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

Maud Borensztein, Laurène Syx, Katia Ancelin, Patricia Diabangouaya, Christel Picard, Tao Liu, Jun-Bin Liang, Ivaylo Vassilev, Rafael Galupa, Nicolas Servant, et al.

Fully *Xist*-dependent initiation of imprinted X inactivation and the early developmental consequences of its failure.

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

[1] Institut Curie, PSL Research University, CNRS UMR3215, INSERM U934, 26 Rue d'Ulm, 75248 Paris Cedex 05, France

[2] Institut Curie, PSL Research University, Mines Paris Tech, Bioinformatics and Computational Systems Biology of Cancer, INSERM U900, F-75005, Paris, France

[3] Wellcome Trust Cancer Research UK Gurdon Institute, Department of Physiology, Development and Neuroscience, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, United Kingdom

[4] Annoroad Gene Technology Co., Ltd, Beijing, China

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 pre-implantation embryo. The non-coding *Xist* RNA is expressed only from the paternal allele leading to paternal X (Xp) inactivation[2]. The Xp remains inactive in extra-embryonic tissues, but is reactivated in the inner cell mass followed by random XCI in the embryo proper[3,4]. Although *Xist* is essential for random XCI, the precise requirements for imprinted XCI during pre-implantation development have been debated[5]. Here we report that the initiation of imprinted XCI absolutely requires *Xist* using single-cell RNA-sequencing (scRNAseq) of early mouse embryos. Lack of paternal *Xist* leads to genome-wide transcriptional misregulation in the early blastocyst, with failure to activate the extra-embryonic pathway that is essential for post-implantation development. We also demonstrate that the expression dynamics of X-linked genes depends both on strain and parent-of-origin, as well as on location along the X chromosome. This study demonstrates that dosage compensation failure has an impact as early as the blastocyst stage and reveals the genetic and epigenetic basis for orchestrating the transcriptional silencing of the X chromosome during early embryogenesis.

In early mouse embryos, XCI has been shown to be dynamic and its requirements, both in *cis* at the level of the X-inactivation center *(*Xic*)* and in *trans*, have been debated[6]. Imprinted XCI has been proposed to initiate *de novo*[2,9] following the onset of zygotic genome activation (ZGA) and *Xist* expression. One study proposed that Xp inactivation is initially *Xist*-independent and that *Xist* may only be required for early maintenance of silencing[5], while another reported a lack of Xp gene silencing in the absence of *Xist*[7]. These studies were all

based on the analysis of just a few genes, however. A recent transcriptomic study in female pre-implantation embryos derived from inter-specific Cast/B6 (CB) crosses[8] revealed that paternal XCI indeed initiates between the 4-8-cell stage. However, whether initiation of Xp-linked gene silencing relies on *Xist* RNA, or is influenced by strain- or parent-of-origin (*eg* imprinted X-linked genes) were not explored.

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, using scRNAseq. F1 embryos were derived from inter-specific crosses, of either wild-type (*wt*) or *Xist* paternally deleted mutant ($Xist^{pat\Delta}$) origin, between the 2-cell and blastocyst (approximately 60-64-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 (Figure 1a) and a minimum of 5 embryos, and 6 single cells per stage for BC and CB *wt* embryos (Supplementary Table 1). Of 24,499 referenced mouse genes, 15,581 were found 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 analyses (PCA) and hierarchical clustering (Figure 1b and Supplementary Figure 1). The primary source of variability between all cells was developmental stage, as expected based on previous studies[8], 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 Figure 1), with the differences between the sexes reaching a minimum by the 32-cell and blastocyst stages, presumably due to dosage compensation.

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[9] 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:Autosomes (X:A) expression ratios in single blastomeres of each sex (Figure 1c). Expected X:A ratios would be 1 in females and 0.5 in males in the absence of any dosage compensation (*ie* no X overexpression compared to autosomes, and no XCI). We found that the X:A ratios were significantly above the expected ratios as early as the 4-cell stage ($p<9*10^{-4}$ for males and females after 4-cell stage, t-test) and continued to rise until the 32-cell stage, suggesting that there is a progressive increase in expression of the X compared to autosomes at the same time as, or soon after ZGA. In females, the X:A ratio rose to 1.58, by the 32 cell stages and then significantly dropped to 1.37 by the early blastocyst stage ($p=1.96*10^{-2}$ between 32-cell and blastocyst, Kruskal-Wallis (KW) test), presumably due to XCI by this stage (see below). This suggests that X:A ratios in female blastocysts progressively reach 1, although even at the early blastocyst stage, they were still slightly higher compared to males ($p=2.03*10^{-3}$, KW), in agreement with previously published data[10].

We next investigated allele-specific X-linked gene expression and the timing of XCI in BC and CB female embryos. At the 2-cell stage, ZGA and massive degradation of the maternal pool of mRNAs occur. Here, transcripts are maternally biased genome-wide as expected given the residual maternal pool (Figure 1d). At subsequent stages, while autosomal transcripts reached parity for both parental genomes, with a parental ratio in blastocysts of 0.5 in both crosses, X-linked transcripts displayed maternal skewing even at the 16-cell stage. By the blastocyst stage global transcription of the Xp was significantly reduced in both crosses ($p<2.2*10^{-16}$, KW) indicating that XCI was fairly complete, as previously reported[7,8,11]. We compared the kinetics of Xp silencing for 13 X-linked genes previously analyzed by nascent RNA-FISH[11] and found that most (12/13) genes showed very similar patterns (Figure 1e and

Supplementary Figure 2), giving us confidence that our scRNAseq data, bioinformatics pipeline and expression thresholds were valid. The one gene (out of the 13) for which slightly earlier Xp silencing was found by scRNAseq compared to previous reports was *Atrx*. We confirmed that *Atrx* is inactivated on the Xp in most cells by the morula stage using RNA FISH with a gene-specific probe (Supplementary Figure 3a). We also confirmed its previously reported Xp reactivation in the blastocyst[11] (Supplementary Figure 2).

We established an *in vivo* chromosome-wide map of X-linked gene activity between the 4-cell and blastocyst stages. Of the 580 X-linked genes expressed in our scRNAseq, we focused on the 164 (BC cross) and 134 (CB cross) most highly expressed and informative genes (RPRT>4 and expressed in at least 25% of the cells of each stage and cross with a minimum of 2), for which we could establish allelic expression profiles with confidence (Supplementary Figure 3b and Figure 2 for the 125 common genes between BC and CB crosses, see Online Methods for allelic expression threshold details). A striking switch from biallelic (grey, pale pink or pale blue) expression at the 4-cell stage, to monoallelic, maternal (red) expression at the blastocyst stage can be observed for most X-linked genes. Several genes underwent only partial or no XCI (escapees) and will be discussed later. As expected, *Xist* expression was exclusively of paternal origin throughout (Figure 2 and Supplementary Figure 3b). 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[12]. By the blastocyst stage *Fthl17f* expression was no longer detectable, presumably due to XCI (Figure 2 and Supplementary Figure 3b).

We categorized genes into different groups with respect to their timing of XCI for each cross (*early ≤ 16-cell; intermediate ≤32-cell; late = blastocyst*; Figure 3a, SI Table 2 and Online Methods). Even by the 8-cell stage, XCI is complete for a few genes (*eg Rnf12*,

Pnma5, Figure 2 and Supplementary Figure 3b). By the blastocyst stage, Xp reached a very similar state of inactivation in both BC and CB crosses (respectively 83.5% and 84.3% of the 164 and 134 X-linked informative and expressed genes are either silenced or maternally biased at the blastocyst stage, Figures 2, 3a and 3b). However, when comparing gene expression in embryos derived from BC and CB crosses (125 common genes), marked differences were seen between crosses, with only 71.2% (89 of 125 genes) of X-linked genes falling into the same or a similar category between BC and CB crosses (*eg* early/mid or late/biased). The degree 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 (Supplementary Table 2 and Online Methods for classification details). Furthermore, several genes show strain-specific escape, some of which have previously been described[13] or reported to escape in a tissue-specific fashion at later stages of development or in somatic tissues (*eg Ddx3x, Idh3g)*[13–15]. Several genes remain biallelically expressed, even at the blastocyst stage and tend to show escape independent of strain (Figure 3b and Supplementary Table 3). Many of these also show escape in somatic tissues[16] (*eg Eif2s3x, Kdm5c*, *Utp14a*).

We then assessed whether gene-silencing kinetics was correlated with 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). Although early and intermediate silenced genes do tend to lie closer to the Xic compared to late silenced genes (Figure 3c), gene silencing does not appear to occur as a simple linear gradient from the Xic according to our allele-specific expression heatmap, with presence of escapees close to the Xic (Figure 2). Rather, we noted several regions across the X chromosome, containing early-silenced genes, (*eg Pnma5*, *Kif4, Magt1*) from which silencing appeared to "spread" locally (Figure 2). A recent study in ES cells showed that *Xist* RNA initially binds to specific genomic regions (*Xist* "entry" sites) along the X chromosome, dependent on 3D proximity to

the *Xist* locus[17]. This binding has been hypothesized to silence genes locally and to then spread along the rest of the X chromosome by Engreitz *et al*[17]. We found that X-linked genes lying within (8 and 11 genes respectively in 32-cell and blastocysts), or close to (20 and 23 respectively in 32-cell and blastocysts) these regions showed the earliest silencing and strongest maternal imbalance (p=0.02 and p=0.03, KW, respectively in 32-cell and blastocysts, Figure 3d). Thus, we show that *Xist* RNA "entry" sites as defined in ESCs[17] could correspond to XCI initiation sites *in vivo*, during imprinted XCI.

These findings suggest that *Xist* RNA plays a key early role in triggering gene silencing during imprinted XCI. This contrasts with a previous report suggesting that initiation of imprinted XCI could be is *Xist*-independent[5]. Indeed, although $Xist^{pat\Delta}$ females die around E10.5, with major growth delay[18], mutant and *wt* females appear morphologically indistinguishable during pre-implantation development (data not shown). To evaluate whether XCI is established in the absence of *Xist* expression, as previously reported[5], we examined X-linked gene expression profiles in single cells of pre-implantation female embryos carrying a paternal *Xist* deletion ($Xist^{pat\Delta}$)[18,19]. *Xist* is normally expressed exclusively from the Xp in pre-implantation embryos[2] (Figure 2). Transcriptomes of single blastomeres from hybrid F1 embryos (Cast females crossed with $Xist^{mat\Delta}$ B6 males) were compared to CB *wt* embryos between the 8-cell stage (when XCI normally initiates for some genes) to the blastocyst stage. We found similar X:A expression ratios between female mutants and controls up to the 32-cell stage (Figure 4a). However at the blastocyst stage, X:A ratios remained much higher in mutants compared to *wt* embryo*s* where this ratio normally decreases due to XCI ($p=1.77*10^{-4}$, KW). This demonstrates that Xp silencing is not initiated in $Xist^{pat\Delta}$ female blastocysts. Bioinformatics analysis on the *Xist*-mutant single cell transcriptomes was used to produce an allele-specific expression heatmap (see 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 (Figure 4b). Only 1.6% genes (2 out of 122 genes: *Rgn* and *Tktl1*) showed maternal monoallelic expression in *Xist*^{patΔ} mutant blastocysts, compared to 84.3% in CB controls. One of these, *Tktl1,* has been hypothesized to be imprinted[20]. Moreover, *Fthl17f*, a well-known imprinted gene was aberrantly expressed in *Xist*^{patΔ} blastocysts, suggesting a lack of Xp silencing.

We thus found no evidence for *Xist*-independent XCI (Supplementary Figure 5a), even for the X-linked genes previously proposed to be silenced independently of *Xist*[5] (11 out of 14 assayed, only *Rnf12, Abcb7* and *Atrx* were found to be Xist-dependent in the Kalantry *et al* study). 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 (left column, Supplementary Figure 5a). However this is probably due to variability in their parental-origin expression, also observed in CB controls (*Abcb7* and *Fmr1,* Supplementary Figures 3 and 5a) or low parental-origin information (*Pgk1*, data not shown) rather than to Xp silencing. Instead, our data is in agreement with the Namekawa et al study[7] where *Xist*-dependent Xp silencing was proposed to occur based on nascent RNA-FISH on 2-cell to blastocyst stage embryos, although their study was only based on 8 genes, 4 of which were in common with ref 5. The discrepancies between these previous studies were likely due to technical differences. The scRNAseq analysis we provide here represents chromosome-wide evidence for *Xist*-dependent gene silencing during pre-implantation embryogenesis.

The transcriptome of *Xist*^{patΔ} embryos also provides us with a unique opportunity to explore the molecular defects that occur in the absence of paternal XCI. A genome-wide differentially expressed (DE) gene analysis was performed in *wt* and *Xist*^{patΔ} embryos (Supplementary Table 4). Expression profiles of single blastomeres of controls and mutants were still found to cluster according to developmental stage by PCA (data not shown). However, at the 8-cell and 32-cell stages, a surprisingly elevated number of DE autosomal

genes (FDR<0.05) was found in *Xist*^{patΔ} embryos compared to *wt* (Supplementary Figure 5b). By the blastocyst stage, when paternal XCI is normally complete in *wt* females, 30% of the total up-regulated genes in *Xist*^{patΔ} embryos were found to be X-linked, corroborating an XCI defect in the absence of *Xist*. DE genes included *Tsix* (the antisense transcript to *Xist*) which is normally not expressed from the Xp at the 32-64 cell stage[21] (Figures 4b and d and Supplementary Figure 5b). The absence of *Xist* on the paternal X thus releases paternal *Tsix* repression in *cis* (without affecting the maternally imprinted *Xist/Tsix* alleles).

We explored the degree to which transcriptomes were perturbed in the *Xist* mutant embryos using Ingenuity Pathway Analysis software. We found that many aberrantly down-regulated genes in *Xist*^{patΔ} female blastocysts were associated with extra-embryonic tissue pathways, embryonic growth and cell viability (Figure 4c and Supplementary Table 5). Key extra-embryonic development genes that were aberrantly down-regulated included *Tead4* (trophectoderm)[22], *Sox17*[23] (primitive endoderm PrE and ExE), *Bmp4*[24] (trophectoderm TE and PrE), *Arid3a*[25] (TE specification) and *Socs3*[26] (placental development) (Figure 4d). To confirm the aberrant decrease of Sox17-positive cells in the PrE in *Xist*^{patΔ} females, we performed immunofluorescence on late blastocysts (Supplementary Figure 6a, c, e and g). In *Xist*^{patΔ} females, fewer cells express Sox17 compared to their male littermates and the intensity of fluorescence of Sox17 is slightly decreased (Supplementary Figure 6g), which corroborates the decrease in mRNA expression that we observe by scRNAseq.

Importantly, in addition to aberrant down regulation or repression of extra-embryonic genes, we also found abnormal overexpression of several pluripotency genes including *Prdm14, Esrrb* and *Tcl1* in *Xist*^{patΔ} embryos. This suggested an inappropriate activation or lack of repression of such factors in the absence of XCI (Figure 4d). This is relevant to our recent findings showing that the presence of two active X chromosomes delays exit from pluripotency in ESCs, by preventing down-regulation of key genes, such as *Prdm14* or

Esrrb[19]. Moreover, aberrant over-expression of *Prdm14, Esrrb* and *Tcl1* was observed in *Xist*$^{-/-}$ female ESCs induced to differentiate[19]. Intriguingly, the most significantly up-regulated gene (10 log fold change) in *Xist*^{patΔ} female blastocysts was the imprinted *Rhox5* gene, also known as *Pem-1*. *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) pre-implantation embryos (Xp only)[27]. Following implantation, its expression switches to the maternal allele and becomes restricted to extra-embryonic tissues[27]. The human *RHOXF1* gene that is hypothesized to be related to the murine *Rhox5*[28] shows similar sex-specific and lineage-specific expression in human pre-implantation embryos[29]. Importantly, previous *in vitro* studies demonstrated that over-expression of *Rhox5* can block differentiation of ESCs by preventing exit from pluripotency[30,31]. We validated *Rhox5* up-regulation at the protein level using immunofluorescence and found that *Xist*^{patΔ} female blastocysts show significantly higher Rhox5 staining, particularly in the polar trophectoderm and inner cell mass region of the embryo, compared to *wt* blastocysts (Supplementary Figure 6b, d, f and g). Quantification of Rhox5 immunofluorescence showed a significant increase in Rhox5 protein levels ($p=0.0171$, Kolmogorov-Smirnov KS test, Supplementary Figure 6h) and in the number of cells stained by Rhox5 antibody (Supplementary Figure 6f). This correlates well with our scRNAseq data.

We conclude that even by the early blastocyst stage, the lack of initiation of Xp inactivation in Xist^{patΔ} embryos leads to inappropriate down-regulation of several key genes involved in extra-embryonic development, overexpression of several pluripotency genes and massive overexpression of Rhox5. Together some or all of these defects must ultimately result in compromised extra-embryonic development and redirection towards what appears to be a more pluripotent state, or at least a state from which further differentiation is perturbed. Importantly, aberrant induction of maternal *Xist* and Xm inactivation in extra-

241 embryonic tissues of blastocysts carrying a maternal *Tsix* deletion demonstrates that the
242 presence of two active X chromosomes at the blastocyst stage can still be rescued in some
243 females, and suggests that the major defect associated with a lack of paternal XCI is initially
244 in the extra-embryonic lineage[32].

245 In conclusion, we have demonstrated here the key role that *Xist* RNA plays in
246 initiating imprinted XCI. Our scRNAseq has enabled us to identify the molecular defects in
247 developmental pathways that emerge from this absence of dosage compensation and result in
248 lethality a few days later. We also reveal the influence of chromosomal location, as well as
249 genetic background and parent-of-origin on XCI kinetics. Our finding that *Xist's* predicted
250 initial binding sites on the X chromosome correspond to the earliest regions silenced, between
251 the 8-16 cell stage, with evidence for local spreading in *cis* at the 32-blastocyst stage will
252 enable exploration of the local features that underlie the spread of silencing along the X
253 chromosome in an in vivo context. Our study demonstrates the critical requirement for
254 accurate X-chromosome gene dosage during early embryo development and uncovers some
255 of the key pathways and factors that are affected in the absence of XCI. Future dissection of
256 these pathways and their relationship to X-linked gene dosage should provide a better
257 understanding of the important role that small changes in RNA and protein levels can play,
258 not only development but also in disease.

259

References

1. Lyon, M. F. Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature* **190,** 372–3 (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,** 644–9 (2004).
4. Mak, W. *et al.* Reactivation of the paternal X chromosome in early mouse embryos. *Science* **303,** 666–9 (2004).
5. 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).
6. Galupa, R. & Heard, E. X-chromosome inactivation: New insights into cis and trans regulation. *Curr. Opin. Genet. Dev.* **31,** 57–66 (2015).
7. 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,** 3187–205 (2010).
8. Deng, Q., Ramskold, D., Reinius, B. & Sandberg, R. Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells. *Science (80-.).* **343,** 193–196 (2014).
9. Brockdorff, N. & Turner, B. M. Dosage compensation in Mammals. *Cold Spring Harb Perspect Biol* **7:a019406,** (2015).
10. Nguyen, D. K. & Disteche, C. M. Dosage compensation of the active X chromosome in mammals. *Nat. Genet.* **38,** 47–53 (2006).
11. Patrat, C. *et al.* Dynamic changes in paternal X-chromosome activity during imprinted

X-chromosome inactivation in mice. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 5198–203 (2009).

12. 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).

13. Calabrese, J. M. *et al.* Site-Specific Silencing of Regulatory Elements as a Mechanism of X Inactivation. *Cell* **151,** 951–963 (2012).

14. Berletch, J. B. *et al.* Escape from X Inactivation Varies in Mouse Tissues. *PLOS Genet.* **11,** e1005079 (2015).

15. Marks, H. *et al.* Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. *Genome Biol.* **16,** 149 (2015).

16. Balaton, B. P. & Brown, C. J. Escape Artists of the X Chromosome. *Trends Genet.* **32,** 348–359 (2016).

17. Engreitz, J. M. *et al.* The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341,** 1237973 (2013).

18. 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).

19. 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).

20. Nesbitt, A. M. Genomic imprinting of the X-linked gene transketolase-like 1 in mouse and human. *ProQuest Diss. Theses A&I.* **884624661,** (2010).

21. Lee, J. T., Davidow, L. S. & Warshawsky, D. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat. Genet.* **21,** 400–4 (1999).

22. Nishioka, N. *et al.* Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mech. Dev.* **125,** 270–283 (2008).

23. 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).

24. 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).

25. 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).

26. Takahashi, Y. *et al.* SOCS3: An essential regulator of LIF receptor signaling in trophoblast giant cell differentiation. *EMBO J.* **22,** 372–384 (2003).

27. 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).

28. 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).

29. Petropoulos, S. *et al.* Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **165,** 1012–1026 (2016).

30. 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,** 481–96. (1999).

31. Cinelli, P. *et al.* Expression profiling in transgenic FVB/N embryonic stem cells overexpressing STAT3. *BMC Dev. Biol.* **8,** 57 (2008).

335 32. Sado, T., Wang, Z., Sasaki, H. & Li, E. Regulation of imprinted X-chromosome
336 inactivation in mice by Tsix. *Development* **128,** 1275–1286 (2001).
337

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Bao and N. Grabole for experimental help in single blastomere RNA sequencing and M. Guttman for sharing the *Xist* "entry" site coordinates. We are grateful to P. Gestraud and V. Sibut respectively for the help in statistical and IPA pathway analysis. 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 Institut Curie for help with light microscopy. We acknowledge E. Schulz, E. Nora, I. Okamoto and the members of E.H. laboratory for help, feedbacks and critical input. This work was funded by a fellowship of Région Ile-de-France (DIM STEMPOLE) to M.B., the Paris Alliance of Cancer Research Institutes (PACRI-ANR) to LS and 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 and France Genomique National infrastructure (ANR-10-INBS-09) to EH, NS, EB.

Author Contributions

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

data with M.B. M.B. and E.H. wrote the paper.

Author Information ScRNAseq data are deposited in Gene Expression Omnibus under accession number GSE80810. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for material should be addressed to E.H. (edith.heard@curie.fr).

Figure Legends

Figure 1: Single cell RNA sequencing of early hybrid embryos and dosage compensation mechanisms.

(a) Schematic illustration of the single cell experiment and the harvested stages during pre-implantation mouse development. Time windows showing the persistence of maternal mRNA pool, activation of zygotic gene expression and Xp inactivation are indicated.

(b) Principal component analysis (PCA) of single oocytes and pre-implantation blastomeres (2C to blastocysts) based on scRNA data. BC and CB Females and BC and CB males are respectively represented as ☐, ■ and ◮, ▲. Different stages are designed by different colors. n= 6 to 30 cells per stage (details of each single cell is listed in Supplementary Table 1).

(c) Dosage compensation during preimplantation development. Level ratio of X chromosome expression is measured over autosomal one (XX:AA or X:AA) between 2-cell stage to blastocyst embryos, by Dunn's test (Kruskal-Wallis), $p<0.001$ to **.

(d) Allele-specific gene ratios from autosomes (plain red or yellow) and from X chromosomes (dash red or yellow) in female single blastomeres (2-cell to blastocyst) from BC and CB crosses. Allele-specific proportion represents the number of reads mapped on paternal genome divided by the total number of paternal and maternal reads mapped for each gene.

e) Example of expression dynamics of three X-linked genes with their classification as "early inactivated", "intermediate inactivated" or "escapee" as described in Patrat *et al*, 2009[11] (see also Supplementary Figure 2). Percentage of parental origin transcripts is represented between oocytes and blastocyst.

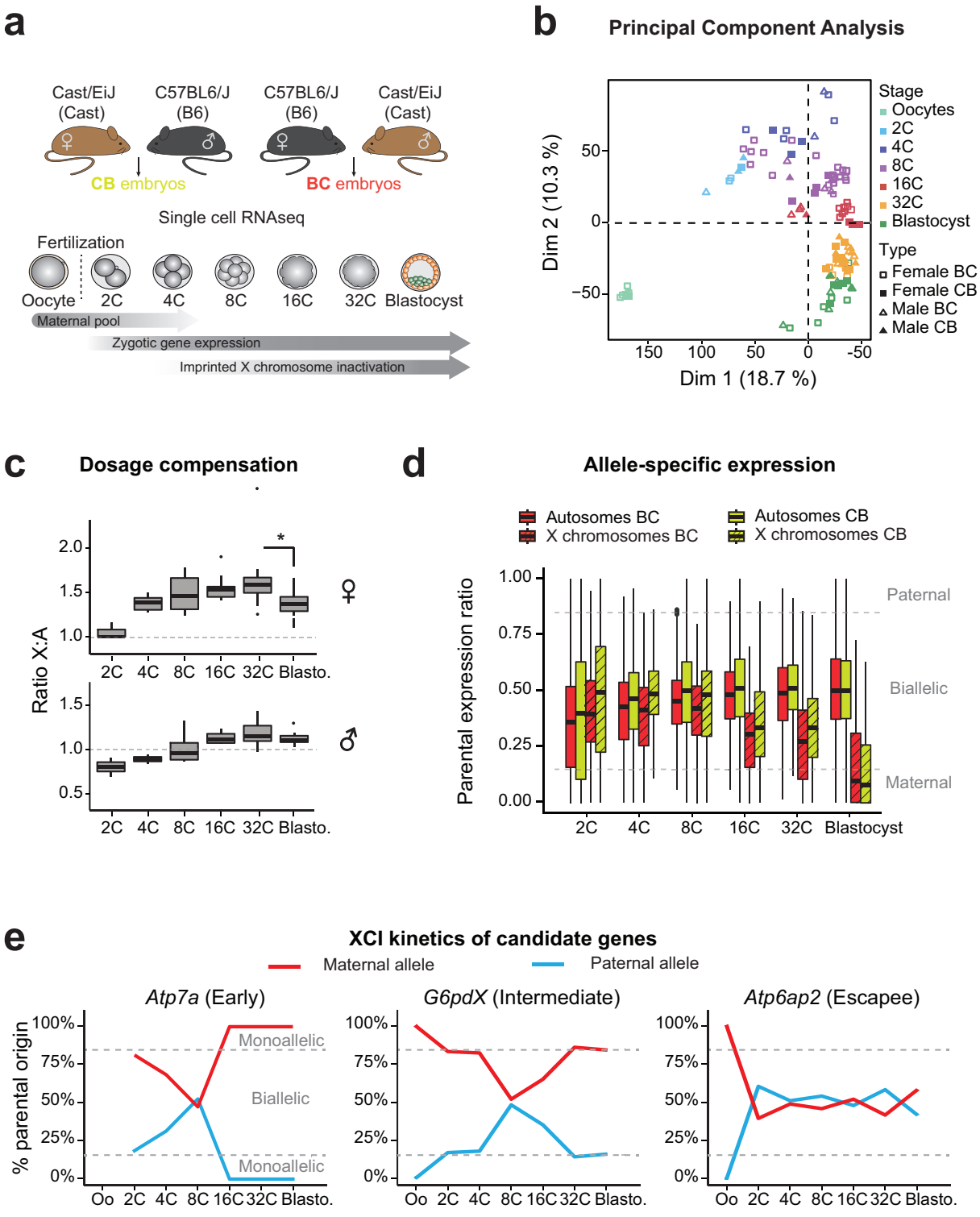


Figure 1- Borensztein *et al*

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 4-cell to blastocyst stage female embryos are represented as heatmaps, 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 by genomic position (centromere top, telomere bottom). Data from CB (left) and BC (right) female embryos are shown (for thresholds see Online Method) and arrows are highlighting example of early silenced and escapee genes. n= 125 informative X-linked genes in common for CB and BC crosses.

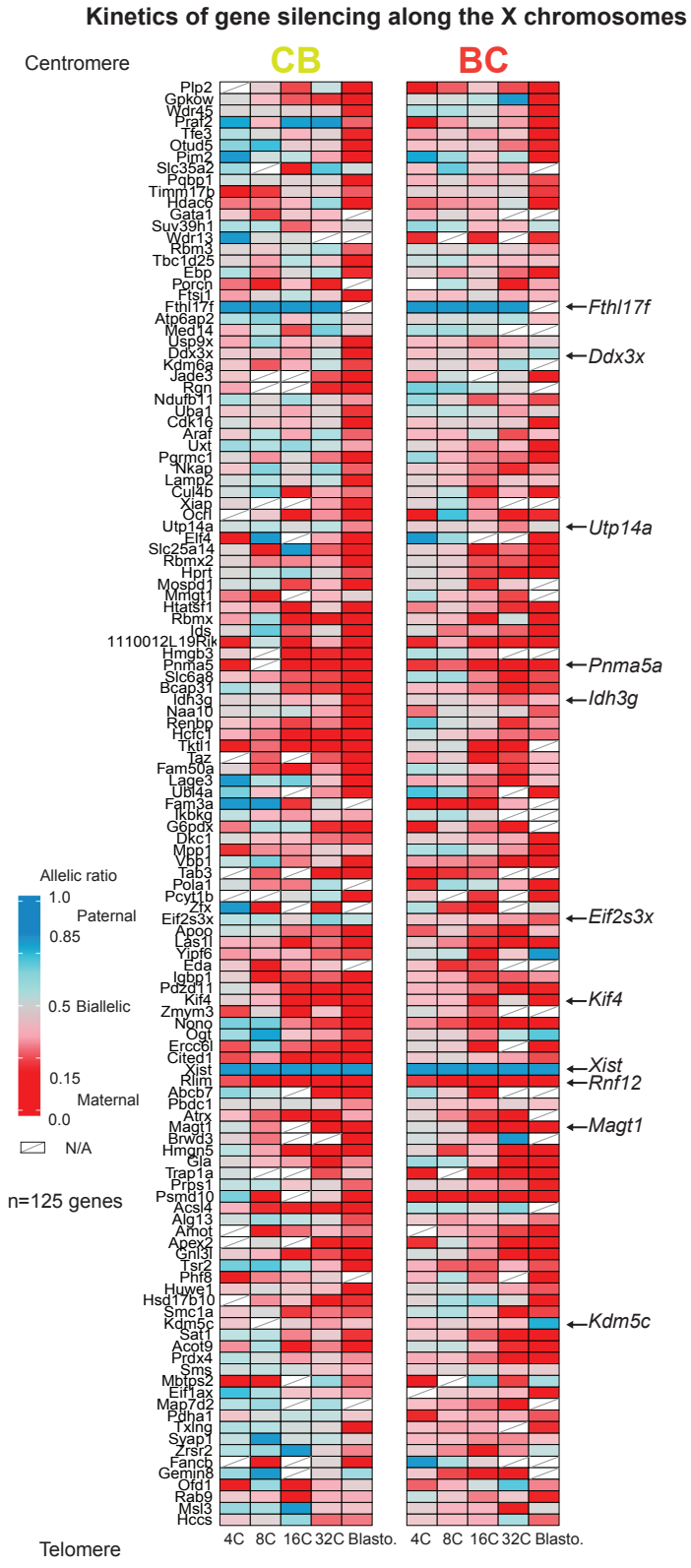


Figure 2- Borensztein *et al*

Figure 3: Strain of origin and chromosomal positions are associated with different kinetics of silencing.

(a) X-linked genes are clustered based on their silencing kinetics as "early" (0.15≤allelic ratio at 16-cell or earlier), "intermediate" (Inter: 0.15≤allelic ratio at 32-cell), "late" (0.15≤allelic ratio at blastocyst), "biased" (Bias: maternally biased in blastocyst, 0.15<allelic ratio ≤0.3) 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 paternal and maternal reads mapped and is represented at 4-cell, 16-cell, 32-cell and blastocyst stages from single female blastomeres. Further information is provided in Supplementary Table 2 and Online Methods. n= 137 X-linked genes (89 with consistent silencing kinetics between BC and CB crosses and 48 BC or CB-specific).

(b) Parental expression ratios of X-linked genes in female blastocysts in BC and CB strains. Each dot represents a single gene. The upper section represents data from BC embryos; lower section is the data from CB embryos (reverse cross). *Xist* is represented by a red dot. Green dots represent genes that escape from early XCI in both strains; orange dots represent genes that are strain-specific escapees. Some genes with biallelic ratios are represented by black dots and correspond to genes that previously underwent Xp silencing, prior to the blastocyst stage, but then became re-expressed. Further information on escapees is found in Supplementary Table 3. n= 125 common X-linked genes.

(c) Box plot representing the distribution of the genomic distances to *Xist* locus (in Mb) for the different clusters of genes. Transcription Start Site (TSS) of each gene has been used to measure the distance, by Dunn's test (Kruskal-Wallis); $p<0.05$ corresponds to *.

(d) Allelic expression of X-linked genes classified by their relative position to *Xist* "early" binding sites (as identified during XCI induction in ESCs[17]): inside (TSS located in a *Xist* "entry" site), next to (TSS located less than 100 kb to an "entry" site) and outside (over 100

kb of an “entry” site). Parental expression ratio has been plotted at 32cell-stage, when paternal XCI is initiated for most of the genes, by Dunn’s test (Kruskal-Wallis); $p<0.05$ corresponds to *. Consistent and strain-specific genes have been used.

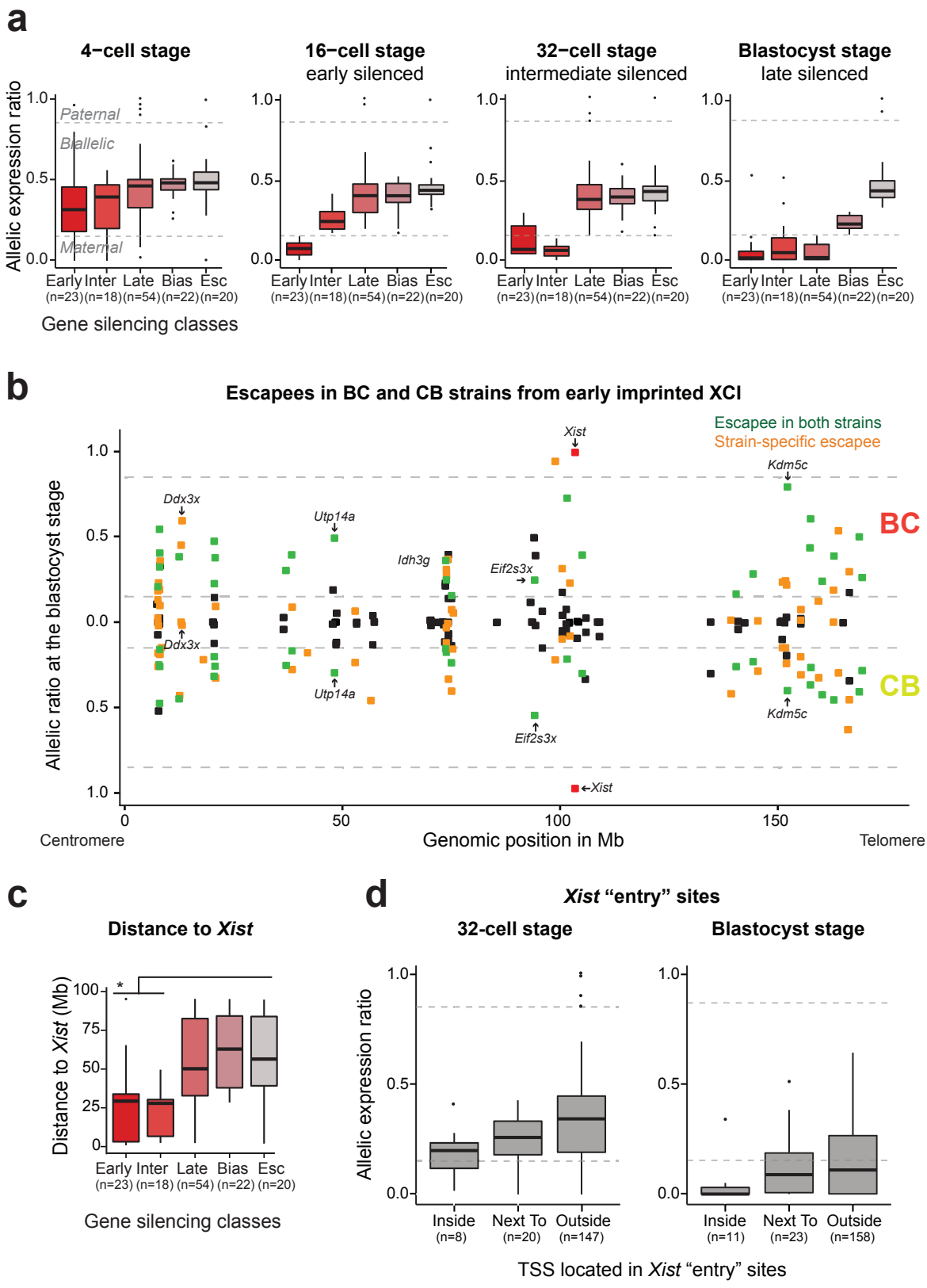


Figure 3- Borensztein *et al*

437

Figure 4: Paternal knockout of *Xist* impaired XCI, dosage compensation and differentiation pathways.

(a) Dosage compensation between X chromosomes and autosomes is measured by X:Autosome expression ratio between 8-cell to blastocyst in CB females (left panel) and $Xist^{pat\Delta}$ CB females (with a paternally inherited knock-out allele) (right panel).

(b) Heatmap representing allele-specific mean expression from 8-cell to blastocyst stage of X-linked genes (from Figure 2) in $Xist^{pat\Delta}$ mutant single cells. Strictly maternally expressed genes (allelic ratio ≤0.15) are represented in red and strictly paternally expressed genes (allelic ratio ≥0.85) in blue. Color gradients are used in between and genes have been ordered by genomic position. *Tsix* data have been included in the heatmap if expressed in at least 2 single cells per stage. = 122 genes.

(c) Major down-regulated genes and pathways detected between CB *wt* and CB $Xist^{pat\Delta}$ females extracted from Supplementary Table 4, using QIAGEN's Ingenuity Pathway Analysis (IPA) software (Supplementary Table 5). Pathways were visualized using Cytoscape software. Arrows color code: red: leads to inhibition; blue: leads to activation; orange: findings consistent with state of downstream molecule; grey: effect not predicted.

(d) Expression data of candidate genes from *wt* CB (black) and $Xist^{pat\Delta}$ CB (red) females, extracted from scRNAseq. Expression is represented in Reads Per Retro-Transcribed length per million mapped reads (RPRT) during early development (8-cell to blastocyst stages). *Gapdh* gene is a control housekeeping gene. n= 4 to 30 cells per stage and genotype. By Kruskal-Wallis test; $p<0.05$ corresponds to *.

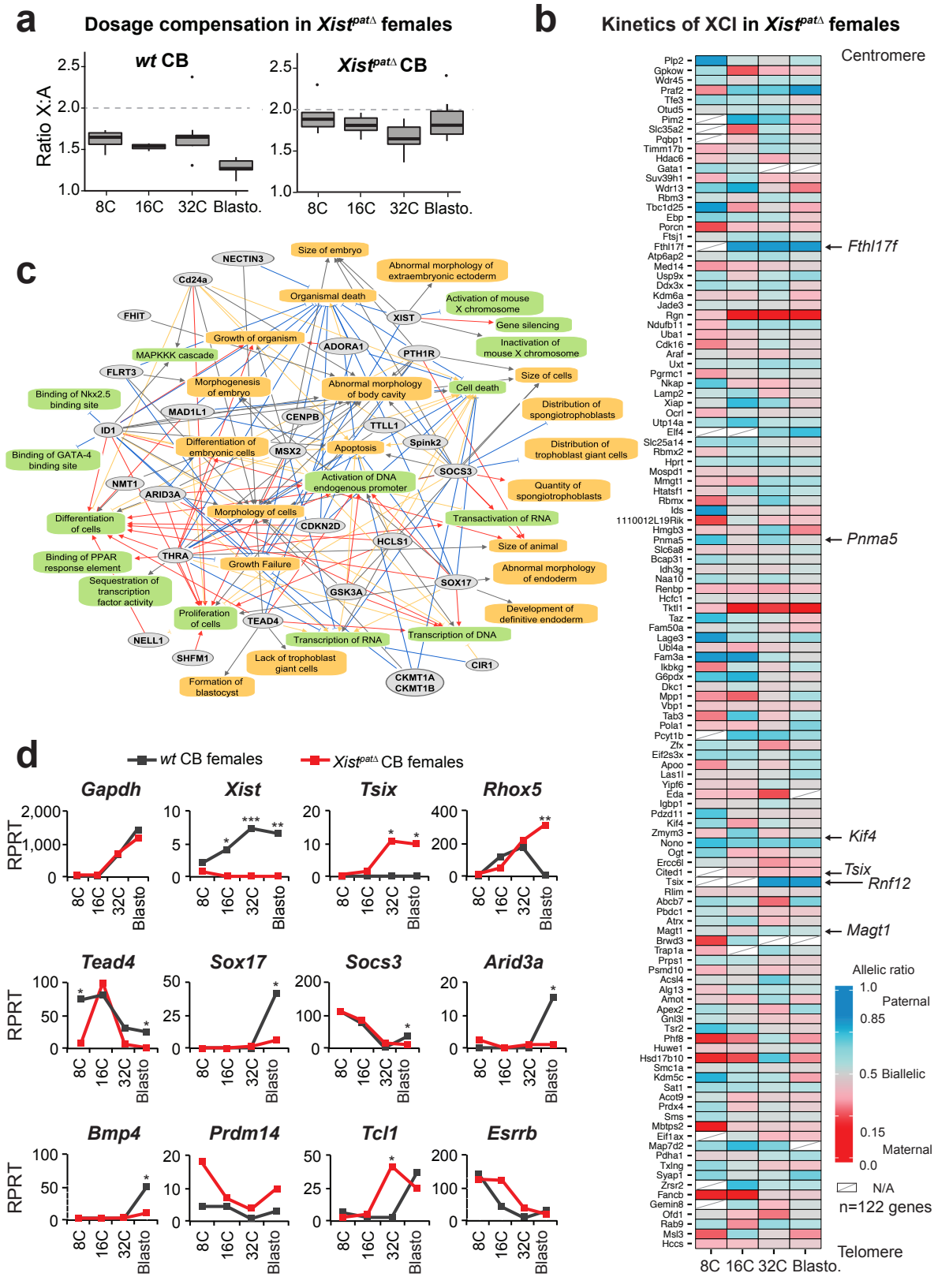
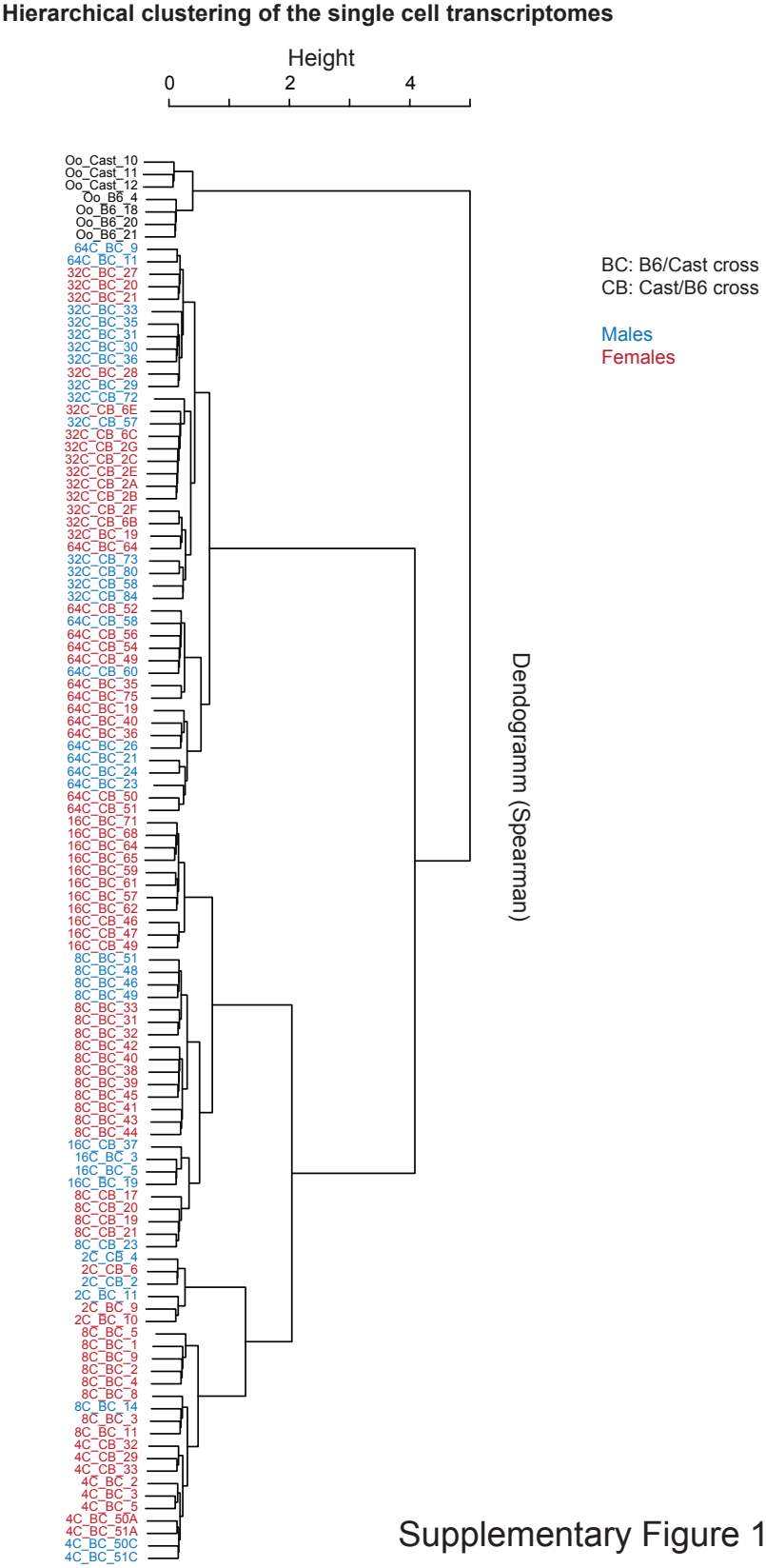


Figure 4- Borensztein *et al*

Supplementary Figure Legends

Supplementary Figure 1: Hierarchical clustering of scRNAseq

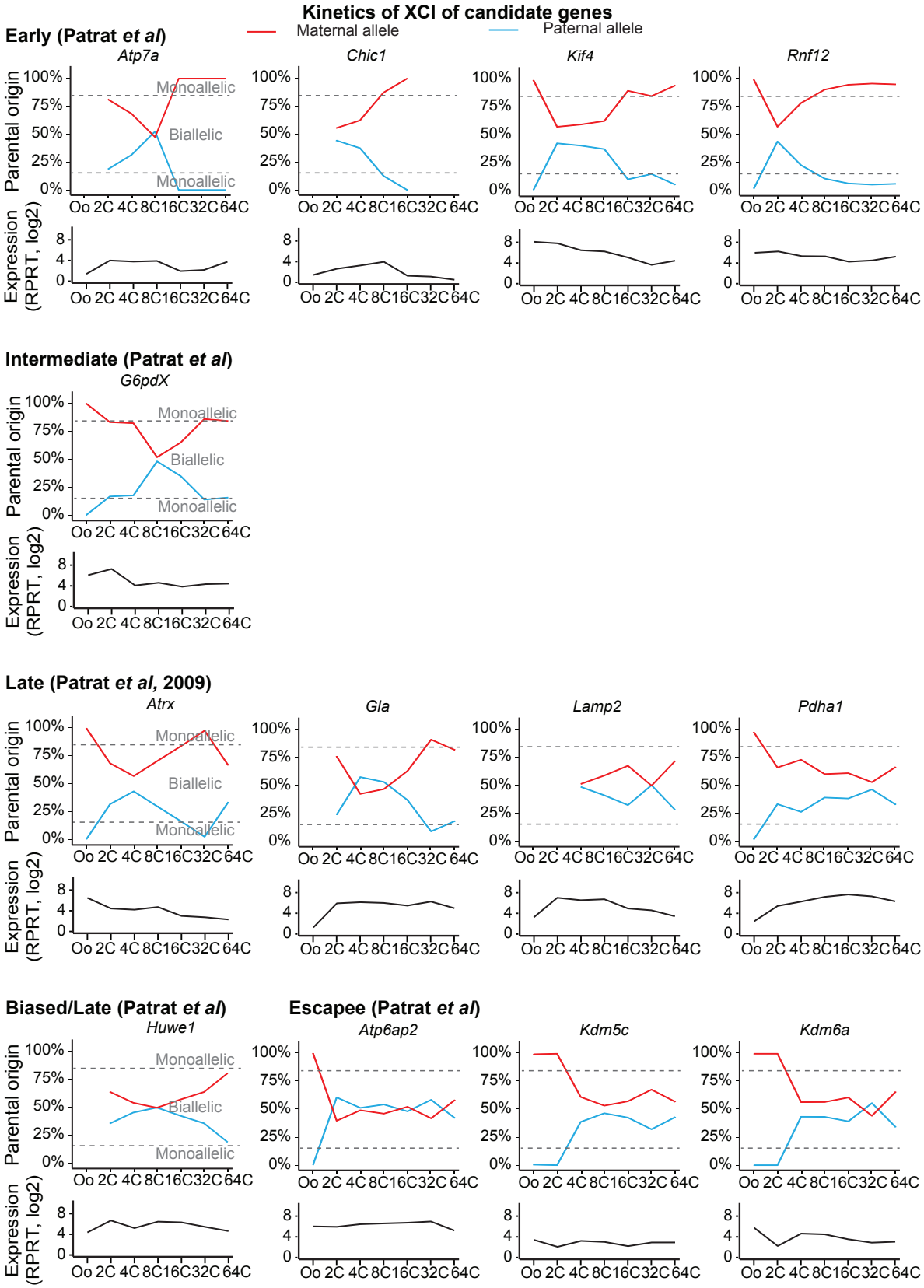
Hierarchical clustering of single cell transcriptomes, based on Spearman's correlation. Cells were clustered by developmental stage, then by cross (BC or CB), and finally according to their sex. ScRNAseq samples called 64C are related to the early blastocyst stage. n=184 single cell samples



Supplementary Figure 1- Borensztein *et al*

Supplementary Figure 2: Single cell RNA sequencing data corroborate RNA-FISH based silencing of kinetics.

Candidate X-linked gene silencing kinetics has been compared to previously known kinetics studied by RNA-FISH in Patrat *et al*[11]. Maternal and paternal reads are respectively represented as a red or a blue line. Percentage of parental origin transcript (top panel) and level of expression in RPRT (bottom panel) of each gene are represented between oocytes and blastocysts.

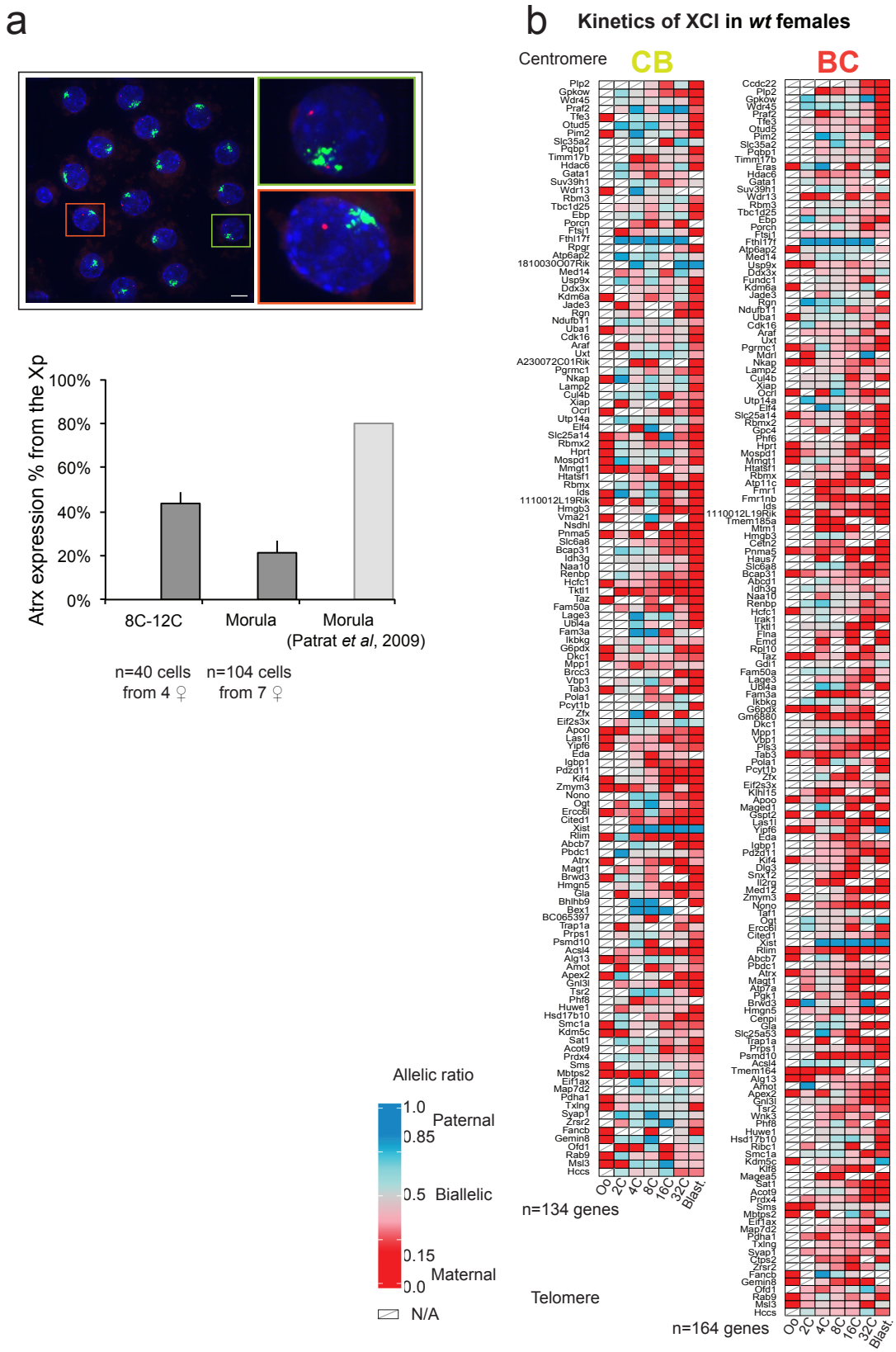


Supplementary Figure 2- Borensztein *et al*

Supplementary Figure 3: Gene validation by RNA-FISH and chromosome-wide representation of parental expression ratio.

(a) RNA FISH using probes recognizing *Xist* (signal in green) and *Atrx* (signal in red) nascent RNAs on a 16C female embryo. DAPI is in blue. Right pictures are enhanced pictures of two individual blastomeres. Percentage of nuclei showing pinpoints of nascent transcripts by RNA FISH from Xp and Xm has been assessed and summarized as the median under the picture. Normalization of the primary transcript detection frequency obtained for the paternal (*Xist* RNA-associated) allele in female embryos has been done thanks to that obtained for the maternal allele in male embryos at the same stage. Number of embryos and single cell processed are indicated under each genotype. Scale bar represents 10µm.

(b) Heatmaps are representing the mean of allele-specific expression of X-linked genes between Oocytes to blastocyst in CB and BC crosses. Strictly maternally expressed genes (allelic ratio ≤0.15) are represented in red and strictly paternally expressed genes (allelic ratio ≥0.85) in blue. Color gradients are used in between and genes have been ordered by genomic position. Oocytes and 2C data have been included to Figure 2 heatmaps as well as strain-specific genes.

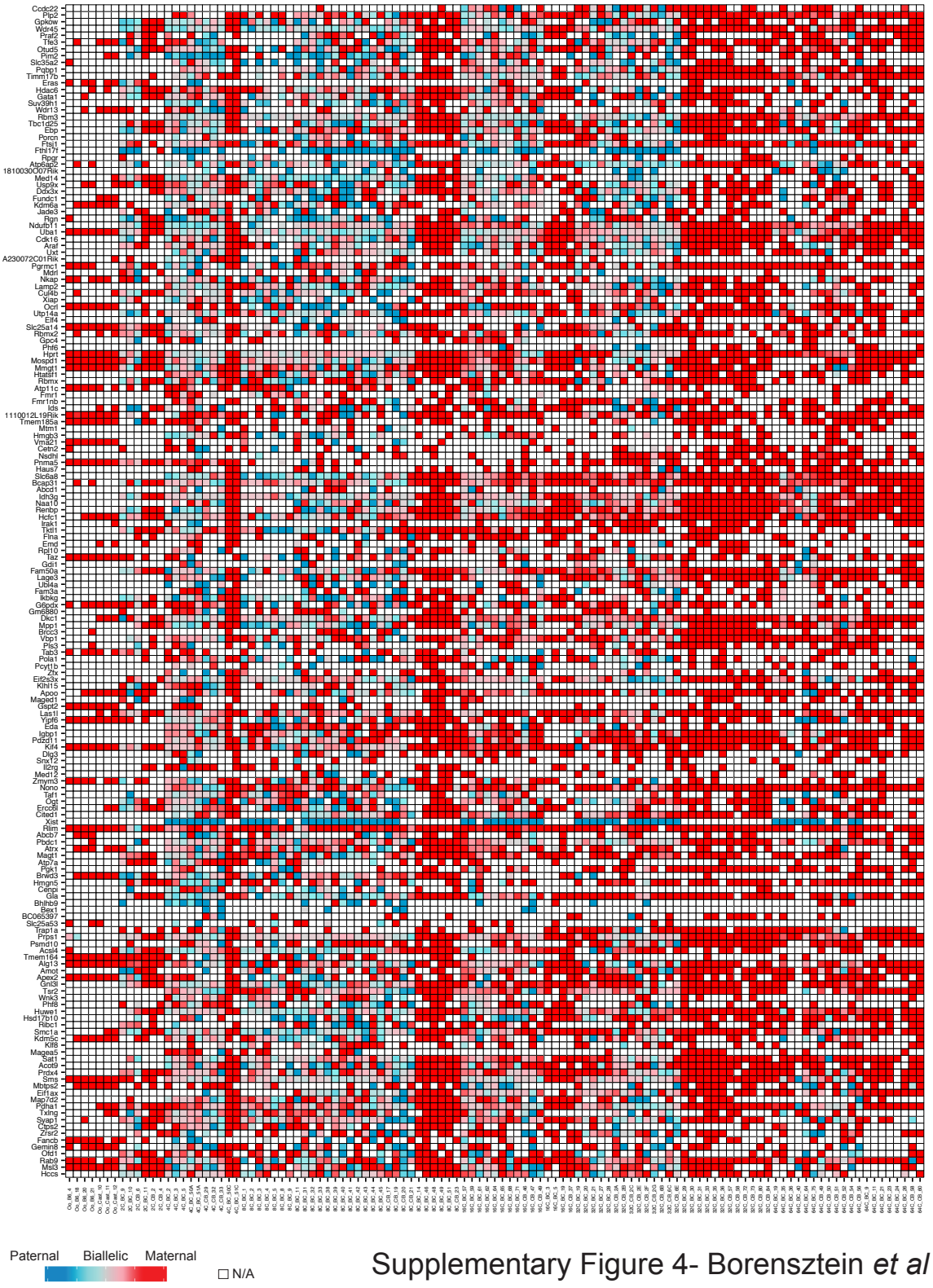


Supplementary Figure 3- Borensztein *et al*

Supplementary Figure 4: Complete single cell information of X-linked gene expression in pre-implantation development

Heatmap representing the allele-specific expression of informative and well expressed X-linked genes in each single cell, from oocytes to blastocysts, in male and female embryos derived from BC and CB crosses. Strictly maternally expressed genes (allelic ratio ≤0.15) are represented in red and strictly paternally expressed genes (allelic ratio ≥0.85) in blue. Color gradients are used in between and genes have been ordered by genomic position. Genes expressed in one or both crosses have been included to the single cell heatmap. n=173 genes.

Allelic expression ratio on X chromosome in males and females

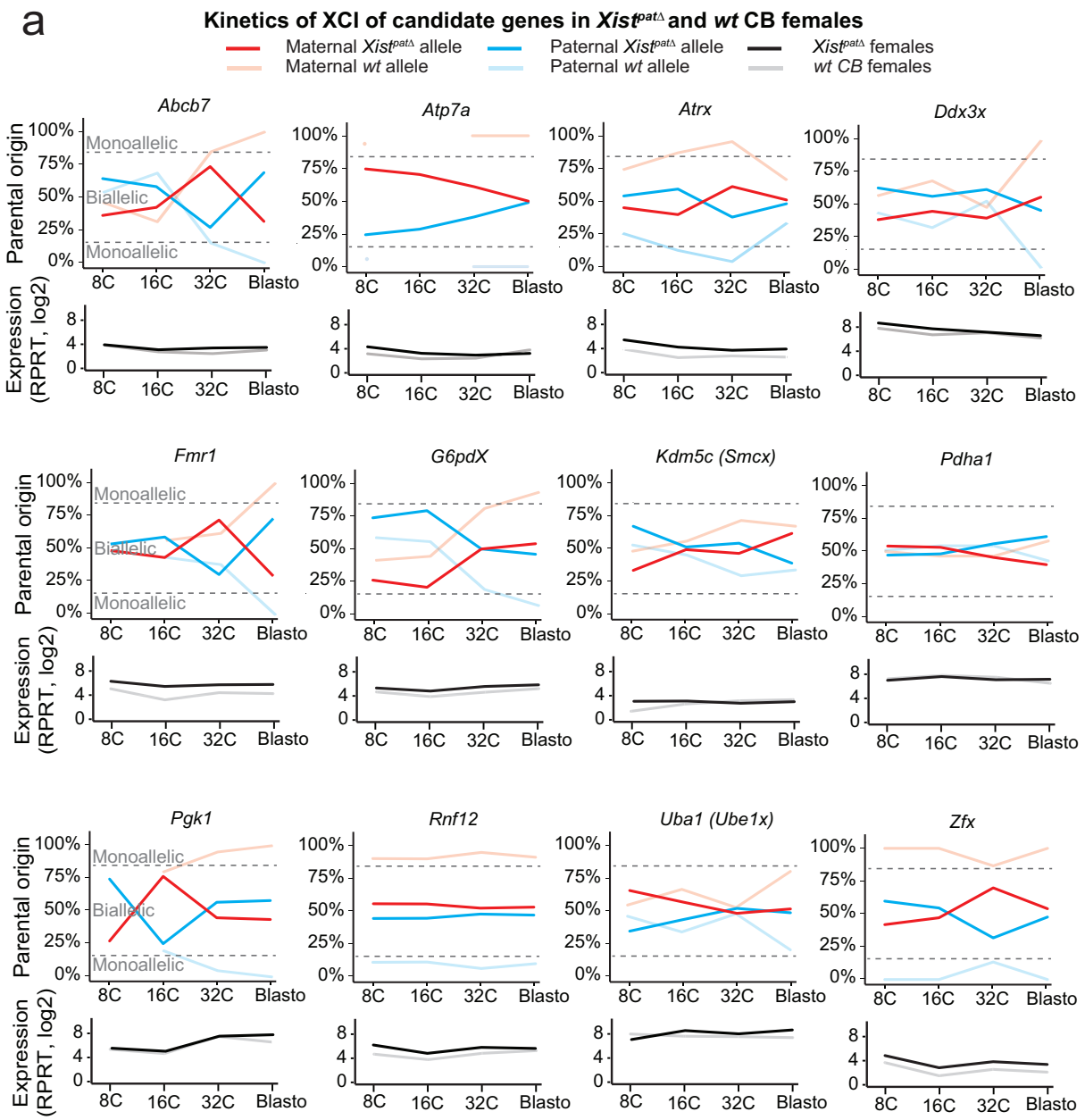


Supplementary Figure 4- Borensztein *et al*

Supplementary Figure 5: Absence of X-chromosome inactivation in absence of *Xist*.

(a) Each plot represents a candidate X-linked gene that has been previously studied in Kalantry *et al*, 2009[5]. Allele-specific expression ratio of each gene represents the number of reads mapped on paternal genome divided by the total number of paternal and maternal reads. Maternal and paternal reads are respectively represented as a red or a blue line. Percentage of parental origin transcript (top panel) and level of expression in RPRT (bottom panel) of each gene are represented between 8-cell stage and blastocysts.

(b) Table is summarizing the total differentially expressed (DE) genes between CB *wt* and CB $Xist^{pat\Delta}$ female embryos during early development and their localization in X chromosomes or autosomes. The percentages shown correspond to the repartition of total DE genes between autosomes and X chromosomes. The blastocyst stage has been highlighted in red and candidate DE genes from this stage have been analyzed in Figure 4.



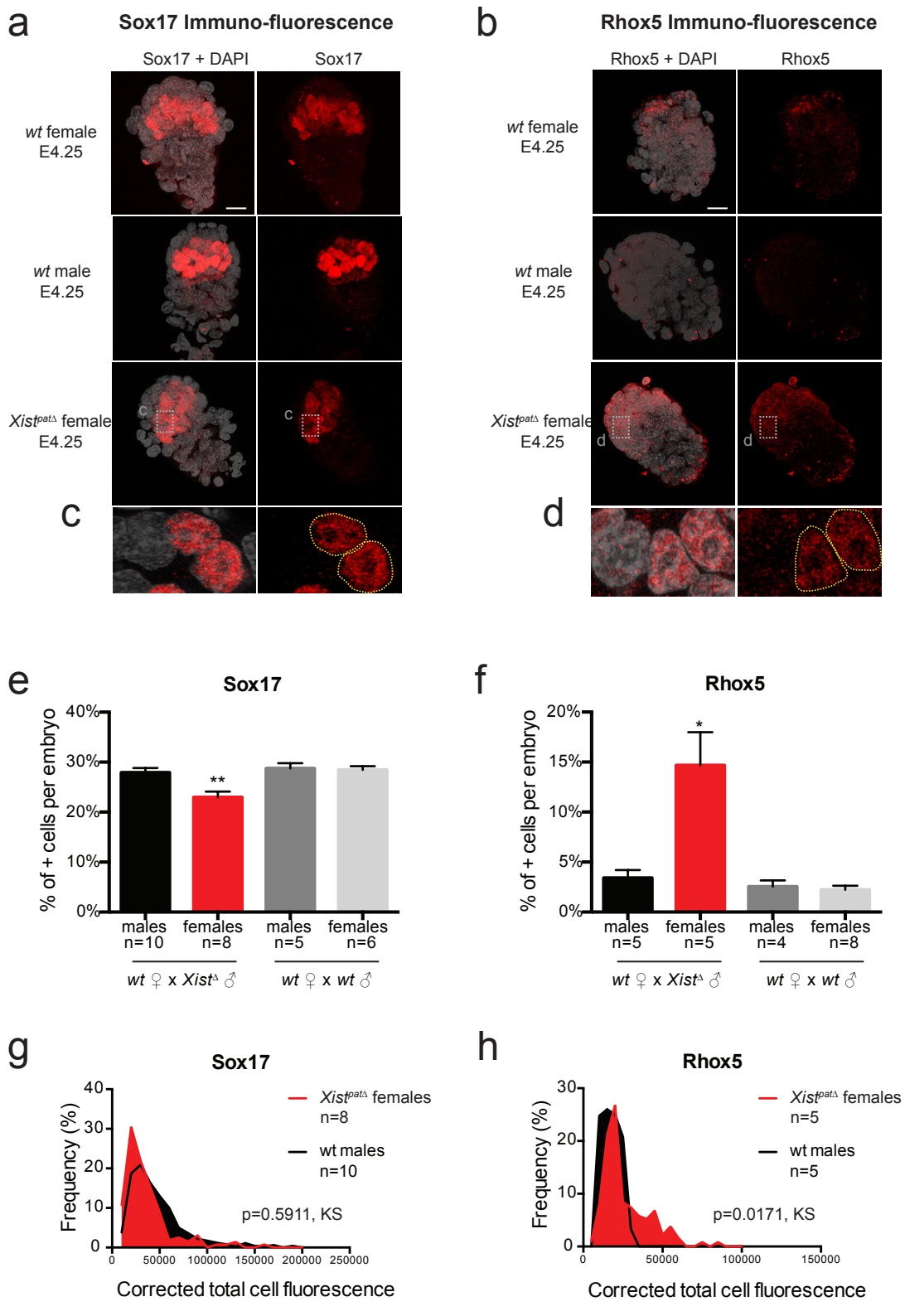
b **Total differentially expressed genes in *Xist*^{patΔ} CB females**

Developmental stages	DE genes from the X chromosome				DE genes from the autosomes			
	Total Upregulated		Total Downregulated		Total Upregulated		Total Downregulated	
	%	n	%	n	%	n	%	n
8-cell	5.0%	27	1.9%	14	95.0%	514	98.1%	541
16-cell	11.6%	15	1.0%	1	88.4%	114	99.0%	129
32-cell	4.7%	29	2.3%	12	95.3%	591	97.7%	620
Blastocyst	29.6%	16	3.6%	2	70.4	38	96.4%	54

Supplementary Figure 5 - Borensztein *et al*

Supplementary Figure 6: Abnormal Sox17 and Rhox5 pattern in *Xist*$^{pat\Delta}$ female blastocysts.

Maximum intensity projection of *wt* and *Xist*$^{pat\Delta}$ E4.25 blastocysts analyzed by immunofluorescence against Sox17 **(a, c**) or Rhox 5 **(b, d)**. Staining for Sox17 or Rhox5 is in red, DAPI is in grey. Scale bar represents 20µm. Percentage of positive cells have been assessed and summarized as the mean+sem for Sox17 **(e)** and Rhox5 **(f)**. Numbers of embryos are indicated under each genotype. By Kruskal-Wallis test; $p<0.05$ and $p<0.001$ correspond respectively to * and **. Quantification of immunofluorescence has been done by measuring the corrected total cell fluorescence using ImageJ software (Fiji, NIH) for Sox17 **(g)** and Rhox5 **(h)** and tested by Kolmogorov-Smirnov test.



Supplementary Figure 6- Borensztein *et al*

538

Supplemental Information

Supplementary Table 1: Summary of single cell RNAseq samples.

For each library is provided: single cell's name, stage, embryo number, gender, cross and the raw read number, filtered ones and percentage of mapping.

Supplementary Table 2: Silencing gene classes

Silencing classes for the 178 informative and well-expressed X-linked genes in BC and/or CB crosses.

Supplementary Table 3: Summary of escapees and their status in other studies.

Information for escapees is provided for CB and BC crosses. Status of each escapee gene has been evaluated in other studies using hybrid cell lines or tissues[13–15,33].

Supplementary Table 4: Misregulated genes in absence of paternal *Xist*

Table summarizing all the genes significantly different between CB *wt* and CB $Xist^{pat\Delta}$ female embryos during early development (8-cell to blastocyst).

Supplementary Table 5: Pathways and candidate genes down-regulated in $Xist^{pat\Delta}$ female blastocysts and highlighted by IPA analysis

Summary of all the pathway and candidate genes and relationships, which have emerged from IPA analysis between CB *wt* and CB $Xist^{pat\Delta}$ female blastocysts.

Methods

Mouse crosses and collection of embryos

All experimental designs and procedures were in agreement with the guidelines from French legislation and institutional policies.

All BC and CB embryos were respectively derived from natural meetings between *C57BL/6J* (B6) females crossed with *CAST/EiJ* (Cast) males or by the reciprocal cross. The $Xist^{pat\Delta}$ mutant embryos ($Xist^{+/-}$) have been obtained by mating between Cast females and $Xist^{-/Y}$ males (mixed background: B6D2F1: *C57BL/6J* and DBA/2J, 129S1/SvlmJ and BALB/cJ). Embryos have been harvested at 2-cell, 4-cell, 8-cell, 16-cell, 32-cell and blastocyst (approximately 60 to 64-cell) stages, respectively at E1.5, E2.0, E2.25, E2.75, E3.25 and E3.5. B6 and Cast pure oocytes have been collected at E0.5 after matings of females with vasectomized males (Figure 1a). The collected embryos were only included in the analysis if they showed a normal morphology and the right number of blastomeres in relation with their developmental stage.

RNA Fluorescent In Situ Hybridization

RNA FISH on preimplantation embryos was performed as previously described[3] using 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 using a wide-field Deltavision core microscope (Applied Precision – GE Healthcase) with a 60× objective (1,42 oil PL APO N) and 0.2 µm Z-sections. Images were analyzed using ImageJ software (Fiji, NIH).

Immunofluorescence staining

Immunofluorescence was essentially carried out as described previously[34] with an additional step of blocking in 3% FCS before the primary antibody incubation. Immunofluoroscence of embryos either from mutant or control male progeny have been always performed in parallel and in suspension. The following antibody were used: anti-goat Pem1 (Rhox5)/Santacruz sc-21650/1:50 and anti-goat Sox17/R&D Systems AF1924/1:100. Images were acquire using Inverted laser scanning confocal microscope with spectral detection (LSM700 - Zeiss) equipped with a 260nm laser (RappOpto), with a 60X objective and 2 µm Z-sections. Maximum projections and total corrected fluorescence measurements (=integrated density – (area of selected cell x mean fluorescence of background readings) were performed with Image J software (Fiji, NIH) using previously described methodology[35]. The total corrected cellular fluorescence (TCCF) = integrated density – (area of selected cell × mean fluorescence of background readings), was calculated.

Single cell dissociation from pre-implantation mouse embryos

Oocytes and embryos were collected by flushing oviducts (E0.5 to E2.75) or uterus (E3.25 and E3.5) with M2 medium (Sigma). The zona pellucida was removed using 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+}, Mg^{2+} free M2 medium for 5 to 20 minutes, depending on the embryonic stage. For the blastocyst stage, Ca^{2+}, Mg^{2+} free M2 free medium was replaced by a 5-minute incubation in TrypLE (Invitrogen). After incubation, each blastomere was mechanically dissociated by mouth pipetting with a thin glass capillary. Single cells were then washed 3 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 amplification or kept the single cells at -80°C for future preparation.

Single cell RNA amplification*:*

PolyA$^+$ mRNA extracted from each single cell was reverse transcribed from the 3'UTR and amplified following the *Tang et al* protocol[36]. Care was taken to process only embryos and single blastomeres of the highest quality based on morphology, number of cells and on amplification yield. A total of 72 BC and 110 CB (or 113 *wt* and 69 *Xist*$^{pat\Delta}$ mutant) blastomeres have been processed and passed quality controls.

Quality and sex determination

After cDNA amplification and before size selection and library preparation, quality of cDNAs from each of the 182 samples were validated by studying expression level of three housekeeping genes: *Gapdh*, *Beta-Actin* 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: cctgtggccatctgcctagt; Hprt_R: gggacgcagcaactgacatt. Care was taken to process only single cells with consistent amplification rate of the three housekeeping genes in the same developmental stage.

The sex of each embryo was accessed by 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 according to Illumina manufacturer protocol and deeply sequenced on an Illumina HiSeq 2500 instruments in single-end 50bp reads (Supplementary Table 1).

Quality control and filtering of raw data

Quality control was applied on raw data as previously described in (Ancelin et al, 2016)[34]. Sequencing reads characterized by at least one of the following criteria were discarded from the analysis:

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

SNP calling and allele-specific origin of the transcripts

SNPs 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. Using SNPsplit tool (v0.3.0)[37], these SNPs were filtered based on their quality values (FI value) and used to reconstruct the Cast genome from mm10 genome assembly.

Allele-specific alignments of RNAseq reads

To study the allele-specific gene expression, reads were processed using a pipeline adapted from Gendrel *et al,* 2014[33]. Single-end reads were first aligned to the mouse mm10 and CAST genomes using the TopHat2 software (v2.1.0)[34]. Only random best alignments with less than 2 mismatches were reported for downstream analyses. The resulting mapping files for both parental genomes were then merged for each sample, using these following rules:

1. If a read mapped at the same genomic position on the two genomes with the same number of mismatches, this read will be considered as a common read.
2. If a read is aligned with less mismatches on one genome, the best alignment will be

retained and this read will be considered as a specific read for the corresponding strain.

3. If a read is aligned with the same number of mismatches on both genomes but at different genomic positions, this read will be discarded.

Allelic imbalance in gene expression and gene classification

SNPs between *C57BL/6J* (B6) and *CAST/EiJ* (Cast) were extracted from the VCF file used to reconstruct the Cast genome. After removing common exonic SNPs between *Xist* and *Tsix* genes, 20,220,776 SNPs were retained.

The SAMtools mpileup utility (v1.1)[38] was then used to extract base-pair information at each genomic position from the merged alignment file. At each SNP position, the number of paternal and maternal allele was counted. Threshold used to call a gene informative was 5 reads mapped per single SNP with a minimum of 8 reads mapped on SNPs per gene to minimize disparity with low polymorphic gene. The allele-specific origin of the transcripts (or allelic ratio) has been measured by 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 are 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 only allows sequencing on average 3 kilobases from the 3'UTR, half of the expressed genes are partially covered (less than 50% of the gene size in average). To estimate the transcript abundance, read counts are thus normalized based on the

amplification size of each transcript (RPRT for Reads Per Retro-Transcribed length per million mapped reads) rather than the size of each gene (RPKM).

Filtering of biased SNPs

As we observed a bias for some polymorphisms in oocytes (maternal reads only) and male cells (maternal X chromosome only), oocytes (autosomes and X-chromosomes) and males (X-chromosome) were used to fix the issue. Therefore, SNPs covered by at least 5 reads and having an allelic ratio greater than 0.3 (biallelic or paternally expressed) in at least 2 of these samples were discarded. In total, 275 SNPs were filtered out, including 40 sites located on the X-chromosome.

Generation of $Xist^{pat\Delta}$ mutant embryos involved the use of a $Xist^{\Delta/Y}$ stud of mixed background (B6D2F1: *C57BL/6J* and DBA/2J, 129S1/SvlmJ and BALB/cJ). We have then applied another SNP filtration to the KO samples to remove all B6 polymorphisms that could have been lost on the X chromosome by the mix background of the $Xist^{\Delta/Y}$ stud. To do so, all existing SNPs between B6 and DBA/2J, 129S1/SvlmJ and BALB/cJ, on the X chromosome, have been removed from our SNP database (34,397 SNPs, which represent 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 using the RefFlat annotation and the HTSeq software[39] (v0.6.1). Only genes with a RPRT (Reads Per Retro-Transcribed length per million mapped reads) value >1 in at least 25% of the single cells of at least one developmental stage (with a minimum of 2 cells) were kept for the downstream analysis. The TMM method from the edgeR R-package (v3.14.0)[37] was first used to normalize the raw counts data. Principal component analysis (PCA) and hierarchical clustering have then been used to determine how

single cells were clustered to the others though their gene expression profiles, depending of their stage, sex and cross. PCA on normalized data was performed using FactoMineR R-package (v1.33). Hierarchical clustering analysis was based on Spearman correlation distance and the Ward method, using the hclust function implemented in the gplots R-package (v3.0.1). Limma R-package (v3.28.4)[38] was applied to identify the differentially expressed genes between 8-cell stage and blastocyst in control and *Xist*$^{pat\Delta}$ mutant females. Using the Benjamini-Hochberg correction, genes with an adjusted p-value lower than α=0.05 were called as differentially expressed.

Functional enrichment analysis

Down-regulated genes in *Xist*$^{pat\Delta}$ mutant female blastocysts compared to CB female blastocysts were analyzed using QIAGEN's Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City, www.qiagen.com/ingenuity). The Functions and Diseases module has been used to extract the most significantly deregulated pathways and their associated genes)[40].

Heatmap generation for X-chromosome allelic gene expression

For BC and CB heatmaps, 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 2 cells) at this particular stage and cross (Figures 2 and 4b and Supplemental Figure 3). To follow the kinetics of expression, we decided to focus only on genes expressed in at least 3 different stages between the 4-cell to blastocyst stages. Mean of the allelic ratio of each gene is represented for the different stages. The same gene candidate list was used to produce the *Xist*$^{pat\Delta}$ heatmaps (Figure 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 2 cells) per stage and cross.

737

738 **Definition of X-linked gene silencing/escape classes**

739 We have automatically assigned X-linked genes that become strictly maternal (allelic ratio
740 ≤0.15) at the 16-cell stage or before to the early silenced gene class; those that become
741 maternal at the 32-cell stage to the intermediate silenced class (allelic ratio equals NA or
742 >0.15 at 16C and ≤0.15 at 32C) and those that are silenced only by the blastocyst stage, to the
743 late silenced gene class (allelic ratio equals NA or >0.15 at 16C and 32C and ≤0.15 at
744 blastocyst stage). At the blastocyst stage, X-linked genes showing a maternal bias of
745 expression (0.15<allelic ratio≤0.3) are categorized as maternally biased. A final group
746 concerns genes that escape imprinted Xp inactivation (allelic ratio >0.3 at blastocyst stage)
747 (Figure 3a). Genes escaping XCI were separated into two classes: constitutive escapees if they
748 were classified as escapees in both CB and BC stages and as facultative escapees if they were
749 strain-specific escapees (Figure 3b and Supplementary Table 3).

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

754

755 **Dosage compensation, X:A expression ratio**

756 We measured the global X:A expression ratio in females (XX :AA ratio) and males (X :AA
757 ratio) as the level of expression of X-linked genes divided by the global level of expression of
758 the autosomal genes. Only genes with an expression value RPRT >4 were used for subsequent
759 analysis (Figures 1d and 4a). Adjustment of the number of expressed genes between X and
760 autosomes has been published to be critical for X:A expression ratio measurement[41]. We then
761 added a bootstrapping step and randomly selected, for each sample, an autosomal gene set
762 with the same number of expressed genes compared to the X to estimate the global X:A ratio.

763 This step was repeated 1000 times and the X:A expression ratio was estimated as the median
764 of the 1000 values.

765
766 **Data access**

767 The Gene Expression Omnibus (GEO) accession number for the data sets reported in this
768 paper is GSE80810.

Supplemental References :

33. Gendrel, A. V. *et al.* Developmental dynamics and disease potential of random monoallelic gene expression. *Dev. Cell* **28,** 366–380 (2014).

34. Ancelin, K. *et al.* Maternal LSD1 / KDM1A is an essential regulator of chromatin and transcription landscapes during zygotic genome activation. *Elife* **pii: e0885,** (2016).

35. McCloy, R. A. *et al.* Partial inhibition of Cdk1 in G2 phase overrides the SAC and decouples mitotic events. *Cell Cycle* **13,** 1400–1412 (2014).

36. Tang, F. *et al.* RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat. Protoc.* **5,** 516–35 (2010).

37. Rozowsky, J. *et al.* AlleleSeq: analysis of allele-specific expression and binding in a network framework. *Mol. Syst. Biol.* **7,** 522 (2011).

38. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).

39. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).

40. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2498–2504 (2003).

41. 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 (2011).