Human Molecular Genetics, 2001, Vol. 10, No. 6 581–589

Primary non-random X inactivation associated with disruption of *Xist* promoter regulation

Alistair E. T. Newall[1], Sarah Duthie[1], Emma Formstone[1], Tatyana Nesterova[1], Maria Alexiou[2], Colette Johnston[1], Marie-Laure Caparros[1] and Neil Brockdorff[1,+]

[1]X inactivation Group, MRC Clinical Sciences Centre, ICSM, Hammersmith Hospital, Du Cane Road, London W12 ONN, UK and [2]MRC Clinical Sciences Centre Transgenic Facility

Received 15 November 2000; Revised and Accepted 29 January 2001

In this report we demonstrate primary non-random X chromosome inactivation following targeted mutagenesis of a region immediately upstream of *Xist* promoter P_1. In heterozygous animals there is a preferential inactivation of the targeted X chromosome in 80–90% of cells. The phenotype correlates with inappropriate activation of *Xist* in a proportion of the mutant XY embryonic stem cells. Strand-specific analysis revealed increased sense transcription initiating upstream of *Xist* promoter P_1. There was, however, no discernible effect on transcription from the antisense *Tsix* gene. We demonstrate that the *in vitro* and *in vivo* phenotypes are specifically attributable to the presence of a PGKneo cassette at the targeted locus. These findings are discussed in the context of understanding mechanisms of *Xist* gene regulation in X inactivation.

INTRODUCTION

X chromosome inactivation (X inactivation) is the mechanism of sex chromosome dosage compensation in mammals. A single X chromosome, selected at random, is inactivated in all female (XX) cells early in development (1). After initiation of X inactivation the inactive state is clonally inherited through subsequent cell generations.

Although X inactivation is normally random there are examples where partial or complete non-random X inactivation occurs. These can be divided into primary and secondary mechanisms. Primary non-random X inactivation is caused by a bias in the initial choice of which X to inactivate, whilst secondary non-randomness arises through selection against cells having elected to inactivate a given X. Examples of the former are provided by imprinted inactivation of the paternally derived X (Xp) in extraembryonic lineages in mouse (2), and skewed X inactivation caused by heterozygosity at the *Xce* locus (3,4). Examples of the latter include many X:autosome translocations where there is selection against cells that inactivate the translocated X chromosome (see for example ref. 5).

Regulation of X inactivation is mediated by a single *cis*-acting master switch locus, classically referred to as the X inactivation centre (Xic) (6). The Xic is involved in the process whereby cells determine how many (if any) and which X chromosomes to inactivate (counting and choice respectively), and is also required for propagation of X inactivation *in cis* (reviewed in ref. 7). Prevailing models suggest that counting and choice are achieved by cells blocking a single Xic, thereby marking the active X chromosome (8). In recent years Xic functions have been mapped to the X inactive specific transcript (*Xist*) locus and surrounding region (9–15). *Xist* produces a large non-protein coding RNA that is expressed exclusively from the inactive X chromosome and which is thought to propagate inactivation by accumulating along the entire length of the chromosome *in cis* (11,13,16).

Developmental regulation of *Xist* gene expression has been studied extensively in mouse. In XX somatic cells *Xist* is transcribed from two promoters, P_1 and P_2, that appear to be functionally equivalent (11,17). Prior to initiation of random X inactivation in the epiblast and in embryonic stem (ES) cells *Xist* transcripts are unstable (18,19). The basis for this is presently unclear. We have proposed transcription of a distinct isoform produced from an upstream promoter P_0 based on the detection of transcription upstream of promoters P_1 and P_2 in undifferentiated ES cells (17). Subsequent studies have shown however that an antisense RNA, *Tsix*, initiated 15 kb downstream of *Xist* accounts for much of the transcription observed in the upstream region, arguing against a promoter switch (20–22). The above findings led to the hypothesis that *Xist* RNA stability is developmentally regulated and that *Tsix* expression is required to destabilize *Xist* RNA on the active X chromosome-elect at the onset of X inactivation (23). However, the site of initiation of unstable sense transcripts remains uncertain and analysis to date has failed to detect endogenous P_1/P_2 activation in normal ES cells. In addition, stable *Xist* RNA is observed in early pre-implantation stage embryos (17,19), and also following induction of an *Xist* cDNA transgene in undifferentiated ES cells (24). To resolve these apparent contradictions it has been suggested that there is a threshold level for *Xist* transcripts over which stabilization and spreading along the X chromosome occur (24). Regulation of imprinted X inactivation occurs by a different mechanism, with pre-emptive accumulation of *Xist* RNA on the paternal allele apparently underlying preferential Xp inactivation in differentiating trophectoderm and primitive endoderm lineages (17,19). In a recent report the *Tsix* promoter region has been shown to be

+To whom correspondence should be addressed. Tel: +44 208 383 8278; Fax: +44 208 383 8303; Email: nbrockdo@hgmp.mrc.ac.uk

required for repression of the Xm *Xist* allele in imprinted X inactivation (25).

Gene targeting and transgenic analysis have demonstrated that the Xic functions of chromosome counting and propagation map to genetically separable regions of the *Xist* locus. A complex picture is emerging with multiple elements being implicated. Gene targeting in XX ES cells demonstrated that deletion of P_1 and P_2, together with 7 kb of transcribed sequence disrupts *cis*-propagation of X inactivation but not counting (15). Transgene analysis led to narrowing down of the sequences required for counting to a 35 kb region encompassing *Xist* (14,26). Elements located both upstream of P_1/P_2 and downstream of the 3′ end of *Xist* have been implicated in subsequent experiments (27,28).

X chromosome choosing also appears to involve multiple and distinct elements. A targeted *Xist* allele has been reported to result in complete primary non-random X inactivation of the normal X chromosome in heterozygous animals (29). Based on comparison with earlier targeting studies these results were interpreted as evidence for a positive choice element located close to the centre of the *Xist* transcribed locus. In classical experiments heterozygosity at the *Xce* locus has been shown to cause partially skewed X inactivation at the primary level (4). Recent genetic mapping demonstrates that *Xce* is tightly linked to *Xist* but at least 40 kb downstream of the 3′ end of the gene (30). More recently, deletion of the *Tsix* promoter located ~15 kb downstream of *Xist* was shown to result in near complete primary non-random X inactivation of the targeted chromosome in heterozygotes (23,25).

In this report we describe extreme primary non-random X inactivation attributable to insertion of a PGKneo cassette immediately upstream of *Xist* promoter P_1. The phenotype correlates with probabilistic activation of P_1/P_2 promoters in a proportion of undifferentiated targeted XY ES cells and increased sense transcription initiated upstream of P_1. We discuss these findings in the context of understanding mechanisms of *Xist* regulation in X inactivation.

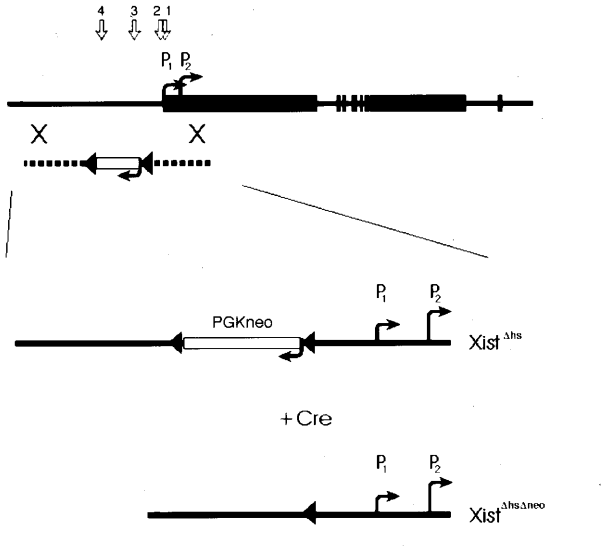

Figure 1. Gene targeting strategy. Illustration of the *Xist* locus showing the position of hypersensitive sites 1–4 relative to the promoters P_1 and P_2. The targeting construct is depicted below. Dashed lines correspond to regions of homology and the open box to the PGKneo selection cassette. The *PGK-1* promoter is oriented in the opposite direction to *Xist* promoters. LoxP sites are depicted by filled arrowheads. Shown below is an expanded view illustrating the $Xist^{\Delta hs}$ targeted allele and below that the $Xist^{\Delta hs\Delta neo}$ allele created by Cre-mediated deletion.

RESULTS

Deletion of hypersensitive sites 3 and 4

We designed a gene targeting strategy to test the function of two DNase1 hypersensitive sites (HS3 and HS4) located 1.4 and 3 kb upstream of the P_1 promoter, respectively (Fig. 1). Both HS3 and HS4 are detectable in XY and XX ES cells and also in XY and XX somatic cells (31 and unpublished data). The targeting construct was designed to replace a 2.5 kb region flanking HS3 and 4 with the neomycin resistance gene (neo^r) driven by the human *PGK-1* promoter (PGKneo). The PGKneo cassette is flanked with loxP sites so that Cre recombinase-mediated deletion can be used to excise the selectable marker in a subsequent step. The construct was transfected into the 129/1 XY ES cell line and a correctly targeted line, designated $Xist^{\Delta hs}$, was identified by Southern blot hybridization at both 3′ and 5′ ends (data not shown). Transient transfection of $Xist^{\Delta hs}$ ES cells with pMC-Cre, a plasmid encoding Cre recombinase was used to derive sub-lines in which the PGKneo cassette was deleted. These lines were designated $Xist^{\Delta hs\Delta neo}$.

Both $Xist^{\Delta hs}$ and $Xist^{\Delta hs\Delta neo}$ cells were injected into blastocysts to produce chimaeric animals. Male chimaeras were then bred to determine whether targeted ES cells are transmitted through the germ line. Previously it has been shown that XY ES cells with an *Xist* null allele contribute to the male germ line of chimaeric animals but that resultant female progeny die in early embryogenesis because of failed imprinted X inactivation in extra-embryonic lineages (32). In contrast to this we found that chimaeras for both $Xist^{\Delta hs}$ and $Xist^{\Delta hs\Delta neo}$ cell lines were transmitted and gave rise to approximately equal numbers of male and female progeny. As all female progeny must carry the targeted Xp chromosome, this result demonstrates that the upstream mutation has not affected propagation of X inactivation from the targeted allele. The fact that female progeny are fully viable also indicates that the counting function of the *Xist* locus has not been disrupted.

Extreme non-random X inactivation of the $Xist^{\Delta hs}$ allele

To determine whether the upstream mutation affects the randomness of X inactivation we quantitated allelic levels of *Xist* and the X-linked *Pgk-1* gene in $Xist^{\Delta hs}$ or $Xist^{\Delta hs\Delta neo}$ female heterozygotes. Single Nucleotide Primer Extension (SNuPE) analysis was carried out using characterized polymorphisms between standard laboratory strain and PGK strain animals which carry a polymorphic X chromosome derived from a wild strain. Data from a number of independent experiments are summarized in Figure 2A. In control (129× PGK) F_1 females we observed slightly higher levels of the 129 strain allele for *Xist* (59%) and lower levels for *Pgk-1* (42%). This mild skewing is attributable to the *Xce* effect (the PGK strain carries a strong Xce^c allele whereas $Xist^{\Delta hs}$ and $Xist^{\Delta hs\Delta neo}$ are on a

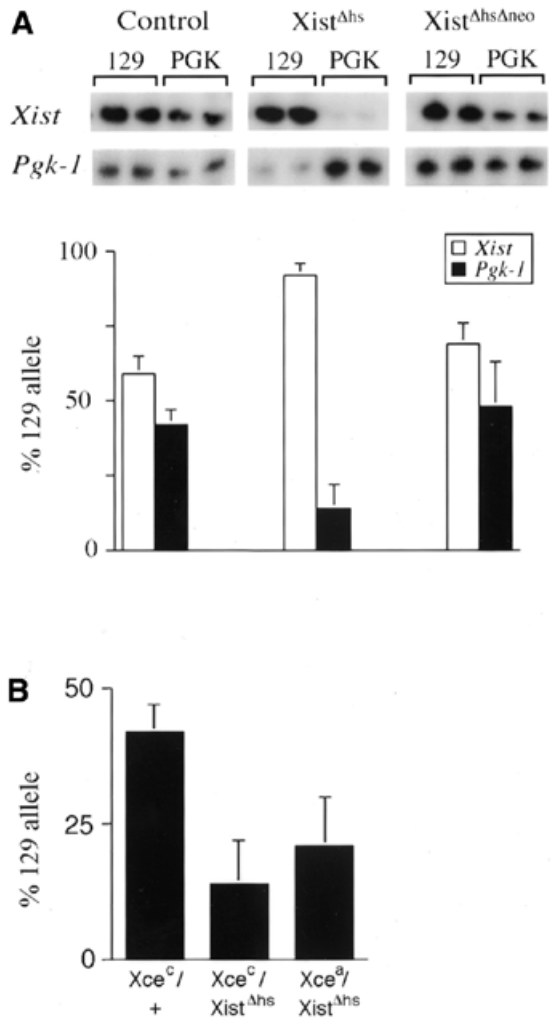

Figure 2. Non-random X inactivation. (**A**) SNuPE analysis of *Xist* and *Pgk-1* alleles in $Xist^{\Delta hs}$ and $Xist^{\Delta hs\Delta neo}$ heterozygotes relative to controls. Above are examples of duplicate loaded samples from a single determination. Graphs illustrate mean values and standard errors for percentage contribution of 129 alleles of *Xist* (open bars) or *Pgk-1* (closed bars) calculated from at least 10 samples from different animals. (**B**) Mean values and standard errors for percentage expression of the 129 allele of *Pgk-1* calculated from eight females heterozygous for $Xist^{\Delta hs}$ and homozygous for Xce^{a} ($Xce^{a}/Xist^{\Delta hs}$). Control values for *Xce* heterozygotes without ($Xce^{c}/+$) or with ($Xce^{c}/Xist^{\Delta hs}$) the $Xist^{\Delta hs}$ allele are from (A) and are included for comparative purposes.

129 strain X chromosome which carries a weak Xce^{a} allele). Targeted ($Xist^{\Delta hs}$ × PGK) F_1 females, however, exhibited extreme non-random X inactivation. The 129 allele of *Xist* and *Pgk-1* represented 92 and 14% of the total signal, respectively. This implies that in heterozygous animals, ~90% of cells inactivate the targeted 129 derived X chromosome. Similar results were obtained in reciprocal crosses demonstrating that parent of origin effects do not play a role (data not shown). Interestingly, $Xist^{\Delta hs\Delta neo}$ heterozygotes showed reversion to a wild-type phenotype (Fig. 2A), indicating that insertion of the PGKneo cassette rather than deletion of HS3 and HS4 underlies the non-random X inactivation.

We went on to test the extent of skewed X inactivation in the absence of *Xce* allelic differences. To do this we crossed $Xist^{\Delta hs}$ males with a female line in which the polymorphic *Pgk-1* allele had been transferred to an Xce^{a} X chromosome (see Materials and Methods). Analysis of *Pgk-1* expression in female progeny from this cross is shown in Figure 2B. A slight decrease in the degree of skewing from ~14% expression of the 129 allele to ~21% was observed. This result indicates that non-random X inactivation attributable to the $Xist^{\Delta hs}$ allele occurs independently of *Xce* effects.

Non-random X inactivation occurs at the primary level

Non-random X inactivation can occur as a result of a bias in the initial choice of which X is inactivated (primary) or through cell selection against cells that keep a given X chromosome active (secondary). Given the extreme degree of skewing in heterozygotes, secondary cell selection events may be predicted to have a discernible effect on the development of $Xist^{\Delta hs}$ female homozygotes. In fact we observed no such effects. Crosses between $Xist^{\Delta hs}/+$ heterozygous females and $Xist^{\Delta hs}/Y$ males produced homozygous $Xist^{\Delta hs}/Xist^{\Delta hs}$ females at the expected frequency (Table 1). This was also the case in equivalent crosses with $Xist^{\Delta hs\Delta neo}$ animals (Table 1). $Xist^{\Delta hs}$ homozygous females and hemizygous males displayed no overt defects and showed no size differences relative to wild-type littermates.

We went on to address the issue of primary versus secondary mechanisms directly by assaying *Xist* and *Pgk-1* alleles in early post-implantation embryos from a cross between $Xist^{\Delta hs}$ heterozygous females and PGK strain males. X inactivation initiates between E5.5 and E6 (33). It has been generally assumed that non-random inactivation due to secondary cell selection events would be detectable as a progressive skewing occurring over the first few days after initiation of X inactivation whereas primary non-random X inactivation would be apparent immediately (see for example ref. 29). Therefore we analysed skewing in embryos as early as possible after the onset of X inactivation (E6.5), and at stages immediately thereafter (E7.5 and E8.5). Results from SNuPE analysis for *Xist* and *Pgk-1* RNA are summarized in Figure 3. At all stages analysed we observed extreme non-random X inactivation of the targeted allele relative to control wild-type littermates. At the E6.5 and E7.5 stages the extent of skewing was less than at E8.5, but in wild-type littermates the *Xce* effect was also reduced (or absent). This difference is probably attributable to the fact that the visceral endoderm is retained in dissected embryonic regions at E6.5 and E7.5. Because cells in this lineage undergo imprinted X inactivation they will express only the 129 strain *Pgk-1* allele and never express the 129 strain *Xist* allele. Taking this into consideration our results indicate that non-random X inactivation in $Xist^{\Delta hs}$ heterozygotes occurs at the primary rather than the secondary level.

Inappropriate activation of *Xist* promoters P_1/P_2 in targeted ES cells correlates with preferential X inactivation

To investigate the basis for extreme non-random X inactivation we analysed *Xist* expression in $Xist^{\Delta hs}$ XY ES cells using RNA FISH. As reported previously (17), using a probe located in exon 1 in conjunction with an upstream probe, we observed a single colocalizing pinpoint signal in control XY ES cells, corresponding to unstable transcripts (Fig. 4A, left panel). The majority of $Xist^{\Delta hs}$ XY ES cells also showed this pattern.

Table 1. *Xist*$^{\Delta hs}$ and *Xist*$^{\Delta hs\Delta neo}$ mutations do not affect viability in XX or XY animals

Progeny	Cross	
	Xist$^{\Delta hs}$/+ (f) × *Xist*$^{\Delta hs}$/Y (m)	*Xist*$^{\Delta hs\Delta neo}$/+ (f) × *Xist*$^{\Delta hs\Delta neo}$/Y (m)
Xist$^{\Delta hs}$/+ (f)	36	
Xist$^{\Delta hs}$/*Xist*$^{\Delta hs}$ (f)	46	
+/Y (m)	25	
Xist$^{\Delta hs}$/Y (m)	23	
Xist$^{\Delta hs\Delta neo}$/+ (f)		8
Xist$^{\Delta hs\Delta neo}$/*Xist*$^{\Delta hs\Delta neo}$ (f)		9
+/Y (m)		11
Xist$^{\Delta hs\Delta neo}$/Y (m)		8

The table illustrates frequency of progeny of defined genotype after crossing heterozygous mutant females (f) with hemizygous mutant males (m) both for *Xist*$^{\Delta hs}$ and *Xist*$^{\Delta hs\Delta neo}$.

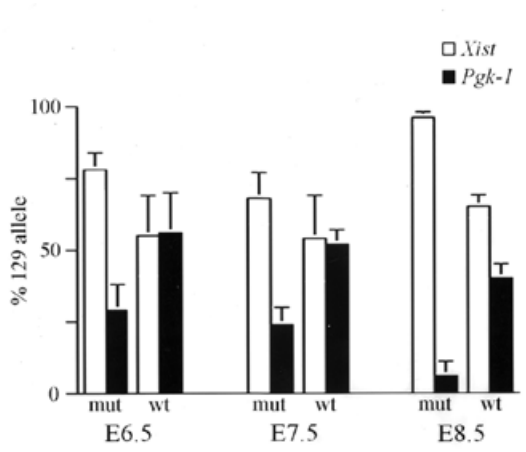

Figure 3. Evidence that non-random X inactivation occurs at the primary level. Results of SNuPE analysis illustrating mean values and standard error for percentage 129 allele *Xist* and *Pgk-1* expression in mutant (mut) and wild-type (wt) littermate embryos obtained at E6.5, E7.5 and E8.5 after mating PGK strain males with *Xist*$^{\Delta hs}$ female heterozygotes. Values were calculated from data obtained from at least five individual embryos.

However, a proportion of cells exhibited a larger *Xist* RNA signal that was detectable only with the downstream probe (Fig. 4A, central and right hand panels). This pattern is similar to that seen in differentiated XX cells (see for example ref. 17), although in *Xist*$^{\Delta hs}$ XY ES cells the signal often appears more diffuse (Fig. 4A, right panel). Accumulation of *Xist* RNA was not seen in *Xist*$^{\Delta hs\Delta neo}$ XY ES cells (data not shown). Therefore, as is the case for skewed X inactivation *in vivo*, the *in vitro* phenotype also correlates with insertion of the PGKneo cassette rather than deletion of HS3 and 4.

To determine the proportion of *Xist*$^{\Delta hs}$ cells with accumulated transcripts, and to ensure that these cells had only a single X chromosome, we used the exon 1 probe for *Xist* RNA FISH in conjunction with a probe specific for X chromosome DNA. In a blind scoring, 12% of *Xist*$^{\Delta hs}$ cells showed high level *Xist* expression compared with a frequency of 0.3% in control XY ES cells (Fig. 4B). Apparent signal in control cells is presumably attributable to non-specific background. No X chromosome aneuploidy was detected.

FISH analysis indicated that cells with accumulated transcripts had inappropriately activated the P_1/P_2 *Xist* promoters. We tested this at the molecular level using nuclease protection. As shown in Figure 4C, transcripts initiated from P_2, the major promoter (17), were undetectable in both control XY ES cells and *Xist*$^{\Delta hs\Delta neo}$ cells. In *Xist*$^{\Delta hs}$ cells, however, P_2 transcripts were clearly detectable, albeit at a lower level than in XX somatic cells. In addition we observed full-length protected product characteristic of P_1 transcripts (17). As P_1/P_2 bands are never detected in control XY ES cells this result points to inappropriate promoter activation in *Xist*$^{\Delta hs}$ XY ES cells.

Accumulation of *Xist* transcripts in ~10% of targeted ES cells should result in steady-state levels approaching 10% of that seen in XX somatic cell lines. To test this we used quantitative RNA slot blot analysis as described previously (17). Combined results from several independent experiments are illustrated in Figure 4D. Using the w7d probe located in exon 1 we observed ~$1.9 \pm 0.4\%$ ($n = 6$) of somatic cell RNA levels in control XY ES cells, consistent with previous results (19), but an elevated level of $5.2 \pm 1.1\%$ ($n = 6$; $P < 0.0001$) in targeted *Xist*$^{\Delta hs}$ XY ES cells. This level is not as high as predicted, presumably reflecting less accumulation relative to XX somatic cells. No difference was observed using the mx3 probe located immediately upstream of P_1/P_2, consistent with elevated levels being attributable to inappropriate P_1/P_2 transcription (Fig. 4D). High levels of accumulated P_1/P_2 transcripts were not seen following differentiation of *Xist*$^{\Delta hs}$ ES cells *in vitro* (data not shown). Presumably this reflects selection as a result of inactivation of the single X chromosome.

Xist/Tsix expression in targeted XY ES cells

As both sense and antisense RNAs contribute to *Xist* transcripts in undifferentiated ES cells we carried out strand-specific RNA FISH analysis on *Xist*$^{\Delta hs}$ ES cells using labelled strand-specific *Xist* exon 1 oligonucleotides in conjuction with

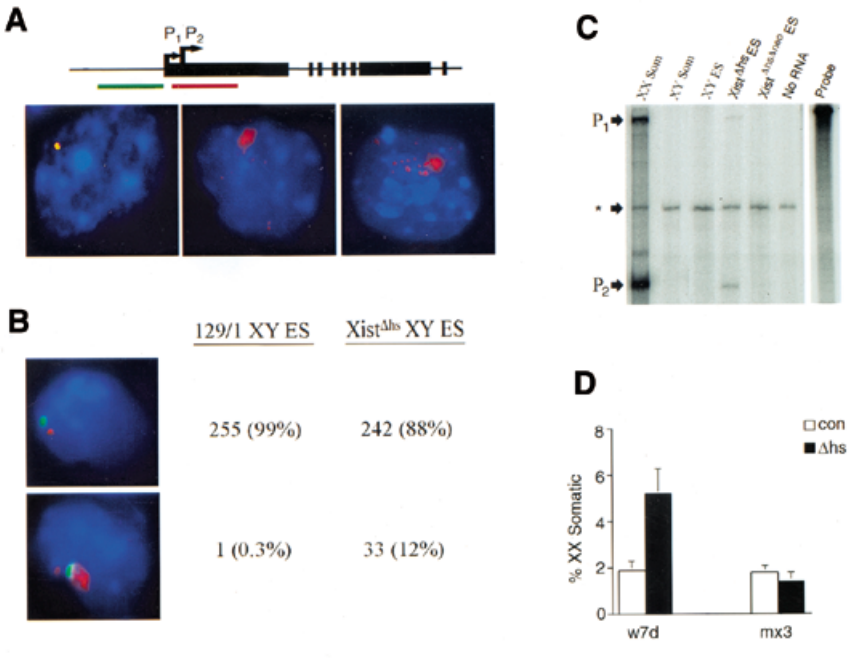

Figure 4. P_1/P_2 activation in *Xist*$^{\Delta hs}$ XY ES cells. (**A**) FISH analysis was carried out using FITC-labelled mx8 probe located upstream of P_1 (green) in conjunction with TRITC-labelled exon 1 probe (red). DNA was counterstained with DAPI (blue). Panels illustrate examples of control 129/1 ES cells (left), where a pinpoint colocalizing signal corresponding to unstable transcripts is detectable in most cells, and *Xist*$^{\Delta hs}$ cells exhibiting accumulated P_1/P_2 transcripts detectable with the exon 1 probe only (centre and right panels). (**B**) Scoring results from DNA/RNA FISH experiment showing the proportion of control 129/1 XY ES cells and *Xist*$^{\Delta hs}$ XY ES cells with accumulated or punctate signal. Examples of each pattern are illustrated to the left. Red signal corresponds to *Xist* RNA and green signal to an X chromosome-specific DNA probe BAC141. DNA was counterstained with DAPI (blue). (**C**) Nuclease protection assay illustrating P_1- and P_2-specific bands. RNA was isolated from C127 XX somatic cells (XX Som), XY somatic cells (XY Som), 129/1 XY ES cells, *Xist*$^{\Delta hs}$ XY ES cells (*Xist*$^{\Delta hs}$ ES) and *Xist*$^{\Delta hs\Delta neo}$ XY ES cells (*Xist*$^{\Delta hs\Delta neo}$ ES). Other lanes show probe digested in the absence of cellular RNA (no RNA) and undigested probe. The band indicated with an asterisk is detectable in the absence of cellular RNA demonstrating that it is an artefact resulting from self-hybridization of the probe. (**D**) *Xist* RNA steady-state analysis in *Xist*$^{\Delta hs}$ XY ES cells (Δhs, filled bars) relative to untargeted 129/1 XY ES cells (con, open bars), following detection with exon 1 probe w7d and the 5′ probe mx3. Data are shown as percentage C127 XX somatic cell RNA level, and represent the mean and standard error from three independent experiments, each performed in duplicate.

a double-stranded exon 1 probe (Fig. 5A). In cells with punctate RNA signal we could detect no overt difference in antisense expression between control and targeted ES cells. Punctate sense transcript was detectable in only 12% of control cells. This low frequency is probably attributable to the fact that we used short oligonucleotide probes as opposed to the larger single-stranded cDNA probes used in some other studies (for example ref. 20). In targeted cells however, punctate sense transcript was seen at a significantly higher frequency (44%).

To analyse this in more detail we carried out strand-specific RT–PCR. Antisense transcription was readily detectable in the upstream region in control and targeted XY ES cells (amplicons 2 and 4, Fig. 5B). In *Xist*$^{\Delta hs}$ cells amplicon 2 product probably represents readthrough transcription from *Tsix* or the PGKneo cassette (which is transcribed in the opposite direction to *Xist*). In control XY ES cells sense transcription was only detected with amplicon 6 located in exon 1 (Fig. 5B). In targeted cells, however, upstream sense transcripts were also readily detectable using amplicon 4. Cre-mediated deletion of the PGKneo cassette in *Xist*$^{\Delta hs\Delta neo}$ cells restored the pattern seen in control ES cells. This result was fully reproducible in independent experiments. 5′ RACE was carried out to determine where the upstream sense transcripts in *Xist*$^{\Delta hs}$ cells initiated. Southern analysis of RACE PCR products using oligonucleotide probe 4s revealed a major band which sequence analysis demonstrated corresponds to initiation ~550 bp upstream of P_1 (Fig. 5C).

DISCUSSION

This study demonstrates that the deletion of two DNase1 hypersensitive sites immediately upstream of *Xist*, together with insertion of a PGKneo selection cassette, causes extreme non-random X inactivation of the chromosome bearing the targeted allele. Non-random X inactivation occurs at the primary level and correlates with premature and inappropriate activation of the *Xist* P_1/P_2 promoters in a proportion of undifferentiated targeted XY ES cells. These cells exhibit high level accumulation of *Xist* RNA similar to that seen in XX somatic cells. Additionally, we observed increased sense transcription from an initiation site upstream of P_1. Both *in vitro* and *in vivo* phenotypes were effectively reversed by Cre-mediated deletion of PGKneo.

Our data indicate that non-random X inactivation *in vivo* occurs at the primary level and therefore corresponds to a bias in the choice of which *Xist* allele is expressed in a given cell. Whilst the basis for this effect is not clear, premature activation of P_1/P_2 promoters in undifferentiated XY ES cells *in vitro* appears to provide a possible model. Specifically, in heterozygous

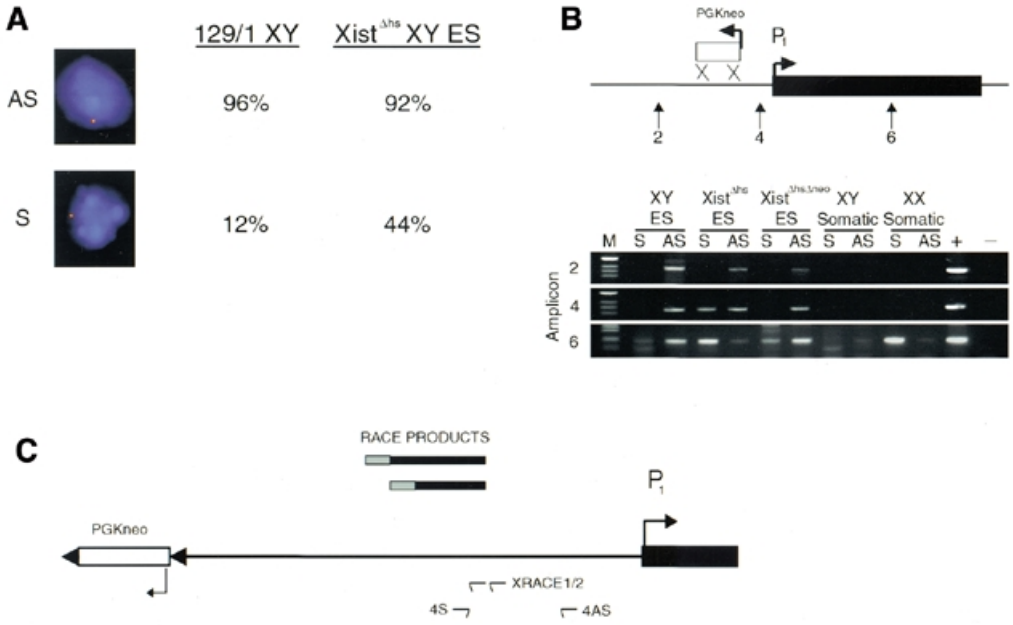

Figure 5. Strand-specific analysis. (**A**) Two colour RNA FISH was performed on 129/1 and *Xist*$^{\Delta hs}$ XY ES cells using double-stranded exon 1 probe (red) in conjunction with sense (S) or antisense (AS) oligonucleotide probes (green). Cells with punctate red signal were scored for the presence of co-localizing green signal (visualized as yellow). Examples of *Xist*$^{\Delta hs}$ XY ES cells with sense and with antisense signal are shown. One-hundred cells from each line were scored for each probe. (**B**) Strand-specific RT–PCR. Total RNA from 129/1 XY ES cells (XY ES), *Xist*$^{\Delta hs}$ XY ES, *Xist*$^{\Delta hs\Delta neo}$ XY ES and wild-type adult male and female kidney (XY and XX somatic) was reverse transcribed using strand-specific primers at each of three amplicons. cDNA was then PCR amplified using the appropriate primer pair (32 or 35 cycles). Also shown are genomic DNA (+) and no template (–) controls, and a 1 kb ladder size marker (M). –RT samples processed in parallel gave no signal (data not shown). Consistent with the report by Debrand *et al.* (21) low level antisense transcription was detected within exon 1 (amplicon 6) in both XX and XY somatic tissue. (**C**) 5′ RACE analysis of upstream sense transcript. *Xist*$^{\Delta hs}$ XY ES cell cDNA specifically primed with the antisense primer 4AS was used as a template for 5′ RACE. Specific product was amplified by nested PCR using the Xist 5′ antisense primers XRACE1 and XRACE2 in conjunction with primers located within the RACE adaptor (grey box). Southern hybridization using the sense primer 4S followed by sequence analysis identified two major products, corresponding to transcriptional initiation at –539 and –616 bases relative to Xist promoter P_1. No RACE product was amplified using control 129/1 XY ES cell 4AS primed cDNA (data not shown).

XX embryos, cells that have prematurely activated P_1/P_2 on the targeted allele could have effectively predetermined which X to inactivate. This would be equivalent to the situation in pre-implantation mouse embryos where accumulation of P_1/P_2 *Xist* transcripts from the Xp allele occurs prior to cellular differentiation (17,19). In this situation early P_1/P_2 expression apparently pre-selects for complete non-random inactivation of Xp in all cells during differentiation of trophectoderm and primitive endoderm lineages.

Disturbance of interactions between distant regulatory elements following insertion of selectable markers has been reported to occur at a number of targeted loci including the β-globin LCR (34) and the immunoglobulin heavy chain constant region locus (35). In the present study it is possible that the inserted PGK promoter acts as a decoy for a negative regulatory element that normally acts either directly or indirectly to repress P_1/P_2 activity in ES cells. P_1/P_2 activation and non-random X inactivation are not observed when PGKneo is targeted 6.5 kb upstream of P_1 in conjunction with a deletion of the putative P_0 promoter (C. Johnston, manuscript in preparation). Similarly, insertion of PGKneo in intron 3 of *Xist* has no effect (M.-L. Caparros, manuscript in preparation), indicating that any such decoy effect must be highly position dependent. In this context it is interesting to note that a mutation in the region of the human *XIST* promoter has been correlated with non-random X inactivation (36).

Interference with either sense or antisense transcripts could also play a causative role in non-random X inactivation, either alone or in combination with premature P_1/P_2 activation. Previously it has been shown that deletion of the major *Tsix* antisense promoter in ES cells results in extreme but not complete non-random X inactivation of the targeted allele (23). In *Xist*$^{\Delta hs}$ cells, however, we detected no difference in the antisense *Tsix* transcript relative to control cells but interestingly we saw a similar X inactivation phenotype associated with increased sense transcription from an upstream site. Although this transcription correlates with insertion of the PGKneo casette, the PGK promoter is oriented in the antisense rather than sense direction (Fig. 1). We mapped the initiation site for these transcripts to a region ~550 bp upstream of P_1, indicating that PGKneo insertion has either activated or hyperactivated an endogenous promoter sequence. It is unclear whether the elevated upstream sense transcription detected by RT–PCR occurs in all cells or whether it is confined to the proportion of *Xist*$^{\Delta hs}$ ES cells exhibiting accumulation of *Xist* RNA. Our data argue against the latter possibility since accumulated *Xist* RNA was never detected using an upstream RNA FISH probe. Additionally, the enhanced RNA FISH detection of punctate sense transcript relative to untargeted control cells is consistent with

an increase in the level of unstable upstream sense transcript. Theoretically, increased sense transcription could be associated with an altered chromatin structure that in turn could underlie the non-random X inactivation phenotype.

Activation of P_1/P_2 transcription in *Xist*$^{\Delta hs}$ ES cells correlates with *cis*-accumulation of high levels of *Xist* RNA. This provides a further example demonstrating that *Xist* RNA stability is not developmentally regulated, but does not allow us to distinguish between the RNA isoform model (17) and the transcript threshold model (24). Clonal analysis of *Xist*$^{\Delta hs}$ XY ES cells indicated that P_1/P_2 activation is not maintained through mitosis (data not shown). This may reflect negative selection as a result of inactivating the single X chromosome or alternatively may indicate that P_1/P_2 activation is dynamic and reversible. Evidence to date has indicated that *cis*-accumulation of stable *Xist* transcripts in undifferentiated cell types does not result in X inactivation. This is based on observations on cells of early pre-implantation embryos (17,19) and on mouse ES cells expressing human *Xist* RNA (37). We observed expression of X-linked loci in *Xist*$^{\Delta hs}$ ES cells with accumulated P_1/P_2 RNA (data not shown), although quantitation of the data indicated a reduction in X-linked gene expression relative to cells without P_1/P_2 transcripts. Thus, partial inactivation of the X chromosome in response to *Xist* RNA accumulation may occur in undifferentiated ES cells. This is consistent with a recent study which demonstrated that *cis*-accumulation of *Xist* RNA and long-range gene silencing can occur following induction of an *Xist* cDNA transgene in ES cells (24).

In summary this study provides evidence for a novel mechanism in which a heterologous promoter upstream of the *Xist* locus can act as a decoy, perturbing normal interactions of regulatory elements that govern X chromosome choice at the onset of random X inactivation.

MATERIALS AND METHODS

Cell lines

Gene targeting was carried out using the 129/1 XY ES cell line isolated in house (38). Cells were maintained and cultured on STO feeders. The XX somatic cell line C127 was maintained and cultured as described previously (39).

Generation of *Xist*$^{\Delta hs}$ and *Xist*$^{\Delta hs\Delta neo}$ alleles

The 5′ and 3′ homologous regions correspond to nucleotides –7484 to –3623 and –1106 to +2380 relative to promoter P_1. The construct was based on a standard positive(neo^r)/negative(thymidine kinase) strategy. The positive selection cassette PGKneo was flanked with loxP sites. 10^7 ES cells were electroporated with 10 µg linearized vector (200 V, 960 µF), and grown for 10 days under geneticin (400 µg/ml) and gancyclovir (2 µM) selection. Individual colonies were picked into 96-well plates and screened by Southern analysis. The recombinant *Xist*$^{\Delta hs}$ line was recovered from 251 +/– selected colonies. To remove the PGKneo cassette, 10^7 *Xist*$^{\Delta hs}$ cells were electroporated with 25 µg supercoiled pMC-Cre plasmid (Gibco BRL) (200 V, 960 µF). After plating out at low density, 163 Cre-transfected clones were analysed and four recombinant *Xist*$^{\Delta hs\Delta neo}$ lines (Cre 1.7D, 2.3E, 2.5F and 2.12E) were recovered. Both *Xist*$^{\Delta hs}$ and Cre 1.7D (*Xist*$^{\Delta hs\Delta neo}$) were found to be karyotypically stable (data not shown), and were used in all subsequent experiments. Male chimaeras produced by blastocyst injection of targeted ES cells were mated to C57BL/6 females. *Xist*$^{\Delta hs}$ and *Xist*$^{\Delta hs\Delta neo}$ colonies were established by breeding with (C57BL/6 × CBA) F_1 animals. Mice were genotyped by PCR analysis of tail-tip DNA; the *Xist*$^{\Delta hs}$ allele was detected by the primer pair TNX35-NeoF2, the *Xist*$^{\Delta hs\Delta neo}$ allele by TNX35-SX2 and the wild-type *Xist* allele by TNX35-HS22 (TNX35: –4109 to –4088; SX2: –475 to –496; HS22: –3078 to –3098; NeoF2: 5'-CCCGTGATATTGCTGAAGAGC-3'). PCR reactions were given 35 cycles of 94°C for 1 min; 53°C for 1 min and 72°C for 1 min. All *Xist* sequence positions are given relative to +1 of *Xist* cDNA.

Analysis of X inactivation *in vivo*

PGK strain animals were bred in house. *Xce*^{a} animals with the polymorphic *Pgk-1*^{a} allele were kindly provided by Bruce Cattanach, MRC Harwell.

DNA and RNA were isolated from fresh tissues as described previously (15). Quantitative *Xist* and *Pgk-1* SNuPE analysis was performed using primers and conditions described previously (15,40). To account for variations in specific activity of nucleotides, analysis of genomic DNA from a (PGK × 129) F_1 female animal was used to normalize the data.

For embryo analysis timed matings were set up between *Xist*$^{\Delta hs}$ heterozygous females and PGK strain males. Embryos were dissected out and separated into embryonic and extra-embryonic components. DNA was prepared from the extra-embryonic component and used for genotyping by PCR. Samples were incubated overnight in 10 µl TE/Tween with 40 µg proteinase K under mineral oil, before being heated at 95°C for 10 min. A total of 2 µl was used for genotyping by the TNX35-NeoF2 PCR as described above. Female embryos were genotyped as carrying either the *Xist*$^{\Delta hs}$ or wild-type maternal X chromosome together with the PGK strain paternal X chromosome. RNA prepared from the embryonic part was used for quantitative SNuPE analysis. RT–PCR analysis of *Xist* expression using a defined *Hin*dIII polymorphism between the PGK strain and standard laboratory strains (38) was used to sex individual embryos.

Analysis of *Xist* expression in ES cells *in vitro*

For FISH analysis preparation of slides, labelling of probes and RNA FISH were carried out as described previously (19). Probes used were GPT16 (*Xist* exon 1) and mx8 (*Xist* upstream region) (17). RNA/DNA FISH was performed as described previously (41). The BAC141 probe is a Bac genomic DNA clone isolated by screening with the Smh141 probe corresponding to an X-specific repeat sequence (42). Strand-specific RNA FISH was performed using exon 1 oligonucleotide probes as described previously (21). Probes were kindly provided by E. Heard.

Steady-state RNA analysis was performed as described previously (17,19), using the upstream probe mx3 and the exon 1 probe w7d. Samples of total RNA (10 µg) were loaded in duplicate slots and 2 ng *Xist* cosmid was loaded as a control. Duplicate samples were subtracted for background, and then normalized to 28S rRNA. Hybridization to *Xist* cosmid was used to normalize the data. Quantitation was performed using a PhosphorImager (Molecular Dynamics; ImageQuant).

Nuclease protection analysis was performed on 10 µg total RNA using the antisense probe RP2 as described previously (17).

Strand-specific RT–PCR was performed essentially as described previously (20). RNA was prepared from ~50 mg freshly dissected tissue or from 10^6 to 10^7 cultured cells, using RNAzol B (Biogenesis). Total cellular RNA (10 µg) was treated with 1 U RNase-free DNase I (Ambion) for 10 min at 37°C. The RNA was then re-extracted and resuspended in a volume of 50 µl RNase-free water. First strand cDNA synthesis was performed using 6 pmol each of the strand-specific *Xist* primers 2S, 2AS, 4S, 4AS, 6S and 6AS described previously (20). Reactions were incubated at 50°C for 1 h and terminated by heating at 80°C for 10 min. Water (10 µl) was added to each tube, and 1 µl cDNA used for PCR amplification with the appropriate sense and antisense primer pairs as described previously (20) with the exception that amplicon 2 was amplified using the pair 2S and TNX6 (5'-GCA AGA TCT TAG CCA AAC GAG-3'), since in our hands this combination gave a cleaner product. Amplicons 2 and 4 were amplified using an annealing temperature of 51°C and were given 35 cycles, amplicon 6 was annealed at 54°C and given 32 cycles.

5′ RACE used the Marathon cDNA amplification kit (Clontech). Double-stranded cDNA was synthesized from 10 µl 4AS-primed first strand cDNA (see above). Following adaptor ligation, first round PCR product was generated between XRACE1 (located –386 to –359 upstream of Xist promoter P_1) and the adaptor primer AP1. Second round product was then generated between XRACE2 (–415 to –385) and adaptor primer AP2.

ACKNOWLEDGEMENTS

We would like to thank members of the X inactivation group for critical reading of this manuscript and Graham Reed and Richard Newton for help with preparing figures. This work was supported by the Medical Research Council, UK. T.N. is supported by an international development award from the Wellcome Trust.

REFERENCES

1. Lyon, M.F. (1961) Gene action in the X chromosome of the mouse (*Mus musculus* L.). *Nature*, **190**, 372–373.
2. Takagi, N. and Sasaki, M. (1975) Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature*, **256**, 640–642.
3. Cattanach, B.M., Pollard, C.E. and Perez, J.N. (1969) Controlling elements in the mouse X-chromosome. I. Interaction with the X-linked genes. *Genet. Res.*, **14**, 223–235.
4. Rastan, S. (1982) Primary non-random X-inactivation caused by controlling elements in the mouse demonstrated at the cellular level. *Genet. Res.*, **40**, 139–147.
5. McMahon, A. and Monk, M. (1983) X-chromosome activity in female mouse embryos heterozygous for Pgk-1 and Searle's translocation, T(X; 16) 16H. *Genet. Res.*, **41**, 69–83.
6. Russell, L.B. and Montgomery, C.S. (1965) The use of X-autosome translocations in locating the X-chromosome inactivation centre. *Genetics*, **52**, 470–471.
7. Rastan, S. and Brown, S.D. (1990) The search for the mouse X-chromosome inactivation centre. *Genet. Res.*, **56**, 99–106.
8. Rastan, S. (1983) Non-random X-chromosome inactivation in mouse X-autosome translocation embryos—location of the inactivation centre. *J. Embryol. Exp. Morphol.*, **78**, 1–22.
9. Borsani, G., Tonlorenzi, R., Simmler, M.C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawrence, C. *et al.* (1991) Characterization of a murine gene expressed from the inactive X chromosome. *Nature*, **351**, 325–329.
10. Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D. and Rastan, S. (1991) Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature*, **351**, 329–331.
11. Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S. and 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.
12. Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R. and Willard, H.F. (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature*, **349**, 38–44.
13. Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafreniere, R.G., Xing, Y., Lawrence, J. and 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.
14. Lee, J.T., Strauss, W.M., Dausman, J.A. and Jaenisch, R. (1996) A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell*, **86**, 83–94.
15. Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S. and Brockdorff, N. (1996) Requirement for Xist in X chromosome inactivation. *Nature*, **379**, 131–137.
16. Clemson, C.M., McNeil, J.A., Willard, H.F. and 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.
17. Johnston, C.M., Nesterova, T.B., Formstone, E.J., Newall, A.E., Duthie, S.M., Sheardown, S.A. and Brockdorff, N. (1998) Developmentally regulated Xist promoter switch mediates initiation of X inactivation. *Cell*, **94**, 809–817.
18. Panning, B., Dausman, J. and Jaenisch, R. (1997) X chromosome inactivation is mediated by Xist RNA stabilization. *Cell*, **90**, 907–916.
19. Sheardown, S.A., Duthie, S.M., Johnston, C.M., Newall, A.E., Formstone, E.J., Arkell, R.M., Nesterova, T.B., Alghisi, G.C., Rastan, S. and Brockdorff, N. (1997) Stabilization of Xist RNA mediates initiation of X chromosome. *Cell*, **91**, 99–107.
20. Lee, J.T., Davidow, L.S. and Warshawsky, D. (1999) Tsix, a gene antisense to Xist at the X-inactivation centre. *Nature Genet.*, **21**, 400–404.
21. Debrand, E., Chureau, C., Arnaud, D., Avner, P. and Heard, E. (1999) Functional analysis of the DXPas34 locus, a 3′ regulator of Xist expression. *Mol. Cell. Biol.*, **19**, 8513–8525.
22. Warshawsky, D., Stavropoulos, N. and Lee, J.T. (1999) Further examination of the Xist promoter-switch hypothesis in X inactivation: Evidence against the existence and function of a P(0) promoter. *Proc. Natl Acad. Sci. USA*, **96**, 14424–14429.
23. Lee, J.T. and Lu, N. (1999) Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell*, **99**, 47–57.
24. Wutz, A. and Jaenisch, R. (2000) A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell*, **5**, 695–705.
25. Lee, J.T. (2000) Disruption of imprinted X inactivation by parent-of-origin effects at Tsix. *Cell*, **103**, 17–27.
26. Herzing, L.B., Romer, J.T., Horn, J.M. and Ashworth, A. (1997) Xist has properties of the X-chromosome inactivation centre. *Nature*, **386**, 272–275.
27. Clerc, P. and Avner, P. (1998) Role of the region 3′ to Xist exon 6 in the counting process of X- chromosome inactivation. *Nature Genet.*, **19**, 249–253.
28. Lee, J.T., Lu, N. and Han, Y. (1999) Genetic analysis of the mouse X inactivation center defines an 80-kb multifunction domain. *Proc. Natl Acad. Sci. USA*, **96**, 3836–3841.
29. Marahrens, Y., Loring, J. and Jaenisch, R. (1998) Role of the Xist gene in X chromosome choosing. *Cell*, **92**, 657–664.
30. Simmler, M.C., Cattanach, B.M., Rasberry, C., Rougeulle, C. and Avner, P. (1993) Mapping the murine Xce locus with (CA)n repeats. *Mamm. Genome*, **4**, 523–530.
31. Sheardown, S.A., Newall, A.E., Norris, D.P., Rastan, S. and Brockdorff, N. (1997) Regulatory elements in the minimal promoter region of the mouse Xist gene. *Gene*, **203**, 159–68.
32. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. and Jaenisch, R. (1997) Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.*, **11**, 156–166.

33. Rastan, S. (1982) Timing of X-chromosome inactivation in postimplantation mouse embryos. *J. Embryol. Exp. Morphol.*, **71**, 11–24.
34. Fiering, S., Epner, E., Robinson, K., Zhuang, Y., Telling, A., Hu, M., Martin, D.I., Enver, T., Ley, T.J. and Groudine, M. (1995) Targeted deletion of 5′HS2 of the murine beta-globin LCR reveals that it is not essential for proper regulation of the beta-globin locus. *Genes Dev.*, **9**, 2203–2213.
35. Seidl, K.J., Manis, J.P., Bottaro, A., Zhang, J., Davidson, L., Kisselgof, A., Oettgen, H. and Alt, F.W. (1999) Position-dependent inhibition of class-switch recombination by PGK-neor cassettes inserted into the immunoglobulin heavy chain constant region locus. *Proc. Natl Acad. Sci. USA*, **96**, 3000–3005.
36. Plenge, R.M., Hendrich, B.D., Schwartz, C., Arena, J.F., Naumova, A., Sapienza, C., Winter, R.M. and 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.
37. Heard, E., Mongelard, F., Arnaud, D., Chureau, C., Vourc′h, C. and 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.
38. Kay, G.F., Penny, G.D., Patel, D., Ashworth, A., Brockdorff, N. and Rastan, S. (1993) Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell*, **72**, 171–182.
39. Jeppesen, P. and Turner, B.M. (1993) The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell*, **74**, 281–289.
40. Singer-Sam, J., LeBon, J.M., Dai, A. and Riggs, A.D. (1992) A sensitive, quantitative assay for measurement of allele-specific transcripts differing by a single nucleotide. *PCR Methods Appl.*, **1**, 160–163.
41. Duthie, S.M., Nesterova, T.B., Formstone, E.J., Keohane, A.M., Turner, B.M., Zakian, S.M. and Brockdorff, N. (1999) Xist RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material *in cis*. *Hum. Mol. Genet.*, **8**, 195–204.
42. Nasir, J., Fisher, E.M., Brockdorff, N., Disteche, C.M., Lyon, M.F. and Brown, S.D. (1990) Unusual molecular characteristics of a repeat sequence island within a Giemsa-positive band on the mouse X chromosome. *Proc. Natl Acad. Sci. USA*, **87**, 399–403. [Erratum (1990) *Proc. Natl Acad. Sci. USA*, **87**, 9508.]