Defects in dosage compensation impact global gene regulation in the mouse trophoblast

酒田, 祐佳

https://doi.org/10.15017/1931751

出版情報：九州大学, 2017, 博士（医学）, 課程博士
バージョン：

© 2017. Published by The Company of Biologists Ltd | Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

RESEARCH ARTICLE

Defects in dosage compensation impact global gene regulation in the mouse trophoblast

Yuka Sakata[1,2,*,‡], Koji Nagao[3,4,‡], Yuko Hoki[2], Hiroyuki Sasaki[2], Chikashi Obuse[3,4] and Takashi Sado[1,2,§]

ABSTRACT

Xist RNA, which is responsible for X inactivation, is a key epigenetic player in the embryogenesis of female mammals. Of the several repeats conserved in *Xist* RNA, the A-repeat has been shown to be essential for its silencing function in differentiating embryonic stem cells. Here, we introduced a new *Xist* allele into mouse that produces mutated *Xist* RNA lacking the A-repeat (*Xist*$^{CAGΔ5'}$). *Xist*$^{CAGΔ5'}$ RNA expressed in the embryo coated the X chromosome but failed to silence it. Although imprinted X inactivation was substantially compromised upon paternal transmission, allele-specific RNA-seq in the trophoblast revealed that *Xist*$^{CAGΔ5'}$ RNA still retained some silencing ability. Furthermore, the failure of imprinted X inactivation had more significant impacts than expected on genome-wide gene expression. It is likely that dosage compensation is required not only for equalizing X-linked gene expression between the sexes but also for proper global gene regulation in differentiated female somatic cells.

KEY WORDS: X chromosome inactivation, *Xist* RNA, Trophoblast, Mouse, Transcriptome

INTRODUCTION

Dosage compensation of X-linked genes is a crucial epigenetic event that occurs at around the peri-implantation stages to allow successful development of female mice. The X inactive-specific transcript (*Xist*) gene encodes a long noncoding RNA that is monoallelically expressed from one of the two X chromosomes and coats the chromosome that it originates from, which triggers chromosome-wide heterochromatinization to compensate for dosage differences of X-linked genes between the sexes (Heard and Disteche, 2006; Lyon, 1961; Marahrens et al., 1997; Penny et al., 1996). In the mouse embryo, there are two waves of X chromosome inactivation (XCI) depending on the developmental lineages. XCI first takes place with a bias to silencing the paternal X chromosome (Xp) during preimplantation stages (imprinted XCI) and is maintained in the trophectoderm and primitive endoderm lineages (Takagi and Sasaki, 1975). The absence of dosage compensation causes an arrest in proliferation of the progenitor cells of the trophoblast and results in the failure of placental differentiation (Mugford et al., 2012; Roberts et al., 2004). It is likely that the defects in the extraembryonic tissues lead to deterioration in the subsequent development of embryonic tissues (Takagi and Abe, 1990). In contrast to the cells in the extraembryonic lineages, those specified to the epiblast lineage in the blastocyst transiently reactivate Xp to erase the memory of imprinted XCI and subsequently undergo XCI in a random fashion, such that either Xp or the maternal X chromosome (Xm) is inactivated regardless of parental origin.

X-linked noncoding *Xist* RNA plays a pivotal role in both imprinted and random XCI. It is monoallelically upregulated on either X chromosome at the onset of XCI and coats it in cis to cause a series of changes in epigenetic state, such as histone modifications and DNA methylation. It has been shown that an X chromosome deficient for *Xist* never undergoes inactivation, demonstrating that *Xist* is essential for XCI to occur in cis (Marahrens et al., 1997). The X chromosome coated with *Xist* RNA forms a unique silencing domain characterized by the accumulation of inactive histone marks such as trimethylation of H3K27 (H3K27me3) (Plath et al., 2003; Rougeulle and Avner, 2003; Silva et al., 2003) and exclusion of active marks such as hypoacetylation of H3K9 and H4 (Jeppesen and Turner, 1993) as well as hypomethylation of H3K4 (Chaumeil et al., 2006). These changes in histone modifications take place relatively early in the process of XCI, suggesting that they are involved in the initiation and/or early maintenance of the inactive state.

Much of our knowledge about the molecular mechanisms of XCI have been derived from studies using not only female embryonic stem cells (ESCs), in which XCI can be initiated by the induction of *Xist* expression, but also male ESCs. The A-repeat, which is one of the repeats in the 5′ region of *Xist* RNA that are conserved among many eutherian mammals, is essential for the silencing function of the RNA but not for its localization to the X chromosome (Wutz et al., 2002). Intriguingly, the *Xist* cloud formed by the RNA deleted for the A-repeat, although defective in X-linked gene silencing, retains a histone modification status indistinguishable from that of the cloud formed by wild-type *Xist* RNA, such as hypoacetylation of histone H4 and enrichment of H3K27me3 and H4K20me1 (Chaumeil et al., 2006; Kohlmaier et al., 2004; Pullirsch et al., 2010). It has been suggested that this apparent discrepancy is reconciled by the finding that X-linked genes that fail to be silenced are located at the outside or periphery of the cloud formed by the *Xist* RNA lacking the A-repeat, and are therefore not incorporated into the apparent heterochromatin domain (Chaumeil et al., 2006), although the molecular basis of how the A-repeat exerts its effect on this spatiotemporal regulation is not known. When we previously attempted to explore this issue by targeted deletion of the A-repeat in mouse, it completely abolished the expression of *Xist*, precluding evaluation of the RNA lacking the A-repeat (Hoki et al., 2009).

In this study, we introduced a mutated *Xist* allele into mouse that is expected to constitutively express *Xist* RNA lacking the

[1]Department of Bioscience, Graduate School of Agriculture, Kindai University, 3327-204, Nakamachi, Nara 631-8505, Japan. [2]Division of Epigenomics, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan. [3]Graduate School of Life Science, Hokkaido University, Sapporo, Hokkaido 001-0021, Japan. [4]Department of Biological Sciences, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan.
*Division of Cancer Biology, Cancer Institute, Japanese Foundation for Cancer Research, 3-8-31, Ariake, Koto-ku, Tokyo 135-8550, Japan.
‡These authors contributed equally to this work

§Author for correspondence (tsado@nara.kindai.ac.jp)

Y.S., 0000-0003-4254-2417; T.S., 0000-0002-1232-0250

Received 16 January 2017; Accepted 27 June 2017

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

5′ sequence including the A-repeat, and studied the effect of this mutated *Xist* on XCI during embryogenesis. The mutated *Xist* RNA was successfully expressed and coated the X chromosome in cis in the embryo, but failed to silence genes on the chromosome in embryonic and extraembryonic lineages. As has been reported in differentiating ESCs, the X chromosome coated with this silencing-defective *Xist* RNA in both lineages apparently acquired histone modifications typical for the properly inactivated X chromosome. Paternal transmission of this allele resulted in embryonic lethality at the peri-implantation stages due to the failure of imprinted XCI in the extraembryonic tissues, suggesting the importance of the A-repeat for the silencing function of *Xist* RNA in the embryo. Transcriptomic analysis of this mutant trophoblast in comparison with another mutant trophoblast carrying a paternally derived *Xist* null allele revealed that, although the dysfunctional *Xist* RNA expressed in the trophoblast indeed failed to inactivate many genes on the Xp, it still retained some ability to silence a subset of genes, suggesting that the silencing function of the *Xist* RNA lacking the A-repeat is not completely abolished. Furthermore, we found that the failure of imprinted paternal XCI had more significant impacts than we expected on the genome-wide regulation of gene expression. Based on these findings, the biological significance of dosage compensation in female mammals is discussed.

RESULTS

A novel *Xist* allele that produces RNA lacking the A-repeat in mouse

Since our previous study revealed that a simple deletion of the A-repeat abolished *Xist* upregulation during embryonic development (Hoki et al., 2009), we decided instead to express *Xist* RNA lacking the A-repeat using a CAG promoter, using the same approach as we used previously to create the $Xist^{CAG}$ allele (Amakawa et al., 2015). The 5′ region of *Xist* spanning from its endogenous promoter to the *Xho*I site 0.9 kb downstream of the major transcription start site in exon 1 was replaced with a fragment containing the CAG promoter and a floxed selection marker by gene targeting to generate the $Xist^{CAGΔ5'\text{-}2L}$ allele (Fig. S1A,B). The presence of the loxP-flanked selection marker inhibits expression of the mutated *Xist* allele, similar to a stop cassette. Chimeric males that were subsequently produced were crossed with females heterozygous for $Xist^{CAG}$ and eventually the targeted allele was successfully transmitted to offspring ($Xist^{CAG}/Xist^{CAGΔ5'\text{-}2L}$, where the maternal allele precedes the paternal allele – the same convention is used hereafter) (Fig. S1C). These $Xist^{CAG}/Xist^{CAGΔ5'\text{-}2L}$ females were further crossed with males carrying a *Pgk2-cre* transgene, which is specifically expressed in spermatocytes (Kido et al., 2005), to produce males carrying $Xist^{CAGΔ5'\text{-}2L}$ in combination with *Pgk2-cre* [$Xist^{CAGΔ5'\text{-}2L}$/Y; Tg(*Pgk2-cre*)]. These males, in which Cre-mediated conversion from $Xist^{CAGΔ5'\text{-}2L}$ to $Xist^{CAGΔ5'}$ occurs during spermatogenesis, allowed us to study the effect of the mutated, most likely dysfunctional, *Xist* RNA ($Xist^{CAGΔ5'}$ RNA) on XCI in female offspring (+/$Xist^{CAGΔ5'}$).

Imprinted XCI is compromised upon paternal transmission of $Xist^{CAGΔ5'}$

In a previous study, we found that all +/$Xist^{CAG}$ embryos obtained from a cross between wild-type females and $Xist^{CAG2L}$/Y; Tg(*Pgk2-cre*) males were apparently normal and expressed the full-length *Xist* RNA under control of the CAG promoter (Amakawa et al., 2015). It was therefore reasonable to expect that $Xist^{CAGΔ5'}$ RNA, although lacking the A-repeat, would be expressed and thus we would be able to evaluate its silencing function. Unless $Xist^{CAGΔ5'}$ RNA has proper silencing function, no female pups would be obtained upon paternal transmission of $Xist^{CAGΔ5'}$ owing to the strict requirement for paternally expressed functional *Xist* RNA for imprinted XCI in the extraembryonic tissues. Accordingly, we crossed wild-type females with $Xist^{CAGΔ5'\text{-}2L}$/Y; Tg(*Pgk2-cre*) males. Of 141 pups born, no females were obtained, suggesting that female embryos had been lost *in utero* (Fig. S1D).

When embryos were dissected out at E7.5 from females crossed with $Xist^{CAGΔ5'\text{-}2L}$/Y; Tg(*Pgk2-cre*) males, a subset of the embryos were stunted and had abnormal morphology, all of which were females carrying the $Xist^{CAGΔ5'}$ allele on the Xp ($XX^{CAGΔ5'}$, where the maternal X precedes the paternal X carrying $Xist^{CAGΔ5'}$) (Fig. 1A,B). Histological analysis revealed that the ectoplacental cone and extraembryonic ectoderm were severely affected in the morphologically abnormal embryos, which were most probably females carrying the paternal $Xist^{CAGΔ5'}$ allele (Fig. 1C). They were reminiscent of female embryos carrying the paternally transmitted *Xist*-deficient X ($XX^{ΔA}$, where $X^{ΔA}$ carries $Xist^{ΔA}$) (Hoki et al., 2009), implying failure of imprinted XCI.

$Xist^{CAGΔ5'}$ RNA coats the X chromosome but fails to induce its silencing in the embryo

We next studied the transcriptional status of the mutated paternal $X^{CAGΔ5'}$ in the extraembryonic tissues. RNA-FISH was performed for *Xist* expression using two different probes (Fig. 1D-F) or one of these probes in combination with a probe for another X-linked gene, either *Hprt* or *Atrx*, in the trophoblast of E7.5 $XX^{CAGΔ5'}$ compared with that in wild-type and XX^{CAG} embryos (Fig. 2A-D). It was evident that an *Xist* RNA cloud was formed in the nucleus of all of these types of embryos, indicating that the $Xist^{CAGΔ5'}$ allele was upregulated as expected and that its transcript coated the X chromosome (Fig. 1E,F). Although the majority of cells in the extraembryonic tissues of $XX^{CAGΔ5'}$ exhibit dispersed *Xist* signals, such dispersed signals as well as the typical condensed signals were found in the wild-type and XX^{CAG} controls. Apparent differences in the morphology of the representative *Xist* cloud in $XX^{CAGΔ5'}$ and that in wild type and XX^{CAG} can most probably be ascribed to the difference in the cell type that composes the extraembryonic tissues. Given that the ectoplacental cone and extraembryonic ectoderm are essentially missing in $XX^{CAGΔ5'}$ embryos (Fig. 1C), we suspect that those cells with the condensed signals in wild type and XX^{CAG} most likely belong to the tissues that are missing in $XX^{CAGΔ5'}$ embryos.

The hybridization signals of other X-linked genes were detected in the cloud of $Xist^{CAGΔ5'}$ RNA in addition to those on the active X chromosome (91% for *Hprt* and 70% for *Atrx*; Fig. 2A-D). This contrasted with the pattern observed in the trophoblast of wild-type and XX^{CAG} embryos, where the expression of *Xist* and X-linked genes were mutually exclusive in the majority of nuclei positive for *Xist* (83% for *Hprt* and 73% for *Atrx* in wild type; 79% for *Hprt* and 75% for *Atrx* in XX^{CAG}). This demonstrates that, as expected from a previous study using differentiating ESCs (Wutz et al., 2002), $Xist^{CAGΔ5'}$ RNA fails to silence genes on the X chromosome on which it accumulates in the embryo. To investigate whether the silencing defects of $Xist^{CAGΔ5'}$ RNA could already be detected during the preimplantation stage, the same analysis was carried out in E3.5 blastocysts (Fig. 3A-F). The prevalence of nuclei with the mutated *Xist* RNA cloud in $XX^{CAGΔ5'}$ blastocysts (81%) was comparable to that with the *Xist* cloud in wild-type XX (89%) and XX^{CAG} (81%) blastocysts (Fig. 3A,B). Whereas *Hprt* and *Atrx* were apparently silenced on the X coated with the *Xist* RNA in 73-80% of the nuclei in wild-type and XX^{CAG} blastocysts, expression of these genes was detected in the cloud formed by $Xist^{CAGΔ5'}$ RNA in the

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

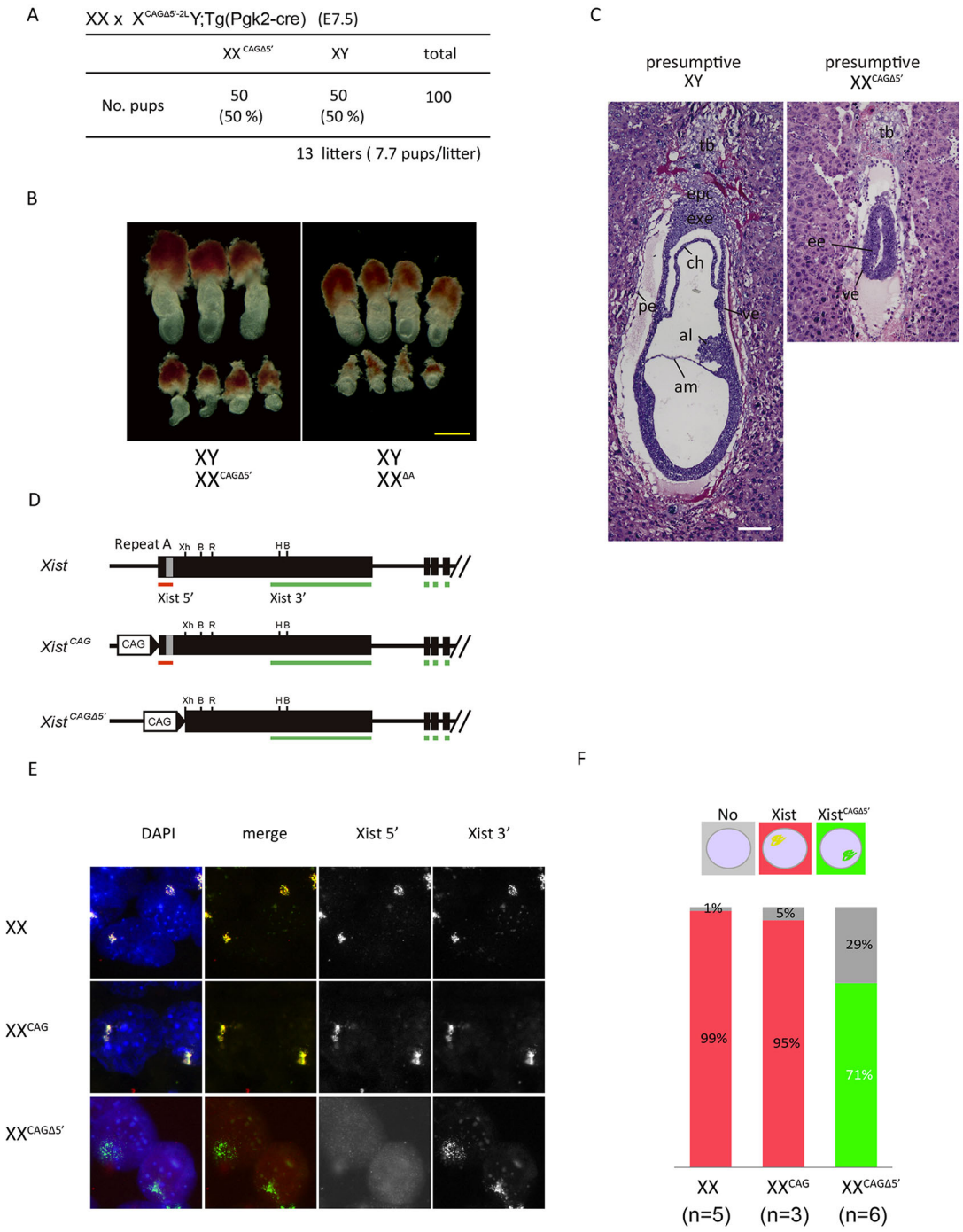

Fig. 1. Imprinted XCI is compromised in the embryo with the paternal X coated with *Xist*$^{CAGΔ5'}$ RNA. (A) The number of E7.5 mouse embryos recovered from wild-type females (XX) crossed with X$^{CAGΔ5'-2L}$Y; *Pgk2-cre* males. The *Xist*$^{CAGΔ5'}$ allele converted from the *Xist*$^{CAGΔ5'-2L}$ allele in the spermatocytes is transmitted to female embryos at the expected ratio. (B) Gross morphology of E7.5 embryos carrying the paternally transmitted *Xist*$^{CAGΔ5'}$ allele as compared with those carrying the paternally transmitted *Xist*^{ΔA} allele. Female embryos (bottom), which carry paternally transmitted *Xist*$^{CAGΔ5'}$ (left) or *Xist*^{ΔA} (right), are shown in comparison with wild-type male littermates (top). The phenotype of females obtained from the respective crosses is similar. Scale bar: 500 µm. (C) Histological analysis of a presumptive female embryo carrying the paternal *Xist*$^{CAGΔ5'}$ recovered at E7.5 comparison with that of a presumptive wild-type male littermate. al, allantois; am, amnion; ch, chorion; ee, embryonic ectoderm; epc, ectoplacental cone; exe, extra-embryonic ectoderm; pe, parietal endoderm; tb, trophoblast; ve, visceral endoderm. Scale bar: 200 µm. (D) The wild-type *Xist*, *Xist*^{CAG} and *Xist*$^{CAGΔ5'}$ alleles, and positions of the probes used for the detection of *Xist* RNA. The Xist3′ probe (green, pXist_SS12.5) detects both wild-type *Xist* and *Xist*$^{CAGΔ5'}$ RNA, whereas the Xist5′ probe (red, pX21Xh) detects only wild-type *Xist* RNA containing the 5′ region. Xh, *Xho*I; B, *Bgl*II; R, *Eco*RI; H, *Hind*III. (E) Accumulation of *Xist*$^{CAGΔ5'}$ RNA was confirmed by RNA-FISH in the E7.5 trophoblast. XX, wild-type female embryo; XX^{CAG}, female embryo carrying paternal *Xist*^{CAG}; XX$^{CAGΔ5'}$, female embryo carrying paternal *Xist*$^{CAGΔ5'}$. Whereas RNA expressed from wild-type *Xist* and *Xist*^{CAG} is detected by both Xist5′ and Xist3′ probes to yield a yellow hybridization signal when the images are merged, the mutant RNA expressed from *Xist*$^{CAGΔ5'}$ is detected only by the Xist3′ probe to yield a green hybridization signal. (F) Summary of RNA-FISH analysis in E7.5 trophoblast. Prevalence of the nuclei showing the respective expression patterns of *Xist* is shown: light gray, no *Xist*; gray, *Xist* RNA detected by both Xist5′ and Xist3′ probes; black, *Xist* RNA detected by only Xist3′ probe. *n*, the number of embryos analyzed. At least 100-150 nuclei were counted in each case.

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

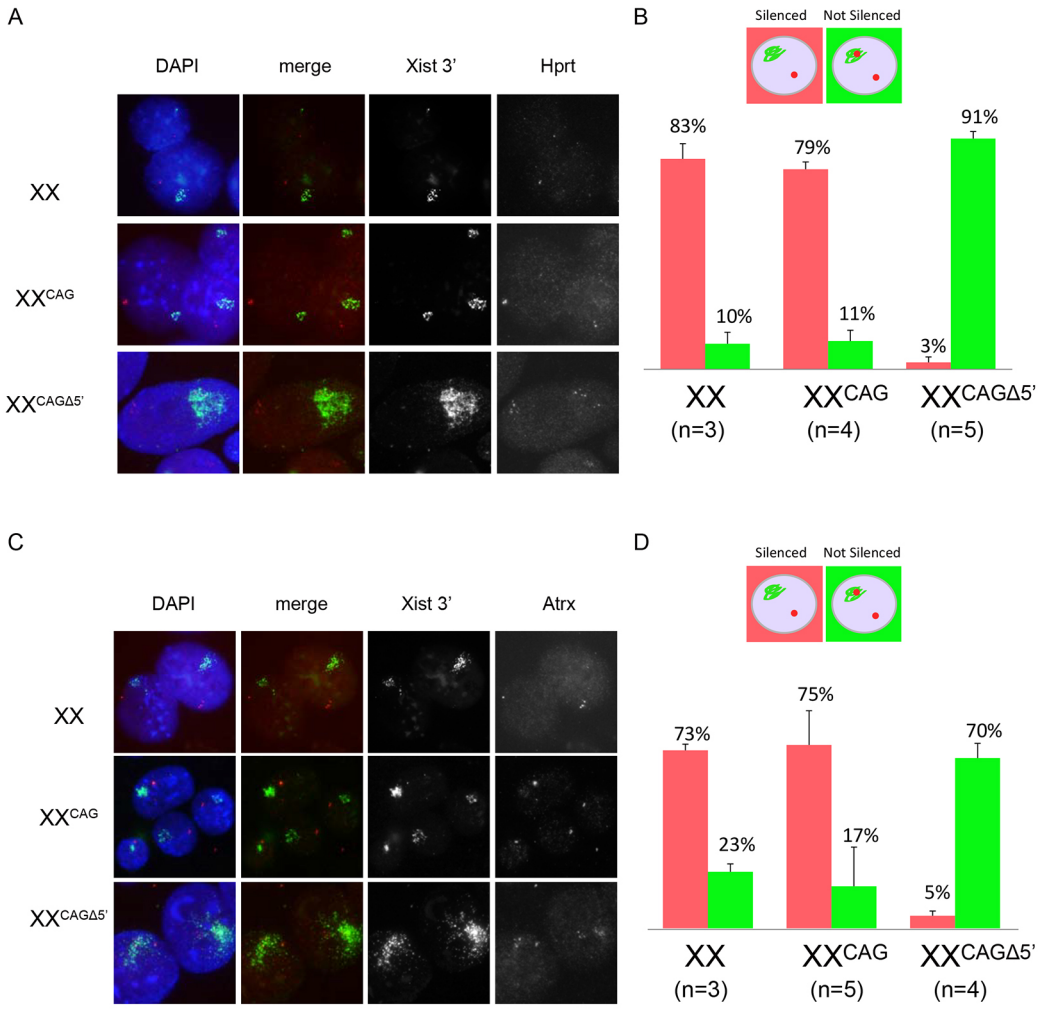

Fig. 2. *Xist*$^{CAGΔ5'}$ RNA is defective in X-linked gene silencing in the trophoblast. (A,C) Representative images of two-color RNA-FISH in E7.5 XX$^{CAGΔ5'}$ trophoblast detecting (A) *Hprt* or (C) *Atrx* in combination with *Xist*. Neither *Hprt* nor *Atrx* was silenced in the cloud formed by *Xist*$^{CAGΔ5'}$ RNA in the majority of the nuclei. (B,D) Summary of (B) *Hprt* or (D) *Atrx* silencing in the *Xist* cloud in XX, XX^{CAG} and XX$^{CAGΔ5'}$. *n*, number of embryos analyzed. Error bars indicate s.d.

majority of the nuclei in XX$^{CAGΔ5'}$ blastocysts (83% for *Hprt* and 93% for *Atrx*), indicating the failure of gene silencing on X$^{CAGΔ5'}$ from the early phase of imprinted XCI at the preimplantation stage (Fig. 3C-F). We therefore concluded that *Xist*$^{CAGΔ5'}$ RNA upregulated on Xp at the onset of imprinted XCI could coat the X chromosome but failed to induce proper silencing.

Xist$^{CAGΔ5'}$ RNA still retains some silencing ability in the trophoblast

To further investigate the extent to which the silencing function of *Xist* RNA was abolished in the absence of the A-repeat, we carried out allele-specific RNA sequencing (RNA-seq) using the trophoblast of E7.5 embryos. In addition to wild-type and XX$^{CAGΔ5'}$ embryos, XX^{ΔA} embryos carrying the functionally null *Xist*^{ΔA} allele (Hoki et al., 2009) were also included in the assay. Females of the JF1 strain were used for the respective crosses, so that the parental origin of the transcripts could be distinguished according to differences in the single-nucleotide polymorphisms (SNPs) and insertions/deletions (indels) present between JF1 and relevant males of B6 background.

Among X-linked genes expressed in the trophoblast, 319 informative genes were selected for the evaluation of misexpression from Xp in the mutants. Fig. 4A shows the number of genes with respective percentage paternal reads in increments of 5%, which represents the frequency of the paternal reads that appeared in the total allele-specific reads mapped at the genomic loci of these genes. The vast majority of genes in the wild-type trophoblast were included in the 0 to 10% category, indicating that the expression of these genes was substantially repressed on Xp, as expected. Several genes exhibiting more than 10% frequency were classified as being expressed from the inactive Xp and could, therefore, be referred to as 'escapees'.

The same analysis was carried out in XX^{ΔA} and XX$^{CAGΔ5'}$ trophoblasts. XX^{ΔA}, which does not express *Xist* RNA at all in the extraembryonic tissues (Fig. S2C), served as a control for the status of 'no dosage compensation' for the X-linked genes. Expression levels of paternal *Xist* were ~4-fold higher in XX$^{CAGΔ5'}$ than in wild type, ruling out the possibility that the lower expression of the mutant *Xist* RNA contributes to the defective silencing (Fig. S2C). A large proportion of the selected X-linked genes were distributed at around 50% in the XX^{ΔA} trophoblast, suggesting that these genes were expressed from the paternal allele at levels comparable to those from the maternal allele. The presence of a minor fraction of genes expressed more highly on the Xp than on Xm (up to 80%) could be due not only to the underlying genetic differences but also to indirect effects of the failure in dosage compensation (see below).

By contrast, percentage paternal reads in the XX$^{CAGΔ5'}$ trophoblast were rather variable, with a bimodal distribution at

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

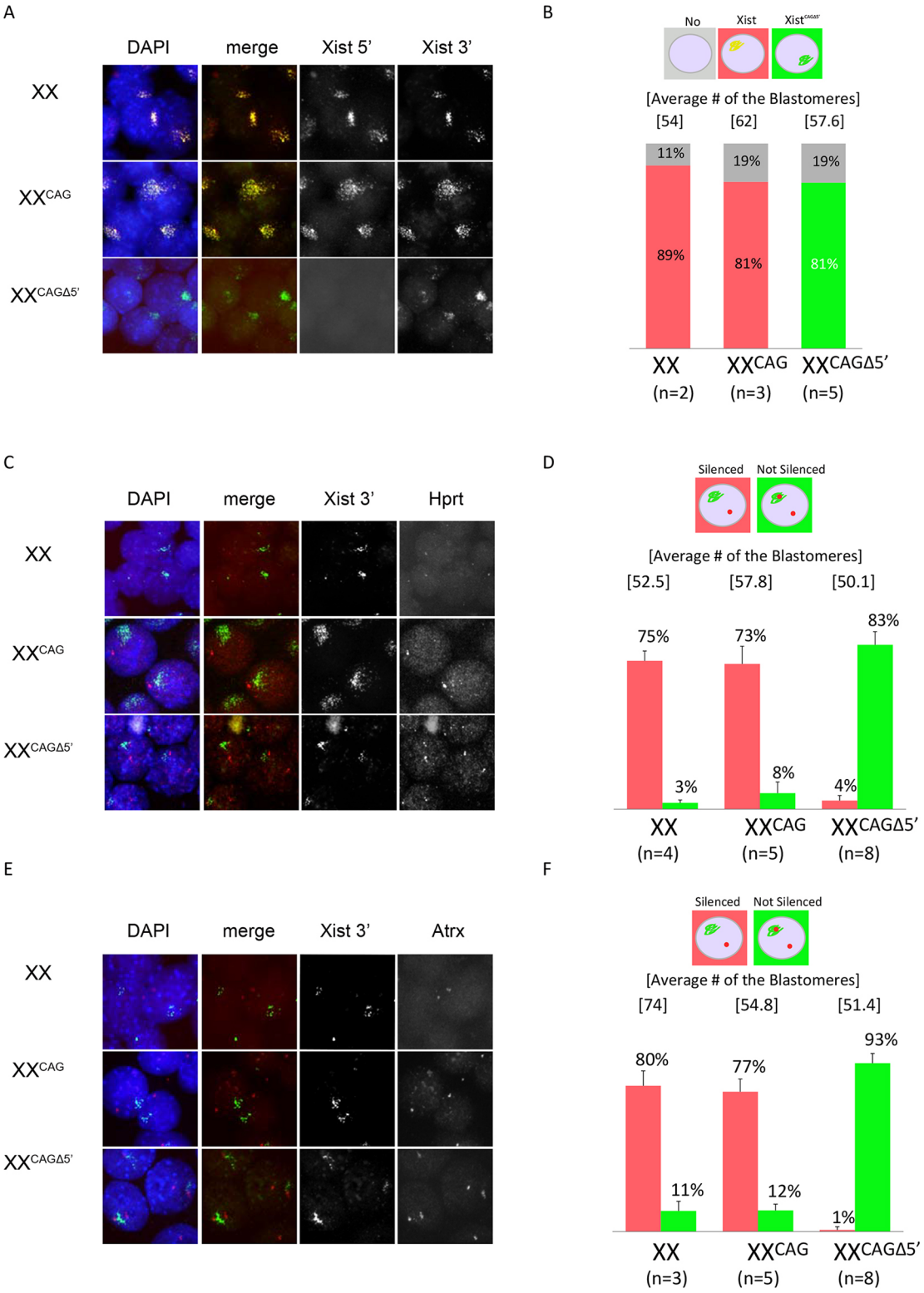

Fig. 3. Defective X-linked gene silencing by *Xist*$^{CAGΔ5'}$ RNA is evident in the blastocyst. (A) Representative images of two-color RNA-FISH detecting wild-type *Xist*, *Xist*^{CAG} and *Xist*$^{CAGΔ5'}$ RNA using the probes shown in Fig. 1C. (C,E) Representative images of two-color RNA-FISH in blastocysts detecting (C) *Hprt* or (E) *Atrx* in combination with *Xist*. Neither *Hprt* nor *Atrx* was silenced in the cloud formed by *Xist*$^{CAGΔ5'}$ RNA in the majority of the nuclei. (B,D,F) Summary of RNA-FISH detecting (B) wild-type and the mutated *Xist* RNAs, (D) *Hprt* and (E) *Atrx* in combination with *Xist*. The average number of blastomeres examined in each case is shown above the bar. *n*, number of embryos analyzed. *Xist*$^{CAGΔ5'}$ RNA fails to silence the X-linked genes in the preimplantation stages. Error bars indicate s.d.

around 0% and 50% (Fig. 4A). This is clearly distinct from the distribution of percentage paternal reads in XX^{ΔA}. Such a bimodal distribution in XX$^{CAGΔ5'}$ is specific for X-linked genes, as autosomal genes are expressed from both parental alleles at comparable levels, as shown by a peak at 50% (Fig. S2A,B). This demonstrates that a subset of X-linked genes are silenced or nearly

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

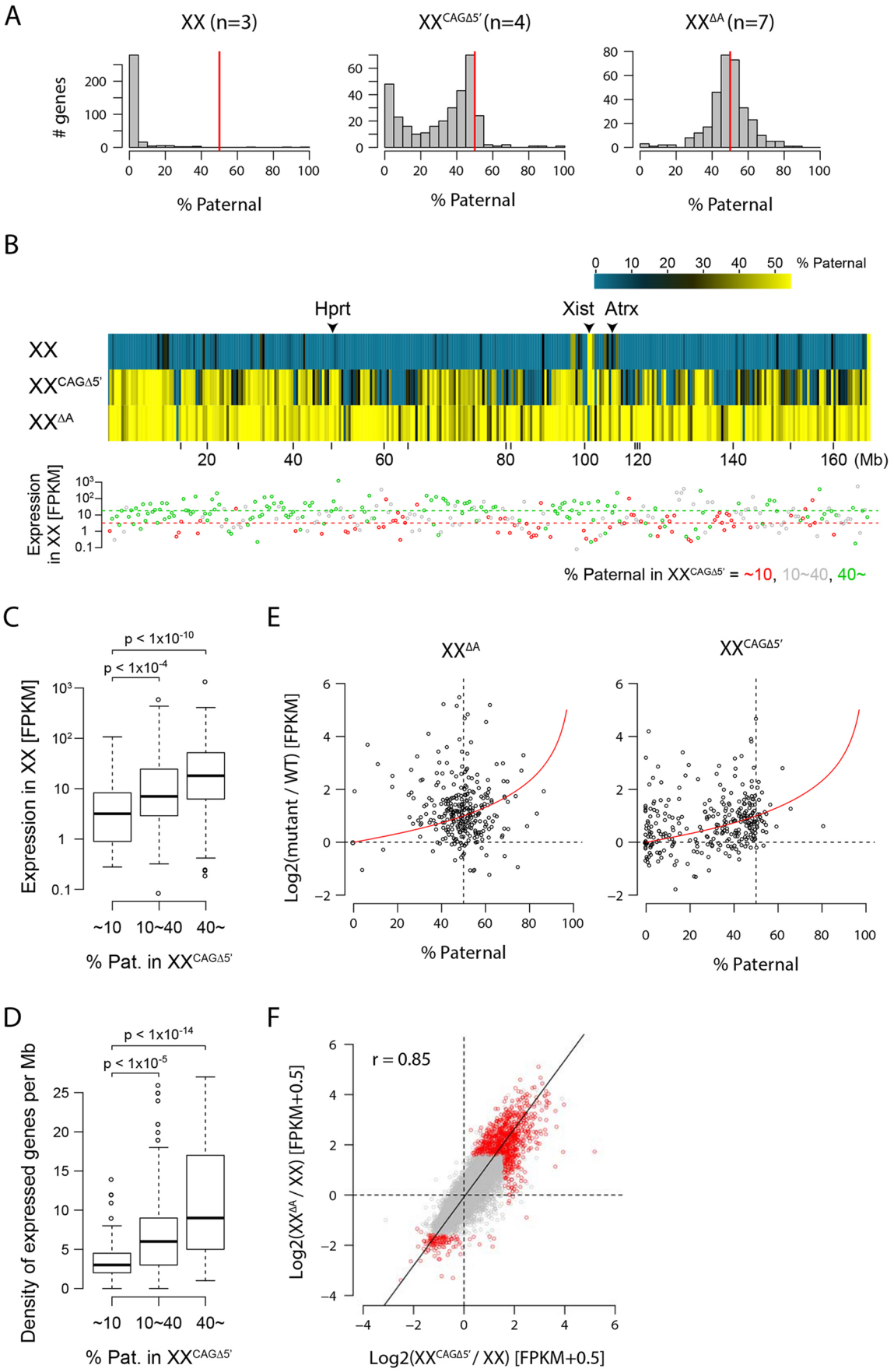

Fig. 4. See next page for legend.

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

Fig. 4. Failure of imprinted XCI impacts the expression of not only X-linked but also autosomal genes in the trophoblast. (A) The number of genes with respective percentage paternal reads among 319 X-linked genes analyzed. The vast majority of genes on the Xp are classified into 0-5% paternal reads in wild type (XX), indicating that they are substantially silenced (left). Most of the genes on $X^{\Delta A}$ are classified into 40-60% paternal reads, indicating that they are expressed from both X chromosomes at comparable levels in $XX^{\Delta A}$ (right). Although many genes are expressed at relatively high levels on $X^{CAG\Delta 5'}$, a subset of genes are classified into 0-10% paternal reads, indicating that they are inactivated by $Xist^{CAG\Delta 5'}$ RNA in $XX^{CAG\Delta 5'}$ (middle). (B) Genes that are inactivated and those that stay active often form clusters on $X^{CAG\Delta 5'}$. Data shown in A are delineated as a heatmap according to the position of 319 genes along the X chromosome. Those genes manifesting fewer than 20% paternal reads are classified as inactivated genes (cyan), whereas all others are classified as active (yellow). Positions of *Xist*, *Hprt* and *Atrx* are indicated. Expression levels of the respective genes in XX are shown beneath the heatmap. Dashed lines indicate the median of FPKM values. (C) Whether the genes on $X^{CAG\Delta 5'}$ are silenced or not in $XX^{CAG\Delta 5'}$ correlates with their expression levels in wild-type female embryos. The 319 genes analyzed are classified into three groups according to their percentage paternal reads (<10%, 71 genes; 10-40%, 103 genes; >40%, 145 genes) and their average expression levels (FPKM, log scale) in the wild-type XX trophoblast are compared among the three groups. Genes expressed more highly on $X^{CAG\Delta 5'}$ appear to be more highly expressed in XX. The significance of the difference was evaluated by the Wilcoxon test. (D) Whether the genes on $X^{CAG\Delta 5'}$ are silenced or not in $XX^{CAG\Delta 5'}$ correlates with gene (FPKM ≥1 in XX) density. The gene densities are compared among the three groups classified in C. (E) The expression levels of X-linked genes are variably affected in the absence of proper XCI in the trophoblast. For 295 genes that are silenced in wild-type Xp (percentage paternal reads ≤10), their percentage paternal reads and expression levels in $XX^{\Delta A}$ trophoblast relative to the values in wild-type trophoblast are plotted (left). The same analysis in $XX^{CAG\Delta 5'}$ trophoblast is also shown (right). The red curve indicates the expected relationship between the percentage paternal reads and expression levels assuming that the genes on the Xm are expressed at the same levels in wild type and mutants and those on the Xp in wild type are completely silenced. (F) The expression levels of autosomal genes are also affected in both $XX^{CAG\Delta 5'}$ and $XX^{\Delta A}$. Among autosomal genes, those with FPKM greater than 1 in any one of wild type, $XX^{CAG\Delta 5'}$ or $XX^{\Delta A}$ (11,688 genes) were analyzed. Changes in their expression levels between $XX^{CAG\Delta 5'}$ and $XX^{\Delta A}$ relative to wild-type trophoblast are compared. Differentially expressed genes in either $XX^{CAG\Delta 5'}$ or $XX^{\Delta A}$ (change of expression ≥3-fold and q-value<0.05) are indicated in red. The major axis regression line is shown (solid line).

silenced on $X^{CAG\Delta 5'}$ in the trophoblast, suggesting that $Xist^{CAG\Delta 5'}$ RNA still retains some residual silencing ability. Interestingly, those genes silenced on $X^{CAG\Delta 5'}$ are not randomly distributed but tend to be clustered along the chromosome and often found in regions of low gene density, and their expression levels are generally low in the wild type (Fig. 4B-D, Fig. S2D). By contrast, the X-linked genes that $Xist^{CAG\Delta 5'}$ RNA failed to silence were expressed at relatively high levels in wild type (Fig. 4C).

Lack of dosage compensation affects gene expression globally in the trophoblast

We further explored the relationship between percentage paternal reads of the X-linked genes silenced in the wild type and the changes in their expression levels (FPKM) relative to those in $XX^{CAG\Delta 5'}$ and $XX^{\Delta A}$ trophoblasts (Fig. 4E). The red curve in Fig. 4E indicates the relationship between the relative expression levels and percentage paternal reads, assuming that the expression levels of Xm genes are the same between wild type and the respective mutants, and that Xp genes in wild type are completely repressed. Many genes make a cluster on the red curve at around 50% paternal reads in $XX^{\Delta A}$, where the expression levels of these genes are around twice those from Xm in wild type, suggesting that the increase in their expression levels in $XX^{\Delta A}$ can be essentially ascribed to the misexpression from Xp. There are, however, many other genes that deviate from the red curve, indicating that their expression is misregulated on Xm as well as on Xp, which fails to undergo inactivation in $XX^{\Delta A}$. Expression levels of those genes shifted upward from the red curve are increased on both Xp and Xm relative to the respective X chromosomes in wild type, whereas those shifted downward represent genes whose expression on Xm is decreased compared with that on Xm in wild type. This suggests that the failure of imprinted X inactivation in the trophoblast results in the aberrant expression of genes on not only Xp but also Xm.

The same analysis in $XX^{CAG\Delta 5'}$ revealed that, in contrast to $XX^{\Delta A}$, many of the misregulated genes were distributed between 0% and 50% paternal reads at variable expression levels, most probably indicating the residual silencing ability of $Xist^{CAG\Delta 5'}$ RNA. It should be noted that although many genes manifest 0-10% paternal reads they are not necessarily silenced but, in fact, those that deviate upward from the red curve are expressed more strongly on the paternal $X^{CAG\Delta 5'}$ than those on Xp in wild type. Genes clustering on the red curve at 0% paternal reads represent those silenced on $X^{CAG\Delta 5'}$ to the same extent as on Xp in wild type. The failure of appropriate silencing of the paternal $X^{CAG\Delta 5'}$ in the trophoblast therefore results in misregulation of genes on not only Xp but also Xm, as was the case in $XX^{\Delta A}$.

These findings prompted us to examine the interesting possibility that this failure also causes global changes in gene expression. Fig. 4F and Fig. S2E show a comparison between the changes in autosomal gene expression in $XX^{CAG\Delta 5'}$ and those in $XX^{\Delta A}$. The changes in expression for 11,688 autosomal genes were very similar in each mutant relative to wild type, as plotted along the diagonal with a correlation coefficient of 0.85. Among these autosomal genes, those expressed at levels at least 3-fold higher or lower in the mutants relative to wild type with q-value <0.05 (red dots) were considered to be significantly upregulated or downregulated in each mutant. The number of upregulated genes was much greater than that of downregulated genes. It should be noted that upregulation is more prominent in $XX^{\Delta A}$ than in $XX^{CAG\Delta 5'}$ (Fig. 4F, Fig. S2C), suggesting more severe effects on autosomal gene regulation in the former. These findings raise the intriguing possibility that dosage compensation impacts not only X-linked genes but also the regulation of gene expression genome-wide in the trophoblast.

We carried out an in-depth analysis of Gene Ontology (GO) functional annotation for autosomal genes dysregulated in the trophoblast of $XX^{CAG\Delta 5'}$ and found that specific pathways were affected (Tables S1 and S2). Genes involved in developmental processes (e.g. cell motility, cell proliferation, cell death, and signal transduction) and in protein translation (e.g. rRNA and tRNA metabolic processes) were overrepresented in the upregulated and downregulated genes, respectively. Interestingly, downregulation of genes related to translation is one of the responses in gene expression commonly observed in aneuploid cells in diverse species (Sheltzer et al., 2012), suggesting that the presence of an additional active X chromosome has effects reminiscent of autosomal aneuploidy.

Xist RNA is biallelically expressed in embryonic tissue of $XX^{CAG\Delta 5'}$

To further investigate the silencing ability of $Xist^{CAG\Delta 5'}$ RNA, we examined XCI status in the embryonic lineage. RNA-FISH was carried out for the expression of *Xist* in the distal part of E7.5 $XX^{CAG\Delta 5'}$ embryos, in which embryonic ectoderm was enriched. Unexpectedly, we noted that most of the nuclei contained two *Xist* clouds. To confirm that these two *Xist* clouds represented biallelic expression from the wild-type X and $X^{CAG\Delta 5'}$, we carried out RNA-

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

FISH using the two *Xist* probes described above (Fig. 5A,B). The results demonstrated that one cloud represents $Xist^{CAG\Delta5'}$ RNA and the other the wild-type *Xist* RNA. RNA-FISH for the expression of either *Hprt* or *Atrx*, when examined simultaneously with a probe common to both wild-type and mutated *Xist* RNA (Xist3′), revealed that expression of one of the two alleles was detected as a single pinpoint in one of the two *Xist* clouds in the majority of the nuclei (70% for *Hprt* and 68% for *Atrx*; Fig. 5C-F). Although we could not distinguish which of the wild-type and mutant *Xist* clouds overlapped with a pinpoint signal of the X-linked genes, it would be reasonable to assume that *Xist* RNA expressed from the wild-type allele could trigger the silencing of the X-linked genes in cis, whereas the mutant $Xist^{CAG\Delta5'}$ RNA failed to do so.

Both X chromosomes acquire histone modifications typical of the normally inactivated X in $XX^{CAG\Delta5'}$ embryonic tissue

An X chromosome that is to be inactivated undergoes a series of changes in histone modifications during the initiation phase of XCI, which apparently depends on *Xist* RNA accumulation on the X chromosome. To examine whether $Xist^{CAG\Delta5'}$ RNA was competent to induce histone modifications, we carried out whole-mount immunofluorescence staining of E7.5 $XX^{CAG\Delta5'}$ embryos using antibodies against various histone modifications. Immunofluorescence staining was carried out with antibody against H3K27me3 in combination with anti-Oct3/4 (Pou5f1) antibody, which allowed us to distinguish the cells of the embryonic tissue from those of the extraembryonic tissues. As was shown by RNA-FISH for *Xist*, two H3K27me3 domains were detected in the nucleus of Oct3/4-positive cells (representing the embryonic tissue), whereas only one such domain was detected in the nucleus of Oct3/4-negative cells (representing the extraembryonic tissues) (Fig. 6A). We also analyzed the acetylation status of the H3K27me3 domains, which we assumed represented the X chromosome coated with *Xist* RNA. Immunofluorescence of acetylated histone H4 was excluded from the H3K27me3 domains in both the embryonic and extraembryonic tissues, suggesting hypoacetylation of the X chromosome highlighted with H3K27me3 (Fig. 6B). These analyses indicate that the X chromosome coated with $Xist^{CAG\Delta5'}$ RNA, which is defective in silencing, is apparently indistinguishable from the normally inactivated X chromosome in terms of the status of these histone modifications.

To further analyze the epigenetic state of the X chromosome coated with $Xist^{CAG\Delta5'}$ RNA, other histone modifications or the localization of proteins known to be enriched on the inactive X chromosome were examined in newly generated ESCs from $XX^{CAG\Delta5'}$ blastocysts. Embryoid body formation of $XX^{CAG\Delta5'}$ ESCs successfully induced both mutant and normal *Xist* RNA, as observed in the epiblast of E7.5 embryos (Fig. 6C). Immunofluorescence staining or immuno-RNA-FISH showed that there were no obvious differences in the epigenetic modifications between the wild-type X and $X^{CAG\Delta5'}$, both of which were highlighted with *Xist* RNA and H3K27me3 (Fig. 6D-F). This is consistent with a previous finding in ESCs that histone modifications such as H3K27me3 and H4K20me1 are still induced by *Xist* RNA lacking the A-repeat (Kohlmaier et al., 2004; Pullirsch et al., 2010). We concluded that $Xist^{CAG\Delta5'}$ RNA, although defective in chromosome silencing, could induce global changes in histone modifications on the X chromosome from which it originated, in a manner similar to the wild-type *Xist* RNA.

DISCUSSION

The A-repeat is essential for XCI during embryogenesis

The importance of the A-repeat in the silencing function of *Xist* RNA was first shown by Wutz and colleagues using a transgenic approach in differentiating ESCs (Wutz et al., 2002). Although their assay system has provided significant insights into *Xist* RNA-mediated chromosome silencing, the possible impact of *Xist* RNA lacking the A-repeat on XCI has not been addressed in the context of embryonic development. In this study, we showed for the first time that $Xist^{CAG\Delta5'}$ RNA, which lacks the 5′ region including the A-repeat, is dysfunctional and fails to induce proper XCI in both the embryonic and extraembryonic lineages, and we further analyzed its effects on XCI in the embryo. Upon paternal transmission of the $Xist^{CAG\Delta5'}$ allele, female embryos died at the early postimplantation stage due to severe developmental defects in the extraembryonic tissues caused by the failure of imprinted XCI. Allele-specific RNA-seq revealed that many genes on the Xp coated with $Xist^{CAG\Delta5'}$ RNA in the trophoblast were significantly misexpressed, but some of the genes located on this chromosome were still properly silenced, suggesting that although the silencing function of $Xist^{CAG\Delta5'}$ RNA was severely compromised, it retains some silencing ability.

Although the extraembryonic ectoderm is essentially missing, the trophoblast is relatively well formed in $XX^{CAG\Delta5'}$ embryos, as compared with that in $XX^{\Delta A}$ embryos. Thus, partial dosage compensation mediated by the apparent residual silencing ability of $Xist^{CAG\Delta5'}$ RNA may allow relatively normal development of the trophoblast, but this extent of dosage compensation might not be sufficient to support the development of the extraembryonic ectoderm in $XX^{CAG\Delta5'}$ embryos. Accordingly, the trophoblast might be more tolerant than the extraembryonic ectoderm to the failure of imprinted X inactivation.

Embryonic tissue of $XX^{CAG\Delta5'}$ shows biallelic expression of *Xist* RNA to accomplish dosage compensation

One of the unexpected findings in this study is the biallelic expression of *Xist* in the embryonic tissues of $XX^{CAG\Delta5'}$ embryos. Since *Xist* RNA is exclusively expressed from the $Xist^{CAG}$ allele in the embryonic tissue of XX^{CAG} (Amakawa et al., 2015), we expected that only the $Xist^{CAG\Delta5'}$ allele would be upregulated in $XX^{CAG\Delta5'}$ embryonic tissue. This was not the case, however, and the wild-type *Xist* allele was also upregulated in addition to the mutated $Xist^{CAG\Delta5'}$ allele, implying that the wild-type *Xist* allele was secondarily upregulated so as to achieve dosage compensation in $XX^{CAG\Delta5'}$ embryos, in which $Xist^{CAG\Delta5'}$ RNA failed to induce a sufficient level of XCI.

A strict requirement for dosage compensation during embryonic development might underlie upregulation of the second, wild-type *Xist* allele. A recent model proposes the action of a putative X-linked activator for the initiation of XCI, which can stochastically upregulate every *Xist* allele regardless of the number present in a cell in a dose-dependent manner. X-linked Rnf12 (Rlim) is a known candidate for such an activator (Gontan et al., 2012; Monkhorst et al., 2008). According to this model, the dose of Rnf12 produced from two active X chromosomes is high enough to stochastically activate either of the *Xist* alleles or both, but once one of the X chromosomes subsequently undergoes inactivation the dose of Rnf12 produced from the remaining active copy would never surpass the threshold for *Xist* activation, and therefore could induce no more *Xist*. This would be consistent with our current observation of biallelic expression of *Xist* in $XX^{CAG\Delta5'}$ embryos. Since the X chromosome coated with $Xist^{CAG\Delta5'}$ RNA does not undergo inactivation, it would be reasonable to assume that the dose of X-linked Rnf12 would remain high enough to activate the second *Xist* allele on the wild-type X, resulting in biallelic expression of *Xist* in the embryonic tissue. However, our preliminary result using the

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

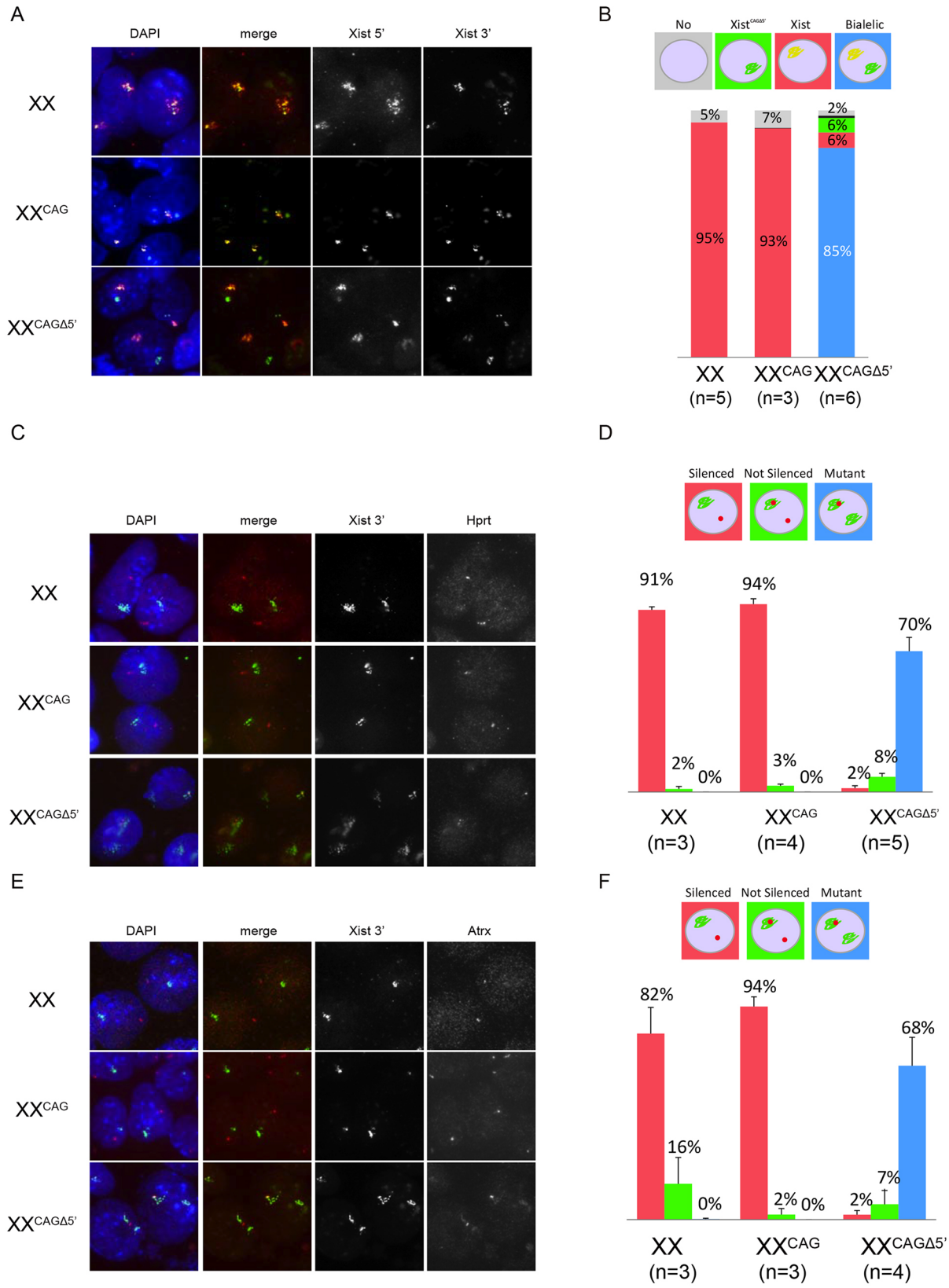

Fig. 5. Defective X-linked gene silencing by *Xist*$^{CAGΔ5'}$ RNA in E7.5 XX$^{CAGΔ5'}$ embryonic tissue. (A) Representative images of RNA-FISH using two different probes detecting *Xist* RNA. Accumulation of both wild-type and *Xist*$^{CAGΔ5'}$ RNA is evident in XX$^{CAGΔ5'}$. (B) Summary of the expression pattern of *Xist* in the embryonic tissue of the indicated genotypes. (C,E) Representative images of two-color RNA FISH for (C) *Hprt* or (E) *Atrx* expression in combination with *Xist*. (D,F) Summary of the expression pattern of (D) *Hprt* or (F) *Atrx* in embryonic tissue of the indicated genotypes. In the majority of the nuclei, expression of these X-linked genes was evident on one of the Xs, presumably X$^{CAGΔ5'}$. Error bars indicate s.d.

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

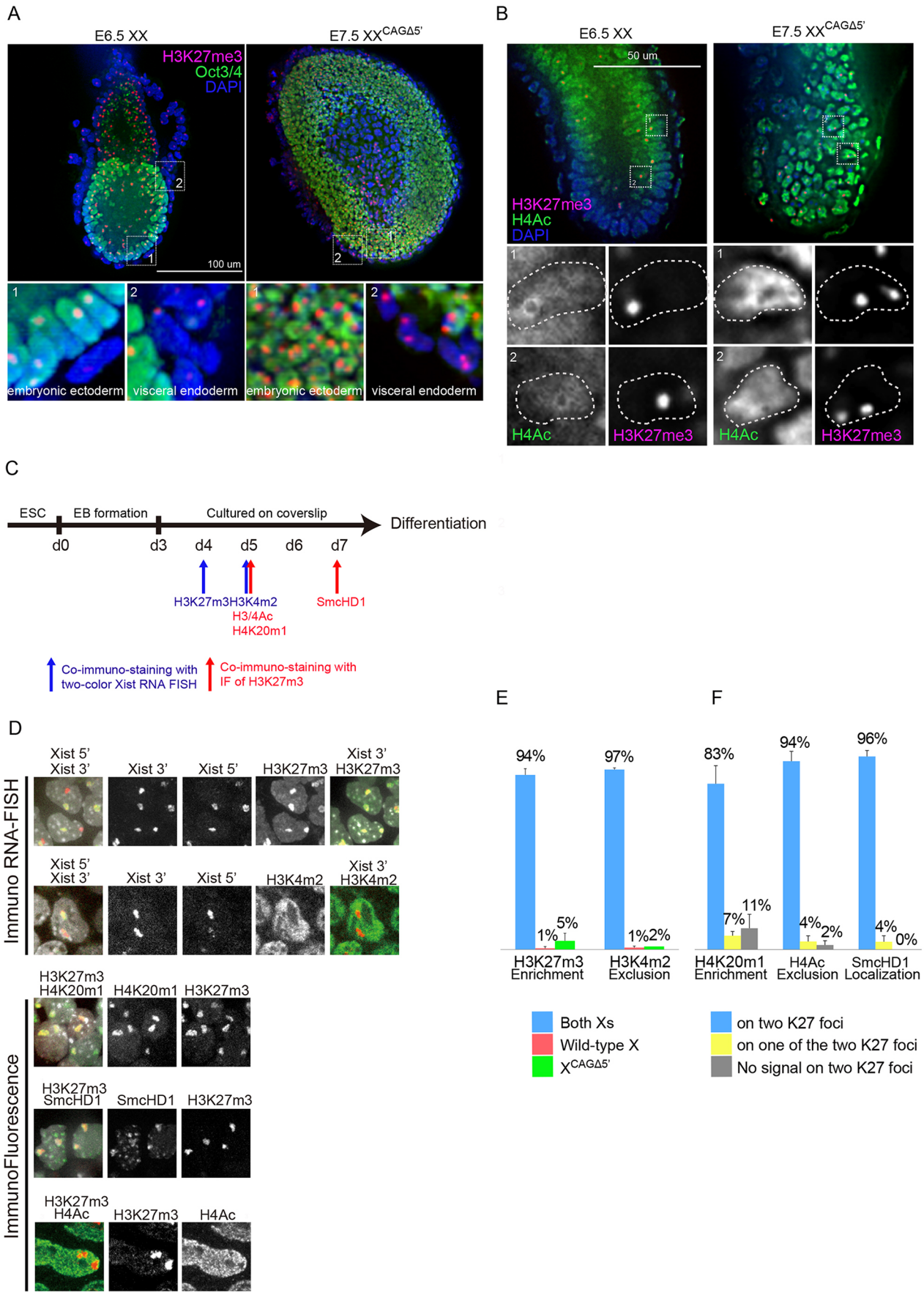

Fig. 6. See next page for legend.

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

Fig. 6. An X chromosome coated with *Xist*$^{CAGΔ5′}$ RNA assumes chromatin modifications typical of the normally inactivated X chromosome.
(A) Representative images of whole-mount immunofluorescence staining of E7.5 XX$^{CAGΔ5′}$ (n=3) and size-matched E6.5 XX (n=3) embryos with an antibody against H3K27me3 in combination with an antibody against Oct3/4. Cells positive for Oct3/4 represent the embryonic tissue, whereas those negative for Oct3/4 represent the extraembryonic tissues. In the XX$^{CAGΔ5′}$ embryo, each nucleus contains a single H3K27me3 domain in the extraembryonic tissues but two domains in the embryonic tissue, suggesting that *Xist*$^{CAGΔ5′}$ RNA is competent to induce H3K27me3. Boxed regions are magnified beneath. (B) Whole-mount immunofluorescence staining of E6.5 XX (n=3) and E7.5 XX$^{CAGΔ5′}$ (n=3) embryos with an antibody against acetylated histone H4 (H4ac) in combination with an antibody against H3K27me3. H4ac is excluded from the domain enriched for H3K27me3. Boxed regions are magnified beneath. (C) Timecourse of embryoid body (EB) differentiation. Arrows indicate days when histone modifications and localization of SmcHD1 (which localizes on the inactive X chromosome) were examined. (D) Representative images of immuno-RNA-FISH and immunofluorescence analyses of differentiated XX$^{CAGΔ5′}$ ESCs. Immuno-RNA-FISH was performed using two different probes for *Xist* RNA and an antibody against either H3K27me3 or H3K4me2 on ESCs differentiated for 4 days (H3K27me3) or 5 days (H3K4me2). H3K27m3 was enriched on, and H3K4me2 was excluded from, the X and X$^{CAGΔ5′}$ chromosomes coated with wild-type *Xist* and *Xist*$^{CAGΔ5′}$ RNA, respectively. Immunofluorescence staining was performed using one of antibodies against H4K20me1, SmcHD1 or H4ac in combination with an antibody against H3K27me3 on ESCs differentiated for 5 days (H4K20me1 and H4ac) or 7 days (SmcHD1). Two H3K27me3 domains that formed in differentiated XX$^{CAGΔ5′}$ ESCs, representing the wild-type X and X$^{CAGΔ5′}$ chromosomes, were enriched for H4K20me1 and SmcHD1 but devoid of H4ac. (E,F) Summary of (E) immuno-RNA-FISH and (F) immunofluorescence analyses of differentiated XX$^{CAGΔ5′}$ ESCs. The X$^{CAGΔ5′}$ chromosome assumes the properties of the normally inactivated X chromosome in the vast majority of nuclei.

distal part of early postimplantation XX$^{CAGΔ5′}$ embryos, which is enriched with epiblast cells, suggests that the upregulation of wild-type *Xist* is not necessarily preceded by the upregulation of *Xist*$^{CAGΔ5′}$ (Fig. S3). When RNA-FISH was carried out for *Xist* expression using two probes, which allowed us to differentiate wild-type and *Xist*$^{CAGΔ5′}$ RNA, we observed monoallelic upregulation of either wild-type or *Xist*$^{CAGΔ5′}$ RNA during the early phase of XCI in the epiblast cells. This suggests that a mechanism for the random choice of monoallelic upregulation is still effective on either the wild-type or *Xist*$^{CAGΔ5′}$ allele despite the fact that *Xist*$^{CAGΔ5′}$ is driven by the CAG promoter. This appears consistent with our recent finding that, in undifferentiated ESCs, the CAG promoter driving the *Xist*^{CAG} allele does not efficiently promote transcription but rather behaves in a manner similar to the endogenous *Xist* promoter, causing only a basal level of transcription (Amakawa et al., 2015). We anticipate that, given the constitutively active nature of the CAG promoter in many differentiated cells, monoallelic upregulation of the wild-type *Xist* allele at the onset of XCI should be followed by upregulation of *Xist*$^{CAGΔ5′}$, resulting in biallelic expression of *Xist* in the embryonic tissues of XX$^{CAGΔ5′}$ later on. Detailed analyses of XX$^{CAGΔ5′}$ ESCs would provide further insights into the mechanisms of action of putative activators of *Xist*, such as Rnf12, in the initiation of XCI.

X$^{CAGΔ5′}$ acquires chromatin modifications typical of the inactive X chromosome

An X chromosome acquires a series of epigenetic modifications following the accumulation of *Xist* RNA and forms a silencing compartment, into which X-linked genes that are to be silenced relocate (Escamilla-Del-Arenal et al., 2011; Patrat et al., 2009). Formation of this silencing compartment appeared to be independent of the A-repeat. Although repetitive sequences present in intergenic regions are incorporated in the silencing compartment formed by *Xist* RNA lacking the A-repeat, X-linked genes that remain active do not relocate into it, explaining why the mutated *Xist* RNA fails to induce XCI (Chaumeil et al., 2006).

Our immunofluorescence analysis demonstrated that the X chromosome coated with *Xist*$^{CAGΔ5′}$ RNA manifested histone modification patterns typical for the normally inactivated X chromosome in XX$^{CAGΔ5′}$, suggesting that *Xist*$^{CAGΔ5′}$ RNA could also form a silencing compartment in the embryos that is apparently indistinguishable from that of the properly silenced X chromosome, as previously described in differentiating ESCs (Chaumeil et al., 2006; Kohlmaier et al., 2004). It is likely that the histone modification patterns observed by immunofluorescence analysis in XX$^{CAGΔ5′}$ represent those distributed in the intergenic regions, which occupy most of the X chromosome, but not in gene regions that do not undergo inactivation. In this case, however, some genes juxtaposing the intergenic region that acquires repressive histone modifications might not need to relocate into the silencing compartment to become silenced. Alternatively, although the patterns of immunofluorescence were apparently indistinguishable, since it does not provide a quantitative measure there could be reduction of one or more histone modifications on the X coated with the mutated *Xist* RNA in XX$^{CAGΔ5′}$. This could also contribute to inefficient silencing. In the present study, we could not address these issues, since the mutant embryos were too small to prepare enough chromatin for a ChIP-seq analysis. Given the fact, however, that the X chromosome coated with the mutated *Xist* RNA in differentiating ESCs established from XX$^{CAGΔ5′}$ blastocysts recapitulates the effects seen in the embryos, such ESCs would provide an alternative source material for the assay. It would be of particular interest to compare the distribution of histone modifications between the wild-type X and X$^{CAGΔ5′}$, which are coated with the wild-type and the mutated *Xist* RNA, respectively, when present in the same nucleus in XX$^{CAGΔ5′}$ ESCs.

Biological significance of dosage compensation in female mammals

It was intriguing to find that the failure of proper imprinted XCI in the trophoblast did not result in a simple overexpression of X-linked genes but rather aberrant upregulation and downregulation of many genes not only on the X chromosome but also autosomes. This suggests that dosage compensation has a much greater impact on gene regulation genome-wide than previously thought in the trophoblast. This is reminiscent of a study demonstrating the improvement of somatic cell nuclear transfer (SCNT) by impeding *Xist* expression on the active X (Inoue et al., 2010). In SCNT embryos, *Xist* becomes upregulated on both X chromosomes in females and on the single X in males at the early preimplantation stages, resulting in remarkable downregulation of many X-linked genes at the blastocyst stage. The use of somatic cells that are either heterozygous or hemizygous for the *Xist* null mutation as a donor of the nucleus significantly improves the development of SCNT embryos, in which the expression levels of the X-linked genes are restored to relatively normal levels. Intriguingly, this is accompanied by a significant improvement in the expression levels of many autosomal genes, which are also misregulated in the SCNT embryos. Insufficiency of X-linked gene expression, in this case, apparently causes aberrant expression of autosomal genes. In XX^{ΔA} and XX$^{CAGΔ5′}$ trophoblasts, by contrast, an overexpression of X-linked genes due to the failure of imprinted XCI apparently affects autosomal gene expression.

These findings suggest that inappropriate levels of X-linked gene expression directly or indirectly affect the expression levels of autosomal genes. It is therefore likely that dosage compensation in female mammals is required not only for equalizing the dosage

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

difference of X-linked genes between the sexes but also for proper gene expression genome-wide.

MATERIALS AND METHODS

Construction of the targeting vector

A targeting vector was constructed by modifying pCAG-CΔM20, which was previously constructed to generate the *Xist*^{CAG2L} allele (Amakawa et al., 2015). pCAG-CΔM20 was digested with *Sal*I and *Xho*I to release a 3.7 kb CAG-Pac cassette and a 0.9 kb *Xist* 5′ fragment (nt 21-912). The reaction products were directly subjected to self-ligation and a plasmid containing only the 3.7 kb CAG-Pac cassette was isolated to derive pCAGΔA.

Gene targeting and mice

The targeting vector was introduced into R1 ESCs (Nagy et al., 1993) by electroporation using Gene Pulser (240 V, 500 µF; Bio-Rad). Selection was applied 24 h later in the presence of 2 µg/ml puromycin. Of 405 colonies picked up, one harbored the expected homologous recombination (*Xist*$^{CAGΔ5′-2L}$). Chimeric males were generated and crossed with females heterozygous for *Xist*^{CAG} (Amakawa et al., 2015) to facilitate the germline transmission of the *Xist*$^{CAGΔ5′-2L}$ allele. Female mice carrying *Xist*$^{CAGΔ5′-2L}$ in combination with *Xist*^{CAG} thus generated were crossed with *Pgk2-cre* transgenic males, which express Cre recombinase specifically in the spermatocyte, to recover X$^{CAGΔ5′-2L}$Y; *Pgk2-cre* males. These males were subsequently crossed with wild-type females to produce female embryos that inherited the *Xist*$^{CAGΔ5′}$ allele converted from *Xist*$^{CAGΔ5′-2L}$ during spermatogenesis. Correct targeting events and germline transmission were verified by Southern blotting.

XX^{CAG} and XX^{ΔA} embryos were obtained at E7.5 from crosses between wild-type females with X^{CAG2L}Y; *Pgk2-cre* and X^{ΔA}Y males, respectively (Amakawa et al., 2015; Hoki et al., 2009). For RNA-seq, JF1 females were used for crossing with the respective wild-type B6, X$^{CAGΔ5′-2L}$Y; *Pgk2-cre*, or X^{ΔA}Y males.

All mice were maintained and used in accordance with the Guidelines for the Care and Use of Laboratory Animals of Kindai University (KDAS-26-006).

RNA-FISH

To differentiate between signals for wild-type *Xist* and *Xist*$^{CAGΔ5′}$ RNA, two plasmids pXist_SS12.9 and pX21Xh that contain different parts of the *Xist* cDNA were labeled with either Cy3-dUTP (GE Healthcare) or Green-dUTP (Abbott Molecular) using a Nick Translation Kit (Abbott Molecular). Probes for *Hprt* and *Atrx* were prepared in the same way using BAC clones RP24-335G16 and RP23-260I15, respectively. Cytological preparations of postimplantation embryos and blastocysts were made as described by Takagi et al. (1982) and Okamoto et al. (2000), respectively. RNA-FISH was performed as described previously (Sado et al., 2001).

Histology

E7.5 embryos in the decidua were fixed with Bouin's fixative, dehydrated, embedded in paraffin, sectioned at 5 µm, and stained with Hematoxylin and Eosin.

Whole-mount immunofluorescence

Embryos dissected out from the decidua were fixed with 1% paraformaldehyde for 10 min at room temperature, permeabilized in PBS containing 0.5% Triton X-100 and 0.5% BSA for 15-20 min, and incubated with primary antibodies in PBS containing 0.5% BSA and 0.5% Tween 20 for 1 h at room temperature. Primary antibodies were: anti-H4ac rabbit polyclonal antibody (Millipore, 06-598; 1:200), anti-H3K27me3 mouse monoclonal antibody (a gift from Hiroshi Kimura; 1:1000) (Hayashi-Takanaka et al., 2011), anti-H3K27me3 rabbit polyclonal antibody (Millipore, 07-449; 1:100) and anti-H4K20me1 (a gift from Hiroshi Kimura; 1:100) (Hayashi-Takanaka et al., 2015). Secondary antibodies were Alexa Fluor 488-conjugated rabbit anti-goat IgG, CF594-conjugated goat anti-mouse IgG and Alexa Fluor 488-conjugated goat anti-mouse IgG (all Invitrogen; 1:1000). Fluorescent images were captured using an Olympus Disk Scanning Unit (DSU) mounted on an inverted microscope (IX71, Olympus) and an EM-CCD camera (iXon, Andor), using MetaMorph imaging software (Molecular Devices).

RNA-seq

Total RNA was isolated from the trophoblast of E7.5 embryos using Trizol (Invitrogen). Libraries were prepared using the TruSeq RNA Sample Prep Kit v2 (Illumina) using at least 20 ng total RNA, and were sequenced using an Illumina HiSeq2500 or 1500 instrument to generate 101 bp single-end reads. The numbers of biological replicates were three for wild type, seven for *Xist*$^{CAGΔ5′}$, four for *Xist*^{ΔA}, and one for *Xist*^{CAG}.

Genome annotations

For a set of the SNPs and indels (referred to as variants) of the JF1 strain, we applied a whole-genome *de novo* assembly strategy to JF1 genomic reads generated by Takada et al. (2013). Reads from the JF1-2 library [DDBJ Sequence Read Archive (DRA) accession numbers DRX000502 and DRX000503] were assembled into scaftigs using fermi v1.1-r751 (Li, 2012) with option '-k 60'. Those scaftigs were aligned to a reference genome (UCSC mm9), including unlocalized/unplaced scaffolds and one unit of rDNA repeats (GenBank accession number BK000964.1) as decoys, using bwa mem v0.7.10 (Li, 2013 preprint) with option '-D 0 -c 10000 -w 1000'. Scaftigs with low mapping quality (MAPQ<40) or overlapping satellite repeats (annotated as GSAT_MM and SYNREP_NM in the UCSC RepeatMasker track) were filtered out to avoid misidentification of variants. Using outputs of SAMtools v0.1.19 mpileup (Li et al., 2009), we extracted informative positions satisfying the following criteria in the reference genome: (1) a depth of scaftigs ≥1 and ≤3; (2) all bases among scaftigs were consistent when multiple scaftigs were overlapped; (3) indel length ≤40 bp. By comparing with the reference genome, we called variants for the JF1 strain. Note that our mutant mice were originally derived from ESCs of 129 background and retained the genomic sequence of the 129 strain in the vicinity of the mutated *Xist* locus even after extensive backcrosses into the B6 background. We also created a set of variants on the X chromosome of the 129S1 strain using genomic reads derived from the 129S1/SvImJ strain (Keane et al., 2011) [NCBI Sequence Read Archive (SRA) accession numbers ERX113419, ERX113423, ERX113427, ERX113431, ERX113435, ERX113439, ERX113447, and ERX113443] as above. These variants located at informative positions on the X chromosome between the JF1 and 129S1 strains were used for further analysis. We constructed the strain-specific genomes by incorporating the JF1 or 129S1 variants into the B6 genome (only the X chromosome was considered for the 129S1 allele). For gene annotation, RefSeq genes were obtained from the UCSC genome browser (accessed January 21, 2015), and genes corresponding to small RNAs were excluded.

Allele-specific RNA-seq alignment

We adapted our pipeline (Nozawa et al., 2013) to allow allele-specific mapping. We prepared a concatenated custom reference, including the strain-specific genomes and sequences spanning splice junctions based on the RefSeq genes for all possible alleles (B6, JF1, and 129S1 for the *Xist* mutants; B6 and JF1 for wild type), which allowed us to trace the possible allelic origins of each read easily and handle splicing during RNA-seq alignment. We filtered out RNA-seq reads derived from rRNA and mitochondrial transcripts by mapping to 45S RNA, 5S RNA, and mitochondrial DNA (GenBank accession numbers BK000964.1 and NR_030686.1; UCSC mm9 chrM). After trimming of the first two bases of each read, an error-rich stretch caused by random hexamer priming during library preparation, the filtered reads were aligned to the custom reference using bwa samse (Li and Durbin, 2009), and the best hits of each read were extracted. The coordinates of hits mapping to splice junctions and/or non-B6 sequences were converted to mm9 genomic coordinates, and redundancy in the coordinates for each read was removed while retaining their allelic information. As a result, each read had a set of coordinates in the mm9 reference and each coordinate had a set of possible allelic origins. Reads unmapped anywhere or with a number of coordinates >20 were excluded, and the remaining were classified into two categories: unique hits (number of coordinates=1) and multiple hits (all others).

Development (2017) 144, 2784-2797 doi:10.1242/dev.149138

Parental imbalance and abundance in gene expression

To assign the parental origin using a set of allelic origins at each variant site on the X chromosome, the haplotype of the Xp chromosome in each sample with the mutated *Xist* was required. We inferred two recombination boundaries between B6-derived and 129S1-derived regions on the Xp chromosome proximal and distal to the *Xist* locus using the distribution of unique hits occurring only in the B6 or 129S1 genome. We referred to unique hits with only one of the maternal or paternal origins as allele-specific reads, and used these to measure parental imbalance in gene expression. For each gene, we counted the number of maternal or paternal allele-specific reads (Nmat and Npat, respectively) overlapping exonic variants within the gene, while ignoring variants overlapping multiple hits in order to control for mapping bias caused by differences in uniquely mappable positions among strain-specific genomes (Degner et al., 2009). Allele-specific reads overlapping distinct genes were basically ignored, but used in certain cases to create union gene models merging inseparable genes. For each gene, when the number of allele-specific reads was ≥10, the value of percentage paternal reads was defined as 100 Npat/(Nmat+Npat); otherwise, as a missing value. The representative value of percentage paternal reads for each *Xist* genotype was the average of biological replicates when at least half of the replicates were available. To facilitate comparison among *Xist* genotypes, we selected genes that were defined from the representative values in all wild type, $Xist^{CAGΔ5'}$ and $Xist^{ΔA}$ as informative genes. In initial analysis, we noticed that the X-linked *Las1l* and *Dynlt3* genes, which have not been reported to escape XCI, showed skewed expression toward the inactive X even in wild type. The RNA-seq alignment revealed that the aberrant data were caused by reads overlapping a single SNP for each gene, and close inspection of the variant call data showed that these SNPs were incorrectly called due to misalignment of pseudogenes. Therefore, those two genes were omitted from further allele-specific analysis.

To estimate gene expression level as the FPKM (fragments per kilobase of exon per million fragments mapped) in a non-allelic manner, the alignments including both unique and multiple hits were processed using Cufflinks v2.2.1 cuffdiff (Trapnell et al., 2010) with option '–library-type ff-firststrand –library-norm-method classic-fpkm –multi-read-correct'.

GO analysis

GO functional annotation was performed using DAVID (Huang et al., 2009). Only autosomal genes were considered as both differentially expressed and background genes. We selected 563 upregulated genes (≥3-fold change and q-value≤0.05) and 181 downregulated genes (≥2-fold change and q-value≤0.05) in $XX^{CAGΔ5'}$. GO terms with Benjamini-Hochberg-corrected $P<1\times10^{-4}$ are shown in Tables S1 and S2.

Derivation of ESCs

ESCs were derived from blastocysts recovered from JF1 females crossed with $X^{CAGΔ5'\text{-}2L}Y$; *Pgk2-cre* males and maintained according to Ying et al. (2008). ESCs were induced to differentiate as embryoid bodies for 3 days in conventional ES medium [DMEM (Sigma-Aldrich) containing 15% FBS, 1× non-essential amino acid (Nacalai), 0.1 mM beta-mercaptoethanol and 1× Penicillin/Streptomycin] without LIF, and resultant embryoid bodies were plated on coverslips to facilitate further differentiation. Immunofluorescence analysis was then carried out in essentially the same manner as described for whole-mount immunofluorescence above.

Acknowledgements

We are grateful to Hiroshi Kimura and Naohito Nozaki for the generous gift of antibodies against H3K27me3, H4K20me1 and H3K4me2.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: T.S.; Investigation: Y.S., K.N., H.S., C.O., T.S.; Resources: Y.H., T.S.; Writing - original draft: Y.S., T.S.; Writing - review & editing: T.S.; Supervision: T.S.; Project administration: T.S.; Funding acquisition: T.S.

Funding

This work was supported by Grants-in-Aid for Scientific Research on Innovative Areas (26113714 and 16H01320 to T.S.; 15K06942 and 15H01462 to K.N.; and 25116004 to C.O.) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT).

Data availability

RNA-seq data have been deposited in the Gene Expression Omnibus under accession number GSE93031.

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.149138.supplemental

References

Amakawa, Y., Sakata, Y., Hoki, Y., Arata, S., Shioda, S., Fukagawa, T., Sasaki, H. and Sado, T. (2015). A new Xist allele driven by a constitutively active promoter is dominated by Xist locus environment and exhibits the parent-of-origin effects. *Development* **142**, 4299-4308.

Chaumeil, J., Le Baccon, P., Wutz, A. and Heard, E. (2006). A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev.* **20**, 2223-2237.

Degner, J. F., Marioni, J. C., Pai, A. A., Pickrell, J. K., Nkadori, E., Gilad, Y. and Pritchard, J. K. (2009). Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. *Bioinformatics* **25**, 3207-3212.

Escamilla-Del-Arenal, M., da Rocha, S. T. and Heard, E. (2011). Evolutionary diversity and developmental regulation of X-chromosome inactivation. *Hum. Genet.* **130**, 307-327.

Gontan, C., Achame, E. M., Demmers, J., Barakat, T. S., Rentmeester, E., van IJcken, W., Grootegoed, J. A. and Gribnau, J. (2012). RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature* **485**, 386-390.

Hayashi-Takanaka, Y., Yamagata, K., Wakayama, T., Stasevich, T. J., Kainuma, T., Tsurimoto, T., Tachibana, M., Shinkai, Y., Kurumizaka, H. and Nozaki, N. (2011). Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. *Nucleic Acids Res.* **39**, 6475-6488.

Hayashi-Takanaka, Y., Maehara, K., Harada, A., Umehara, T., Yokoyama, S., Obuse, C., Ohkawa, Y., Nozaki, N. and Kimura, H. (2015). Distribution of histone H4 modifications as revealed by a panel of specific monoclonal antibodies. *Chromosome Res.* **23**, 753-766.

Heard, E. and Disteche, C. M. (2006). Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev.* **20**, 1848-1867.

Hoki, Y., Kimura, N., Kanbayashi, M., Amakawa, Y., Ohhata, T., Sasaki, H. and Sado, T. (2009). A proximal conserved repeat in the Xist gene is essential as a genomic element for X-inactivation in mouse. *Development* **136**, 139-146.

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44-57.

Inoue, K., Kohda, T., Sugimoto, M., Sado, T., Ogonuki, N., Matoba, S., Shiura, H., Ikeda, R., Mochida, K., Fujii, T. et al. (2010). Impeding Xist expression from the active X chromosome improves mouse somatic cell nuclear transfer. *Science* **330**, 496-499.

Jeppesen, P. and Turner, B. M. (1993). The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* **74**, 281-289.

Keane, T. M., Goodstadt, L., Danecek, P., White, M. A., Wong, K., Yalcin, B., Heger, A., Agam, A., Slater, G., Goodson, M. et al. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289-294.

Kido, T., Arata, S., Suzuki, R., Hosono, T., Nakanishi, Y., Miyazaki, J.-I., Saito, I., Kuroki, T. and Shioda, S. (2005). The testicular fatty acid binding protein PERF15 regulates the fate of germ cells in PERF15 transgenic mice. *Dev. Growth Differ.* **47**, 15-24.

Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T. and Wutz, A. (2004). A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. *PLoS Biol.* **2**, e171.

Li, H. (2012). Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly. *Bioinformatics* **28**, 1838-1844.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*, arXiv:1303.3997.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079.

Lyon, M. F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature* **190**, 372-373.

Marahrens, Y., Panning, B., Dausman, J., Strauss, W. and Jaenisch, R. (1997). Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* **11**, 156-166.
Monkhorst, K., Jonkers, I., Rentmeester, E., Grosveld, F. and Gribnau, J. (2008). X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. *Cell* **132**, 410-421.
Mugford, J. W., Yee, D. and Magnuson, T. (2012). Failure of extra-embryonic progenitor maintenance in the absence of dosage compensation. *Development* **139**, 2130-2138.
Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. and Roder, J. C. (1993). Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **90**, 8424-8428.
Nozawa, R.-S., Nagao, K., Igami, K.-T., Shibata, S., Shirai, N., Nozaki, N., Sado, T., Kimura, H. and Obuse, C. (2013). Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBiX1 pathway. *Nat. Struct. Mol. Biol.* **20**, 566-573.
Okamoto, I., Tan, S. and Takagi, N. (2000). X-chromosome inactivation in XX androgenetic mouse embryos surviving implantation. *Development* **127**, 4137-4145.
Patrat, C., Okamoto, I., Diabangouaya, P., Vialon, V., Le Baccon, P., Chow, J. and Heard, E. (2009). Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. *Proc. Natl. Acad. Sci. USA* **106**, 5198-5203.
Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131-137.
Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., de la Cruz, C. C., Otte, A. P., Panning, B. and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300**, 131-135.
Pullirsch, D., Hartel, R., Kishimoto, H., Leeb, M., Steiner, G. and Wutz, A. (2010). The Trithorax group protein Ash2l and Saf-A are recruited to the inactive X chromosome at the onset of stable X inactivation. *Development* **137**, 935-943.
Roberts, R. M., Ezashi, T. and Das, P. (2004). Trophoblast gene expression: transcription factors in the specification of early trophoblast. *Reprod. Biol. Endocrinol.* **2**, 47.
Rougeulle, C. and Avner, P. (2003). Controlling X-inactivation in mammals: what does the centre hold? *Semin. Cell Dev. Biol.* **14**, 331-340.
Sado, T., Wang, Z., Sasaki, H. and Li, E. (2001). Regulation of imprinted X-chromosome inactivation in mice by Tsix. *Development* **128**, 1275-1286.
Sheltzer, J. M., Torres, E. M., Dunham, M. J. and Amon, A. (2012). Transcriptional consequences of aneuploidy. *Proc. Natl. Acad. Sci. USA* **109**, 12644-12649.
Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T. B., Webster, Z., Peters, A. H. F. M., Jenuwein, T., Otte, A. P. and Brockdorff, N. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev. Cell* **4**, 481-495.
Takada, T., Ebata, T., Noguchi, H., Keane, T. M., Adams, D. J., Narita, T., Shin-I, T., Fujisawa, H., Toyoda, A., Abe, K. et al. (2013). The ancestor of extant Japanese fancy mice contributed to the mosaic genomes of classical inbred strains. *Genome Res.* **23**, 1329-1338.
Takagi, N. and Abe, K. (1990). Detrimental effects of two active X chromosomes on early mouse development. *Development* **109**, 189-201.
Takagi, N. and Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256**, 640-642.
Takagi, N., Sugawara, O. and Sasaki, M. (1982). Regional and temporal changes in the pattern of X-chromosome replication during the early post-implantation development of the female mouse. *Chromosoma* **85**, 275-286.
Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J. and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511-515.
Wutz, A., Rasmussen, T. P. and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* **30**, 167-174.
Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P. and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523.