Disruption of a conserved region of Xist exon 1 impairs Xist RNA localisation and X-linked gene silencing during random and imprinted X chromosome inactivation

Claire E. Senner1,2,*, Tatjana B. Nesterova1, Sara Norton2, Hamlata Dewchand2, Jonathan Godwin1, Winifred Mak2 and Neil Brockdorff1,†

SUMMARY
In XX female mammals a single X chromosome is inactivated early in embryonic development, a process that is required to equalise X-linked gene dosage relative to XY males. X inactivation is regulated by a cis-acting master switch, the Xist locus, the product of which is a large non-coding RNA that coats the chromosome from which it is transcribed, triggering recruitment of chromatin modifying factors that establish and maintain gene silencing chromosome-wide. Chromosome coating and Xist RNA-mediated silencing remain poorly understood, both at the level of RNA sequence determinants and interacting factors. Here, we describe analysis of a novel targeted mutation, XistINV, designed to test the function of a conserved region located in exon 1 of Xist RNA during X inactivation in mouse. We show that XistINV is a strong hypomorphic allele that is appropriately regulated but compromised in its ability to silence X-linked loci in cis. Inheritance of XistINV on the paternal X chromosome results in embryonic lethality due to failure of imprinted X inactivation in extra-embryonic lineages. Female embryos inheriting XistINV on the maternal X chromosome undergo extreme secondary non-random X inactivation, eliminating the majority of cells that express the XistINV allele. Analysis of cells that express XistINV RNA demonstrates reduced association of the mutant RNA to the X chromosome, suggesting that conserved sequences in the inverted region are important for Xist RNA localisation.

KEY WORDS: X inactivation, Xist, Epigenetics, Mouse

INTRODUCTION
X inactivation is the dosage compensation mechanism that has evolved in mammals to ensure equal levels of X-linked gene products in females (XX) and males (XY). The master regulator of this process is a large non-coding RNA termed Xist (inactive X specific transcripts) (Brown et al., 1991; Brockdorff et al., 1991; Borsani et al., 1991). X inactivation is initiated by the upregulation of Xist expression. This upregulation leads to accumulation of Xist RNA that can be detected by RNA fluorescence in situ hybridisation (FISH) as a large domain coating the entire X chromosome from which it is expressed (Sheardown et al., 1997; Panning et al., 1997). Chromosome-wide transcriptional silencing rapidly ensues. The X chromosome becomes heterochromatic, exhibiting marks of transcriptionally silent chromatin such as H3K27me3 (Silva et al., 2003; Plath et al., 2003), H3K9me2 (Heard et al., 2001; Peters et al., 2002; Mermod et al., 2002), H4K20me1 (Kohlmaier et al., 2000; Tada et al., 2000) and promoter CpG island methylation (Norris et al., 1991). The inactive X chromosome (Xi) also loses marks associated with active transcription, such as H3K4 methylation (Heard et al., 2001) and acetylation of the four major histones H2A, H2B, H3 and H4 (Jeppesen and Turner, 1993; Belyaev et al., 1996).

In mouse there are two waves of X inactivation: imprinted X inactivation of the paternal X chromosome initiated in early preimplantation embryos and persisting in extra-embryonic trophoblast and primitive endoderm lineages; and random X inactivation that occurs in cells of the embryo proper at embryonic day (E)5.5-6.5. Embryo progenitor cells in the inner cell mass of the blastocyst reverse imprinted X inactivation in order to prepare for subsequent random X inactivation (Mak et al., 2004; Okamoto et al., 2004). Initial inactivation of the paternal X chromosome occurs as a result of a repressive imprint inherited by the maternal Xist allele, with the paternal allele being set to a default on state (Kay et al., 1994; Goto and Takagi, 2000; Tada et al., 2000). At the onset of random X inactivation, Xist is expressed only in females and from only one of the two X chromosomes present. The mechanism regulating this is not fully understood but involves precise integration of Xist repressors (for a review, see Senner and Brockdorff, 2009) and activators (Barakat et al., 2010), as well as trans-interactions of the two Xist loci. Xist expression is absolutely required for X inactivation to occur in cis. XX embryonic stem (ES) cells or female mice carrying one mutant Xist allele always inactivate the X chromosome carrying the wild-type Xist allele (Penny et al., 1996; Marahrens et al., 1997). Further to this, Xist expression is sufficient to trigger the inactivation process. Ectopic expression of Xist cDNA transgenes leads to transcriptional silencing and establishment of the epigenetic modifications associated with the Xi chromosome (Wutz and Jaenisch, 2000; Kohlmaier et al., 2004).
One approach to understanding how Xist carries out its silencing function has been to investigate the structure of the Xist gene (Brown et al., 1992; Brockdorff et al., 1992; Nesterova et al., 2001). Although primary sequence conservation is poor within the Xist gene, the overall gene structure is well conserved and, in particular, the conservation of several blocks of tandem repeats, termed A-F, between mammalian species suggests that they might be involved in Xist function.

In order to establish which parts of the Xist RNA transcript are important for its function, Wutz et al. (Wutz et al., 2002) generated ES cell lines with inducible Xist cDNA transgenes carrying deletion of different regions of the transcript. This study concluded that only the 5’ region of the transcript containing the A-repeats was necessary for gene silencing. Further to this, it appeared that sequences 3’ of the A-repeats act cooperatively to localise the RNA to the chromosome but they are redundant as no single sequence is absolutely required. The function of the A-repeats was subsequently investigated in vivo using a gene targeting approach. Interestingly, deleting the A-repeats altered the regulation of the mutant Xist allele such that it was never expressed. As a consequence it was not possible to examine the functionality of an A-repeat deficient Xist in vivo (Hoki et al., 2009).

Although transgene studies (Wutz et al., 2002) have yielded valuable information on functionally important regions of Xist RNA there are disadvantages to this approach. First, overexpression of Xist transgenes could mask subtle loss-of-function phenotypes and second, partial silencing by mutant Xist RNAs might be insufficient to cause functional nullism of the X chromosome, the basis of the silencing assay used by Wutz et al. (Wutz et al., 2002), and might therefore lead to hypomorphic Xist RNA mutations being scored as unaffected. With this in mind, it is clearly important to complement transgene studies by generating mutations of the endogenous Xist locus and analysing their phenotype in vivo. The majority of knockout alleles produced to date have deleted Xist promoter sequences and have therefore ablated transcription of mutant alleles. These mutations have all resulted in primary non-random X inactivation of the wild-type allele in the epiblast of heterozygous female embryos. This is true also for the specific deletion of the A-repeats, which is not predicted to affect known Xist promoter sequences (Hoki et al., 2009). The only exception is the targeted deletion of exon 4 that includes a highly conserved sequence predicted to form a stable stem loop structure (Caparros et al., 2002). Surprisingly, deleting exon 4 in mice did not cause detectable effects on X inactivation, as evidenced by viability and complete absence of skewing of random X inactivation in heterozygous females. Mutant alleles produced lower levels of Xist RNA but this did not impact Xist function (Caparros et al., 2002).

To extend analysis of functional elements within Xist RNA by gene targeting, we have used the XistINV allele to generate a targeted inversion, XistINV+, disrupting conserved sequences beginning 6 kb into exon 1 and extending to intron 5. Female mice carrying a maternally inherited XistINV allele exhibited secondary non-random X inactivation, demonstrating that the inversion affects Xist-mediated silencing but not Xist gene regulation. Female embryos inheriting XistINV on the paternal X chromosome are lost at ~E11.5 as a result of failed imprinted X inactivation. This phenotype is less severe than that reported for Xist null alleles (Marahrens et al., 1997), suggesting that XistINV is a hypomorphic allele. Consistent with this interpretation, molecular analyses suggest that XistINV RNA localises to the X chromosome in cis, but at reduced efficiency, and that silencing of X-linked genes is limited.

**MATERIALS AND METHODS**

### Gene targeting

129<sup>th</sup>-genomic sequences for the arms of homology were isolated from a λDASHIII genomic library made in house. A NotI-XhoI 9.2 kb fragment (+2189 nt to +11,382 nt relative to the Xist start site) was purified from pMLClI phage clone and cloned into pBluescript SK+. The plasmid was digested with Hpal, which linearised the plasmid and separated the 3.8 kb 5’ homology region from the 5.4 kb 3’ homology region. A blunt-ended LoxP-PGKNeo-LoxP cassette was then ligated into the dephosphorylated Hpal site in the vector. Plasmid DNA was digested with XhoI to accept a SalI/XhoI fragment of the diptheria-toxin A gene. The latter was used for negative selection of random integrants (Yagi et al., 1990). The final construct pNBXT1 was linearised with Nael and electroporated into the XistEx4del XY ES cell line as described elsewhere (Caparros et al., 2002).

### Sexing embryos

Embryos were added to embryo lysis buffer (50 mM Tris HCl pH 8, 1 mM EDTA, 0.5% Tween 20, 200 μg/ml proteinase K) and incubated overnight at 55°C to lyse cells and digest proteins. The samples were then incubated at 95°C for 10 minutes to lyse any remaining cells and inactivate the proteinase K. The embryo lysate was then pipetted up and down rapidly to dissociate any remaining clumps and then spun briefly to pellet debris. The supernatant was used for the sexing PCR reaction. The sexing PCR amplifies the UBEX gene on the X chromosome as well as the homologous UBEY gene on the Y chromosome (F 5’-TTGGTCTGGACCCAAAAGCTGTCCACA-3’, R 5’-GGCCAGGAGCTACATACTCCAGATG-3’). The same primers yield a larger product from the X linked gene than from the Y. Accordingly, XX females have one band and XY males have two bands (Chuma and Nakatsuji, 2001). Reaction conditions were: 94°C for 4 minutes, 35 cycles of 94°C for 30 seconds, 64°C for 30 seconds, 72°C for 30 seconds, and a final 72°C for 5 minutes.

### Derivation and culture of mouse embryonic fibroblasts (MEFs)

E14.5 embryos were obtained from crosses between males carrying an X-linked GFP transgene and XistINV carrying females. The heads and organs were removed and carcasses were transferred to 0.25% trypsin/0.04% EDTA/2% chick serum and the tissues roughly dissociated by passing them through an 18 gauge needle (BD Biosciences). The dissociated carcasses were then incubated in the 0.25% trypsin/0.04% EDTA/2% CS for 10 minutes at 37°C. The trypsin was inactivated by adding an equal volume of MEF culture medium (DMEM supplemented with 10% foetal calf serum, 100 units/ml penicillin, 100 μg/ml streptomycin, 2 mM L-glutamine, 1× NEAA and 50 μM β-mercaptoethanol). Clumps of tissue were further dissociated by passing them through a 21 gauge needle (BD Biosciences). The resulting cells were washed in fresh MEF culture medium and transferred to a tissue culture vessel at a density of roughly one embryo per 25 cm<sup>2</sup>. All tissue culture reagents were from Invitrogen.

### RT-PCR and northern blot

RNA was isolated from MEFs using TRIzol reagent (Sigma) according to the manufacturer’s instructions. RNA was routinely treated with Turbo RNA-free reagent (Ambion) to exclude the possibility of DNA contamination. cDNA synthesis was primed from random hexamers with Superscript III reverse transcriptase (Invitrogen). PCR was then carried out for primers located in exon 1 (a, 5’-GTGCTGTGAGTGAACCTATGG-3’; b, 5’-TGCACCCAGAC- AAAGATTGGG-3’; e, 5’-CAATCCCTATGTGAGACTCA-3’; h, 5’-TGGTCTGGACCCAAAAGCTGTCCACA-3’; l, 5’-ATTATATGGATGAGTGATTGTGGTGAG-3’; m, 5’-CCATTTGCTTCAAGACGAT-3’; n, 5’-GTCATCCAGTCAGCT-3’; o, 5’-GTCATCCAGTCAGCT-3’; p, 5’-GTCATCCAGTCAGCT-3’; q, 5’-GTCATCCAGTCAGCT-3’). Reaction conditions were: 15°C for 4 minutes, 30 cycles of 94°C for 30 seconds, specific annealing temperature (T<sub>a</sub>) for 30 seconds, 72°C for 30 seconds. The PCR products were resolved on agarose gels and stained with ethidium bromide. The reactions were repeated at least twice to verify the same results.
Fig. 1. Strategy for generating a targeted inversion of Xist exon 1 conserved sequences. (A) Dotplot analysis of mouse and human Xist/XIST cDNA. Full length mouse and human Xist/XIST cDNA sequences were compared by dotplot analysis using the EMBOSS dotmatcher program (Rice et al., 2000). Two sets of parameters with different stringencies were used to demonstrate overall homology between two sequences (window size 50, threshold 45) or to highlight longer stretches with high homology (window size 150, threshold 95). Exon structure for each transcript is shown above (mouse) or alongside (human) the dotplots. Mouse intron 7, which contributes to a proportion of splice variant transcripts, is shown as a white box. Exon 4, the most conserved region, is shown in blue on the schematic and circled in blue on the dotplot. The region of inversion is indicated with red bracketed arrows and encircled in red on the dotplot. Note the extended homology over the whole region of the inversion. (B) A schematic representation of the previously targeted Xist allele with a deletion of exon 4 (Xist\textsuperscript{neo}) (Caparros et al., 2002) and the pNBXT1 targeting construct. Conserved repetitive elements A-F within exons 1 and 7 are shown as shaded boxes. LoxP sequences (grey triangles) flank a Pgk promoter-driven neomycin (neo) selection cassette (pale grey rectangle) between 5' and 3' arms of homology (dark grey rectangles). The positions of the arms of homology are indicated with black lines below the schematic. A diptheria-toxin A negative selection cassette (white rectangle) was included in the targeting construct. (C) The resulting homologous recombinant allele (Xist\textsuperscript{inv}) carrying both a deletion of Xist exon 4 (replaced by a LoxP sequence) and a floxed neo cassette in reverse orientation. (D) Transient expression of Cre-lox mediated recombinase generates a deletion of the neo cassette (Xist\textsuperscript{inv}) and/or inversion (Xist\textsuperscript{inv}) of the region between exon 1 and intron 4 (5947 bp to 13,670 bp downstream of the transcriptional start site).

**RESULTS**

**Generation of a targeted inversion in Xist exon 1**

Comparison of Xist cDNA sequences from mouse and human at high stringency reveals conservation in only a limited number of short regions (Fig. 1A). In addition to exon 4 and the A-repeats, both of which have been functionally analysed in previous studies (Caparros et al., 2002; Wutz et al., 2002; Chaumeil et al., 2006; Hoki et al., 2009), high sequence conservation occurs over a region of ~3 kb beginning 6 kb downstream of the transcriptional start site in the large exon 1 (highlighted in red in Fig. 1A, right panel). To determine the importance of this region in native Xist RNA transcripts we used homologous recombination in ES cells to engineer an inducible inversion of the conserved exon 1 sequence through to intron 5 (Fig. 1B-D). A loxP site was targeted 6 kb upstream of the transcriptional start site in XY ES cells and in the
opposite orientation to a loxP site present in the parent cell line \(Xist\)Ex4del in which \(Xist\) exon 4 had previously been deleted (Caparros et al., 2002). Correctly targeted cells carrying this allele, designated \(Xist^{\Delta neo}\), were then transiently transfected with Cre recombinase to generate the \(Xist^{\Delta neo}\) allele in which the neo cassette was excised and the floxed region of \(Xist\) exon 1 was in the correct orientation.

\(Xist^{\Delta neo}\) cells were injected into blastocysts and germ line transmission was achieved from chimeric offspring. \(Xist^{\Delta neo}\) hemizygous males, heterozygous and homozygous females were normal and fertile, indicating that insertion of a loxP site in exon 1 did not affect \(Xist\) gene function. This was confirmed by analysing allelic ratios of \(Xist\) and \(Pgk-1\) in female progeny from \(Xist^{\Delta neo} \times Pgk-1\) strain crosses (data not shown). To induce inversion of \(Xist\) exon 1 we used transient expression of Cre-recombinase in fertilised oocytes of \(Xist^{\Delta neo}\) homozygous animals. Resultant progeny were analysed to identify males and female heterozygotes in which sequences between the two loxP sites were stably inverted, designated \(Xist^{INV}\).

**Females with a maternally inherited \(Xist^{INV}\) allele exhibit secondary non-random X inactivation**

To determine whether the \(Xist^{INV}\) mutation affects X inactivation we first assayed expression of wild-type and \(Xist^{INV}\) alleles in tissues of adult \(Xist^{INV}\) heterozygous females. As illustrated in Fig. 2A, expression of the \(Xist^{INV}\) allele was detected but at a significantly reduced level relative to the wild-type allele. This observation suggested extreme non-random X inactivation of the X chromosome carrying the wild-type \(Xist\) allele.

Skewed \(Xist^{INV}\) expression could result from primary non-random X inactivation at the time X inactivation is initiated (\(-E5.5-6.5\)), or alternatively from secondary non-random X inactivation in which cells expressing the \(Xist^{INV}\) allele are selected against subsequent to the onset of random X inactivation. To address this, we crossed heterozygous \(Xist^{INV}\) females with males carrying an X-linked GFP transgene (XGFP) and a wild-type \(Xist\) allele. The GFP transgene is located on the proximal part of the X chromosome and has previously been shown to undergo silencing both in imprinted X inactivation in extra-embryonic lineages and in random X inactivation in the embryo (Hadjantonakis et al., 1998; Hadjantonakis et al., 2001). Primary non-random X inactivation is predicted to lead to absence of GFP in both extra-embryonic and embryonic cells immediately following initiation of random X inactivation (E5.5-6.5). Conversely, in secondary non-random X inactivation loss of GFP expression is predicted to occur more gradually in cells of the embryo proper, the rate of loss being determined by the degree to which dosage compensation is compromised.

XX embryos with a maternally inherited \(Xist^{INV}\) allele exhibited progressive loss of GFP expression cells between E6.5 and E14.5, consistent with secondary non-random X inactivation. Figure 2B shows representative embryos at E8.5 where significant GFP expression is seen throughout the embryo and at E11.5 where GFP expression is restricted to the heart.

FACS analysis of MEFs from E14.5 embryos identified a small proportion (1-10%) of GFP positive cells (Fig. 2C,D). We carried out RT-PCR on sorted populations using primer pairs to detect expression of wild-type and \(Xist^{INV}\) RNAs. As a control we analysed MEFs expressing the \(Xist^{-}\) allele (Caparros et al., 2002). Using primers flanking one of the break points in the \(Xist^{INV}\) allele we confirmed that only the GFP positive MEFs express this allele (Fig. 2E). Primers (located in exon 3 and exon 5) that are specific for the wild-type allele amplified products of the predicted size from both the GFP negative MEFs and the \(Xist^{-}\) fibroblasts and, as expected, did not amplify a product from the \(Xist^{INV}\) allele. Primers located in exon 1, upstream of the inversion, detected transcripts in all samples. A full RT-PCR analysis of \(Xist^{INV}\) transcripts using primers across the locus is shown in Fig. S1 in the supplementary material.
Northern blot analysis of GFP-positive MEFs revealed that the XistINV allele transcribes two major RNA species, approximately 12 and 20 kb in length (see Fig. S2A in the supplementary material). The 20 kb species corresponds to the maximum length transcript predicted from the sequence of the genomic locus and RT-PCR analysis (see Fig. S1 and Fig. S2B in the supplementary material). The 12 kb species probably results from premature termination. Consistent with this, there were potential polyadenylation signals clustered ~12 kb downstream of the promoter within inverted exon 1 sequence (see Fig. S2B in the supplementary material). The northern blot analysis also indicated that the level of XistINV RNA is similar to that seen with the Xistexed allele, both of which are lower than Xist RNA levels in wild-type cells (see Fig. S2A in the supplementary material). It should be noted that because the Xistexed mutation does not affect X inactivation (Caparros et al., 2002), the phenotype of the XistINV mutation is unlikely to be attributable to reduced RNA levels. In summary, the XistINV allele produces a contiguous transcript (including inverted sequences) up to Xist exon 5, subsequent to which processing is identical to wild-type Xist RNA. A proportion of XistINV transcripts truncate prematurely at polyadenylation signals 12Kb downstream of the promoter.

Taken together, these results demonstrate that random X inactivation occurs in female embryos with a maternally inherited XistINV allele and that subsequently there is a progressive selection against cells that express the XistINV allele. A small proportion of cells expressing XistINV RNA persist through ontogeny and in adult animals.

**Embryonic lethality of a paternally inherited XistINV allele**

We observed that male animals carrying the XistINV allele are fertile but fail to sire daughters, suggesting failure of imprinted X inactivation in the extra-embryonic tissues where only the paternally transmitted Xist allele is expressed. To investigate this further, litters fathered by XistINV carrying males were dissected at various gestational timepoints. The morphology and size of all embryos in a litter was noted and then PCR was carried out to determine the sex of each embryo. Examples are shown in Fig. 3. XistINV female embryos were indistinguishable from their wild-type male littermates at E6.5 but appeared growth retarded as early as E7.5 (Fig. 3). At E8.5 XistINV female embryos appeared markedly smaller. At E9.5 XistINV female embryos were much smaller and were often developmentally delayed. For example, at E9.5 neural tube closure is usually complete but some XistINV female embryos had not completed caudal fusion at this stage. Furthermore, E9.5 XistINV female embryos showed no signs of developing a placenta whereas placental development of male wild-type littermates appeared normal. By E10.5 XistINV female embryos were disintegrating.

Loss of female embryos with a paternally inherited XistINV allele during the first half of gestation and in particular the failure to develop a placenta is likely to be due to a failure of imprinted X inactivation in the extra-embryonic tissues where only the paternally inherited XistINV allele is expressed. However, the embryonic phenotype observed was not as severe as that noted previously in studies analysing a null allele, dXist (Marahrens et al., 1997).

**Silencing of X-linked genes is compromised in cells expressing XistINV RNA**

To assess the expression of X-linked genes in cells expressing XistINV RNA, we first crossed wild-type females with male mice carrying both XGFP and the XistINV allele on their single X chromosome and analysed GFP expression in tissues that undergo imprinted X inactivation. As shown in Fig. 4A, E3.5 blastocyst stage embryos, in which all cells have undergone imprinted X inactivation, have high levels of GFP expression relative to wild-type controls. There was, however, some variability between embryos and between patches of cells within each individual embryo and overall GFP expression levels appeared lower than in E3.5 embryos, in which XGFP is located on the active maternally derived X chromosome (Fig. 4A; arrowhead). We also analysed E6.5 embryos (Fig. 4B). In wild-type controls GFP expression was undetectable in extra-embryonic ectoderm, consistent with maternally imprinted X inactivation, whereas in XistINV females we observed low-level GFP expression throughout with some patches of stronger GFP expression. In embryonic ectoderm there was strong GFP expression both in wild-type and XistINV embryos, consistent with random X inactivation. These results suggest that cells expressing the XistINV allele fail to initiate or maintain silencing of the XGFP transgene appropriately.

Next, we analysed the expression of endogenous X-linked genes. E3.5 embryos were obtained from crosses between either wild-type or XistINV males carrying the XGFP transgene and PGK strain females. Polymorphisms arising between the PGK strain and the targeted 129 strain allowed expression from the maternally derived (PGK) and paternally derived (129) alleles to be distinguished by single nucleotide primer extension (SNuPE). Female embryos were identified by GFP expression as above and then pooled together for analysis. We assessed expression of three X-linked genes, Pgk-1,
Gla and Smc111, for which there are known polymorphisms (Fig. 4C-D). As expected, in wild-type embryos expression of Pgk-I and Gla was predominantly from the maternally derived allele. Only very low levels of paternally derived transcripts were detected, reflecting the inactive state of the paternally inherited X chromosome. In Xist\textsuperscript{INV} carrying embryos, however, significantly more expression was detected from the paternally derived allele (Student’s t-test, P<0.005 for Pgk-I and Gla) indicating that the process of imprinted X inactivation is indeed impaired. Although we cannot rule out that a proportion of cells in the Xist\textsuperscript{INV} blastocyst have reversed imprinted X inactivation patterns and then proliferated under selection, we consider this to be unlikely both because failure to develop a placenta argues against significant rescue and because reversion of imprinted X inactivation under selective pressure was not observed in previous studies using null Xist alleles (Marahrens et al., 1997). In wild-type embryos it appeared, as previously reported, that Smc111 is only partially inactivated at this stage (Mak et al., 2004). However, Xist\textsuperscript{INV} carrying embryos still had significantly higher expression levels from the paternally derived allele compared with wild-type embryos (Student’s t-test, P<0.05).

Reduced H3K27me3 and H2AK119u1 domains in cells expressing Xist\textsuperscript{INV} RNA

To determine whether cells expressing Xist\textsuperscript{INV} RNA exhibit chromosomal features of X inactivation we analysed H3K27me3 and H2AK119u1, histone modifications that are enriched on Xi as a consequence of Xist-dependent recruitment of polycomb repressor complexes PRC2 and PRC1, respectively (Silva et al., 2003; Plath et al., 2003; de Nepoles et al., 2004; Fang et al., 2004). E6.5 XX embryos carrying either a wild-type or Xist\textsuperscript{INV} allele on the paternal X chromosome were analysed by immunofluorescence using antibodies specific for H3K27me3 and H2AK119u1 (Fig. 5).

Similar numbers and sizes of domains of colocalising H3K27me3 and H2AK119u1 were observed in wild-type embryonic and extra-embryonic cells. A significant increase in the number of cells with no domains was observed in Xist\textsuperscript{INV} embryonic cells (Student’s t-test, P<0.05). However, large bright domains similar to those seen in the wild-type embryos were observed. Significantly fewer domains were seen in Xist\textsuperscript{INV} extra-embryonic tissues (Student’s t-test, P<0.05) and those domains that were seen were smaller and of reduced intensity (Fig. 5). More frequent bright domains seen in embryonic cells probably represent cells that inactivated the wild-type X chromosome. As a control, NB18 TS cells carrying the Xist\textsuperscript{inv} allele were analysed and the domains were found to be of a comparable size and intensity to those observed in the wild-type embryonic and extra-embryonic cells (see Fig. S3 in the supplementary material), indicating that the reduction in these modifications is attributable to the inversion of Xist RNA sequences and not to the deletion of exon 4. The reduced H2AK119u1 domains observed in extra-embryonic cells of mutant embryos were associated with underlying histone H4 hypoacetylation in many, but not all, instances and domains were again relatively small (see Fig. S4 in the supplementary material). These observations are consistent with compromised gene silencing by Xist\textsuperscript{INV} RNA.

Localisation of Xist\textsuperscript{INV} RNA to the X chromosome is impaired

Next, we analysed localisation of Xist\textsuperscript{INV} RNA to the Xi chromosome by RNA FISH on whole-mount E6.5 female embryos (Fig. 6). Again, these embryos had a paternally inherited Xist\textsuperscript{INV} allele and so extra-embryonic tissues expressed only the mutant allele. When compared with wild-type embryonic and extra-embryonic and Xist\textsuperscript{INV} embryonic cells, the size and intensity of Xist domains in the Xist\textsuperscript{INV} extra-embryonic cells were reduced.

The apparent reduction in Xist\textsuperscript{INV} RNA domains could underlie the compromised silencing of X-linked genes, either because overall transcript levels are insufficient for chromosome wide silencing or because of compromised localisation of Xist\textsuperscript{INV} RNA to Xi in cis. Consistent with the former possibility, Xist\textsuperscript{INV}
transcript levels are lower than wild-type Xist RNA (see Fig. S2A in the supplementary material). However, levels are equivalent to the Xist∆ex4 allele that was shown previously to function normally (Caparros et al., 2002). RNA FISH analysis of interphase nuclei in MEF cell lines, however, indicated that XistINV RNA domains are reduced in size compared with both wild type and Xist∆ex4 controls, suggesting that in cis spreading of XistINV RNA might occur over only a limited region of the chromosome. To test this, we analysed in cis localisation of XistINV RNA on metaphase chromosomes. As illustrated in Fig. 7, XistINV RNA localised along the entire chromosome and there was no detectable concentration of transcripts within a smaller region of the chromosome. However, overall levels were substantially lower than those seen for wild-type and Xist∆ex4 controls. These results suggest that XistINV RNA does associate with the X chromosome, but with reduced affinity, and that this underlies defective silencing of X-linked genes and female embryo lethality.

DISCUSSION
Here we describe analysis of a novel targeted mutation, XistINV, designed to test the function of a conserved region located in exon 1 of Xist RNA during X inactivation. We show that XistINV is regulated appropriately but is compromised in its ability to silence X-linked genes in cis. Inheritance of XistINV on the paternal X chromosome results in embryonic lethality due to failure of imprinted X inactivation in extra-embryonic lineages. Female embryos inheriting XistINV on the maternal X chromosome undergo extreme secondary non-random X inactivation, eliminating the majority of cells that express the XistINV allele. Analysis of cells that express XistINV demonstrates that the mutant RNA shows reduced association with the X chromosome, suggesting that conserved sequences in the inverted region are important for Xist RNA localisation.

XistINV compromises X chromosome silencing
Paternally inherited XistINV was appropriately expressed in extra-embryonic ectoderm of E6.5 embryos, indicating that the mutation does not affect the regulation of Xist imprinting. There was, nevertheless, a failure of imprinted X inactivation demonstrated by inefficient silencing of an X-linked GFP transgene both at the blastocyst stage (E3.5) and at E6.5. Endogenous X-linked genes Pgk-1, Gla and Smc1l1 also showed impaired silencing at the blastocyst stage. This impairment contrasts with a recently published study in which female embryos with a paternally inherited Xist knockout allele were suggested to initiate X inactivation normally (Kalantry et al., 2009). In these experiments, blastocysts with an identical paternally inherited GFP transgene silenced the transgene to the same degree in mutant and wild-type female embryos. It was only later that the GFP was reactivated. Based on these results, it was suggested that initiation of imprinted X inactivation is independent of Xist. Here, GFP silencing was clearly not observed at the blastocyst stage to the same extent as in wild-type embryos with a paternally inherited Xist knockout allele, suggesting that initiation of silencing is dependent on Xist. Further studies will be required to resolve this apparent paradox.
The phenotype of XistINV is not limited to imprinted X inactivation. During random X inactivation female embryos exhibited progressive loss of cells in which XistINV RNA was expressed. Thus, like in imprinted X inactivation, appropriate regulation of Xist expression is unaffected by the presence of the XistINV allele but downstream silencing by XistINV RNA is compromised. This contrasts with previously studied Xist mutants that in most cases trigger primary non-random X inactivation of the non-mutant allele or, in the case of deletion of Xist exon 4, show no discernible phenotype. XistINV is the first Xist mutant allele, to our knowledge, shown to cause secondary non-random X inactivation. It is interesting to note that a small number of XistINV RNA-expressing cells are viable both in MEF cultures and in adult tissues. These could represent cells in which failure of dosage compensation has been tolerated and/or cells in which silencing by XistINV RNA is more robust than is generally the case.

**Impaired localisation of XistINV transcripts**

RNA FISH on whole-mount embryos and metaphase chromosomes revealed that XistINV RNA is not localising efficiently to the X chromosome from which it is transcribed. This does not appear to relate to the low expression of XistRNA, as RT-PCR analysis of the transcript showed equivalent levels compared with XistΔex4 RNA, which localises and executes X inactivation normally. Moreover, Xist RNA levels in some wild-derived mouse strains are significantly lower than in laboratory mice (Brockdorff et al., 1991; Buzin et al., 1994) but X inactivation appears to be unaffected. There does not appear to be a problem with spreading of XistINV transcripts as chromosome coating is detected all along the chromosome. However, there are clearly lower levels of XistINV RNA associated with the X chromosome at metaphase, suggesting that the mutant transcript adheres with reduced affinity.

Previously, the entire XistINV region was deleted in an Xist cDNA transgene and, upon ectopic expression in ES cells, silencing and localisation appeared approximately the same as observed with the full length transgene (Wutz et al., 2002). This suggests that the sequence of the inverted region is not directly involved in localisation. It is possible that expression of the transgene was greater than endogenous Xist and that this partially overrides reduced localisation efficiency. Alternatively, inclusion of additional reversed exons and introns between the LoxP sites in the XistINV transcript might have changed the structure of the molecule in such a way as to inhibit interactions required for localisation.

Wutz et al. (Wutz et al., 2002) previously reported that redundant sequences throughout the 3’ end of the molecule mediate localisation, including a region in exon 1 immediately adjacent to the inverted region described here. It is possible that our targeted inversion impairs the interactions that this adjacent region is involved in.

Levels of enrichment of the repressive histone modifications H2AK119u1 and H3K27me3 catalysed by recruitment of PRC1 and PRC2 polycomb complexes to Xi are strongly reduced in cells of E6.5 embryos expressing XistINV RNA. This is consistent with previous data analysing blastocysts expressing XistINV RNA from the paternal X chromosome (Silva et al., 2003). In both cases, reduced polycomb complex recruitment is likely to be a direct reflection of reduced Xist RNA domains. However, we cannot rule out that XistINV RNA recruits PRC1 and PRC2 less efficiently. Previous work has demonstrated that sequences throughout Xist RNA have a role in recruiting PRC1 and PRC2 and, moreover, that the complexes are recruited by a silencing deficient Xist RNA transgenie in which the A-repeats have been deleted (Kohlmaier et al., 2004; Chaumeil et al., 2006).

**XistINV is a hypomorphic allele**

Female mice carrying a paternally inherited XistINV allele are lost during embryogenesis. However, the phenotype of these mice is less severe than female embryos carrying a paternally inherited Xist knockout allele. Female embryos with a paternally inherited XistINV allele are indistinguishable from male littermates at E6.5 and appear smaller at E8.5 but have no developmental phenotype at that stage. By contrast, female embryos with a paternally inherited Xist null allele are reduced in size as early as E6.5 and are severely dysmorphic by E8.5 (Marahrens et al., 1997). This disparity raises the possibility that XistINV is a hypomorphic allele. XistINV is clearly not capable of efficiently silencing the X chromosome. However, in blastocysts with a paternally inherited XistINV on the same
chromosome as a GFP transgene there was evidence for some degree of silencing of the GFP transgene. In some places the GFP expression was in the same range as that of the maternally inherited GFP but in others it was almost as low as in the wild-type embryos. Although endogenous genes also showed increased expression from the paternally inherited X chromosome, expression levels did not reach the levels detected from the maternally inherited X chromosome. This observation indicates partial silencing of the paternally inherited X chromosome and again supports the idea that Xist\textsuperscript{INV} is a hypomorphic allele. In further support of this is the observation that enrichment of the Xist\textsuperscript{INV} extra-embryonic tissues at E6.5 with H3K27me3 and H2AK119u1 was substantially reduced but not completely abolished. Underlying H4 hypoacetylation was found coincident with H2AK119u1 foci at interphase, but this was only in a proportion of cells and did not extend beyond the relatively small H2AK119u1 foci observed in Xist\textsuperscript{INV} extra-embryonic tissues. This is suggestive of reduced efficiency of silencing in response to Xist\textsuperscript{INV} RNA.

A small proportion of genes might be silenced in each cell and this silencing maintained, for example, through the recruitment of polycomb complexes. Partial silencing would not be sufficient for normal embryonic development but it could perhaps carry the embryos slightly further along embryonic development than a full Xist knockout. The low level of Xist\textsuperscript{INV} RNA localisation observed could constitute a smaller repressive compartment into which some genes could be recruited. Given the variability of the GFP expression levels throughout each embryo, different genes might be silenced in different cells in a stochastic manner. Analysis of endogenous genes in pooled blastocysts, however, suggests that individual loci are subject to different degrees of silencing and that this might relate to distance from the Xist locus. Future studies, for example using global analysis of gene expression in Xist\textsuperscript{INV} mutants, might discriminate weak silencing of all or most genes from relatively efficient silencing of only a subset of genes.

In summary, this study suggests that conserved sequences located in the distal half of Xist exon 1 are important for in cis localisation of Xist RNA. A recent study has demonstrated that scaffold attachment factor A (SAF-A; HNRNPU – Human Gene Nomenclature Database) mediates Xist RNA localisation, interacting with Xist RNA through an RRM RNA binding domain, and immunoprecipitation of UV cross-linked SAF-A preferentially enriched sequences at the distal end of Xist exon 1 (Hasegawa et al., 2010). Conserved exon 1 sequences included in the Xist\textsuperscript{INV} region therefore provide good candidates for further analysis to determine the precise RNA sequence/structure required for SAFA binding which in turn might provide significant insight into the mechanism of in cis localisation of Xist RNA.

Acknowledgements

We thank members of the lab for helpful discussion and Vasso Episkopou for assistance with embryo studies. This work was funded by the Medical Research Council, UK, and the Wellcome Trust. Deposited in PMC for release after 6 months.

Competing interests statement

The authors declare no competing financial interests.

Supplementary material

Supplementary material for this article is available at http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.056812/-/DC1

References


A

```
inv
∆
ex4wt
H2O++ +- - -

b - i
f - n
g - o
i - l
c - j
d - k
e - m
a - h
```

```
wt

exon 1 2 3 4 5 6 7

h

j

k

l

m

o
```

```
Dex4

exon 1 2 3 4 5 6 7

h

j

k

l

m

o
```

```
INV

exon 1 2 3 4 5 6 7

h

j

k

l

m

o
```

B

```
wt  Dex4  INV
+-  +  -  +  -  H2O

a - h
```

```
wt  Dex4  INV
+-  +  -  +  -  H2O

f - n
g - o
```

```
wt  Dex4  INV
+-  +  -  +  -  H2O

d - k
b - d
e - m
```

```
b - i
i - l
c - j
```

Figure S3

H3K27me3  H2AK119u1  DAPI  merge
Table S1. Sequences of the primers and annealing temperatures (Ta) used in RT-PCR analysis

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Ta (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>GTGCTGTGTGAGTGAAACCTATGG</td>
<td>65</td>
</tr>
<tr>
<td>b</td>
<td>AAGTGTTG6CATATCTGGTTCCT</td>
<td>66</td>
</tr>
<tr>
<td>c</td>
<td>CATCAGGCTTGGAGCAGCAA</td>
<td>63</td>
</tr>
<tr>
<td>d</td>
<td>TGAATTGGAGTCTGTTTTGTCC</td>
<td>61</td>
</tr>
<tr>
<td>e</td>
<td>CAATCCCTATGTGAGACTCAAGGACT</td>
<td>63</td>
</tr>
<tr>
<td>f</td>
<td>CTTGCCACGGCTTCTATCTTTCC</td>
<td>62</td>
</tr>
<tr>
<td>g</td>
<td>TTCAGCGTGATCAAGCAATTG</td>
<td>64</td>
</tr>
<tr>
<td>h</td>
<td>TACACACCAAGAAGGATTTGGG</td>
<td>65</td>
</tr>
<tr>
<td>i</td>
<td>CTTGAAGGACCATTGACGGTATT</td>
<td>66</td>
</tr>
<tr>
<td>j</td>
<td>GCACTCTTACACTCTCAAATCGAAG</td>
<td>63</td>
</tr>
<tr>
<td>k</td>
<td>CAGAGTACCGAGGACTTGAAGAG</td>
<td>63</td>
</tr>
<tr>
<td>l</td>
<td>ATTATAGGCAATGAGGTTAGCAGTG</td>
<td>66</td>
</tr>
<tr>
<td>m</td>
<td>GCCTTCCCTAGCTCTGGAGAGAGA</td>
<td>63</td>
</tr>
<tr>
<td>n</td>
<td>ATGTTGATCTCGGGTCAATTAT</td>
<td>64</td>
</tr>
<tr>
<td>o</td>
<td>TAACAAAGAAATTAGACACACCAGAC</td>
<td>60</td>
</tr>
</tbody>
</table>