A dynamic view of sex chromosome evolution
Doris Bachtrog

Sex chromosomes are derived from ordinary autosomes. The X chromosome is thought to maintain most of its ancestral genes over evolutionary time, whereas its Y counterpart degenerates, owing to its lack of recombination. Genomic analyses of young sex chromosome pairs support this view and have shed light on the evolutionary processes underlying loss of gene function on the Y. Studies of ancestral sex chromosomes, however, have also revealed that the process of sex chromosome evolution can be more dynamic than traditionally appreciated. In particular, ancient Y-chromosomes are characterized not only by a loss of genes relative to the X but also by recurrent gains of individual genes or genomic regions, and they often accumulate genes beneficial to males. Furthermore, X chromosomes are not passive players in this evolutionary process but respond both to their sex-biased transmission and to Y-chromosome degeneration, through feminization and the evolution of dosage compensation.

Addresses
Division of Biological Sciences, University of California, San Diego, 9500 Gilman Drive, MC 0116, La Jolla, CA 92093, USA

Corresponding author: Bachtrog, Doris (dbachtrog@ucsd.edu)

Introduction
In many eukaryotic organisms, the two sexes have different chromosomal constitutions. Typically, one sex is heterogametic and has a pair of morphologically different chromosomes, whereas the other sex is homogametic and has two identical members of each chromosomal pair. Morphologically and genetically distinct X and Y chromosomes (or Z and W chromosomes in systems with female heterogamy) involved in sex determination have evolved independently many times in both animals and plants [1,2]. A striking common feature of many taxa is the almost complete erosion of genes on the Y chromosome. The Y also often contains an unusual abundance of repetitive DNA sequences. In several cases, the possession of genetically eroded Y-chromosomes is known to be associated with dosage compensation of the X chromosome, such that the activity of most X-linked genes is effectively the same in males and females [3].

The X and Y chromosome are thought to have evolved from an ordinary pair of autosomes that stopped recombining with each other after acquiring a sex-determining role [1,2]. The accumulation of sexually antagonistic genes (i.e. genes that are beneficial in one sex but detrimental in the other) linked to the sex-determining genes favors the evolution of suppression of recombination between the nascent sex chromosomes. In the absence of recombination, these originally homologous chromosomes continue to differentiate. It is generally thought that the morphological differentiation between the sex chromosomes is a by-product of the degeneration of the chromosome that is present only in the heterogametic sex (i.e. the Y or W) and is thus completely sheltered from genetic recombination. Evolutionary theory predicts that a non-recombining genome is vulnerable to the accumulation of deleterious mutations, and will show lower levels of adaptation [4]. Both of these processes can, over evolutionary time, lead to the loss of function of most genes on Y chromosomes. The X chromosome, however, can still recombine in females and is thought to maintain most of its original genes. Thus, in this traditional view of sex chromosome evolution, Y chromosomes are characterized by a continuous loss of active genes, whereas little change is thought to have occurred on the X.

In this article, I cover recent advances in our understanding of sex chromosome evolution. In particular, molecular characterization of young Y-chromosomes in several non-model species has uncovered events in the initial stages of sex chromosome formation that support our traditional view on how sex chromosomes originated and evolved from ordinary autosomes. Research on more ancient sex chromosomes, however, has also added some unexpected twists to our knowledge of the dynamic nature of Y chromosome evolution and has challenged our notion that Y chromosomes merely represent degenerate Xs. Comparative genomic and expression analyses have also shown that the X chromosome is not simply a passive player in the process of sex chromosome evolution. Instead, several adaptive events have shaped the evolution of the X chromosome, in response to both its sex-biased transmission and the degeneration of the Y.

Loss of gene function on young Y-chromosomes
Well-studied Y chromosomes, such as those of humans or Drosophila melanogaster, retain few, if any, of their original genes and thus reveal little about the evolutionary pro-
cesses shaping their initial degeneration. Recent genomic analyses of diverse organisms with very young Y-chromosomes (i.e. Y chromosomes that still carry most of their original genes) have enabled researchers to tackle the question of how and why Y chromosomes degenerate. The molecular characterization of several young Y-chromosomes, originating from between 1 and 10 million years ago, has begun in both plants (e.g. papaya [5] and Silene latifolia [6]) and animals (e.g. three-spine sticklebacks [7], medaka fish [8], and Drosophila miranda [9]).

Probably the most extensively studied young Y-chromosome to date is the neo-Y of D. miranda [9,10]. This neo-sex chromosome was only formed about one million years ago by the fusion of an ordinary autosome to the ancestral Y-chromosome, causing more than 2000 formerly autosomal genes to become sex-linked. Genomic regions on the neo-X show few changes, and most genes are highly constrained at the protein level. Although the neo-Y of D. miranda still retains many of its original genes, it shows clear signs of degeneration. About one-third of all genes are non-functional on the neo-Y chromosome, containing frame-shift mutations or premature stop codons. Most other neo-Y genes show various degrees of degeneration, including an accumulation of amino-acid and unpreferred codon changes. In addition, non-coding DNA — both intergenic regions and intron sequences — shows less functional constraint on the neo-Y [9]. Transposable element insertions also accumulate on the neo-Y relative to the neo-X, suggesting that they play a key role in the early process of Y-chromosome degeneration [9]. Thus, a diverse array of molecular changes — from transposable element insertions and null-mutations at protein-coding genes to more subtle changes at coding and non-coding sites — are contributing to the observed degeneration of Y chromosomes. Most of these mutations accumulate in genes that are still transcribed from the neo-Y of D. miranda, suggesting that they probably have had deleterious effects on fitness [11]. Similarly detailed molecular analyses have yet to be performed in other systems, but a common finding from investigating young Y-chromosomes is that the male-specific region of nascent Y-chromosomes is enriched with repetitive sequences in comparison with other genomic regions. For example, recent sequence analyses in sticklebacks, medaka fish and papaya all show that these young Y-chromosomes are accumulating repetitive elements faster than their homologous X-linked regions are [5,7,8]. This suggests that transposable elements are an early colonizer of new Y-chromosomes.

Although progress has been made in identifying how Y-chromosomes degenerate (i.e. which molecular changes occur), an important question to be answered concerns the question of why they do so. The accumulation of deleterious mutations on the Y, and in the long term its complete degeneration, is thought to result from interference between selected mutations on the Y, a process that is linked to the lack of recombination on this chromosome (see Box 1). Models of either purifying selection against deleterious mutations (i.e. background selection, Muller’s ratchet) or the fixation of beneficial mutations on the Y (i.e. selective sweeps) can both, in principle, result in the observed fixation of deleterious mutations on evolving Y-chromosomes. Consistent with selection-driven models of Y-chromosome degeneration, neutral diversity at several Y chromosome systems is reduced in such diverse systems as plants [12], Drosophila [13,14], birds [15] and mammals [16,17]. In an attempt to distinguish between these models of Y-chromosome degeneration, researchers have employed population genetics techniques, because different models are expected to leave characteristic signatures in patterns of DNA sequence variability [18,19]. For example, patterns of molecular variation on the neo-Y chromosome of D. miranda appear to be incompatible with deleterious mutation models but are expected under recent positive selection [19]. Ironically, this suggests that Y-chromosome degeneration is the consequence of evolutionary improvements at a few loci — possibly ones that specialize in male-specific functions (see below) — at the expense of most other genes on this chromosome. It will be of great interest to study patterns of population-level variation on several other young Y-chromosomes to determine the relative importance of positive versus negative selection in causing Y-chromosome degeneration.

**Maintenance and gain of genes on old Y-chromosomes**

Although studies of very young sex chromosomes support the notion that Y chromosomes largely represent degenerate Xs, research on ancient Y-chromosomes has revealed a more dynamic view. In particular, the evolution of Y chromosomes appears to be characterized by the interplay of losses and gains of individual genes, large genomic regions and even entire chromosomes (Figure 1). For example, the Y chromosome shared by all mammals, including humans, originated about 300 million years ago (mya). However, comparative mapping and DNA sequence analysis clearly demonstrate that human sex chromosomes were not derived from a single evolutionary event, but instead comprise a mosaic of different strata that became incorporated into the sex chromosomes at different times [20,21]. The oldest stratum contains the SRY gene (the sex-determining gene in mammals) and is part of the original autosome that became the sex chromosomes in all mammals. The younger strata represent added chromosomes or chromosomal regions that became incorporated into the non-recombining region more recently [20,21]. Once recombination ceased for these new Y-linked regions, they became prone to the accumulation of deleterious mutations, resetting the degeneration process. Thus, younger
Population genetic models to explain the observed degeneration of Y-chromosomes have three important components: mutation, selection and genetic drift. Natural populations are subject to recurrent mutations of beneficial and deleterious effects, and natural selection will act to incorporate beneficial mutations while eliminating deleterious ones. However, in the absence of recombination, the action of directional selection at one locus can interfere with the action of directional selection at a linked locus. Thus, the net efficacy of natural selection on a non-recombining chromosome, such as the Y, is impaired; deleterious mutations are less efficiently purged, and beneficial mutations are less efficiently incorporated. This forms the basis for most population genetics models that have been proposed to explain the observed degeneration of Y-chromosomes.

(a) Accumulation of weakly deleterious mutations by background selection. In a large, non-recombining population at mutation-selection balance, only Y chromosomes free of strongly deleterious mutations will contribute to the ancestry of future generations. The effective population size \( N_e \) of the Y can therefore be greatly reduced in the absence of recombination [2]. The intensity of selection is proportional to \( N_e \) (i.e. selection is more efficient in a larger population). Thus, reductions in \( N_e \) increase the rate of fixation of weakly deleterious mutations on a proto-Y (that would be eliminated from the proto-X), leading — in the long-term — to its degeneration [2,48].

(b) Muller’s ratchet. This process involves the stochastic loss of the class of Y chromosomes carrying the fewest number of deleterious mutations from a finite population [49–51]. In the absence of recombination and back mutation, this class of chromosomes cannot be restored. The next best class then replaces it (i.e. the class of chromosomes with the next fewest number of deleterious mutations). This class can in turn be lost, in a succession of irreversible steps. Each such loss of a class of chromosomes is quickly followed by the fixation of a specific deleterious mutation on the Y [52].

(c) Genetic hitchhiking by favorable mutations. The spread of a favorable mutation in a population of non-recombining Y-chromosomes can drag to fixation any deleterious mutant alleles initially associated with it, as long as the chromosome still has a net fitness advantage [53,54]. Thus, the hitchhiking model requires that selection coefficients for beneficial mutations be larger than for deleterious alleles [53,54]. Successive adaptive substitutions on an evolving Y chromosome can lead to the fixation of deleterious mutations at many loci. Under this model, Y-chromosome degeneration reflects adaptation at a few loci, at the cost of most other genes on this chromosome.

(d) Lack of adaptation on the non-recombining Y chromosome. The rate of adaptation on a non-recombining chromosome can be greatly reduced, owing to interference of positive mutations with linked deleterious alleles [55]. If selection coefficients for beneficial mutations are of the same magnitude or smaller than those for deleterious mutations, only beneficial mutations on Y-chromosomes free of deleterious alleles can contribute to adaptation [47]. If loci on a proto-X continue to adapt, while their homologs on the proto-Y fail to do so, it can be advantageous to upregulate the well-adapted X-linked genes and to downregulate or eliminate their maladapted Y-linked homologs [47].

Models (a–c) assume that purifying selection against deleterious mutations is reduced on the Y, whereas model (d) assumes that positive selection for beneficial mutations is less efficient on the Y. Any accumulation of deleterious alleles on the proto-Y will reduce the fitness of their carriers (models (a–c)). As a result, there is an advantage to enhancing the activity of genes transcribed from the proto-X (i.e. dosage compensation), and reducing the activity of their maladapted homologues on the proto-Y. This is also true if the proto-Y simply fails to adapt as rapidly as the proto-X (model (d)).
strata are at an earlier stage of their degeneration and typically contain a larger fraction of genes that have not yet degenerated. Analogous strata of different ages have also been found on the sex chromosomes of the plant Silene [22] and in chicken [23].

Ancestral Y-chromosomes not only incorporate new material from autosomes but might even be completely replaced by a novel Y [24]. In Drosophila pseudoobscura, the second Drosophila species to have its genome sequenced, the X chromosome consists of two chromosomal arms, representing the fusion of the ancestral X chromosome arm — common to all members of the genus Drosophila — to an autosome about 18 mya. The homolog (the neo-Y chromosome) of this fused autosome has been subject to the same degeneration process suffered by the
true Y and has only a few dozen of its originally more than 2000 genes left [24]. The current Y-chromosome of the
D. pseudoobscura group appears to be derived mostly from this relatively recently formed neo-Y, because 10 of the 15
genes identified from the Y are homologous to the auto-
some that fused to the X [24]. In an unexpected twist,
genes originally on the ancestral Y-chromosome appear to
have become translocated to another chromosome, and
these genes are now autosomal. Strikingly, these translo-
cated genes have since purged much of the repetitive
DNA that they accumulated during their tenure on the Y-
chromosome [24], demonstrating the remarkable fluidity
of genome evolution.

Y chromosomes can also gain individual genes. For exam-
ple, genome sequence analyses of the human Y have
uncovered seven gene families whose closest paralogs are
autosomal, suggesting that these genes transposed to the
Y from an autosome. Most of these autosomal-derived
genes are predominantly expressed in testis, and their
location on the Y chromosome might confer a male-
specific benefit [21]. Two additional gene families with
testis-specific expression have evolved from X-linked
homologs [21]. Thus, the Y chromosome might become ‘masculinized’ by functional specialization of ancestral Y-
linked genes and preferential retention of transposed
genes with a male-benefit. An unexpected finding emer-
ging from the human Y-chromosome sequence was that
most gene families are organized as palindromes (i.e. two
similar repeats pointing in opposite directions, connected
by a spacer) [21,25]. Ongoing gene conversion between
these palindromes enables members of the gene families
to recombine (not with the X, but with their Y homologs),
counterbalancing the evolutionary forces responsible for
the degeneration of Y chromosomes [25]. Genome ana-
lysis of the D. melanogaster Y chromosome, which was
formed at least 60 mya, has revealed that all Y-linked
genes are probably derived from autosomes [26,27]. Most
of the Y genes in D. melanogaster are involved in sperma-
togenesis, and their location on the Y chromosome prob-
ably reflects their male-specific advantage, with possibly
deleterious effects for females. Gene transpositions are
also apparent on the Y of D. pseudoobscura, because five of
the fifteen genes identified on its Y come from a variety of
autosomal locations, and four of these recently transposed
genes have become pseudogenes [24]. This finding
highlights the fact that only a small fraction of genes transposed to the Y will be able to acquire a male-specific
function and be maintained by selection, whereas the vast
majority degenerates.

The not-so-passive X chromosome
Responding to sex-biased transmission
The X chromosome is characterized by two features that
distinguish it from ordinary autosomes. First, males only
carry a single X-chromosome (i.e. the X is hemizygous).
This means that each recessive allele that arises on the X
and which gives males a reproductive advantage is imme-
diately visible to positive natural selection, resulting in
the masculinization of the X chromosome (i.e. an accumu-
lation of male-beneficial genes [28]). Second, the X
chromosome spends two-thirds of its time in females, but
only one-third in males. This gives selection more oppor-
tunity to act on genes benefiting females, and predicts the
‘feminization’ of the X (i.e. an accumulation of female-
beneficial genes [28]). Thus, hemizygosity in males, and
sex-bias in transmission implies that the gene content of
the X chromosome might not simply reflect that of the
ancestral autosome from which it arose, but might instead
have evolved an excess of male-beneficial and/or female-
beneficial genes.

Whole-genome expression studies in different species
have revealed that the X chromosomes of mice, D. mel-
anogaster and Caenorhabditis elegans are enriched for genes
showing female-biased expression, consistent with the
idea that the X chromosome becomes feminized in
response to its female-biased transmission [29,30,31].
Interestingly, there is also an apparent paucity of genes
on the X chromosome showing male-biased expression
[29,30,31], contrary to the expected accumulation of
positively selected recessive male-beneficial genes.

The reason for the deficit of male-biased genes on the
X might be the consequence of a process known as
meiotic sex-chromosome inactivation (MSCI), in which
the sex chromosomes become heterochromatic and tran-
scriptionally inactive during male meiosis [29,30,31].
MSCI might create a selective pressure for male-biased
genomes that are transcribed during male meiosis to be
located on autosomes, which are not inactivated. Support
for this hypothesis comes from a study in mice focusing on
cells in two different stages of spermatogenesis: mitotic
male germ cells in which the sex chromosomes are not yet
inactivated, and meiotic male germ cells that display sex
chromosome inactivation [29]. Genes expressed in male
germ cells before the onset of MSCI are over-represented
on the X chromosome, whereas genes expressed after
MSCI are under-represented [29]. This suggests that the
X chromosome might indeed be masculinized, but that
this effect is counterbalanced by MSCI, preventing genes
that are expressed later in spermatogenesis from residing
on the X. Whether a similar enrichment for early-sper-
matogenesis genes exists on the D. melanogaster and C.
elegans X remains to be seen. However, in Drosophila,
somatically expressed male-biased genes are also under-
represented on the X [32], which cannot be explained by
MSCI. This suggests the action of additional evolutionary
forces causing a deficit of male-biased genes on the X, at
least in Drosophila.

Feminization of the X chromosome could be achieved in
two different ways: through gene transpositions (i.e.
recruiting female-specific genes to the X and relocating
male-specific genes to the Y or autosomes) or through the
specialization in function of genes already present on the ancestral X-chromosome. Genome analyses in humans, mice and Drosophila have revealed that genes with testis-biased expression exhibit a unidirectional excess of retrotransposition from the X to the autosome [33,34]. These patterns of gene movements can result in X chromosomes becoming masculinized in comparison with the rest of the genome. However, the gene content of the X chromosome in both mammals and Drosophila is generally very stable [35], suggesting that feminization should mainly reflect changes to native X-linked genes, rather than physical movements of genes on and off the X. Empirical evidence for the functional specialization hypothesis comes from D. miranda [10,14]. Two genes in D. miranda that show an extremely high rate of protein evolution on the neo-X driven by recurrent positive selection have potentially sexually antagonistic functions (i.e. both genes show high expression levels in testis and ovaries and are related to male and female fertility). The adaptive evolution detected at these genes supports a model in which sex-related genes are free to undergo sex-specific functional specialization on newly evolving sex chromosomes, with the X-linked copy of such a gene specializing in its female-specific role [10,14]. More empirical work is necessary to establish whether most sex-related genes indeed undergo selection for functional specialization on newly evolving sex chromosomes.

Responding to Y-chromosome degeneration

The degeneration of the Y chromosome creates the problem of reduced gene-dosage of X-linked genes in males. Although females have two doses of each X-linked gene, males — after the Y-linked copy degenerates — only have one. Many genes on the X chromosome, however, will frequently have the same optimal level of expression in both sexes. This implies strong selective pressure to evolve compensatory mechanisms to upregulate the level of products of X-linked genes in males. Indeed, in several cases, the possession of genetically eroded Y-chromosomes is known to be associated with dosage compensation, such that the activity of most X-linked genes is effectively the same in males and females [36]. Dramatically different dosage-compensation mechanisms have evolved in different organisms: mammals upregulate the X approximately twofold in both sexes but inactivate one of their two X chromosomes in females [37,38]; XX hermaphrodite Caenorhabditis elegans effectively halve the expression from each X [39]; and male Drosophila increase the transcription of their single X approximately twofold [40,41].

In the genus Drosophila, the age of a Y chromosome — and thus its level of degeneration — broadly correlates with the extent to which dosage compensation on the X has evolved [42,43]. The homologs of the almost completely degenerated Y of D. melanogaster and neo-Y of D. pseudoobscura appear to be fully dosage compensated, whereas the neo-X chromosome of D. miranda shows only partial dosage compensation, consistent with partial degeneration of its neo-Y [42,43]. A similar relationship has been detected among the evolutionary strata of the human sex chromosomes [44*]; about 15% of X-linked genes escape X-inactivation in females, and most of them were found on the younger strata of the X-chromosome. This suggests that these regions have not had sufficient time to fully acquire the X-inactivation system.

Despite these general trends, little is know about the evolution of dosage compensation and the link between Y-chromosome degeneration and the acquisition of dosage compensation. In principal, dosage compensation could evolve on a ‘gene-by-gene’ basis (i.e. whenever a gene becomes malfunctional on the evolving Y, the homologous copy on the X acquires dosage compensation [45]), or by upregulating entire blocks of adjacent genes — perhaps owing to changes in chromatin structure [46]. Under the latter ‘block-by-block’ model of dosage compensation, many genes might not be expressed at their optimal level; a gene might be dosage compensated on the X but still transcribed from the Y. The process by which dosage compensation evolves has important implications for the dynamics of Y-chromosome degeneration. Under the gene-by-gene model, the evolution of dosage compensation would have little influence on the process or the speed of Y-chromosome degeneration, whereas under the block-by-block model, large regions of the Y chromosome would be rendered functionally redundant. Fully functional Y-linked genes that are dosage compensated on the X can accumulate deleterious mutations in a neutral manner, or might even be inactivated by positive selection to restore proper gene-dosage [47]. In the latter case, Y-chromosome degeneration might actually become adaptive, after genes have achieved dosage compensation on the X. Thus, to fully understand the evolutionary processes governing the fate of Y chromosomes, it is necessary to study the X in concert with the Y. We are only beginning to ask how the X responds to degenerative changes on the Y.

Conclusions

Recent research has substantially increased our understanding of the evolutionary processes shaping sex chromosomes. In particular, the organization of Y chromosomes reflects a combination of several evolutionary processes: the degeneration of the bulk of genes that were originally common on the primeval XY chromosome pair; subsequent additions of genes, genomic regions or chromosomes to the non-recombining region of the sex-chromosomes; and the accumulation of genes with male-specific functions. Although genome analyses have enabled us to reconstruct the evolutionary history of sex chromosomes and to describe the degeneration process at the molecular level, our understanding of why Y
chromosomes degenerate is still limited. The X chromosome plays an active part in sex chromosome evolution, and the gene content of the X has evolved to reflect its female-biased transmission. To counterbalance the deleterious effects associated with degeneration of Y-linked genes, the X chromosome has evolved dosage compensation. The evolutionary interplay between Y-chromosome degeneration and dosage compensation, however, has not been intensely studied to date.

Acknowledgements

I thank Peter Andolfatto, Brian Charlesworth, Andy Clark, Isabel Gordo, Kevin Thornton and Amanda Larracuente for valuable suggestions on the manuscript. This work was supported by the National Institutes of Health grant GM076007-01.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest


meiosis were found to be enriched on the mouse X-chromosome, as predicted under models of sexually antagonistic alleles, whereas those expressed later in spermatogenesis are depleted, probably owing to sex chromosome inactivation in meiosis.


Dosage compensation in mammals involves the inactivation of an X chromosome in females. This study shows that approximately 15% of X-linked genes that cluster in the younger regions of the X chromosome escape inactivation to some degree. This suggests that these genes have not had enough time to fully evolve dosage compensation.


