Commentary

LINE-1 elements and X chromosome inactivation: A function for "junk" DNA?

Mary F. Lyon*

Medical Research Council, Mammalian Genetics Unit, Harwell, Didcot, Oxfordshire OX11 0RD, United Kingdom

X chromosome inactivation (XCI) is the mechanism used in mammals for dosage compensation of X-linked genes between chromosomally XX females and XY males. One of the two X chromosomes of females becomes transcriptionally inactive in every cell of the early embryo and remains so in all somatic cells throughout life. The phenomenon is highly unusual among mechanisms of gene regulation in that the whole, or almost the whole, of a chromosome is silenced. Most gene-silencing mechanisms, such as the position effect variegation system of *Drosophila*, operate over much shorter distances (1–2 megabases in that case). XCI is initiated from an X inactivation center on the X chromosome, and from there, it spreads throughout the chromosome. Segments of X chromosome without a center, through deletion or translocation, remain active. Conversely, if autosomal segments are attached to the X chromosome by translocation, the inactivating signal can spread for long distances into the autosome, up to 100 megabases in a recently described case (1). However, spread into autosomal material extends less far and is less effective in gene silencing than in the X chromosome itself (1–3). Recently, there have been great advances in knowledge of the X inactivation center. The *Xist* gene (*Xist* in mouse and *XIST* in human), located at the center, is expressed from the inactive X chromosome (Xi) and is necessary in *cis* for X inactivation. It encodes a nontranslated RNA. Before the onset of XCI in the embryo, *Xist* RNA is transcribed but is unstable. At the time of XCI, it becomes stable and accumulates. It then remains close to the Xi and spreads to coat the whole chromosome (reviewed in refs. 4 and 5). When *Xist* is introduced as a transgene into an autosome, its RNA will also coat the autosome, although less efficiently than on the X chromosome (6). How the RNA spreads and how it brings about transcriptional silencing are unknown, but the spreading of RNA and of silencing into attached autosomes provides some clue. Clearly, any DNA sequences involved in spreading are not specific to the X chromosome. However, because the spread and inactivation are less efficient in autosomes, there must be something about X chromatin that promotes them. Riggs (7) suggested the concept of "way stations" or "boosters" along the X chromosome that act as promoters of spreading. Recently, I suggested that interspersed elements are candidates for the boosters (8). In particular, I proposed that LINE-1 elements (L1s) are those concerned, based on evidence then available that the X chromosomes of human and mouse seemed to be rich in L1s. This idea has now received support from work of Bailey *et al.* (9) published in this issue. These authors made use of sequence data accruing from the human genome project to compare the relative content of various repetitive elements in the human X chromosome and autosomes. The authors found that the X chromosome was indeed rich in L1 elements—there being about 26% of L1s on the X chromosome and only 13% in autosomal DNA.

In earlier work of others, fluorescence *in situ* hybridization with antibodies to L1s had been used to map the distribution of these elements in the mouse and human genomes. Boyle *et al.* (10) found that, in the mouse genome as a whole, L1s were located preferentially in dark G bands, whereas the whole of the X chromosome, including light and dark G bands, stained brightly (i.e., it was rich in L1s). Korenberg and Rykowski (11) showed the human X chromosome to stain brightly and remarked that the paracentric R band of the X chromosome was the brightest region of the genome. The work of Bailey *et al.* (9) thus confirms and extends the earlier studies but is more precise and more detailed. It lends important support to the concept that L1s are candidates for booster elements. L1s clearly satisfy the necessary conditions for boosters—that they must be present throughout the genome but must be increased in frequency on the X chromosome.

Bailey *et al.* (9) were able to compare the distribution of L1s in specific segments of the X chromosome. The density was greatest in the region Xq13–Xq21. This region contains the X inactivation center, which maps in Xq13. Previous work with fluorescence *in situ* hybridization suggested that, in the mouse, the region also richest in L1s was near or included the X inactivation center (10). This suggestion is consistent with the idea that the booster elements aid in establishing the onset of spread of *Xist* RNA at the time when XCI first occurs in the embryo and the accumulation of stable *Xist* RNA begins.

L1s can be divided into subgroups inserted into the genome at various evolutionary time points. Bailey *et al.* (9) found that the subgroups for which the X chromosome was particularly enriched were the younger elements, inserted <100 million years ago, around the time of the radiation of eutherian groups and continuing later. This finding is significant in relation to ideas on the evolution of XCI, which is thought to have arisen before the split of the mammalian lineage into the Metatheria and Eutheria. Both these groups have XCI, but the details differ. In eutherians, XCI is more complete and more stable than in Metatheria (12). Thus, L1s may have a role in establishing the completeness, the stability, or both in eutherians. Clearly, it would be valuable if similar studies could be carried out in Metatheria, but opportunities for such studies may well not arise.

A further point studied by Bailey *et al.* (9) concerned genes that escape inactivation. Around 10% of genes on the human X chromosome escape inactivation, being expressed from both the active X chromosome and the Xi, and these are preferentially located in certain regions (13). Bailey *et al.* (9) compared the L1 content of segments including genes that escape XCI with the L1 content of segments including typically inactivated genes. There were significantly fewer L1s in segments, particularly Xp22, containing genes that escape XCI. Once again, this result is consistent with L1s having a role in establishing or maintaining XCI and with present ideas on the evolution of

See companion article on page 6634.

*E-mail: m.lyon@har.mrc.ac.uk.

mammalian sex chromosomes. According to Ohno's Law, all mammalian X chromosomes carry the same genes. However, there is a difference between eutherians and metatherians. The eutherian X chromosome is larger and carries genes thought to have been added from autosomes to both X and Y chromosomes, followed by differentiation into nonhomologous X and Y chromosome genes, and by acquisition of XCI by the recently added X chromosome genes (14, 15). The recently added region on the human X chromosome lies in distal Xp (14, 15). Thus, this region became subject to XCI more recently in evolutionary time and could be still in the process of achieving complete XCI, which in turn could involve acquiring L1 elements. It would be very valuable to have similar data on the distribution of L1s on the mouse X chromosome, which carries genes corresponding to those on the human X chromosome, but these are much rearranged in order (16).

If L1s are indeed part of the mechanism of XCI, the question of how they fulfil this function is open and highly intriguing. At the onset of XCI in the embryo, the Xi takes on a set of properties typical of heterochromatin. It is condensed; it replicates its DNA late in the S phase; its histones are hypoacetylated; and in eutherians, the cytosines in CpG islands are methylated. This latter property is not seen in metatherians and is thought to have a role in stabilizing XCI in eutherians. In addition, the Xi is enriched in an unusual histone, macroH2A1 (17). Asynchronous DNA replication occurs first in the process of XCI, followed by hypoacetylation of histones and then by differential methylation (18). The acquisition of these heterochromatic properties by the Xi is thought to be brought about by its coating with *Xist* RNA. Bailey *et al.* (9) suggest that the L1 elements serve as binding sites for the *Xist* RNA, perhaps acting through an RNA–protein complex. It is relevant that, in *Drosophila*, RNA–protein complexes are involved in spreading of dosage compensation (19), although the mechanism of compensation is different, involving enhancement of transcription of the male X chromosome. It is also interesting that Duthie *et al.* (20), studying the detailed distribution of *Xist* RNA on the mouse Xi, found it to be localized to light G bands, rather than being uniformly spread. The work of Bailey *et al.* (9) suggests that, on the human X chromosome, both light and dark G bands are enriched for L1s, and according to Boyle *et al.* (10), such an enrichment is true for the mouse also. Thus, the work of Duthie *et al.* suggests that there may be some other feature of chromatin, in addition to L1s, that determines the tightness or the density of *Xist* RNA binding.

XCI in eutherians is thought to be a highly complex process. For instance, a functional *Xist* gene is needed for initiation of XCI in the embryo, but loss of *Xist* later does not lead to reactivation (reviewed in refs. 4 and 5). Thus, other factors, particularly methylation of CpGs, must be involved in stabilizing the inactive state. Perhaps L1s also are involved in stabilization. Escape of genes from XCI could be due either to resistance to the original inactivating signal or to the genes undergoing inactivation and then being reactivated (21, 22). Conceivably, a paucity of L1s weakens the stabilizing process, and this weakened stability accounts for the apparent failure of XCI in L1-poor regions.

I speculated recently (8) that XCI was a form of repeat-induced gene silencing. In various organisms, reiterated transgenes are expressed poorly. As the copy number increases, the transcription of individual copies decreases markedly, and this decrease is known as repeat-induced gene silencing. Wolffe (23) suggested that this type of gene silencing is the result of a cell defense mechanism that senses the presence of invading genomic parasites, such as transposable elements, and silences them. I suggested that, in XCI, the cell has made use of this mechanism to provide dosage compensation by silencing one X chromosome, with the L1s acting as the repeats sensed. However, this suggestion remains speculative.

The work of Bailey *et al.* (9) considerably strengthens the case for considering L1s as candidates for booster sequences facilitating the spread and possibly the stabilization of XCI. However, as they point out, one cannot entirely discount the possibility that the insertion of L1s is a consequence of XCI, rather than part of its mechanism. There is now a need for experimental work in the mouse, where XCI has been very much studied, addressed at comparing the onset and stability of XCI in LINE-rich and LINE-poor regions. If a role for L1s in XCI can indeed be established, then, as Bailey *et al.* (9) mention, it will be a very interesting example of DNA regarded as "junk" having developed a function in the organism.

1. White, W. M., Willard, H. F., van Dyke, D. L. & Wolff, D. J. (1998) *Am. J. Hum. Genet.* **63,** 20–28.
2. Rastan, S. (1983) *J. Embryol. Exp. Morphol.* **78,** 1–22.
3. Russell, L. B. (1983) in *Cytogenetics of the Mammalian X-Chromosome: Part A, Basic Mechanisms of X-Chromosome Behavior*, ed. Sandberg A. A. (Liss, New York), pp. 205–250.
4. Brockdorff, N. (1998) *Curr. Opin. Genet. Dev.* **8,** 328–333.
5. Heard, E., Clerc, P. & Avner, P. (1997) *Annu. Rev. Genet.* **31,** 571–610.
6. Lee, J. T. & Jaenisch, R. (1997) *Nature (London)* **386,** 275–279.
7. Riggs, A. D. (1990) *Aust. J. Zool.* **37,** 419–441.
8. Lyon, M. F. (1998) *Cytogenet. Cell Genet.* **80,** 133–137.
9. Bailey, J. A., Carrel, L., Chakravarti, A. & Eichler, E. E. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6634–6639.
10. Boyle, A. L., Ballard, S. G. & Ward, D. C. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 7757–7761.
11. Korenberg, J. R. & Rykowski, M. C. (1988) *Cell* **53,** 391–400.
12. Cooper, D. W., Johnston, P. G., Watson, J. M. & Graves, J. A. M. (1993) *Dev. Biol.* **4,** 117–128.
13. Carrel, L., Cottle, A. A., Goglin, K. C. & Willard, H. F. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 14440–14444.
14. Graves, J. A. M., Disteche, C. M. & Toder, R. (1998) *Cytogenet. Cell Genet.* **80,** 94–103.
15. Lahn, B. T. & Page, D. C. (1999) *Science* **286,** 964–967.
16. Boyd, Y., Blair, H. J., Cunliffe, P., Masson, W. K. & Reed, V. (2000) *Genome Res.* **10,** 277–292.
17. Costanzi, C. & Pehrson, J. R. (1998) *Nature (London)* **393,** 599–601.
18. Keohane, A. M., Lavender, J. S., O'Neill, L. P. & Turner, B. M. (1998) *Dev. Genet.* **22,** 65–73.
19. Kelley, R. L., Meller, V. H., Gordadze, P. R., Roman, G., Davis, R. L. & Kuroda, M. I. (1999) *Cell* **98,** 513–522.
20. Duthie, S. M., Nesterova, T. B., Formstone, E. J., Keohane, A. M., Turner, B. M., Zakian, S. M. & Brockdorff, N. (1999) *Hum. Mol. Genet.* **8,** 195–204.
21. Disteche, C. M. (1995) *Trends Genet.* **11,** 17–22.
22. Lingenfelter, P. A., Adler, D. A., Poslinski, D., Thomas, S., Elliott, R. W., Chapman, V. M. & Disteche, C. M. (1998) *Nat. Genet.* **18,** 212–213.
23. Wolffe, A. P. (1997) *Curr. Biol.* **7,** R796–R798.