Chromosomal silencing and localization are mediated by different domains of *Xist* RNA

Anton Wutz, Theodore P. Rasmussen & Rudolf Jaenisch

Published online: 7 January 2002, DOI: 10.1038/ng820

The gene *Xist* initiates the chromosomal silencing process of X inactivation in mammals. Its product, a noncoding RNA, is expressed from and specifically associates with the inactive X chromosome in female cells. Here we use an inducible *Xist* expression system in mouse embryonic stem cells that recapitulates long-range chromosomal silencing to elucidate which *Xist* RNA sequences are necessary for chromosomal association and silencing. We show that chromosomal association and spreading of *Xist* RNA can be functionally separated from silencing by specific mutations. Silencing requires a conserved repeat sequence located at the 5′ end of *Xist*. Deletion of this element results in *Xist* RNA that still associates with chromatin and spreads over the chromosome but does not effect transcriptional repression. Association of *Xist* RNA with chromatin is mediated by functionally redundant sequences that act cooperatively and are dispersed throughout the remainder of *Xist* but show little or no homology.

Introduction

Evidence of mammalian X-chromosome inactivation was first obtained in 1949, through the observation of a dense-staining nuclear structure in neuronal nuclei of female cats[1]. Later it was determined that this Barr body contained one of the two X-chromosomes of the female mammalian nucleus. X inactivation equalizes the expression of X-linked genes between XY males and XX females through transcriptional silencing of one of the two female X chromosomes in a random manner[2–4]. A more recent study showed that a single locus on human Xq13, the X inactivation center (*XIC*), is required for X inactivation[5]. This observation facilitated the isolation of *XIST* in humans[6] and mice[7] as a gene expressed specifically from the otherwise silent inactive X chromosome. *Xist* is an unusual gene; it does not encode any protein, but its processed RNA localizes to the chromatin of the inactive X[8]. The unique localization of the transcript suggested that it might be involved in the silencing process. Genetic manipulations in the mouse have shown that *Xist* is required for silencing[9,10]. X inactivation has also been re-created in transgenic systems with genomic fragments of the X-inactivation center (*Xic*)[11,12] and an *Xist* cDNA, establishing *Xist* as a crucial regulator of X inactivation[13]. The mechanism by which *Xist* RNA localizes to chromatin, spreads over the inactive X chromosome and causes transcriptional silencing of X-linked genes is, however, largely unknown. *Xist* is regulated by a counting-and-choosing mechanism that ensures that one of the two X chromosomes in females become inactivated in a random manner[2]. The complexity of this regulation has complicated the investigation of the RNA. For example, the fact that a deletion within *Xist* affects choice—leading to primary nonrandom inactivation of the other X chromosome—complicates the study of the function of the endogenous RNA[14].

We previously repeated the process of *Xist*-mediated transcriptional silencing based on inducible expression of an *Xist* cDNA transgene, separating *Xist* expression from counting and choice[13]. Our results indicated that X inactivation has to be initiated early in embryonic stem (ES) cell differentiation. In undifferentiated ES cells, *Xist*-mediated silencing occurred independently of any of the known chromosomal modifications of the inactive X in differentiated cells, such as late-replicating chromatin, histone H4 hypoacetylation[13] and localization of macroH2A1 (ref. 15). This suggested that *Xist* used an as-yet-unknown mechanism to initiate silencing. Notably, transcriptional repression followed localization of *Xist* RNA with little or no delay, suggesting that silencing is directly caused by localization of *Xist*.

It is known that *Xist* produces an RNA of approximately 17 kb and shows intermediate sequence conservation between mouse and humans[8,16]. Several direct repeats have been noted as motifs that could mediate some functions of *Xist*, but experimental evidence of the functionality of any sequences of *Xist* is so far lacking. We therefore carried out a deletion analysis of mouse *Xist* RNA to unequivocally define sequences required for localization and silencing.

Results

Homing single-copy *Xist* transgenes to the *Hprt* locus

Xist expression in undifferentiated ES cells and early-differentiated cells leads to chromosomal transcriptional repression[13]. In differentiated cell types, by contrast, *Xist* is no longer able to initiate silencing[13,17]. ES cells are thus a useful cell-culture system for studying initiation of *Xist* RNA–mediated gene silencing. To explore the function of *Xist* sequences, we used a transgene-based approach. Comparison of different *Xist* mutations was facilitated by the integration of the constructs at a predetermined

Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA. Correspondence should be addressed to R.J. (e-mail: jaenisch@wi.mit.edu).

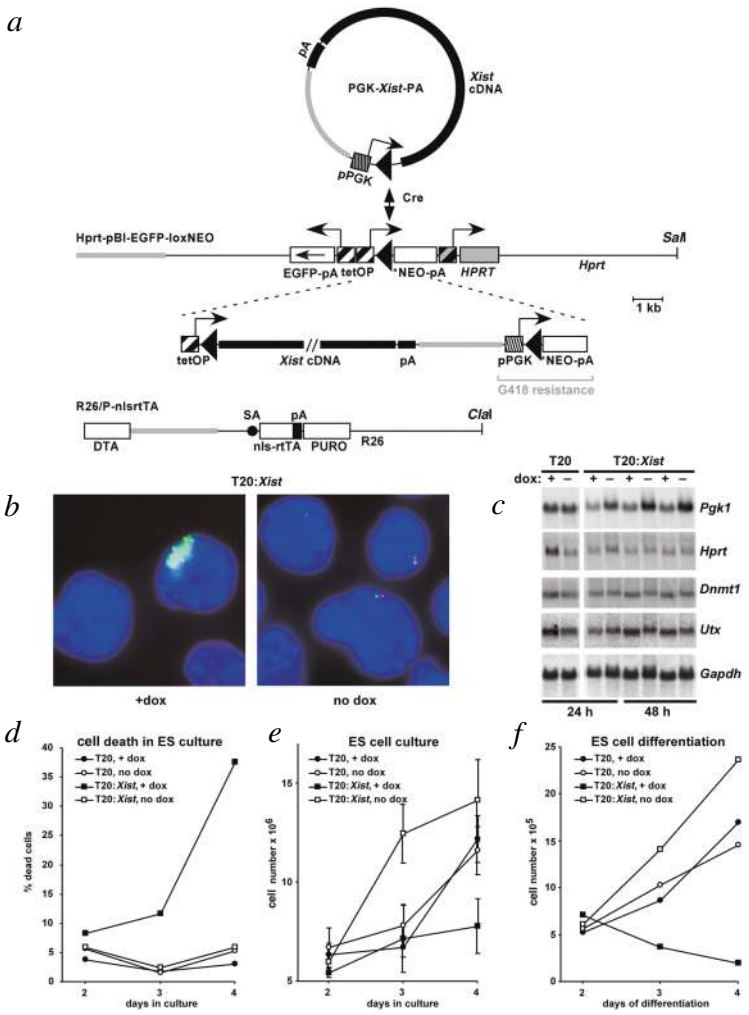


Fig. 1 Homing of single-copy *Xist* cDNA transgenes to the *Hprt* locus. ***a***, Schematic representations of the *Hprt* targeting vector Hprt-pBI-EGFP-lox-neo, the *Xist* cDNA construct PGK-*Xist*-PA and the targeting vector used to introduce the tetracycline-regulated transactivator (nls-rtTA) into the *ROSA26* targeting vector R26/P-nlsrtTA. Cre-mediated integration of the *Xist* cDNA plasmid into the homing site in the *Hprt* locus restores a *PGK-neo*^{r}-pA gene that confers resistance to G418. Black triangles represent *lox*P sites; the inducible promoter is shown as a hatched box. Vector sequences are shown as a gray line, the human *Hprt* promoter as a gray hatched box and exon 1 as a gray box. DTA, diphtheria toxin A chain counter selection marker; SA, splice acceptor. ***b***, Analysis of *Xist* (green) and *Pgk1* (red) expression in T20:*Xist* cells grown in the presence or absence of 1 μg ml^{-1} doxycycline (dox) for 24 h by two-color RNA FISH. ***c***, Northern-blot analysis of expression of *Pgk1* and *Hprt* in T20:*Xist* and control T20 ES cells grown in the presence or absence of 1 μg ml^{-1} dox for 24 h or 48 h. The autosomal *Gapdh* and *Dnmt1* as well as *Utx*, which escapes X inactivation in mice, are included as controls. Note that because of the close proximity to the inducible promoter, *Hprt* is activated in the control T20 cells upon induction with dox. ***d***, Cell death was measured in ES cell cultures of T20:*Xist* and control T20 ES cells grown in the presence or absence of 1 μg ml^{-1} dox. Dead cells in the culture supernatant were counted after trypan blue staining and normalized to the total number of cells in culture. ***e***, The number of cells in ES cell cultures of T20:*Xist* and control T20 ES cells grown in the presence or absence of 1 μg ml^{-1} dox were determined. ***f***, The number of cells in cultures of T20:*Xist* and control T20 ES cells in the presence or absence of 1 μg ml^{-1} doxycycline (dox) during retinoic acid–induced differentiation were determined by cell counts.

homing site. We selected the X-linked *Hprt* locus as a transgene integration site for several reasons: (i) it allowed us to study the mutant *Xist* RNAs in the native X-chromosomal context (ii) inactivation of the single male X chromosome led to cell death owing to silencing of X-linked genes[13] and (iii) the *Hprt* locus is favorable for the expression of transgenes[18].

To analyze a large number of *Xist* mutations, we used a transgene integration strategy mediated by site-specific recombination[19]. Transgenes were integrated into a homing site containing a *lox*P site in the presence of Cre recombinase. This restored an antibiotic resistance marker, which facilitated efficient selection of transgenic cells. To introduce the transgene homing site into the *Hprt* locus, we generated the targeting vector Hprt-pBI-EGFP-loxNEO containing a bidirectional, tetracycline-inducible promoter[20] for expression of a control gene, *EGFP* (encoding enhanced green fluorescent protein), and the *Xist* transgenes introduced later by site-specific recombination (Fig. 1*a*). We introduced the nlsrtTA cDNA[21], encoding the tetracycline-inducible transactivator, into the ubiquitously expressed *ROSA26* locus[22] of Hprt-deficient E14-Tg5 male mouse ES cells[18] by gene targeting using the R26/P-nlsrtTA targeting vector. This targeting vector was similar to the R26/N-nlsrtTA construct[13] but contained a puromycin-resistance cassette. We inserted the transgene homing site into the resulting E14-nlsrTA-

Table 1• *Xist* and *Pgk1* expression

		Pgk1 expression			*Xist* expression						
Cell line	dox	Pgk1		total	clusters		point		no signal		total
Xist	+	53	23%	227	93	60%	52	34%	10	6%	155
	–	173	**83%**	209							
ΔSSX	+	222	78%	285	60	56%	41	38%	6	6%	107
	–	238	**81%**	195							
ΔEvN:XCR 7.5× sense	+	55	23%	242	86	67%	39	30%	3	2%	128
	–	148	**74%**	199							
ΔEvN:XCR 7.5× AS	+	179	77%	232	83	56%	60	41%	5	3%	148
	–	189	**77%**	247							
control	+	190	83%	229	0	0%	133	92%	11	8%	144
	–	145	**78%**	186							

Expression of *Xist* and *Pgk1* was assayed using RNA FISH in undifferentiated transgenic ES cells in the absence or presence of 1 μg ml^{-1} doxycycline (dox) for 24 h. *Pgk1* was identified as a dot-like transcription focus. *Xist* expression was scored into three classes: clusters, indicating stable large *Xist* accumulation similar to somatic female cells; point, indicating a point-like transcription focus seen in undifferentiated wildtype ES cells; and no signal, indicating that no *Xist* expression was observed. The number of nuclei for each category is given, and the percentage was calculated. Control cells are the T20 ES cell line without the *Xist* transgene.

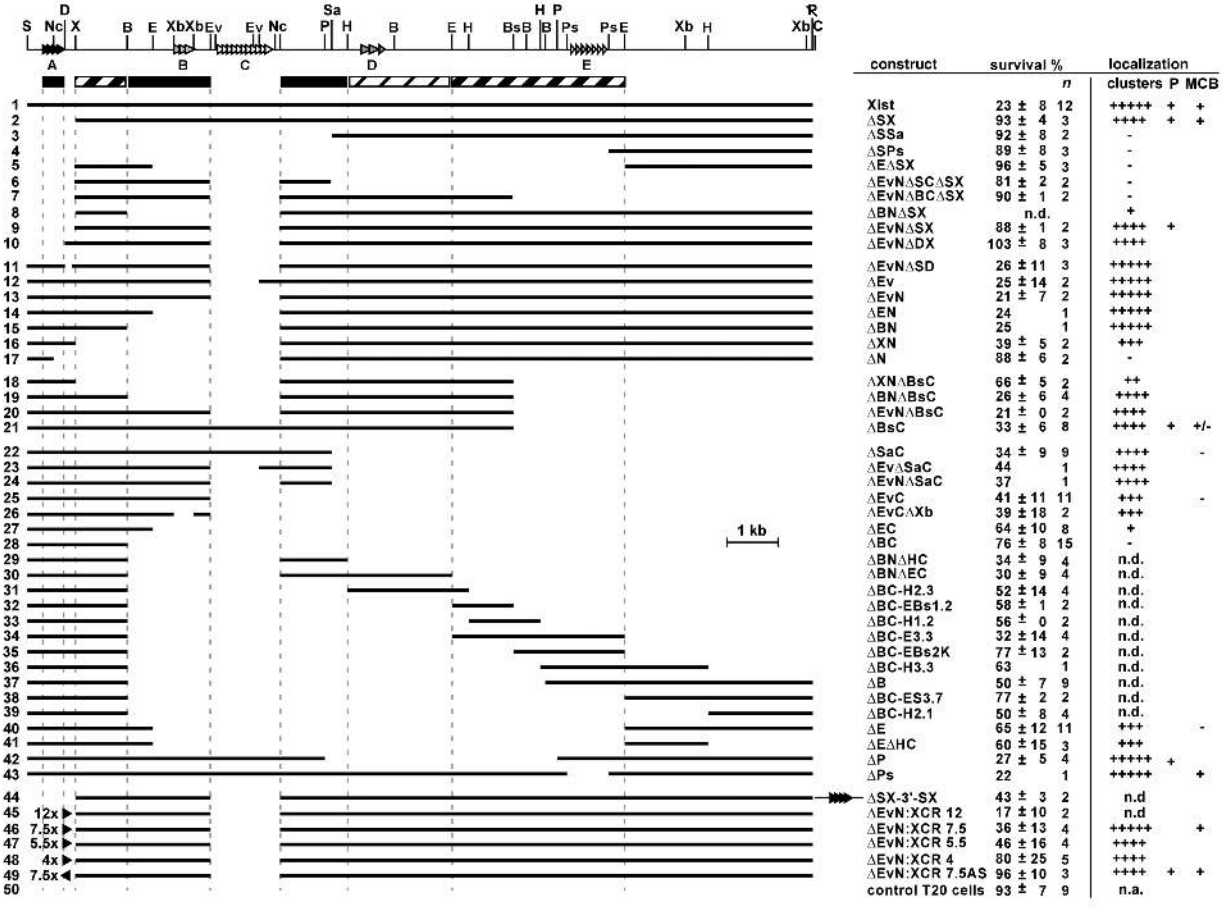
7 ES cells using the Hprt-pBI-EGFP-loxNEO vector, which restored Hprt activity by introducing the human *HPRT* promoter and exon 1 in targeted cells, allowing growth in HAT medium[18]. We verified induction of the control gene *EGFP* in the presence of doxycycline in the resulting T20 ES cells by FACS analysis and fluorescence microscopy (data not shown).

The transgene homing site in the *Hprt* locus contained a *lox*P site linked to a truncated neomycin-resistance gene (*neo*^{r}) lacking a promoter and translation initiation codon. We integrated the *Xist* cDNA transgene into the homing site by co-electroporation of plasmid Pgk-*Xist*-PA and a Cre expression plasmid (Fig. 1*a*). Integration of the plasmid by site-specific recombination of the *lox*P site linked to a translation initiation codon and a *Pgk1* promoter restored the *neo*^{r} gene, conferring resistance to G418. A single copy of the *Xist* cDNA transgene was thus integrated under the control of the inducible promoter. This was confirmed by Southern-blot analysis (data not shown). We induced *Xist* expression in the resulting T20:Xist cells by adding doxycycline (Fig. 1*b*). Transgenic *Xist* expression caused transcriptional repression of *Pgk1* and *Hprt* but not of the control *Dnmt1*, *Utx* and *Gapdh*, as shown by northern-blot analysis (Fig. 1*c*). RNA FISH confirmed that *Pgk1* was significantly repressed after 24 h of induction of transgenic *Xist* expression (Table 1). *Xist*-mediated repression of the single male X caused cell death in undifferentiated ES cells (Fig. 1*d*) and reduced the cell numbers in undifferentiated and differentiating T20:*Xist* cultures as compared with noninduced cultures (Fig. 1*e*,*f*). This confirmed our previous observation that expression of *Xist* from an X-chromosomal integration of the inducible *Xist* cDNA transgene leads to cell death[13].

Silencing requires a conserved 5′ element of *Xist*

We generated a panel of *Xist* constructs containing specific deletions (Fig. 2). We introduced the constructs into T20 ES cells by Cre-mediated recombination as outlined before, and isolated ES cell lines carrying a single integration of the transgenes in the *Hprt* locus. We studied expression of the mutant *Xist* RNAs in ES cells by RNA FISH (Fig. 3*a*) and northern-blot analysis after inducing expression with doxycycline for 24 h (Fig. 3*c*).

To assay the effect of expression of different *Xist* mutations on chromosomal silencing, we measured cell survival. ES cells were differentiated for five days in the presence and absence of doxycycline in triplicate cultures. The cell number was then determined by spectroscopic absorbance measurement. We used as controls cultures differentiated in the absence of doxycycline, defining their



construct	survival %	n	localization clusters	P	MCB
Xist	23 ± 8	12	+++++	+	+
ΔSX	93 ± 4	3	++++	+	+
ΔSSa	92 ± 8	2	-		
ΔSPs	89 ± 8	3	-		
ΔEΔSX	96 ± 5	3	-		
ΔEvNΔSCΔSX	81 ± 2	2	-		
ΔEvNΔBCΔSX	90 ± 1	2	-		
ΔBNΔSX	n.d.		+		
ΔEvNΔSX	88 ± 1	2	++++	+	
ΔEvNΔDX	103 ± 8	3	++++		
ΔEvNΔSD	26 ± 11	3	+++++		
ΔEv	25 ± 14	2	+++++		
ΔEvN	21 ± 7	2	+++++		
ΔEN	24	1	+++++		
ΔBN	25	1	+++++		
ΔXN	39 ± 5	2	+++		
ΔN	88 ± 6	2	-		
ΔXNΔBsC	66 ± 5	2	++		
ΔBNΔBsC	26 ± 6	4	++++		
ΔEvNΔBsC	21 ± 0	2	++++		
ΔBsC	33 ± 6	8	++++	+	+/-
ΔSaC	34 ± 9	9	++++		-
ΔEvΔSaC	44	1	++++		
ΔEvNΔSaC	37	1	++++		
ΔEvC	41 ± 11	11	+++		-
ΔEvCΔXb	39 ± 18	2	+++		
ΔEC	64 ± 10	8	+		
ΔBC	76 ± 8	15	-		
ΔBNΔHC	34 ± 9	4	n.d.		
ΔBNΔEC	30 ± 9	4	n.d.		
ΔBC-H2.3	52 ± 14	4	n.d.		
ΔBC-EBs1.2	58 ± 1	2	n.d.		
ΔBC-H1.2	56 ± 0	2	n.d.		
ΔBC-E3.3	32 ± 14	4	n.d.		
ΔBC-EBs2K	77 ± 13	2	n.d.		
ΔBC-H3.3	63	1	n.d.		
ΔB	50 ± 7	9	n.d.		
ΔBC-ES3.7	77 ± 2	2	n.d.		
ΔBC-H2.1	50 ± 8	4	n.d.		
ΔE	65 ± 12	11	+++		-
ΔEΔHC	60 ± 15	3	+++		
ΔP	27 ± 5	4	+++++	+	
ΔPs	22	1	+++++		+
ΔSX-3′-SX	43 ± 3	2	n.d		
ΔEvN:XCR 12	17 ± 10	2	n.d		
ΔEvN:XCR 7.5	36 ± 13	4	+++++		+
ΔEvN:XCR 5.5	46 ± 16	4	++++		
ΔEvN:XCR 4	80 ± 25	5	++++		
ΔEvN:XCR 7.5AS	96 ± 10	3	++++	+	+
control T20 cells	93 ± 7	9	n.a.		

Fig. 2 Summary of *Xist* mutations. Schematic representation of *Xist* RNA (top) and the mutations generated in this study (rows 1–50). Repeat elements A–E of *Xist* and restriction sites are indicated. The boxes below the map summarize the regions localizing *Xist* RNA to chromatin (black boxes represent the greatest activity). *Xist* sequences are represented by solid lines, and synthetic repeats are indicated by a triangle (pointing to the right for sense and to the left for antisense orientation) preceded by the number of repeat units. The percentage of cells (± s.d.; *n* is the number of measurements in triplicates; typically 2–4 independent clones were analyzed for each construct) surviving in differentiating cultures in the presence as compared to the absence of dox is given as a measure of silencing (20% corresponds to full-length *Xist* RNA, 90% corresponds to no silencing activity). The localization pattern of the mutant *Xist* RNAs is indicated on a scale. '+++++' indicates wildtype *Xist* pattern; '–' indicates that RNA does not form nuclear clusters; and '+' in column P indicates that painted metaphase chromosomes were observed (see Fig. 3). The formation of MCBs in differentiated cells is given. S, *Sac*I; Nc, *Nco*I; D, *Dra*I; X, *Xho*I; B, *Bam*HI; E, *Eco*RI; Xb, *Xba*I; Ev, *Eco*RV; P, *Pvu*II; Sa, *Sac*I; H, *Hin*dIII; Bs, *Bss*HII; Ps, *Pst*I; C, *Cla*I; n.d., no data was collected; n.a., not applicable.

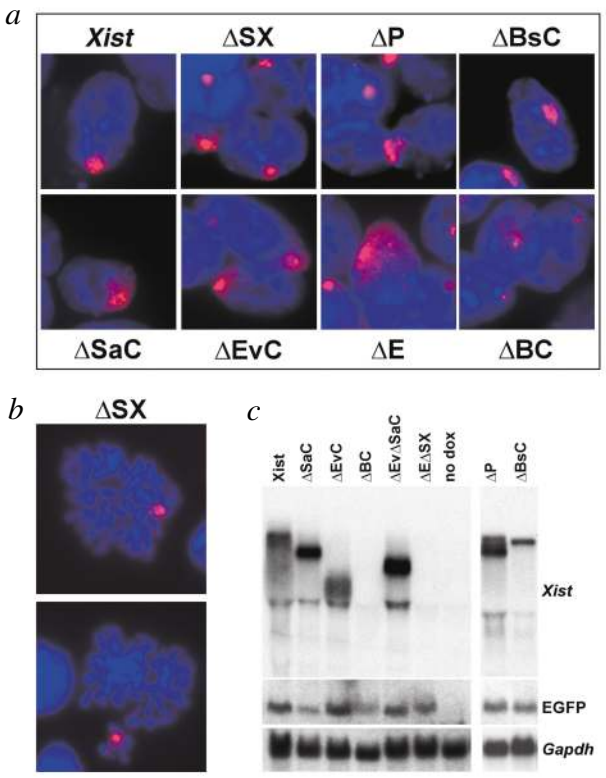


Fig. 3 Analysis of expression and localization of mutated *Xist* RNAs. ***a***, To study localization of *Xist* RNA in interphase nuclei of transgenic ES cells by RNA FISH, a Cy3-labeled probe (red) spanning the entire *Xist* cDNA was used and nuclei were counterstained with DAPI (blue). ***b***, Localization of *Xist* RNA lacking the 5′ repeat element of *Xist* (ΔSX) on chromosomes in metaphase spreads of transgenic ES cells grown in the presence of doxycycline (dox). ***c***, Northern-blot analysis of *Xist* RNA and GFP expression in different transgenic ES cells. Non-induced cells (no dox) were used as a negative control and *Gapdh* was used to control for loading.

cell number as 100%. The number of cells surviving in the presence of doxycycline was expressed relative to this value: 23% ± 8% of T20:*Xist* ES cells expressing the full-length *Xist* RNA survived (Fig. 1*f* and Fig. 2, row 1) compared with 93% ± 7% of control T20 cells, which contained no *Xist* transgene (Fig. 2, row 50). Thus, the range for measurements of silencing activity using this assay was between 20% (for full-length *Xist*) and 90% (for no *Xist*).

Silencing activity in *Xist* mutations falls over a broad and continuous range. We isolated several mutations with intermediate (40–60%) cell survival (Fig. 2). Large deletions in the middle and 3′ end of *Xist* did not affect *Xist* function (Fig. 2, rows 15,21,42), which suggests that these sequences are either nonessential or compensated for by redundant sequences. Notably, four of the conserved repeats (Fig. 2, repeats B–E) identified in human and mouse *XIST* were not required for function (Fig. 2, rows 13,14,19,42,43). Deletion of 0.9 kb of the 5′ end of *Xist* containing the highly conserved A repeat[8], however, completely abolished silencing activity (ΔSX construct; Fig. 2, row 2).

Localizing *Xist* RNA by FISH analysis showed that full-length transgenic *Xist* accumulated in dense clusters in the nucleus, similar to the pattern of *Xist* expression observed in female somatic cells (Fig. 3*a*). *Xist* mutations with an intermediate silencing function generally formed more diffuse clusters (Figs 2 and 3*a*), with an inverse correlation between the size of the deletion and function with respect to localization and silencing. Several *Xist* mutations that lacked the A repeat and contained additional deletions did not lead to stable RNA, as seen by northern-blot analysis (Fig. 2, rows 3–7). Only a pinpoint signal was observed by RNA FISH for these constructs. However, the ΔSX construct, lacking 900 bp of *Xist* 5′ sequences, expressed a transcript that localized in strong clusters in interphase nuclei and associated with metaphase chromosomes (Figs 2 and 3*a*,*b*). Despite correct localization of the ΔSX RNA, cell survival was not affected by ΔSX expression (Fig. 2). RNA FISH showed that *Pgk1* expression was not reduced in ES cells when ΔSX expression was induced (Table 1). These results indicate that the ΔSX RNA contains elements sufficient for proper localization but not for silencing.

To further characterize the regions of *Xist* that determine localization to chromatin, we fused small parts of *Xist* to a 2.3-kb 5′ sequence of *Xist* that was minimally active (Fig. 2, rows 28–41). We identified three spatially separated domains within *Xist* that contribute to RNA localization (Fig. 2). Two of these domains were sufficient for *Xist* function when the 5′ repeat element was present (Fig. 2, rows 11–28 and 42). In the absence of the 5′

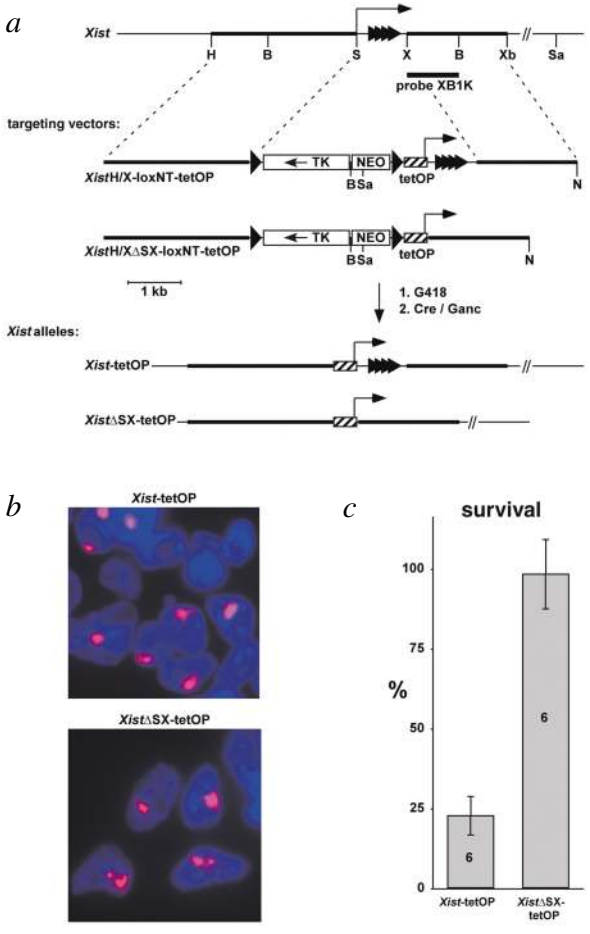


Fig. 4 Targeting the endogenous *Xist* locus in male mouse J1 ES cells. ***a***, The targeting vectors *Xist*H/X-loxNT-tetOP and *Xist*H/XΔSX-loxNT-tetOP, the relevant part of the *Xist* locus and the structure of the targeted alleles are shown. H, *Hin*dIII; B, *Bam*HI; S, *Sac*I; X, *Xho*I; Xb, *Xba*I; Sa, *Sac*I. Black triangles represent *lox*P sites and the tetracycline responsive promoter is shown as a hatched box. The 5′ repeat element of *Xist* is shown as a series of black triangles. ***b***, *Xist* expression (red) from the *Xist*H/X-tetOP and the *Xist*H/XΔSX-tetOP alleles was analyzed by RNA FISH in J1 ES cells expressing the nlsrtTA transactivator from the *ROSA26* locus. Nuclei were counter-stained with DAPI (blue). ***c***, Cell survival in cultures of cells containing the *Xist*H/X-tetOP and the *Xist*H/XΔSX-tetOP allele differentiated in the presence compared to the absence of 1 μg ml^{-1} doxycycline (dox). The error bar spans 2 s.d.

repeat element, all three domains were required for proper localization of the transcript (Fig. 2, rows 2–10). This suggested that the *Xist* 5′ element contributed to the localization of the RNA in addition to its role in transcriptional silencing.

To further exclude the possibility that silencing mediated by different *Xist* constructs was not a result of repression or activation of the tetracycline-inducible promoter by the introduced sequences, we analyzed the expression of the control gene *EGFP* driven by the same bidirectional promoter as the *Xist* cDNA. We did not observe a significant difference in EGFP expression between clones of ES cells carrying different *Xist* transgenes by northern-blot analysis (Fig. 3*c*) or by FACS analysis (data not shown). We conclude that the transcription level of all *Xist* transgenes was equal.

The 5′ element is required for silencing by the endogenous *Xist*

To ascertain the relevance of the results obtained with the transgenic system to the endogenous *Xist* gene, we introduced the tetracycline-inducible promoter into the *Xist* locus by gene targeting in male J1 ES cells (Fig. 4*a*). We generated the *Xist*-tetOP allele by inserting the inducible promoter at the transcription initiation site of *Xist*. We generated a second allele, *Xist*-ΔSX-tetOP, that also harbors a deletion of the 5′ repeat element of *Xist*. The selection markers were removed from the targeted alleles by Cre-mediated excision to avoid misregulation of *Xist* expression. We subsequently introduced the nlsrtTA cDNA into the *ROSA26* locus of these targeted cells using the R26/P-nlsrtTA targeting vector, allowing inducible expression of the endogenous *Xist* gene. *Xist* RNA was expressed in ES cells from the *Xist*-tetOP and *Xist*-ΔSX-tetOP alleles after the addition of doxycycline, as seen by northern blot analysis and RNA FISH (Fig. 4*b* and data not shown). *Xist* RNA was stable in ES cells even when expressed from the endogenous locus, confirming our previous results with *Xist* expression from ectopically integrated transgenes[13]. Thus, the presence of genomic sequences, introns and the *Tsix*[23] transcript, which has been shown to regulate *Xist*, did not lead to destabilization of *Xist* RNA in ES cells when *Xist* was expressed under control of the inducible promoter.

We analyzed the effect of *Xist* expression on cell survival. Expression of the full-length *Xist*-tetOP allele resulted in cell death (Fig. 4*c*); however, induction of *Xist* expression from the *Xist*-ΔSX-tetOP allele, which lacks the critical 5′ repeat elements of *Xist*, did not cause cell death. These data are consistent with our results for the *Xist* cDNA transgenes in the *Hprt* locus and suggest that the results obtained with transgenic *Xist* RNA are relevant for the endogenous *Xist* gene.

Localization of macroH2A1 can occur independently of silencing

One particularly interesting chromosomal modification of the inactive X chromosome is the accumulation of histone macroH2A1 (ref. 24). This histone contains an H2A histone core and a large C-terminal domain of unknown function. In ES cells, macroH2A1 mainly localizes to the centrosome and re-localizes to the inactive X chromosome upon differentiation, where it forms a macrochromatin body (MCB)[25]. This event occurs relatively late in X inactivation, suggesting a role for macroH2A1 in the maintenance of the inactive state rather than in the initiation of silencing. *Xist* RNA has been shown to be required for localization of macroH2A1 to the inactive X chromosome[26]. Deletion of *Xist* by Cre-mediated excision from an inactive X chromosome in mouse female fibroblasts leads to the disappearance of macroH2A1 from the inactive X, despite the maintenance of chromosomal silencing. MacroH2A1 localizes to autosomes carrying an inducible *Xist* cDNA transgene even if *Xist* is expressed at a time when transcriptional silencing can no longer be initiated[15].

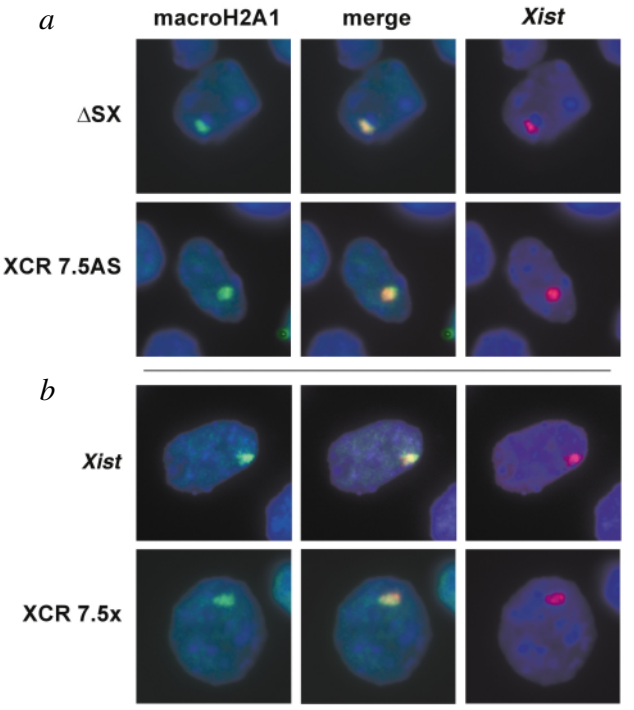


Fig. 5 Transgenic *Xist* expression leads to MCB formation. ***a***, Combined *Xist* RNA FISH (red) and macroH2A1 immunofluorescence (green) reveals MCBs in ES cells that were differentiated for 6 d in the presence of doxycycline to induce expression of ΔSX RNA or XCR 7.5 AS RNA, which lack the 5′ repeat of *Xist* or have an antisense synthetic repeat, respectively. Nuclei were counterstained with DAPI. ***b***, Combined *Xist* RNA FISH (red) and macroH2A1 immunofluorescence (green) on cells either expressing full-length transgenic *Xist* or having the 5′ element replaced by a synthetic repeat (XCR 7.5×). The cells were differentiated for 2 d in the absence of doxycycline to avoid cell death; *Xist* expression was subsequently turned on by addition of doxycycline. Analysis was carried out on day 6 of differentiation.

To determine whether a single copy of the *Xist* cDNA transgene in the *Hprt* locus would lead to the formation of MCBs, we induced *Xist* expression in ES cells that had been differentiated for two days in the presence of retinoic acid and cultured the cells for four more days in differentiation medium. This late induction of *Xist* expression did not lead to silencing of the X chromosome, allowing us to circumvent the cell lethality. When these cells were analyzed by combined *Xist* RNA FISH and macroH2A1 immunofluorescence, MCBs had formed in 26% of the cells whose nuclei also showed *Xist* expression (Fig. 5*b*). This was similar to the proportion of cells with MCBs (30%) observed in differentiated female cells using this method[15]. No MCBs were observed when *Xist* expression was not induced in control T20 cells.

We further analyzed the ability of different *Xist* mutations to mediate macroH2A1 localization. *Xist* RNA containing a 4-kb deletion in the middle (ΔPv mutation; Fig. 2, row 42) caused MCBs in 4% of the cells that also showed *Xist* expression, and a mutation that had been truncated from the 3′ end of *Xist* (ΔBsC mutation; Fig. 2, row 21) formed less prominent MCBs in 1% of *Xist*-expressing cells. Both the ΔPv and the ΔBsC mutations produced RNA that localized to the chromatin of metaphase spreads and mediated efficient chromosomal silencing (Fig. 2, rows 21 and 42). This may indicate a specific role for the 3′ region of *Xist* in macroH2A1 localization. Notably, the ΔSX mutation localized correctly to the chromosome but did not mediate chromosomal silencing. We found that the ΔSX RNA

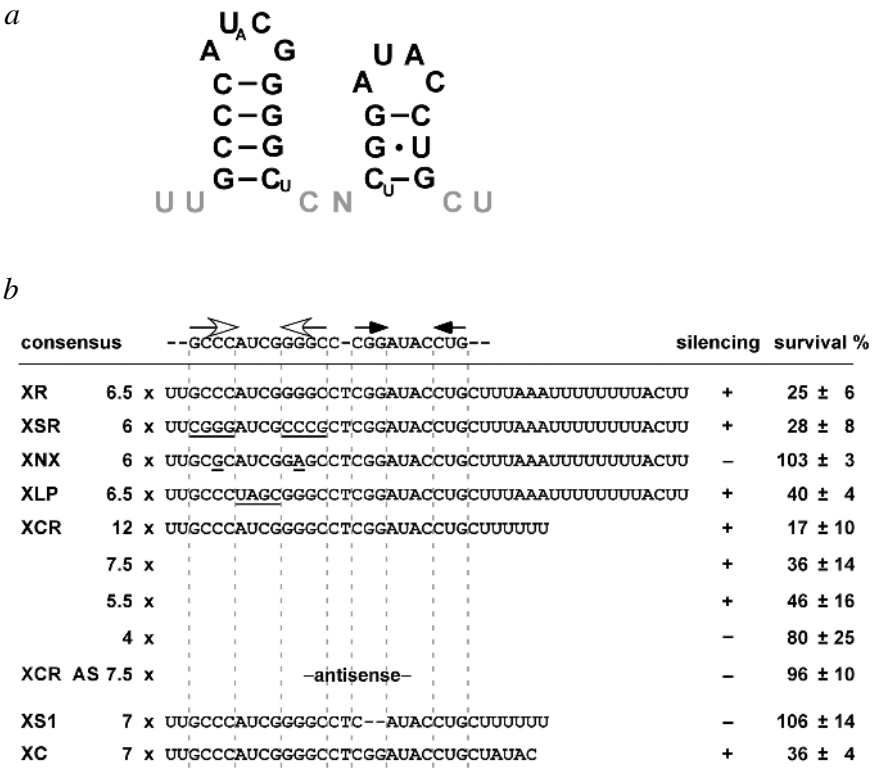


Fig. 6 Analysis of the 5′ repeat element of *Xist*. ***a***, RNA secondary structure prediction of the repeat sequence. Nonconserved sequence is in gray. Small characters indicate less frequent base changes. GC base pairs are represented by a line, GU base pairs by a dot. ***b***, Sequences of the synthesized repeat monomers and mutations along with the number of repeats analyzed are given. Arrows indicate sequences involved in stem loop formation. The ability of the sequences to complement the endogenous *Xist* 5′ element for silencing is indicated (+ or –), and cell survival upon *Xist* induction in differentiating cells is given.

was able to efficiently recruit macroH2A1 to the X chromosome to form MCBs in 17% of *Xist*-expressing cells (Fig. 5*a*). Thus, macroH2A1 accumulation and silencing require different domains of *Xist*. This finding also supports the notion that *Xist* expression is sufficient for the formation of MCBs in the absence of transcriptional silencing.

Two stem loops in the 5′ repeat of *Xist* RNA mediate silencing

Our results show that *Xist* RNA localization is not sufficient for silencing and that the conserved 5′ element of *Xist* is required to initiate transcriptional repression. To assess the significance of the position of these sequences on the 5′ end, we moved the 900-bp 5′ end of *Xist* to the 3′ end. The resulting construct, ΔSX-3′-SX (Fig. 2, row 44), was active in silencing, indicating that the sequence was functionally independent of its position in the RNA. Computational structure prediction indicated that the conserved direct repeats might fold into a secondary structure comprised of two stem loops (Fig. 6*a*). The 5′ region of *Xist* contains 7.5 repeat units, comprising 8 of the first and 7 of the second smaller stem loops. These repeats are separated by AU-rich spacers whose sequence is not conserved. We synthesized a minimal version of these repeats by concatemerization of oligonucleotides and inserted them into *Xist* cDNA constructs from which the first 900 bp were deleted (Fig. 6*b*, construct XR). Replacing the *Xist* 5′ element by the synthetic repeat complemented the sequence and restored silencing (Fig. 6*b*). To test whether the AU- rich spacer is required for the function of the repeat, we shortened the sequence to only eight uracil residues (Fig. 6*b*, construct XCR). The resulting repeat was still functional and mediated silencing. We further mutated the uracil residues in the spacer between the repeat monomers (Fig. 6*b*, construct XC) and found that the sequence of the spacer has no consequence for silencing.

To address the effect of different numbers of repeat units, we created constructs that contained 4, 5.5, 7.5 and 12 repeat units. The constructs of 12 and 7.5 repeat units resulted in silencing at a level close to that of the wild type, whereas constructs of 5.5 units were significantly weaker and constructs of 4.4 repeat units were nearly inactive in silencing (Fig. 2, rows 45–48). If constructs of 7.5 repeat units were inserted in antisense orientation (construct XCR 7.5 AS), the resulting construct did not cause silencing, as determined by cell survival (Fig. 2, row 49) and RNA FISH for *Pgk1* expression (Table 1). The localization of the RNA was normal and caused formation of MCBs in differentiated ES cells (Fig. 5*a*). These results demonstrate the specificity of the assay.

To address the significance of the stem loops, we designed several specific mutations of the repeats (Fig. 6*b*). Inversion of the base pairs in the first stem loop (Fig. 6*b*, construct XSR) did not abolish the function of the repeat. If, however, the stem loop was destroyed by two base changes predicted to disrupt base pairing (Fig. 6*b*, construct XNX), silencing was not initiated. This indicated that the first stem loop is essential but the sequence of its stem is not crucial for function. Changing the bases in the loop of the first stem loop (Fig. 6*b*, construct XLP) did not greatly affect the function of the repeats, which indicated that the sequence of the loop is not crucial. Removal of the second stem loop (Fig. 6*b*, construct XS1) resulted in loss of function, indicating that both stem loops of the repeat monomers are essential for silencing. Taken together, these data support the structural model of the 5′ repeat of *Xist* that is based on two stem loops, which constitute the motif recognized by a putative binding protein. The less-conserved spacers are probably required for both proper RNA folding of the individual repeat modules and the access of putative binding factors.

Discussion

We have identified specific sequences of *Xist* that mediate silencing and chromosomal localization and have separated these functions to different domains of the RNA. In contrast to the defined sequence requirement for silencing, regions of *Xist* that mediate localization do not have any common motifs. These sequences are scattered throughout *Xist* and are functionally redundant. This parallels the relatively poor conservation among species[8,16,27]. A likely explanation for the absence of obvious sequence motifs is the presence of many low-affinity binding sites that are coopera-

tively bound by factors, a situation comparable to the binding of rev to a motif of human immunodeficiency virus RNA[28]. Localization of *Xist* RNA might be influenced by chromatin and by proximity to other ribonucleoprotein complexes. The stability of the complexes would thus be influenced by the local concentration of *Xist* RNA. This may provide a mechanistic basis for the observed spreading of the RNA in *cis* along the chromosome, although specific transport cannot be ruled out as an alternative.

Sequences mediating the localization of *Xist* were sufficient to recruit macroH2A1 to the chromosome, suggesting that MCB formation can occur by a pathway independent of silencing. This is consistent with previous reports that silencing and macroH2A1 recruitment can be separated at various stages of development or differentiation[15,26]. In addition, localization is also required for the stability of *Xist* RNA. *Xist* constructs carrying large deletions did not give rise to stable transcripts. Despite our extensive analysis, we did not observe *Xist* mutants that did not localize but were stable. Thus, Xist RNA is probably subject to degradation when it does not attach to chromatin. This is important for the *cis*-limited spreading of *Xist* RNA, as diffusible RNA that fails to bind to chromatin near its site of synthesis is prevented from binding to neighboring chromosomes because of rapid degradation.

Correct localization of *Xist* is required for *cis*-limited silencing, in agreement with a recent study using PNA mapping[29]. In contrast to the poorly defined sequence requirements for localization, we clearly defined the 5′ repeat A of *Xist* as an essential motif responsible for silencing. Computational structure prediction suggests that this sequence folds into two stem loops reiterated 7.5 times, which might represent binding sites for interacting factors. We further generated specific mutations in the stems that disturb silencing, suggesting that the repeats can function as binding sites for factors involved in setting up heterochromatin. The 5′ sequences of *Xist* have been suggested to mediate transcriptional repression[30] in a transient transgenic system. In contrast to our findings, that study concluded that the full sequence, including the spacer region of the A repeat, is required for repression of the reporter constructs that are not chromosomally integrated. Future work is needed to identify the factors that specifically interact with these sequences. The high sequence conservation between mammalian species and invariant 5′ location merit further study of their addition to *Xist* RNA in a modular way during the evolution of mammals. Although no relatives of *Xist* have been identified in non-mammalian species, the roX RNAs in flies have been shown to associate with chromatin of the dosage-compensated X chromosome in males[31] and could be regarded as a potential distant ancestor of *Xist*.

Methods

Plasmid construction. To create the targeting vector Hprt-pBI-EGFP-loxNEO, we ligated a 2-kb *Xho*I–*Bam*HI fragment from plasmid pSF1 (Gibco) to the *Sac*I/*Bam*HI-cut vector tetOP-H/X, containing the tetracycline-inducible promoter[13]. A 2.5-kb *Xho*I–*Bam*HI fragment was excised and ligated into *Pvu*II-cut vector pBIEGFP (Clontech), in which the SV40 polyadenylation signal had been replaced by the *Pgk1* polyadenylation signal. We then inserted the 4.7-kb *Bgl*II–*Mlu*I cassette into the *Not*I site of the *Hprt* targeting vector[18]. To create targeting vector R26/P-nlsrtTA, we cloned a 1.6-kb *Eco*RI–*Sca*I fragment containing the Pgk-*pur*^{r} gene (*Cla*I site destroyed) into the *Xba*I site of plasmid ROSA26-1 (*Sal*I site destroyed; ref. 32), giving plasmid R26/P. We inserted the nlsrtTA cDNA with a splice acceptor and polyadenylation signal[13] into the *Sal*I site, resulting in R26/P-nlsrtTA. To create plasmid PGK-*Xist*-PA, we replaced the 730-bp *Xba*I fragment from plasmid pBS226, containing the CMV promoter, by the 500-bp *Eco*RI–*Xho*I promoter fragment of PGK-*pur*^{r}[13], giving plasmid pBS226/PGK. The 540-bp *Sal*I–*Hin*dIII fragment was inserted into the *Xho*I–*Sac*I sites of pUHD19-1 (ref. 33). We ligated the *Pvu*I–*Sac*II fragment to the *Sac*II/*Pvu*I-cut plasmid tetOP-*Xist*[13], giving plasmid PGK-*Xist*-PA. We generated deletion mutants using the restriction sites indicated in Fig. 2. Synthetic versions of *Xist* repeat A were generated by concatemerization of overlapping oligonucleotides (Research Genetics; see Fig. 6). The oligonucleotides were phosphorylated using T4 polynucleotide kinase, denatured at 95 °C for 5 min, annealed at 65 °C, 50 °C and 37 °C for 15 min each and then filled in with Klenow DNA polymerase. We purified ligation products of 300–400 bp from 2% agarose gels and cloned them into the *Eco*RV site of pBluescript (Stratagene). After sequence verification, the fragments were excised with *Sac*II and *Xho*I and cloned into the *Sac*II/*Xho*I–cut plasmid PGK-*Xist*-ΔEvN. To create targeting vectors *Xist*-tetOP and *Xist*-ΔSX-tetOP, we subcloned a 5.4-kb *Hin*dIII–*Xba*I fragment, containing *Xist* upstream sequences and part of exon 1, into pBluescript (*Sac*II and *Xho*I sites destroyed), giving plasmid *Xist*H/X. A 600-bp *Xho*I–*Sac*II fragment containing the tetracycline-inducible promoter was inserted into plasmid pGHA, containing the *lox*P site–flanked *neo-TK* cassette[34]. The 3.3-kb *Xba*I fragment was ligated into the *Sac*II- or *Sac*II/*Xho*I-cut plasmid *Xist*H/X, giving *Xist*-tetOP and *Xist*-ΔSX, respectively.

ES cell culture and generation of transgenic cell lines. We grew ES cells in high-glucose DMEM (Gibco) supplemented with 15% fetal calf serum, 250 units LIF (Leukemia Inhibitory Factor) per ml, penicillin/streptomycin, 2-mercaptoethanol and nonessential amino acids (Gibco) on irradiated mouse embryonic fibroblast feeders as described[13]. We electroporated E14-Tg5 ES cells[18] with 30 μg *Cla*I-linearized R26/P-nlsrTA targeting vector. After 8 d of selection in ES medium containing 2 μg ml^{-1} puromycin (Sigma), we picked colonies and identified homologous recombination by Southern-blot analysis[30]. Resulting E14-nlsrtTA-7 ES cells were electroporated with 30 μg of *Sal*I-linearized Hprt-pBI-EGFP-loxNEO targeting vector. After selection for 8 d in ES medium containing HAT (Stratagene), we picked colonies and confirmed targeting by Southern blotting[18]. We analyzed inducible EGFP expression in the targeted T20 ES cells by fluorescence microscopy and FACS. To introduce the *Xist* constructs, we electroporated T20 ES cells with 60–100 μg of plasmid PGK-*Xist*-PA (or other constructs) and 30 μg of pMC-Cre[25]. Cells were selected for 12 d in ES medium containing 0.3 g ml^{-1} G418 (Gibco). We picked colonies and confirmed the integrity and copy number of the integration by Southern blotting using the probes XB1K (see Fig. 4*a*) and HP2K (containing a 2-kb *Hin*dIII–*Pst*I fragment of *Xist* exon 7). To measure cell survival in differentiating cultures, we removed feeders by preplating and then differentiated ES cells on gelatin-coated six-well dishes in medium containing all-*trans*-retinoic acid[13] with or without doxycycline for 5 d (three wells contained 1 μg ml^{-1} doxycycline and three wells served as controls). We trypsinized cells and resuspended them in 2 ml medium, and measured the optical density (OD) at 600 nm using a spectrophotometer (Beckman). We set the OD value of the control wells without doxycycline to 100%; the OD of the wells differentiated in the presence of doxycycline was expressed relative to that. For targeting *Xist* in J1, we electroporated ES cells with 30 μg of linearized *Xist*-tetOP or *Xist* ΔSX-tetOP targeting vector. We picked colonies after selection in ES medium containing 0.3 mg ml^{-1} G418. We confirmed targeting by Southern blotting using the probe XB1K (see Fig. 4*a*). We electroporated targeted ES cell lines with 30 μg pMC-Cre plasmid and picked colonies after selection in ES medium containing 2 μM gancyclovir[34].

RNA FISH and immunofluorescence. Cells for RNA FISH were grown on multiwell slides (ROBOZ) or cytospun onto Superfrost Plus slides (VWR). We prepared metaphase spreads after methanol/acetic acid (3:1) fixation and carried out RNA FISH as previously described[13]. We carried out detection of macroH2A1 on cells grown on multiwell slides after RNA FISH as described[15], using a rabbit polyclonal antiserum directed against the nonhistone domain of mouse macroH2A1 (T.P.R., manuscript in preparation) at a dilution of 1:2,000, followed by detection with fluorescein-conjugated goat anti-rabbit antiserum. For MCB counts, we scored at least 400 nuclei from several fields for *Xist* expression and colocalizing macroH2A1 immunostaining (see Fig. 5).

Acknowledgments

We are grateful for the ability to use the microscopes of the W.M. Keck Biological Imaging Facility. This work was supported by grants from the National Institutes of Health and the Max Kade Foundation, and by the Human Frontiers Science Program Organization.

Received 17 September; accepted 21 November 2001.

1. Barr, M.L. & Bertram, E.G. A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature* **163**, 676–677 (1949).
2. Avner, P. & Heard, E. X-chromosome inactivation: counting, choice and initiation. *Nature Rev. Genet*. **2**, 59–67 (2001).
3. Lyon, M.F. Gene action in the X-chromosome of the mouse (*Mus musculus L.*). *Nature* **190**, 372–373 (1961).
4. Russell, L.B. Genetics of mammalian sex chromosomes. *Science*, 1795–1803 (1961).
5. Brown, C.J. & Willard, H.F. Localization of the X inactivation center (XIC) to Xq13. *Cytogenet. Cell. Genet*. **51**, 971 (1989).
6. Brown, C.J. *et al*. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**, 38–44 (1991).
7. Brockdorff, N. *et al*. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **351**, 329–331 (1991).
8. Brown, C.J. *et al*. 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 (1992).
9. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. & Jaenisch, R. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes. Dev.* **11**, 156–166 (1997).
10. Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S. & Brockdorff, N. Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131–137 (1996).
11. Lee, J.T., Strauss, W.M., Dausman, J.A. & Jaenisch, R. A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell* **86**, 83–94 (1996).
12. Herzing, L.B., Romer, J.T., Horn, J.M. & Ashworth, A. Xist has properties of the X-chromosome inactivation centre. *Nature* **386**, 272–275 (1997).
13. Wutz, A. & Jaenisch, R. A shift from reversible to irreversible X inactivation is triggered during ESCell differentiation. *Mol. Cell*. **5**, 695–705 (2000).
14. Marahrens, Y., Loring, J. & Jaenisch, R. Role of the Xist gene in X chromosome choosing. *Cell* **92**, 657–664 (1998).
15. Rasmussen, T.P., Wutz, A., Pehrson, J.R. & Jaenisch, R. Xist expression is sufficient for macroH2A localization. *Chromosoma* **110**, 411–420 (2001).
16. Brockdorff, N. *et al*. 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 (1992).
17. Tinker, A.V. & Brown, C.J. Induction of XIST expression from the human active X chromosome in mouse/human somatic cell hybrids by DNA demethylation. *Nucleic Acids Res*. **26**, 2935–2940 (1998).
18. Bronson, S.K. *et al*. Single-copy transgenic mice with chosen-site integration. *Proc. Natl Acad. Sci. USA* **93**, 9067–9072 (1996).
19. Fukushige, S. & Sauer, B. Genomic targeting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells. *Proc. Natl Acad. Sci. USA* **89**, 7905–7909 (1992).
20. Baron, U., Freundlieb, S., Gossen, M. & Bujard, H. Co-regulation of two gene activities by tetracycline via a bidirectional promoter. *Nucleic Acids Res*. **23**, 3605–3606 (1995).
21. Gossen, M. *et al*. Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**, 1766–1769 (1995).
22. Zambrowicz, B.P. *et al*. Disruption of overlapping transcripts in the ROSA β geo 26 gene trap strain leads to widespread expression of β-galactosidase in mouse embryos and hematopoietic cells. *Proc. Natl Acad. Sci. USA* **94**, 3789–3794 (1997).
23. Lee, J.T., Davidow, L.S. & Warshawsky, D. *Tsix*, a gene antisense to *Xist* at the inactivation centre. *Nature Genet*. **21**, 400–404 (1999).
24. Costanzi, C. & Pehrson, J.R. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**, 599–601 (1998).
25. Rasmussen, T.P., Mastrangelo, M.A., Eden, A., Pehrson, J.R. & Jaenisch, R. Dynamic relocalization of histone MacroH2A1 from centrosomes to inactive X chromosomes during X inactivation. *J. Cell. Biol*. **150**, 1189–1198 (2000).
26. Csankovszki, G., Panning, B., Bates, B., Pehrson, J.R. & Jaenisch, R. Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. *Nature Genet*. **22**, 323–324 (1999).
27. Nesterova, T.B. *et al*. Characterization of the genomic xist locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. *Genome Res*. **11**, 833–849 (2001).
28. Karn, J. *et al*. Control of immunodeficiency virus gene expression by the RNA-binding proteins tat and rev. in *RNA-Protein Interactions* (eds Nagai, K. & Mattaj, I.W.) 192–220 (Oxford University Press, New York, 1994).
29. Beletskii, A., Hong, Y.-K., Pehrson, J., Egholm, M. & Strauss, W.M. PNA interface mapping demonstrates functional domains in the noncoding RNA *Xist*. *Proc. Natl Acad. Sci. USA* **98**, 9215–9220 (2001).
30. Allaman-Pillet, N., Djemai, A., Bonny, C. & Schorderet, D.F. The 5′ repeat elements of the mouse *Xist* gene inhibit the transcription of X-linked genes. *Gene. Expr*. **9**, 93–101 (2000).
31. Kelley, R.L. & Kuroda, M.I. Noncoding RNA genes in dosage compensation and imprinting. *Cell* **103**, 9–12 (2000).
32. Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature Genet*. **21**, 70–71 (1999).
33. Baron, U., Gossen, M. & Bujard, H. Tetracycline-controlled transcription in eukaryotes: novel transactivators with graded transactivation potential. *Nucleic Acids Res*. **25**, 2723–2729 (1997).
34. Wutz, A. *et al*. Non-imprinted Igf2r expression decreases growth and rescues the Tme mutation in mice. *Development* **128**, 1881–1887 (2001).