nature structural & molecular biology

Xist-dependent imprinted X inactivation and the early developmental consequences of its failure

Maud Borensztein[1,5], Laurène Syx[1,2], Katia Ancelin[1], Patricia Diabangouaya[1], Christel Picard[1], Tao Liu[3], Jun-Bin Liang[3], Ivaylo Vassilev[1,2], Rafael Galupa[1], Nicolas Servant[2], Emmanuel Barillot[2], Azim Surani[4], Chong-Jian Chen[3] & Edith Heard[1]

The long noncoding RNA Xist is expressed from only the paternal X chromosome in mouse preimplantation female embryos and mediates transcriptional silencing of that chromosome. In females, absence of Xist leads to postimplantation lethality. Here, through single-cell RNA sequencing of early preimplantation mouse embryos, we found that the initiation of imprinted X-chromosome inactivation absolutely requires Xist. Lack of paternal Xist leads to genome-wide transcriptional misregulation in the early blastocyst and to failure to activate the extraembryonic pathway that is essential for postimplantation development. We also demonstrate that the expression dynamics of X-linked genes depends on the strain and parent of origin as well as on the location along the X chromosome, particularly at the first 'entry' sites of Xist. This study demonstrates that dosage-compensation failure has an effect as early as the blastocyst stage and reveals genetic and epigenetic contributions to orchestrating transcriptional silencing of the X chromosome during early embryogenesis.

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

In mammals, differences in sex-chromosome constitution between males (XY) and females (XX) have led to the evolution of dosage-compensation strategies, including transcriptional silencing of one X chromosome in females[1]. In mice, X-chromosome inactivation (XCI) first initiates in the preimplantation embryo. The noncoding RNA Xist is expressed only from the paternal allele, thus leading to paternal X-chromosome (Xp) inactivation[2]. The Xp remains inactive in extraembryonic tissues but is reactivated in the inner cell mass, and this reactivation is followed by random XCI in the embryo proper[3,4]. In early mouse embryos, XCI has been shown to be very dynamic, and its requirements, both in *cis* at the level of the X-inactivation center (Xic) and in *trans*, have been debated[5]. Imprinted XCI has been proposed to initiate *de novo*[2,6] after the onset of zygotic genome activation (ZGA) and Xist expression. One study has proposed that Xp inactivation is initially Xist independent and that Xist may be required only for early maintenance of silencing[7], whereas another study has reported a lack of Xp gene silencing in the absence of *Xist*[8]. However, these studies were based on the analysis of just a few genes. Two recent single-cell transcriptomic studies have exploited interspecific crosses to investigate XCI in female preimplantation embryos[6] and differentiating embryonic stem cells (ESCs)[9]. These studies have revealed that imprinted XCI indeed initiates between the four- and eight-cell stages[6] and that progression of random XCI is correlated with differentiation[9]. However, the extent to which initiation of Xp-linked gene silencing is dependent on Xist RNA or is influenced by the strain or parent of origin (e.g., imprinted X-linked genes) has not been investigated.

In this study, we set out to explore the precise kinetics of paternal and maternal X-linked gene expression during preimplantation embryogenesis, by using interspecific crosses and single-cell RNA sequencing (scRNA-seq). We investigated the differences in the dynamics of imprinted XCI that were attributable to genetic background and/or to parental origin. By investigating X-linked gene expression in female embryos derived from *Xist*-knockout (KO) males, we also demonstrated the absolute *Xist* dependence of early imprinted XCI and observed the genome-wide transcriptional consequences induced by a lack of dosage compensation. Overall, this study provides insight into the transcriptional and allelic dynamics of XCI as well as the nature of the requirement for dosage compensation during the first stages of mammalian development.

RESULTS

Allele-specific scRNA-seq during preimplantation development

To investigate the extent and requirements of gene silencing during imprinted XCI in early embryogenesis, we profiled the expression kinetics of genes on the Xp and Xm chromosomes, by using scRNA-seq[10]. F_1 embryos were derived from interspecific crosses of either wild type (WT) or *Xist* paternally deleted mutant ($Xist^{pat\Delta}$) origin, between the two-cell and blastocyst (approximately 60- to 64-cell) stages. Reciprocal crosses between highly polymorphic *Mus musculus castaneus* (Cast/EiJ) and *Mus musculus domesticus* (C57BL6/J) strains, herein referred to as Cast and B6, respectively, were used (**Fig. 1a**), and a minimum of five embryos and six single cells per stage for

[1]Institut Curie, PSL Research University, CNRS UMR3215, INSERM U934, Paris, France. [2]Institut Curie, PSL Research University, Mines Paris Tech, Bioinformatics and Computational Systems Biology of Cancer, INSERM U900, F-75005, Paris, France. [3]Annoroad Gene Technology Co., Ltd, Beijing, China. [4]Wellcome Trust Cancer Research UK Gurdon Institute, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK. [5]Present address: Wellcome Trust Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, UK. Correspondence should be addressed to E.H. (edith.heard@curie.fr).

Received 23 October 2016; accepted 20 December 2016; published online 30 January 2017; doi:10.1038/nsmb.3365

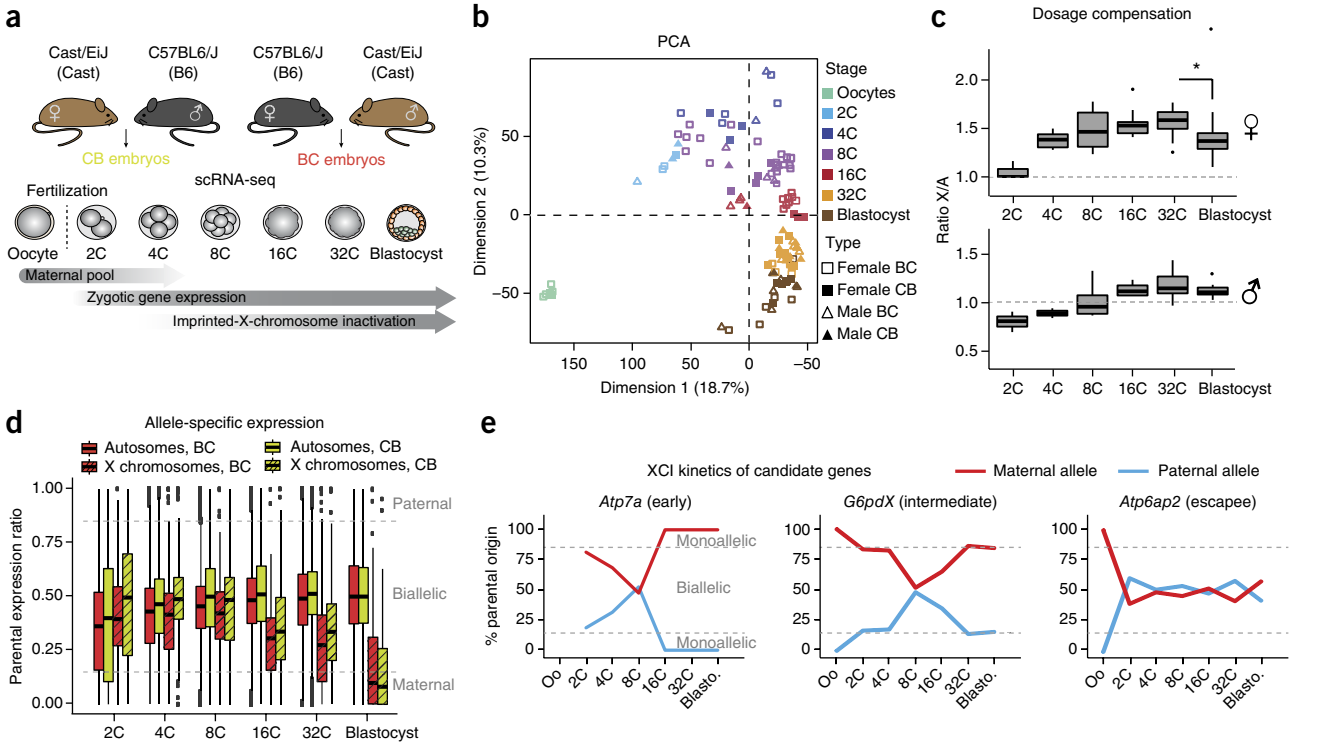


Figure 1 scRNA-seq of early hybrid embryos and dosage-compensation mechanisms. (**a**) Schematic illustration of the single-cell experiment and the harvested stages during preimplantation mouse development. Time windows showing the persistence of the maternal mRNA pool, activation of zygotic gene expression and Xp inactivation are indicated. C, cell stage. (**b**) PCA of single oocytes and preimplantation blastomeres (two-cell to blastocyst stages), on the basis of scRNA-seq data. Different stages are designated by different colors. Numbers of cells analyzed per stage and further details of scRNA-seq samples are in **Supplementary Data Set 1**. (**c**) Differences in ratios of X-chromosome expression levels to autosomal expression levels, from two-cell to blastocyst stages. $*P < 0.005$ by Dunn's test (KW). Box plots represent medians (center lines) with lower and upper quartiles (box limits). Whiskers represent 1.5× the interquartile range. Outliers are represented as dots. Numbers of cells analyzed per stage are given in **Supplementary Data Set 1**. (**d**) Allele-specific expression ratios for genes on autosomes (plain red, BC; yellow, CB) and on X chromosomes (dashed red, BC; yellow, CB) in female single blastomeres (two-cell to blastocyst stages) from BC and CB crosses. Allele-specific proportion represents the number of reads mapped to the paternal genome divided by the total number of paternal and maternal reads mapped for each gene. Box plots are as in **c**. Numbers of cells analyzed per stage are given in **Supplementary Data Set 1**. (**e**) Examples of scRNA expression dynamics of three X-linked genes classified as early inactivated, intermediate inactivated or escapee (as used in ref. 14) (additional data in **Supplementary Fig. 2**). Mean percentages of parental-origin transcripts are represented between oocytes (oo) and blastocysts (blasto.). Source data for **b**–**e** are available online.

WT embryos from B6 female × Cast male (BC) or Cast female × B6 male (CB) crosses (**Supplementary Data Set 1**). Of 24,499 referenced mouse genes, 15,581 were expressed in at least one developmental stage, including 580 X-linked genes.

We first assessed the extent to which transcriptomes of single cells were associated by stage, sex or cross, by performing principal component analysis (PCA) and hierarchical clustering (**Fig. 1b** and **Supplementary Fig. 1**). The primary source of variability among all cells was developmental stage, as expected on the basis of previous studies[6], thus validating the quality of our data. Single-cell transcriptomes clustered to a lesser extent by cross (BC and CB) and then by sex (XX and XY) (**Supplementary Fig. 1**), and the differences between the sexes reached a minimum by the 32-cell and blastocyst stages, presumably because of dosage compensation.

Timing of dosage compensation and imprinted XCI

To assess the precise timing of dosage compensation in male and female embryos, we examined autosomal and X-linked transcripts at each stage in both sexes. According to Ohno's law[11] average X-linked gene expression should be equivalent to the expression of autosomal genes. Furthermore, equal expression of X-linked genes between females and males is expected through XCI. We compared X/autosome (X/A) expression ratios in single blastomeres of each sex (**Fig. 1c**). The expected X/A ratios were 1 in females and 0.5 in males in the absence of any dosage compensation (i.e., no X overexpression compared with autosomes, and no XCI). We found that the X/A ratios were significantly higher than expected as early as the four-cell stage ($P < 9 \times 10^{-4}$ for males and females after the four-cell stage, by two-sided t test) and continued to rise until the 32-cell stage, thus suggesting that there is a progressive increase in expression of the X chromosome compared with autosomes at the same time as, or soon after ZGA. In females, the X/A ratio rose to 1.58 by the 32-cell stage and then significantly dropped to 1.37 by the early blastocyst stage ($P = 1.96 \times 10^{-2}$ between 32-cell and blastocyst stages, by Kruskal–Wallis (KW) test), presumably because of XCI by this stage (described below). This finding suggests that X/A ratios in female blastocysts progressively reach 1, although even at the early blastocyst stage, the ratios were still slightly higher than those in males ($P = 2.03 \times 10^{-3}$ by KW), in agreement with previously published data[12,13].

We next investigated allele-specific X-linked gene expression and the timing of XCI in BC and CB female embryos. At the two-cell stage, ZGA and massive degradation of the maternal pool of mRNAs occur. Transcripts were maternally biased genome wide, as expected given the residual maternal pool (**Fig. 1d**). At subsequent stages,

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

autosomal transcripts reached parity for both parental genomes, with a parental ratio in blastocysts of 0.5 in both crosses, whereas X-linked transcripts displayed maternal skewing even at the 16-cell stage. By the blastocyst stage, global transcription of the Xp was significantly decreased in both crosses ($P < 2.2 \times 10^{-16}$ by KW), thus indicating that XCI was fairly complete, as previously reported[6,8,14]. We compared the kinetics of Xp silencing for 13 X-linked genes previously analyzed by fluorescence *in situ* hybridization targeting nascent RNA (RNA-FISH)[14] and found that most (12/13) genes showed very similar patterns (**Fig. 1e** and **Supplementary Fig. 2**), thus giving us confidence that our scRNA-seq data, bioinformatics pipeline and expression thresholds were valid. The one gene (out of the 13) for which slightly earlier Xp silencing was found by scRNA-seq, as compared with results from previous reports, was *Atrx*. By using RNA-FISH with a gene-specific probe, we confirmed that *Atrx* was inactivated on the Xp in most cells by the morula stage (**Supplementary Fig. 3a**). We also confirmed its previously reported Xp reactivation in the blastocyst[14] (**Supplementary Fig. 2**).

Strain-specific XCI and escape

We established an *in vivo* chromosome-wide map of X-linked gene activity between the four-cell and blastocyst stages. Of the 580 X-linked genes expressed in our scRNA-seq, we focused on the 164 (BC cross) and 134 (CB cross) most highly expressed and informative genes (i.e., those with reads per retrotranscribed length per million mapped reads (RPRT) >4 and expressed in at least 25% of the cells of each stage and cross, with a minimum of two cells), for which we were able to establish allelic expression profiles with confidence (125 common genes between BC and CB crosses in **Supplementary Fig. 3b** and **Fig. 2**; details of allelic expression thresholds in Online Methods). A striking switch from biallelic (gray, light pink or light blue) expression at the four-cell stage to monoallelic maternal (red) expression at the blastocyst stage was observed for most X-linked genes. Several genes underwent only partial or no XCI (termed escapees) and will be discussed later. As expected, Xist expression was exclusively of paternal origin throughout (**Fig. 2** and **Supplementary Figs. 3b** and **4**). Another gene showing only paternal expression was *Fthl17f*, part of the ferritin, heavy polypeptide-like 17 family (also known as *Gm5635*), which has previously been reported to be exclusively paternally expressed and imprinted[15]. By the blastocyst stage, *Fthl17f* expression was no longer detectable, presumably because of XCI (**Fig. 2** and **Supplementary Figs. 3b** and **4**).

We categorized genes into different groups with respect to their timing of XCI for each cross (early, ≤16-cell stage; intermediate, ≤32-cell stage; late, blastocyst stage; **Fig. 3a**, **Supplementary Table 1** and Online Methods). Even by the eight-cell stage, XCI was complete for a few genes (e.g., *Rnf12* (official symbol *Rlim*) and *Pnma5*; **Fig. 2** and **Supplementary Fig. 3b**). By the blastocyst stage, Xp reached a very similar state of inactivation in both BC and CB crosses (83.5% and 84.3% of the 164 and 134 X-linked informative and expressed genes, respectively, were either silenced or maternally biased at the blastocyst stage; **Figs. 2** and **3a**,**b**). However, when comparing gene expression in embryos derived from BC and CB crosses (125 common genes), we observed marked differences between crosses, with just 71.2% (89 of 125 genes) of X-linked genes falling into the same or a similar category between BC and CB crosses (e.g., early and mid or late and biased). The degree of inconsistency in silencing kinetics between crosses was evaluated if more than one developmental stage separated the same gene between BC and CB crosses (**Supplementary Table 1** and classification details in Online Methods). Several genes also showed strain-specific escape (**Fig. 3b**). Some of these genes have previously

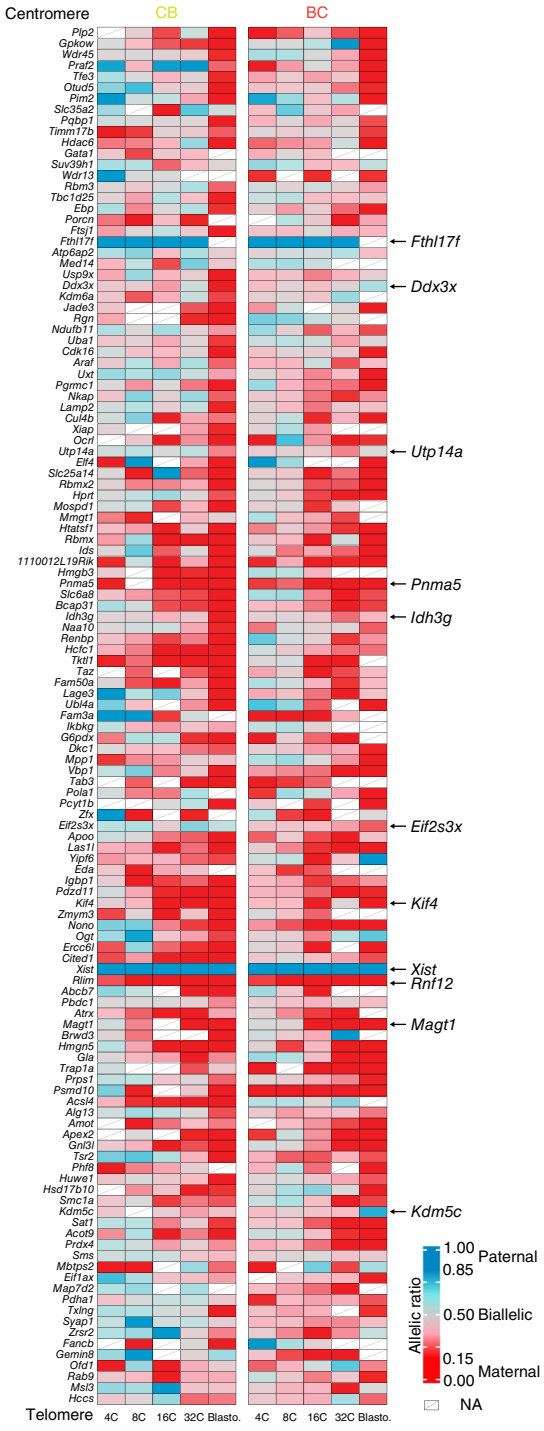


Figure 2 Kinetics of silencing of X-linked genes over the entire X chromosome during imprinted XCI in different strains. The mean allele-specific expression ratios per embryonic stage for each informative and expressed X-linked gene in female embryos at four-cell to blastocyst stages are represented as heat maps, with strictly maternal expression (ratio ≤0.15) in red and strictly paternal expression (ratio ≥0.85) in blue. Color gradients are used in between these two values, as shown in the key. Genes are ordered on the basis of genomic position (top, centromere; bottom, telomere). Data from CB (left) and BC (right) female embryos are shown (thresholds in Online Methods), and arrows highlight examples of early silenced or escapee genes. $n = 125$ informative X-linked genes in common for CB and BC crosses. NA, data not available (below threshold).

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

been described[16] or reported to escape in a tissue-specific fashion at later stages of development or in somatic tissues (e.g., *Ddx3x* and *Idh3g*)[16–18]. However, several genes remained biallelically expressed even at the blastocyst stage and tended to show escape independently of strain (**Fig. 3b** and **Supplementary Table 2**). Many of these genes also have been reported to show escape in somatic tissues[19]

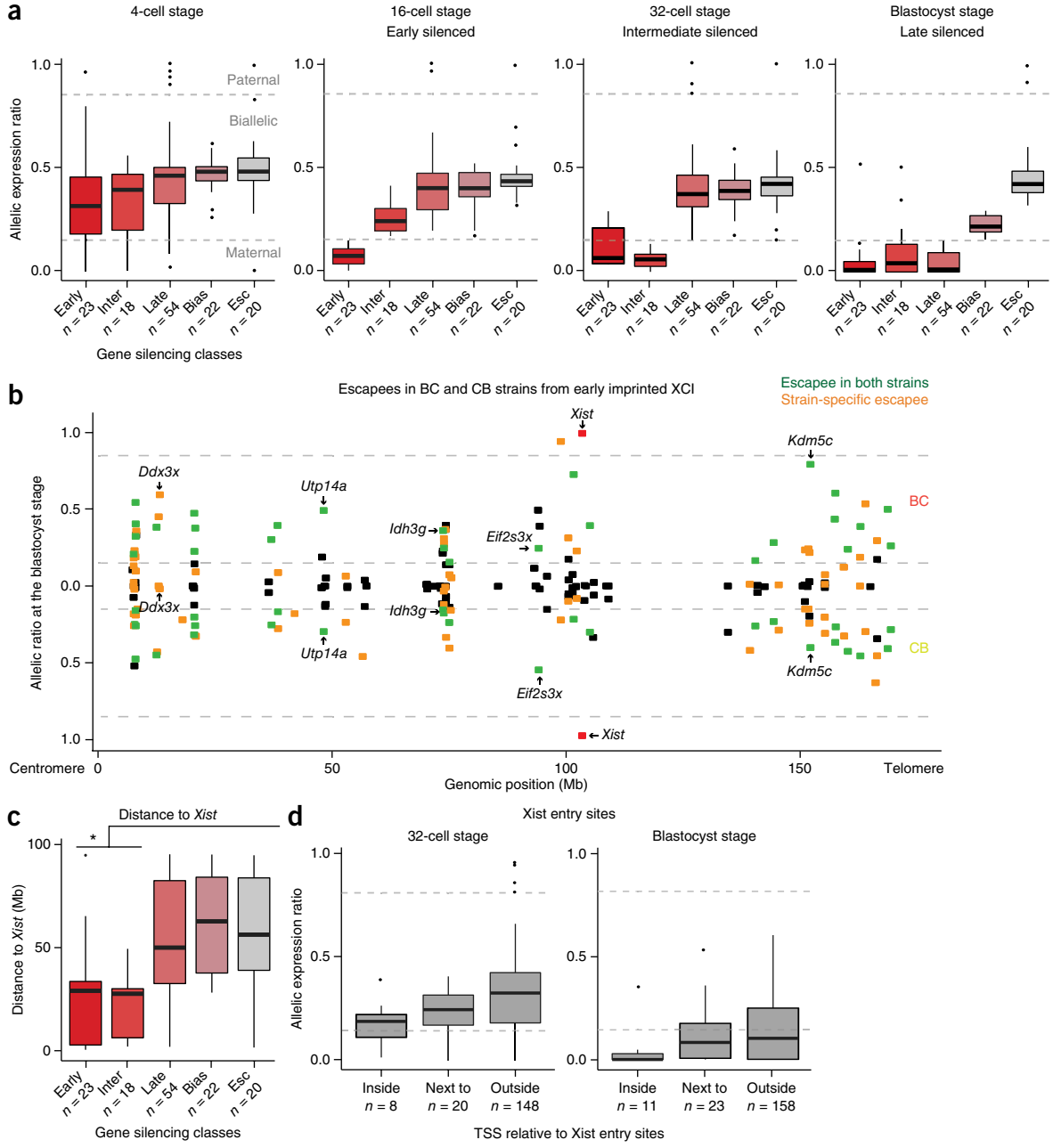


Figure 3 Different genes show different kinetics of silencing associated with chromosomal position and Xist entry site localization. (**a**) X-linked genes are clustered on the basis of their silencing kinetics as early (silenced at the 16-cell stage or earlier), intermediate (inter; silenced at the 32-cell stage), late (silenced at the blastocyst stage), biased (bias; maternally biased) and escapee (esc; not silenced). The allelic ratio of each gene represents the number of reads mapped on the paternal genome divided by the total number of reads mapped, and is shown for 4-cell, 16-cell, 32-cell and blastocyst stages from single female blastomeres. Further information is provided in **Supplementary Table 1** and Online Methods. $n = 137$ X-linked genes (89 with consistent silencing kinetics between BC and CB crosses and 48 BC or CB specific). Box plots are as in **Figure 1c**. Numbers of cells analyzed per stage are given in **Supplementary Data Set 1**. (**b**) Parental expression ratios of X-linked genes in female blastocysts in BC and CB strains. Each dot represents a single gene. The upper and lower sections represent data from BC and CB embryos, respectively. *Xist* is represented by a red dot. Green and orange dots represent genes that escape from early XCI in both strains or in a strain-specific manner, respectively. Further information on escapees can be found in **Supplementary Table 2**. $n = 125$ common X-linked genes. (**c**) Box plot representing the distribution of the genomic distances to the *Xist* locus for the different clusters of genes. The transcription start site (TSS) of each gene was used to measure the distance to *Xist*. $^*P < 0.05$ by Dunn's test. (**d**) Allelic expression of X-linked genes classified on the basis of their relative position to Xist entry sites (as identified during XCI induction in ESCs[20]): inside (TSS located in an Xist entry site), next to (TSS located less than 100 kb from an entry site) and outside (over 100 kb from an entry site). $^*P < 0.05$ by Dunn's test. Consistent or strain-specific genes were used. Box plots are as in **Figure 1c**. Source data are available online.

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

(e.g., *Eif2s3x*, *Kdm5c* and *Utp14a*). Finally, some genes with biallelic ratios (represented as black dots in **Fig. 3b**) corresponded to genes that previously underwent Xp silencing before the blastocyst stage but then became reexpressed, as previously described for *Atrx*[14].

Xist entry sites and early silenced genes

We next assessed whether the gene-silencing kinetics was correlated with the genomic position along the X chromosome. We first focused on the 71.2% ($n = 89$ genes) of genes with correlated kinetics between crosses and the strain-specific genes ($n = 48$). According to our allele-specific-expression heat map, although early and intermediate-silenced genes tended to lie closer than late-silenced genes to the Xic (**Fig. 3c**), gene silencing did not appear to occur as a simple linear gradient from the Xic, with some escapees close to the Xic (**Fig. 2**). Instead, we noted that several regions across the X chromosome contained early silenced genes (e.g., *Pnma5*, *Kif4* and *Magt1*), from which silencing appeared to 'spread' locally (**Fig. 2**). A recent study in ESCs has shown that Xist RNA initially binds to specific genomic regions (Xist entry sites) along the X chromosome, in a manner dependent on 3D proximity to the *Xist* locus[20]. This binding has been hypothesized by Engreitz *et al.* to silence genes locally and to then spread along the rest of the X chromosome[20]. We found that X-linked genes located within the predicted Xist entry regions (8 and 11 genes, respectively, in 32-cell embryos and blastocysts) or close to these regions (20 and 23, respectively, in 32-cell embryos and blastocysts) showed the earliest silencing and the strongest maternal imbalance ($P = 0.02$ and $P = 0.03$, respectively, by KW, in 32-cell embryos and blastocysts; **Fig. 3d**). Thus, the Xist-RNA entry sites, as defined in ESCs[20] may potentially correspond to XCI-initiation sites *in vivo* during imprinted XCI.

Fully *Xist*-dependent imprinted XCI

The above findings suggested that Xist RNA plays an early role in triggering gene silencing during imprinted XCI. This finding contrasts with results from a previous report suggesting that initiation of imprinted XCI is *Xist* independent[7]. Indeed, although $Xist^{patΔ}$ females die around embryonic day (E)10.5 and exhibit marked growth delays[21], mutant and WT females appeared morphologically indistinguishable during preimplantation development (data not shown). To evaluate whether XCI could be established, even in the absence of Xist expression, as previously reported[7], we examined X-linked gene expression profiles in single cells of preimplantation female embryos carrying a paternal *Xist* deletion ($Xist^{patΔ}$)[21,22]. Xist is normally expressed exclusively from the Xp in preimplantation embryos[2] (**Fig. 2**). Transcriptomes of single blastomeres from hybrid F_1 embryos (Cast females crossed with B6 males carrying a maternal *Xist* deletion ($Xist^{matΔ}$)) were compared with CB WT embryos from the eight-cell stage (when XCI normally initiates for some genes) to the blastocyst stage. We found similar X/A expression ratios between mutant and control female embryos up to the 32-cell stage (**Fig. 4a**). However, at the blastocyst stage, the X/A ratios remained much higher in mutants than in WT embryos, whereas this ratio normally decreases because of XCI ($P = 1.77 \times 10^{-4}$ by KW). This result indicated that Xp silencing was not initiated in $Xist^{patΔ}$ female blastocysts. Bioinformatics analysis on the *Xist*-mutant single-cell transcriptomes was used to produce an allele-specific expression heat map (Online Methods), and, as expected given the absence of apparent dosage compensation in the mutants, we found that X-linked genes remained significantly biallelically expressed in $Xist^{patΔ}$ embryos (**Fig. 4b**). Only two genes (*Rgn* and *Tktl1*) out of 122 assessed (i.e., 1.6%) showed maternal monoallelic expression in $Xist^{patΔ}$ mutant blastocysts, compared with 84.3% in CB WT controls. One of these genes, *Tktl1,* has been hypothesized to be imprinted[23]. Moreover, *Fthl17f*, a well-known imprinted gene, was aberrantly expressed in $Xist^{patΔ}$ blastocysts, thus suggesting a lack of Xp silencing.

We therefore found no evidence of *Xist*-independent XCI (**Supplementary Fig. 5a**), even for X-linked genes previously proposed by Kalantry *et al.* to be silenced independently of *Xist*[7] (11 out of 14 genes assayed by Kalantry *et al.*, of which only *Rnf12*, *Abcb7* and *Atrx* were *Xist* dependent). Three of the genes assayed by Kalantry *et al.* (*Abcb7*, *Fmr1* and *Pgk1*) showed a slight maternal bias at the 16-cell or 32-cell stages in the $Xist^{patΔ}$ cells in our study (**Supplementary Fig. 5a**, left). However this result was probably due to variability in parental-origin expression, as also observed in CB controls (*Abcb7* and *Fmr1,* **Supplementary Figs. 3**, **4** and **5a**) rather than to Xp silencing. Instead, our data are in agreement with those from a study by Namekawa *et al.*[8], in which *Xist*-dependent Xp silencing has been proposed to occur on the basis of nascent RNA-FISH on two-cell-stage to blastocyst-stage embryos and allele-specific RT–PCR on morulae, although that study was based on 14 genes, 10 of which were in common with those in ref. 5. The discrepancies between these previous studies are probably due to technical differences. The scRNA-seq analysis in the present study provides chromosome-wide evidence of *Xist*-dependent gene silencing during preimplantation embryogenesis and corroborates recent findings of *Xist*-dependent X-linked gene dosage[13].

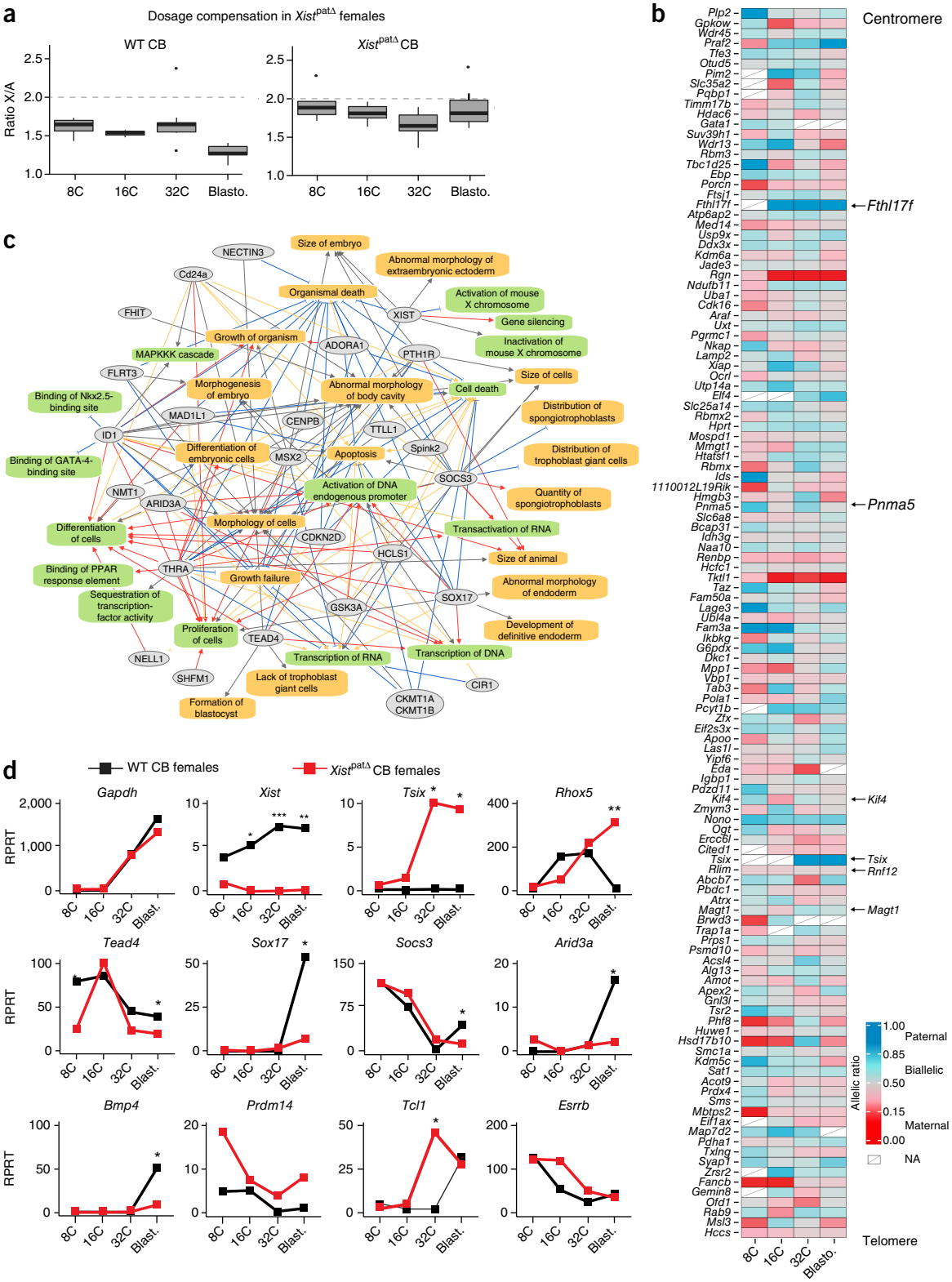
Improper gene expression in *Xist*-mutant embryos

The transcriptome of $Xist^{patΔ}$ embryos provided us with a unique opportunity to explore the molecular defects that occur in the absence of paternal XCI. A genome-wide analysis of differentially expressed (DE) genes was performed in WT and $Xist^{patΔ}$ embryos (**Supplementary Data Set 2**). Expression profiles of single blastomeres of controls and mutants still clustered according to developmental

Figure 4 Paternal knockout of *Xist* impairs XCI, dosage compensation and differentiation pathways. (**a**) Differences in ratios of X-chromosome expression levels to autosomal expression levels, from eight-cell to blastocyst stages in CB females (left) and $Xist^{patΔ}$ CB females (with a paternally inherited knockout allele) (right). Box plots are as in **Figure 1c**. Numbers of cells analyzed per stage are given in **Supplementary Data Set 1**. (**b**) Heat map representing allele-specific mean expression from eight-cell to blastocyst stages of X-linked genes (as in **Fig. 2**) in $Xist^{patΔ}$-mutant single cells. Red, strictly maternally expressed genes (allelic ratio ≤0.15); blue, strictly paternally expressed genes (allelic ratio ≥0.85). Color gradients are used in between, and genes are ordered on the basis of genomic position. *Tsix* was included in the heat map if it was expressed in at least two single cells per stage, even though it did not reach the expression threshold used (RPRT >4 and expressed in at least 25% of the cells of each stage and cross, with a minimum of two cells). $n = 122$ genes. *Rhox5* is not included, because of a lack of SNPs. (**c**) Major downregulated genes and pathways detected between CB WT and CB $Xist^{patΔ}$ females, as extracted from misregulated genes in the absence of paternal *Xist* (**Supplementary Data Set 2**) through Qiagen's Ingenuity Pathway Analysis (IPA) software (full list in **Supplementary Data Set 3**). Color code for arrows: red, leads to inhibition; blue, leads to activation; orange, findings consistent with state of downstream molecule; gray, effect not predicted. (**d**) Expression data of candidate genes from WT CB (black) and $Xist^{patΔ}$ CB (red) females, extracted from scRNA-seq data. Mean expression is represented in RPRT during early development (eight-cell to blastocyst stages). *Gapdh* is a control housekeeping gene. Box plots are as in **Figure 1c**. Numbers of cells analyzed per stage are given in **Supplementary Data Set 1**. $^*P < 0.05$ by KW. Source data for **a** and **d** are available online.

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

stage by PCA (data not shown). However, at the 8-cell and 32-cell stages, a markedly elevated number of DE autosomal genes (false discovery rate <0.05) was found in *Xist*^{patΔ} embryos compared with WT (**Supplementary Fig. 5b**). By the blastocyst stage, when paternal XCI is normally complete in WT females, 30% of the total upregulated genes in *Xist*^{patΔ} embryos were found to be X-linked, thus corroborating an XCI defect in the absence of *Xist*. DE genes included *Tsix* (encoding the antisense transcript to *Xist*) which is normally not expressed from the Xp at the 32- to 64-cell stages[24] (**Fig. 4b**,**d** and **Supplementary Fig. 5b**). The absence of *Xist* on the paternal X chromosome thus releases paternal *Tsix* repression in *cis* (without affecting the maternally imprinted *Xist*/*Tsix* alleles).



© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

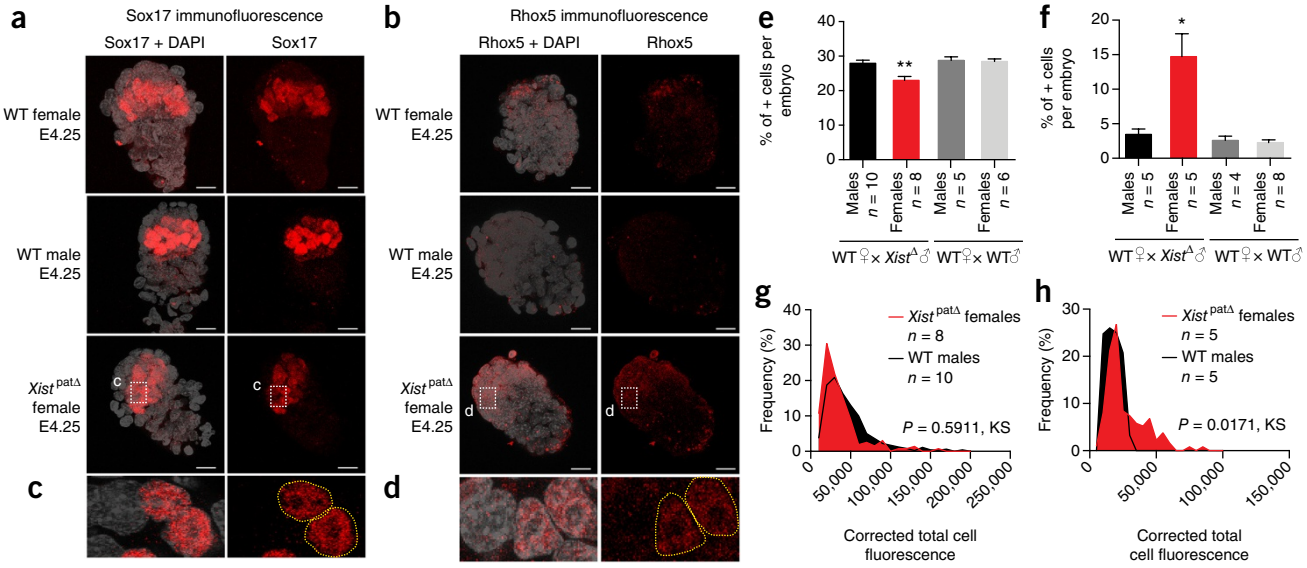


Figure 5 Abnormal Sox17 and Rhox5 patterns in *Xist*^{patΔ} female blastocysts. (**a**–**d**) Maximum intensity projection of WT and *Xist*^{patΔ} E4.25 blastocysts analyzed on the basis of immunofluorescence against Sox17 (**a**,**c**) or Rhox 5 (**b**,**d**). Red, Sox17 or Rhox5 staining; gray, DAPI nuclear staining. Scale bars, 20 µm. (**e**,**f**) Percentages of positive cells were assessed and are shown as the median + s.e.m. for Sox17 (**e**) and Rhox5 (**f**). Numbers of embryos analyzed are indicated under each genotype. $*P < 0.05$ and $**P < 0.001$ by KW. (**g**,**h**) Average distribution of positive single-cell fluorescence, represented by measurement of corrected total cell fluorescence (in arbitrary units) with ImageJ software for Sox17 (**g**) and Rhox5 (**h**) and tested by KS test. All cells with total cell fluorescence values under 10,000 and 5,000 for Sox17 and Rhox5, respectively, were considered negative.

We explored the degree to which transcriptomes were perturbed in the *Xist*-mutant embryos by using Ingenuity Pathway Analysis software. We found that many aberrantly downregulated genes in *Xist*^{patΔ} female blastocysts were associated with extraembryonic-tissue pathways, embryonic growth and cell viability (**Fig. 4c** and **Supplementary Data Set 3**). Key extraembryonic-development genes that were aberrantly downregulated included *Tead4* (ref. 25; trophectoderm (TE)), *Sox17* (ref. 26; primitive endoderm (PrE) and extraembryonic endoderm), *Bmp4* (ref. 27; TE and PrE), *Arid3a*[28] (TE specification) and *Socs3* (ref. 29; placental development) (**Fig. 4d**). To confirm the aberrant decrease in Sox17-positive cells in the PrE in *Xist*^{patΔ} females, we performed immunofluorescence analysis on late blastocysts (**Fig. 5a**,**c**,**e**,**g**). In *Xist*^{patΔ} females, compared with their male littermates, fewer cells expressed Sox17, and the intensity of fluorescence of Sox17 was slightly decreased (**Fig. 5g**), thus corroborating the decrease in mRNA expression that we observed by scRNA-seq.

Importantly, in addition to aberrant downregulation or repression of extraembryonic genes, we also found abnormal overexpression of several pluripotency genes including *Prdm14*, *Esrrb* and *Tcl1* in *Xist*^{patΔ} embryos. This result suggested an inappropriate activation or lack of repression of such factors in the absence of XCI (**Fig. 4d**) and was relevant to our recent findings showing that the presence of two active X chromosomes delays exit from pluripotency in ESCs by preventing downregulation of key genes, such as *Prdm14* and *Esrrb*[22]. Moreover, aberrant overexpression of *Prdm14*, *Esrrb* and *Tcl1* has been observed in *Xist*$^{-/-}$ female ESCs induced to differentiate[22]. Intriguingly, the most significantly upregulated gene (log fold change of 10) in *Xist*^{patΔ} female blastocysts was the imprinted *Rhox5* gene, also known as *Pem-1* (**Fig. 4d** and legend to **Fig. 4b**). *Rhox5* is a member of the reproductive X-linked Hox (*Rhox*) cluster and is expressed exclusively in the male germ line and in female (but not male) preimplantation embryos (Xp only)[30]. After implantation, its expression switches to the maternal allele and becomes restricted to extraembryonic tissues[30]. The human *RHOXF1* gene, which has been hypothesized to be related to mouse *Rhox5* (ref. 31) shows similar sex-specific and lineage-specific expression in human preimplantation embryos[32]. Importantly, previous *in vitro* studies have demonstrated that overexpression of *Rhox5* blocks differentiation of ESCs by preventing exit from pluripotency[33,34]. We validated *Rhox5* upregulation at the protein level by using immunofluorescence and found that *Xist*^{patΔ} female blastocysts, compared with WT blastocysts, showed significantly higher Rhox5 staining, particularly in the polar trophectoderm and inner-cell-mass region of the embryo (**Fig. 5b**,**d**,**f**,**g**). Quantification of Rhox5 immunofluorescence showed a significant increase in Rhox5 protein levels ($P = 0.0171$ by Kolmogorov–Smirnov (KS) test; **Fig. 5h**) and in the number of cells stained by anti-Rhox5 (**Fig. 5f**). This result correlated well with our scRNA-seq data.

We conclude that even by the early blastocyst stage, the lack of initiation of Xp inactivation in *Xist*^{patΔ} embryos leads to inappropriate downregulation of several key genes involved in extraembryonic development, overexpression of several pluripotency genes and massive overexpression of Rhox5, all of which may interfere with appropriate subsequent differentiation.

DISCUSSION

Here we demonstrated the key role of Xist RNA in initiating imprinted XCI. Although its role in triggering random XCI has previously been established, our study provides evidence that *Xist* is also essential for initiating early paternal XCI. Furthermore, our scRNA-seq analysis enabled us to identify the molecular defects in developmental pathways that emerge from this absence of dosage compensation and result in lethality several days later. Absence of *Xist* leads to an inappropriate downregulation of extraembryonic-development genes, a lack of downregulation of some pluripotency genes and massive overexpression of Rhox5. Together, some or all of these defects must ultimately result in compromised extraembryonic development and redirection toward what appears to be a more pluripotent state, or at

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

least a state from which further differentiation is perturbed. Previous studies[22] and a recent scRNA analysis of differentiating ESCs[9] have found that XCI progression is negatively correlated with pluripotency and positively correlated with differentiation. The gene expression perturbations that we observed in *Xist*-mutant embryos and their subsequent lethality are consistent with these previous observations and implicate some of the factors that are potentially involved.

It is also notable that the previously reported[33] aberrant induction of maternal Xist and Xm inactivation in extraembryonic tissues of blastocysts carrying a maternal *Tsix* deletion indicates that the presence of two active X chromosomes at the blastocyst stage can still be rescued in some females and suggests that the major defect associated with a lack of paternal XCI is initially in the extraembryonic lineage.

In this study, we also defined the influence of chromosomal location, as well as genetic background and parent of origin, on XCI kinetics. Our finding that Xist's predicted initial binding sites on the X chromosome correspond to the earliest regions silenced, between the 8- and 16-cell stages, with evidence for local spreading in *cis* at the 32-cell blastocyst stage, should enable exploration of the local features that underlie the spread of silencing along the X chromosome in an *in vivo* context. Finally, our study demonstrated the critical requirement for accurate X-chromosome gene dosage during early embryonic development and uncovered some of the key pathways and factors that are affected in the absence of XCI. Future dissection of these pathways and their relationship to X-linked gene dosage should provide a better understanding of the important roles of even small changes in RNA and protein levels, not only in development but also in disease.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank S. Bao and N. Grabole for experimental assistance in single-blastomere RNA sequencing and M. Guttman for sharing the coordinates of Xist entry sites. We are grateful to P. Gestraud and V. Sibut for help in statistical and IPA pathway analysis, respectively. We thank the pathogen-free barrier animal facility of the Institut Curie and J. Iranzo for help with the animals, and the Cell and Tissue Imaging Platform (PICT-IBiSA; member of France–Bioimaging) of the Genetics and Developmental Biology Department (UMR3215/U934) of the Institut Curie for assistance with light microscopy. We acknowledge E. Schulz, E. Nora, I. Okamoto and the members of E.H.'s laboratory for assistance, feedback and critical input. This work was funded by a fellowship from the Région Ile-de-France (DIM STEMPOLE) to M.B.; the Paris Alliance of Cancer Research Institutes (PACRI-ANR) to L.S.; an ERC Advanced Investigator award (ERC-2010-AdG, no. 250367), EU FP7 grants SYBOSS (EU 7th Framework G.A., no. 242129) and MODHEP (EU 7th Framework G.A., no. 259743), La Ligue, Fondation de France, Labex DEEP (ANR-11-LBX-0044), part of the IDEX Idex PSL (ANR-10-IDEX-0001-02 PSL), and ABS4NGS (ANR-11-BINF-0001) to E.H.; France Genomique National infrastructure (ANR-10-INBS-09) to E.H., N.S. and E.B; and CELLECTCHIP (ANR-14-CE10-0013) to E.H. and M.B.

AUTHOR CONTRIBUTIONS

M.B., A.S. and E.H. conceived the study. M.B. performed most of the experiments. K.A. performed the immunofluorescence, and P.D., C.P., M.B. and R.G. performed the RNA-FISH experiments. T.L., J.-B.L. and C.-J.C. performed single-cell transcriptome library preparation and sequencing. L.S., M.B., C.-J.C., I.V., N.S. and E.B. defined the data-processing and bioinformatics analysis. L.S. built the computational pipeline for scRNA-seq and analyzed the data with M.B.; M.B. and E.H. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Lyon, M.F. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* **190**, 372–373 (1961).
2. Okamoto, I. *et al.* Evidence for *de novo* imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* **438**, 369–373 (2005).
3. Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D. & Heard, E. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* **303**, 644–649 (2004).
4. Mak, W. *et al.* Reactivation of the paternal X chromosome in early mouse embryos. *Science* **303**, 666–669 (2004).
5. Galupa, R. & Heard, E. X-chromosome inactivation: new insights into *cis* and *trans* regulation. *Curr. Opin. Genet. Dev.* **31**, 57–66 (2015).
6. Deng, Q., Ramskold, D., Reinius, B. & Sandberg, R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* **343**, 193–196 (2014).
7. Kalantry, S., Purushothaman, S., Bowen, R.B., Starmer, J. & Magnuson, T. Evidence of Xist RNA-independent initiation of mouse imprinted X-chromosome inactivation. *Nature* **460**, 647–651 (2009).
8. Namekawa, S.H., Payer, B., Huynh, K.D., Jaenisch, R. & Lee, J.T. Two-step imprinted X inactivation: repeat versus genic silencing in the mouse. *Mol. Cell. Biol.* **30**, 3187–3205 (2010).
9. Chen, G. *et al.* Single-cell analyses of X chromosome inactivation dynamics and pluripotency during differentiation. *Genome Res.* **26**, 1342–1354 (2016).
10. Tang, F. *et al.* RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat. Protoc.* **5**, 516–535 (2010).
11. Brockdorff, N. & Turner, B.M. Dosage compensation in mammals. *Cold Spring Harb. Perspect. Biol.* **7**, a019406 (2015).
12. Nguyen, D.K. & Disteche, C.M. Dosage compensation of the active X chromosome in mammals. *Nat. Genet.* **38**, 47–53 (2006).
13. Wang, F. *et al.* Regulation of X-linked gene expression during early mouse development by *Rlim*. *eLife* **5**, e19127 (2016).
14. Patrat, C. *et al.* Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. *Proc. Natl. Acad. Sci. USA* **106**, 5198–5203 (2009).
15. Kobayashi, S. *et al.* The X-linked imprinted gene family Fthl17 shows predominantly female expression following the two-cell stage in mouse embryos. *Nucleic Acids Res.* **38**, 3672–3681 (2010).
16. Calabrese, J.M. *et al.* Site-specific silencing of regulatory elements as a mechanism of X inactivation. *Cell* **151**, 951–963 (2012).
17. Berletch, J.B. *et al.* Escape from X inactivation varies in mouse tissues. *PLoS Genet.* **11**, e1005079 (2015).
18. Marks, H. *et al.* Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. *Genome Biol.* **16**, 149 (2015).
19. Balaton, B.P. & Brown, C.J. Escape artists of the X chromosome. *Trends Genet.* **32**, 348–359 (2016).
20. Engreitz, J.M. *et al.* The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**, 1237973 (2013).
21. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. & Jaenisch, R. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* **11**, 156–166 (1997).
22. Schulz, E.G. *et al.* The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell* **14**, 203–216 (2014).
23. Nesbitt, A.M. Genomic imprinting of the X-linked gene transketolase-like 1 in mouse and human. *ProQuest Diss. Theses A&I.* 884624661 (2010).
24. Lee, J.T., Davidow, L.S. & Warshawsky, D. *Tsix*, a gene antisense to *Xist* at the X-inactivation centre. *Nat. Genet.* **21**, 400–404 (1999).
25. Nishioka, N. *et al.* Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mech. Dev.* **125**, 270–283 (2008).
26. Niakan, K.K. *et al.* Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. *Genes Dev.* **24**, 312–326 (2010).
27. Graham, S.J.L. *et al.* BMP signalling regulates the pre-implantation development of extra-embryonic cell lineages in the mouse embryo. *Nat. Commun.* **5**, 5667 (2014).
28. Rhee, C. *et al.* Arid3a is essential to execution of the first cell fate decision via direct embryonic and extraembryonic transcriptional regulation. *Genes Dev.* **28**, 2219–2232 (2014).
29. Takahashi, Y. *et al.* SOCS3: an essential regulator of LIF receptor signaling in trophoblast giant cell differentiation. *EMBO J.* **22**, 372–384 (2003).
30. Kobayashi, S. *et al.* Comparison of gene expression in male and female mouse blastocysts revealed imprinting of the X-linked gene, Rhox5/Pem, at preimplantation stages. *Curr. Biol.* **16**, 166–172 (2006).
31. Li, Q., O'Malley, M.E., Bartlett, D.L. & Guo, Z.S. Homeobox gene Rhox5 is regulated by epigenetic mechanisms in cancer and stem cells and promotes cancer growth. *Mol. Cancer* **10**, 63 (2011).
32. Petropoulos, S. *et al.* Single-cell RNA-seq reveals lineage and x chromosome dynamics in human preimplantation embryos. *Cell* **165**, 1012–1026 (2016).
33. Fan, Y., Melhem, M.F. & Chaillet, J.R. Forced expression of the homeobox-containing gene Pem blocks differentiation of embryonic stem cells. *Dev. Biol.* **210**, 481–496 (1999).
34. Cinelli, P. *et al.* Expression profiling in transgenic FVB/N embryonic stem cells overexpressing STAT3. *BMC Dev. Biol.* **8**, 57 (2008).

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

ONLINE METHODS

Mouse crosses and collection of embryos. All experimental design and procedures were in agreement with the guidelines from French legislation and institutional policies (French ethical committee of animal experimentation: Institut Curie #118 and agreement C75-05-17 for the animal facility). All BC and CB embryos were derived from natural mating between C57BL/6J (B6) females crossed with CAST/EiJ (Cast) males or by the reciprocal cross, respectively. The *Xist*^{patΔ} mutant embryos (*Xist*$^{+/-}$) were obtained by mating Cast females and *Xist*$^{-/Y}$ males (mixed background: B6D2F1(C57BL/6J;DBA/2J);129S1/SvlmJ;BALB/cJ). Embryos were harvested at 2-cell, 4-cell, 8-cell, 16-cell, 32-cell and blastocyst stages (approximately 60 to 64 cells), at E1.5, E2.0, E2.25, E2.75, E3.25 and E3.5, respectively. B6 and Cast pure oocytes were collected at E0.5 after mating of females with vasectomized males (**Fig. 1a**). The collected embryos were included in the analyses only if they showed a normal morphology and the correct number of blastomeres for their developmental stage. No statistical method was used to predetermine sample size. A minimum of five embryos from at least two different litters were used for each developmental stage. The sex of the embryos was characterized only after single-cell RNA amplification (performed in a blinded manner).

RNA fluorescence *in situ* hybridization. RNA-FISH on preimplantation embryos was performed as previously described[3] with the intron-spanning Fosmid probe WI1-2039P10 (BacPac Consortium at Children's Hospital Oakland Research Institute) for *Atrx* and the intron-spanning plasmid probe p510 for Xist. Images were acquired with a wide-field Deltavision core microscope (Applied Precision, GE Healthcare) with a 60× objective (1.42 oil PL APO N) and 0.2-µm *z* sections. Images were analyzed with ImageJ software (Fiji; NIH).

Immunofluorescence staining. Immunofluorescence was carried out essentially as described previously[35] with an additional step of blocking in 3% FCS before the primary-antibody incubation. Immunofluorescence analyses of embryos either from mutant or control male progeny were always performed in parallel and in suspension. The following antibodies were used: goat anti-mouse Pem-1 (Rhox5) (Santa Cruz, sc-21650; 1:50) and goat anti-human Sox17 (R&D Systems, AF1924; 1:100). Antibody validation is provided on the manufacturers' websites. Images were acquired with an Inverted laser scanning confocal microscope with spectral detection (LSM700, Zeiss) equipped with a 260-nm laser (RappOpto), with a 60× objective, and 2-µm *z* sections. Maximum projections and total corrected fluorescence measurements (integrated density – (area of selected cell × mean fluorescence of background readings) were performed in **Figure 5g,h** with ImageJ software, through previously described methodology[36]. The total corrected cellular fluorescence (TCCF) was calculated as integrated density – (area of selected cell × mean fluorescence of background readings).

Single-cell dissociation from preimplantation mouse embryos. Oocytes and embryos were collected by flushing oviducts (E0.5 to E2.75) or uteri (E3.25 and E3.5) with M2 medium (Sigma). The zona pellucida was removed with acid Tyrode's solution (Sigma), and embryos were washed twice with M2 medium (Sigma). To isolate individual cells, we then incubated embryos in Ca^{2+}- and Mg^{2+}-free M2 medium for 5 to 20 min, depending on the embryonic stage. For the blastocyst stage, incubation with Ca^{2+}- and Mg^{2+}-free M2 medium was replaced with a 5-min incubation in TrypLE (Invitrogen). After incubation, each blastomere was mechanically dissociated by mouth pipetting with a thin glass capillary. Single cells were then washed three times in PBS/acetylated BSA (Sigma) before being manually picked into PCR tubes with a minimum amount of liquid. We either directly prepared the cDNA amplifications or kept the single cells at –80 °C for future preparation.

Single-cell RNA amplification. Poly(A)$^+$ mRNA extracted from each single cell was reverse transcribed from the 3′ UTRs and amplified according to a previously described protocol[10]. Care was taken to process only embryos and single blastomeres of the highest quality, on the basis of morphology, number of cells and amplification yield. A total of 72 BC and 110 CB (including 113 WT and 69 *Xist*^{patΔ} mutant) blastomeres were processed and passed quality controls.

Quality and sex determination. After cDNA amplification and before size selection and library preparation, the quality of cDNAs from each of the samples was validated on the basis of the expression levels of three housekeeping genes: *Gapdh*, *Actb* and *Hprt*. Primers used for real-time PCR were as follows: Gapdh_F, ccccaacactgagcatctcc; Gapdh_R, attatgggggtctgggatgg; ActB_F, aagtgacgttgacatccg; ActB_R, gatccacatctgctggaagg; Hprt_F, cctgttggccatctgcctagt; and Hprt_R, gggacgcagcaactgacatt. Care was taken to process only single cells with consistent amplification rates compared with those of the three housekeeping genes in the same developmental stage.

The sex of each embryo was assessed through expression-level analysis of Xist (female-specific transcript) and *Eif2s3y* (male-specific transcript) by real-time PCR. Primers used were: Eif2s3y_F, aattgccaggttattttcattttc; Eif2s3y_R, agttcagtggtgcacagcaa; Xist_F, ggttctctctccagaagctaggaa; and Xist_R, tggtagatggcattgtgtattatatgg.

Single-cell libraries and deep sequencing. Single-cell libraries were prepared from the 182 samples that passed QC according to the manufacturer's protocol (Illumina). Sequencing to produce single-end 50-bp reads was then performed on an Illumina HiSeq 2500 instrument (**Supplementary Data Set 1**).

Quality control and filtering of raw data. Quality control was applied to raw data, as previously described[35]. Sequencing reads characterized by at least one of the following criteria were discarded from the analysis:

1. More than 50% low-quality bases (Phred score <5)
2. More than 5% N bases
3. More than 80% A/T bases
4. At least 30% (15 bases) of continuous A and/or T

SNP calling and allele-specific origin of the transcripts. *SNP collection and strain-specific genome construction.* The VCF file (mgp.v5.merged.snps_all.dbSNP142.vcf) reporting all SNP sites from 36 mouse strains, based on mm10, was downloaded from the Sanger database. With SNPsplit tool (v0.3.0)[37], these SNPs were filtered on the basis of their quality values (FI values) and were used to reconstruct the Cast genome from the mm10 genome assembly.

Allele-specific alignments of RNA-seq reads. To study the allele-specific gene expression, reads were processed with a pipeline adapted from ref. 38. Single-end reads were first aligned to the mouse mm10 and Cast genomes with TopHat2 software (v2.1.0)[39]. Only random best alignments with fewer than two mismatches were reported for downstream analyses. The resulting mapping files for both parental genomes were then merged for each sample, with the following rules:

1. If a read mapped at the same genomic position on the two genomes and had the same number of mismatches, it was considered to be a common read.
2. If a read aligned with fewer mismatches on one genome, the best alignment was retained, and the read was considered to be a specific read for the corresponding strain.
3. If a read aligned with the same number of mismatches on both genomes but at different genomic positions, it was discarded.

Allelic imbalance in gene expression and gene classification. SNPs between B6 and Cast were extracted from the VCF file used to reconstruct the Cast genome. After removal of common exonic SNPs between the *Xist* and *Tsix* genes, 20,220,776 SNPs were retained.

The SAMtools mpileup utility (v1.1)[40] was then used to extract base-pair information at each genomic position from the merged alignment file. At each SNP position, the numbers of paternal and maternal alleles were counted. The threshold used to call a gene informative was five reads mapped per single SNP, with a minimum of eight reads mapped on SNPs per gene, to minimize disparity with low-polymorphic genes. The allele-specific origin of the transcripts (or allelic ratio) was measured as the total number of reads mapped on the paternal genome divided by the total number of paternal and maternal reads for each gene: allelic ratio = paternal reads/(paternal + maternal) reads.

Genes were thus classified into two categories:

1. Monoallelically expressed genes: allelic-ratio value ≤0.15 or ≥0.85
2. Biallelically expressed genes: allelic-ratio value >0.15 or <0.85

Estimation of gene expression levels. Given that our RNA reverse transcription allowed sequencing only up to an average of 3 kb from the 3′ UTR, half of the

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

doi:10.1038/nsmb.3365

expressed genes were only partially covered (less than 50% of the gene size in average). To estimate transcript abundance, read counts were thus normalized on the basis of the amplification size of each transcript (RPRT) rather than on the basis of the size of each gene (RPKM).

Filtering of biased SNPs. Because we observed a bias for some polymorphisms when we should not have (e.g., in oocytes with only maternal expression or male cells with only maternal expression for the X chromosome), we decided to remove biased SNPs from our study. Oocytes and male cells were used to address this issue. Therefore, SNPs covered by at least five reads and having an allelic ratio greater than 0.3 (biallelic or paternally expressed) in at least two of these samples were discarded. In total, 275 SNPs were filtered out, including 40 sites located on the X chromosome.

Generation of *Xist*^{patΔ}-mutant embryos involved the use of an *Xist*$^{Δ/Y}$ sire of mixed background (mixed background: B6D2F1(C57BL/6J;DBA/2J);129S1/SvlmJ;BALB/cJ.). We therefore had to apply another SNP filtration to the KO samples to remove all B6 polymorphisms that might have been lost on the X chromosome because of the mixed background of the *Xist*$^{Δ/Y}$ sire. To this end, all existing SNPs between B6 and DBA/2J, 129S1/SvlmJ and BALB/cJ on the X chromosome were removed from our SNP database (34,397 SNPs, representing 5.5% of X-chromosome SNPs between B6 and Cast).

Principal component analysis, hierarchical clustering and differentially expressed genes. Count tables of gene expression were generated with the refSeq annotation and HTSeq software (v0.6.1)[41]. Only genes with an RPRT value >1 in at least 25% of the single cells of at least one developmental stage (with a minimum of two cells) were kept for the downstream analysis. The TMM method from the edgeR R package (v3.14.0)[42] was first used to normalize the raw count data. PCA and hierarchical clustering were then used to determine how single cells were clustered to the others, on the basis of their gene expression profiles, depending on their stage, sex and cross. PCA on normalized data was performed with the FactoMineR R package (v1.33). Hierarchical clustering analysis was based on Spearman correlation distance and the Ward method, with the hclust function implemented in the gplots R package (v3.0.1). The limma R package (v3.28.4)[43] was applied to identify the differentially expressed genes between the eight-cell and blastocyst stages in control and *Xist*^{patΔ}-mutant females. With the Benjamini–Hochberg correction, genes with an adjusted *P* value lower than $\alpha = 0.05$ were called as differentially expressed.

Functional enrichment analysis. Downregulated genes in *Xist*^{patΔ}-mutant female blastocysts compared with CB female blastocysts were analyzed with Qiagen's Ingenuity Pathway Analysis (IPA, http://www.qiagen.com/ingenuity). The Functions and Diseases module was used to extract the most significantly deregulated pathways and their associated genes.

Heat-map generation for X-chromosome allelic gene expression. For BC and CB heat maps, data from informative genes were analyzed at each developmental stage only if the gene was expressed (RPRT >4) in at least 25% of single blastomeres (with a minimum of two cells) at this particular stage and cross (**Figs. 2** and **4b** and **Supplementary Fig. 3**). To follow the kinetics of expression, we decided to focus only on genes expressed in at least three different stages between the four-cell to blastocyst stages. The mean of the allelic ratio of each gene was represented for the different stages. The same gene candidate list was used to produce the *Xist*^{patΔ} heat maps (**Fig. 4b**). A value was given only if the gene reached the threshold of RPRT >4 in at least 25% of single cells (with a minimum of two cells) per stage and cross.

Definition of X-linked gene silencing/escape classes. We automatically assigned X-linked genes that become strictly maternal (allelic ratio ≤0.15) at the 16-cell stage or earlier to the 'early silenced' gene class, assigned those that become maternal at the 32-cell stage to the 'intermediate silenced' class (allelic ratio NA or >0.15 at the 16-cell stage and ≤0.15 at the 32-cell stage), and assigned those silenced only by the blastocyst stage to the 'late silenced' gene class (allelic ratio NA or >0.15 at the 16-cell and 32-cell stages and ≤0.15 at the blastocyst stage). At the blastocyst stage, X-linked genes showing a maternal bias of expression ($0.15 <$ allelic ratio ≤ 0.3) were categorized as maternally biased. A final group comprised genes that escaped imprinted Xp inactivation (allelic ratio >0.3 at the blastocyst stage) (**Fig. 3a**). Genes escaping XCI were separated into two classes: 'constitutive escapees' if they were classified as escapees in both CB and BC stages and 'strain-specific escapees' if they were escapees in only one cross (**Fig. 3b** and **Supplementary Table 2**).

Existence of consistency in silencing kinetics between crosses was evaluated if no more than one developmental stage separated the same gene between BC and CB crosses. If the consistent genes belonged to two different classes, a class for all (BC + CB) was assigned, owing to their parental ratio mean of (BC mean + CB mean) in **Figure 3a,d**.

Dosage compensation and X/A expression ratio. We measured the global X/A expression ratio in females (XX/AA ratio) and males (X/AA ratio) as the level of expression of X-linked genes divided by the global level of expression of the autosomal genes. Only genes with an expression value RPRT >4 were used for subsequent analysis (**Figs. 1d** and **4a**). Adjustment of the number of expressed genes between the X chromosome and autosomes has been reported to be critical in X/A expression ratio measurement[44]. We then added a bootstrapping step and randomly selected, for each sample, an autosomal gene set with the same number of expressed genes compared with the X chromosome to estimate the global X/A ratio. This step was repeated 1,000 times, and the X/A expression ratio was estimated as the median of the 1,000 values.

Statistics. Statistical significance was evaluated through Dunn's multiple comparison test with Benjamini–Hochberg correction and Kruskal–Wallis analysis of variance. *P* values are provided in the figure legends and/or main text. Kruskal–Wallis and *post hoc* tests were used to analyze nonparametric and unrelated samples.

Data availability. The Gene Expression Omnibus accession number for the data sets reported in this paper is GSE80810. Source data for **Figure 1** (**Fig. 1b–e**), **Figure 3** (**3a–d**) and **Figure 4** (**4a,d**) are available with the paper online. All other data are available from the corresponding author upon reasonable request.

35. Ancelin, K. *et al.* Maternal LSD1/KDM1A is an essential regulator of chromatin and transcription landscapes during zygotic genome activation. *eLife* **5**, e0885 (2016).
36. McCloy, R.A. *et al.* Partial inhibition of Cdk1 in G 2 phase overrides the SAC and decouples mitotic events. *Cell Cycle* **13**, 1400–1412 (2014).
37. Rozowsky, J. *et al.* AlleleSeq: analysis of allele-specific expression and binding in a network framework. *Mol. Syst. Biol.* **7**, 522 (2011).
38. Gendrel, A.V. *et al.* Developmental dynamics and disease potential of random monoallelic gene expression. *Dev. Cell* **28**, 366–380 (2014).
39. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
40. Krueger, F. & Andrews, S.R. SNPsplit: allele-specific splitting of alignments between genomes with known SNP genotypes. *F1000Res.* **5**, 1479 (2016).
41. Anders, S., Pyl, P.T. & Huber, W. HTSeq: a Python framework to work with high-throughput sequencing data. *Bioinformat. ics* **31**, 166–169 (2015).
42. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
43. Ritchie, M.E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
44. Kharchenko, P.V., Xi, R. & Park, P.J. Evidence for dosage compensation between the X chromosome and autosomes in mammals. *Nat. Genet.* **43**, 1167–1169, author reply 1171–1172 (2011).

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.