Functional intergenic transcription: a case study of the X-inactivation centre

Jeannie T. Lee

Howard Hughes Medical Institute, Department of Molecular Biology, Massachusetts General Hospital,
Department of Genetics, Harvard Medical School, Boston, MA 02114, USA (lee@frodo.mgh.harvard.edu)

Long known to be riddled with repetitive elements and regarded as 'junk', intergenic regions in the mammalian genome now appear to be more than incidental spacers between coding sequences. Here, I review the example of *Xite*, an intergenic region at the X-inactivation centre which was recently shown to regulate the X-chromosome choice decision. *Xite* contains a series of DNaseI-hypersensitive sites and harbours two intergenic transcription start sites. These intergenic transcription elements act at the onset of X-chromosome inactivation (XCI) to bias the selection of the active X. It has been proposed that *Xite* acts *in cis* on *Tsix* by promoting its persistence during XCI. *Xite* has also been proposed to be a candidate for the X-controlling element, a naturally occurring modifier of XCI ratios in mice and possibly also in humans. It seems likely that intergenic transcription will turn out to be a widespread phenomenon in mammals and that, more importantly, it will emerge as a significant regulatory mechanism for the expression of coding sequences.

Keywords: non-coding RNA; intergenic transcription; *Xite*; *Xist*; *Tsix*; X inactivation

1. NOVEL PLAYERS IN GENE REGULATION: TRANSCRIBED SEQUENCES WITHIN INTERGENIC SPACE

With the human and mouse genomes largely sequenced, computational biologists have predicted the existence of *ca.* 30 000 protein-coding genes in the mammalian genome, a number much smaller than originally thought (Mouse Genome Sequencing Consortium 2002; Lander *et al.* 2003). If each gene occupies only 30–50 kb of sequence on average, this prediction implies great stretches of unoccupied 'real estate' between genetic units. As much as two-thirds of the genome may be 'intergenic DNA'. In so far as repetitive elements like LINEs and SINEs are found in great numbers in intergenic regions, it has been suggested (and even well accepted) that intergenic space is largely 'junk' devoid of functional information.

However, several recent discoveries challenge this idea that non-coding is necessarily equated with junk. Microarray analyses of human chromosomes 21 and 22 demonstrate that there may be as much as 10 times more transcription coming from intergenic spaces as from coding sequences (Shoemaker *et al.* 2001; Kapranov *et al.* 2002; Rinn *et al.* 2003). Because quantitative information could not be gathered from the analysis, the transcript abundance at intergenic positions relative to known protein-coding regions could not be estimated. Nonetheless, the observation of large-scale transcriptional activity on the two human chromosomes in itself raises many interesting questions. Are they the products of spurious transcription initiation or possibly of read-through transcription

from coding sequences? Or are they functional transcription elements? If so, could they play a role in gene regulation or maintenance of chromosome structure?

In the past 3 years, there have been several descriptions of intergenic transcription elements with apparent function in metazoans. Curiously, these elements tend to lie in loci with complex patterns of gene expression. One example lies in the human β -globin locus, where genes for four developmentally specific β -globin isoforms (ϵ , γ , δ , β) are expressed in sequence according to their proximity to the LCR (Ashe *et al.* 1997; Gribnau *et al.* 2000; Plant *et al.* 2001; figure 1a). Progressive switching from foetal to adult haemoglobin genes is proposed to depend on sequential activation of contiguous intergenic transcripts. Early in embryogenesis, a low level transcript initiated by the LCR is believed to enable expression of the ϵ and γ genes. Later in embryogenesis, a switch to a second intergenic promoter lying downstream of the γ genes permits the developmentally timed progression to the δ and β (adult) isoforms. How intergenic transcription works remain unclear, but Gribnau *et al.* (2000) favour the idea that transcription opens up the chromatin locally to permit activation of downstream genes.

A second example can be found in the non-coding transcripts of the *Drosophila* BX-C (Bae *et al.* 2002; Bender & Fitzgerald 2002; Drewell *et al.* 2002b; Hogga & Karch 2002; Rank *et al.* 2002). The BX-C contains several homeotic genes, *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*), each genetically arranged in order of the insect body segment that the genes control (figure 1b). Their pattern of expression is determined by a series of *cis*-regulatory elements (*bx*, *bxd* and *iab-2* to *iab-9*) that together establish the specific parasegment expression pattern of the three homeotic genes along the anterior–posterior axis of the fly larvae. Both PREs and trithorax

Figure 1. Known examples of intergenic transcription at complex metazoan loci. (a) The human beta-globin locus. DHSs are indicated as downward arrows. Intergenic transcripts are shown as broken horizontal arrows. (b) The *Drosophila* BX-C. Intergenic transcripts occur all across the region and can be in either orientation. (c) The mouse *Xic*. The *Xite* intergenic transcripts initiate at two sites and may cross into the *Tsix* locus in some cases.

response elements are found in these elements and are thought to play a role in setting up the transcriptionally silenced as well as activated states. Although the association of non-coding RNAs with the *cis*-elements has been known for 15 years (Lipshitz *et al.* 1987), the roles of these intergenic transcripts were not clear until this past year. In 2002, several groups simultaneously found that intergenic transcription drives the switching of transcription states from off to on within the parasegment controlled by the associated *cis*-element. Both on transgene arrays and at the endogenous locus, a pulse of transcription across PREs was sufficient to derepress previously silent BX-C genes. Just as at the globin locus, transcriptional activity or the transcripts themselves are believed to provide a favourable chromatin conformation for the developmentally regulated activation of genes in the complex locus.

2. NON-CODING RNAs IN XCI

We now believe that a similar mechanism regulates the switching on and off of dosage compensation in mammals (Ogawa & Lee 2003; figure 1c). XCI achieves dosage compensation in mammals between males with one X chromosome (XY) and females with two (XX) (Lyon 1961). XCI is controlled *in cis* by a master regulatory locus called the *Xic* (Brown *et al.* 1991), a locus both necessary and sufficient to direct the initiation of silencing in a developmentally specific and numerically correct manner (Lee *et al.* 1996). In diploid cells, the random form of XCI takes place during peri-implantation development in the epiblast (embryo proper) and only when more than one X is present (Rastan & Robertson 1985). Because XCI is

not necessarily sex-specific, males with two X chromosomes (XXY) will also silence one X and females with only one (XO) will forego dosage compensation. In cells with two Xs or more, a choosing mechanism designates one active X and designates all remaining Xs as inactive (Avner & Heard 2001; Boumil & Lee 2001). Once chosen, the future inactive X undergoes a series of heterochromatic changes, beginning with the painting of that X by the non-coding RNA made by *Xist* (Brockdorff *et al.* 1992; Brown *et al.* 1992) at the *Xic* (Clemson *et al.* 1996), the recruitment of silencing factors such as Eed and Ezh2 proteins along the X (Mak *et al.* 2002), distinct histone modifications such as H3 K4 demethylation, H3 K9 methylation (Boggs *et al.* 2002; Peters *et al.* 2002), H4 deacetylation (Keohane *et al.* 1996), and DNA hypermethylation of CpG dinucleotides at 5' regions of X-linked genes (Li 2002). How X chromosomes are 'counted' and then chosen as Xa or Xi are two of the most fascinating questions in the field. Both properties have been ascribed to the *Xic*, which, when integrated into autosomes in an XY mouse ES cell line, will confer upon those autosomes the ability to be counted as additional X chromosomes and then be chosen for silencing (Lee *et al.* 1996; Heard *et al.* 1999; Migeon *et al.* 1999).

Several years ago, a systematic deletional analysis of the mouse *Xic* led us to identify an antisense partner to *Xist* RNA (Lee *et al.* 1999). Named '*Tsix*', the antisense gene regulates XCI in two ways. First, *Tsix* is a key determinant of which X will become the Xa and which Xi during the choice step (Lee & Lu 1999; Lee 2002). This function apparently depends on DNA sequences at the CpG-rich 5' end of the gene, which are currently believed to contain

Figure 2. Asymmetric *Tsix* expression marks the onset of XCI and designates choice. See text for details. Depicted are female ES cell nuclei prior to XCI (undifferentiated), during XCI, or after (differentiated). The three nuclei undergoing XCI are believed to occur in sequence, although they can be observed simultaneously in any population of differentiating female ES cells. Low-level *Xist* RNA is shown as small filled black circles, high-level *Xist* RNA as large filled circles, and *Tsix* RNA as white circles.

DNA motifs for binding *trans*-acting transcription factors. One such factor is CTCF (Chao *et al.* 2002), an 11-zinc finger DNA-binding protein with functions as diverse as transcriptional repression/activation (Lobanenkov *et al.* 1990; Ohlsson *et al.* 2001), chromatin insulation (Bell & Felsenfeld 2000; Hark *et al.* 2000; Kanduri *et al.* 2000) and boundary element formation (Bell *et al.* 1999). We have hypothesized that the binding of CTCF and other as yet unidentified factors to *Tsix* represents the first step of the choice decision (Chao *et al.* 2002). This binding then leads to expression of the antisense transcript. A role for *Tsix* in choice is supported by deletional analysis in which knocking out the CpG island led to a total inability of the mutated X chromosome to be chosen as the Xa (Lee & Lu 1999). In heterozygous female cells, the mutated X is always silenced.

Antisense transcription forms the second arm of *Tsix*'s regulatory mechanism (Luikenhuis *et al.* 2001; Stavropoulos *et al.* 2001). We now suspect that *Tsix*'s transcriptional activity plays a critical role apart from the occurrence of crucial DNA motifs at its 5' end. Antisense transcription is indeed sufficient to block the spread of *Xist* RNA and all subsequent steps of silencing. Prior to XCI, *Tsix* and *Xist* are expressed together on all X chromosomes and this concurrent antisense expression is thought to prevent the accumulation of *Xist* RNA (figure 2). As cells enter the XCI pathway, *Tsix* is turned off on the future Xi, an event that is believed to be a prerequisite for *Xist* RNA accumulation along the Xi. On the future Xa, *Tsix* RNA persists until *Xist* is turned off. This persistence of *Tsix* blocks *Xist* RNA spread and the recruitment of the silencing machinery. We have proposed that *Tsix* could work by titrating *Xist* RNA through RNA:RNA base-pairing, in which duplex formation destabilizes *Xist* RNA or masks *Xist* RNA domains involved in binding silencing proteins (Lee *et al.* 1999).

Together, these studies demonstrate that one of the earliest events to take place during XCI is the generation of asymmetric *Tsix* expression (figure 2). Indeed, the choice decision depends on stochastic allelic inequality at *Tsix*, with loss of antisense expression on the Xi and persistence on the Xa. Therefore, one might imagine that, during the choice step, the X chromosome that keeps *Tsix* on longer than its homologue will have a competitive

advantage in staying active. But how is the *Tsix* allelic asymmetry generated? We recently proposed that non-coding transcription within the intergenic space at the *Xic* may provide part of the regulation (Ogawa & Lee 2003).

3. *Xite*: X-INACTIVATION INTERGENIC TRANSCRIPTION ELEMENTS THAT REGULATE THE PROBABILITY OF X-CHROMOSOME CHOICE

The presence at the *Xic* of additional required elements for the choice decision has been suggested by earlier work that deleted a 65 kb region downstream of *Xist* and showed complete skewing of XCI to favour the mutated X chromosome (Clerc & Avner 1998). Although part of the phenotype can now be attributed to the loss of *Tsix* in this deletion, adding back the 15 kb of *Tsix* sequence deleted by the 65 kb knockout did not restore random choice (Morey *et al.* 2001). This result suggested that there must be other regulatory elements within the 50 kb of remaining sequence. This region is known to contain *Tsx*, a testis-specific gene that is poorly conserved between mouse, humans and cows, and is not likely to play a role in XCI (Simmler *et al.* 1996; Chureau *et al.* 2002). The region also contains a small CpG island whose knockout does not apparently disrupt any aspect of XCI, despite the fact that the knockout deletes a minor alternate upstream promoter for the *Tsix* transcript (Simmler *et al.* 1996; Sado *et al.* 2001).

The fact that non-coding RNAs dominate the *Xic* prompted us to search for additional non-coding transcripts at the *Xic*. Indeed, we found that there is extensive non-coding transcription in the 25 kb region between *Tsx* and *Tsix* (Ogawa & Lee 2003). The transcripts originate at two start sites located *ca.* 6 kb apart. The first start site coincides with the small CpG island previously deleted without a phenotype (Sado *et al.* 2001). The second is located just 10 kb upstream of the *Tsix* start site. Quantitative analysis by both conventional and real-time RT-PCR showed that the steady-state level of the intergenic transcripts is one to two orders of magnitude lower than that of *Tsix*. The two transcription start sites are associated with specialized chromatin, as indicated by six DHSs, with

two weak ones at the first start site and four very strong sites at the second.

Named *Xite* for 'X-inactivation intergenic transcription elements', these elements span more than 12.5 kb of sequence upstream of *Tsix*. Although hypersensitive sites and associated transcripts are broadly distributed, the critical elements seem to be located around the second of the two transcription starts. Indeed, deletions of the first start site in our laboratory (Ogawa & Lee 2003) and by others (Sado *et al.* 2001) do not display measurable phenotypes. By contrast, deletions that include the second start site alter the probability with which the linked X chromosome will be chosen as the Xa and exhibit properties reminiscent of that of a classically defined X-linked modifier known as the *Xce* (Cattanach & Isaacson 1967; Avner *et al.* 1998).

In wild-type female cells of inbred strains, two X-chromosomes normally have an equal chance of inactivation (50 : 50). Random choice, however, is known to be affected by the *Xce*, which displays strain-specific haplotypes of varying strengths (Cattanach & Isaacson 1967; Avner *et al.* 1998). In F₁ hybrids produced by mating wild strains (e.g. *Mus musculus castaneus*) to domestic strains (e.g. 129, C57BL/6), there can be dramatic skewing of XCI allele ratios to favour inactivation of the X of domestic origin. This skewing can be as significant as 90 : 10 or 80 : 20 but is manifested only when two Xs of different *Xce* haplotypes are present in the same nucleus. Homozygosity at the *Xce* results in completely random choice (50 : 50). A human *XCE* equivalent has been proposed based on an occasional familial clustering of very skewed XCI ratios in otherwise normal women (Naumova *et al.* 1996, 1998; Plenge *et al.* 1997). In the clinical setting, such skewing has been linked to penetrance of X-linked disease in heterozygous women (Verkerk *et al.* 1991; Azofeifa *et al.* 1995)—in principle, if the disease locus occurs on an X chromosome with a 'strong' *XCE* (i.e. on an X that is more likely to stay active), the disease phenotype would be more likely to be manifested.

Xite has been proposed to be the *Xce* based on a number of genetic and molecular observations (Ogawa & Lee 2003). *Xite*'s location some 20–32 kb downstream of *Xist* is roughly consistent with prior genetic studies that place the mouse *Xce* downstream of *Xist*, possibly more than 12–27 kb distal. Most importantly, when 12.5 kb of *Xite* sequence is deleted on one X chromosome in heterozygous females, there is a resulting bias towards inactivating the mutated X chromosome. Importantly, however, *Xite* knockouts increase the likelihood of XCI *in cis* without precluding inactivation of the X *in trans*. Such is the property of the *Xce* modifier, which merely alters the inactivation ratio between two X chromosomes and does not prevent inactivation of either. This contrasts with all *Tsix* deletions made to date, which leads to exclusive inactivation of the *Tsix*-deficient X chromosome. Thus, *Xite* does not determine choice but influences the likelihood with which an X will be inactivated, a property reminiscent of that of the *Xce*. The *Xite* deletion may be regarded as a very weak *Xce* haplotype.

If *Xite* were the *Xce*, one would predict the existence of strain-specific differences at this locus. Indeed, several strain-specific differences have been found. DNA sequencing of the wild (*M. castaneus*) and domestic (129) strains

revealed numerous *Xite* polymorphisms including nucleotide substitutions, insertions and deletions, some of which may underlie *Xce* differences. There is also a correlation between *Xce* strength and *Xite* expression levels: the X chromosome with higher *Xite* transcript levels is more likely to remain active. For example, the *M. castaneus* X shows two to three times greater *Xite* RNA levels at steady state as compared with the 129 X (weaker *Xce*, more likely to be inactivated). Furthermore, there is also a correlation between *Xce* strength and the openness of chromatin. The *Xite* locus contains at least six DHSs, two associated with the first transcription start site and four very strong sites associated with the second. Interestingly, the degree of hypersensitivity is measurably greater on the *M. castaneus* X chromosome than on the 129 X. Therefore, the likelihood of XCI on one X may depend on the activity state and chromatin conformation of the *Xite* locus.

4. MECHANISMS OF *Xite* ACTION

How does *Xite* work to regulate choice? Given that an asymmetry of *Tsix* expression designates X-inactivation choice, we considered from the outset that *Xite* works through *Tsix*. Indeed, genetic analysis supports this epistatic relationship. When *Tsix* levels were measured in mutant female ES cells, an interesting effect was observed (Ogawa & Lee 2003). Prior to XCI, *Tsix* expression was nearly equal from the wild-type and *Xite*-deficient X chromosomes, indicating that *Xite* exerts a minimal effect on *Tsix* at this stage. However, as cells go through the XCI process (induced by differentiation of ES cells), there was a dramatic decrease in *Tsix* expression from the *Xite*-deficient X. This result suggested that not only does *Xite* affect *Tsix* but that *Xite* acts on *Tsix* only once cells enter the XCI pathway. This temporally specific effect on *Tsix* led us to hypothesize that *Xite*'s function might be to promote the persistence of *Tsix* expression during XCI. That is, *Xite* is required to keep *Tsix* on once cells enter the XCI pathway. In this way, *Xite* would be directly responsible for generating the asymmetry of *Tsix* expression during XCI (figure 2). In this model, when two Xs harbour identical *Xce* haplotypes, the *Xite* alleles are equivalent and do not provide a competitive advantage to either *Tsix* allele. However, when one X carries a stronger *Xce* (*Xite*), *Tsix* from that X has the competitive edge and is more likely to persist during XCI.

How does *Xite* act on *Tsix*? Could *Xite* act as a non-coding RNA? To address its potential role as RNA, we introduced three polyadenylation cassettes downstream of the second transcription start site, leading to 10-fold reductions in RNA levels (Ogawa & Lee 2003). However, this RNA downregulation did not skew the choice decision. This finding suggests that *Xite*'s effect does not rely on an intact *Xite* transcript. However, it does not rule out the potential significance of transcriptional activity, as polyadenylation signals do not provide transcriptional termination activity. Thus, this mutation in *Xite* may still allow the RNA polymerase complex to proceed past the poly(A) insertion and even allow transcription to continue into *Tsix*.

Three possible mechanisms of action are currently being tested (figure 3). In the first, *Xite* could work as a classical enhancer of *Tsix* expression but would be one that operates specifically at the onset of XCI. Another

Figure 3. Models of *Tsix* regulation by *Xite*. See text for details.

possibility is that *Xite* is an LCR similar to what has been described at the globin locus. Interestingly, the globin LCR is also associated with strong DHSs and developmentally regulated intergenic transcripts that direct temporal switching between foetal and adult globin genes. As an enhancer or LCR, *Xite* could open up the chromatin locally and propagate this favourable chromatin structure into *Tsix* to facilitate binding of transcription factors to the *Tsix* allele *in cis*. Because the globin LCR induces specific chromatin modifications believed to stimulate globin gene expression, it would be interesting in the future to test if *Xite* might have similar effects on *Tsix* and *Xist*. Alternatively, *Xite* itself may be a loading site for *Tsix* transcription factors which could then be 'piggybacked' into the *Tsix* allele *in cis* by the moving RNA polymerase complex.

5. CAN FUNCTIONAL INTERGENIC TRANSCRIPTION BE GENERALIZED?

The *Xic* now joins the beta-globin locus as a rare case of mammalian intergenic transcription with demonstrated and specific biological function. However, intergenic transcription has been detected at other mammalian loci, including a large fraction of human chromosomes 21 and 22 (Shoemaker *et al.* 2001; Kapranov *et al.* 2002; Rinn *et al.* 2003), the IL-4/IL-13 gene cluster in human cells (Rogan *et al.* 1999) and the *H19/Igf2* locus in mouse (Drewell *et al.* 2002a). In the IL-4/IL-13 cluster, low-level transcripts could be detected constitutively along the entire locus in primary human T cells and T-cell lines: they are present regardless of whether IL-4 and IL-13 are activated. By contrast, they are absent in HeLa cells, a cell type that cannot express the IL genes at all. These observations raise the possibility that intergenic transcription may distinguish stably repressed and poised chromatin.

In the mouse *H19/Igf2* locus, very short tandem transcripts named HUC1 and HUC2 have been found in between the reciprocally imprinted *H19* and *Igf2* genes (Drewell *et al.* 2002a). They seem to differ from *Xite* in several ways. They are very short (less than 500 bp), are expressed at high levels and are transcribed biallelically. The biallelic expression pattern suggests that they do not directly cause the parent-of-origin-specific expression pattern. However, with evidence that they can drive high-level expression of a reporter transgene in mesodermal

tissues of mice, these intergenic transcription elements may be enhancers that regulate expression of *H19* and *Igf2* in mesodermal tissues.

Interestingly, evidence so far indicates that intergenic transcription tends to occur in complex loci. One might speculate that intergenic transcription will be a general property of all mammalian chromatin. Far from spurious, this type of non-coding transcription may soon be appreciated as an integral part of gene regulation and maintenance of chromosome structure. In mammals, low-level transcription across intronic regions in immunoglobulin genes is also known to stimulate recombinational activity at VDJ joints and during heavy-chain class switching (Bottaro *et al.* 1994; Harriman *et al.* 1996; Seidl *et al.* 1998). Although this class of transcription is intronic rather than intergenic, it may be part of the same molecular strategy to usher in a change in chromatin state. Finally, given that intergenic transcripts also regulate gene expression in *Drosophila* (Bender & Fitzgerald 2002; Bae *et al.* 2002; Drewell *et al.* 2002b; Hogga & Karch 2002; Rank *et al.* 2002), the likelihood that functional transcription of intergenic space will be a universal phenomenon among metazoans, plants and fungi must now be considered seriously.

I thank Welcome Bender for helpful discussion of intergenic transcription and the BX-C. I also thank Yuya Ogawa for critical reading of the manuscript. J.T.L. is supported by a grant from the National Institutes of Health, USA, and is an assistant investigator of the Howard Hughes Medical Institute and a Pew Scholar.

REFERENCES

- Ashe, H. L., Monks, J., Wijgerde, M., Fraser, P. & Proudfoot, N. J. 1997 Intergenic transcription and transinduction of the human beta-globin locus. *Genes Dev.* **11**, 2494–2509.
- Avner, P. & Heard, E. 2001 X-chromosome inactivation: counting, choice, and initiation. *Nature Rev. Genet.* **2**, 59–67.
- Avner, P., Prissette, M., Arnaud, D., Courtier, B., Cecchi, C. & Heard, E. 1998 Molecular correlates of the murine *Xce* locus. *Genet. Res. Camb.* **72**, 217–224.
- Azofeifa, J., Voit, T., Hubner, C. & Cremer, M. 1995 X-chromosome methylation in manifesting and healthy carriers of dystrophinopathies: concordance of activation ratios among first degree female relatives and skewed inactivation as cause of the affected phenotypes. *Hum. Genet.* **96**, 167–176.

- Bae, E., Calhoun, V. C., Levine, M., Lewis, E. B. & Drewell, R. A. 2002 Characterization of the intergenic RNA profile at *abdominal-A* and *Abdominal-B* in the *Drosophila* bithorax complex. *Proc. Natl Acad. Sci. USA* **99**, 16 847–16 852.
- Bell, A. & Felsenfeld, G. 2000 Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* **405**, 482–485.
- Bell, A. C., West, A. G. & Felsenfeld, G. 1999 The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387–396.
- Bender, W. & Fitzgerald, D. P. 2002 Transcription activates repressed domains in the *Drosophila* bithorax complex. *Development* **129**, 4923–4930.
- Boggs, B. A., Cheung, P., Heard, E., Spector, D. L., Chinault, A. C. & Allis, C. D. 2002 Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes. *Nature Genet.* **30**, 73–76.
- Bottaro, A., Lansford, R., Xu, L., Zhang, J., Rothman, P. & Alt, F. W. 1994 S region transcription *per se* promotes basal IgE class switch recombination but additional factors regulate the efficiency of the process. *EMBO J.* **13**, 665–674.
- Boumil, R. M. & Lee, J. T. 2001 40 years of decoding the silence in X-chromosome inactivation. *Hum. Mol. Genet.* **10**, 2225–2232.
- Brockdorff, N., Ashworth, A., Kay, G. F., McCabe, V. M., Norris, D. P., Cooper, P. J., Swift, S. & Rastan, S. 1992 The product of the mouse *Xist* gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **71**, 515–526.
- 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 Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature* **349**, 82–84.
- Brown, C. J., Hendrich, B. D., Rupert, J. L., Lafreniere, R. G., Xing, Y., Lawrence, J. & Willard, H. F. 1992 The human *XIST* gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **71**, 527–542.
- Cattanach, B. M. & Isaacson, J. H. 1967 Controlling elements in the mouse X chromosome. *Genetics* **57**, 331–346.
- Chao, W., Huynh, K. D., Spencer, R. J., Davidow, L. S. & Lee, J. T. 2002 CTCF, a candidate *trans*-acting factor for X-inactivation choice. *Science* **295**, 345–347.
- Chureau, C., Prissette, M., Bourdet, A., Barbe, V., Cattolico, L., Jones, L., Eggen, A., Avner, P. & Duret, L. 2002 Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. *Genome Res.* **12**, 894–908.
- Clemson, C. M., McNeil, J. A., Willard, H. & Lawrence, J. B. 1996 *XIST* RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.* **132**, 259–275.
- Clerc, P. & Avner, P. 1998 Role of the region 3' to *Xist* in the counting process of X-chromosome inactivation. *Nature Genet.* **19**, 249–253.
- Drewell, R. A., Arney, K. L., Arima, T., Barton, S. C., Brenton, J. D. & Surani, M. A. 2002a Novel conserved elements upstream of the H19 gene are transcribed and act as mesodermal enhancers. *Development* **129**, 1205–1213.
- Drewell, R. A., Bae, E., Burr, J. & Lewis, E. B. 2002b Transcription defines the embryonic domains of *cis*-regulatory activity at the *Drosophila* bithorax complex. *Proc. Natl Acad. Sci. USA* **99**, 16 853–16 858.
- Gribnau, J., Diderich, K., Pruzina, S., Calzolari, R. & Fraser, P. 2000 Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. *Mol. Cell* **5**, 377–386.
- Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M. & Tilghman, S. M. 2000 CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**, 486–489.
- Harriman, G. R., Bradley, A., Das, S., Rogers-Fani, P. & Davis, A. C. 1996 IgA class switch in I alpha exon-deficient mice. Role of germline transcription in class switch recombination. *J. Clin. Invest.* **97**, 477–485.
- Heard, E., Mongelard, F., Arnaud, D., Chureau, C., Vourc'h, C. & Avner, P. 1999 Human *XIST* yeast artificial chromosome transgenes show partial X inactivation center function in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA* **96**, 6841–6846.
- Hogga, I. & Karch, F. 2002 Transcription through the *iab-7 cis*-regulatory domain of the bithorax complex interferes with maintenance of polycomb-mediated silencing. *Development* **129**, 4915–4922.
- Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C. F., Wolffe, A., Ohlsson, R. & Lobanenko, V. V. 2000 Functional association of CTCF with the insulator upstream of the *H19* gene is parent of origin-specific and methylation-sensitive. *Curr. Biol.* **10**, 853–856.
- Kapranov, P., Cawley, S. E., Drenkow, J., Bekiranov, S., Strausberg, R. L., Fodor, S. P. & Gingeras, T. R. 2002 Large-scale transcriptional activity in chromosomes 21 and 22. *Science* **296**, 916–919.
- Keohane, A. M., Belyaev, N. D., Lavender, J. S., O'Neill, L. P. & Turner, B. M. 1996 X-inactivation and histone H4 acetylation in embryonic stem cells. *Dev. Biol.* **180**, 618–630.
- Lander, E. S. (and 48 others) 2003 Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.
- Lee, J. T. 2002 Homozygous *Tsix* mutant mice reveal a sex-ratio distortion and revert to random X-inactivation. *Nature Genet.* **32**, 195–200.
- Lee, J. T. & Lu, N. 1999 Targeted mutagenesis of *Tsix* leads to nonrandom X-inactivation. *Cell* **99**, 47–57.
- Lee, J. T., Strauss, W. M., Dausman, J. A. & Jaenisch, R. 1996 A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell* **86**, 83–94.
- Lee, J. T., Davidow, L. S. & Warshawsky, D. 1999 *Tsix*, a gene antisense to *Xist* at the X-inactivation center. *Nature Genet.* **21**, 400–404.
- Li, E. 2002 Chromatin modification and epigenetic reprogramming in mammalian development. *Nature Rev. Genet.* **3**, 662–673.
- Lipshitz, H. D., Peattie, D. A. & Hogness, D. S. 1987 Novel transcripts from the Ultrabithorax domain of the bithorax complex. *Genes Dev.* **1**, 307–322.
- Lobanenko, V. V., Nicolas, R. H., Adler, V. V., Paterson, H., Klenova, E. M., Polotskaja, A. V. & Goodwin, G. H. 1990 A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5' flanking sequence of the chicken *c-myc* gene. *Oncogene* **5**, 1743–1753.
- Luikenhuis, S., Wutz, A. & Jaenisch, R. 2001 Antisense transcription through the *Xist* locus mediates *Tsix* function in embryonic stem cells. *Mol. Cell Biol.* **21**, 8512–8520.
- Lyon, M. F. 1961 Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* **190**, 372–373.
- Mak, W., Baxter, J., Silva, J., Newall, A. E., Otte, A. P. & Brockdorff, N. 2002 Mitotically stable association of polycomb group proteins Eed and Enx1 with the inactive X chromosome in trophoblast stem cells. *Curr. Biol.* **12**, 1016–1020.
- Migeon, B. R., Kazi, E., Haisley-Royster, C., Hu, J., Reeves, R., Call, L., Lawler, A., Moore, C. S., Morrison, H. & Jeppesen, P. 1999 Human X inactivation center induces random X chromosome inactivation in male transgenic mice. *Genomics* **59**, 113–121.
- Morey, C., Arnaud, D., Avner, P. & Clerc, P. 2001 *Tsix*-mediated repression of *Xist* accumulation is not sufficient for normal random X inactivation. *Hum. Mol. Genet.* **10**, 1403–1411.

- Mouse Genome Sequencing Consortium 2002 Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562.
- Naumova, A. K., Plenge, R. M., Bird, L. M., Leppert, M., Morgan, K., Willard, H. F. & Sapienza, C. 1996 Heritability of X chromosome-inactivation phenotype in a large family. *Am. J. Hum. Genet.* **58**, 1111–1119.
- Naumova, A. K., Olien, L., Bird, L. M., Simth, M., Verner, A. E., Leppert, M., Morgan, K. & Sapienza, C. 1998 Genetic mapping of X-linked loci involved in skewing of X chromosome inactivation in the human. *Eur. J. Hum. Genet.* **6**, 552–562.
- Ogawa, Y. & Lee, J. T. 2003 Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol. Cell* **11**, 731–743.
- Ohlsson, R., Renkawitz, R. & Lobanenko, V. V. 2001 CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet.* **7**, 520–527.
- Peters, A. H. F. M., Mermoud, J. E., O'Carroll, D., Pagani, M., Schweizer, D., Brockdorff, N. & Jenuwein, T. 2002 Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nature Genet.* **30**, 77–80.
- Plant, K. E., Routledge, S. J. & Proudfoot, N. J. 2001 Intergenic transcription in the human beta-globin gene cluster. *Mol. Cell Biol.* **21**, 6507–6514.
- Plenge, R. M., Hendrich, B. D., Schwartz, C., Arena, J. F., Naumova, A., Sapienza, C., Winter, R. M. & Willard, H. F. 1997 A promoter mutation in the *XIST* gene in two unrelated families with skewed X-chromosome inactivation. *Nature Genet.* **17**, 353–356.
- Rank, G., Prestel, M. & Paro, R. 2002 Transcription through intergenic chromosomal memory elements of the *Drosophila* Bithorax complex correlates with an epigenetic switch. *Mol. Cell Biol.* **22**, 8026–8034.
- Rastan, S. & Robertson, E. J. 1985 X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. *J. Embryol. Exp. Morph.* **90**, 379–388.
- Rinn, J. L., Eukirchen, G., Bertone, P., Martone, R., Luscombe, N. M., Hartman, S., Harrison, J., Nelson, F. K., Miller, P. & Gerstein, M. 2003 The transcriptional activity of human chromosome 22. *Genes Dev.* **17**, 529–540.
- Rogan, D. F., Cousins, D. J. & Staynov, D. Z. 1999 Intergenic transcription occurs throughout the human IL4/IL-13 gene cluster. *Biochem. Biophys. Res. Commun.* **255**, 556–561.
- Sado, T., Wang, Z., Sasaki, H. & Li, E. 2001 Regulation of imprinted X-chromosome inactivation in mice by *Tsix*. *Development* **128**, 1275–1286.
- Seidl, K. J., Bottaro, A., Vo, A., Zhang, J., Davidson, L. & Alt, F. W. 1998 An expressed neo(r) cassette provides required functions of the Igamma2b exon for class switching. *Int. Immunol.* **10**, 1683–1692.
- Shoemaker, D. D., Schadt, E. E., Armour, C. D., He, Y. D., Garrett-Engle, P., McDonagh, P. D., Loerch, P. M., Leonardson, A., Lum, P. Y. & Cavet, G. 2001 Experimental annotation of the human genome using microarray technology. *Nature* **409**, 922–927.
- Simmler, M.-C. (and 10 others) 1996 A 94 kb genomic sequence 3' to the murine *Xist* gene reveals an AT rich region containing a new testis specific gene *Tsx*. *Hum. Mol. Genet.* **5**, 1713–1726.
- Stavropoulos, N., Lu, N. & Lee, J. T. 2001 A functional role for *Tsix* transcription in blocking *Xist* RNA accumulation but not in X-chromosome choice. *Proc. Natl Acad. Sci. USA* **98**, 10 232–10 237.
- Verkerk, A. J., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M. F. & Zhang, F. P. 1991 Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905–914.

Discussion

K. Arney (MRC Clinical Sciences Centre, Imperial College, London, UK). Can you link up the *Xite* transcripts? Where are the *Xist/Tsix/Xite* enhancers?

J. T. Lee. There are at least two initiation sites for the *Xite* transcripts. At present, it is not clear how long each transcript is, but we know that there is transcription going across the entire region from *Xite* to *Tsix*. Some of the *Xite* transcripts even cross into the *Tsix* locus. However, the *Xite* effect does not depend on an additive effect of *Xite* RNA to *Tsix* RNA because truncation of the *Xite* RNA does not produce the skewing phenotype seen by deleting the *Xite* locus. We have not yet found enhancers for *Xist* and *Tsix*, but the models predict that there is one.

T. Nesterova (MRC Clinical Sciences Centre, Imperial College, London, UK). What is the developmental profile of *Xite*? Is there any correlation between randomness of X-inactivation and *Xite/Tsix* transcriptional level?

J. T. Lee. *Xite* RNA and associated hypersensitive sites are expressed prior to and during XCI. They disappear at the end of X inactivation. While we know that *Tsix* transcription is important for keeping the X on, there is no correlation between the level of *Tsix* RNA and the likelihood of inactivating the linked X chromosome. However, there is a good correlation between *Xite* RNA level and inactivation of the X chromosome in *cis*: the X with the higher *Xite* level is more likely to stay on.

T. Bestor (Department of Genetics and Development, Columbia University, New York, NY, USA). Where are the breakpoints that separate *Xist* and *Xce*?

J. T. Lee. The *Xce* was previously mapped by Avner's group to be more than 12–27 kb downstream of *Xist*. This is based on one published meiotic recombinant in mice. *Xite* is 20–30 kb downstream. We propose that *Xite* is the *Xce*, based on its location and its knockout phenotype. When *Xite* is deleted, the probability increases that the X chromosome in *cis* will be silenced. Thus, *Xite* influences the likelihood of inactivation but does not preclude inactivation of either X. This is a hallmark of the *Xce*. Furthermore, we have shown that the X chromosome with a stronger *Xce* (which is more likely to stay active) is more hypersensitive to DNaseI at *Xite* and also expresses more *Xite* transcripts. Numerous DNA polymorphisms between mouse strains occur at *Xite*, any of which may underlie *Xce* differences.

GLOSSARY

- BX-C: bithorax complex
 CpG: cytosine–guanosine dinucleotide
 CTCF: CCCTC-binding factor
 DHS: DNaseI-hypersensitive site
 ES: embryonic stem
 IL: interleukin
 LCR: locus control region
 LINE: long interspersed transposable element
 PRE: polycomb-response element
 SINE: short interspersed transposable element
 Xa: active chromosome
 Xce: X-controlling element
 XCI: X-chromosome inactivation
 Xi: inactive chromosome
 Xic: X-inactivation centre