Human Molecular Genetics, 1996, Vol. 5, No. 4 451–459

Xist expression from an *Xist* YAC transgene carried on the mouse Y chromosome

S. Matsuura, V. Episkopou, Renata Hamvas[1] and S. D. M. Brown*

Department of Biochemistry and Molecular Genetics, St. Mary's Hospital Medical School, Imperial College of Science, Technology and Medicine, London W2 1PG, UK and [1]Genome Analysis Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, UK.

Received November 30, 1995; Revised and Accepted January 17, 1996

We have constructed mouse transgenic lines carrying a YAC clone encompassing the *Xist* gene in order to investigate the factors influencing *Xist* expression and the initiation of X-inactivation. Two transgenic lines were derived, one carrying four copies integrated at an autosomal site and a second line carrying four copies integrated at a single site on the Y chromosome. *Xist* expression was not observed in mice carrying the autosomal insertion. However, *Xist* expression from the Y-inserted transgenes was observed and at levels commensurate with that found in normal female mice. Methylation sites in the autosomal transgene both 5′ and 3′ of the *Xist* gene are hypermethylated and appear to reflect methylation patterns observed on the active X chromosome. For the Y-linked transgene, methylation sites 5′ and 3′ of the *Xist* gene are hypomethylated reflecting patterns found on the inactive X chromosome. However, the 5′ and 3′ methylation levels have been decoupled at the active transgenic locus. The data suggest that sequences in the vicinity of *Xist* can initiate some of the features that are associated with the initiation process of X-inactivation.

INTRODUCTION

The *Xist* gene is the only known locus to be expressed from the inactive X chromosome in mouse and human (1–3). Assessment of a variety of chromosomal rearrangements in mouse and human have defined a narrow critical region in both species required for the initiation of X-inactivation—the X-inactivation centre or *Xic* region (4–6). The *Xist* gene maps within the *Xic* region. The map location of *Xist* along with its unusual expression pattern have suggested that *Xist* plays a critical role in the initiation of X-inactivation. The *Xist* gene sequence is conserved between mouse and human (7) but does not appear to encode a gene product. Rather, it has been proposed that the RNA may function directly in initiating the process of heterochromatisation that underlies X-inactivation. Recently, definitive evidence has been obtained that *Xist* is required for the process of X chromosome inactivation (8). Gene targeting of one of the two copies of *Xist* in female ES cells demonstrates that the targeted X chromosome fails to undergo X chromosome inactivation in contrast to the normal X chromosome.

X-inactivation in mouse is first observed in the extra-embryonic tissues where X-inactivation is non-random and *Xist* expression occurs only from the paternal X chromosome. Paternal *Xist* expression can be detected as early as the 4–8 cell stage in mouse where it is associated with an apparent methylation imprint—sites 5′ to the *Xist* gene are hypomethylated (9,10) and this hypomethylation reflects the undermethylation of the *Xist* gene observed in spermatogenesis (10). Subsequently, this paternal imprint is lost at the blastocyst stage and random *Xist* expression in the embryo proper starts prior to gastrulation (9,11). The random X-inactivation process is associated on the inactive X chromosome with hypomethylation around the expressed *Xist* gene at both 5′ (10) and 3′ sites (12). The *Xist* gene on the active X chromosome remains hypermethylated.

In the mouse, a second locus, *Xce* (X-chromosome controlling element), affects the randomness of X-inactivation and three alleles have been defined—Xce^a, Xce^b and Xce^c. In heterozygote mice carrying Xce^a and Xce^b, X chromosomes carrying the Xce^a allele are more likely to be inactivated. In heterozygote mice carrying the Xce^b and Xce^c alleles, X chromosomes carrying the Xce^b allele are likely to be inactivated. The *Xce* locus maps in the vicinity of *Xic* (13) but recent genetic mapping evidence suggests that *e* is a discrete locus separable from the *Xist* locus but lying 3′ to the *Xist* gene in close proximity (14). From this point of view, it is intriguing that the methylation of 3′ sites that are normally hypermethylated on the active X chromosome shows a strong correlation between the level of hypermethylation and the strength of the *Xce* allele (12).

In order to investigate the sequences involved with the initiation of X-inactivation and *Xist* expression and their relationship to the status of methylation sites in the vicinity of the *Xist* locus, we have constructed transgenic lines carrying YAC clones encompassing the *Xist* locus. An autosomal transgenic line failed to express *Xist* in contrast to a transgenic line where the YAC transgenes had inserted onto the Y chromosome.

RESULTS

Structure of YAC transgene and derivation of transgenic lines

A 350 kb YAC clone, yXist1, containing the *Xist* gene has previously been characterised (8; see Fig. 1a). *Xist* lies close to the

*To whom correspondence should be addressed at present address: MRC Mouse Genome Centre, Harwell, Oxon OX11 ORD, UK

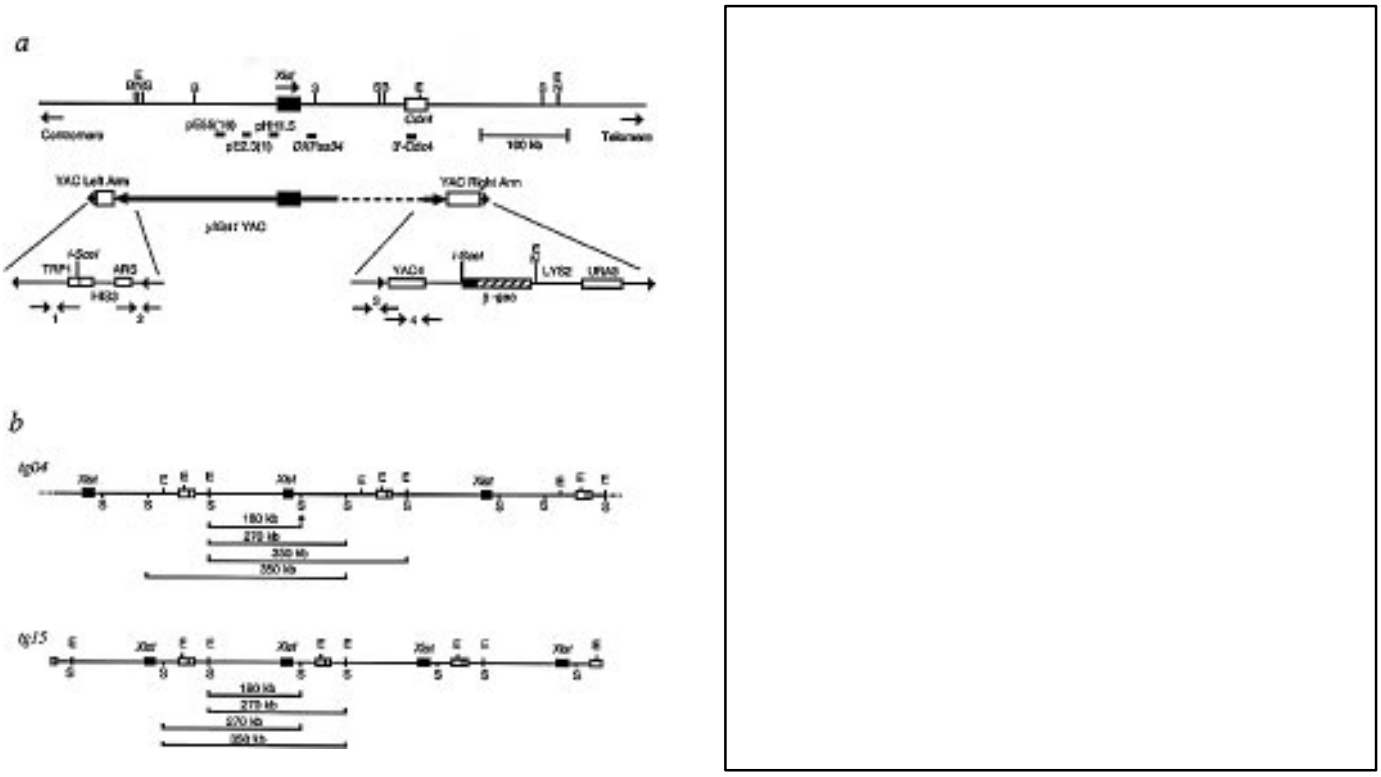

Figure 1. (**a**) A genomic map surrounding the mouse *Xist* locus. The mouse *Cdx4* gene is located 100 kb distal to the *Xist* gene. Probes used in this study are shown as black boxes and indicated. The restriction sites are: B, *Bss*HII; E, *Eag*I; N, *Not*I; S, *Sal*I. Lower panel shows the structure of a 350 kb *Xist* YAC (yXist1). The YAC was modified for transgenesis by the introduction of the i-*Sce*I restriction sites into both arms and the βgeo reporter gene into the right arm. The size of the YAC arms are not drawn to size. The arrows represent PCR primers used for screening of the founder mice. (**b**) Structures of tg04 and tg15 transgenes. In both cases, three to four copies of the *Xist* YAC were integrated at a single site in a tandem head-to-tail fashion. tg15 carried an approximately 100 kb deletion distal to the *Xist* locus, which is indicated as a dotted line within the YAC insert in Fig. 1a. *Sal*I restriction fragments are shown below the maps. A hypermethylated *Sal*I site observed in tg04 is indicated as an asterisk. (**c**) Fluorescent *in situ* hybridization of tg15. The whole *Xist* YAC DNA was used as a probe. Two signals are seen on a single X chromosome (small arrowhead) and on a distal long arm of the Y chromosome (large arrowhead). The intensity of the signal seen on the Y chromosome is stronger than that of the X chromosome.

centre of this YAC which also contains the recently reported *Cdx4* gene (15). yXist1 was modified for transgenesis by the introduction of i-*Sce*I sites into both arms (see Materials and Methods). The introduction of an i-*Sce*I site into the right arm was accompanied by the insertion of a β*geo* reporter gene driven off a PGK promoter. However, this reporter was not used in the present study to assess the *cis* effects of *Xist* expression given the down-regulation that occurs from the PGK promoter during development.

YAC transgenics were recovered by pronuclear injection of mouse fertilised eggs recovered from FVB/N and B6CBAF1 mice. From 1581 injections, 81 mice were born and four transgenic lines were recovered. Potential founder mice were tested by PCR for the presence of both ends of the 350 kb YAC transgene by utilising primers for YAC insert-vector junction sequences (primer set 2 and 3, see Fig. 1a and Materials and Methods). Four founder mice were identified (see Table 1). tg74 was negative for primer set 2 (as well as primer set 1, see Fig. 1a) and had presumably lost the YAC left arm and was not investigated further. The three remaining founders, tg04, tg15 and tg59, were positive for primer sets 1–4 (Fig. 1a) and had retained both YAC arms. One founder, tg59, failed to pass the transgene onto its progeny and was presumably mosaic and was not investigated further. A transgenic line was established from tg04 and the inheritance pattern of the transgene indicated that it had inserted autosomally. The transgene segregated with a dominant phenotype demonstrating a curly tail. The curly tail phenotype was very variable (data not shown). A second transgenic line was established from tg15. From 93 animals recovered from the tg15 transgenic line, the transgene was without exception observed only in males. The transgene has therefore integrated on the Y chromosome. FISH analysis of tg15 using the yXist1 YAC shows two signals (see Fig. 1c), one signal towards the centre of a relatively long acrocentric, presumably the single X chromosome, and a second strong signal on one of the smallest acrocentrics. The transgene would appear to have integrated at a single site on the distal long arm of the Y chromosome (see also below).

Table 1. Summary of transgenic lines obtained by pronuclear microinjection

Transgenic line	tgo4	tg15	tg59	tg74
Sex	F	M	M	M
Phenotype	Curly tail	–	–	–
Transgene	Germ line	Germ line	Mosaic	n.d.
Chromosome	Autosome	Y chromosome	n.d.	n.d.
Copy number	4	4	n.d.	n.d.
Size of transgene	350 kb	240 kb	n.d.	n.d.
Xist transcription from transgene	–	+	n.d.	n.d.

n.d., not done; M, male; F, female. tgo4 and tg74 were derived from microinjections of fertilised eggs from B6CBAF1 mice. tg15 was recovered from injections of fertilised eggs from a B6CBAF1 ×FVB/N cross and tg59 was recovered from injections of eggs from FVB/N mice.

Analysis of the organisation of the transgenic loci, tg04 and tg15

Southern blot analysis was carried out (Fig. 2a) to estimate the copy number of the inserted transgenes in both tg04 and tg15. Digestion with *Eco*RI and probing with a mouse *Xist* probe (W7D) demonstrated that both transgenic lines contained around 3–4 copies of the *Xist* YAC (see Fig. 2 legend). This was confirmed by i-*Sce*I partial digestion of the two transgenic lines followed by pulsed field gel electrophoresis and probing with the same *Xist* probe (Fig. 2b). In both cases, a regular ladder of bands is observed suggesting that there are multiple copies which have integrated at a single site in tandem. However, it was also clear from this analysis that whereas tg04 retained the cognate 350 kb structure, the tg15 transgene carried a deletion as the i-*Sce*I ladder was based upon a 240 kb band (Fig. 2b). Digestion with *Eag*I and *Sal*I (Fig. 2c) also confirmed that tg15 carried an approximately 110 kb deletion and allowed us to map the extent of the deletion to a region 3′ of *Xist* removing the *Cdx4* gene (Fig. 1). The 200 kb band in *Eag*I digests of tg15 hybridised to pE55 (16), pE2.3 (1), pHH1.5 and *DXPas34* probes but failed to hybridise to *Cdx4* (data not shown). The major *Sal*I fragment of 180 kb encompassing *Xist* is intact in both tg04 and tg15 (Fig. 1b and 2c). In tg15 a prominent 180 kb band is visible but it is noteworthy in tg04 that the predominant band is 270 kb (see below). *Sal*I fragments of 270 kb and 440 kb are observed in normal mouse digests, these larger fragments arising from methylation at the *Sal*I sites distal to *Xist* (14). Other fragments are seen in *Sal*I digests of tg04 and tg15 that represent fragments arising from the tandem nature of the inserted transgenes and methylation of *Sal*I sites either in the YAC clone insert or in the right vector arm of the YAC. Specifically, both tg04 and tg15 demonstrate 350 kb fragments. The origin of these fragments is presented diagramatically in Figure 1b and this interpretation agrees with the integration of the transgenes in a tandem head-to-tail fashion. Finally, both tg04 and tg15 show *Sal*I fragments of around 100 kb that may result from one copy of the transgene array, possibly the most 5′ copy, being deleted upstream of the *Xist* gene. In conclusion, each array contains multiple copies of the *Xist* region, most copies containing a substantial amount of sequence both 5′ and 3′ of *Xist* and covering at least 200 kb (see Fig. 1 for summary). In addition, the restriction mapping data confirm the integrity of sequences surrounding *Xist* in both of the YAC transgenes.

Expression of *Xist* from the tg04 and tg15 transgenes

RT-PCR analysis of liver from adult males carrying the tg04 transgene did not show any *Xist* expression (data not shown). Preliminary RT-PCR analysis demonstrated that tg15 adult males did show *Xist* expression. In order to confirm that the observed expression was from the tg15 transgene rather than the parental X chromosome, the tg15 males were crossed to the PGK inbred strain. yXist1 and PGK demonstrate a *Hin*dIII restriction site polymorphism in the *Xist* sequence recovered by RT-PCR (Fig. 3a). Like C57BL/6, the 578 bp *Xist* RT-PCR product of yXist1 is digested to two fragments of 320 bp and 183 bp (plus a smaller fragment of 75 bp, not observed on the gel in Figure 3a). PGK however gives only two products of 503 bp and 75 bp. The data clearly demonstrates that tg15 when crossed into the PGK strain gives expression only from the tg15 transgene and that no expression is observed from the PGK X chromosome.

In addition, we explored the expression of *Xist* from the tg15 transgene at earlier time points. Expression of the tg15 transgene was also observed in 10.5 day old male embryos (data not shown). Furthermore, we looked for evidence of *Xist* expression at the blastocyst (3.5 d.p.c.) stage. *Xist* expression at this stage is exclusively from the imprinted paternally-inherited *Xist* allele (9,11). Analysis of pools of blastocysts derived from tg15/PGK males crossed to PGK females clearly demonstrated expression from the tg15 *Xist* allele (Fig. 3b). In addition, as expected expression is also observed in the pool from the PGK *Xist* allele, presumably from PGK/PGK female blastocysts. Slot blot analysis of RNA from both normal females and tg15 transgenic males, demonstrated that the levels of *Xist* expression from the tg15 transgenic locus were commensurate with the levels of expression seen in normal female mice (Fig. 3c). Finally, Northern analysis of poly(A)$^+$ RNA from tg15 mice demonstrated the presence of an approximately 15 kb *Xist* transcript and an identical hybridisation pattern to CBA female RNA (Fig. 3d).

Methylation of the tg04 and tg15 transgenes

We investigated the methylation patterns in the tg04 and tg15 transgenes at sites both 5′ (10) and 3′ to the *Xist* gene (12). Again we took advantage of polymorphisms between the transgene and the cognate *Xist* gene in the PGK strain in order to identify methylation changes associated specifically with the transgenes. On the inactive X chromosome, it has been observed that several

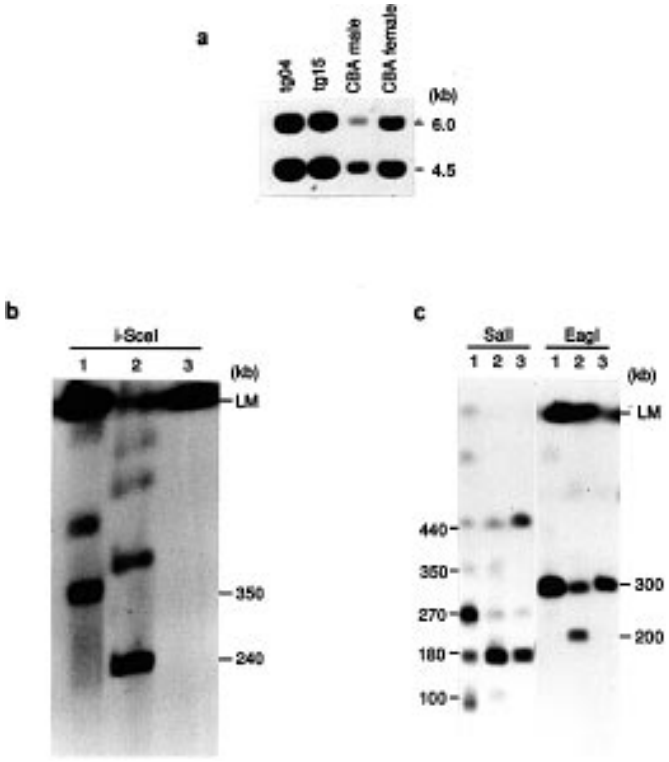

Figure 2. Structural analysis of the tg04 and tg15 transgenes. (**a**) Copy number of the *Xist* YAC. Genomic DNAs in agarose blocks from tg04 male, tg15 male, CBA male and CBA female mouse were digested with *Eco*RI and hybridized to the *Xist* cDNA probe, W7D (see Materials and Methods). The W7D probe detects two bands of 6.0 kb and 4.5 kb. DNA loading was assessed by subsequent hybridisation of the filter to a β-actin probe. Band intensities were measured with a PhosphorImager (see Materials and Methods) and following normalisation for DNA loading the copy number of the tg04 and tg15 transgenes was estimated as 4.14 and 3.58 respectively. (**b**) Analysis of i-*Sce*I partial digestion of tg04 (lane 1), tg15 (lane 2) and CBA female (lane 3) mouse DNA with the *Xist* cDNA probe. Genomic DNAs digested with i-*Sce*I were fractionated by PFGE using a pulse time of 45 s at 170 V for 24 h and hybridized to the W7D probe. In both tg04 and tg15, a ladder of bands is detected. tg04 showed an intact 350 kb band, whereas tg15 showed a deleted band of 240 kb. (**c**) Analysis of *Sal*I and *Eag*I genomic fragments detected by the *Xist* W7D probe. Genomic DNAs from tg04 male, tg15 male and CBA female mouse were digested with *Sal*I and *Eag*I and analyzed by PFGE Southern blot as usual. In control female mouse, the *Xist* probe detected a single 300 kb *Eag*I fragment. *Eag*I of tg15 mouse DNA generated a 200 kb band in addition to the band of normal size. In a normal female mouse, *Sal*I fragments of 180 kb, 270 kb, and 440 kb are observed. The 180 kb *Sal*I fragment encompassing the *Xist* gene was detected in both tg04 and tg15 mouse. In tg15 the prominent *Sal*I fragment was 180 kb, whereas in tg04 the predominant band was 270 kb.

sites, including *Sac*II, *Mlu*I and *Hae*II, just 5′ of the *Xist* gene are hypomethylated (10; see Fig. 4a,b). These sites are fully hypermethylated on the active X chromosomes in males. In C57BL/6 a 9.5 kb *Pvu*II fragment encompasses these sites and is detected by probe 2 (10) while on the PGK chromosome a 4.7 kb fragment is detected by probe 2 and this smaller fragment does not encompass the methylation sites. In tg04 the digestion pattern is typical of a C57BL/6 male and there is no evidence of cutting of the transgenic 9.5 kb *Pvu*II band. However, in tg15 there is clear evidence of hypomethylation. A pattern of digestion is seen that is similar to a typical C57BL/6 female with identical *Sac*II and *Mlu*I bands. Interestingly, a hypomethylated *Hae*II site is detected in tg15 that corresponds to a hypomethylated site detected by Norris *et al*. using probes downstream of the *Xist* promoter (10). The detection of this site by probe 2 however indicates that the *Hae*II site just left of probe 2 must be methylated to some degree in tg15. The intensity of the digestion products in tg15 is much lower than the undigested 9.5 kb product suggesting that only one copy of *Xist* is hypomethylated at the transgenic locus. This would indeed correlate with our estimate that only one copy of the tg15 transgenic array is expressing *Xist*. However, first, in order to assess whether the hypomethylated sites are spread across the transgenic array or confined to one copy we have carried out a sequential digestion experiment in tg15 and C57BL/6 female of a 4.5 kb *Eco*RI fragment encompassing this region. If the hypomethylated sites *Mlu*I and *Sac*II are largely confined together in the same copy, double digestion with these enzymes would result in a single band the same size as *Sac*II alone (see Fig. 4a) with little or no residual *Mlu*I-sized band. This is indeed the case (Fig. 4c)—only a faint *Mlu*I band can be observed in the double digest confirming that the hypomethylation is largely confined to one copy in the transgenic array. In addition, using densitometry from a PhosphorImager (see Materials and Methods), we have estimated the relative degree of undermethylation. If the hypomethylated copy in the transgenic array were completely unmethylated we would expect the intensity of the digested *Sac*II and *Mlu*I bands to be 1/4 the intensity of the 9.5 kb *Pvu*II band and equivalent in intensity to the 4.7 kb *Pvu*II band. Densitometry measurements show that the ratio of digested bands to the 9.5 kb *Pvu*II bands is 0.28 and 0.21 for the *Sac*II and *Mlu*I digests respectively. In addition, the ratio of digested to the 4.7 kb *Pvu*II band is 0.82 and 0.81 for the *Sac*II and *Mlu*I digests respectively. So while we cannot conclude that the undermethylated copy is completely unmethylated, substantial hypomethylation has taken place.

As well as the 5′ sites, a 3′ *Sal*I site is hypomethylated on the inactive X chromosome (12). Using the *DXPas34* probe, we have been able to assess the methylation status of this site at both transgenic loci also taking advantage of a polymorphism between the transgene and the PGK strain (see Fig. 4d). On the inactive chromosome, there is extensive hypomethylation at this site which is observed in female DNA where there is marked cutting of the 4.6 kb *Eco*RI fragment by *Sal*I. tg04 demonstrates a typical male pattern with little cutting of the 4.6 kb fragment from the transgene. However, for tg15 there is noticeable hypomethylation at the 3′ site which reflects the pattern seen in female C57BL/6 (Fig. 4d). It is noteworthy that the level of hypomethylation is comparable to normal females, with equivalent intensity of the 4.6 and 3.2 kb bands indicating that, unlike the 5′ sites (above), the 3′ *Sal*I site is hypomethylated in most copies of the tg15 transgene.

DISCUSSION

We have generated two transgenic lines each carrying multiple copies of a YAC *Xist* transgene. In one line four copies were integrated at a single autosomal site and in the other line four copies were integrated on the Y chromosome. Only the latter line expressed *Xist* and at a level equivalent to that observed in a normal female.

The methylation status of the autosomal and Y-linked transgenes reflects *Xist* expression and activity and mimics the methylation changes found on the active and inactive X

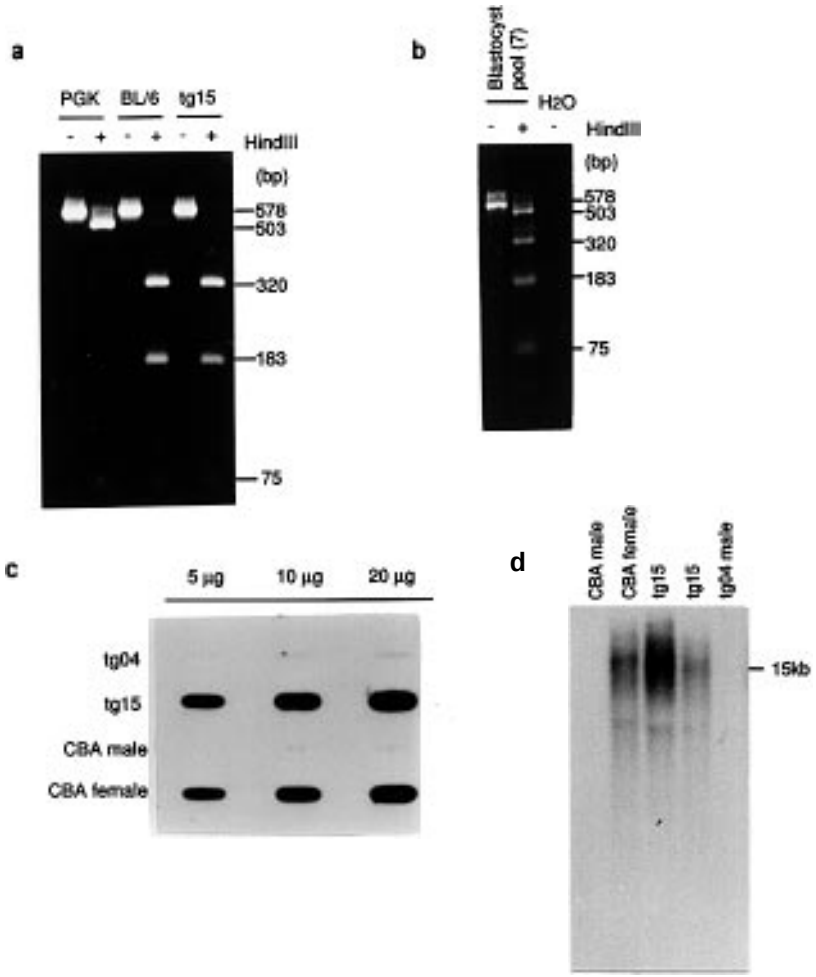

Figure 3. Analysis of *Xist* expression from the tg15 transgene. (**a**) Allele-specific RT-PCR analysis. Total RNAs were extracted from adult liver of a PGK female, a C57BL/6 female mouse, and a tg15 male arising from crosses of the tg15 line to PGK females. *Xist* RNA sequences spanning exon 3 and exon 6 were amplified by RT-PCR using primers of MX23b and MIX20. PCR products were digested with *Hin*dIII and analyzed by VisiGel Separation Matrix. PCR products derived from the PGK X chromosome generate a 503 bp band, whereas PCR products derived from a BL/6 X chromosome generate a 320 bp band plus a 183 bp band. The pattern of tg15 crossed into PGK strain was identical to that of C57BL/6, clearly indicating that in tg15 *Xist* RNA is transcribed only from the tg15 transgene on the Y chromosome. (**b**) Allele-specific RT-PCR of pooled blastocysts. Allele-specific RT-PCR of *Xist* expression in a pool of seven blastocysts from tg15/PGK male crossed to a PGK female. *Xist* expression from a tg15 transgene and a PGK X chromosome was detected at a 3.5 d.p.c. stage. (**c**) RNA slot blot. Total RNAs from adult liver of tg04 male, tg15 male, CBA male and CBA female were blotted onto the filter, and hybridized to the *Xist* cDNA probe (W7D, see Materials and Methods). To assess RNA loading, the filter was stripped and hybridized to a mouse β-actin cDNA probe. The intensity of signals was measured by a PhosphorImager. The relative value of *Xist* expression from the tg15 transgene was estimated as 0.9 when compared to that of a normal female mouse. (**d**) Northern analysis of *Xist* transcript from tg15 males. 5–10 µg of poly(A)$^+$ RNA was probed with the mXist1 *Xist* cDNA. tg15 gives an identical major transcript to CBA female. Two separate but different loadings of tg15 RNA were applied to the gel.

chromosome respectively. Methylation sites in the 5′ promoter region of *Xist* as well as a methylation site some distance 3′ of *Xist* were examined in both transgenic lines. For the autosomal transgene, both 5′ and 3′ sites were hypermethylated reflecting their usual status on the active X chromosome where *Xist* is not expressed. For the Y-linked transgene, 5′ sites showed hypomethylation to a degree consistent with one copy being substantially undermethylated. Indeed, for two of the 5′ methylation sites we demonstrated that the hypomethylation appeared to be largely confined to one copy of the transgenic array consistent with the notion that only one copy of *Xist* within the Y-linked transgene is expressed and it is this copy that has undergone hypomethylation with the other copies remaining hypermethylated. However, although the data is consistent with one copy of the array being expressed, it is not possible to determine whether the same copy is expressed in all cells. The possibility of mosaicism remains. Nevertheless, the methylation status of the 3′ methylation site has clearly become decoupled from that of the 5′ site since the 3′ site appears to be hypomethylated in the bulk of the copies of the Y-linked transgene (Fig. 5).

These results suggest a model for the mechanisms surrounding *Xist* expression from the Y-linked transgene. We propose that the cell recognises and counts the Y-linked transgene as a separate but single X-inactivation centre. Only one copy of the Y-linked transgenic array appears to be expressed. We propose that initially one copy is activated accompanied by 5′ hypomethylation. It may be that additional *Xist* copies in a single tightly-linked array are not recognised as separate centres. Alternatively, and maybe

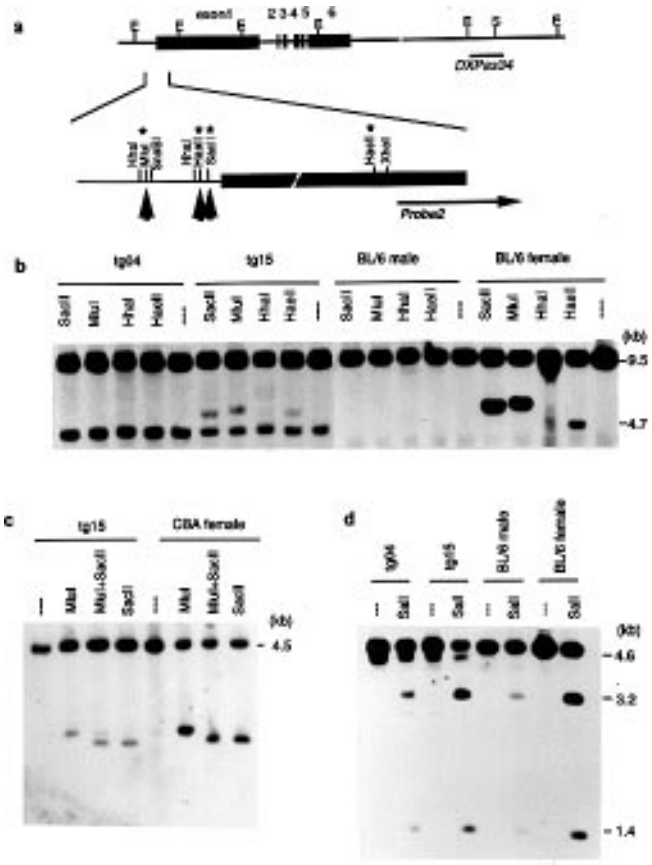

Figure 4. (**a**) A map of the 5′ end, the transcription unit and the 3′ region of the *Xist* locus showing the position of the methylation-sensitive restriction sites. The locations of probes are shown under the map. *Eco*RI and *Sal*I sites are denoted as E and S respectively. Sites that are hypomethylated on the inactive X chromosome are indicated by asterisks. Sites hypomethylated in tg15 are indicated by arrowheads. (**b**) Methylation analysis of 5′ end of the *Xist* gene. Genomic DNAs in agarose blocks from tg04 male crossed to a PGK female mouse, tg15 male crossed to a PGK female, C57BL/6 male and C57BL/6 female mouse were digested with *Pvu*II (broken line) and additionally with *Sac*II, *Mlu*I, *Hha*I and *Hae*II, separated on a 1% agarose gel, blotted and hybridized to Probe 2 [see (a) and Materials and Methods]. *Sac*II, *Mlu*I and *Hae*II sites are hypomethylated on the inactive X chromosome, and fully hypermethylated on the active X chromosome (11). A 4.7 kb *Xist* allele on the PGK X chromosome is not cut with the methylation-sensitive restriction enzymes. However, the 9.5 kb *Xist* tg15 allele on the Y chromosome was clearly cut with these enzymes, showing a similar pattern to the normal female mouse. Densitometry using a PhosphorImager was used to estimate the degree of hypomethylation occurring in the transgenic array (see text for full discussion). (**c**) Hypomethylated sites in tg15 reside in the same copy. *Eco*RI digests of tg15 male and CBA female sequentially digested with *Mlu*I, *Sac*II and *Mlu*I and *Sac*II together followed by hybridisation to probe 2 (see Materials and Methods). (**d**) Methylation analysis of the *Sal*I site 15 kb distal to the *Xist* locus. Genomic DNAs were digested with *Eco*RI and re-digested with *Sal*I, blotted and hybridized to a *DXPas3′* probe. In the male C57BL/6 mouse the predominant band is 4.6 kb, showing marked hypermethylation of the *Sal*I site on the active X chromosome. In the female C57BL/6 mouse, a 4.6 kb band is of approximately the same intensity as the total of a 3.2 kb and a 1.4 kb band, indicating the hypomethylation of this site on the inactive X chromosome. As the *Eco*RI fragment from a PGK X chromosome was slightly smaller than the 4.6 kb fragment of a C57BL/6 X chromosome, methylation status of PGK and C57BL/6 alleles can be distinguished. tg04 male crossed to a PGK female showed the typical male pattern with little cutting of the 4.6 kb fragment from the transgene. However, for tg15 male crossed to a PGK female there is noticeable hypomethylation at this site which reflects the pattern seen in female C57BL/6.

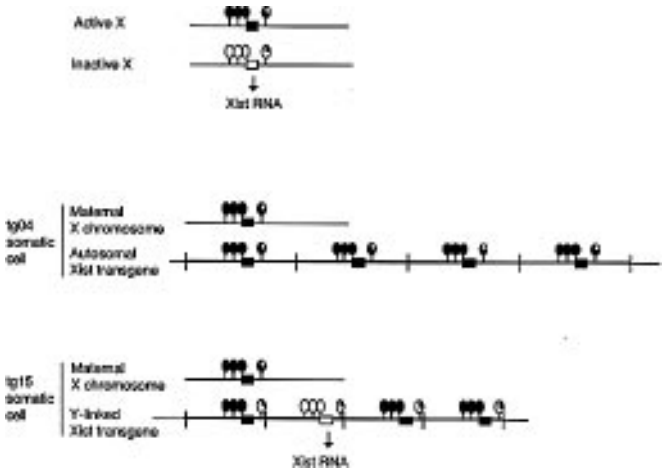

Figure 5. Summary of methylation status and RNA transcription of *Xist* transgenes. Methylated sites are indicated by closed circles and unmethylated sites are indicated by open circles. In tg04 all copies of transgene were hypermethylated at all sites and no *Xist* RNA was transcribed. In tg15 *Xist* was expressed and only one of the multiple copies of the transgene appears to be hypomethylated 5′. In contrast, most copies appear to be hypomethylated at the 3′ *Sal*I site.

more likely, *Xist* expression from one copy may inactivate adjacent copies of the *Xist* gene and prevent their expression, the adjacent copies becoming hypermethylated at the 5′ sites. We observed that hypomethylation at the 3′ site is decoupled from the methylation status of sequences 5′ of *Xist*. Generalised hypomethylation of the 3′ site may be required for *Xist* expression. Alternatively, methylation of the 3′ site may be downstream of the major events that initiate *Xist* expression. It may be that 3′ hypomethylation represents a generalised effect of *Xist* expression on neighbouring sequences resulting in a 3′ hypomethylated site in most copies of the Y-linked transgenic array.

There are two possible but distinct explanations for the difference in the expression patterns of the autosomal and Y-linked transgenes. Firstly, it is possible that the autosomal transgene has inserted into a critical autosomal region. Expression of *Xist* and consequent X-inactivation of surrounding sequences may result in a lethal autosomal monosomy accompanied by cell selection against cells undergoing X-inactivation from the transgenic locus. In the case of the Y-linked transgene, which has integrated towards the telomere of the long arm in a region that is heterochromatic and devoid of genes, expression of *Xist* may have little deleterious effect. X-inactivation may have difficulty spreading through the long arm heterochromatic region and indeed, consistent with this, we have seen no evidence of sex-reversal in the tg15 transgenic line indicating that the short arm *Tdy* gene at least is unaffected by *Xist* expression from the long arm transgene. If for the Y-linked transgene a counting mechanism is in operation, cells expressing *Xist* from the normal X chromosome are likely to be selected against and consistent with this we have not been able to detect *Xist* expression from the normal X chromosome later in development. If counting is occurring at the Y transgenic locus, it might be expected to be occurring also at the autosomal tg04 transgenic locus. However, this conclusion would be inconsistent with the notion that expression from the autosomal transgene has been selected against due to lethal autosomal monosomy. The remaining cells

in the tg04 line would have to express *Xist* from the single X chromosome, presumably also a lethal event. Evidence has been presented that suggests that the efficiency of generation of autosomal *Xist* YAC transgenics does not differ from that observed for other YACs (28). Heard *et al.* (28) conclude that the lack of *Xist* expression observed from the autosomal YAC transgenes along with normal expression of adjacent genes would suggest that these autosomal YAC transgenes are not silenced by position effects. Thus, there is no evidence for selection against the generation of autosomal *Xist* YAC transgenes expressing *Xist*. If this is the case, then it may be that we need to consider a second explanation for lack of expression from autosomal transgenes. Expression of the Y-linked transgene may be facilitated in some way by its position in relation to the surrounding sequence environment—there may be potential interactions between the heterochromatic environment of the Y chromosome long arm and the initiation process of *Xist* expression and X-inactivation. It would be interesting to assess *Xist* expression from both autosomal and Y-linked YACs carrying varying amounts of 5′ and 3′ sequence from the *Xist* region. However, given the impracticality of directing YAC integration to the Y chromosome, no answers are likely to be forthcoming on the role of flanking sequences in Y-linked expression.

In conclusion, it would appear that the Y-linked transgenic mouse line carries features associated with normal *Xist* expression in female cells including the phenomenon of methylation. Given recent evidence (8) demonstrating the importance of *Xist* in the process of X-inactivation, this mouse line represents an important model for the further investigation of the processes leading to X-inactivation. It should be possible to use this mouse line to investigate a number of the features of X chromosome inactivation subsequent to *Xist* expression, including spreading and heterochromatisation.

MATERIALS AND METHODS

Mice

B6CBAF1, FVB/N, CBA and C57BL/6 mice were obtained from Charles River and B&K Universal (UK). PGK mice were a gift from N. Brockdorff.

Retrofitting YAC yXist1

A 350 kb *Xist* YAC, yXist1, was derived from a *rad52* YAC library from female C57BL/10 mice (16), and described previously (7). Two YAC targetting vectors, yRP17his3SceI and pLUSSceIβgeo, were constructed as follows.

yRP17his3SceI vector. The 1.7 kb *Bam*HI his3 fragment was ligated into the *Bgl*II site of yRP17 vector (17) creating yRP17his3 (courtesy of M. Rubock, GenPharm International). A synthetic i-*Sce*I linker formed using two complementary oligos: 5′-TAAGGTAGGGATAACAGGGTAATCC-3′ and 5′-TTAGGATTACCCTGTTATCCCTACC-3′ was introduced into the *Bsu*36I site of yRP17his3 vector. The 3.1 kb *Aat*II-*Eco*RI fragment of the yRP17his3SceI vector was targetted to the TRP1-ARS arm of yXist1 by the spheroplasting procedure. His^+ transformants were selected and screened by PCR and PFGE analysis.

pLUSSceIβgeo vector. An i-*Sce*I linker with *Hin*dIII site on one side and *Sac*I site on the other was formed with two synthetic oligos: 5′-AGCTTTAGGGATAACAGGGTAATGAGCT-3′ and 5′-CATTACCCTGTTATCCCTAA-3′, and inserted into *Hin*dIII-*Sac*I sites of pLUS vector (21) to give pLUSSceI. A 4.7 kb βgeo cassette was released from pPGKβgeopbA (18) by digestion of *Hin*dIII, following introduction of a *Not*I-*Hin*dIII adaptor to the unique *Not*I site of the plasmid, and cloned into *Hin*dIII-cut pLUSSceI vector. The resulting plasmid, pLUSSceIβgeo, was linearised with *Sal*I and targetted to the right arm of yXist1 YAC. Lys+ transformants were isolated and screened.

Production of transgenic mice

YAC DNA was purified as described (19). In brief, agarose blocks of yeast strains were prepared at a final concentration of 4×10^9 cells/ml. YAC DNA in agarose blocks was separated on a preparative CHEF gel (LKB Pulsaphor Electrophoresis Unit) containing 1% SeaPlaque GTG Agarose (FMC BioProducts) in 0.5 × TBE. The gel was run at 170 V for 36 h, with a field switching time of 30 s. A slice containing the YAC was excised from the unstained central section of the gel, equilibrated in agarase buffer (10 mM Bis Tris-HCl pH 6.5, 0.2 mM EDTA pH 8.0, 100 mM NaCl) and digested with β-Agarase I (New England Biolabs) at 40°C for 2 h. Undigested agarose was removed by centrifugation at 12 000 *g* for 30 min at room temperature. The supernatant was transferred to a floating 0.05 μm dialysis filter (Millipore) against injection buffer (10 mM Tris-HCl pH 7.5, 0.2 mM EDTA pH 8.0, 100 mM NaCl) overnight. The isolated YAC DNA was at a concentration of approximately 0.2 ng/μl. The integrity and purity of the YAC DNA was tested by submarine PFGE prior to microinjection. 2pl (approximately one YAC copy) was microinjected into the pronuclei of B6CBAF1 or FVB/N zygotes as described (20). Injected oocytes were transferred to the oviducts of pseudopregnant B6CBAF1 females and allowed to develop to term.

PCR and Southern blot analysis

Genomic DNA was extracted from mouse tail as described (20), and PCR was used for screening of founder mice. In order to design PCR primers, end clones from yXist1 YAC were rescued using the pLUS and pICL plasmids (21), and nucleotide sequences of YAC insert-vector junction were determined. YAC left insert-vector junction region was amplified using primer set 2: 5′-CTTGCGGGATATCGTCCATT-3′ and 5′-GCTTGCATGCATACACACAT-3′ (434 bp). Right junction region was amplified using primer set 3: 5′-AGGGTCTCCTGTCACGAAAC-3′ and 5′-ACTGGGTTGAAGGCTCTCAA-3′ (515 bp).

Founder mice positive for PCR primer set 2 or 3 were further analyzed by left arm PCR primer set 1: 5′-CAGGGTTATTGTCTCATGAGCGGAT-3′ and 5′-TTAGGATTACCCTGTTATCCCTACC-3′ (450 bp) and right arm PCR primer set 4: 5′-CTTGAGATCGGGCGTTCGACTCGC-3′ and 5′-TGAACGGTGATCCCCACCGGAATTG-3′ (1855 bp).

Location of primers in yXist1 are indicated in Figure 1.

For Southern blot analysis, genomic DNA was prepared from mouse spleen cells embedded in agarose blocks at a concentration of 10^7 cells/ml as described (19). DNA in agarose blocks were digested with *Eco*RI, i-*Sce*I, *Sal*I or *Eag*I according to the instructions of the manufacturer, equilibrated in 10 mM Tris–HCl pH 7.5, 1 mM EDTA pH 8.0, electrophoresed on 1% agarose gels,

and blotted to nylon membranes (Hybond-N^+, Amersham) by alkaline transfer. Probes were radioactively labelled with [α-^{32}P]dCTP by the random hexamer priming (22). Hybridization and autoradiography were carried out as described (23). Radioactivity in the bands was quantified using a PhosphorImager (Molecular Dynamics). W7D probe is a 2.1 kb *Eco*RI fragment of an *Xist* cDNA clone (7). pHH1.5 probe is a 1.5 kb *Hin*dIII fragment in the *Xist* promoter region. pE2.3 (1) probe is a 1.0 kb *Eco*RI fragment located 25–30 kb upstream from the *Xist* start site, and pE55 (16) probe is a 1.6 kb *Eco*RI fragment located approximately 60 kb upstream from the *Xist* start site. Probe 2 is a 1.5 kb *Eco*RI-*Xho*I fragment isolated from a 4.5 kb *Eco*RI subclone pGPT2 which contains the start site. A 660 bp 3′ Cdx4 probe was generated by PCR using primers of 5′-TCTGGTTTC-AGAATCGCAGA-3′ and 5′-GTAATCACCTCCTGATGC-TG-3′ (24). DXPas34 is a 1.6 kb *Eco*RI fragment containing a *Sal*I site 15 kb downstream of the *Xist* gene (12).

Fluorescent *in situ* hybridization

Probe labelling, suppression of repetitive sequences, denaturation, hybridization, and fluorescence detection were carried out as described (25). Gel-purified yXist1 DNA was labelled with biotin-14-dATP by nick translation (BioNick Labelling System, GIBCO BRL). Avidin-FITC labeled (SIGMA) was used to detect biotin probe. Nuclei were counterstained in 1 μg/ml propidium iodide (SIGMA) in Vectashield mounting medium (Vector Laboratories). Interphase nuclei were revealed by fluorescence microscopy using a Leiz Aristoplan microscope.

RNA analysis

Total cellular RNA was extracted from adult mouse liver by the AGPC method (26). Expression of the transgene was analyzed by RNA slot blot. Twenty μg of total RNA was denatured in formaldehyde/formamide, applied to a nylon membrane (Hybond-N^+, Amersham) with mild vacuum and hybridized with the W7D probe as for Southern hybridisations (described above). To assess RNA loading, the blot was stripped and reprobed with mouse β-actin cDNA probe (27).

Poly(A)$^+$ RNA was purified for adult liver using Oligotex-dT mRNA kits (Qiagen). Approximately 5–10 μg of poly(A)$^+$ RNA was denatured in formaldehyde/formamide, electrophoresed in 1% agarose containing formaldehyde in 1×MOPS buffer and transferred to a nylon membrane (Hybond N^+, Amersham) followed by hybridisation to the mXist1 probe (7) as for Southern hybridisations (see above). The blot was stripped and reprobed with mouse β-actin cDNA.

Allele specific RT-PCR was performed as described (9). Ten μg of total RNA was transcribed using 1 000 U M-MLV reverse transcriptase (GIBCO BRL) with random hexamer primers. *Xist* cDNA was amplified using primers MX23b and MIX20 (9), which span exon 3 and exon 6 (578 bp). PCR amplifications of *Xist* cDNA were performed as follows: one fifteenth of cDNA generated from reverse transcription was mixed with 200 mM of each dNTP, 10 pmol of each primer, 2.5 U of *Taq* DNA polymerase in 20 μl PCR buffer. The samples were incubated at 95°C, 59°C, and 72°C for 30 s each in a DNA Thermal Cycler (Perkin-Elmer Cetus) for a total of 30 cycles. PCR products were digested with an excess of *Hin*dIII and separated by VisiGel Matrix electrophoresis. The gel was stained with ethidium bromide and photographed.

Allele-specific RT-PCR of pooled blastocysts

Blastocysts were collected by flushing oviducts of female mice 3.5 d.p.c. Blastocysts were washed with PBS without Ca^{2+} and Mg^{2+} and lysed in 10 μl of water containing 0.01% diethyl pyrocarbonate. RNA was isolated from a pool of seven blastocysts as described (9). First strand cDNA was prepared with random hexamer primers as described above . First round PCR was carried out with primers MIX10 and MX20 in a total volume of 20 μl. Then nested second round PCR was performed with 1 μl of the first PCR product and primers MIX20 and MX23b in a total volume of 20 μl. A half of the nested second round PCR products were digested with *Hin*dIII and fractionated on VisiGel matrix.

ACKNOWLEDGEMENTS

S.M. was supported by a Human Frontier Science Fellowship. This work was also partly supported by grant G9201373 from the Medical Research Council, UK. We thank Phil Avner and Neil Brockdorff for generously providing DNA probes and for discussion.

REFERENCES

1. Brown, C. J. *et al.* (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**, 38–44.
2. Borsani, G. *et al.* (1991) Characterization of a murine gene expressed from the inactive X chromosome *Nature* **351**, 325–329.
3. Brockdorff, N. *et al.* (1991) Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **351**, 329–331.
4. Cattanach, B. M. *et al.* (1991) X-chromosome deletion spanning the tabby (*Ta*) and testicular ferminization (*Tfm*) loci in the mouse. *Cytogenet. Cell Genet.* **56**, 137–143.
5. Lafreniere, R. G. *et al.* (1993) 2.6 Mb YAC contig of the human X inactivation center region in Xq13: physical linkage of the RPS4X, PHKA1, XIST and DXS128E genes. *Hum. Mol. Genet.* **2**, 1105–1115.
6. Rastan, S. and Brown, S. D. M. (1990) The search for the mouse X-chromosome inactivation centre. *Genet. Res., Camb.* **56**, 99–106.
7. Brockdorff, N. *et al.* (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.
8. 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.
9. Kay, G. F. *et al.* (1993) Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell* **72**, 171–182.
10. Norris, D. P. *et al.* (1994) Evidence that random and imprinted Xist expression is controlled by preemptive methylation. *Cell* **77**, 41–51.
11. Kay, G. F. *et al.* (1994) Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. *Cell* **77**, 639–650.
12. Courtier, B., Heard, E. and Avner, P. Xce haplotypes show modified methylation in a region of the active X chromosome lying 3(to Xist. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 3531–3535.
13. Cattanach, B. M. *et al.* (1989) Further *Xce* linkage data. *Mouse News Lett.* **83**, 165.
14. Simmler, M. C. *et al.* (1993) Mapping the murine Xce locus with (CA)n repeats. *Mamm. Genome* **4**, 523–530.
15. Horn, J. M. and Ashworth, A. (1995) A member of the caudal family of homeobox genes maps to the X-inactivation centre region of the mouse and human X chromosome. *Hum. Mol. Genet.* **4**, 1041–1047.
16. Chartier, F. L. *et al.* (1992) Construction of a mouse yeast artificial chromosome library in a recombination-deficient strain of yeast. *Nature Genet.* **1**, 132–136.
17. Mann, C. and Davis, R. W. (1986) Structure and sequence of the centromeric DNA of chromosome 4 in Saccharomyces cerevisiae. *Mol. Cell Biol.* **6**, 241–245.

18. Friedrich, G. and Soriano, P. (1991) Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes and Dev.* **5**, 1513–1523.
19. Gnirke, A. *et al*. (1993) Microinjection of intact 200- and 500-kb fragments of YAC DNA into mammalian cells. *Genomics* **15**, 659–667.
20. Hogan, B., Costantini, F. and Lacy, E. (1986) *Manipulating the mouse embryo: A Laboratory Manual*, Cold Spring Harbor Lab., Cold Spring Harbor, NY, USA.
21. Hermanson, G. G. *et al*. (1991) Rescue of end fragments of yeast artificial chromosomes by homologous recombination in yeast. *Nucleic Acids Res.* **19**, 4943–4948.
22. Feinberg, A. P. and Vogelstein, B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6–13.
23. Church, G. M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl Acad. Sci. USA* **81**, 1991–1995.
24. Gamer, L.W. and Wright, C.V.E. (1993) Murine *Cdx4* bears a striking similarities to Drosophila caudal gene in its homeodomain sequence and early expression pattern. *Mech. Dev.* **43**, 71–81.
25. Lichter, P. *et al.* (1990) High resolution mapping of human chromosome 11 by in situ hybridisation with cosmid clones. *Science* **247**, 64–69.
26. Chomozynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156–159.
27. Minty, A. J. *et al*. (1981) Mouse actin messenger RNAs. Construction and characterization of a recombinant plasmid molecule containing a complementary DNA transcript of mouse alpha-actin mRNA. *J. Biol. Chem.* **256**, 1008–1014.
28. Heard, E., Kress, C., Mongelard, F., Courtier, B., Rougeulle, C., Ashworth, A., Vourc'h, C., Babinet, C. and Avner, P. (1996) Transgenic mice carrying an *Xist*-containing YAC. *Hum. Mol. Genet.* **5**, 441–450.