Commentary

Making sense (and antisense) of the X inactivation center

Huntington F. Willard* and Laura Carrel

Department of Genetics, Case Western Reserve University and Center for Human Genetics and Research Institute, University Hospitals of Cleveland, Cleveland, OH 44106

The concept of the X inactivation center (Xic) as the master regulatory locus for X inactivation dates back to the mid-1960s (1, 2), not long after X chromosome inactivation itself was proposed as the mammalian dosage compensation mechanism (3–5). Defined genetically as the cis-acting locus required for an X chromosome to undergo inactivation early in female embryogenesis, the Xic defied molecular characterization for nearly 30 years until the discovery of the *XIST* gene (6), whose noncoding RNA product is transcribed from and remains intimately associated with the inactive X chromosome in female somatic cells (6–9). *Xist* in both humans and mice maps to the Xic region on the X chromosome and thus became a compelling candidate for a component of the Xic itself (10–12). Biology is rarely as simple as it first appears, however, and X inactivation embodies this principle fully! Only a few years ago, Lee and colleagues (13) described another component of the Xic just downstream of *Xist,* the *Tsix* gene, so-named because it consists of an antisense transcript of *Xist* and whose pattern of expression suggested a potential role as a regulator of *Xist.* Now, in this issue of PNAS, Stavropoulos *et al.* (14) extend these studies to address a fundamental question concerning the potential function of *Tsix*—does its antagonistic relationship to *Xist* depend on its own noncoding RNA product or does it reflect actions at the DNA level, independent of transcription?

What evidence implicates *Xist* or *Tsix* as functional components of the Xic? The process of random X inactivation consists of multiple steps, responsible for assessing the number of X chromosomes in the developing mammalian embryo (counting), designating one or the other X as the future active or inactive X chromosome (choice), initiating and promulgating this choice along the length of the inactive X chromosome in cis to silence most of the genes on that X (initiation), and stably maintaining the inactive chromatin state through subsequent cell divisions (maintenance) (reviewed in refs. 15–17). [The choice step is predetermined in the imprinted form of X inactivation, in which the paternally inherited X is always the inactive X, as is seen in marsupials and in extraembryonic tissues in eutherian mammals (18).]

Transgenic experiments have demonstrated that DNA from the Xic (including *Xist* and *Tsix* sequences) can largely recapitulate X inactivation (e.g., ref. 19), and more refined targeted mutagenesis studies have implicated specific sequences in the various steps of the X inactivation pathway. Specific regions in the Xic contribute to counting and choice (e.g., refs. 20 and 21), in addition to the role of the Xist transcripts themselves in initiation of silencing on the inactive X. But what controls expression of *Xist* itself and determines whether it produces abundant RNA product (localized on the inactive X) or whether it is itself suppressed (on the active X)? This is where the antisense *Tsix* gene appears to come in.

Transcription of *Tsix* initiates principally from a promoter near a CpG island about 15 kb downstream from the 3′ terminus of *Xist* and extends across *Xist* in the antisense direction (13, 22). In differentiating female mouse embryonic stem cells, Tsix RNA (which is initially transcribed along with *Xist* from both X chromosomes) is turned off on one X that then accumulates Xist transcripts and becomes the inactive X (Fig. 1*A*). In contrast, Tsix RNA persists longer on the other X, on which *Xist* is then turned off, enabling it to become established as the active X. Because of this dynamic and complementary relationship, *Tsix* has been proposed to be the regulatory "switch" that determines whether *Xist* is on or off and thus whether the X becomes the inactive or active X (16, 17). This model predicts that targeted mutation of *Tsix* should influence *Xist* expression and thus the choice of whether one or the other X becomes inactive. A number of studies have tested this hypothesis directly and have implicated a 2-kb region at the 5′ end of *Tsix* (22–24). For example, when sequences near the *Tsix* CpG island are deleted (ΔCpG; Fig. 1*B*), the targeted X expresses *Xist* constitutively, leading to that X being designated the inactive X in all cells (23). Although these and similar experiments indicate the involvement of these sequences in the cis regulation of *Xist* expression and thus in the choice step of X inactivation, the precise mechanism of their involvement has remained obscure. The question remains whether they act through DNA as a long-range silencer to repress *Xist* or whether it is the Tsix transcripts themselves that act to antagonize *Xist* in cis. Tsix RNA could function in a number of ways, either through a mechanism that prevents *Xist* transcription or via direct interactions between the Xist and Tsix RNAs themselves (17).

Tsix has been proposed to be the regulatory "switch" that determines whether _Xist_ is on or off.

It is this question that Lee and her colleagues (14) have addressed. They have generated a constitutive "knock-in" allele of *Tsix,* by targeting an active human EF1a promoter into one *Tsix* allele in female mouse embryonic stem cells. Upon differentiation, the persistent high-level expression of Tsix transcripts from this allele is sufficient to block the usual accumulation of Xist transcripts on that X chromosome, thus preventing it from becoming an inactive X (Fig. 1*C*). [In cells in which the other (nontargeted) X is chosen and the normal *Tsix* allele is shut off, *Xist* expression continues unperturbed only on that X, and X inactivation proceeds normally.] This effect of persistent high-level *Tsix* expression seems likely to be caused by the

See companion article on page 10232.

*To whom reprint requests should be addressed at: Lakeside 1400, Research Institute, University Hospitals of Cleveland, 11100 Euclid Avenue, Cleveland, OH 44106. E-mail: Willard@UHRI.org.

COMMENTARY

Downloaded by guest on April 9, 2020

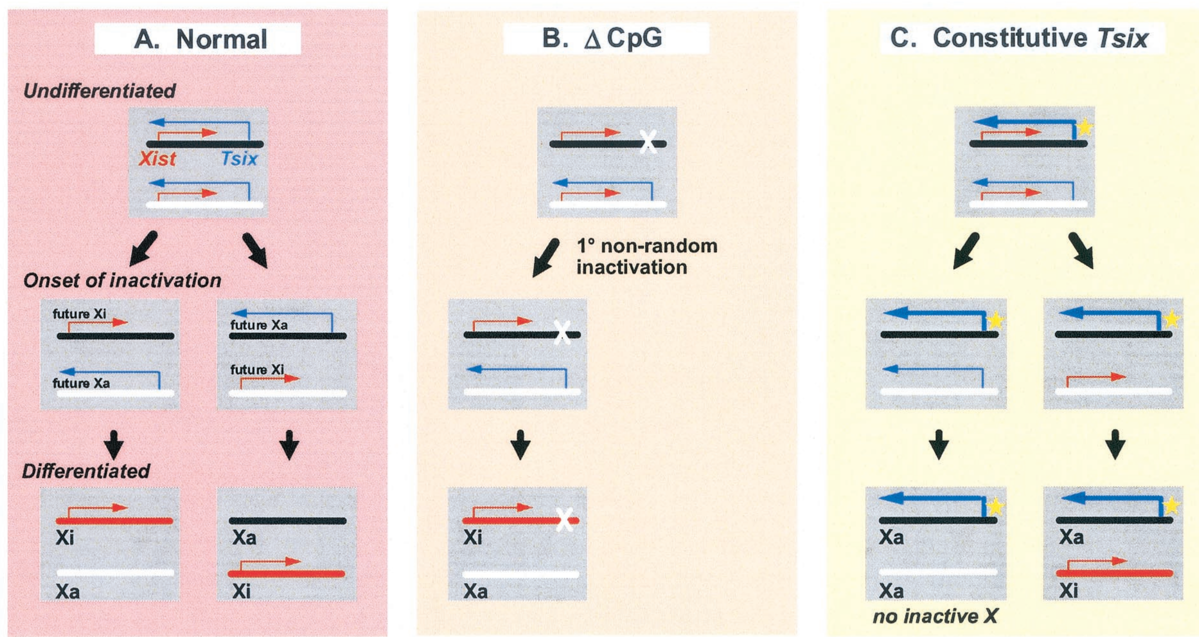

Fig. 1. Schematic representation of proposed events at the *Xist* and *Tsix* loci in the X inactivation center. (*A*) In normal female embryonic cells, both genes are expressed in the early embryo, before the onset of inactivation. *Tsix* is then suppressed on one X, which continues to express *Xist* (the future inactive X, Xi). On the other X, *Tsix* continues to be expressed, whereas *Xist* is suppressed (the future active X, Xa). As a result of the choice step of inactivation, there are two populations of cells (*Left* and *Right*), in which one or the other X is chosen to be the future Xa (or Xi). (*B*) In cells in which the promoter region of *Tsix* has been deleted (ΔCpG in ref. 23; see also refs. 21 and 22), X inactivation is nonrandom. The targeted X is the Xi in all cells, suggesting that *Tsix* expression and/or this region of the *Tsix* gene is responsible for suppressing *Xist* in cis. (*C*) In cells in which *Tsix* is constitutively expressed, as described in this issue (14), due to targeting at the *Tsix* promoter (yellow star), *Tsix* shows high-level expression (large blue arrow). This prevents expression of *Xist* from that X in both cell lineages. These data indicate that it is the Tsix RNA itself that is responsible for suppression of *Xist*, providing mechanistic insight into this critical early step in X inactivation.

transcripts themselves, because the targeting event doesn't remove any *Tsix* sequences, although indirect effects on chromatin caused by the knock-in are difficult to exclude completely.

The story doesn't end there, however. The earlier *Tsix* knockout mutations led to nonrandom inactivation of the mutated X chromosome, providing critical evidence that *Tsix* was required to silence the *Xist* allele on the same X (21, 23). Without *Tsix*, *Xist* expression continued unabated, thus ensuring that that X chromosome would be the inactive X in all cells (Fig. 1*B*). By comparison, then, one might have predicted that the constitutive *Tsix* allele reported (14) would result in complete nonrandom inactivation of the nonmutated X, secondary to the constitutive active state of the mutated X necessitated by persistent *Tsix* expression. However, this was not the case. In fact, choice appears not to have been affected by this targeted mutation. At the expected frequency, in some cells, the nonmutated X was chosen to be the inactive X, and X inactivation proceeded normally. But the remaining cells, in which the mutated X would be predicted to be chosen to be the inactive X, in fact show no X inactivation at all (Fig. 1*C*). In these cells, persistent high-level *Tsix* expression is apparently epistatic to X inactivation, and the steps of the X inactivation pathway downstream of choice are blocked. This result establishes that *Tsix* transcripts have a functional role in random X inactivation, suppressing *Xist*, and that shutting off of *Tsix* expression is a necessary step to permit inactivation of that X. These data support and complement data indicating a role for *Tsix* in imprinted X inactivation as well (22, 25).

So, does the Xic consist of just these two interacting genes, the sense *Xist* gene and its antisense antagonist *Tsix*? Again, the story is not apt to be that simple. At a minimum, there are additional sequences yet further downstream of *Xist* (i.e., upstream of *Tsix*) that appear to be involved in X inactivation. The X controlling element locus, *Xce,* of Cattanach has been defined genetically (26) and maps at least 40 kb away from the *Xist* 3′ end or the *Tsix* promoter (27). *Xce* influences the choice step of X inactivation, and mice heterozygous for variants at this locus show a nonrandom pattern of inactivation that reflects the relative strength of different *Xce* alleles (26). Although the molecular nature of *Xce* has not been determined, targeted deletion and add-back studies have implicated sequences in this vicinity in counting or choice, independent of *Tsix* transcription (22, 28). Whether these constitute the *Xce* locus has not been established. Further, whether they represent long-range regulatory sites for either *Tsix* or *Xist* and, if so, what their trans-acting binding partners might be remains unknown. Further, as X inactivation exists in both an imprinted and nonimprinted (i.e., random) form, it is conceivable that regulatory sites within the Xic differ between the two.

A long-standing model for X inactivation has posited a trans-acting, autosomally encoded "blocking factor" that interacts with sequences in the Xic to designate the future active X, thus allowing the other X (or all remaining Xs in the case of aneuploid cells) to be inactivated (29, 30). The paper reported in this issue (14), together with other recent studies (22, 23, 28), begin to delimit the location of possible binding sites within the Xic for such a blocking factor. The factor could be a protein involved directly in the regulation of *Tsix* or less directly in binding to the *Xce* sequences located upstream of *Tsix.*

Tsix "on?" No *Xist!* Whichever direction you read it, *Xist* and *Tsix* both appear to be critical components of the Xic and the X inactivation pathway. Although the

Downloaded by guest on April 9, 2020

cis interactions within the Xic and between the Xic and the X chromosome itself are becoming clearer, it remains to define the thus far elusive trans-acting factors that set this cascade in motion in the first place. Only then will we be able to claim to make sense (or antisense) of this complex developmental and chromosomal phenomenon.

1. Lyon, M. F. (1963) in *Second International Conference on Congenital Malforrmations,* (International Medical Congress, New York), pp. 67–68.
2. Russell, L. B. & Montgomery, C. S. (1965) *Genetics* **52,** 470–471.
3. Lyon, M. F. (1961) *Nature (London)* **190,** 372–373.
4. Russell, L. B. (1961) *Science* **133,** 1795–1803.
5. Beutler, E., Yeh, M. & Fairbanks, V. F. (1962) *Proc. Natl. Acad. Sci. USA* **48,** 9–16.
6. Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R. & Willard, H. F. (1991) *Nature (London)* **349,** 38–44.
7. Brockdorff, N., Ashworth, A., Kay, G. F., McCabe, V. M., Norris, D. P., Cooper, P. J., Swift, S. & Rastan, S. (1992) *Cell* **71,** 515–526.
8. Brown, C. J., Hendrich, B. D., Rupert, J. L., Lafreniere, R. G., Xing, Y., Lawrence, J. & Willard, H. F. (1992) *Cell* **71,** 527–542.
9. Clemson, C. M., McNeil, J. A., Willard, H. F. & Lawrence, J. B. (1996) *J. Cell Biol.* **132,** 1–17.
10. Brown, C. J., Lafreniere, R. G., Powers, V. E., Sebastio, G., Ballabio, A., Pettigrew, A. L., Ledbetter, D. H., Levy, E., Craig, I. W. & Willard, H. F. (1991) *Nature (London)* **349,** 82–84.
11. Borsani, G., Tonlorenzi, R., Simmler, M. C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawrence, C., *et al.* (1991) *Nature (London)* **351,** 325–329.
12. Brockdorff, N., Ashworth, A., Kay, G. F., Cooper, P., Smith, S., McCabe, V. M., Norris, D. P., Penny, G. D., Patel, D. & Rastan, S. (1991) *Nature (London)* **351,** 329–331.
13. Lee, J. T., Davidow, L. S. & Warshawsky, D. (1999) *Nat. Genet.* **21,** 400–404.
14. Stavropoulos, N., Lu, N. & Lee, J. T. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 10232–10237. (First Published July 31, 2001; 10.1073/pnas.171243598)
15. Heard, E., Clerc, P. & Avner, P. (1997) *Annu. Rev. Genet.* **31,** 571–610.
16. Avner, P. & Heard, E. (2001) *Nat. Rev. Genet.* **2,** 59–67.
17. Boumil, R. M. & Lee, J. T. (2001) *Hum. Mol. Genet*., in press.
18. Takagi, N. & Sasaki, M. (1975) *Nature (London)* **256,** 640–642.
19. Lee, J. T., Lu, N. & Han, Y. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 3836–3841.
20. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. (1996) *Nature (London)* **379,** 131–137.
21. Clerc, P. & Avner, P. (1998) *Nat. Genet.* **19,** 249–253.
22. Sado, T., Wang, Z., Sasaki, H. & Li, E. (2001) *Development (Cambridge, U.K.)* **128,** 1275–1286.
23. Lee, J. T. & Lu, N. (1999) *Cell* **99,** 47–57.
24. Debrand, E., Chureau, C., Arnaud, D., Avner, P. & Heard, E. (1999) *Mol. Cell. Biol.* **19,** 8513–8525.
25. Lee, J. T. (2000) *Cell* **103,** 17–27.
26. Cattanach, B. M. (1970) *Genet. Res.* **16,** 293–301.
27. Simmler, M.-C., Cattanach, B., Rasberry, C., Rougeulle, C. & Avner, P. (1993) *Mamm. Genome* **4,** 523–530.
28. Morey, C., Arnaud, D., Avner, P. & Clerc, P. (2001) *Hum. Mol. Genet* **10,** 1403–1411.
29. Lyon, M. F. (1971) *Nat. New Biol.* **232,** 229–232.
30. Brown, S. W. & Chandra, H. S. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 195–199.

Downloaded by guest on April 9, 2020