The evolution of dosage-compensation mechanisms

Ignacio Marín,[1] Mark L. Siegal,[2] and Bruce S. Baker[2]*

Summary
Dosage compensation is the process by which the expression levels of sex-linked genes are altered in one sex to offset a difference in sex-chromosome number between females and males of a heterogametic species. Degeneration of a sex-limited chromosome to produce heterogamety is a common, perhaps unavoidable, feature of sex-chromosome evolution. Selective pressure to equalize sex-linked gene expression in the two sexes accompanies degeneration, thereby driving the evolution of dosage-compensation mechanisms. Studies of model species indicate that what appear to be very different mechanisms have evolved in different lineages: the male X chromosome is hypertranscribed in drosophilid flies, both hermaphrodite X chromosomes are downregulated in the nematode *Caenorhabditis elegans*, and one X is inactivated in mammalian females. Moreover, comparative genomic studies demonstrate that the *trans*-acting factors (proteins and non-coding RNAs) that have been shown to mediate dosage compensation are unrelated among the three lineages. Some tantalizing similarities in the fly and mammalian mechanisms, however, remain to be explained. ***BioEssays*** **22:1106–1114, 2000.**

Introduction

In numerous eukaryotic organisms, the two sexes have different chromosomal constitutions. Typically, one sex, termed heterogametic, has a pair of morphologically different chromosomes, whereas the other sex, termed homogametic, has two identical members of each chromosomal pair. Morphologically distinct sex chromosomes are believed to derive from an initially identical chromosome pair, with morphological differentiation a by-product of the gradual loss of gene functions on the chromosome that is present only in the heterogametic sex (e.g., the Y in species with XY males and XX females).[1] This chromosome degeneration creates a significant genetic problem. Those genes on the sex chromosome that is not sex-specific (e.g., the X in XY/XX species) will frequently have the same optimal level of expression in both sexes.[2] The homogametic sex, however, has two copies of these sex-linked genes, while the heterogametic sex has, for many or all of them, only one. For reasons that we detail below, this situation creates a strong pressure to evolve compensatory mechanisms to equalize the level of products of sex-linked genes in both sexes. First discovered by Muller in Drosophila[2,3] such mechanisms are known as ''dosage compensation''. Dosage compensation is a particularly interesting problem in terms of gene regulation because it brings about the sex-specific coordinate regulation of a large group of genes whose only common feature is that they are physically linked but, for the most part, are functionally unrelated.

How dosage-compensation mechanisms have evolved in different lineages is the subject of this review. Dosage-compensation has been extensively studied in the fly *Drosophila melanogaster*, the nematode *Caenorhabditis elegans* and mammalian species. Our understanding of how dosage-compensation has evolved rests heavily on the information derived from studies of these model species. However, the knowledge provided by population genetics and comparative genomics is also necessary to understand the evolution of dosage compensation. The available evidence suggests that the extant dosage-compensation systems are relatively ancient (e.g., the mammalian dosage-compensation system arose at least 170 million years ago). They have been reused multiple times within a lineage to accommodate changes in the sex chromosomes that led to successive, evolutionarily independent, episodes of chromosome degeneration. Recent data show that several of the dosage-compensation genes have orthologs in many other species, including yeasts. Some of these genes have been co-opted for dosage compensation while retaining their ancestral functions. Others, which encode dosage-compensation-specific proteins, seem to have arisen by duplication of genes involved in other processes. The fact that the genes encoding components of the dosage-compensation machinery are different in flies, nematodes and mammals suggest that these mechanisms arose after these three lineages diverged. However, some profound similarities have been discovered in the mammalian and drosophilid mechanisms. It remains possible that a common dosage-compensation mechanism was present before the split between protostomes and deuterostomes.

[1]Departamento de Genética, Universidad de Valencia, Valencia, Spain.
[2]Department of Biological Sciences, Stanford University, Stanford, California.
The first two authors contributed equally to this paper.
*Correspondence to: Bruce Baker, Department of Biological Sciences, Stanford University, Stanford, CA 94305. E-mail: bbaker@cmgm.stanford.edu

Chromosome degeneration leads to selective pressure to dosage compensate

Since chromosome degeneration creates the need for dosage compensation, it is necessary to understand how degeneration occurs to grasp the conditions under which a dosage-compensation mechanism evolves. Although the forces underlying chromosome degeneration are still not fully understood, a general outline, supported by both theoretical and empirical studies, is clear.

A first step for chromosomal degeneration is for a chromosome to become limited to one sex, which may happen in several ways. Sometimes, this is due to changes in the sex-determination hierarchies.[4] In other cases, the cause is a chromosomal rearrangement (a Robertsonian fusion or translocation involving the sex chromosomes, see below). A simple example is the emergence of a new, dominant sex-determining gene. Obviously, in this situation, the chromosome carrying such a dominant gene becomes sex-limited. Once a chromosome becomes sex-limited, whatever the proximate cause, restriction of recombination with its homolog is likely to ensue.[1] Restricted recombination in turn favours the accumulation of mutations on the sex-limited chromosome, although the precise evolutionary forces leading to this accumulation are still under discussion and may differ between species.[1,5–7] The end result, after millions of years, is the loss of essentially all genes on the sex-limited chromosome. Two of the best studied cases of chromosome degeneration concern neo-Y chromosomes in drosophilid species (Fig. 1) and the human Y chromosome (Fig. 2).

Because dosage compensation will be selected for to the extent that loss of genes on the sex-limited chromosome is deleterious, it is instructive to examine classical cytogenetic and population-genetic studies that have addressed the fitness costs associated with such loss. Studies in diverse eukaryotic organisms have shown that the balance of products derived from the different genes is critical. Thus, aneuploidy (e.g., monosomy for single chromosomes) is highly deleterious, as is hemizygosity for subregions of chromosomes (i.e., heterozygous deficiencies), while polyploidy, where all the chromosomes remain in equal number, is much better tolerated. The deleterious effects of aneuploidy are, in general, not due to the presence of occasional individual genes for which aneuploidy has severe effects, but rather are due to the cumulative, small effects of aneuploidy at many (perhaps most) genes. Indeed, in a diploid, reducing the copy number of a wild-type allele at one locus from two to one is frequently sufficient to have negative fitness consequences. In a series of studies, Crow, Mukai and their colleagues examined the heterozygous effects of recessive lethal mutations. As such mutations are usually loss-of-function alleles, they provide a good model for what is likely to happen during sex-chromosome degeneration. These studies found that, in general, recessive lethals are in fact partially dominant, with selection coefficients against heterozygotes averaging 2–3%.[8] Thus, reducing the dosage of a wild-type gene by half is generally deleterious to a degree easily acted on by selection. Based on these findings one would expect that, during sex-chromosome degeneration, selection acts to ameliorate the effects of the loss of genes from the degenerating chromosome on a gene-by-gene basis as such losses occur across evolutionary time. (As a historical aside, it is worth noting that long before the data discussed in this section existed, Muller[2] argued that dosage compensation had been selected for at individual genes, based on phenotypic differences not readily discernable to us, but detected by natural selection, between individuals with one versus two copies of a gene.) This reasoning, while asserting the sufficiency of single-gene hemizygosity to create selective pressure for compensation, does not preclude the acquisition of dosage compensation simultaneously by closely linked genes. For example, if large deletions encompassing several genes contribute to chromosome degeneration, then perhaps dosage compensation will be acquired simultaneously by all genes within the newly hemizygous region.

Dosage compensation in flies, nematodes, and mammals

Here we briefly summarize the salient features of dosage compensation in flies, nematodes and mammals as background for discussion of the evolution of dosage-compensation mechanisms.

D. melanogaster

In *D. melanogaster* (males XY, females XX), dosage compensation occurs by doubling the transcription rate of X-linked genes in males.[9,10] Dosage compensation in flies is mediated by the protein products of five known genes [*maleless* (*mle*), the *male-specific lethal* (*msl*) genes *msl-1*, *msl-2*, and *msl-3*, and *males absent on the first* (*mof*)], collectively referred to as the *msl* genes, together with the non-coding RNA products of two additional genes, *roX1* and *roX2* (*roX* = *RNA on the X*).[11–13] The *msl* genes were discovered in screenings for male-specific lethals,[14–17] while the ROX RNAs were found when looking for sex-specific transcripts in the nervous system.[18,19]

There is substantial molecular and genetic evidence that the products of all of these genes function together in a complex, termed the compensasome,[20] to mediate dosage compensation by altering the chromatin structure of the male X chromosome.[9,10,12] The products of these genes are all specifically associated with the same set of hundreds of positions along the male X chromosome. Moreover, all of the MSL proteins and at least one of the ROX RNAs must be present for the association of the compensasome with the characteristic set of sites along the X chromosome[20] (at a small number of sites [approx. 30–40] incomplete compensasomes can form in the absence of particular MSL proteins;

Refs. 21–23). The requirement that all the MSL proteins be present for the compensasome to form allows the process to be sex-specifically regulated by controlling the production of just one component. Indeed, translation of *msl-2* transcripts is repressed by the protein product of the female-specific, master sex-determination gene, *Sex-lethal* (*Sxl*).[24] Therefore, in females, MSL-2 protein is not generated, so active compensasomes do not form.

A possible mechanism by which the compensasome alters chromatin structure (thereby leading to hypertranscription) was revealed by the discovery that an isoform of histone H4 acetylated at lysine 16, H4Ac16, is enriched at a set of sites on the male X chromosome whose locations correlate with the sites bound by compensasomes.[25,26] Moreover, the *mof* gene encodes a histone acetyltransferase[17] and, recently, it has been shown that a partially purified complex containing the five MSL proteins and ROX2 RNA is able to acetylate histone H4 specifically at lysine 16,[27] as is recombinant MOF alone.[28] Post-translational modification of histones, and particularly their acetylation, has been implicated in chromatin activation leading to increased transcription.[29,30]

Possible insight into how compensasomes locate the male X chromosome has come from the findings that, when a *roX* transgene is inserted on an autosome, compensasomes are recruited to this site, and neighboring sites sometimes show compensasome binding as well. These and other findings have led to the proposal that the *roX1* and *roX2* genes themselves may act as chromatin-entry sites from which epigenetic spreading of the compensasomes proceeds along the chromosome.[31] However, it is known that some X-linked genes are not compensated, despite the fact that their Y-linked homologs have degenerated.[9] If the model of epigenetic spreading of the compensasome from a limited number of chromatin-entry sites is correct, it must somehow incorporate these cases in which genes appear to be skipped over by the spreading.

The molecular nature of compensasome-binding sites (or that of sites facilitating spreading) remains mysterious. It is known that the *Drosophila* X chromosome has some molecular characteristics that make it different from the autosomes, particularly enrichment in some mono- and dinucleotide repeats[32–34] and certain satellite-related repeats.[35,36] Whether any of these features is related to compensasome binding or spreading is, however, unknown. An alternative approach for finding compensasome-binding sequences would be to identify an X-chromosome-derived transgene that, when inserted into an autosome, is dosage compensated (and subsequently to identify those sequences within the transgene that are necessary and sufficient for compensation). However, most experiments using this approach were inconclusive.[9] An important problem is that those experiments did not take into account the existence of *cis*-acting, repressive effects of the autosomal chromatin on transgene expression. These effects became evident once the transgenes were insulated.[37] However, a systematic screening for compensasome-binding sequences using insulated transgenes remains to be performed.

Finally, we note there is evidence that not all dosage compensation is achieved through compensasome action.[9,10] For example, the X-linked gene *runt,* which is expressed in early embryos, has been shown to be compensated in an *msl*-independent, but *Sxl*-dependent manner. Interestingly, a search for matches to the optimal SXL binding site in the 3′ untranslated regions of 1324 *Drosophila* genes yielded 21 genes with three or more such sites, 20 of which are X-linked (one of which is *runt*; the one autosomal gene was *msl-1*).[38] Thus, SXL may directly compensate some genes by down-regulating their expression in females.

C. elegans

In *C. elegans* (males X0, hermaphrodites XX), dosage compensation is achieved by downregulating transcription of genes on both hermaphrodite X chromosomes.[10,39,40] Among the regulators of dosage compensation are a subset of *dumpy* (*dpy*) genes (*dpy-21*, *dpy-26*, *dpy-27*, *dpy-28* and *dpy-30*). Mutations in these genes are maternal-effect XX-specific lethals and lead to elevated X-linked transcript levels in hermaphrodites.

Similar to the situation in flies, the *C. elegans* dosage-compensation genes are under the control of the master sex-determining gene, *XO lethal-1 (xol-1)*. However, this control is not direct in *C. elegans*. Instead, *xol-1*, which is only expressed in males, negatively regulates the *sex determination and dosage compensation* (*sdc*) genes *sdc-1*, *sdc-2*, and *sdc-3*. These three genes in turn negatively regulate *her-1*, which is at the top of the sex-determination branch of the regulatory hierarchy. As their name implies, the *sdc* genes also regulate the dosage-compensation branch of the hierarchy. In this respect, the role of *sdc-1* is unclear because null mutations in it do not cause XX-specific lethality although they do lead to overexpression of X-linked genes. However, the role of the other *sdc*'s is clear from molecular and genetic analysis. Together with *dpy-30*, which encodes a ubiquitously expressed nuclear protein, *sdc-2* activates *sdc-3*. Both SDC-2 and SDC-3 localize specifically to the X chromosome in hermaphrodites, but SDC-2 is able to bind in absence of SDC-3, suggesting a critical role of SDC-2 in X-chromosome targeting. Moreover, only when SDC-2 and SDC-3 are present, are DPY-26, DPY-27, DPY-28, and the product of another gene, MIX-1, also localized to the X chromosome. Levels of these proteins are mutually dependent, suggesting they form a complex or they interact in some way that improves their stability.[39,40]

It is not known how these proteins recognize the X, nor how they effect down-regulation of X-linked genes in hermaphrodites. However, the DPY-27 and MIX-1 proteins are members of the SMC (structural maintenance of chromosomes) family

of proteins (see below). Yeast SMC proteins are required for condensation and segregation of mitotic chromosomes, and vertebrate SMCs act in chromosome condensation as well. Thus, it is reasonable to think that dosage compensation in *C. elegans* is achieved by increasing condensation of the hermaphrodite X chromosomes.

Mammals

In mammals (most often, males XY, females XX), dosage compensation is achieved by inactivation of one of the female X chromosomes.[39,41–43] However, marsupials (and probably monotremes) have a system of X inactivation with some differences from eutherian mammals.[41] In particular, in embryonic tissues the X chromosome to be inactivated is chosen randomly in eutherians, while in marsupials it is the paternal X that becomes inactivated (in marsupials and at least the mouse the paternal X is inactivated in extraembryonic tissues). Moreover, inactivation is incomplete in marsupials, with different loci of the paternal X being inactive in different tissues. Thus, the following details apply to eutherians, where the process is best understood.

In eutherian mammals, X chromosome inactivation has several steps. First, X chromosomes are counted and the chromosome to inactivate is chosen. Then, local initiation of the inactivation process and subsequent spreading of inactivation along the chromosome occurs. Finally, the inactive state must be maintained. Inactivation requires the presence of the X inactivation center (Xic), a locus that, when active, functions in *cis* to inactivate the X chromosome. Within the Xic is the *Xist* gene, which produces a large, non-coding RNA, which spreads in *cis* along the X that expresses it, coating that chromosome and leading to its inactivation. The processes, previous to inactivation itself, of X chromosome counting and choice of the chromosome to inactivate are still incompletely understood.[39,43] An intact *Xist* gene is not required for correct counting. Choice involves a switch from unstable transcription of *Xist* from both X chromosomes to stable transcription from one X and silencing of the other. Recently, a regulator of *Xist*-mediated silencing, the *Tsix* gene, located 15 kb 3′ to Xist and so-named because it produces a non-coding, antisense transcript with respect to *Xist*, has been discovered.[44,45] A mutation in *Tsix* does not affect chromosomal counting, but it biases choice: the chromosome that carries the mutation is preferentially inactivated.[44] The available data suggest that, preceding inactivation, there is a transition from *Tsix* being expressed biallelically to its being expressed only from the allele on one X chromosome, and that this blocks *Xist* accumulation on that chromosome, the future active X.[45]

From studies of deletions of the *Xist* locus and ectopic expression from autosomal transgenes, it has been demonstrated that *Xist* is both necessary and sufficient to establish X inactivation. However, *Xist* is not required to maintain inactivation.[46] Thus, it is likely that *Xist* acts during a narrow period of development to establish an inactive chromatin state that is then maintained epigenetically by other factors. The inactive X has several properties distinguishing it from the other mammalian chromosomes: it is replicated late in S phase, methylated at CG dinucleotides, enriched in the histone H2 variant macroH2A, and relatively impoverished of acetylated isoforms of histones H2A, H3, and H4.[39,47] In particular, promoters of the inactivated genes are hypoacetylated.[48] It is likely that some of these features are related to either establishment or maintenance of the inactive state.

Further elucidation of the mode of action of *Xist* and *Tsix* and of the local patterns of chromatin modification will no doubt shed light on the mechanism of chromosome inactivation. Particularly important is the characterization of the *trans*-acting factors that must mediate counting, influence chromosome choice and/or determine spreading of *Xist*. Models of chromosome inactivation must also explain how a significant number of genes seem to be skipped over by the spreading of X inactivation. In a recent survey of 224 X-linked human genes, 34 were found to be transcribed from both the active and the inactive X chromosome.[49] The genes escaping inactivation are not randomly distributed (Fig. 2), possibly reflecting the evolutionary history of the mammalian X (discussed below). Interestingly, a recent study[50] suggests that the distribution of the non-LTR retrotransposon LINE-1 may contribute to the determination of which regions are inactivated and which ones are skipped by inactivation along the X chromosome.

Origin and evolution of the model dosage-compensation mechanisms

When the dosage-compensation mechanisms observed in the model species arose is still unclear. Direct data on the phylogenetic ranges of the extant systems are scant. In the case of drosophilid species, it has been shown that the *msl*-based dosage-compensation system is at least 50–60 million years old, being present in at least one drosophilid genus (*Chymomyza*) other than *Drosophila*[51] [*Zaprionus* and *Hirtodrosophila*, which we considered as different genera in our original report, may belong in the *Drosophila* genus, Ref. 52]. It is unknown whether other dipteran species have a similar system. All mammalian species, including marsupials and monotremes, have X-chromosome inactivation (although with slightly different characteristics, see above), but birds are believed to lack the *Xist*-based dosage-compensation system, so this system most likely emerged after the split of the avian and mammalian lineages, but before the differentiation of the mammalian Orders (between 170 and 310 million years ago). Finally, comparative data for nematodes are unavailable at present.

An interesting alternative approach is to ask whether the genes known to act in dosage compensation in the model species can be detected in other organisms by comparative

genomic analysis. Of particular interest is whether there are orthologs of the dosage-compensation genes in distantly related species.

Three of the five *Drosophila msl* genes belong to characterized gene families. The exceptions are *msl-1* and *msl-2*, for which clear homologies to other genes have not been found. MLE belongs to the DEAH family of proteins, some of which have roles in splicing.[53] It has been shown that both a close relative of MLE, mammalian RNA helicase A, and MLE itself, have DNA and RNA helicase activities.[54,55] The characterization of *mof* revealed similarity to histone acetyltransferases of the MYST family.[17] Finally, in a recent study, it has been shown that *msl-3* is a member of a previously unknown gene family.[56]

Phylogenetic analyses of the *mle*, *mof*, and *msl-3* gene families have shown that related genes exist in the yeast *S. cerevisiae*, and are found in nematodes and mammals as well[56] (Sanjuán and I. M., unpublished data). It is unclear, however, whether a MSL complex exists in any of these species and, at least for yeasts and nematodes, it seems unlikely. MLE belongs to a monophyletic subfamily of the DEAH proteins, with a single member in yeasts, but several in animal species. Thus, it appears that diversification of this DEAH subfamily may have made possible the specialization of the function of MLE and its orthologs in animals (Sanjuán and I. M., unpublished data). The existence of a gene duplication of *msl-3* in *Drosophila*, but not yeasts or *C. elegans*, reveals another potentially important functional distinction. In particular, the single *msl-3*-related gene in *C. elegans* is more similar to the *Drosophila* paralog of *msl-3* (called *MRG15)* than to *msl-3* itself, suggesting that nematodes lack a key component of the MSL complex of *Drosophila*.[56] A similar situation is seen for *mof*, which has a closely related duplicate gene in *Drosophila*, but only single yeast and nematode homologs, (Sanjuán and I. M., unpublished data).

An altogether different situation is found in mammals. We have confirmed that two previously characterized genes (*MSL3-L1* and *RNA helicase A*) are the mammalian orthologs of *msl-3* and *mle*, respectively[56] (Sanjuán and I. M., unpublished data). Moreover, a likely ortholog of *mof* also exists in mammals.[57] The three putative mammalian orthologs of the *Drosophila* dosage-compensation genes encode proteins structurally similar to their *Drosophila* counterparts. Thus, unlike the situation in yeasts or nematodes, and assuming as-yet-unidentified *msl-1* and *msl-2* orthologs are present in mammals, it is possible that a complex identical to that found in *Drosophila* is formed in mammalian species.

With respect to *C. elegans*, dosage-compensation regulators DPY-27[58] and MIX-1[59] have clear similarities to proteins of the SMC family. This family is generally represented in eukaryotes by four genes (orthologous to the yeast *Smc1*, *Smc2*, *Smc3* and *Smc4* genes).[60–62] *mix-1* is a *Smc2* ortholog, while *dpy-27* is a *Smc4* ortholog. In addition to *dpy-27*, there is also a second SMC4 class gene in *C. elegans*.[61,63] So far, this is the only duplicated gene found in the whole SMC family in any eukaryotic species (unpublished data). This finding suggests that *dpy-27* has a relatively recent origin. Significantly, although complexes (called "condensins") containing both SMC2 and SMC4 proteins have been shown to be involved in mitotic chromosome condensation both in yeasts and *Xenopus,* the lack of a vital function for DPY-27 outside of dosage compensation suggests that it is not participating in a mitotic-chromosome protein complex. Instead, the *dpy-27* paralog is the likely candidate to be the SMC4 protein involved in mitosis in *C. elegans*. In contrast, *mix-1*, which encodes the sole *C. elegans* SMC2 class gene, has roles in both dosage compensation and mitotic chromosome segregation.[39,40]

In sum, the evidence suggests that the dosage-compensation complexes of *Drosophila* and *Caenorhabditis* arose as novel combinations of proteins, rather than by co-opting complete, preexisting complexes. Apparently, some proteins acquired new functions in dosage compensation while retaining an ancestral function (becoming pleiotropic; e.g., MIX-1, as well as DPY-26 and DPY-28, which are also required in meiosis) while others became specialized in dosage compensation following gene duplication (e.g., DPY-27 and MSL-3).

Insights into dosage compensation from studies of sex-chromosome evolution

Further insight into the evolution of dosage-compensation systems comes from an examination of the interplay of sex-chromosome evolution and dosage compensation in Diptera and mammals. One way that new sex chromosomes form is when translocations occur between a preexisting sex chromosome and an autosome. For example, a Robertsonian centric fusion between an X chromosome and an autosome restricts the untranslocated homolog of this autosome to males, thus creating a neo-Y chromosome. In such cases, one would expect that, as the neo-Y chromosome degenerates, its homolog will become dosage compensated.

In *Drosophila*, several instances of Robertsonian centric fusions are known (Fig. 1) and the degree to which the new arm of the X chromosome is dosage compensated is strikingly correlated with the amount of degeneration on its homologous neo-Y. At one extreme is *D. americana*, where the neo-Y chromosome is of sufficiently recent origin that it has not degenerated (Fig. 1B). Thus, there is no selective pressure to dosage compensate the homologous arm of the X. Consistent with this expectation is the observation that the MSL proteins bind only to the old arm of the *D. americana* X chromosome, and not to the new arm.[51,64] At the other extreme, in *D. pseudoobscura*, *D. willistoni* and *D. robusta*, species where evolutionarily independent X-A translocations occurred, the neo-Y fully degenerated and was lost (Fig. 1C). In all three species, it has been demonstrated that the MSL proteins bind

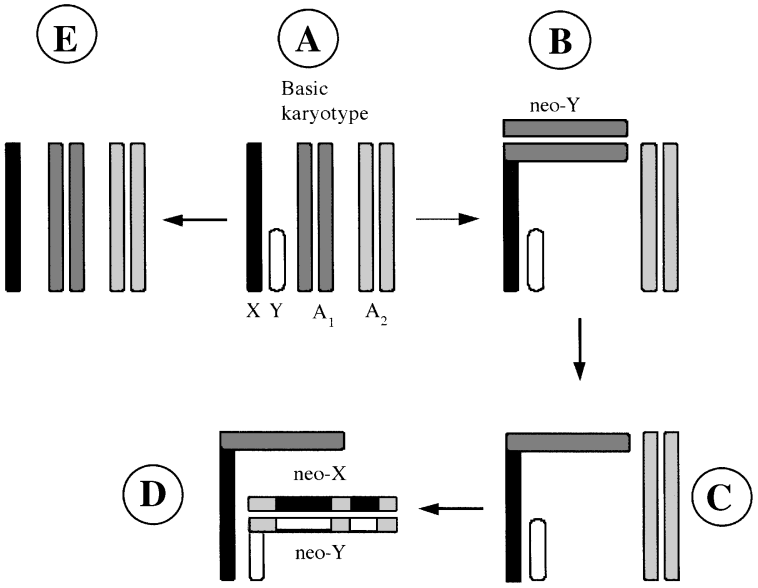

Figure 1. Evolution of the sex chromosomes in the *Drosophila* genus. Male karyotypes are shown. **A:** The basic karyotype of *Drosophila* species. For simplification only two of the four autosomal arms are shown. **B:** the translocation of an autosomal arm to the X chromosome generates a metacentric X and a neo-Y chromosome. This karyotype is found in *Drosophila americana*, where the translocation is so recent that the neo-Y chromosome has not yet degenerated. Thus, there is no selective pressure to dosage compensate the homologous arm of the X. The MSL proteins only bind to the old arm of the *D. americana* X chromosome, not to the new arm.[51,64] **C:** after the X-A translocation, and once degeneration has been completed, the neo-Y chromosome may disappear. This karyotype is found in several *Drosophila* species. Three of them (*D. pseudoobscura*, *D. willistoni* and *D. robusta*) suffered independent translocations and, in all three, it has been shown that the MSL proteins bind to the new arm, and that the density of sites on the new arm stained by anti-MSL antibodies is comparable to that observed for the original X chromosome arm.[51] **D:** This peculiar karyotype has been observed in a single species, *Drosophila miranda*. This is a sibling of *D. pseudoobscura*; and thus it suffered the same X-A translocation and lost the corresponding neo-Y chromosome. However, in *D. miranda*, a second translocation, this time Y-A, occurred more recently, generating another neo-Y chromosome. Parts of this second neo-Y have degenerated and the corresponding parts of its homologous neo-X show increased transcription in males. Furthermore, the compensated regions of the neo-X show MSL binding, whereas the uncompensated regions do not.[51] **E:** while in *D. melanogaster* the Y chromosome carries genes required for male fertility, in some other species the Y chromosome has become completely dispensable and has been lost. Figure modified from.[51]

to the new arm, and that the density of sites on the new arm stained by anti-MSL antibodies is comparable to that observed for the *D. melanogaster* X chromosome.[51] Finally, *D. miranda* represents an intermediate case with a Y chromosome undergoing degeneration. *D. miranda*, which, like its sibling *D. pseudoobscura*, has lost a fully degenerated neo-Y, in addition has a more recent Y-autosome translocation that has produced a second neo-Y. Parts of this second neo-Y have degenerated and the corresponding parts of its homologous neo-X show increased transcription in males (Fig. 1D). Furthermore, the compensated regions of the neo-X show MSL binding, whereas uncompensated regions do not.[51] Similar results were obtained in *D. pseudobscura* and *D. miranda* using anti-H4Ac16 antibodies.[65] The finding that, in four independent cases where a need for dosage compensation of a new X chromosome occurred, the MSL-based system was recruited to new X chromosomes suggest that the recruitment of a preexisting system may be much more likely than the development of a new one. In addition, these results establish that the *Drosophila* dosage-compensation system can be older than the particular X chromosome on which it operates.

Mammalian sex chromosome evolution is more complex than the *Drosophila* cases just cited, because several translocations of autosomal material to the sex chromosomes have occurred since the eutherian–marsupial divergence about 170 million years ago (Fig. 2).[41] The probability of degeneration of Y-linked genes of modern eutherian species depends then on two factors. First, on whether those genes belong to the most ancestral section of the X chromosome, or were translocated to the sex chromosomes more recently. Second, on when the region that contains a particular gene became unable to recombine with the homologous region of the X chromosome, and thus started degenerating. By comparing the level of divergence of genes that exist in both X and Y chromosomes, Lahn and Page[66] concluded that the human Y chromosome has at least four distinct regions that started their degeneration at different times. The oldest region started degeneration shortly after the bird–mammal split (310 million years ago). The youngest may have started degenerating only 30–50 million years ago. Not surprisingly, most of the genes that are still active in both X and Y chromosomes are in the regions that most recently began degenerating (Fig. 2). Since all regions of the X that are dosage compensated are governed by the *Xist*-based dosage-compensation system, in mammals (as in the particular fly species described in Fig. 1) the extant dosage-compensation system predates the appearance of some regions of the sex chromosomes.

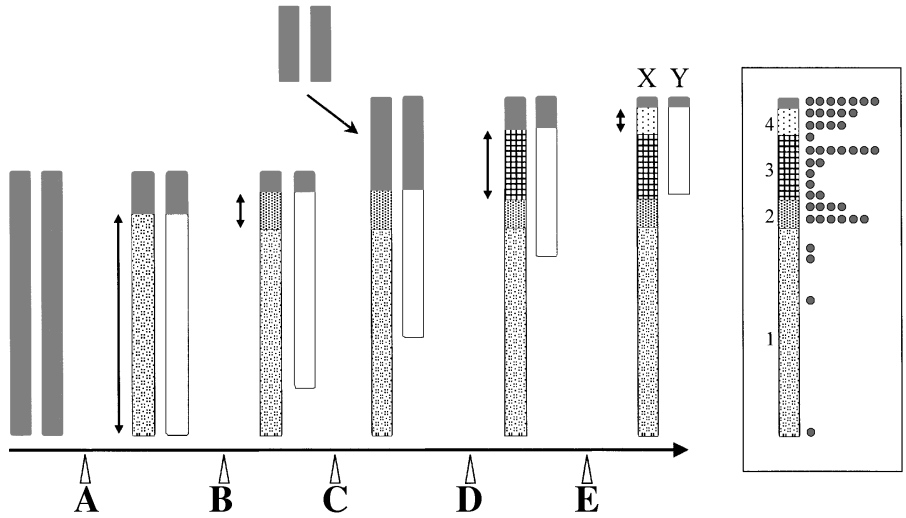

Figure 2. A model of the evolution of the sex chromosomes in the lineage giving rise to humans. The horizontal arrow at the bottom represents time. The pair of rods correspond to the X (left) and Y (right) chromosomes. Dark-grey shading corresponds to autosomal (or what is now defined as pseudoautosomal) material (for simplification only one pseudoautosomal region is shown). Vertical lines correspond to the regions that suffer chromosomal inversions. The evolution of these chromosomes starts about 320 million years ago, with a pair of autosomes (left pair of chromosomes, in grey). **A:** They evolve into sex chromosomes and suffer a first inversion, which restricts recombination between the X and Y. The non-recombining region of theY chromosome starts degenerating (light grey shows the degenerating regions). This first process has been estimated to have occurred about 240 to 320 million years ago. **B:** Second inversion, restricts further recombination between the X and the Y (130 to 170 million years ago). **C:** Addition of autosomal material to X and Y (80 to 130 million years ago). **D:** Third inversion (80–130 million years ago). **E:** Fourth inversion (30 to 50 million years ago). Today (right pair of chromosomes), four regions of the X can be recognized, according to the moment recombination with the Y chromosome became restricted, while the Y is almost totally degenerated. Adapted from Ref. 66. In the X chromosome shown in the box, we have schematized the locations of the genes known to escape inactivation in humans. Almost all them are in the pseudoautosomal region or in the more recently inverted regions (regions 2, 3 and 4).[49,68]

Final theoretical considerations and speculations

There are three aspects of the evolution of dosage compensation that deserve special discussion. First, if selection for dosage compensation occurs on a gene-by-gene basis, one can ask why are general mechanisms, acting on many genes, observed in the model species. Second, if dosage compensation is selected for, as discussed above, to ameliorate the deleterious effects of hemizygosity in the heterogametic sex, then the mechanisms of dosage compensation that would be expected to arise are ones that increase the level of expression of genes on the X chromosome in the heterogametic sex, as opposed to those that decrease the expression of X-linked genes in the homogametic sex. We may ask why then downregulating mechanisms in the homogametic sex are observed in *C. elegans* and mammals. Finally, we may ask whether certain unexpected similarities found between the drosophilid and mammalian mechanisms may have an ancient common origin or, alternatively, are the result of convergence.

The favoring of global solutions may seem at odds with the data that suggest that dosage compensation should be acquired on a gene-by-gene basis. However, recall that two elements are required for dosage compensation of an X chromosome: a *trans*-acting machinery that alters gene expression in one sex, and *cis*-acting sequences that attract this machinery (or facilitate its spreading) to appropriate X-chromosomal sites. Thus, gene-by-gene acquisition of dosage compensation may reflect local acquisition of such *cis*-acting sequences, a process that (as the mammalian and *Drosophila* data discussed in the previous section suggest) is likely to be molecularly simpler, and thus more probable, than the development of several independent dosage-compensation machineries. Thus, the extant systems may have arisen as gene-specific mechanisms that, by virtue of their being readily co-opted as degeneration proceeded, provided an important advantage to the organisms that possessed them. Therefore, they may have progressively spread to dosage compensate many genes, appearing today as general, chromosome-wide mechanisms.

With regard to the second issue raised above — that dosage-compensation mechanisms would have first evolved to increase the level of expression of genes on the single X chromosome in the heterogametic sex to restore the balance

of products of such chromosome with respect to the autosomes—the upregulating system in flies is compatible with this view, but the *C. elegans* and mammalian systems are not. These two downregulating mechanisms only exacerbate the deleterious effects of lowering gene dosage, by extending X chromosome low expression to both sexes. Nevertheless, one can envision an evolutionary route that could have led to the current mammalian and *C. elegans* systems.[1] Selection for dosage compensation in males in response to Y-chromosome degeneration could have led to an increase in X-linked gene expression that was not sex-specific. Such a situation, while beneficial to males, would be detrimental to the other sex, as the latter would be overexpressing X-linked genes. This situation would in turn have led to selection for downregulating compensation mechanisms that turned back down the X chromosomes in the homogametic sex, generating the systems seen in mammals and nematodes.[1] It is worth noting that such a scenario predicts that both nematodes and mammals should still have regulatory systems that upregulate the expression of the X chromosomes in both sexes. Some evidence for an X-specific upregulating system has been obtained in mouse species.[67]

Although the molecular mechanisms of dosage compensation in *Drosophila*, *C. elegans* and mammals are different enough to appear to have independent origins, there are some similarities that may reflect either an ancient, conserved mechanism, or convergence constrained by a limited set of possible solutions to the problem of coordinately regulating gene expression on an entire chromosome. Among the striking mechanistic similarities is the finding in *Drosophila* and mammals of genes that (1) are transcribed into noncoding RNAs that are essential components of the dosage-compensation machinery and (2) may act as entry sites from which epigenetic spreading of the machinery proceeds. In addition, in both flies and mammals, histone acetylation appears to play a central role in dosage compensation. It is not evident why dosage-compensation mechanisms found in such distant species would share such properties. Comparative data from other species, as well as more detailed mechanistic understandings of dosage compensation, may shed light on the significance of these suggestive findings.

References

1. Charlesworth B. The evolution of chromosomal sex determination and dosage compensation. Curr Biol 1996;6:149–162.
2. Muller HJ. Evidence of the precision of genetic adaptation. Harvey Lect 1950;43:165–229.
3. Muller HJ. Further studies on the nature and causes of gene mutations. Proc 6th Int Congr Genet 1932;1:213–255.
4. Marín I, Baker BS. The evolutionary dynamics of sex determination. Science 1998;281:1990–1994.
5. Charlesworth B, Charlesworth D. Rapid fixation of deleterious alleles can be caused by Muller's ratchet.Genet Res 1997;70:63–73.
6. Orr HA, Kim Y. An adaptive hypothesis for the evolution of the Y chromosome. Genetics 1998;150:1693–1698.
7. Yi S, Charlesworth B. Contrasting patterns of molecular evolution of the genes on the new and old sex chromosomes of *Drosophila miranda*. Mol Biol Evol 2000;17:703–717.
8. Crow JF, Simmons MJ. The mutation load in *Drosophila*. In: Ashburner M, Carson HL, Thompson JN. editors. The Biology and Genetics of *Drosophila* vol. 3c. London: Academic Press; 1983. p. 1–35.
9. Baker BS, Gorman M, Marín I. Dosage compensation in *Drosophila*. Annu Rev Genet 1994;28:491–521.
10. Cline TW, Meyer BJ. Vive la différence: males vs. females in worms vs. flies. Annu Rev Genet 1996;30:637–702.
11. Lucchesi JC. Dosage compensation: roX marks the spot. Curr Biol 1999; 9:R807–R808.
12. Stuckenholz C, Kageyama Y, Kuroda MI. Guilt by association. Noncoding RNAs, chromosome-specific proteins and dosage compensation in *Drosophila*. Trends Genet 1999;15:454–458.
13. Franke A, Baker BS. Dosage compensation rox!. Curr Opin Cell Biol 2000;12:351–354.
14. Belote JM, Lucchesi JC. Control of X chromosome transcription by the *maleless* gene in *Drosophila*. Nature 1980;285:573–575.
15. Belote JM, Lucchesi JC. Male-specific lethal mutations of *Drosophila melanogaster*. Genetics 1980;96:165–186.
16. Uchida S, Uenoyama T, Oishi K. Studies on the sex-specific lethals of *Drosophila melanogaster*. III. A third chromosome male-specific lethal mutant. Jpn J Genet 1981;56:523–527.
17. Hilfiker A, Hilfiker-Kleiner D, Pannuti A, Lucchesi JC. *mof*, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in *Drosophila*. EMBO J 1997;16:2054–2060.
18. Amrein H, Axel R. Genes expressed in neurons of adult male *Drosophila*. Cell 1997;88:459–469.
19. Meller VH, Wu KH, Roman G, Kuroda MI, Davis RL. roX1 RNA paints the X chromosome of male *Drosophila* and is regulated by the dosage compensation system. Cell 1997;88:445–457.
20. Franke A, Baker BS. The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in *Drosophila*. Mol Cell 1999;4:117–122.
21. Palmer MJ, Richman R, Richter L, Kuroda MI. Sex-specific regulation of the *male-specific lethal-1* dosage compensation gene in *Drosophila*. Genes Dev 1994;8:698–706.
22. Gorman M, Franke A, Baker BS. Molecular characterization of the *male-specific lethal-3* gene and investigations of the regulation of dosage compensation in *Drosophila*. Development 1995;122:463–475.
23. Lyman LM, Copps K, Rastelli L, Kelley RL, Kuroda MI. *Drosophila male-specific lethal-2* protein: structure/function analysis and dependence on MSL-1 for chromosome association. Genetics 1997;147: 1743–1753.
24. Bashaw GJ, Baker BS. Dosage compensation and chromatin structure in *Drosophila*. Curr Opin Genet Dev 1996;6:496–501.
25. Turner BM, Birley AJ, Lavender J. Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. Cell 1992;69:375–384.
26. Bone JR, Lavender J, Richman R, Palmer MJ, Turner BM, Kuroda MI. Acetylated histone H4 on the male X chromosome is associated with dosage compensation in *Drosophila*. Genes Dev 1994;8:96–104.
27. Smith ER, Pannuti A, Gu W, Steurnagel A, Cook RG, Allis CD, Lucchesi JC. The *Drosophila* MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. Mol Cell Biol 2000;20:312–318.
28. Akhtar A, Becker PB. Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. Molecular Cell 2000;5:367–375.
29. Kuo MH, Allis CD. Roles of histone acetylatransferases and deacetylases in gene regulation. Bioessays 1998;20:615–626.
30. Strahl BD, Allis CD. The language of covalent histone modifications. Nature 2000;403:41–45.
31. Kelley RL, Meller VH, Gorgadze PR, Roman G, Davis RL, Kuroda MI. Epigenetic spreading of the *Drosophila* dosage compensation complex from roX RNA genes into flanking chromatin. Cell 1999;98:513–522.
32. Pardue ML, Lowenhaupt K, Rich A, Nordheim A. $(dC\text{-}dA)_n \cdot (dG\text{-}dT)_n$ sequences have evolutionarily conserved chromosomal locations in

Drosophila with implications for roles in chromosome structure and function. EMBO J 1987;6:1781–1789.
33. Huijser P, Hennig W, Dijkhof R. Poly(dC-dA/dG-dT) repeats in the *Drosophila* genome: a key function for dosage compensation and position effects?. Chromosoma 1987;95:209–215.
34. Lowenhaupt K, Rich A, Pardue ML. Nonrandom distribution of long mono- and dinucleotide repeats in *Drosophila* chromosomes: correlations with dosage compensation, heterochromatin, and recombination. Mol Cell Biol 1989;9:1173–1182.
35. Waring GL, Pollack JC. Cloning and characterization of a dispersed, multicopy, X chromosome sequence in *Drosophila melanogaster*. Proc Natl Acad Sci USA 84:2843–2847.
36. DiBartolomeis S, Tartof KD, Jackson FR. A superfamily of *Drosophila* satellite related (SR) DNA repeats restricted to the X chromosome euchromatin. Nucl Acids Res 1992;20:1113–1116.
37. Roseman RR, Swan JM, Geyer PK. A *Drosophila* insulator protein facilitates dosage compensation of the X chromosome mini-*white* gene located at autosomal insertion sites. Development 1995;121:3573–3582.
38. Kelley RL, Solovyeva I, Lyman M, Richman R, Solovyev V, Kuroda MI. Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in *Drosophila*. Cell 1995;81:867–877.
39. Meller VH. Dosage compensation: making 1X equal 2X. Trends Cell Biol 2000;10:54–59.
40. Meyer BJ. Sex in the worm. Counting and compensating X-chromosome dose. Trends Genet 2000;16:247–253.
41. Graves JAM. Mammals that break the rules: genetics of marsupials and monotremes. Annu Rev Genet 1996;30:233–260.
42. Brockdorff N. The role of *Xist* in X-inactivation. Curr Opin Genet Dev 1998;8:328–333.
43. Panning B, Jaenisch R. RNA and the epigenetic regulation of X chromosome inactivation. Cell 1998;93:305–308.
44. Lee JT, Lu N. Targeted mutagenesis of *Tsix* leads to nonrandom X inactivation. Cell 1999;99:47–57.
45. Lee JT, Davidow LS, Warshawsky D. *Tsix*, a gene antisense to *Xist* at the X-inactivation center. Nat Genet 1999;21:400–404.
46. Brown CJ, Willard HF. The human X-inactivation center is not required for maintenance of X-chromosome inactivation. Nature 1994;368: 154–156.
47. Heard E, Clerc P, Avner P. X-chromosome inactivation in mammals. Annu Rev Genet 1997;31:571–610.
48. Gilbert SL, Sharp PA. Promoter-specific hypoacetylation of X-inactivated genes. Proc Natl Acad Sci USA 1999;96:13825–13830.
49. Carrel L, Cottle AA, Goglin KC, Willard HF. A first-generation X-inactivation profile of the human X chromosome. Proc Natl Acad Sci USA 1999;96:14440–14444.
50. Bailey JA, Carrel L, Chakravarti A, Eichler EE. Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proc Natl Acad Sci USA 2000;97: 6634–6639.
51. Marín I, Franke A, Bashaw GJ, Baker BS. The dosage compensation system of *Drosophila* is co-opted by newly evolved X chromosomes. Nature 1996;383:160–163.
52. Tatarenkov A, Kwiatowski J, Skarecky D, Barrio E, Ayala FJ. On the evolution of *Dopa decarboxylase* (*Ddc*) and *Drosophila* systematics. J Mol Evol.1999;48:445-462
53. Kuroda MI, Kernan MJ, Kreber R, Ganetzky B, Baker BS. The maleless protein associates with the X chromosome to regulate dosage compensation in *Drosophila*. Cell 1991;66:935–947.
54. Zhang S, Grosse F. Nuclear DNA helicase II unwinds both DNA and RNA. Biochemistry 1994;33:3906–3912.
55. Lee CG, Chang KA, Kuroda MI, Hurwitz J. The NTPase/helicase activities of *Drosophila maleless*, an essential factor in dosage compensation. EMBO J 1997;16:2671–2681.
56. Marín I, Baker BS. Origin and evolution of the regulatory gene *male-specific lethal-3*. Mol Biol Evol 2000;17:1240–1250.
57. Neal KC, Pannuti A, Smith ER, Lucchesi JC. A new human member of the MYST family of histone acetyl transferases with high sequence similarity to *Drosophila* MOF. Biochem Biophys Acta 2000;1490:170–174.
58. Chuang PT, Albertson DG, Meyer BJ. DPY-27: a chromosome condensation protein homolog that regulates *C. elegans* dosage compensation through association with the X chromosome. Cell 1994;79: 459–474.
59. Lieb JD, Albrecht MR, Chuang PT, Meyer BJ. MIX-1: an essential component of the *C. elegans* mitotic machinery executes X chromosome dosage compensation. Cell 1998;92:265–277.
60. Strunnikov AV. SMC proteins and chromosome structure. Trends Cell Biol 1998;8:454–459.
61. Jessberger R, Frei C, Gasser SM. Chromosome dynamics: the SMC protein family. Curr Opin Genet Dev 1998;8:254–259.
62. Hirano T. SMC-mediated chromosome mechanics: a conserved scheme from bacteria to vertebrates?. Genes Dev. 1999;13:11–19.
63. Melby TE, Ciampaglio CN, Briscoe G, Erickson HP. The symmetrical structure of structural maintenance of chromosomes (SMC) and MukB proteins: long, antiparallel coiled coils, folded at a flexible hinge. J Cell Biol 1998;142:1595–1604.
64. Bone JR, Kuroda MI. Dosage compensation regulatory proteins and the evolution of sex chromosomes in *Drosophila*. Genetics 1996;144: 705–713.
65. Steinemann M, Steinemann S, Turner BM. Evolution of dosage compensation. Chrom Res 1996;4:185–190.
66. Lahn BT, Page DC. Four evolutionary strata on the human X chromosome. Science 1999;286:964–967.
67. Adler DA, Rugarli EI, Lingenfelter PA, Tsuchiya K, Poslinsk D, Liggitt HD, Chapman VM, Elliott RW, Ballabio A, Disteche CM. Evidence of evolutionary up-regulation of the single active X chromosome in mammals based on Clc4 expression leveles in *Mus spretus* and *Mus musculus*. Proc Natl Acad Sci USA 1997;94:9244–9248.
68. Disteche CM. Escapees on the X chromosome. Proc Natl Acad Sci USA 1999;96:14180–14182.