X chromosome [68, 69] though genes with hermaphrodite somatic-biased expression are enriched on the X [69]. Although seemingly unrelated to the sexual antagonism model, the gene complements of the X and

autosomes in *C. elegans* also differ in other respects; the X chromosome is also relatively devoid of genes essential for basic cellular and developmental processes in the embryo [70].

The distribution of genes with sexually antagonistic consequences in mammalian systems is less straightforward. While genes with male-specific expression patterns consistently appear to be less frequent on the X chromosome in humans, there does not appear to be an overrepresentation of X-linked female-specific genes [71]. Genes with sex- and reproduction-related functions appear to be enriched on the human X as well [72]. In mouse, both a dearth of male-specific and an overabundance of female-specific genes on the X have been documented [73], though genes expressed in spermatogonia in this system are more likely to be X-linked than autosomal [74].

One component of the explanation for the conflicting results from *Drosophila*, *C. elegans*, and mammals in relation to one another as well as to theoretical predictions [30] is related to the inactivation of the X chromosome, which occurs during meiosis in the male germline of mammals and many insect taxa [75]. Thus, genes required in late spermatogenesis are at a selective disadvantage if they are X-linked rather than autosomal, although the same is not the case for genes expressed in the male germline prior to inactivation. Thus, we might not expect to find an accumulation of X-linked male-biased genes for those genes that are expressed after X inactivation (for review see [76]). This is supported by data from mouse, which indicate that the X has a depletion of male-biased genes that are expressed late in spermatogenesis and an enrichment of male-biased genes expressed earlier [73], thus supporting the intersection of the sexual-antagonism [30] and X-inactivation models. Moreover, of the 26 genes identified as acting in late spermatogenesis none appears to be X-linked [77].

Patterns of Gene Traffic on the X and the Autosomes

In addition to differing with respect to rates and patterns of molecular evolution as well as gene content, the X and the autosomes also differ with respect to patterns of gene movement via retrotransposition. While duplicate genes can arise from several mutational mechanisms such as small-scale duplication or whole genome duplication, novel genes created through retrotransposition are somewhat more readily identifiable. The process leading to the formation of these retroposed gene duplicates involves reverse transcription of the mRNA from the parental gene and subsequent insertion of this new DNA sequence into an ectopic location in the genome. As a direct consequence, duplicate genes generated through this mechanism bear signatures of this process, which include the lack of introns, poly-A tracts as well as direct flanking repeats.

These latter two characteristics may erode over evolutionary time due to the accumulation of single nucleotide, insertion and deletion mutations. Although there do appear to be cases in which retroposed genes have recruited new introns [78], the lack of introns is a more stable feature of retrotransposed genes, and is thus generally used as a criterion for identifying retroposed genes.

Studies of the fate of retrotransposed genes have been carried out extensively in *Drosophila* and mammals (for review see [79, 80]). In *Drosophila*, there appears to be a significant excess of retrotransposition of X-linked genes to autosomal locations [77]. This excess of X-linked retrogene origination is similarly documented in mammals [78, 81], although in this system the X chromosome also disproportionately recruits duplicate genes arising through retrotransposition [81]. While the X chromosome also shows increased recruitment of retropseudogenes relative to expectation, which implicates a mutational bias, this mutational explanation is not wholly sufficient to explain the patterns of gene traffic of functional retrogenes in mammals [81].

Interestingly, a large fraction of these retrogenes with X-linked parents derive testis-specific expression patterns [77, 81]. In *Drosophila*, five out of six retrogenes originating from the X chromosome are expressed in testis while their parental gene is not (see [77]). In mammals, a higher percentage of X-originating autosomal retrogenes are expressed in testis than autosomal retrogenes that originate from autosomal genes [81], though, it is not yet clear whether these testis-biased expression patterns are predominantly derived or ancestral. More recent analysis of functional retrogenes in the human genome revealed seven functional duplicate genes arising through retrotransposition, three of which originated from X-linked genes, and all of which had acquired novel testis-biased or testis-specific expression patterns [82]. Larger sampling of retrogenes in humans also supports the hypothesis that retrogenes tend to be expressed in testis [78].

Beyond gene retrotransposition, rates of interchromosomal gene movement for all mechanisms of genic translocation also vary substantially between the X chromosome and the autosomes in *Drosophila* [Davis, Singh and Petrov, unpublished]. By comparing physical map locations between pairs of orthologous genes in *D. pseudoobscura* and *D. melanogaster*, asymmetries in gene movement rates between the sex chromosomes and the autosomes can be explored, with emphasis on the newly formed sex chromosomes in the *D. pseudoobscura* lineage. Preliminary analysis of comparative map locations of orthologous genes in these two species has provided tantalizing evidence that rates of gene movement differ between the sex chromosomes and the autosomes [Davis, Singh and Petrov, unpublished]. In particular, the autosome-X translocation has led to a strong bias toward overall gene loss from the neo-X chromosome in *D. pseudoobscura*. Specifically, it was estimated using a maximum

likelihood framework that while the rate of gene emigration from the neo-X increased by as much as 8-fold, the rate of gene immigration to the neo-X declined to undetectable levels. In addition, these preliminary results suggest that rates of gene movement between the ancestral X and autosomes are higher than the rates of interautosomal gene movement.

There are several models that have been put forward to explain patterns of gene movement between the X and the autosomes, none of which can fully account for all of the described patterns (for review see [79]). The X-inactivation hypothesis suggests that selection favors autosomal locations for retroposed genes with functions requiring expression during male meiosis because of the inactivation of the X chromosome in male germline cells. A related hypothesis, the SAXI hypothesis [76], predicts the redistribution of genes functioning in late spermatogenesis from the X chromosome to the autosomes and the gradual demasculinization of the X chromosome perhaps as a consequence of interactions among sexually antagonistic alleles. Finally, formal population genetic models suggest that mutations that are at least partially dominant may accumulate at higher rates on the autosomes [18], and it seems likely that mutations in genes with sexually antagonistic or sex-limited effects would be predominantly gain-of-function mutations and therefore be at least partially dominant [67]. Each of these models can explain certain aspects of the observed data, but no single model appears to be consistent with all previous reports. Consequently, further investigation of the relative importance of these models in generating the observed patterns of gene movement among chromosomes is warranted.

Summary

Sex chromosomes have evolved from autosomes independently in diverse lineages. The prominent feature of the transition between the ancestral, autosomal state and the sex chromosome state is the reduction in gene dosage of sex-linked alleles in the heterogametic sex. The effective haploidy of the sex chromosome in heterogametes is thus the foundation for major differences between the sex chromosomes and the autosomes, which are of tremendous consequence for the evolution of these sets of chromosomes. A great deal of theoretical attention has been devoted to the evolution of sex-linked versus autosomal alleles [18, 25]. These results suggest that rates of evolution may differ between the X and the autosomes both with respect to adaptive evolution and purifying selection, although the magnitude and direction of the difference depends on parameters such as the coefficient of dominance [18, 25] and on other features of the model such as, for example, whether evolution primarily acts on novel

mutational variants or standing variation [29]. In addition, theoretical considerations of genes with different selective effects in the two sexes predict that gene complements between the X and the autosomes may differ depending on the fitness consequences in each sex and the dominance of mutations affecting these genes [30].

The bulk of the data on contrasting patterns of X chromosomal and autosomal evolution comes from *Drosophila* and mammals. While the data are somewhat conflicting, overall there appears to be some evidence in support of a faster-X model of evolution, suggesting that positive selection plays a role in the evolution of sex chromosomes, although may not be a prominent feature of X-linked genes overall. Increased rates of adaptive evolution for X-linked genes may indeed be restricted to genes that are evolving rapidly under positive selection. Within species polymorphism data from X-linked and autosomal loci are also consistent with increased action of positive selection on the X as well [44, 50–53, 56, 57]. Data from codon bias evolution studies also suggest that purifying selection is more effective on the X chromosome than on the autosomes [63–65]. Together, these data lend support to the recessivity of both beneficial and deleterious alleles.

Although the gene complements of the X and the autosomes tend to differ, there do not appear to be any systematic trends across taxa. In *Drosophila* and *C. elegans*, the X chromosome is relatively devoid of male-biased genes while female or hermaphrodite-biased genes appear to be overrepresented [21, 66–69]. The mammalian X also shows a deficiency of genes expressed late in spermatogenesis, and shows a putative enrichment of female-biased genes as well as an overrepresentation of male-biased genes expressed early in the germline [73].

Similarities and differences between mammals and *Drosophila* are also found with respect to patterns of gene traffic of X-linked and autosomal genes, as well as in the functional characteristics of these retrogenes (for review see [79]). While in both mammals and *Drosophila* the X chromosome disproportionately exports new retrogenes, in mammals the X chromosome also recruits retrogenes in excess of expectation. Retrogenes in both systems tend to be expressed in testis, although it remains to be seen how much of this effect is due to acquisition of novel expression pattern in the derived retroposed gene. While patterns of gene traffic in general appear to differ between the X and the autosomes in *Drosophila* [Davis, Singh and Petrov, unpublished], it remains to be seen whether this is also the case in mammalian genomes. Although several models have been proposed to explain these patterns, none can sufficiently account for all of the observations.

Thus, the sex chromosomes of diverse species share several salient evolutionary features. The similarities among patterns of X-linked and autosomal evolution in these systems likely result from similar evolutionary forces acting

on sex-linked genes in spite of their independent origin. However, there are marked differences between the systems presented here, which may speak to differences in the relative roles of the evolutionary forces of mutation, random genetic drift, and natural selection among these organisms. Clearly, the appropriate framework for understanding the evolution of coding and noncoding sequences on sex chromosomes and the autosomes will integrate general features of sex chromosome evolution as well as lineage-specific effects.

References

1 Charlesworth B: The evolution of sex chromosomes. Science 1991;251:1030–1033.
2 Bull JJ: Evolution of Sex Determining Mechanisms. Benjamin/Cummings Pub. Co. Advanced Book Program, Menlo Park 1983.
3 Solari AJ: Sex Chromosomes and Sex Determination in Vertebrates. CRC Press, Boca Raton 1994.
4 Charlesworth B: The evolution of chromosomal sex determination and dosage compensation. Curr Biol 1996;6:149–162.
5 Marais G, Galtier N: Sex chromosomes: how X-Y recombination stops. Curr Biol 2003;13: R641–R643.
6 Lahn BT, Page DC: Four evolutionary strata on the human X chromosome. Science 1999;286: 964–967.
7 Charlesworth B: Model for evolution of Y chromosomes and dosage compensation. Proc Natl Acad Sci USA 1978;75:5618–5622.
8 Charlesworth B, Charlesworth D: The degeneration of Y chromosomes. Phil Trans R Soc Lond B Biol Sci 2000;355:1563–1572.
9 Muller HJ: The relation of recombination to mutational advance. Mutat Res 1964;1:2–9.
10 Felsenstein J: The evolutionary advantage of recombination. Genetics 1974;78:737–756.
11 Charlesworth B, Morgan MT, Charlesworth D: The effect of deleterious mutations on neutral molecular variation. Genetics 1993;134:1289–1303.
12 McVean GAT, Charlesworth B: The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. Genetics 2000;155:929–944.
13 Maynard Smith J, Haigh J: The hitch-hiking effect of a favourable gene. Genet Res 1974;23: 23–35.
14 Rice WR: Genetic hitch-hiking and the evolution of reduced genetic activity of the Y chromosome. Genetics 1987;116:161–167.
15 Orr HA, Kim Y: An adaptive hypothesis for the evolution of the Y chromosome. Genetics 1998;150:1693–1698.
16 Caballero A: On the effective size of populations with separate sexes, with particular reference to sex-linked genes. Genetics 1995;139:1007–1011.
17 Laporte V, Charlesworth B: Effective population size and population subdivision in demographically structured populations. Genetics 2002;162:501–519.
18 Charlesworth B, Coyne JA, Barton NH: The relative rates of evolution of sex chromosomes and autosomes. Am Nat 1987;130:113–146.
19 Baker BS, Gorman M, Marin I: Dosage compensation in *Drosophila*. Ann Rev Genet 1994;28: 491–521.
20 Marin I, Siegal ML, Baker BS: The evolution of dosage-compensation mechanisms. Bioessays 2000;22:1106–1114.
21 Swanson WJ, Clark AG, Waldrip-Dail HM, Wolfner MF, Aquadro CF: Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. Proc Natl Acad Sci USA 2001;98:7375–7379.

22 Bachtrog D, Charlesworth B: On the genomic location of the *exuperantia1* gene in *Drosophila miranda* – the limits of in situ hybridization experiments. Genetics 2003;164:1237–1240.
23 Yi S, Charlesworth B: A selective sweep associated with a recent gene transposition in *Drosophila miranda*. Genetics 2000;156:1753–1763.
24 Carlini DB, Stephan W: In vivo introduction of unpreferred synonymous codons into the *Drosophila Adh* gene results in reduced levels of ADH protein. Genetics 2003;163:239–243.
25 Avery PJ: The population genetics of haplo-diploids and X-linked genes. Genet Res 1984;44: 321–341.
26 Betancourt AJ, Kim Y, Orr HA: A pseudohitchhiking model of X vs. autosomal diversity. Genetics 2004;168:2261–2269.
27 Aquadro CF, Begun DJ, Kindahl EC: Selection, recombination, and DNA polymorphism in *Drosophila*; in Golding B (ed): Non-Neutral Evolution. Chapman & Hall, London, 1994, pp 46–56.
28 Fisher RA: The Genetical Theory of Natural Selection. Dover, New York, 1958.
29 Orr HA, Bentancourt AJ: Haldane's sieve and adaptation from the standing genetic variation. Genetics 2001;157:875–884.
30 Rice WR: Sex chromosomes and the evolution of sex dimorphism. Evolution 1984;38:735–742.
31 Betancourt AJ, Presgraves DC, Swanson WJ: A test for faster X evolution in *Drosophila*. Mol Biol Evol 2002;19:1816–1819.
32 Counterman BA, Ortiz-Barrientos D, Noor MA: Using comparative genomic data to test for fast-X evolution. Evolution 2004;58:656–660.
33 Thornton K, Long M: Rapid divergence of gene duplicates on the *Drosophila melanogaster* X chromosome. Mol Biol Evol 2002;19:918–925.
34 Thornton K, Bachtrog D, Andolfatto P: X chromosomes and autosomes evolve at similar rates in *Drosophila*: no evidence for faster-X protein evolution. Genome Res 2006;16:498–504.
35 Lu J, Wu CI: Weak selection revealed by the whole-genome comparison of the X chromosome and autosomes of human and chimpanzee. Proc Natl Acad Sci USA 2005;102:4063–4067.
36 Nielsen R, Bustamante C, Clark AG, Glanowski S, Sackton TB, et al: A scan for positively selected genes in the genomes of humans and chimpanzees. PLoS Biol 2005;3:976–985.
37 Wang ET, Kodama G, Baldi P, Moyzis RK: Global landscape of recent inferred Darwinian selection for *Homo sapiens*. Proc Natl Acad Sci USA 2006;103:135–140.
38 Torgerson DG, Singh RS: Sex-linked mammalian sperm proteins evolve faster than autosomal ones. Mol Biol Evol 2003;20:1705–1709.
39 Torgerson DG, Singh RS: Enhanced adaptive evolution of sperm-expressed genes on the mammalian X chromosome. Heredity 2006;96:39–44.
40 Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, et al: Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. Science 2005;309:1850–1854.
41 Wang XX, Zhang JZ: Rapid evolution of mammalian X-linked testis-expressed homeobox genes. Genetics 2004;167:879–888.
42 Begun DJ, Whitley P: Reduced X-linked nucleotide polymorphism in *Drosophila simulans*. Proc Natl Acad Sci USA 2000;97:5960–5965.
43 Charlesworth B: Background selection and patterns of genetic diversity in *Drosophila melanogaster*. Genet Res 1996;68:131–149.
44 Mousset S, Derome N: Molecular polymorphism in *Drosophila melanogaster* and *D. simulans*: what have we learned from recent studies? Genetica 2004;120:79–86.
45 Andolfatto P: Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *Drosophila simulans*. Mol Biol Evol 2001;18:279–290.
46 Schoefl G, Schloetterer C: Patterns of microsatellite variability among X chromosomes and autosomes indicate a high frequency of beneficial mutations in non-african *D. simulans*. Mol Biol Evol 2004;21:1384–1390.
47 Irvin SD, Wetterstrand KA, Hutter CM, Aquadro CF: Genetic variation and differentiation at microsatellite loci in *Drosophila simulans*: evidence for founder effects in new world populations. Genetics 1998;150:777–790.
48 Sheldahl LA, Weinreich DM, Rand DM: Recombination, dominance and selection on amino acid polymorphism in the *Drosophila* genome: contrasting patterns on the X and fourth chromosomes. Genetics 2003;165:1195–1208.

49 Kauer M, Zangerl B, Dieringer D, Schlotterer C: Chromosomal patterns of microsatellite variability contrast sharply in African and non-African populations of *Drosophila melanogaster*. Genetics 2002;160:247–256.
50 Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, et al: A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 2001;409: 928–933.
51 Dib C, Faure S, Fizames C, Samson D, Drouot N, et al: A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 1996;380:152–154.
52 Payseur BA, Cutter AD, Nachman MW: Searching for evidence of positive selection in the human genome using patterns of microsatellite variability. Mol Biol Evol 2002;19:1143–1153.
53 Payseur BA, Nachman MW: Natural selection at linked sites in humans. Gene 2002;300:31–42.
54 Dietrich WF, Miller J, Steen R, Merchant MA, Damron-Boles D, et al: A comprehensive genetic map of the mouse genome. Nature 1996;380:149–152.
55 Jacob HJ, Brown DM, Bunker RK, Daly MJ, Dzau VJ, et al: A genetic linkage map of the laboratory rat, *Rattus norvegicus*. Nat Genet 1995;9:63–69.
56 Borge T, Webster MT, Andersson G, Saetre GP: Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. Genetics 2005;171: 1861–1873.
57 Sundstrom H, Webster MT, Ellegren H: Reduced variation on the chicken Z chromosome. Genetics 2004;167:377–385.
58 Andolfatto P: Adaptive evolution of non-coding DNA in *Drosophila*. Nature 2005;437: 1149–1152.
59 Sharp PM, Li WH: An evolutionary perspective on synonymous codon usage in unicellular organisms. J Mol Evol 1986;24:28–38.
60 Bulmer M: The selection-mutation-drift theory of synonymous codon usage. Genetics 1991;129: 897–908.
61 Akashi H: Codon bias evolution in *Drosophila*. Population genetics of mutation-selection drift. Gene 1997;205:269–278.
62 McVean GAT, Charlesworth B: A population genetic model for the evolution of synonymous codon usage: patterns and predictions. Genet Res 1999;74:145–158.
63 Hambuch TM, Parsch J: Patterns of synonymous codon usage in *Drosophila melanogaster* genes with sex-biased expression. Genetics 2005;170:1691–1700.
64 Comeron JM, Kreitman M, Aguade M: Natural selection on synonymous sites is correlated with gene length and recombination in *Drosophila*. Genetics 1999;151:239–249.
65 Singh ND, Davis JC, Petrov DA: X-linked genes evolve higher codon bias in *Drosophila* and *Caenorhabditis*. Genetics 2005;171:145–155.
66 Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, et al: Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. Science 2003;299:697–700.
67 Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL: Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. Science 2003;300:1742–1745.
68 Reinke V, Smith HE, Nance J, Wang J, Van Doren C, et al: A global profile of germline gene expression in *C. elegans*. Mol Cell 2000;6:605–616.
69 Reinke V, Gil IS, Ward S, Kazmer K: Genome-wide germline-enriched and sex-biased expression profiles in *Caenorhabditis elegans*. Development 2004;131:311–323.
70 Piano F, Schetter AJ, Morton DG, Gunsalus KC, Reinke V, et al: Gene clustering based on RNAi phenotypes of ovary-enriched genes in *C. elegans*. Curr Biol 2002;12:1959–1964.
71 Lercher MJ, Urrutia AO, Hurst LD: Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. Mol Biol Evol 2003;20:1113–1116.
72 Saifi GM, Chandra HS: An apparent excess of sex and reproduction-related genes on the human X chromosome. Proc R Soc Biol Sci B 1999;270:53–59.
73 Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD: The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. Nat Genet 2004;36:642–646.
74 Wang PJ, McCarrey JR, Yang F, Page DC: An abundance of X-linked genes expressed in spermatogonia. Nat Genet 2001;27:422–426.

75 Lifschytz E, Lindsley DL: The role of X chromosome inactivation during spermatogenesis (Drosophila-allocycly-chromosome evolution-male sterility-dosage compensation). Proc Natl Acad Sci USA 1972;69:182–186.
76 Wu CI, Xu EY: Sexual antagonism and X inactivation: the SAXI hypothesis. Trends Genet 2003;19: 243–247.
77 Betran E, Thornton K, Long M: Retroposed new genes out of the X in *Drosophila*. Genome Res 2002;12:1854–1859.
78 Vinckenbosch N, Dupanloup I, Kaessmann H: Evolutionary fate of retroposed gene copies in the human genome. Proc Natl Acad Sci USA 2006;103:3220–3225.
79 Betran E, Emerson JJ, Kaessmann H, Long M: Sex chromosomes and male functions: where do new genes go? Cell Cycle 2004;3:873–875.
80 Khil PP, Oliver B, Camerini-Otero RD: X for intersection: retrotransposition both on and off the X chromosome is more frequent. Trends Genet 2005;21:3–7.
81 Emerson JJ, Kaessmann H, Betran E, Long M: Extensive gene traffic on the mammalian X chromosome. Science 2004;303:537–540.
82 Marques AC, Dupanloup I, Vinckenbosch N, Reymond A, Kaessmann H: Emergence of young human genes after a burst of retroposition in primates. PLoS Biol 2005;3:1970–1979.

Nadia D. Singh
371 Serra Mall
Stanford, CA 94305 (USA)
Tel. +1-607-254-4839, Fax +1-607-255-6249, E-Mail nds25@cornell.edu