Dosage Compensation Mechanisms: Evolution

Ignacio Marín, *Universidad de Valencia, Valencia, Spain*

Degeneration of sex-limited chromosomes generates the need for dosage compensation mechanisms able to equalize in both sexes the amount of products derived from sex-linked genes. Evolution of these mechanisms is understood in some detail in the fly *Drosophila melanogaster*, the nematode *Caenorhabditis elegans* and in eutherian mammals.

Advanced article

Article contents

- Dosage Compensation Mechanisms: An Evolutionary Perspective
- Human Dosage Compensation in an Evolutionary Framework
- Relationships among Dosage Compensation Systems: Common Origin or Convergence?

doi: 10.1038/npg.els.0006128

Dosage Compensation Mechanisms: An Evolutionary Perspective

In many species, sex chromosomes that originated as an identical, homologous pair currently show morphological differentiation in one of the sexes. An individual of the homogametic sex has two identical sex chromosomes, whereas an individual of the heterogametic sex has two distinct chromosomes. The chromosome restricted to the heterogametic sex often has a low number of genes. Loss of gene function is caused by the combined effect of chromosome restriction to one sex (often due to the emergence of a dominant sex-determining gene) and inhibition of recombination with its homolog. Under these circumstances, mutations cannot be eliminated by recombination. They therefore accumulate, leading to the progressive loss of genetic function in a process known as chromosome degeneration (reviewed in Charlesworth, 1996). (***See*** Sex Chromosomes.)

When chromosome degeneration starts, a mechanism known as dosage compensation evolves to equalize the levels of sex chromosome-derived products in both sexes. The discoverer of dosage compensation, H. J. Muller, reasoned more than 50 years ago that once the genes of a chromosome start accumulating mutations, a progressive need arises to compensate for the loss of their products. In particular, lowering by half the products of genes required in similar levels in both sexes must lead to dysfunctions in the heterogametic sex. Thus, a dosage compensation mechanism is strongly favored by natural selection. (***See*** Muller, Herman Joseph.)

There are two ways of generating compensatory mechanisms. As the need for compensation arises in parallel to mutation accumulation, and thus in principle gene-by-gene, independent solutions for each gene could emerge. The alternative is the emergence of a general mechanism able to regulate a twofold difference in the expression of many unrelated genes. This option appears unlikely. Genes in a degenerating chromosome are functionally diverse. They are often expressed at very different levels, in different times and places. However, all available data suggest that general dosage compensation mechanisms are the rule, not the exception. We have molecular information for three dosage compensation systems, those of the fly *Drosophila melanogaster*, the nematode *Caenorhabditis elegans* and eutherian mammalian species. In these three cases, the key feature is the existence of a mechanism able to modify transcriptional levels by modulating chromatin structure along whole chromosomes.

A consideration of the mechanisms found in protostome species is relevant here. In *D. melanogaster*, females are the homogametic sex (XX), whereas males are heterogametic (XY). Dosage compensation occurs via increased expression of X-linked genes in males. It is achieved by the action of a male-specific ribonucleoproteic complex, called a compensasome, that is formed by at least five proteins and two noncoding RNAs. Compensasomes bind in hundreds of positions along the male X chromosome. This binding correlates with hypertranscription of X-linked genes. A known biochemical function of the compensasome is its contribution to the enrichment along the male X chromosome of an acetylated isoform of histone H4 (acetylated at lysine 16). One of the proteins of the compensasome is a histone acetyltransferase. It is still unknown how compensasomes detect their binding targets on the X chromosome, although there is evidence of 'entry sites', sequences that recruit compensasomes and from which the complexes would spread in *cis*.

In *C. elegans* (hermaphrodites XX, males X0), gene expression in both hermaphrodite X chromosomes is decreased by half. This is the result of the binding of a complex of proteins unrelated to those found in the *Drosophila* compensasome. Which sequences are detected on the X chromosomes for the complex to bind is not yet understood. As in *Drosophila*, chromatin modification is thought to be critical. Two of the

proteins of the *C. elegans* dosage compensation complex belong to a family of proteins involved in chromatin condensation in other processes, including mitosis. However, how condensation is biochemically achieved is still unknown.

Comparison of these two systems reveals some features that are relevant for understanding the origin and evolution of the mammalian system: (1) compensation is established by global modification of the degree of compaction of sex chromosomes; (2) it requires the binding of *trans*-acting factors to sequences on the X chromosomes and (3) the *trans*-acting factors are different in flies and nematodes, suggesting that these mechanisms evolved independently.

Human Dosage Compensation in an Evolutionary Framework

In most eutherian mammals, including humans, females are the homogametic sex (XX) and males are heterogametic (XY). The evolutionary history of human sex chromosomes is complex. They originated as a pair of homologous autosomes that can be called proto-X and proto-Y. At some moment after the split of the lineages (about 320 million years ago), the proto-Y chromosome became restricted to the male sex, probably by the acquisition of a dominant male-determining function. There is evidence that, 80–130 million years ago, regions of autosomal origin were added to both the proto-X and proto-Y chromosomes. Several chromosomal inversions also occurred, leading to different gene arrangements in the proto-X and proto-Y, and progressively inhibiting recombination between them. The combination of sex-specific restriction and lack of recombination led to chromosome degeneration of the proto-Y. Mutations that arose on the proto-Y could not be eliminated by recombination in the heterogametic sex. The proto-Y chromosome progressively lost most of its genes and became the Y chromosome of today (Lahn and Page, 1999). (***See*** Chromosome Rearrangement Patterns in Mammalian Evolution; Chromosome X; Chromosome Y; Mammalian Sex Chromosome Evolution.)

Dosage compensation in embryonic and adult tissues of eutherians is achieved by random inactivation of one of the female X chromosomes. As happened in the protostome species discussed above, this occurs by changes in chromosome condensation. In eutherians, the change is radical: one of the female X chromosomes becomes highly compacted, or heterochromatic. It is still not fully understood how inactivation is achieved, but some of the main factors in the process have been characterized. It is known that the X-linked gene *Xist* is essential in *cis* for inactivation. A large noncoding RNA transcribed from the *Xist* locus accumulates and 'coats' the future inactive chromosome. The gene *Tsix*, involved in regulation of *Xist* function, is transcribed into another noncoding RNA from the strand opposite *Xist*. As *Tsix* overlaps with *Xist*, its RNA is antisense of the *Xist* RNA (hence the peculiar name '*Tsix*'). The most current model suggests that *Tsix* expression antagonizes *Xist* expression, protecting one of the chromosomes from inactivation. Recently, the first trans-acting regulator has been found. The transcription factor CTCF binds *Tsix*. It has been suggested that CTCF binding occurs on only one of the two X chromosomes, owing to methylation differences between *Tsix* alleles. Once CTCF binds and activates *Tsix* on an X chromosome, *Xist* function is inhibited, allowing that chromosome to remain active (Chao *et al*., 2002). (***See*** Noncoding RNAs: A Regulatory Role?; Nonprotein-coding Genes; X-chromosome Inactivation; X-chromosome Inactivation and Disease.)

However, several lines of evidence suggest that models including only *Xist* and *Tsix* are insufficient. For instance, the choice between X chromosomes for inactivation is known to be influenced by the poorly understood X-controlling element (*Xce*), which is located nearby but does not overlap with *Xist* or *Tsix*.

How *Xist* RNA spreads along the X chromosome is also not fully understood. In particular, what kind of *cis*-acting sequences contribute to this spreading has not been determined completely. An involvement of LINE-1 repetitive elements has been suggested. These elements are enriched in the X chromosome. Moreover, some genes are able to escape inactivation and, in humans, are concentrated in the regions added most recently to the sex chromosomes. It has been suggested that low concentrations of LINE-1 might define which regions escape inactivation (Bailey *et al*., 2000). (***See*** Long Interspersed Nuclear Elements (LINEs).)

Once *Xist* RNA coats an X chromosome, it is assumed that it recruits proteins, contributing to the dramatic change in chromatin structure that follows. Soon after *Xist* spreading, histones H3 and H4 become hypoacetylated, the unusual histone macroH2A1 accumulates, and gene promoters become hypermethylated. Methylation of histone H3 in its lysine 9 residue, characteristic of heterochromatin in many eucaryotic organisms, also occurs soon after *Xist* coating, preceding transcriptional inactivation of X-linked genes. This may be the first modification to appear on the future inactive X chromosome (Heard *et al*., 2001). It is likely that all of these modifications are directly or indirectly targeted to the X chromosome by the presence of the *Xist* transcript. (***See*** Histone Acetylation: Long-range Patterns in the Genome; Methylation-mediated Transcriptional Silencing in Tumorigenesis.)

Current data support a relatively ancient origin of the inactivation-based mechanism. *Xist* has been characterized in humans and rodents, having diverged about 80 million years ago, and is known to exist in other eutherians. Moreover, female X-chromosome inactivation occurs also in marsupial mammals. In marsupials, inactivation is not random but, as also happens in extraembryonic tissues in at least some eutherians (e.g. mouse), it is always the paternal X chromosome that is inactivated. In addition, X inactivation is incomplete in marsupials; different sets of paternal genes are inactivated in different tissues. There is so far no evidence of gene orthologs of *Xist* or *Tsix* in these organisms. However, it is known that *Xist* evolves rapidly (Nesterova *et al*., 2001). Thus, the presence of an *Xist*-related gene in marsupials cannot be dismissed at present. It is therefore possible that marsupial and eutherian dosage compensation systems, although somewhat different, have the same origin. It is still unclear whether monotreme mammals have X-chromosome inactivation. The other organisms for which some information exists are birds, where females are the heterogametic sex (ZW) and males are homogametic (ZZ). It is thought that these chromosomes also arose from a homomorphic pair of autosomes, although different from the one that became the XY pair in mammals. Birds had previously been thought to lack dosage compensation. However, recent data suggest that dosage compensation, at least for some Z-linked genes, exists (McQueen *et al*., 2001). In any case, there is no evidence of Z-chromosome inactivation in birds. These data suggest that the X-inactivation mechanism originated at least 130, but less than 320, million years ago. (***See*** Homologous, Orthologous and Paralogous Genes.)

Relationships among Dosage Compensation Systems: Common Origin or Convergence?

The previous sections highlight the differences among the dosage compensation systems in humans and those found in other species. However, the speculative idea that all dosage compensation systems found today in protostomes and deuterostomes may be variations of an ancestral mechanism has some indirect experimental support. First, functional similarities, in particular when *Drosophila* and mammals are compared, clearly exist. Both *Drosophila* and mammalian dosage compensation systems require noncoding RNAs that spread along the sex chromosomes and also involve covalent modifications of histones. Second, data exist showing that, once a dosage compensation mechanism has been established, it is co-opted when subsequent degenerative processes occur. One example is the previously mentioned addition of autosomal material to the proto-X and proto-Y chromosomes in mammals. A second example is the translocations between sex chromosomes and autosomes found in a few drosophilid fly species (Marín *et al*., 1996).

In these two situations, degenerative processes started in new chromosomal regions. In both cases, it has been found that the range of action of the preexisting dosage compensation systems was expanded to include those chromosomal regions. These results suggest that reusing a dosage compensation system is simpler than finding a new way to dosage-compensate.

However, these data are far from conclusive. First, mechanistic similarities may be the result of convergence. Multiple processes require chromatin modulatory changes that involve histone modifications, and it is likely that any global system of gene expression regulation requires those modifications. In addition, studies are increasingly detecting the involvement of noncoding RNAs in the regulation of gene expression (Blencowe, 2002; Sleutels *et al*., 2002). (***See*** Evolution: Convergent and Parallel; Noncoding RNAs: A Regulatory Role?)

It is unclear whether finding RNAs that are involved in chromatin regulatory processes is as exceptional as it seems today. The lack of similarity of the *C. elegans* system to the other two mechanisms suggests that fully innovative systems may emerge. Perhaps the main argument in favor of independent origins of dosage compensation mechanisms is historical. It is unlikely that a species without heteromorphic chromosomes has dosage compensation. Thus, when two lineages derive from such a species, there is no ancestral mechanism they must share. If dosage compensation arises, it is an independent occurrence. An example of this may be the differences in the avian and mammalian lineages. Following this argument, it is notable that all cases in which dosage compensation mechanisms have been co-opted for new regions involve additions to preexisting heteromorphic sex chromosomes, but not degenerative processes started from scratch on a homomorphic pair of chromosomes. That X-chromosome inactivation emerged in mammalian species is still the most parsimonious hypothesis.

See also

Chromosome X: General Features
Sex Chromosomes
X-chromosome Inactivation and Disease

References

Bailey JA, Carrel L, Chakravarti A and Eichler EE (2000) Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis.

Proceedings of the National Academy of Sciences of the United States of America **97**: 6634–6639.

Blencowe BJ (2002) Transcription: surprising role for an elusive small nuclear RNA. *Current Biology* **12**: R147–R149.

Chao W, Huynh KD, Spencer RJ, Davidow LS and Lee JT (2002) CTCF, a candidate trans-acting factor for X-inactivation choice. *Science* **295**: 345–347.

Charlesworth B (1996) The evolution of chromosomal sex determination and dosage compensation. *Current Biology* **6**: 149–162.

Heard E, Rougeulle C, Arnaud D, *et al*. (2001) Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell* **107**: 727–738.

Lahn BT and Page DC (1999) Four evolutionary strata on the human X chromosome. *Science* **286**: 964–967.

Marín I, Franke A, Bashaw GJ and Baker BS (1996) The dosage compensation system of *Drosophila* is co-opted by newly evolved X chromosomes. *Nature* **383**: 160–163.

McQueen HA, McBride D, Miele G, Bird AP and Clinton M (2001) Dosage compensation in birds. *Current Biology* **11**: 253–257.

Nesterova TB, Slobodyanyuk SY, Elisaphenko EA, *et al*. (2001) Characterization of the genomic *Xist* locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. *Genome Research* **11**: 833–849.

Sleutels F, Zwart R and Barlow DP (2002) The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415**: 810–813.

Further Reading

Avner P and Heard E (2001) X-chromosome inactivation: counting, choice and initiation. *Nature Reviews Genetics* **2**: 59–67.

Cohen DE and Lee JT (2002) X-chromosome inactivation and the search for chromosome-wide silencers. *Current Opinion in Genetics and Development* **12**: 219–224.

Eddy SR (2001) Non-coding RNA genes and the modern RNA world. *Nature Reviews Genetics* **2**: 919–929.

Ellegren H (2002) Dosage compensation: do birds do it as well? *Trends in Genetics* **18**: 25–28.

Graves JAM and Shetty S (2001) Sex from W to Z: evolution of vertebrate sex chromosomes and sex determining genes. *Journal of Experimental Zoology* **290**: 449–462.

Marín I, Siegal ML and Baker BS (2001) The evolution of dosage-compensation mechanisms. *Bioessays* **22**: 1106–1114.

Meller VH (2000) Dosage compensation: making 1X equal 2X. *Trends in Cell Biology* **10**: 54–59.

Park Y and Kuroda MI (2001) Epigenetic aspects of X-chromosome dosage compensation. *Science* **293**: 1083–1085.

Web Links

X (inactive)-specific transcript (*XIST*); LocusID: 7503. LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink/LocRpt.cgi?l = 7503

X (inactive)-specific transcript, antisense (*TSIX*); LocusID: 9383. LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink/LocRpt.cgi?l = 9383

X (inactive)-specific transcript (*XIST*); MIM number: 314670. OMIM: http://www.ncbi.nlm.nih.gov/htbin-post/Omim/dispmim?314670

X (inactive)-specific transcript, antisense (*TSIX*); MIM number: 300181. OMIM: http://www.ncbi.nlm.nih.gov/htbin-post/Omim/dispmim?300181