MOLECULAR AND CELLULAR BIOLOGY, Apr. 1999, p. 3156–3166 Vol. 19, No. 4
0270-7306/99/$04.00+0

Xist Yeast Artificial Chromosome Transgenes Function as X-Inactivation Centers Only in Multicopy Arrays and Not as Single Copies

EDITH HEARD,[1]* FABIEN MONGELARD,[2] DANIELLE ARNAUD,[1] AND PHILIP AVNER[1]

Unité de Génétique Moléculaire Murine, CNRS URA 1968, Institut Pasteur, F-75724 Paris Cedex 15,[1] and Equipe Dyogen, Institut Albert Bonniot, Faculté de Médecine, 38706 La Tronche Cedex,[2] France

Received 20 July 1998/Returned for modification 2 September 1998/Accepted 17 December 1998

X-chromosome inactivation in female mammals is controlled by the X-inactivation center (Xic). This locus is required for inactivation in *cis* and is thought to be involved in the counting process which ensures that only a single X chromosome remains active per diploid cell. The *Xist* gene maps to the Xic region and has been shown to be essential for inactivation in *cis*. Transgenesis represents a stringent test for defining the minimal region that can carry out the functions attributed to the Xic. Although YAC and cosmid *Xist*-containing transgenes have previously been reported to be capable of *cis* inactivation and counting, the transgenes were all present as multicopy arrays and it was unclear to what extent individual copies are functional. Using two different yeast artificial chromosomes (YACs), we have found that single-copy transgenes, unlike multicopy arrays, can induce neither inactivation in *cis* nor counting. These results demonstrate that despite their large size and the presence of *Xist*, the YACs that we have tested lack sequences critical for autonomous function with respect to X inactivation.

Mammalian X-chromosome inactivation is the process by which one of the two X chromosomes in female cells is transcriptionally silenced, resulting in dosage compensation for X-linked gene products between the sexes (31). This process takes place during early female development and is closely associated with cellular differentiation during early embryogenesis. Similarly, in vitro differentiation of female embryonic stem (ES) cells results in inactivation of one of the two X chromosomes (43).

Classical genetic studies have defined a critical region of the X chromosome, the X-inactivation center (Xic), which controls the initiation of X inactivation (see reference 21 for a review). The Xic not only is required for the spread in *cis* of inactivation but also appears to be involved in a counting process that senses the number of X chromosomes in a cell and ensures that only a single X chromosome remains active per diploid cell, all supplementary X chromosomes being inactivated (see reference 32 for a review).

The *Xist* gene, which lies within the critical Xic region, is expressed exclusively from the inactive X chromosome (4, 6, 9), producing a long, processed transcript which does not appear to be protein coding (7, 10). The *Xist* RNA appears to associate with or "coat" the inactive X chromosome (10, 14), leading to suggestions that it may play a chromatin remodeling role in the initiation and propagation of X inactivation (53). A role for *Xist* in the initiation of X inactivation is supported by its developmental expression pattern (26). The onset of X inactivation is preceded by a marked increase in *Xist* RNA levels in early embryos and differentiating ES cells, and recent studies have shown that this up-regulation is achieved via stabilization of *Xist* transcripts, rather than higher transcription rates (39, 48). Formal proof of *Xist*'s involvement in X inactivation has been provided by two independent targeted deletions of *Xist* sequences (35, 40). Both studies demonstrated that *Xist* expression is essential for inactivation in *cis*, thus supporting the idea that the *Xist* RNA is the propagator of X inactivation. In both cases, however, an undeleted X chromosome could still be inactivated in *Xist* mutant cells, implying that counting had not been disrupted. Thus, counting and *cis* inactivation appear to be controlled by genetically separable elements.

The presence of sequences involved in counting outside the *Xist* gene has recently been suggested by the deletion of a 65-kb region lying 3′ to *Xist* exon 6, where the X chromosome bearing the deletion undergoes inactivation even in the absence of a second X chromosome (15). Furthermore, the genetically defined X-chromosome controlling element (*Xce*), which affects the random nature of X inactivation in the mouse and thus appears to be involved in the choice of X chromosome that is inactivated (13), has been found to be closely linked to but genetically distinct from *Xist*, lying 3′ to it (49). Elements potentially affecting choice have also been reported to lie within the *Xist* gene (34) and 5′ to it (41).

Taken together, the above genetic and knockout data support the idea that several distinct elements, distributed across the region containing *Xist*, may be required to regulate the initiation of X inactivation. Although targeted deletions can establish whether a particular sequence is necessary for X inactivation, such an approach cannot easily address the question of what the extent of the sequences necessary and sufficient to carry out X inactivation might be. A stringent test for determining whether a given region is sufficient to bring about correctly regulated X inactivation is to examine its behavior at ectopic loci in transgenic experiments. A number of such experiments have now been reported (see reference 19 for a review). In one case, ES cell lines carrying multiple copies of a 450-kb yeast artificial chromosome (YAC) including the *Xist* gene were found to show a number of functions attributed to the Xic upon differentiation: the transgenic *Xist* RNA could coat the autosome carrying the transgene; and inactivation in

* Corresponding author. Mailing address: Unité de Génétique Moléculaire Murine, CNRS URA 1968, Institut Pasteur, 25 rue du Docteur Roux, F-75724 Paris Cedex 15, France. Phone: 33 1 456 88653. Fax: 33 1 456 88656. E-mail: eheard@pasteur.fr.

cis was observed, as was inactivation of the single X chromosome (counting) (29, 30). In another study, Matsuura et al. (37) found ectopic *Xist* expression in transgenic mice containing four copies of a 240-kb *Xist* YAC integrated onto the Y chromosome, although inactivation was not demonstrated. More recently, multiple copies of a 35-kb cosmid transgene encompassing *Xist* were also found to be capable of some ectopic Xic function, although the variability in *Xist* expression levels observed between cell lines suggested that these short transgenes were subject to position effects and may therefore be missing elements for autonomous function (24).

In all of the above studies, however, the transgenic lines contained *Xist* transgenes present in multiple linked copies (3 to 24 copies [30]; 4 copies [37]; 2 to 8 copies [24]). The ability of a single-copy transgene encompassing *Xist* to show X-inactivation function has thus not been demonstrated so far, and it could not be ruled out in the above studies that complex interactions among linked YAC copies may have somehow enabled ectopic Xic function to be observed (52).

In a previous report we described several mouse lines carrying single-copy *Xist* YAC transgenes in which *Xist* was not expressed and there was no evidence of inactivation (22). In the study presented here, we wished to establish definitively whether single copies of *Xist*-containing YAC transgenes are capable of inducing inactivation in *cis* and/or counting in ES cells. ES cells and their in vitro-differentiated derivatives were used as the selective pressure against transgene function is likely to be less stringent in vitro than in vivo, where cell selection might occur as a result of autosomal *cis* inactivation (leading to monosomy) or counting (leading to X-chromosome nullisomy). We have generated an extensive series of transgenic cell lines by transferring two different YACs into male and female ES cells. We also derived ES cells from one of our previously described transgenic mouse lines (22), which carries a single intact copy of a 460-kb YAC, to see whether the absence of Xic function that we had previously observed was due to the single-copy nature of the transgene or to in vivo counterselection. Transgene activity was followed at the single-cell level using RNA fluorescent in situ hybridization (FISH) to detect expression both of *Xist* and of *Brx*, a gene susceptible to inactivation which is located 75 kb 3′ to *Xist* (51) and which is expressed from the active X chromosome in both differentiated and undifferentiated ES cells. The degree of mosaicism associated with transgene function in differentiating ES cell populations could thus be assessed.

We report that, unlike multicopy arrays, large, single-copy *Xist* YAC transgenes lead to neither inactivation in *cis* nor counting, despite the presence of all of the sequences defined to date by deletion as being essential for the correct regulation of X inactivation. Our results indicate that certain elements necessary for the initiation of X inactivation at ectopic sites must be lacking from the *Xist* YAC transgenes tested, and that their lack is compensated for by the presence of multiple copies. A number of hypotheses to explain the differences in behavior between single and multicopy *Xist* transgenes are presented, and the implications of our findings for autonomous Xic function are discussed.

MATERIALS AND METHODS

YAC manipulation. The two vector arms of YAC PA-2 were retrofitted with I-*Ppo*I sites, and the right YAC arm was retrofitted with the *pgk1* promoter-neomycin resistance gene (*pgk-neo*) cassette as described elsewhere (18). YAC PA-3 F1n (20) was retrofitted by using the pLNA-1 vector, which carries the *pgk1* promoter-*neo* cassette (45). This vector could integrate into the ampicillin resistance gene present on either of the two YAC vector arms.

Generation of transgenic ES cell lines. High-molecular-weight yeast DNA was prepared in agarose plugs and separated by pulsed-field gel electrophoresis, and the YAC DNA was excised by previously described procedures (22). YAC DNA was lipofected into CK35 ES cells (11) by using the protocol described in reference 28. Transfer of YAC PA-2 by spheroplast fusion into the XX HP310 ES cell line (2a) was performed as previously described (18). In both cases, Neo^r clones were selected 24 h after YAC transfer by treatment with G418 (0.425 mg/ml for HP310-derived F1 clones and 0.25 mg/ml for CK35-derived L1 and L4 clones). ES cell lines were derived from transgenic mouse line 53 (22) by using standard procedures (46) that involved crosses between transgenic females and transgenic males, both of which were on an essentially (>98%) 129 genetic background.

Characterization of transgenic ES cell lines. Southern analysis of *Eco*RI- or *Hin*dIII-digested DNA was carried out by using standard procedures. Transgene copy number was quantitated by PhosphorImager (Molecular Dynamics) analysis of blots hybridized with *Xist* probes generated by PCR (e.g., nucleotides 5379 to 5547) and a 1.1-kb fragment in exon 1 (22), as well as the X-chromosome probe *Xpct* (128E2 probe [16]), which does not lie within the YACs. Other probes used for mapping (e.g., *IVA5*, *DXPas34*, *DXPas19*, and *DXPas27*) have been described previously (20, 23). The *Tsx* probe (50) was generated by PCR using primers GAG ATG TTC ACT GAT CAG AA and CTC AGT GGC TCA TCT GC. Pulsed-field gel analysis of ES cell DNA in agarose blocks digested with rare-cutter enzymes such as *Sal*I was performed by using standard procedures (23). I-*Ppo*I digestion was performed as specified by the manufacturer (Promega).

Culture and differentiation of ES cells. CK35 and HP310 ES cells and transgenic derivatives were maintained in the undifferentiated state by culturing in ES medium (high-glucose Dulbecco modified Eagle medium [DMEM], 2 mM glutaMAX I [GIBCO] 0.1 mM 2-mercaptoethanol, 15% fetal calf serum and 10^3 U of recombinant leukemia inhibitory factor [GIBCO] per ml, 0.05 mg of streptomycin per ml, 50 U of penicillin per ml) on mitomycin C-treated mouse fibroblast feeder cells and appropriate G418 selection for transgenic cells. For differentiation into embryoid bodies (EBs), feeders were first removed by successive adsorptions on gelatinized dishes and then cultivated for 3 days under adherent conditions in ES cell medium (46). Aggregates were formed following mild trypsinization and transferred to suspension culture (day 0) in EB medium (DMEM, 10% newborn calf serum, 10^{-4} M 2-mercaptoethanol), without G418 selection. Half of the EB medium was changed every day. Four day EBs were attached onto LabTek chamber slides (Nunc) for monolayer outgrowth for 3 to 10 days in DMEM–10% fetal calf serum.

DNA and RNA FISH analysis. Metaphase spreads were obtained from ES cells, and FISH was performed as described elsewhere (14). Interphase nuclei were prepared as described previously (30). Briefly, cells grown on chamber slides or cytospun (4 min, 400 rpm) onto baked glass slides were permeabilized with Triton X-100 in ice-cold cytoskeletal buffer for 7 min, fixed with 4% paraformaldehyde for 10 min on ice, and then stored in 70% ethanol. EBs were taken from culture at days 1, 2, 3, 4, and 6 of differentiation and disaggregated by pipetting in phosphate-buffered saline prior to cytospinning. For RNA FISH, slides were dehydrated and used directly for hybridization. For DNA FISH, nuclei were denatured for 2 min in 70% formamide, 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 75°C; for simultaneous RNA-DNA FISH, nuclei were denatured for just 1.5 min. Hybridization and washing conditions were as described previously (14). Nuclei were mounted in Vectashield (Vector) and counterstained with 4′,6-diamidino-2-phenylindole (DAPI). A Quantix change-coupled device camera and IPLab and Photoshop software were used for image acquisition and treatment. All probes were labeled by nick translation (Vysis) with spectrum green-dUTP or spectrum red-dUTP (Vysis). The *Xist* probe was either plasmid p*Xist*3K (30) or lambda 510 (47), which covers most of the *Xist* gene. The *Brx* probe consisted of the partially overlapping lambda clones IIIC2 and IG10 (47). The YAC vector probe was pYAC4 vector DNA, and the bacterial artificial chromosome (BAC) X probe was a BAC isolated by using the microsatellite marker *DXMit158*, which lies outside the X-chromosome region covered by YACs PA-2 and PA-3 F1n.

RESULTS

Generation of transgenic ES cell lines with YACs PA-2 (460 kb) and PA-3 F1n (320 kb). To study the capacity of different regions containing *Xist* to induce the X-inactivation process, we used two YACs, PA-2 (460 kb) and PA-3 F1n (320 kb) (Fig. 1), both of which have previously been mapped in detail (20, 23). Both YACs carry inserts derived from the C3H/He mouse strain. YAC PA-2 contains sequences 130 kb 5′ and 310 kb 3′ to *Xist*, while YAC PA-3 F1n contains sequences 170 kb 5′ and 100 kb 3′ to *Xist*. Using homologous recombination in yeast, both YACs were retrofitted with the *pgk1* promoter-*neo* selectable marker in one vector arm (see Materials and Methods).

Both of these YACs were transferred into CK35 XY ES cells by lipofection of purified YAC DNA. Twenty clones were ob-

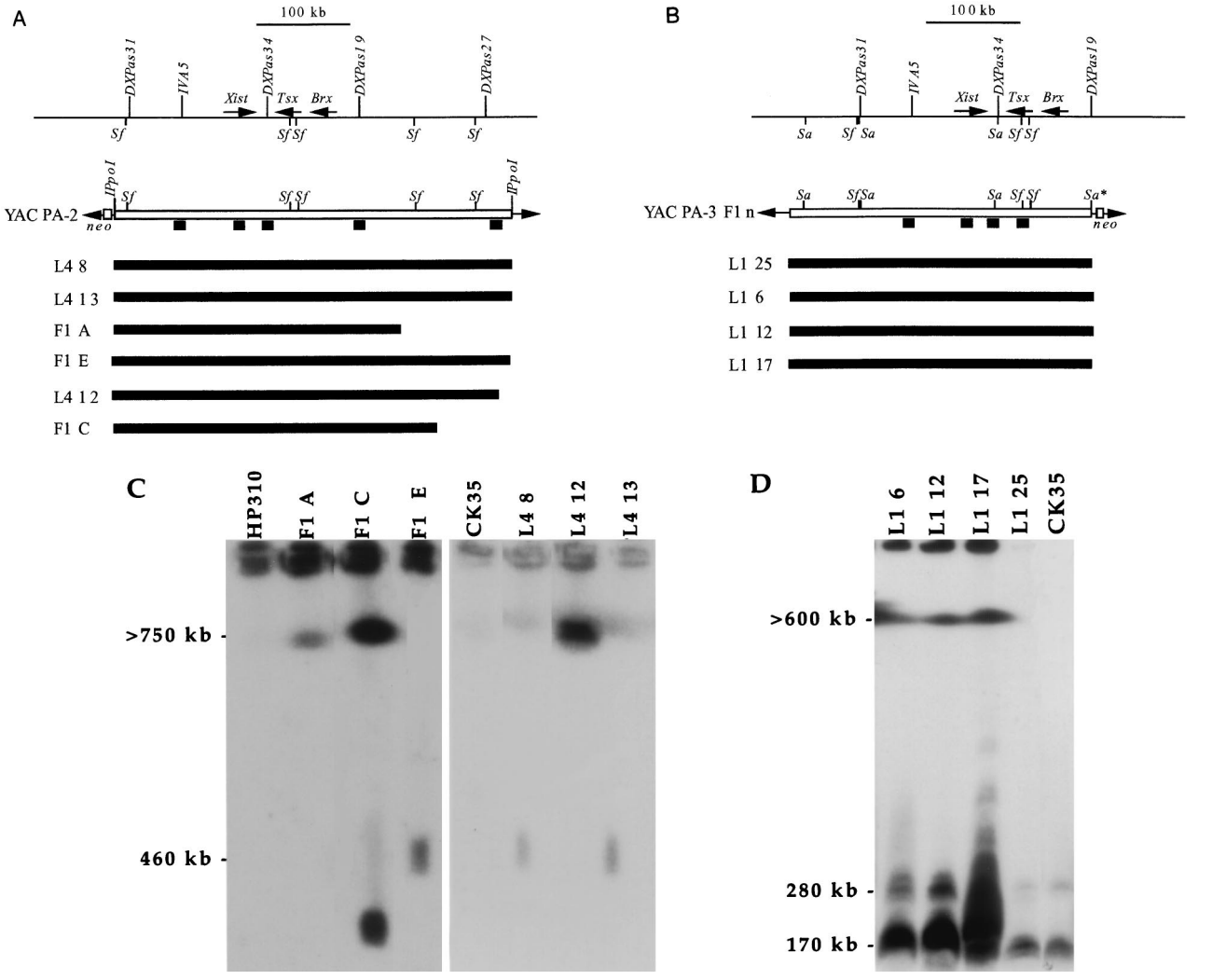

FIG. 1. (A and B) Structures of the two YACs used in this study, PA-2 (460 kb) and PA-3 F1n (320 kb). More detailed physical maps of these YACs have been published previously (20, 23). The DNA probes used in characterization of the YACs and transgenic clones are shown as black boxes below each YAC map along with the I-*Ppo*I, *Sal*I (*Sa*), and *Sfi*I (*Sf*) sites that were informative in the analysis of the transgenes. (A) Transgene structures of the L4 (lipofection-derived) and F1 (spheroplast fusion-derived) lines created with the 460-kb YAC PA-2. I-*Ppo*I meganuclease sites present at each extremity of this YAC enabled YAC integrity of the transgenic clones to be assessed (see panel C). *Sfi*I mapping (data not shown) provided further resolution concerning the extent of the region covered by the transgenes (particularly in line F1 A). Multicopy transgenic line F1 C showed, in addition to the major YAC fragment illustrated, a minority of smaller transgenic fragments. (B) The transgene structures of L1 (lipofection-derived) lines created with the 320-kb YAC PA-3 F1n. Both *Sal*I and *Sfi*I restriction sites were used to analyze the structures of these transgenes (see panel D for an example of the *Sfi*I analysis). A *Sal*I site within the YAC arm of PA-3 F1n which leads to the generation of a *Sal*I band detected by the *Tsx* and *DXPas34* probes of approximately the same size as that at the endogenous locus is indicated as *Sa**. It should be noted that there are other *Sal*I sites present in this region (23), but these are methylated and thus uncut by *Sal*I in mouse genomic DNA. Multicopy transgenic line L1 17 showed, in addition to the major YAC fragment illustrated, a minority of smaller transgenic fragments. (C) Example of pulsed-field gel analysis of I-*Ppo*I-digested DNA of cell lines carrying YAC PA-2 transgenes and the nontransgenic control lines CK35 and HP310. In lines L4 8, L4 13, and F1 E (single-copy transgenes), a 460-kb fragment detected by all probes tested lying within the YAC (*Xist* probe shown) following I-*Ppo*I digestion suggested that the YAC was intact. In lines F1 A (single copy) and L4 12 (two copy), larger (>750-kb) transgene fragments were detected, suggesting that at least one I-*Ppo*I site had been lost. In line F1 C (three to four copies), detection of a 360-kb band as well as a >750-kb band suggested that at least two copies of the truncated YAC, as shown in panel A, were present, with additional end fragment(s) (i.e., minus a second I-*Ppo*I site) being responsible for the higher-molecular-weight band. (D) Example of pulsed-field gel analysis of *Sfi*I-digested DNA of lines carrying YAC PA-3 F1n transgenes. The internal *Sfi*I bands detected by various probes (B) appear to be intact in lines L1 25 (single copy), L1 6 (two copies) and L1 12 (six to seven copies) and are detected as an increase in the intensity of the endogenous bands compared with the control cell line CK35 (170 and 280 kb with the *Xist* probe shown here). In L1 17 (five to eight copies), the majority of the *Sfi*I fragments appear to be intact, although a minority of smaller and larger fragments, corresponding to truncated fragments and partial digests of neighboring transgene copies, respectively, can also be seen.

tained with YAC PA-3 F1n (L1 series), of which four are described in this report, and eight clones were obtained with YAC PA-2 (L4 series), of which three are reported. A further six clones were obtained with YAC PA-2 via spheroplast fusion with HP310 XX ES cells (F1 series), three of which are described here.

Five single-copy (three XY and two XX) and five multicopy (copy number ranging from 2 to 7, XY and X0) lines generated in this way were analyzed in detail (Tables 1 and 2). Transgene copy number was assessed by Southern blot analysis of *Eco*RI- or *Hin*dIII-digested ES cell DNA, using hybridization of different probes lying within or close to *Xist* and an X-chromosome probe (*Xpct*) (16) that lies outside the YAC, followed by comparative quantitation of *Xist* and *Xpct* bands with a PhosphorImager (Table 1).

In all lines examined, the transgenes were found to have

TABLE 1. Copy number of *Xist* YAC transgenes[a]

Transgenic line	*Xist*/*Xpct* ratio	SD	Transgene copy no.
L4 8	2.1	0.2 (3)	1
L4 12	3.0	0.4 (4)	2
L4 13	2.0	0.3 (3)	1
F1 A	1.5	0.2 (4)	1
F1 E	1.6	0.1 (4)	1
F1 C	4.5	0.4 (4)	3–4
L1 6	3.2	0.4 (3)	2–3
L1 12	7.4	0.8 (4)	6–7
L1 17	7.7	1.5 (4)	5–8
L1 25	1.9	0.3 (4)	1

[a] PhosphorImager readings of *Xist* and *Xpct* band intensities were normalized to those for male (one X), female (two X), line 53.1 (one copy of the *Xist* transgene, one X), and line 53.2 (two copies of the *Xist* transgene, one X) controls. *Xist*/*Xpct* ratio (mean ratio for each transgenic line) was calculated based on at least three separate readings (as indicated in parentheses) involving independent Southern blots. All lines but two contain a single X chromosome (including F1 C, which is XO); the exceptions, F1 A and F1 E, contain two X chromosomes, such that the *Xist*/*Xpct* ratio is $(n + 2)/2$, where n = transgene copy number.

integrated at a single (non-X) chromosomal site by DNA FISH on metaphase spreads (for example, line L1 17 in Fig. 2A).

To facilitate the assessment of transgene integrity following transfer into the mouse genome, sites for the meganuclease I-*Ppo*I had been introduced into both arms of YAC PA-2 prior to its transfer into ES cells (18) (see Materials and Methods). The DNA from transgenic lines containing YAC PA-2 (L4 and F1 series) was I-*Ppo*I digested and analyzed by pulsed-field gel electrophoresis, and the resulting Southern blots were hybridized with a variety of probes (Fig. 1A and C). The hybridization of *Xist* and several probes flanking it to common I-*Ppo*I fragments in each transgenic line showed that a large region centered around *Xist* was present in all the PA-2-derived transgenic lines. In single-copy lines L4 8, L4 13, and F1 E, the presence of a unique 460-kb I-*Ppo*I fragment to which all probes hybridized indicated that the transgene was intact. This conclusion was confirmed by the analysis of *Sfi*I-digested DNA, for which a detailed map in this region is available (23). In multicopy lines L4 12 and F1 C and in single-copy line F1 A, the detection of I-*Ppo*I bands that were much larger than 460 kb was attributed to the loss of one of the I-*Ppo*I sites (Fig. 1A and C). *Sfi*I analysis using several probes suggested that in line L4 12 the two-copy transgene was intact and that in single-copy cell line F1A and multicopy cell line F1C at least 100 kb of upstream and 50 kb of downstream sequence around *Xist* were intact (data not shown; Fig. 1A shows a summary).

Lines containing YAC PA-3 F1n (L1 series) were similarly assessed for transgene integrity, using the enzymes *Sal*I and *Sfi*I for pulsed-field analysis (Fig. 1B and C). The detection of appropriately sized restriction fragments, corresponding to those predicted from the *Sal*I and *Sfi*I map of the region (23), by using a variety of probes suggested that no major rearrangements were present in transgenic lines L1 6 (two to three copies), L1 25 (single copy), and L1 12 (six to seven copies). Although the majority of the YAC copies appeared to be intact in line L1 17 (five to eight copies), some additional *Sal*I and *Sfi*I bands (Fig. 1B and D) were detected, indicating the presence of one or two truncated fragments.

Finally, we also derived two ES cell lines (53.1 and 53.2) from our previously described transgenic mouse line Tg53, which detailed pulsed-field mapping has previously shown to carry a single, intact copy of YAC PA-2 (22). DNA FISH on metaphase spreads of the derived ES cell lines revealed that line 53.1 is XY with the transgene present on a single chromosome, while line 53.2 is XO with the transgene present on both chromosome homologues (data not shown; see Table 2).

YAC transgene function in undifferentiated ES cells. In undifferentiated ES cells, *Xist* is expressed as an unstable transcript from every X chromosome present (whether XX or XY) (39, 48). This unstable transcript can be detected by RNA FISH as a punctate signal (pinpoint) at the site of the *Xist* gene. We wished to determine whether single and multicopy transgenes were capable of expressing *Xist* in its unstable form in undifferentiated ES cells. In single-copy transgenic lines (referred to as single-copy lines hereafter), a *Xist* RNA pinpoint was observed from the transgene locus as well as from the X chromosome. In lines L4 8 and 53.1 (single copy, XY) two pinpoints were observed; in 53.2 (two single copies, XO) and F1 A (single copy, XX), three pinpoints were observed (Fig. 2B to D). Confirmation that the transgene was indeed able to express *Xist* was provided by performing YAC-specific DNA FISH on nuclei following *Xist* RNA FISH detection (as shown, for example, in line 53.2 [Fig. 2D]). No mosaicism in transgenic *Xist* expression was observed in these undifferentiated cells, with the expected number of *Xist* pinpoints, derived from the transgene(s) and the X chromosome(s), being detected in almost all cells. Similar results were obtained in experiments using a probe to detect *Brx* transcripts in lines L4 8 (single copy, XY) and F1 A (single copy, XX) (Fig. 2G and H), where in 80% of cells, two and three *Brx* pinpoints, respectively, were detectable. Furthermore, *Xist* (and *Brx*) pinpoints from the X chromosome or the transgene could not be distinguished, suggesting that the genes were being expressed at similar levels at both loci.

Three undifferentiated multicopy lines, L1 17, L4 12, and F1 C, were also examined by RNA FISH (Fig. 2E, F, and I). Again *Xist* pinpoint signals were observed from the transgene as well as the X chromosome in the overwhelming majority of cells. The transgenic pinpoint was often larger than the endogenous pinpoint, with its size varying as a function of copy number in these lines, as has been described by Lee et al. (30).

Both single-copy and multicopy YAC PA-2 and PA-3 F1n transgenes thus appear to express *Xist* efficiently in undifferentiated ES cells at the level of RNA FISH, and the size and intensity of the transgene signal appears to be correlated with copy number (note that no amplification was used in our fluorescence detection protocol). More precise quantitation of transgenic *Xist* RNA expression levels was rendered difficult

TABLE 2. *Xist* expression in differentiated transgenic ES cell lines

Cell line	YAC	XY content	Copy no.	% Cells with transgene-induced *Xist* RNA domains ($n > 100$)
53.1	PA-2	XY	1	0[a]
53.2	PA-2	XO	1[b]	0[a]
L4 8	PA-2	XY	1	0
L4 13	PA-2	XY	1	0
L1 25	PA-3 F1n	XY	1	0[a]
F1 A	PA-2	XX	1	0[c]
F1 E	PA-2	XX	1	0[c]
L4 12	PA-2	XY	2	10
L1 6	PA-3 F1n	XY	2–3	13
F1 C	PA-2	XO	3–4	66
L1 12	PA-3 F1n	XY	6–7	65
L1 17	PA-3 F1n	XY	5–8	74

[a] Very rare *Xist* RNA domains (i.e., approximately 0.2%; $n > 500$).
[b] Line 53.2 is homozygous for the single-copy transgene.
[c] *Xist* RNA domains associated only with one X chromosome (as expected for an XX cell).

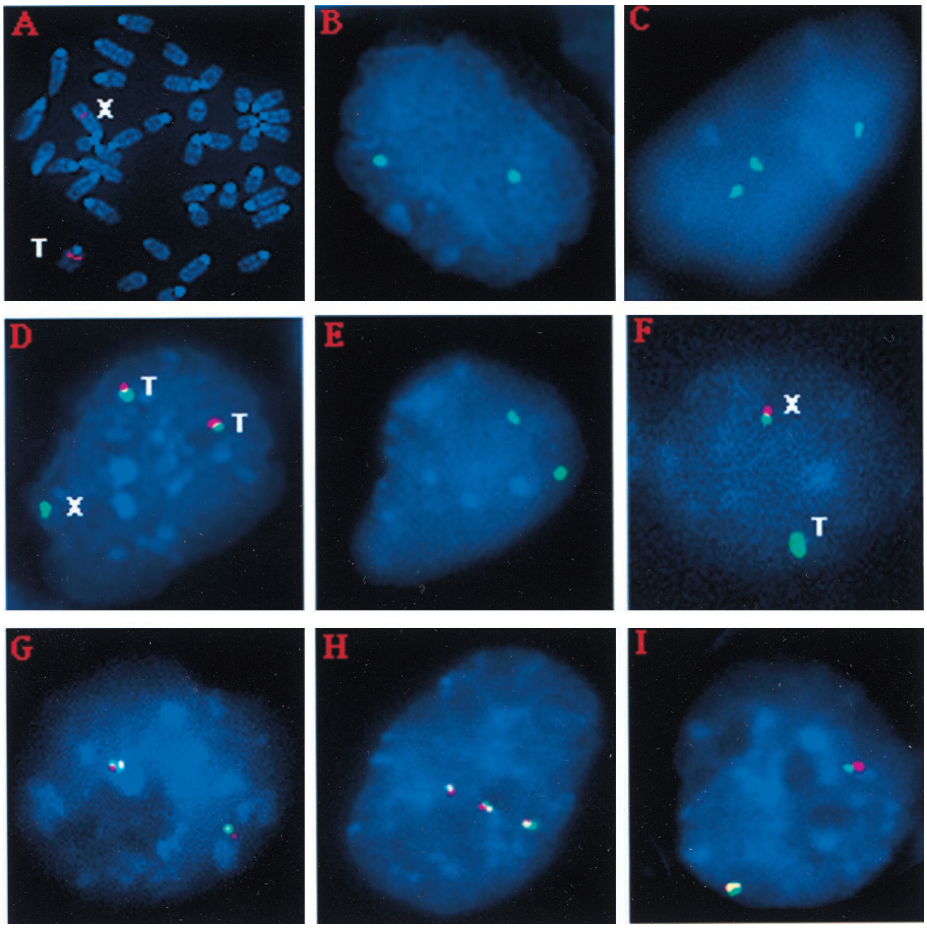

FIG. 2. *Xist* RNA expression in single-copy and multicopy transgenic lines in undifferentiated ES cells. (A) DNA FISH on metaphase chromosomes from line L1 17 showing a single integration site for the multicopy transgene in a mouse autosome. A spectrum red-labeled YAC PA-3 F1n probe hybridized with both the transgenic locus (T) and its endogenous site on the X chromosome (X). The larger signal at the transgene locus reflected its higher copy number (seven in this case). This was confirmed in other hybridizations involving a YAC-specific probe. (B to F) Under conditions that do not denature chromosomal DNA, undifferentiated ES cells were hybridized with a spectrum green labeled *Xist* probe to detect *Xist* RNA. In lines L4 8 (single copy, XY) (B), F1 A (single copy, XX) (C), and 53.2 (two unlinked copies, XO) (D), *Xist* RNA pinpoint signals over the X chromosome and the single-copy transgene loci were indistinguishable. In lines L4 12 (multicopy, XY) (E), and L1 17 (multicopy, XY) (F), the transgenic *Xist* RNA pinpoint tended to be slightly larger than the endogenous signal. (D and F) The *Xist* RNA FISH signals in positioned nuclei were photographed, and the cells were then denatured and DNA FISH was performed. (D) A probe specific for the YAC vector (pYAC4, spectrum red labeled) detected the two single-copy transgenes (T) in line 53.2; (F) a probe specific for the X chromosome (BAC X, spectrum red labeled) demonstrated that the larger *Xist* pinpoint signal corresponded to the seven-copy transgene (T) present in line L1 17. (G to I) Two-color RNA FISH for simultaneous detection of *Xist* (spectrum green) and *Brx* (spectrum red) transcripts under nondenaturing conditions (see above). (G) L4 8 (single copy, XY); (H) F1 A (single copy, XX); (I) L4 12 (two copy, XY).

owing to the unstable nature of the *Xist* transcript present in undifferentiated ES cells, which means that it can be detected only by reverse transcription-PCR (39, 48), which is not strictly quantitative.

Single-copy transgene function in differentiated cells. When induced to differentiate in vitro, female ES cell lines undergo X inactivation, the first sign of this being the coating of the prospective inactive X chromosome with stabilized *Xist* RNA (detected by RNA FISH as a *Xist* RNA domain) (39, 48). The unstable *Xist* transcript (*Xist* pinpoint) on the active X chromosome is still expressed when the *Xist* RNA domain appears but gradually disappears as differentiation progresses. In male or XO ES cells where X inactivation does not normally occur, no *Xist* RNA domain is ever observed upon differentiation and *Xist* expression on the single active X chromosome eventually disappears, as it does on the active X in female cells.

Using RNA FISH, we assessed the ability of single-copy YAC transgenes to produce *Xist* RNA domains over the transgenic locus or the X chromosome following differentiation in vitro. Cells were differentiated by growing EBs in suspension and then attaching them onto slides (total differentiation time, 7 to 12 days; see Materials and Methods). Most cell lines were differentiated more than once, and several hundred cells were analyzed by *Xist* RNA FISH for each.

In the four single-copy XY lines examined, *Xist* RNA domains were absent or extremely rare (Table 2). A control female ES cell line (HP310) differentiated in parallel showed large numbers of cells containing *Xist* RNA domains (Fig. 3A). Although *Xist* RNA domains were absent from differentiated single-copy transgenic cells, *Xist* RNA pinpoints (characteristic of undifferentiated cells) were still found in cells within and close to the attached EB mass, which are presumably in a less advanced differentiation state, but not in more fully differentiated cells, further away from the central mass (Fig. 3C). A similar profile was found in differentiated cells of the CK35 male control line (Fig. 3B).

To determine whether the transcriptional silence of the *Xist* transgene in these differentiated cells might be due to some

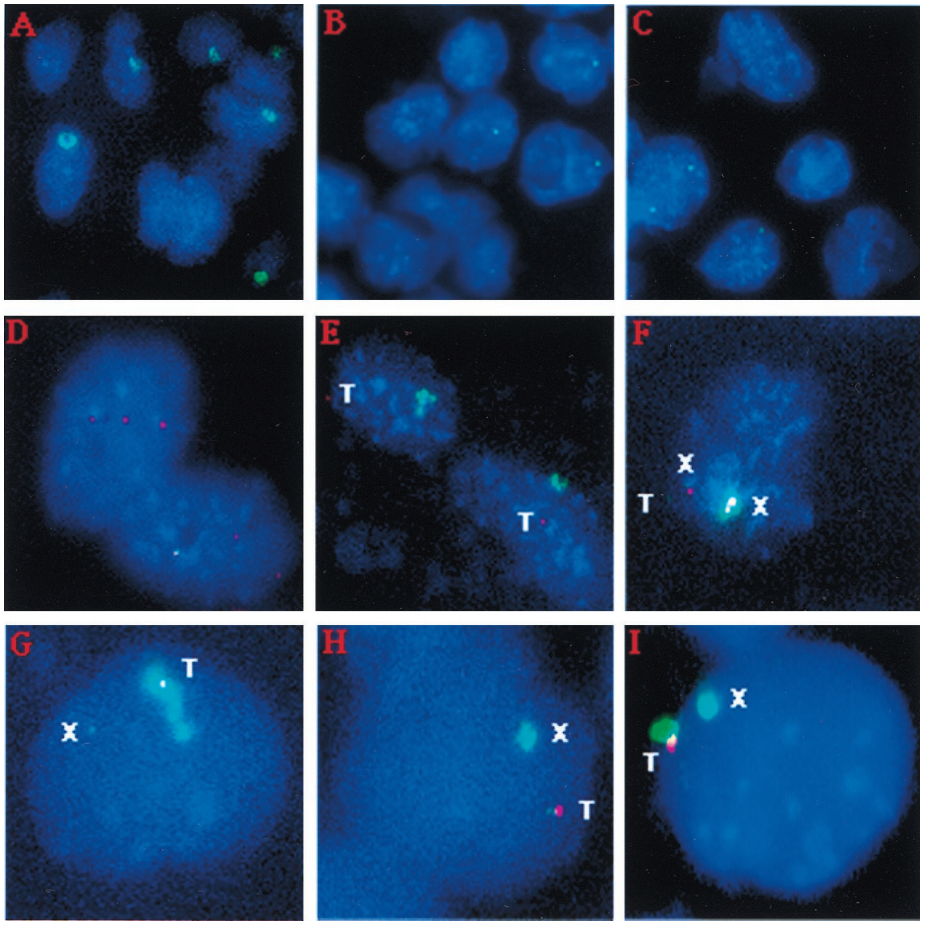

FIG. 3. *Xist* and *Brx* expression in single-copy and multicopy transgenic lines in fully differentiated ES cells. (A to D) Under conditions that do not denature chromosomal DNA, ES cells differentiated by EB attachment and outgrowth on chamber slides for several days, were hybridized with a spectrum green-labeled *Xist* probe to detect *Xist* RNA as well as a spectrum red-labeled *Brx* probe to detect *Brx* RNA in panel D. (A) Several nuclei in the control female cell line (HP310, XX) showing *Xist* RNA domains (inactive X chromosomes) in the majority of cells; in some cells, a single *Xist* RNA pinpoint (on the active X chromosome) is also detected. (B) In differentiated cells of the control male cell line (CK35, XY), either a single *Xist* RNA pinpoint or no *Xist* signal at all is detected. (C) In the single-copy transgenic line L4 8, either one, two, or no *Xist* RNA pinpoints are detected. (D) Two representative differentiated nuclei of line 53.2 (two unlinked single-copy transgenes, X0) are shown, one with no *Xist* RNA signals (green) but three *Brx* signals (red) and the other with three *Brx* signals and a single *Xist* RNA signal (yellow as it overlaps with *Brx*). (E and F) Line F1 A (single-copy transgene, XX) simultaneous RNA-DNA two-color FISH on mildly denatured nuclei (see Materials and Methods), using a spectrum green-labeled *Xist* probe and spectrum red-labeled YAC vector (E) or BAC X (F) probe. The *Xist* RNA domain is associated with the X chromosome (X) and not the single-copy transgene (T). (G to I) RNA FISH on undenatured nuclei of differentiated L4 12 ES cells (multicopy, XY) using a *Xist* probe (spectrum green) was followed by denaturation and DNA FISH using a spectrum red-labeled YAC vector probe. In the majority of nuclei (G), a single *Xist* RNA domain associated with the transgene (T) was observed. Occasionally the *Xist* RNA domain was associated with the X chromosome (H) or with both the X chromosome and the transgene (I).

overall repressive position effect, *Xist* and *Brx* transcripts were simultaneously detected by two-color RNA FISH. *Brx* transcripts from both the transgene and the X chromosome could be detected in most differentiated cells. For example, in line 53.2 (two single copy, XY) 84% ($n = 70$) of differentiated cells had three *Brx* signals but no *Xist* pinpoint signals; in line L4 8 (single copy, XY), 87% ($n = 78$) of cells had two *Brx* signals (and no *Xist* pinpoint). This finding indicates that *Brx* is still expressed from the transgene (Fig. 3D), even though *Xist* expression is absent.

We conclude that single-copy transgenes in XY ES cells are not able to induce efficient *Xist* RNA domain formation and inactivation of either the transgenic locus or the X chromosome upon differentiation.

We also tested the possibility that the presence of a second X chromosome and hence of an inactivating X chromosome upon differentiation might, via some kind of *trans* effect that could compensate for missing elements in *cis*, trigger X-inactivation function (both *cis* inactivation and counting) of single-copy transgenes. Two XX cell lines carrying single-copy transgenes (F1 A and F1 E) were analyzed following 7 to 12 days of differentiation as described above. A single *Xist* RNA domain was found in the majority of cells. The origins of the *Xist* domains were determined by using simultaneous RNA-DNA two-color FISH. With a probe specific for the X chromosome (BAC X) together with *Xist*, over 96% ($n = 75$) of *Xist* RNA domains included a BAC X pinpoint, suggesting that the domains originated only from the X chromosome (Fig. 3F). This conclusion was also reached when a probe specific for the YAC transgene (YAC vector) was used together with *Xist* (Fig. 3E).

Thus, even within an environment that is clearly permissive for X inactivation, single-copy transgenes appear to be unable to initiate inactivation in *cis* or to lead to counting.

We wished to examine whether this lack of function in single-copy transgenic lines was due to early counterselection against autosomal monosomy or X-chromosome nullisomy. *Xist* RNA domain formation has been reported to begin after just 1.3 days of in vitro differentiation in female ES cells (48),

while transcriptional inactivation of the inactive X chromosome is observed by days 2 to 4 (27). *Xist* expression was therefore examined by RNA FISH from the onset of differentiation on EBs (see Materials and Methods). No *Xist* RNA domains were ever observed for the single-copy line L4 8 (XY) when hundreds of cells were analyzed at each stage of differentiation from days 1 through 6. In control female EBs, mature *Xist* RNA domains could be found from day 2 onward (Fig. 4B). On the other hand, what looked like a small cluster of *Xist* pinpoints was sometimes observed in line L4 8 at one of the two *Xist* loci, particularly at the earliest time points (23% [n = 303] at day 1) (Fig. 4C and D), but the proportion had decreased by later stages (9% [n = 89] at day 6). In the control male ES cell line, such *Xist* RNA clusters were not seen (for the distinction made between "clusters" and "domains," see the legend to Fig. 4A). Although these *Xist* pinpoint clusters observed in differentiating L4 8 cells did not resemble RNA domains, they were suggestive of the beginning of formation of domains. Similar pinpoint clusters were also found in cells of the control female ES cell line at the earliest time points (Fig. 4A).

In summary, single-copy transgenes do not appear to give rise to mature *Xist* RNA domains even at the earliest stages of differentiation. Transient enlargement of *Xist* RNA pinpoints is, however, occasionally observed.

Multicopy transgene function in differentiated cells. Five multicopy transgenic ES cell lines were examined for the ability to induce *Xist* RNA domain formation and inactivation following differentiation via EB attachment and outgrowth (Table 2). In these lines, areas of cells with *Xist* RNA domains could easily be detected. The majority of transgenic cells contained only a single *Xist* RNA domain, although two domains were sometimes observed. The results concerning the origins of the *Xist* RNA domains for lines L4 12 (two copies of YAC PA-2, XY) and L1 17 (five to eight copies of YAC PA-3F1n, XY) are summarized in Table 3. The majority of *Xist* domains were associated with the transgene locus alone (Fig. 3G), although in a small percentage of cells only the X chromosome or both the transgene and the X chromosome were affected (Fig. 3H and I). The *Xist* RNA domains associated with the X chromosome were often smaller than those at the transgene. This may be indicative of rapid counterselection against those cells in which the single X chromosome becomes fully inactivated, leading to nullisomy.

Transcriptional inactivation following *Xist* RNA domain formation was shown to have occurred in L4 12 cells by analyzing *Xist* and *Brx* expression simultaneously. In fully differentiated cells, a *Brx* pinpoint was never found to be present in *Xist* domains (n = 22), although a *Brx* signal was present elsewhere in the nucleus (an example of one such nucleus is shown in Fig. 5C).

In summary, our data indicate that both YACs PA-2 and PA-3 F1n when present in a multicopy transgene array show ectopic Xic function associated with both inactivation in *cis* and counting.

In the course of our analysis we noted a variable degree of mosaicism in differentiated multicopy cell lines, represented by patches of cells in which *Xist* RNA domains were either present or absent. This was most striking in cells where differentiation was more advanced (i.e., those furthest from the EB mass) (Fig. 5A and B). Lines containing fewer copies of the transgene, such as L4 12 (two copies), showed a lower proportion of cells with *Xist* RNA domains (10% [n = 200]) than lines carrying higher-copy-number transgenes, such as L1 17 (five to eight copies) (74% [n = 393]) (Table 2). This absence of *Xist* RNA domains in a proportion of cells was unlikely to be due to silencing of the transgene as a whole, as in line L4 12 *Brx* expression from the transgene (and the X chromosome) could be observed in 80% (n = 236) of these *Xist*-negative cells (Fig. 5C).

To determine whether the mosaicism observed in multicopy lines might be due to inefficiency in transgene function or to early counterselection against such function, *Xist* RNA domain formation in line L4 12 (two copies, XY) was examined at the earliest stages of differentiation (day 1 to 6 EBs). Large *Xist* RNA pinpoint clusters (Fig. 4E and G) were observed in up to 67% (n = 138) of cells at day 1 but had decreased to 21% (n = 154) by day 6, being replaced mainly by cells in which only one or no *Xist* pinpoints were present. At day 2, 3% of cells (n = 171) contained mature *Xist* domains (Fig. 4F), and this proportion reached a maximum of 9% (n = 154) by day 6, similar to that observed at later stages of differentiation (10% [see above]). In line L1 17 (five to eight copies), large numbers of cells with enlarged pinpoint clusters were found in differentiating ES cells, but unlike in line L4 12, they appeared to be replaced mainly by cells with mature *Xist* RNA domains at more advanced stages of differentiation (data not shown).

Since the proportion of cells with *Xist* RNA domains in the two-copy transgenic line L4 12 at early stages of differentiation is no greater than that found at later stages, we conclude that there is no obvious counterselection against transgene function in these cells. Rather, in lines with low copy numbers of transgenes, the transient presence of enlarged *Xist* pinpoint clusters at early differentiation stages and the marked mosaicism of cells with *Xist* RNA domains observed at later stages suggests an inefficiency in establishing *Xist* RNA domains, with certain aspects of the inactivation process perhaps being initiated but then being frequently aborted.

DISCUSSION

Based on accumulating genetic and knockout data, it is becoming apparent that the Xic may be a complex locus and that elements situated not only within, but also some distance from the *Xist* gene may be required for the regulation of X-chromosome counting, choice, and *cis* inactivation (see reference 12 for a review). Given the potential complexity of the Xic region, we have addressed the question of whether large, *Xist*-containing transgenes can trigger the X-inactivation process in an autonomous fashion when present as single copies as opposed to multicopy arrays.

ES cell lines carrying single or multiple copies of *Xist* YAC transgenes have been assessed for the capacity to induce inactivation in *cis* and counting. Both single-copy and multicopy transgenes express the *Xist* transcript in its unstable form in undifferentiated ES cells. Upon differentiation, however, none of the six single-copy lines investigated showed any ability to induce the X-inactivation process. Only lines with multicopy transgenes displayed mature *Xist* RNA domain formation and subsequent inactivation at the transgene-carrying autosome and, to a lesser extent, of the X chromosome. Single-copy transgenes did not induce *Xist* RNA domain formation either at the autosome on which they were integrated or at the endogenous X chromosome. Thus, although the sequences enabling *Xist* transcription in undifferentiated cells are clearly present in these single-copy transgenes, other elements necessary for the initiation of X inactivation appear to be lacking. Importantly, this incapacity of single-copy transgenes for Xic function was unchanged by the presence of an X chromosome undergoing inactivation in differentiating XX ES cell lines (F1 A and F1 E), indicating that any *trans*-acting factors that might be required for the initiation of X inactivation were not capable of compensating for this lack of function.

We were also able to address the question of whether any

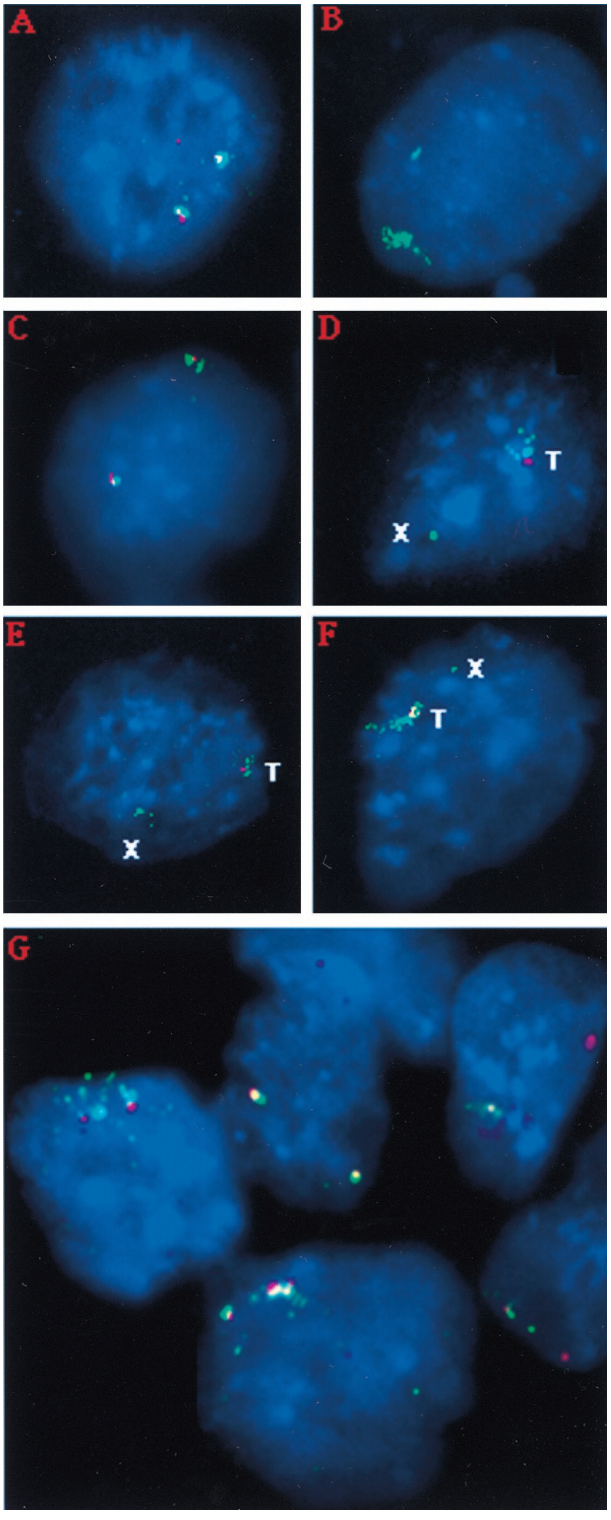

FIG. 4. Evolution of *Xist* RNA signals in differentiating ES cells. (A) One-day EB of control female ES line HP310; two-color RNA FISH detecting *Xist* (spectrum green) and *Brx* (spectrum red) transcripts. *Xist* pinpoint clusters can sometimes be seen at this stage. We considered "clusters" to consist of 2 to 10 distinct pinpoints randomly scattered around the site of transcription, while "domains" were dense, more extensive signals that covered an interphase chromosome-sized area. (B) Two-day EB, control female ES line HP310; *Xist* (spectrum green) RNA FISH showing a *Xist* RNA domain. (C) One day EB of line L4 8 (single-copy transgene, XY); two-color RNA FISH detecting *Xist* (spectrum green) and *Brx* (spectrum red) transcripts. *Xist* pinpoint clusters can sometimes be seen at this stage. (D) L4 8 1-day EB. Performance of *Xist* RNA FISH (spectrum green) was followed by denaturation and DNA FISH using a YAC-specific probe (spectrum red), which shows that *Xist* RNA pinpoint clusters tend to be associated with the transgene loci. (E and F) L4 12 (two-copy transgene, XY), 2-day EB. Performance of *Xist* RNA FISH (spectrum green) was followed by denaturation and DNA FISH using a YAC-specific probe (spectrum red). (E) *Xist* RNA pinpoint clusters are sometimes seen either at the transgenic locus alone or (as shown here) at both the X chromosome and the transgene (T). (F) *Xist* RNA domains are mostly found to be associated with the transgene. (G) Representative view of L4 12 cells (two-copy transgene, XY) from 1-day EBs. Two-color RNA FISH detecting *Xist* (spectrum green) and *Brx* (spectrum red) transcripts shows that the majority of cells contain one or two *Xist* pinpoint clusters.

TABLE 3. Origin of *Xist* RNA domains in fully differentiated transgenic ES cells[a]

Line	No. (%) with:				Total no. of cells with *Xist* RNA domains
	One *Xist* RNA domain		Two *Xist* RNA domains		
	Tg only	X only	Tg only[b]	Tg and X	
L4 12	233 (93)	6 (2)	0 (0)	12 (5)	251
L1 17	82 (80)	2 (2)	9 (9)	9 (9)	102

[a] Simultaneous RNA-DNA FISH was used with either *Xist*/pYAC4 or *Xist*/BAC X probe combinations. Tg, transgene.
[b] Polyploid cells carrying two or more copies of the transgene.

number of *Xist* YAC transgenes greater than one would enable ectopic Xic function. We found that lines containing just two copies of either YAC PA-2 (line L4 12) or PA-3 F1n (line L1 6) were capable of both *cis* inactivation and counting. Furthermore, the efficiency of such function appeared to increase with increasing transgene copy number: compared to lines with seven transgene copies (L1 12 and L1 17), the proportion of differentiated cells with *Xist* RNA domains in the two-copy line, L4 12, was low. We believe that the inefficiency of low-copy-number transgenes is due to an incapacity to establish *Xist* RNA domains, rather than to selective loss of cells in which the transgene was functional. This conclusion is based on the observation that the proportion of cells in which mature *Xist* RNA domains were formed was low and did not decrease over time and that a transiently high proportion of *Xist* pinpoint clusters rather than domains was found at the onset of differentiation.

We have found that multiple transgene copies have to be linked in *cis* in order to induce *cis* inactivation and counting. The presence of two single-copy transgenes on homologous chromosomes in the same cell (line 53.2, YAC PA-2) did not lead to *Xist* RNA domain formation and associated inactivation, unlike the situation in lines L4 12 (YAC PA-2) and L1 6 (YAC PA-3 F1n), where the two transgene copies are *cis* linked. This *cis* requirement suggests that critical elements that may be missing in the YACs can be compensated for in some way by the juxtaposition of two or more YAC copies. Similar requirements for multiple transgene copies in order to establish appropriate genetic control have been described for other systems. For example, β-globin locus control region 5′HS2 transgenes appear to be functional when at least two copies are present but are unable to direct position-independent expression when present as single copies (17). Similarly, in the case of H19 minitransgenes, correct expression and imprinting was observed only when two or more linked copies were present (3). The use of larger transgenes (130 kb) overcomes the incapacity of single-copy H19 transgenes to function autonomously (1) either because sequences essential for correct reg-

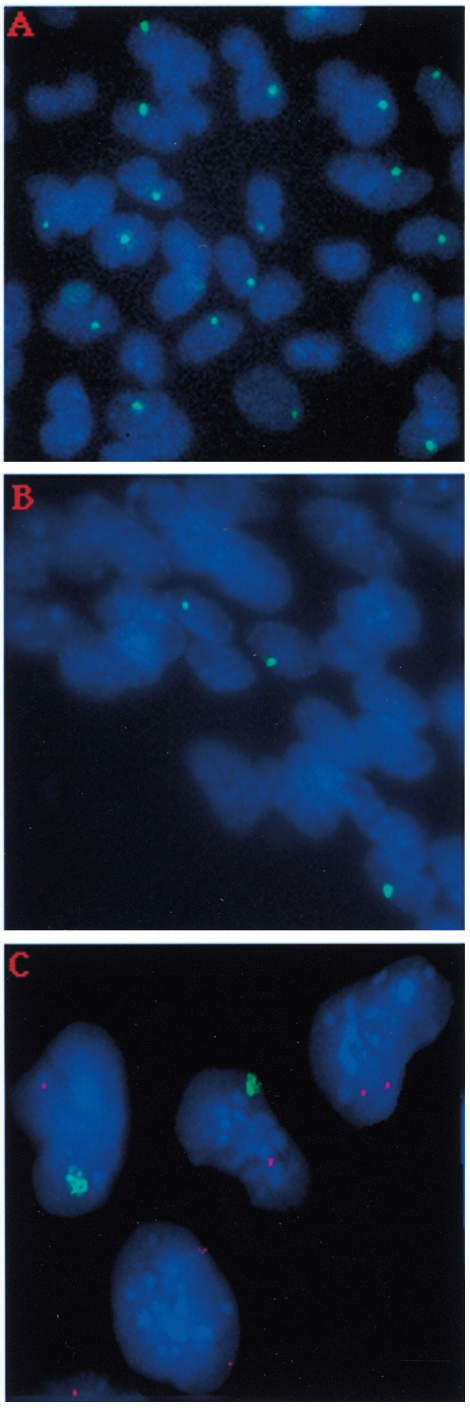

FIG. 5. Mosaicism in the presence of *Xist* RNA domains in fully differentiated ES cells. (A) *Xist* RNA FISH (spectrum green) on differentiated L1 17 cells (five- to eight-copy transgene, XY). Shown is a representative view with *Xist* RNA domains present in the majority of cells. (B) *Xist* RNA FISH (spectrum green) on differentiated L4 12 cells (two-copy transgene, XY). Shown is a representative view with *Xist* RNA domains present in a minority of cells. (C) Two-color RNA FISH detecting *Xist* (spectrum green) and *Brx* (spectrum red) on differentiated L412 cells (two-copy transgene, XY) as in panel B. In cells containing a *Xist* RNA domain, no *Brx* signal is detected within the domain (indicating *Brx* inactivation), while a *Brx* signal is detected on the other (active) Xic in these cells. In cells with no *Xist* domain, two *Brx* signals are detected, indicating that despite the lack of *Xist* expression and domains in these cells, the transgene is still functional (i.e., *Brx* is expressed). In all cases, ES cells were differentiated for a total of 9 days (4 days of EB growth in suspension followed by 5 days of outgrowth after EB attachment onto chamber slides) and fixed in situ on the slides.

ulation of the H19 gene are present in this larger region or because H19 is buffered by the extra sequences from position effects at random integration sites. As discussed below, *Xist* transgenes larger than those tested here may similarly enable autonomous Xic function.

Unlike single-copy transgenes, multicopy transgenes were also found to induce inactivation of the X chromosome (evidence for counting), although the frequency of differentiated cells in which we observed this was lower than that reported by Lee et al. (30) and Herzing et al. (24). This difference in frequency may be due to sequence differences between the transgenes. An *Xce* effect biasing the choice of allele (transgene or X chromosome) that is inactivated may be involved. Genetic mapping suggests that *Xce* is closely linked to, but distinct from, *Xist* (49). Both of the YACs that we used are derived from the C3H/He strain mice, which carry an Xce^a allele, while the YAC and cosmid used by Lee et al. (30) and Herzing et al. (24), respectively, are both derived from strains carrying an Xce^b allele. As the ES cell X chromosome is 129 derived (Xce^a) in all these studies, it would be more likely to be chosen for inactivation in the presence of an Xce^b allele (24, 30) than in the presence of another Xce^a allele (this study) (13), which is consistent with the differences observed. Another possibility that we cannot formally rule out is that the host ES cell lines we used, or our culture conditions, have led to a more severe selection against cells with X-chromosome nullisomy. This seems unlikely, however, since no evidence was found for a higher proportion of X chromosome *Xist* RNA domains at the earliest stages of differentiation.

Hypotheses of X-inactivation function based on the differences in behavior of single-copy and multicopy *Xist* transgenes. (i) Capacity of transgenes to inactivate in *cis*. One hypothesis which could explain the inability of single copy transgenes to lead to *cis* inactivation is that the YACs we have used lack elements that are necessary for *Xist* RNA association in *cis* and/or the spread of inactivation when X inactivation initiates. Thus, although our YACs may actually include all of the sequences necessary to trigger X inactivation, they may lack elements that are, for example, involved in the spread of the inactive state. The existence of X chromosome-specific relay elements that allow the efficient spread of the inactive state from the Xic, via for example the binding of heterochromatin-inducing protein factors, has previously been proposed (38, 44). The potential role of the *Xist* RNA as an effector of propagation may be mediated by its association with chromatin in *cis* and might occur via proteins bound to such relay elements along the X chromosome (8). A high density of such relay elements in the region around *Xist* might be essential to nucleate formation of *Xist* RNA-chromatin protein complexes and to establish a *Xist* RNA domain. The inability of single-copy transgenes to function could be due to the absence of a sufficient number of such elements in the YACs tested. Multiple linked YAC copies would increase the number of such relay elements surrounding the *Xist* genes, allowing more efficient establishment of *Xist* RNA association in *cis*. The *Xist* RNA clusters observed in both single- and two-copy transgenic lines during early differentiation stages may be evidence for the inability of the *Xist* RNA to nucleate the formation of a full domain from a limited set of such elements. The prediction would be that *Xist* transgenes flanked with more extensive X-linked sequences should be functional as single copies at ectopic sites. A 750-kb YAC centered around *Xist* will be analyzed for single-copy function.

Recently, Lyon (33) has proposed that interspersed repetitive elements of the LINE (long interspersed nuclear element) type, in which the X chromosome is particularly rich (5), might

represent the above-mentioned relay elements. The 94-kb region lying immediately 3′ to *Xist* which is present in YACs PA-2 and PA-3F1n, contains relatively few LINE sequences (52), consistent with the hypothesis that there may be insufficient relay elements to allow autonomous transgenic function of these YACs. A testable prediction of this hypothesis would be that single-copy *Xist* transgenes, integrated into LINE-rich autosomal regions (5) would be capable of showing *cis* inactivation. The rare *Xist* RNA domains seen in single-copy lines 53.1, 53.2, and L1 25 (Table 2) may be indicative of transgene integration sites close to LINE elements. Analysis of these and other single-copy lines and the LINE content of the chromosomal regions into which the YACs have integrated will address this point. Alternatively, LINE sequences could be targeted into the region surrounding the *Xist* gene in the YACs prior to transfer into ES cells. Clearly, elements other than LINEs may also be involved in this spreading process.

Another possibility is that higher local concentrations of *Xist* RNA provided by multicopy *Xist* transgenes facilitate the establishment of inactivation via autosomal sequences acting as relay elements. Such autosomal elements may be less frequent or have lower binding affinities for the factors involved in inactivation than X-chromosome elements and may therefore require higher levels of *Xist* expression in order to nucleate *Xist* RNA domain formation at ectopic sites. Accurate quantitation of the *Xist* RNA levels derived from our single-copy and multicopy transgenes in differentiating cells was hampered by the heterogeneity that is inherent to this in vitro cell system, in addition to the fact that our RNA FISH data clearly revealed a high degree of mosaicism in the capacity of transgenes to induce *Xist* RNA domain formation. The possibility that *Xist* expression levels might facilitate *Xist* RNA domain formation could be tested by engineering transgenes from which very high levels of *Xist* expression can be induced and seeing whether they can function as single copies.

(ii) Capacity of transgenes to count. The inability of our single-copy transgenes to induce counting was surprising in the light of currently favored models for this process. It has been proposed that as X inactivation initiates, one Xic per diploid cell is blocked to remain active by a factor present in limiting quantity, and as differentiation proceeds, X inactivation occurs only from remaining unblocked Xics (42). Given the large size of our transgenes, it seemed likely that sequences involved in this hypothetical blocking step would be present, for example, within the 65-kb region 3′ to *Xist* which when deleted eliminates counting (15). It has also recently been suggested that transcription from a newly identified *Xist* upstream promoter, Po, from which only the unstable form of *Xist* RNA is transcribed, may be required for a *Xist* allele to be registered by the counting mechanism (25). Preliminary results obtained using probes that specifically detect Po-derived transcripts (19a) indicate that Po is efficiently used by our single-copy transgenes, thus suggesting that transcription from Po is not in itself sufficient to induce counting in this context.

One hypothesis that could explain the lack of counting, as well as the lack of *cis* inactivation, observed with our single-copy YACs is that they are simply not recognized as Xics at the time of X inactivation. Spatial restriction within the nucleus of factors involved in the first steps of X inactivation could explain the counting process whereby two Xics in the same nucleus are treated differently. Thus, localization of an Xic within (or exclusion from) a specific nuclear compartment might be essential for the initiation of X inactivation, whether it be via a switch in *Xist* promoter usage (25), stabilization of the *Xist* transcript (39, 48), or the binding of specific factors to the region 3′ to *Xist* (15). Nuclear localization might be mediated by long-range elements flanking the Xic, such as sequences that associate with the nuclear envelope that could anchor *Xist* within a defined compartment of the nucleus (36). Indeed, it has recently been demonstrated that perinuclear localization of chromatin facilitates transcriptional silencing in yeast (2). The YACs that we have used may lack such elements or a subset of them, so that single-copy transgenes will not be registered as supplementary Xics, leading neither to counting nor to inactivation in *cis*. When multiple linked copies are present, however, elements from neighboring YACs might allow the appropriate compartmentalization of the transgenic locus.

A three-dimensional FISH analysis of ES cells undergoing differentiation should reveal eventual colocalization of Xic regions present in female XX cells and transgenic cell lines. The prediction is that in multicopy lines the transgenic locus will localize to the same nuclear compartment as a second X chromosome, whereas in a single-copy line the transgene will show absence of such localization. Another prediction of this hypothesis is that longer *Xist* YAC transgenes will contain the appropriate nuclear compartmentalization signals enabling them to function as single copies.

In conclusion, we present evidence that large single-copy *Xist* transgenes are incapable of autonomous function with respect to X inactivation. The fact that two or more YAC copies linked in *cis* can ensure such X-inactivation function suggests that the key elements necessary for inactivation may all be included in the YACs that we have used, but that certain of these elements may not be present in numbers sufficient to allow efficient function of the Xic. Although the precise mechanism by which multicopy transgenes can function remains unclear, the challenge is now to identify the nature of the elements that would allow autonomous Xic function. The large-scale sequencing of the Xic regions in human and mouse genomes combined with functional analysis using modified *Xist* transgenes as well as deletions should provide further insight into Xic function.

ACKNOWLEDGMENTS

This work was conducted in collaboration with the Equipe Dyogen, and the support of C. Vourc'h is gratefully acknowledged. We also thank C. Kress and C. Babinet for the gift of the CK35 ES cell line, R. Anand for the pLNA-1 vector, B. Panning for the p*Xist*3K probe, N. Brockdorff for *Xist* Po probes, and L. Dandolo, P. Clerc, E. Debrand, and C. Rougeulle for critical reading of the manuscript. Special thanks go to V. Colot for helpful discussions.

This work was supported by grants from the Association Francaise contre les Myopathies and the Association de Recherche sur le Cancer.

REFERENCES

1. **Ainscough, J. F.-X., T. Koide, M. Tada, S. Barton, and A. Surani.** 1997. Imprinting of *Igf2* and *H19* from a 130 kb YAC transgene. Development **124:** 3621–3632.
2. **Andrulis, E. D., A. M. Neiman, D. C. Zappulla, and R. Sternglanz.** 1998. Perinuclear localization of chromatin facilitates transcriptional silencing. Nature **394:**592–595.

2a. **Arnaud, D.** Unpublished data.

3. **Bartolomei, M. S., A. L. Webber, M. E. Brunkow, and S. M. Tilghman.** 1993. Epigenetic mechanisms underlying the imprinting of the mouse *H19* gene. Mol. Cell. Biol. **7:**1663–1673.
4. **Borsani, B., R. Tonlorenzi, M.-C. Simmler, L. Dandolo, D. Arnaud, V. Capra, M. Grompe, A. Pizzuti, D. Muzni, C. Lawrence, H. F. Willard, P. Avner, and A. Ballabio.** 1991. Characterization of a murine gene expressed from the inactive X chromosome. Nature **351:**325–329.
5. **Boyle, A. L., S. G. Ballard, and D. C. Ward.** 1990. Differential distribution of long and short interspersed element sequences in the mouse genome: chromosome karyotyping by fluorescence in situ hybridization. Proc. Natl. Acad. Sci. USA **87:**7757–7761.
6. **Brockdorff, N., A. Ashworth, G. F. Kay, P. Cooper, S. Smith, V. M. McCabe, D. P. Norris, G. D. Penny, D. Patel, and S. Rastan.** 1991. Conservation of position and exclusive expression of mouse *Xist* from the inactive X chromosome. Nature **35:**329–331.

7. **Brockdorff, N., A. Ashworth, G. F. Kay, V. M. McCabe, D. P. Norris, P. J. Cooper, S. Swift, and S. Rastan.** 1992. The product of the mouse *Xist* gene is a 15kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell **71:**515–526.
8. **Brockdorff, N., and M. Duthie.** 1998. X chromosome inactivation and the *Xist* gene. Cell. Mol. Life Sci. **54:**104–112.
9. **Brown, C. J., A. Ballabio, J. L. Rupert, R. G. Lafrenière, M. Grompe, R. Tonlorenzi, and H. F. Willard.** 1991. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature **349:**38–44.
10. **Brown, C. J., B. D. Hendrich, J. L. Rupert, R. G. Lafrenière, Y. Xing, R. G. Lawrence, and H. F. Willard.** 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.
11. **Camus, A., C. Kress, C. Babinet, and J. Barras.** 1996. Unexpected behaviour of a gene trap vector comprising a fusion between the *Sh ble* and the *LacZ* genes. Mol. Reprod. Dev. **45:**255–263.
12. **Carrel, L., and H. F. Willard.** 1998. Counting on *Xist*. Nat. Genet. **19:**211–212.
13. **Cattanach, B. M., and C. E. Williams.** 1972. Evidence of non-random X-chromosome activity in the mouse. Genet. Res. **19:**229–240.
14. **Clemson, C. M., J. A. McNeil, H. F. Willard, and J. B. Lawrence.** 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.
15. **Clerc, P., and P. Avner.** 1998. Role of the region 3′ to *Xist* exon 6 in the counting process of X-chromosome inactivation. Nat. Genet. **19:**249–253.
16. **Debrand, E., E. Heard, and P. Avner.** 1998. Cloning and localization of the murine *Xpct* gene: evidence for complex rearrangements during the evolution of the region around the *Xist* gene. Genomics **48:**296–303.
17. **Ellis, J., D. Talbot, N. Dillon, and F. Grosveld.** 1993. Synthetic human β-globin 5′HS2 constructs function as locus control regions only in multicopy transgene concatamers. EMBO J. **12:**127–134.
18. **Fairhead, C., E. Heard, D. Arnaud, P. Avner, and B. Dujon.** 1995. Insertion of unique sites into YAC arms for rapid physical analysis following YAC transfer into mammalian cells. Nucleic Acids Res. **23:**4011–4012.
19. **Goto, T., and M. Monk.** 1998. Regulation of X-chromosome inactivation in development in mice and humans. Microbiol. Mol. Biol. Rev. **62:**362–378.

19a. **Heard, E.** Unpublished data.

20. **Heard, E., P. Avner, and R. Rothstein.** 1994. Creation of a deletion series of mouse YACs covering a 500 kb region around *Xist*. Nucleic Acids Res. **22:**1830–1837.
21. **Heard, E., P. Clerc, and P. Avner.** 1997. X-chromosome inactivation in mammals. Annu. Rev. Genet. **31:**571–610.
22. **Heard, E., C. Kress, F. Mongelard, B. Courtier, C. Rougeulle, A. Ashworth, C. Vourc'h, C. Babinet, and P. Avner.** 1996. Transgenic mice carrying an *Xist*-containing YAC. Hum. Mol. Genet. **5:**441–450.
23. **Heard, E., M. C. Simmler, Z. Larin, C. Rougeulle, B. Courtier, H. Lehrach, and P. Avner.** 1993. Physical mapping and YAC contig analysis of the region surrounding *Xist* on the mouse X chromosome. Genomics **15:**559–569.
24. **Herzing, L. B. K., J. T. Romer, J. M. Horn, and A. Ashworth.** 1997. *Xist* has properties of the X-chromosome inactivation centre. Nature **386:**272–75.
25. **Johnston, C. M., T. B. Nesterova, E. J. Formstone, A. E. Newall, S. M. Duthie, S. A. Sheardown, and N. Brockdorff.** 1998. Developmentally regulated *Xist* promoter switch mediates initiation of X inactivation. Cell **94:**809–817.
26. **Kay, G. F., G. D. Penny, D. Patel, A. Ashworth, N. Brockdorff, and S. Rastan.** 1993. Expression of *Xist* during mouse development suggests a role in initiation of X chromosome inactivation. Cell **72:**171–182.
27. **Keohane, A. M., L. P. O'Neill, N. D. Belyaev, J. S. Lavender, and B. Turner.** 1996. X-inactivation and histone H4 acetylation in embryonic stem cells. Dev. Biol. **180:**618–630.
28. **Lamb, B. T., S. S. Sangram, A. M. Lawler, H. H. Slunt, C. A. Kitt, W. G. Kearns, P. L. Pearson, D. P. Price, and J. D. Gearhart.** 1993. Introduction and expression of the 400 kb precursor amyloid protein gene in transgenic mice. Nat. Genet. **5:**22–29.
29. **Lee, J. T., and R. Jaenisch.** 1997. Long-range *cis* effects of ectopic X-inactivation centres on a mouse autosome. Nature **386:**275–278.
30. **Lee, J. T., W. M. Strauss, J. A. Dausman, and R. Jaenisch.** 1996. A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell **86:**83–94.
31. **Lyon, M. F.** 1961. Gene action in the X chromosome of the mouse (*Mus musculus* L.). Nature **190:**372–373.
32. **Lyon, M. F.** 1996. X-chromosome inactivation. Pinpointing the centre. Nature **379:**116–117.
33. **Lyon, M. F.** 1998. X-chromosome inactivation: a repeat hypothesis. Cytogenet. Cell Genet. **80:**133–137.
34. **Marahrens, Y., J. Loring, and R. Jaenisch.** 1998. Role of the *Xist* gene in X chromosome choosing. Cell **92:**657–664.
35. **Marahrens, Y., B. Panning, J. Dausman, W. Strauss, and R. Jaenisch.** 1997. *Xist*-deficient mice are defective in dosage compensation but not spermatogenesis. Genes Dev. **11:**156–166.
36. **Marshall, W. F., A. F. Dernburg, B. Harmon, D. A. Agard, and J. W. Sedat.** 1996. Specific interactions of chromatin with the nuclear envelope: positional determination within the nucleus in *Drosophila melanogaster*. Mol. Biol. Cell **7:**825–842.
37. **Matsuura, S., V. Episkopou, R. Hamvas, and S. D. M. Brown.** 1996. *Xist* expression from an *Xist* YAC transgene carried on the mouse Y chromosome. Hum. Mol. Genet. **5:**451–459.
38. **McBurney, M. W.** 1988. X chromosome inactivation: a hypothesis. Bioessays **9:**85–88.
39. **Panning, B., J. Dausman, and R. Jaenisch.** 1997. X chromosome inactivation is mediated by *Xist* RNA stabilization. Cell **90:**907–916.
40. **Penny, G. D., G. F. Kay, S. A. Sheardown, S. Rastan, and N. Brockdorff.** 1996. Requirement for *Xist* in X chromosome inactivation. Nature **379:**131–137.
41. **Plenge, R. M., B. D. Hendrich, C. Schwartz, J. F. Arena, A. Naumova, C. Sapienza, R. M. Winter, and H. F. Willard.** 1997. A promoter mutation in the *XIST* gene in two unrelated families with skewed X chromosome inactivation. Nat. Genet. **17:**353–356.
42. **Rastan, S.** 1983. Non-random X chromosome inactivation in mouse X-autosome translocations—location of the inactivation centre. J. Embryol. Exp. Morphol. **78:**1–22.
43. **Rastan, S., and E. J. Robertson.** 1985. X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. J. Embryol. Exp. Morphol. **90:**379–388.
44. **Riggs, A. D., J. Singer-Sam, and D. H. Keith.** 1985. Methylation of the *PGK* promoter and an enhancer way-station model for X-chromosome inactivation, p. 211–222. *In* A. A. Sandberg (ed.), Biochemistry and biology of DNA methylation. Alan R. Liss, New York, N.Y.
45. **Riley, J. H., J. E. N. Morten, and R. Anand.** 1992. Targeted integration of neomycin into yeast artificial chromosomes (YACs) for transfection into mammalian cells. Nucleic Acids Res. **20:**2971–2976.
46. **Robertson, E. J.** 1987. Embryo-derived stem cell lines, p. 71–112. *In* E. J. Robertson (ed.), Teratocarcinomas and embryonic stem cells: a practical approach. IRL Press, Oxford, England.
47. **Rougeulle, C., L. Colleaux, B. Dujon, and P. Avner.** 1994. Generation and characterization of an ordered lambda clone array for the 460 kb region surrounding the murine *Xist* sequence. Mamm. Genome **5:**416–423.
48. **Sheardown, S. A., S. M. Duthie, C. M. Johnston, A. E. T. Newall, E. J. Formstone, R. M. Arkell, T. B. Nesterova, G. C. Alghisi, S. Rastan, and N. Brockdorff.** 1997. Stabilization of *Xist* RNA mediates initiation of X chromosome inactivation. Cell **91:**99–107.
49. **Simmler, M. C., B. M. Cattanach, C. Rasberry, C. Rougeulle and P. Avner.** 1993. Mapping the murine *Xce* locus with (CA)n repeats. Mamm. Genome **4:**523–530.
50. **Simmler, M. C., D. B. Cunningham, P. Clerc, T. Vermat, B. Caudron, C. Cruaud, A. Pawlak, C. Szpirer, J. Weissenbach, J. M. Claverie, and P. Avner.** 1996. A 94kb genomic sequence 3′ to the murine *Xist* gene reveals an AT rich region containing a new testis specific gene *Tsx*. Hum. Mol. Genet. **11:**1713–1726.
51. **Simmler, M. C., E. Heard, C. Rougeulle, C. Cruaud, J. Weissenbach, and P. Avner.** 1997. Localisation and expression analysis of a novel conserved brain expressed transcript, *brx/BRX*, lying within the Xic/XIC candidate region. Mamm. Genome **8:**760–766.
52. **Willard, H. F.** 1996. X chromosome inactivation, *XIST*, and pursuit of the X-inactivation center. Cell **86:**5–7.
53. **Willard, H. F., and H. K. Salz.** 1997. Remodeling chromatin with RNA. Nature **386:**228–229.