



**HAL**  
open science

## Parental-to-embryo switch of chromosome organization in early embryogenesis

Samuel Collombet, Noemie Ranisavljevic, Takashi Nagano, Csilla Varnai,  
Tarak Shisode, Wing Leung, Tristan Piolot, Rafael Galupa, Maud  
Borensztein, Nicolas Servant, et al.

► **To cite this version:**

Samuel Collombet, Noemie Ranisavljevic, Takashi Nagano, Csilla Varnai, Tarak Shisode, et al..  
Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature*, 2020, 580  
(7801), pp.142 - 146. 10.1038/s41586-020-2125-z . hal-03027197

**HAL Id: hal-03027197**

**<https://hal.science/hal-03027197v1>**

Submitted on 27 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Parental-to-embryo switch of chromosome organization in early embryogenesis

<https://doi.org/10.1038/s41586-020-2125-z>

Received: 3 April 2019

Accepted: 16 January 2020

Published online: 25 March 2020

 Check for updates

Samuel Collombet<sup>1,2,10</sup>, Noémie Ranisavljevic<sup>1,3,10</sup>, Takashi Nagano<sup>4,9,10</sup>, Csilla Varnai<sup>4,5</sup>, Tarak Shisode<sup>6</sup>, Wing Leung<sup>4,9</sup>, Tristan Piolot<sup>1</sup>, Rafael Galupa<sup>1,2</sup>, Maud Borensztein<sup>1</sup>, Nicolas Servant<sup>7</sup>, Peter Fraser<sup>4,8,11</sup>✉, Katia Ancelin<sup>1,11</sup>✉ & Edith Heard<sup>1,2,11</sup>✉

Paternal and maternal epigenomes undergo marked changes after fertilization<sup>1</sup>. Recent epigenomic studies have revealed the unusual chromatin landscapes that are present in oocytes, sperm and early preimplantation embryos, including atypical patterns of histone modifications<sup>2–4</sup> and differences in chromosome organization and accessibility, both in gametes<sup>5–8</sup> and after fertilization<sup>5,8–10</sup>. However, these studies have led to very different conclusions: the global absence of local topological-associated domains (TADs) in gametes and their appearance in the embryo<sup>8,9</sup> versus the pre-existence of TADs and loops in the zygote<sup>5,11</sup>. The questions of whether parental structures can be inherited in the newly formed embryo and how these structures might relate to allele-specific gene regulation remain open. Here we map genomic interactions for each parental genome (including the X chromosome), using an optimized single-cell high-throughput chromosome conformation capture (HiC) protocol<sup>12,13</sup>, during preimplantation in the mouse. We integrate chromosome organization with allelic expression states and chromatin marks, and reveal that higher-order chromatin structure after fertilization coincides with an allele-specific enrichment of methylation of histone H3 at lysine 27. These early parental-specific domains correlate with gene repression and participate in parentally biased gene expression—including in recently described, transiently imprinted loci<sup>14</sup>. We also find TADs that arise in a non-parental-specific manner during a second wave of genome assembly. These de novo domains are associated with active chromatin. Finally, we obtain insights into the relationship between TADs and gene expression by investigating structural changes to the paternal X chromosome before and during X chromosome inactivation in preimplantation female embryos<sup>15</sup>. We find that TADs are lost as genes become silenced on the paternal X chromosome but linger in regions that escape X chromosome inactivation. These findings demonstrate the complex dynamics of three-dimensional genome organization and gene expression during early development.

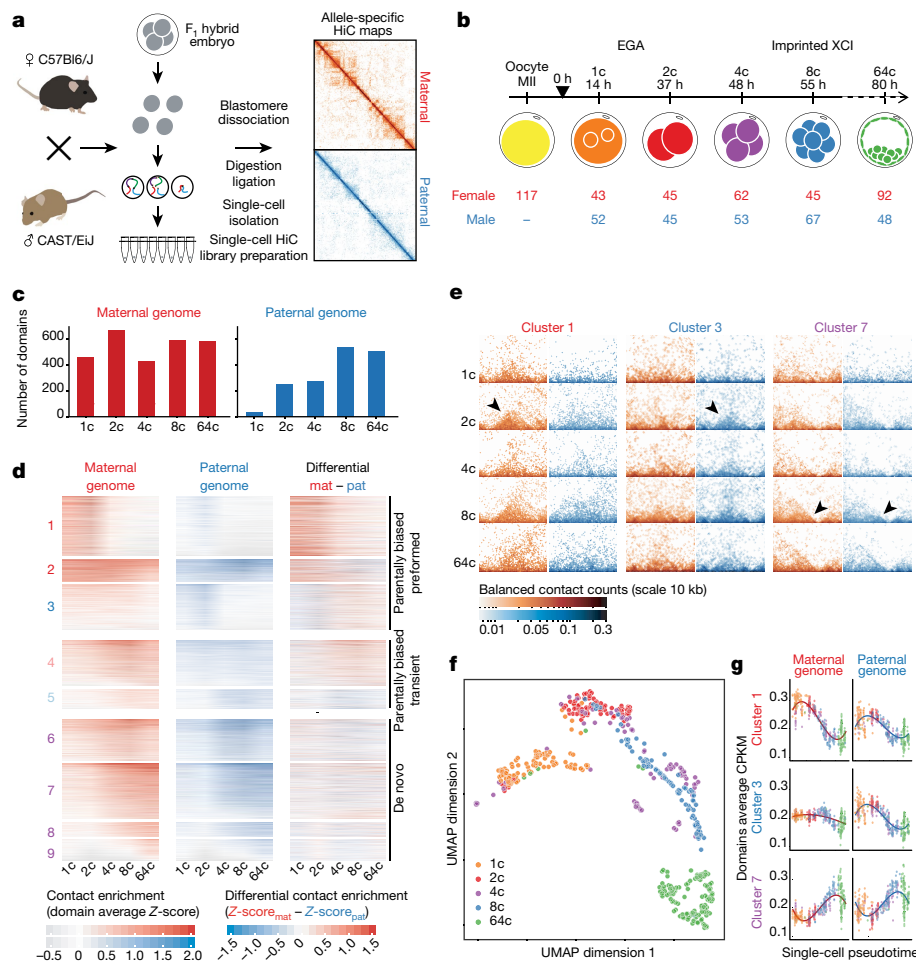
We performed allele-specific single-cell HiC, modified from previous studies<sup>12,13</sup>, on single blastomeres (at the 1-, 2-, 4-, 8- and 64-cell stages, as well as oocytes) from highly polymorphic F<sub>1</sub> hybrid embryos that were obtained by crossing female *Mus musculus domesticus* (C57Bl/6J) with male *Mus musculus castaneus* CAST/Eij) (Fig. 1a, b). After excluding cells with poor data quality (Methods, Extended Data Fig. 1a), we used the relative coverage of the two X chromosomes to investigate sex-specific differences beyond autosomes (Extended Data Fig. 1b). Finally, we used cell cycle phasing<sup>13</sup> to remove cells in the pre-M and M phases, in which chromosomes lose their organization into compartments and/or domains<sup>13,16</sup> (Extended Data Fig. 1c–e). Looking first at the total

contacts (that is, not split between alleles), we detected the formation of TAD-like domains, with clear boundaries that appeared at specific stages of development (Extended Data Fig. 1f). This was confirmed by DNA fluorescence in situ hybridization (FISH) on three-dimensional (3D) preserved embryos using intra- or interdomain-specific probes (Extended Data Fig. 2).

## Asymmetric chromosome architecture

Previous studies have investigated the dynamics of TADs in mouse embryos on the basis of TAD atlases defined in embryonic stem cells,

<sup>1</sup>Institut Curie, PSL Research University, CNRS UMR3215, INSERM U934, UPMC Paris-Sorbonne, Paris, France. <sup>2</sup>EMBL, Heidelberg, Germany. <sup>3</sup>Department of Reproductive Medicine - Gynecology, CHU and University of Montpellier, Montpellier, France. <sup>4</sup>Nuclear Dynamics Programme, The Babraham Institute, Cambridge, UK. <sup>5</sup>Centre for Computational Biology, University of Birmingham, Birmingham, UK. <sup>6</sup>Department of Applied Mathematics, Florida State University, Tallahassee, FL, USA. <sup>7</sup>Institut Curie, PSL Research University, INSERM U900, Mines ParisTech, Paris, France. <sup>8</sup>Department of Biological Science, Florida State University, Tallahassee, FL, USA. <sup>9</sup>Present address: Laboratory for Nuclear Dynamics, Institute for Protein Research, Osaka University, Suita, Japan. <sup>10</sup>These authors contributed equally: Samuel Collombet, Noémie Ranisavljevic, Takashi Nagano. <sup>11</sup>These authors jointly supervised this work: Peter Fraser, Katia Ancelin, Edith Heard. ✉e-mail: pfraser@bio.fsu.edu; Katia.Ancelin@curie.fr; Edith.Heard@embl.org



**Fig. 1 | Single-cell HiC approach to studying chromosome organization in preimplantation embryos in the mouse. a**, Scheme of the single-cell HiC method on mouse F<sub>1</sub> embryos. **b**, Timeline of embryo collection at selected stages. The numbers of blastomeres after quality-filtering and sex assignment are indicated (c refers to cell stage). EGA, embryonic genome activation; XCI, X chromosome inactivation. **c**, Number of domains at different stages, on the maternal (red) or paternal genome (blue). **d**, Clustering of domain dynamics (rows) through stages (columns). Colour scale indicates contact enrichments

and have not attempted to identify any alternative, embryo-specific domains<sup>5–9,11</sup>. Our allelic data revealed that parental genomes display a notably asymmetric structural organization before the eight-cell stage; the maternal genome displays most of the domains called at the one- and two-cell stages (Fig. 1c). We detected two independent gains in domain number—the first at the two-cell stage, and the second at the eight-cell stage. The second round of domain formation at the eight-cell stage correlated with a previously reported progressive acquisition of TADs<sup>8,9</sup> (Extended Data Fig. 3a). To better capture the dynamics of allelic domain organization, we quantified the contact enrichment inside domains (Methods) for both parental genomes at each stage and performed an unsupervised clustering (Fig. 1d, Extended Data Fig. 3a, b). We found that domains fall into three main categories. The first category (clusters 1–3) comprises parentally biased preformed domains, which are present as early as the one-cell stage and show a bias for the maternal (Fig. 1e, left) or paternal genome (Fig. 1e, middle). Most of these domains (those in clusters 1 and 3) disappear by the 4-cell stage, but a subset of maternally preformed domains (cluster 2) becomes balanced by the blastocyst stage (64-cell stage). A second category (clusters 4 and 5) of domain exhibits a more-transient bias for one allele, and generally has a weaker structure. In the third set (clusters 6–9), domains are acquired symmetrically on the two parental genomes at

inside domains (average Z-score (Methods)) and the difference in enrichment between the two alleles. Mat, maternal; pat, paternal. **e**, Snapshots of HiC maps from the maternal (red) and paternal (blue) alleles for 3 regions, in cluster 1, cluster 3 and cluster 7. Arrowheads indicate the domains of interest. **f**, Single-cell projection in reduced space (uniform manifold approximation and projection (UMAP)) based on the quantification of domain contacts (n = 470 cells). **g**, Cluster average contacts per kilobase per million contacts (CPKM) in single cells, ordered by pseudotime from the trajectory in f.

different stages after embryonic genome activation (Fig. 1e, right), as previously described<sup>8,9</sup> (Extended Data Fig. 3c).

We also assessed whether these dynamics were discernible in single cells, and were not an effect of the evaluation on pseudo-bulk data. Notably, the quantification of domain contacts in single cells was sufficient to capture the developmental trajectories of early embryos (Fig. 1f, Extended Data Fig. 3d–i), as well as capturing the dynamics of the clusters that we identified in the pseudo-bulk data (Fig. 1g, Extended Data Fig. 3d–i).

In conclusion, our results identify parent-of-origin-specific levels of chromosome organization as early as the 1-cell stage that are mostly resolved as the 2 genomes mature towards the 64-cell stage, except for cluster 2. These data reconcile those of previous studies<sup>5,8,9</sup>, and provide insights into the early differential organization of the two parental genomes.

## Parental domains and histone modification

To evaluate whether this unusual parental asymmetry in structure might be linked to specific chromatin states, we integrated our data with chromatin immunoprecipitation and sequencing (ChIP-seq) data for histone modifications from early embryos<sup>17,18</sup>. Notably,

parental-specific early domains (clusters 1–3) coincide with large accumulations of the Polycomb-associated mark, trimethylation of histone H3 at Lys27 (H3K27me3); the strongest enrichment of this mark is associated with the maternal genome, whereas the de novo-formed domains (clusters 6–9) are depleted for this mark (Fig. 2a, b, Extended Data Fig. 4a–e). Whereas H3K27me3 domains are maintained up to the eight-cell stage and diminish thereafter (Fig. 2c), the structural domains are lost or transformed by the four-cell stage—concomitantly with a transient gain in the H3K4me3 mark (Fig. 2c, Extended Data Fig. 4f). We note that the enrichment of H3K27me3 occurs during oogenesis (Extended Data Fig. 4g) and that the domains of cluster 2 appear as early as postnatal day 5, but not in sperm (Extended Data Fig. 4h).

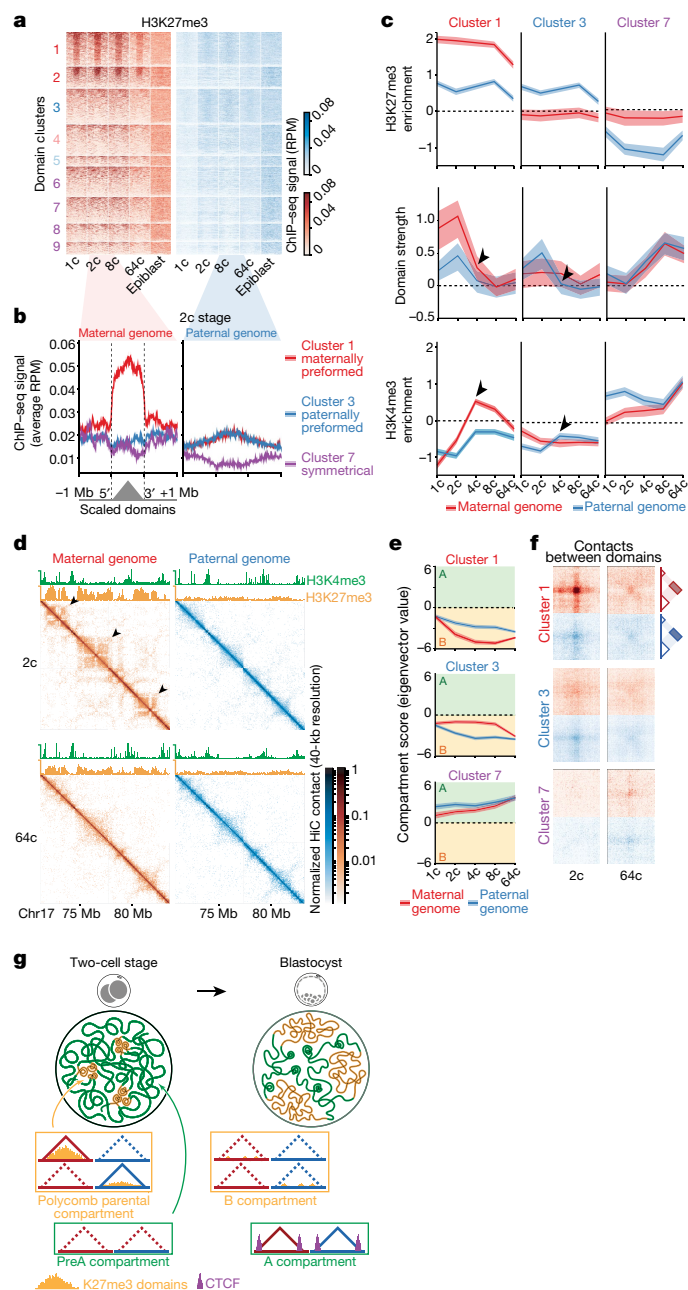
Parentally preformed domains also exhibit interactions between domains similar to the patterns of A and B compartments (Fig. 2d). We found that the parentally preformed domains form allele-specific B-like compartments at the two-cell stage (Fig. 2e, Extended Data Fig. 4i, j). These domains also display stronger interactions between domains at the 2-cell stage than do the de novo domains at the 64-cell stage (Fig. 2f, Extended Data Fig. 4k). Parentally preformed domains are depleted for CTCF motifs flanking their borders (Extended Data Fig. 4i), which points towards an independency for this factor (as has previously been shown for compartments<sup>19</sup>). Altogether, these results suggest that parental-specific domains might form local compartments associated with the Polycomb-repressive mark after fertilization, which later dissolve into the classical A and B compartments (Fig. 2g).

### Parental domains and transient imprint

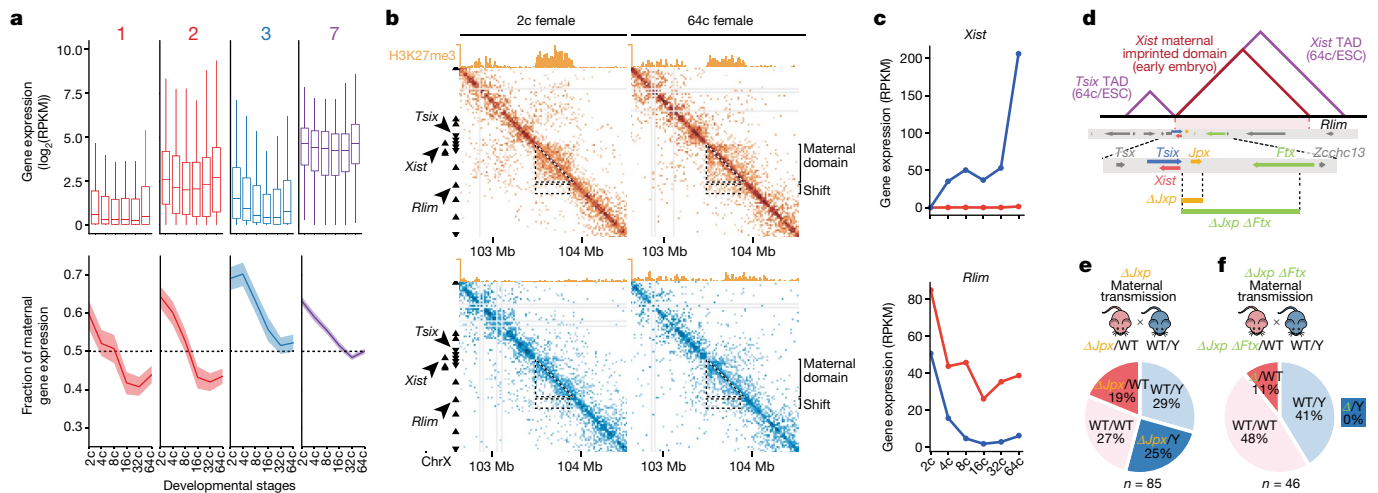
To evaluate how the allele-specific dynamics of chromosome organization relate to gene expression, we examined previously published RNA-sequencing data<sup>20</sup> obtained from equivalent F<sub>1</sub> hybrid preimplantation embryos. We found that parentally preformed domains are associated with generally lower gene expression (Fig. 3a, top, Extended Data Fig. 5a) and an average lower expression on the structured allele (Fig. 3a, bottom), as well as a higher frequency of strongly biased genes (Extended Data Fig. 5b). Gene ontology analysis revealed that silenced genes within early preformed clusters are significantly enriched for terms associated with tissue morphogenesis, such as neurogenesis or osteogenesis (Extended Data Fig. 5c), the expression of which is required only at late developmental stages. Conversely, symmetric de novo clusters were predominantly enriched in genes that drive the patterning of the embryo at preimplantation (such as cell cycle, lineage specification, metabolism and gene regulation).

Maternally preformed domains encompass most genes that have previously been described as transiently maternally imprinted<sup>14,21</sup> (19 out of 27 genes), such as the X inactivation centre locus (Fig. 3b, Extended Data Fig. 6a, b). Indeed, at the two-cell stage *Xist* is encompassed in a maternal-specific domain, the left border of which coincides with the *Xist* TAD that has previously been described in embryonic stem cells<sup>22</sup>; the right border of this maternal-specific domain is slightly shifted with respect to the previously described *Xist* TAD, and excludes the *Xist* transactivator *Rlim* (Extended Data Fig. 6c). Accordingly, *Xist* is maternally repressed, whereas the adjacent *Rlim* is kept expressed on the maternal allele and becomes silenced upon X chromosome inactivation<sup>23</sup> (Fig. 3c). We noticed a similar pattern of shifting from maternal imprinted domains at the two-cell stage to TADs at the blastocyst stages for other transiently imprinted genes, such as *Tle3*, *Enc1* and *Mbnl2* (Extended Data Fig. 6d–h).

To investigate the importance of such domains for imprinted gene regulation, we focused on the maternal 3D domain spanning the *Xist* locus and engineered genetic deletions around the *Jpx* and *Ftx* loci, within a region that has previously been proposed to be sufficient for imprinted X chromosome inactivation<sup>23</sup> (Fig. 3d). *Jpx* is a putative regulator of *Xist*<sup>24</sup>. We found that mice with a deletion encompassing



**Fig. 2 | Early domains are associated with Polycomb and form local compartments.** **a**, Heat maps of H3K27me3 ChIP-seq signal on domains (scaled to 1 Mb) from each cluster, on the maternal (red) (left) and paternal (blue) (right) alleles. Data are taken from Gene Expression Omnibus (GEO) accession GSE76687. RPM, reads per million reads sequenced. **b**, Average ChIP-seq signal at the 2-cell stage on the maternal (left) and paternal (right) alleles, for cluster 1 ( $n = 375$  domains), cluster 3 ( $n = 387$  domains) and cluster 7 ( $n = 287$  domains). **c**, Quantification of H3K27me3 (top) or H3K4me3 (bottom) enrichment (versus mean of the genome (Methods)) or domain strength (middle, average Z-score) for cluster 1, cluster 3 and cluster 7 ( $n$  values as in **b**). Lines represent the mean, and shading represents the 95% confidence interval of the mean. The maternal allele is in red and paternal allele is in blue. H3K4me3 data are taken from GSE71434. **d**, Snapshots of ChIP-seq and HiC maps (40-kb resolution) on the maternal (left) and paternal (right) alleles for a locus on chromosome 17. **e**, Dynamics of compartment scores (principal component analysis first eigenvectors) for cluster 1, cluster 3 and cluster 7 ( $n$  values as in **b**). The A and B compartments are assigned on the basis of gene density (Methods). Line represents the mean, and shading represents the 95% confidence interval of the mean. **f**, Average HiC map enrichment of long-distance interactions (>1 Mb) around the intersection between domain centres ( $n$  values as in **b**). **g**, Model of the parental preformed local compartment to de novo-acquired conventional A and B compartments.



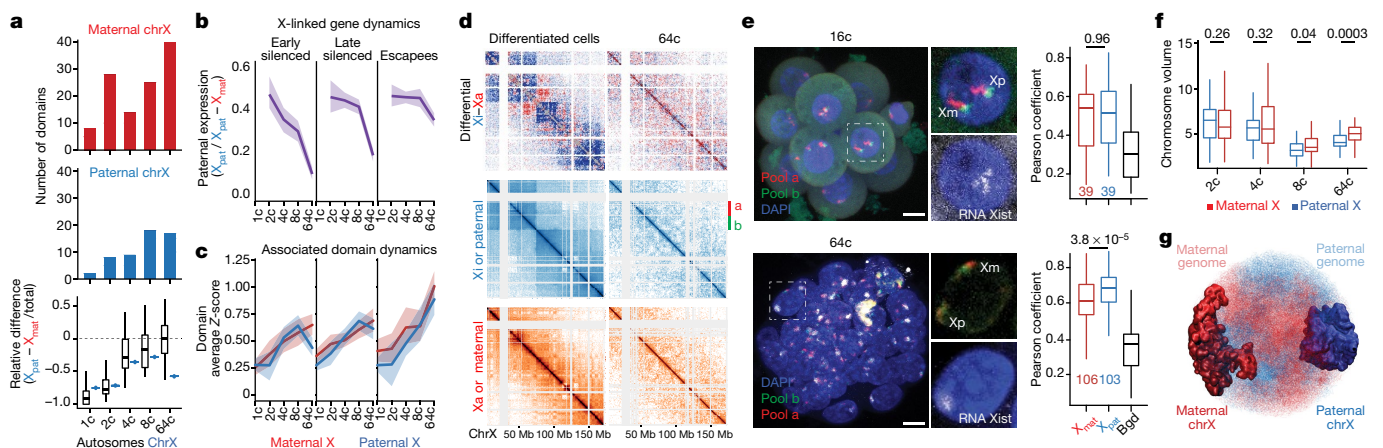
**Fig. 3 | Parental preformed domains are associated with a transient imprint.**

**a**, Distribution of gene expression (top) and fraction of maternal expression (maternal/(maternal + paternal)) (bottom) for genes present within the domains of selected clusters. Box plots (top) represent  $\pm 1.5\times$  interquartile range, 25th and 75th percentiles and median values. Lines represent the mean, and shading represents the 95% confidence interval of the mean. RPKM, reads per kilobase of transcript per million mapped reads. **b**, Snapshots of HiC on the maternal (red) and paternal (blue) genome, and H3K27me3 ChIP-seq, at the

*Xist* locus (female cells only were pooled;  $n = 43$  at the 2-cell stage,  $n = 83$  at the 64-cell stage). **c**, Allele-specific expression of *Rlim* and *Xist* from the 2-cell stage to the 64-cell stage. **d**, Scheme of the X inactivation centre and of the CRISPR deletions that we engineered. ESC, embryonic stem cell. **e**, Genotype distribution after maternal transmission of *Jpx* deletion ( $n = 85$  pups). WT, wild type. **f**, Genotype distribution after maternal transmission of *Jpx* and *Ftx* deletion ( $\Delta$ ) ( $n = 46$  pups).

*Jpx* are viable, and that normal expression of *Xist* occurs in these mice (Fig. 3e, Extended Data Fig. 6i, j). Whereas *Ftx* deletion alone is dispensable for imprinted X chromosome inactivation in preimplantation embryos<sup>25</sup>, the maternal transmission of the deletion containing *Jpx* and *Ftx* strongly compromised female viability (5  $\Delta Jpx$ -*Ftx*/wild-type

female mice out of 46 pups received the deleted allele, corresponding to 11% transmission) and no viable male could be obtained (0% transmission) (Fig. 3f). Taken together, our analysis identifies a minimal control region for imprinting in proximity to *Xist*, and opens up new possibilities for testing other transient imprint regions.



**Fig. 4 | Structural changes at the paternal X chromosome during imprinted X chromosome inactivation.**

**a**, Number of domains called on the maternal X chromosome (red) and on the paternal X chromosome (blue) over preimplantation development, and the relative difference of domains on the paternal versus maternal alleles (autosomes in black, X chromosome in blue). Box plot represents the  $\pm 1.5\times$  interquartile range, 25th and 75th percentiles and median value for all autosomes ( $n = 19$ ). **b**, Expression dynamics for early-silenced, late-silenced and escapee genes ( $n = 40, 76$  and 52, respectively, as in a previous study<sup>20</sup>). **c**, Structural changes in the corresponding domains that contain genes in the categories shown in **b**. **d**, Parental differential (top) and allele-specific (bottom and middle) HiC contact maps (pooled female cells) over the entire X chromosome (resolution of 640 kb) in neural progenitor cells (left) and at 64-cell stage (right). DNA FISH probes (oligonucleotide pools a and b) are indicated in colours over the centromeric megadomain. Lines represent the mean and shading represents the 95% confidence interval of the mean. Xa, active X chromosome; Xi, inactive X chromosome. **e**, Representative

3D RNA-DNA FISH images of 16-cell stage (top) or 64-cell stage (bottom) embryo with corresponding box plot ( $\pm 1.5\times$  interquartile range, 25th and 75th percentiles and median value) quantifications for signal correlations. Statistical significance ( $P < 0.001$ ) was assessed using Wilcoxon's rank-sum test (two-sided).  $n = 39$  nuclei from 8 female embryos for 16-cell stage;  $n = 106$  and 103 signals from 106 nuclei from 4 female embryos for 64-cell stage. DNA is counterstained with DAPI. Scale bar, 10  $\mu\text{m}$ . Bgd, background; Xm, maternal X chromosome; Xp, paternal X chromosome. **f**, Dynamics of the volume of the paternal and maternal X chromosomes. Box plots represent  $\pm 1.5\times$  interquartile range, 25th and 75th percentiles and median value.  $P$  values are indicated above the box plot, and were calculated using Mann-Whitney  $U$  test (two-sided).  $n = 43, 46, 35$  and 83 single cells for the 2-cell, 4-cell, 8-cell and 64-cell stages, respectively. **g**, Three-dimensional model of whole-genome conformation for 64-cell-stage single cell number 118. Maternal chromosomes are in red; paternal chromosomes are in blue; autosomes in thin line and X chromosomes are highlighted. The model was computed at 500-kb resolution.

## Features of imprinted X inactivation

In differentiated female cells, the inactive X chromosome is organized into two megadomains rather than into A and B compartments, and displays a marked weakening of TADs<sup>26,27</sup>; however, little is known of the dynamics of this organization during development. Pooling only female cells, we found that the paternal X chromosome displays a strong deficit in domains compared to its maternal counterpart (Fig. 4a). Whereas preformed maternal domains are lost, domains that are formed de novo become weaker on the paternal genome by the blastocyst stage, with the exception of a small subset of domains (Extended Data Fig. 7a). Comparing the dynamics of structural domains with those of gene expression, we found that early silenced loci on the paternal X chromosome show a marked loss of domain strength only after the eight-cell stage (that is, after silencing initiation), and domains that contain late-silenced genes display little structural change (although imprinted X chromosome inactivation is largely complete) (Fig. 4b, c). Although we cannot formally exclude that this might be due to differences in sensitivity between RNA sequencing and single-cell HiC, these results suggest that the loss of TAD structure on the paternal X chromosome would follow or accompany, rather than precede, gene silencing.

Using 3D modelling of chromosomes, we also found that early silenced genes are localized more at the centre of the paternal X chromosome whereas escapees tend to be located at its periphery (Extended Data Fig. 7b), similar to differentiated cells<sup>28</sup>. However, megadomains do not appear on the paternal X chromosome (Fig. 4d) despite a higher colocalization of intradomain probes by DNA FISH (Fig. 4e), which suggests a global compaction of the inactive paternal X chromosome. Three-dimensional modelling confirmed that the paternal X chromosome was substantially smaller (by approximately a third) than its maternal homologue at the 64-cell stage (Fig. 4f) and adopted a more globular shape (whereas the maternal X chromosome is more elongated) (Fig. 4g), as has previously been reported in somatic cells<sup>26</sup>.

## Conclusions

Here we show that higher-order chromatin structure matures from parental-specific and early repressive compartments towards a progressive establishment of TADs in early development in the mouse (Fig. 2g). This developmental switch might illustrate the autonomous mechanisms at play—cohesin-dependent and -independent—that have previously been observed for the 3D organization of the genome<sup>29</sup> and that might also reflect the unusual chromatin landscape and nuclear organization of the early embryo, compared to later developmental stages<sup>1,30</sup>. Early compartments are Polycomb-marked and are accompanied by contrasting allelic gene-expression states. These parentally preformed repressive domains may be important in counterbalancing genome-wide embryonic genome activation for transiently imprinted genes such as *Tle3* (the dose of which affects the pluripotency programs<sup>31</sup>) or *Xist* (which is central to the process of gene dose compensation in females<sup>32</sup>). Our study also illustrates that, after embryonic genome activation, structures tend to be TAD-like and their appearance is generally linked to active chromatin states. In the case of the paternal X chromosome, the loss of TAD structure during X chromosome inactivation is a late event that seems to follow—rather than precede—gene silencing. Furthermore, we find that there is progressive compaction of the paternal X chromosome, but no megadomain formation, by the blastocyst stage. Local domains are maintained only across escapee loci, suggesting that local structure might require an active chromatin state and/or transcription.

Overall, our study provides broad insights into the intricate interplay between chromosome folding and parental gene activity with the developmental potential of the early embryo.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2125-z>.

- Burton, A. & Torres-Padilla, M.-E. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 723–735 (2014).
- Wu, J. et al. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652–657 (2016).
- Dahl, J. A. et al. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**, 548–552 (2016).
- Liu, X. et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**, 558–562 (2016).
- Flyamer, I. M. et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017).
- Battulin, N. et al. Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach. *Genome Biol.* **16**, 77 (2015).
- Jung, Y. H. et al. Chromatin states in mouse sperm correlate with embryonic and adult regulatory landscapes. *Cell Rep.* **18**, 1366–1382 (2017).
- Ke, Y. et al. 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* **170**, 367–381.e20 (2017).
- Du, Z. et al. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**, 232–235 (2017).
- Wu, J. et al. Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature* **557**, 256–260 (2018).
- Gassler, J. et al. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J.* **36**, 3600–3618 (2017).
- Nagano, T. et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2013).
- Nagano, T. et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017).
- Inoue, A., Jiang, L., Lu, F., Suzuki, T. & Zhang, Y. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* **547**, 419–424 (2017).
- 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).
- Naumova, N. et al. Organization of the mitotic chromosome. *Science* **342**, 948–953 (2013).
- Zheng, H. et al. Resetting epigenetic memory by reprogramming of histone modifications in mammals. *Mol. Cell* **63**, 1066–1079 (2016).
- Zhang, B. et al. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**, 553–557 (2016).
- Nora, E. P. et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944.e22 (2017).
- Borensztein, M. et al. Xist-dependent imprinted X inactivation and the early developmental consequences of its failure. *Nat. Struct. Mol. Biol.* **24**, 226–233 (2017).
- Inoue, A., Jiang, L., Lu, F. & Zhang, Y. Genomic imprinting of *Xist* by maternal H3K27me3. *Genes Dev.* **31**, 1927–1932 (2017).
- Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
- Okamoto, I. et al. Evidence for *de novo* imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* **438**, 369–373 (2005).
- Tian, D., Sun, S. & Lee, J. T. The long noncoding RNA, *Jpx*, is a molecular switch for X chromosome inactivation. *Cell* **143**, 390–403 (2010).
- Soma, M., Fujihara, Y., Okabe, M., Ishino, F. & Kobayashi, S. *Ftx* is dispensable for imprinted X-chromosome inactivation in preimplantation mouse embryos. *Sci. Rep.* **4**, 5181 (2014).
- Giorgetti, L. et al. Structural organization of the inactive X chromosome in the mouse. *Nature* **535**, 575–579 (2016).
- Deng, X. et al. Bipartite structure of the inactive mouse X chromosome. *Genome Biol.* **16**, 152 (2015).
- Chaumeil, J., Le Baccon, P., Wutz, A. & Heard, E. A novel role for *Xist* RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev.* **20**, 2223–2237 (2006).
- Schwarzer, W. et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).
- Borsos, M. et al. Genome–lamina interactions are established de novo in the early mouse embryo. *Nature* **569**, 729–733 (2019).
- Laing, A. F., Lowell, S. & Brickman, J. M. Gro/TLE enables embryonic stem cell differentiation by repressing pluripotent gene expression. *Dev. Biol.* **397**, 56–66 (2015).
- Galupa, R. & Heard, E. X-chromosome inactivation: a crossroads between chromosome architecture and gene regulation. *Annu. Rev. Genet.* **52**, 535–566 (2018).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

### Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

### Mouse embryo collection, single-cell dissociation and formaldehyde fixation

Five-week-old female C57BL/6J mice were purchased from Charles River. Animal care and use for this study were performed in accordance with the recommendations of the European community (2010/63/UE). All experimental protocols were approved by the ethics committee of the Institut Curie CEEA-IC118 under the number APAFIS#8812-2017020611033784v2, given by national authority in compliance with the international guidelines. When stated, intraperitoneal injection of 5 IU pregnant mare's serum gonadotropin, followed 46 h later by injection of 5 IU human gonadotropin, were applied to induce ovulation of female mice. DNA FISH was performed on embryos collected from superovulated C57BL/6J (B6) female mice (except for the blastocyst stage), mated with C57BL/6J (B6) male mice. The single-cell HiC protocol was applied to blastomeres of embryos collected from crosses between C57BL/6J (B6) female mice and CAST/Eij male mice. In the case of the one-cell, two-cell and four-cell stages, some embryos were collected after female superovulation. Embryos were collected from the reproductive tracts in M2 medium at defined time periods according to mating and/or hCG administration (given in this order): 14 h or 21 h for 1-cell stage (pronuclear stage 3 or 4), 37 h or 44 h for late 2-cell stage, 48 h or 55 h for 4-cell stage, 55 h or 62 h for 8-cell stage and 80 h for blastocyst stages (approximately 60 to 64 cells) (64-cell stage). B6 pure oocytes were collected 15 h after hCG injection. Embryos were included in the analyses when they showed a normal morphology and the correct number of blastomeres for their developmental stage. Zona pellucida and polar bodies were removed using acid Tyrode's solution and/or gentle pipetting (except in a few cases for the blastocyst stage). Embryos were incubated in  $\text{Ca}^{2+}$ - and  $\text{Mg}^{2+}$ -free M2 medium for 5 to 30 min to remove the polar body in zygotes or to isolate individual cells at subsequent stages. For the blastocyst stage, incubation with  $\text{Ca}^{2+}$ - and  $\text{Mg}^{2+}$ -free M2 medium was replaced with a 5-min incubation in TrypLE (Invitrogen). During the picking, the origin of the blastomere (inner cell mass or trophoctoderm) was not recorded). Blastomeres were mechanically dissociated, rinsed three times in PBS/acylated BSA (Sigma) before being fixed for 10 min in a 2% formaldehyde solution at room temperature. Fixation was stopped by transferring cells to a 127-mM glycine solution (5 min on ice). Blastomeres from different embryos were pooled from this step onwards to perform the single-cell HiC procedure post-fixation.

### Single-cell HiC procedure

The procedure for embryo blastomeres was optimized from a previous study<sup>13</sup>. Care was taken at every step to reduce putative contamination between solutions. In brief, following fixation, and rapid rinses in 1× PBS solution 1% acetylated BSA (Sigma), blastomeres were permeabilized for 30 min on ice in 10 mM Tris-HCl (pH 8), 10 mM NaCl, 0.2% IGEPAL CA-630 containing complete EDTA-free protease inhibitor cocktail (Roche). Cells were transferred to a protein low binding tube (Sigma) containing 0.3% SDS diluted with 1.24× NEBuffer3 for 60 min at 37 °C with constant agitation. Triton X-100 was added to 2% final and incubation was extended for 60 min, before addition of 625 U of Mbol (New England Biolabs) and overnight incubation. To label the digested DNA ends, a mix containing 28.4 μM final of dCTP, dGTP and dTTP and biotin-14 dATP were added with 25 U DNA polymerase I, large (Klenow) fragment (New England Biolabs) for 60 min with constant agitation. After spinning, blastomeres were treated with 10 U of T4 DNA ligase (Thermo Fisher) in presence of 1× reaction buffer with 1× BSA (both by

New England Biolabs) at 16 °C for at least 4 h. After spinning, blastomeres were resuspended with PBS 1× and BSA 1 mg/ml to dispatch them individually into PCR tubes (in strips; one per tube) before storage at -80 °C until further processing.

### Library preparation and sequencing

To prepare single-cell HiC libraries from single nuclei in PCR strips, 5 μl of PBS was added to each well and crosslinks reversed by incubating at 65 °C overnight. HiC concatemer DNA was fragmented and linked with sequencing adapters using the Nextera XT DNA library preparation kit (Illumina), by adding 10 μl of Tagment DNA buffer and 5 μl of Amplicon Tagment mix, incubating at 55 °C for 20 min, then cooling down to 10 °C, followed by addition 5 μl of Neutralize Tagment buffer and incubation for 5 min at room temperature. HiC ligation junctions were then captured by Dynabeads M-280 streptavidin beads (Thermo Fisher) (20 μl of original suspension per single-cell sample). Beads were prepared by washing with 1× BW buffer (5mM Tris-Cl pH 7.5, 0.5 mM EDTA, 1 M NaCl), resuspended in 4× BW buffer (20 mM Tris-Cl pH 7.5, 2 mM EDTA, 4 M NaCl; 8 μl per sample), and then mixed with the 25-μl sample and incubated at room temperature overnight with gentle agitation. The beads were then washed 4 times with 200 μl of 1× BW buffer, twice with 200 μl of 10 mM Tris-Cl pH 7.5 at room temperature, and resuspended in 25 μl of 10 mM Tris-Cl pH 7.5. Single-cell HiC libraries were amplified from the beads by adding 15 μl of Nextera PCR master mix, 5 μl of i7 Index primer of choice and 5 μl of i5 Index primer of choice. Samples were then incubated at 72 °C for 3 min, 95 °C for 30 s followed by the thermal cycling at 95 °C for 10 s, 55 °C for 30 s and 72 °C for 30 s for 18 cycles, and then incubated at 72 °C for 5 min. The supernatant was separated from the beads and purified one by one with AMPure XP beads (Beckman Coulter; 0.6 times volume of the supernatant) according to manufacturer's instructions and eluted with 30 μl each of 10 mM Tris-Cl pH 8.5. The eluate was purified once more with AMPure XP beads (equal volume to the previous eluate) and eluted with 11 μl of 10 mM Tris-Cl pH 8.5.

Before sequencing, the libraries were quantified by quantitative PCR (Kapa Biosystems) and the size distribution was assessed with Agilent 2100 Bioanalyzer (Agilent Technologies). The libraries were sequenced by 2×150-bp paired-end run using either a HiSeq 1500, HiSeq 2500 or NextSeq 500 (Illumina).

### Bioinformatics analysis

All data were mapped to the mouse genome mm10, using the C57BL/6J/CAST-Eij single nucleotide polymorphisms (SNPs) from the mouse genome project (v.5 SNP142), and the gene annotation from ensembl (v.92). Analyses were performed in R (v.3.4.2) and Bioconductor (v.3.6). Gene ontology was performed using the package ClusterProfiler (v.3.10.1).

### HiC data processing

Data were processed with HiC-Pro<sup>33</sup> (v.2.11.0) in allele-specific mode. The following parameters were used: - For mapping: -very-sensitive -L 30 -score-min L,-0.6,-0.2 -end-to-end -reorder. No minimal fragment size, insert size or contact distance were defined. - For processing: GET\_ALL\_INTERACTION\_CLASSES=0 GET\_PROCESS\_SAM=0 RM\_SINGLETON=1 RM\_MULTI=1 RM\_DUP=1. - for iced scaling: MAX\_ITER=100 FILTER\_LOW\_COUNT\_PERC = 0.02 FILTER\_HIGH\_COUNT\_PERC = 0 EPS = 0.1. Only pairs with both reads having MAPQ > 30 were kept.

### Cell cycle phasing

Cell cycle phasing was done by plotting the proportion of short-range contacts (between 25 kb and 2 Mb) versus long-range contacts (between 2 Mb and 12 Mb) in single cells. An ellipsoid was fitted to the single-cell points, as in a previous publication<sup>34</sup>. The reference in polar coordinates was set to the segment going from the centre of the ellipsoid to the point of coordinates [0.15, 0.35], which corresponds to the beginning

# Article

of the left-ascending part of the single-cell trajectory. Cells in the G1, S and G2 phases were defined as those in the angle between 0 and  $-0.35\pi$  ( $65^\circ$  anticlockwise). For each stage, contacts from all cells phased in G and S were pooled (all contacts or genome-specific contacts independently) and matrices at 10- and 40-kb resolution were created using cooler (v.0.7.9, parameter: `-balance`). Data were visualized in HiGlass<sup>35</sup>.

## Domain calling

Domains were first identified on the 40-kb matrices, independently for each stage, on both the maternal and paternal genomes, using 3dNetMod<sup>36</sup> (v.1.0.10.06.17), with parameters favouring sensitivity over specificity: `-PRE_PROCESSING: region_size 150, overlap 100, logged True, qnorm False`. `-GPS: badregionfilter True, scale genome wide, plateau 8, chaos filter True, chaos_pct 0.85, diagonal_density 0.65, consecutive_diagonal_zero 20`. `-MMCP: num_part 20, plots False, pctile_threshold 0, pct_value 0`. `-HSVM: size_threshold 7, size_s1 600000, size_s2 1200000, size_s3 3000000, size_s4 6000000, size_s5 12000000, var_thresh1 0, var_thresh2 50, var_thresh3 100, var_thresh4 100, var_thresh5 100, boundary_buffer 80000`. For the analysis of the X chromosome in female cells, domains were called from the female pseudo-bulk HiC maps.

## Domain average enrichment

We converted the HiC matrices to Z-score matrices, in which the scores are normalized to the distribution of scores for the same contact distance, as in a previous publication<sup>37</sup>. In brief, for any two loci  $i$  and  $j$  on chromosome  $c$ , separated by a distance  $n$  and with a balanced count of contacts  $C_{i,j}$ , the corresponding Z-score is  $Z_{i,j} = (C_{i,j} - \mu_n) / \sigma_n$ , in which  $\mu_n$  and  $\sigma_n$  are the mean and standard deviation of the distribution of contact counts for any pair of loci distant by  $n$ . Z-score matrices were calculated on the 40-kb matrices with HicExplorer<sup>37</sup> (v.2.1.1) using the HicFindTads function (parameters: `-correctForMultipleTesting None -minDepth 60000000 -maxDepth 200000000 -step 60000000 -thresholdComparisons 1 -delta 0`). For analysis of the X chromosome, contacts from female cells only were pooled and matrices obtained in the same way.

The average contact enrichment of domains was computed by averaging the Z-score over the domain upper triangle, excluding the diagonal. For a domain  $D_{i,j}$  spanning bins  $i$  to  $j$ , the upper triangle  $T$  in matrix  $M$  is the submatrix  $T[a,b]$  with  $a \in \{i, \dots, j-1\}$  and  $b \in \{i+1, \dots, j\}$  and  $b > a$ . This was calculated using the custom function `hicSummarizePerRegion` for HicExplorer, available from the E.H. laboratory GitHub version of HicExplorer at <https://github.com/heard-lab/HicExplorer>, branch `SummarizePerRegion`, or directly from <https://github.com/heard-lab/HicExplorer/blob/SummarizePerRegion/hicexplorer/hicSummarizeScorePerRegion.py>. We kept only domains with an average Z-score  $> 0.5$ .

## Overlapping domain filtering

As largely overlapping domains with very similar boundaries can be called within or between different time points, we further filtered redundancy using a custom script (available on GitHub, from [https://github.com/heard-lab/HicTools/blob/master/FilterRegions\\_MinMutualOverlap\\_maxScore.r](https://github.com/heard-lab/HicTools/blob/master/FilterRegions_MinMutualOverlap_maxScore.r)). In brief, starting from a set of domains  $D_n = 0$  equal to the set of all domains  $D_{\text{all}}$ , and the empty sets  $D_{\text{overlap}}$  and  $D_{\text{highest}}$ , the following steps were used: (1) From  $D_n$ , all pairs of overlapping domains are compared two by two. (2) If their overlap represents more than 70% of each other's lengths, they are added to the set  $D_{\text{overlap}}$ . (3) For each pair of overlapping domains ( $>70\%$ ), only the domain with the highest score is kept and added to the set  $D_{\text{highest}}$ . (4)  $D_{n+1}$  is assigned the union of  $D_{\text{highest}}$  and all domains from  $D_{\text{all}}$  that were not in  $D_{\text{overlap}}$ . The procedure is repeated from step 1 to step 4 until  $D_{n+1} = D_n$ . The reinjection in step (4) of all domains from  $D_{\text{all}}$  that were not in  $D_{\text{overlap}}$  allows us to keep isolated domains, as well as avoiding chains between pairs of domains. For stage-specific analysis (Figs. 1c, 4a) this procedure was applied to the domains called at each stage and on each genome

individually. For the dynamic analysis across stages, sets of all domains called individually at each stage and on each genome (after this redundancy filtering) were pooled together as one set and filtered with the same procedure, resulting in one common set of domains.

## Clustering

Domain dynamics clustering was performed using the R package Mfuzz (2.26.0)<sup>38</sup>, using as input the average Z-score per domain (row) in each stage from the 1-cell stage to the 64-cell stage, on the maternal and paternal genomes (columns). Fuzzification parameter  $m$  was estimated using the `mestimate()` function. The number of clusters was defined as nine, on the basis of the minimal distance between cluster centroids.

## Single-cell analysis

The sum of contacts per domain for each genome per single cell was computed using the function `hicSummarizePerRegion` (as described in 'Domain average enrichment'), excluding the diagonal. The matrix of counts (domains on rows, single-cell maternal genome and single-cell paternal genome on columns) was used as input for monocle3<sup>39</sup>. Data were processed using the `preprocess_cds` function using the first 75 components of the principal components analysis (parameters: `num_dim = 75, method = "PCA", norm_method = "log"`). Dimension reduction was performed using UMAP with the `reduce_dimension` function (`max_components = 2`) and graph for pseudotime inferred using `learn_graph` (parameters: `use_partition = FALSE, learn_graph_control = list(minimal_branch_len = 3)`). For cluster average score, counts per domains were converted to CPKM by dividing the counts by the total number of contacts in domains per allele (divided by  $10^6$ ), and by the domain length in kb.

## Compartments and domain interactions

Compartments were called using HiTC (v.1.26.0)<sup>40</sup>. An aggregate plot of interaction between pairs of domains was performed using a custom function `hicAggregateContact` for HicExplorer (available on GitHub, from <https://github.com/deeptools/HicExplorer>, branch `aggregateGenome`; parameters: `-range 1000000:999000000 -numberOfBins 200 -avgType mean -genome -regionReferencePosotion centre`), which also output the list of pairs of domains with respect to the distance threshold (that is, distance of more than 1 Mb). Only domains that did not contain another domain were used to avoid redundancy between domains that contained one another. The normalized contact counts of the intersection between pairs of domains was calculated using a custom function `hicSsummarizeScorePerRegion` for HicExplorer (available on GitHub, from <https://github.com/heard-lab/HicExplorer>, branch `SummarizePerRegion`, or directly from <https://github.com/heard-lab/HicExplorer/blob/SummarizePerRegion/hicexplorer/hicSummarizeScorePerRegion.py>; parameter: `-summarizeType sum`).

## Chromosome 3D modelling

Three-dimensional models of chromosomes (allele-specific) was performed using the programs Dip-C and HicKit<sup>41</sup>. We performed 3 rounds of 3D reconstruction at 100-kb resolution with 3D haplotype imputation (parameters: `-temps 20 -s 8 4 2 0.4 0.2 0.1`), and then 2 rounds of 3D reconstruction at 20-kb resolution with 3D haplotype imputation (parameter `-temps 20 -s 8 4 2 0.4 0.2 0.1 0.04 0.02`). Chromosome volumes were calculated using the alpha-convex hull algorithms from the R package `alphashape3d` ( $\alpha = 0.6$ ).

## ChIP-seq analysis

Reads were trimmed using TrimGalore (v.0.4.4), mapped using STAR<sup>42</sup> (2.5.3a, parameters: `-outFilterMultimapNmax 1 -outFilterMismatchNmax 999 -outFilterMismatchNoverLmax 0.06 -alignIntronMax 1 -alignMatesGapMax 2000 -alignEndsType EndToEnd -outSAMattributes NH HI NM MD`), and removed when they mapped to the mitochondrial genome. The remaining reads were split by allele using `SNPsplit`



(v.0.3.2). Allele-specific and unassigned .bam files were sorted, duplicates removed using Picard (v.2.18.2, parameters: REMOVE\_DUPLICATES = true ASSUME\_SORTED = true) and pooled as the total reads. BigWig of coverage files were done using DeepTools<sup>43</sup> bamCoverage (parameters: -extendReads -binSize 1, with -extendReads 200 for single-end data). A scaling factor was calculated as  $10^6/\text{total number of reads}$ , and the same factor was given as the parameter '-scaleFactor' for both allelic signals. The heat map and average plots of signal were performed using DeepTools computeMatrix scale-regions (with parameters: -regionBodyLength 1000000 -beforeRegionStartLength 1000000 -afterRegionStartLength 1000000 -binSize 50000) as well as plotHeatmap and plotProfile. For quantification of ChIP-seq in domains, reads were counted using the featureCounts function from Subread<sup>44</sup> (v.1.28.1, parameters: -p -s 0). Data scaling was performed in R using DESeq2 (v.1.18.1), calculating the sizeFactor on the count of total reads and applying it to the allele-specific counts. Enrichment relative to background was calculated as the ChIP-seq signal per domain in RPKM, divided by the average RPKM on the genome calculated in 10-kb bins.

### RNA-sequencing analysis

Single-cell RNA-sequencing data were processed similarly to those from ChIP-seq, except for the mapping, for which the following parameters were used: -outFilterMultimapNmax 1 -outFilterMismatchNmax 999 -outFilterMismatchNoverLmax 0.06 -alignIntronMax 500000 -alignMatesGapMax 500000 -alignEndsType EndToEnd -outSAMattributes NH HI NM MD. The quantification of expression was performed using featureCounts (parameters: -p -s 0 -t exon -g gene\_id). Data were then analysed in R using DESeq2<sup>45</sup> (v.1.18.1), calculating the sizeFactor on the count of total reads and applying it to the allele-specific counts. Filtering was performed similarly that in a previous publication<sup>20</sup>. In single-cell data, a pseudo-RPKM score was calculated as the normalized count  $\times 1,000/\text{gene length in base pairs}$ ; as the previously used protocol<sup>20</sup> is 3'-biased and does not recover more than the last 3 kb of the transcripts (longer genes (>3 kb) were assigned a length of 3 kb). In single-cell data, genes with a pseudo-RPKM value < 5 (not allele-specific) and a count of reads lower than 10 reads on both alleles were assigned as lowly expressed. An allelic *D*-score ( $\text{expression}_{\text{maternal}}/(\text{expression}_{\text{maternal}} + \text{expression}_{\text{paternal}})$ ) was calculated only for genes that were not lowly expressed, to avoid artefactual strong bias due to noisy low-expressed genes. Single-cell data were then pooled in pseudo-bulk by stage, and for each gene an average *D*-score was calculated only when more than 20% of single cells had an allelic *D*-score calculated (that is, did not show too low expression on both alleles). Average pseudo-RPKM values were calculated by averaging the pseudo-RPKM values of all single cells without filtering.

### DNA FISH probes

Probes for DNA FISH on the X chromosome were obtained as previously described<sup>22</sup>, or using BAC DNA for chromosome 13 (purchased from CHORI RP24-278M23; RP23-325G4; RP23-2B17; RP23-222A16; RP24-389D15; RP23-302B3; RP23-359G6; RP23-326J5; RP23-307F19) or were purchased from MYcroarray (fluorescent oligonucleotides, average length 45 bp, 5'-modified with Atto 448 or Atto 550, average density: one oligonucleotide every 3 kb). Oligonucleotides were designed to tile the following consecutive 18-Mb regions: chromosome X: 35,000,000–53,000,000 (termed pool a) and chromosome X: 53,000,000–72,000,000 (termed pool b)<sup>26</sup>. To prepare the probe mix for DNA FISH, 100 ng of labelled BAC DNA was used, along with 5  $\mu\text{g}$  of Cot-1 DNA and resuspended in formamide before adding equal volume of hybridization buffer (2 $\times$ , 20% dextran sulfate; 4 $\times$  SSC; 1 mM EDTA; 0.1% TritonX-100; 0.5 mg/ml BSA; 1 mg/ml PVP). Oligonucleotide probes were used in formamide at 10% final concentration

### DNA FISH procedure on embryonic stem cells

FISH on cells from tissue culture was performed as previously described<sup>22,46</sup>. Feeder-free male mouse embryonic stem cells (E14;

GSM1366337) were cultured on gelatin-coated coverslips no. 1.5 (1 mm) and fixed in 3% paraformaldehyde for 10 min at room temperature. Permeabilization was then performed on ice for 5 min in 1 $\times$  PBS containing 0.5% Triton X-100 and 2 mM vanadyl-ribonucleoside complex (New England Biolabs). Coverslips were preserved in 70% EtOH at -20 °C. Prior to FISH, samples were dehydrated through an ethanol series (80%, 95% and 100%, twice) and air-dried quickly. DNA FISH was preceded by sample denaturation in 50% formamide in 2 $\times$  SSC at pH 7.2 at 80 °C for 40 min. After overnight hybridization at 42 °C, washes were carried out at 45 °C, 3 times 5 min in 50% formamide in 2 $\times$  SSC at pH 7.2 and 3 times 5 min in 2 $\times$  SSC. DAPI at 0.2 mg/ml was used for counterstaining and mounting medium consisting of 90% glycerol, 0.1 $\times$  PBS, 0.1% *p*-phenylenediamine at pH 9 (Sigma).

### Three-dimensional DNA FISH procedure on embryos and *Xist* RNA FISH combined with DNA FISH using oligonucleotide probes

Collected embryos were prefixed for 1 min at room temperature in paraformaldehyde (PFA) 1% 1 mg/ml polyvinylpyrrolidone (PVP), pre-permeabilized for 1 min at room temperature in PFA 0.5% and TritonX-100 0.4% and fixed for 10 min at room temperature in PFA 4%. After a brief wash in PBS 1 $\times$  with PVP 1 mg/ml and TritonX-100 0.05% (PBS-TP), embryos were permeabilized for 1 h at 37 °C in PBS 1 $\times$  with TritonX-100 0.5% (with RNase A 5  $\mu\text{g}/\text{ml}$  in case of DNA FISH). After a brief rinse in PBS-TP, embryos were transferred into hybridization buffer 1 $\times$  and equilibrated overnight with 1 mg/ml Cot-1 DNA mix at 37 °C. Embryos and probes were denatured for 10 min at 83 °C and put back for at least 3 h at 37 °C. After competition in Cot-1 mix, embryos were moved into the probe mix overnight at 37 °C. Excess of probes was eliminated through 3 washes at 45 °C in SSC 2 $\times$  solution and SSC 0.2 $\times$  solution for 10 min each. Embryos were then briefly washed in PBS 1 $\times$  and mounted in a Vectashield drop containing DAPI under oil on a glass-bottomed plate, coated with poly-lysine.

### Microscopy and image analysis

Combined RNA and DNA FISH imaging was performed on an inverted confocal microscope (Zeiss) LSM700 with a Plan apo DICII (numerical aperture 1.4) 63 $\times$  oil objective. *Z*-sections were taken every 0.4  $\mu\text{m}$ . Structured illumination for DNA FISH was performed using an OMX system (Applied Precision) as in a previous publication<sup>22</sup>. Signals from all channels were realigned using fluorescent beads before each session of image acquisition. For colocalization analysis, analysis was restricted to a region of interest of identical volume around the FISH signal. The respective intensities of red and green channels were retrieved semi-automatically using the JACoP ImageJ plugin, and box plot distribution of the Pearson correlation coefficient was compared using Wilcoxon's rank-sum statistics with R.

### Engineering mice

The mouse mutant lines were generated following a previously described strategy<sup>47</sup>, with minor modifications. Single-guide (sg)RNAs were designed using CRISPOR<sup>48</sup>. For deleting the locus containing *Jpx* and *Ftx*, we used sgRNAs no. 57 (GGTACAATTATGCAACCTG), no. 58 (ATACTCCGATTACATACTC), no. 61 (TGCCCAAGCAAAAAGCGTGA) and no. 62 (AAAGTATTGACACCTTACCC). For deleting the *Jpx* locus, we used sgRNAs no. 57, no. 58 and no. 59 (TGCCCAAGCAAAAAGCGTGA) and no. 60 (AGTTAGATACCACCAAGT). T7-sgRNA PCR products were used as the template for in vitro transcription with the MEGAshortscript T7 kit (Life Technologies) and the products were purified using the MEGAclear kit (Life Technologies). sgRNAs were eluted in DEPC-treated RNase-free water, and their quality was assessed by electrophoresis on an agarose gel after incubation at 95 °C for 3 min with denaturing agent provided with the in vitro transcription kits. *Cas9* mRNA (Tebu-bio, L-7206) and sgRNAs were injected at 100 ng/ $\mu\text{l}$  and 50 ng/ $\mu\text{l}$ , respectively, into the cytoplasm of mouse B6D2F1 zygotes from eight-week-old superovulated B6D2F1 (C57BL/6J  $\times$  DBA2) female mice mated to stud male mice of the same background. Zygotes with well-recognized pronuclei were collected in M2 medium (Sigma) at E0.5. Injected embryos were cultured in M16

# Article

medium (Sigma) at 37 °C under 5% CO<sub>2</sub>, until transfer at the one-cell stage the same day or at the two-cell stage the following day to the infundibulum of the oviduct of a pseudogestant CD1 female at E0.5 (25–30 embryos were transferred per female). All weaned mice (N0) were genotyped for presence of deletion (locus covering *Jpx* and *Ftx*, primers RG140.1: TGC-TACCGGTCACAGATATAAGT and RG145: TCTGGGATGCTTGTTCAACA; *Jpx* locus, primers RG140.1 and RG143: ACAAGGTGAGCGATGAGACA). Mice carrying deletion alleles were crossed to B6D2F1 mice and their progeny screened again for the presence of the deletion allele; PCR products were sequenced to determine the exact location of the deletions (locus covering *Jpx* and *Ftx*, chromosome X: 100,683,288–100,801,657, mm9; *Jpx*: 100683306–100702361, mm9). The F<sub>1</sub> mice were considered the ‘founders’ and bred to B6D2F1 mice; their progeny was then backcrossed to B6D2F1 mice, to generate heterozygous mice and lines were kept in heterozygosity. To establish mouse embryonic fibroblasts, single embryos were recovered at day 13.5 of gestation after the confirmation of vaginal plugs on  $\Delta Jpx$ /wild-type females bred with wild-type/Y or  $\Delta Jpx$ /Y males. Head and internal organs were removed and the body cavity was incubated for 1 h at 37 °C in TripLE (Invitrogen). After repetitive pipetting up and down, the resulting chunks were put in culture for 24–48 h until collected to prepare RNA with Trizol extraction for further examination by quantitative PCR. The level of gene expression was normalized to the geometric mean of the expression level of *Ppia* and *Gapdh* housekeeping genes according to geNorm method<sup>49</sup> to assess the relative expression of *Xist* and *Jpx*. The following primers were used and are listed as forward reverse and in 5' to 3': *Gapdh*, ccccaactgagcatctcc/attatgggggtctgggatgg; *Ppia*, ttaccatcaaacctctctg/aacccaagaacttcagtgagagc; *Jpx*, ataaaatggcggtccac/ggccagtttctcactctcc; and *Xist*, gggtctctccagaa-gctagaa/tggtagatggcattgtgtattatgg

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

The HiC data generated and analysed are available in the GEO repository under accession number GSE129029. Previously published data were downloaded from GEO: H3K27me3 in early embryos (GSE76687); H3K27me3 in day-5 post-natal oocytes (GSE93941); single-cell RNA sequencing in early embryos (GSE80810); and HiC in gametes and early embryos (GSE82185). Source Data for Figs. 3, 4 and Extended Data Fig. 2, 6 are provided with the paper. Any other relevant data are available from the corresponding authors upon reasonable request.

## Code availability

The code developed for this study is available on the GitHub repository of the laboratory of E.H. (<https://github.com/heard-lab>).

33. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).
34. Pu, J., Zheng, B., Leader, J. K. & Gur, D. An ellipse-fitting based method for efficient registration of breast masses on two mammographic views. *Med. Phys.* **35**, 487–494 (2008).
35. Kerpedjiev, P. et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol.* **19**, 125 (2018).
36. Norton, H. K. et al. Detecting hierarchical genome folding with network modularity. *Nat. Methods* **15**, 119–122 (2018).
37. Ramirez, F. et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018).
38. Kumar, L. & E. Futschik, M. Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics* **2**, 5–7 (2007).
39. Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
40. Servant, N. et al. HiTC: exploration of high-throughput ‘C’ experiments. *Bioinformatics* **28**, 2843–2844 (2012).
41. Tan, L., Xing, D., Chang, C.-H., Li, H. & Xie, X. S. Three-dimensional genome structures of single diploid human cells. *Science* **361**, 924–928 (2018).
42. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
43. Ramirez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
44. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108 (2013).
45. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
46. Chaumeil, J., Augui, S., Chow, J. C. & Heard, E. in *The Nucleus* (ed. Hancock, R.) 297–308 (Springer, 2008).
47. Wang, H. et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
48. Haeussler, M. et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **17**, 148 (2016).
49. Vandesompele, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, RESEARCH0034 (2002).

**Acknowledgements** We thank the Institut Curie Animal facility for animal welfare and husbandry; the imaging facility PICTiBiSA@BDD for technical assistance; members of the E.H. laboratory for critical input; Y. Komarnitskaya for graphic design; and F. Ramirez and G. Richard for their help with bioinformatic analysis. This work was supported by FRM (FDM20140630223 and FDM40917) to N.R., ERC Advanced Grant DEVOCHROMO to P.F., by ERC Advanced Investigator Awards ERC-ADG-2014 671027 and Labellisation la Ligue, Labex DEEP: ANR-11-LBX-0044, IDEX PSL: ANR-10-IDEX-0001-02 PSL to E.H.

**Author contributions** N.R., T.N., P.F., K.A. and E.H. designed the experiments. N.R., T.N. and W.L. performed the single-cell HiC experiments. S.C., K.A. and N.S. designed and performed the single-cell HiC data analysis and integration. C.V. and T.S. produced the chromosome modelling data. N.R. and K.A. performed DNA FISH on preimplantation embryos. N.R. and T.P. performed structured illumination microscopy and image analysis. R.G. and K.A. engineered CRISPR deletions. M.B. set up single-cell dissociation and collected embryos with N.R. The manuscript was written by S.C., K.A. and E.H. with contributions from N.R., C.V. and P.F., and input from all authors.

**Competing interests** The authors declare no competing interests.

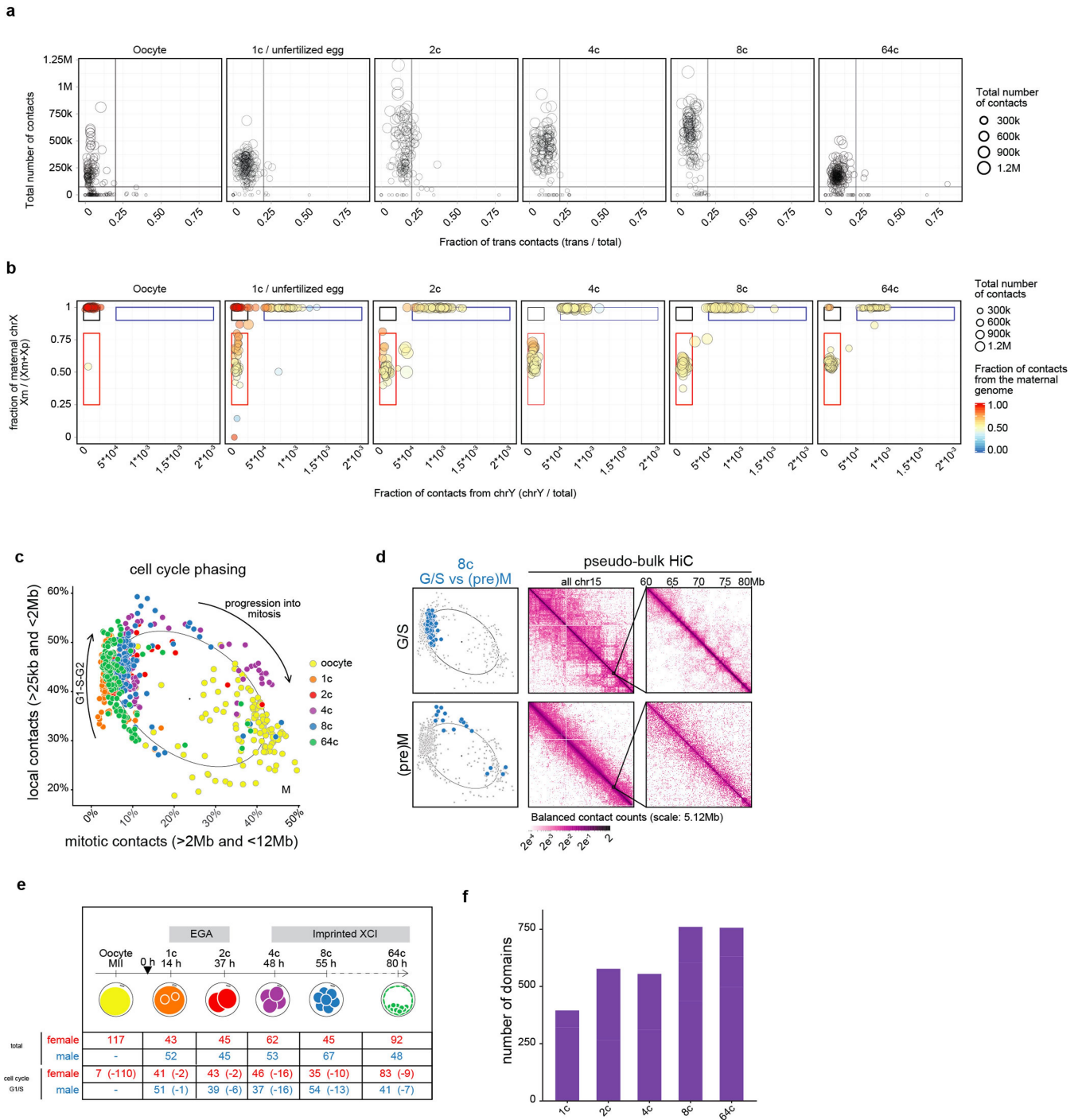
## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2125-z>.

**Correspondence and requests for materials** should be addressed to P.F., K.A. or E.H.

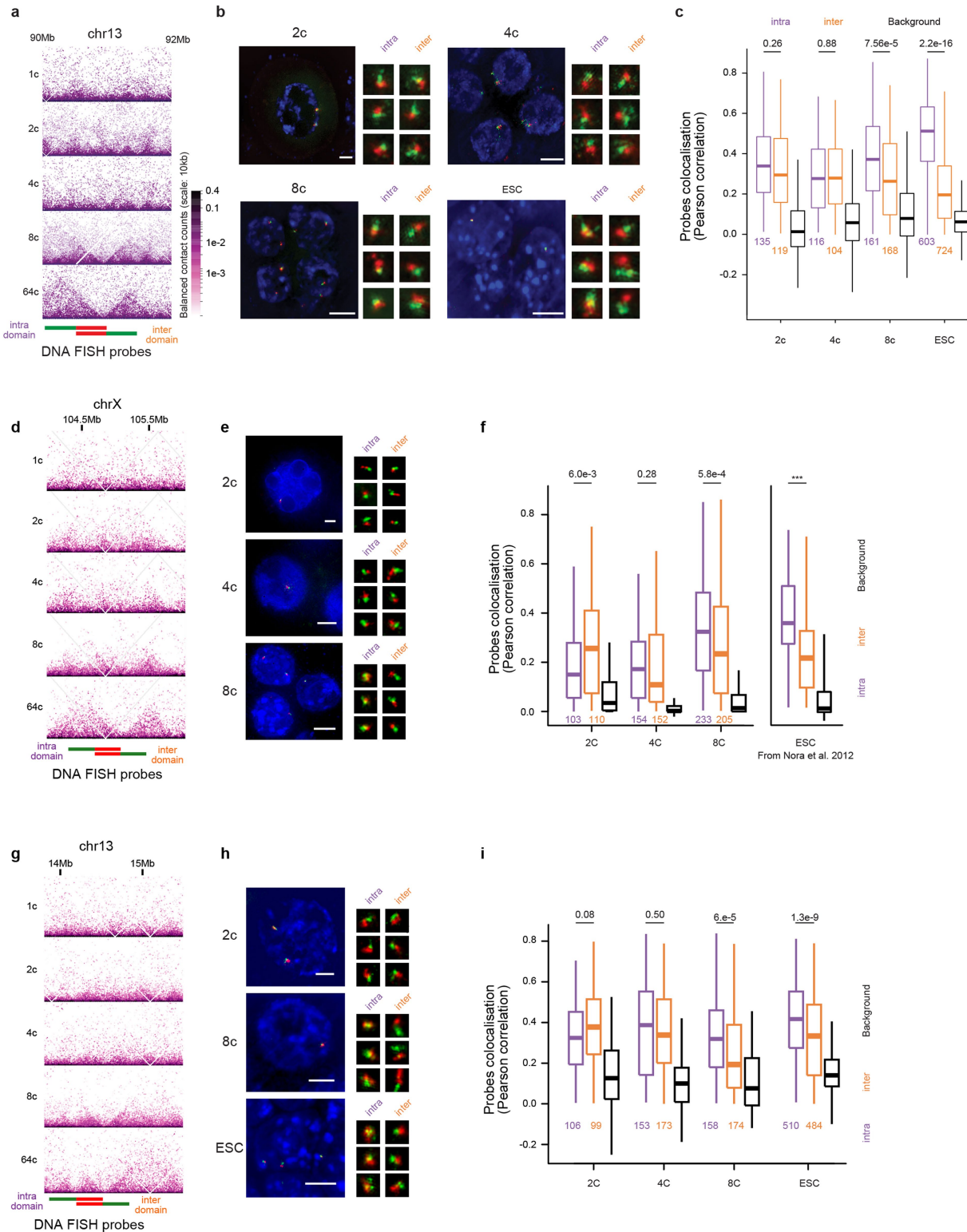
**Peer review information** Nature thanks Takashi Sato and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



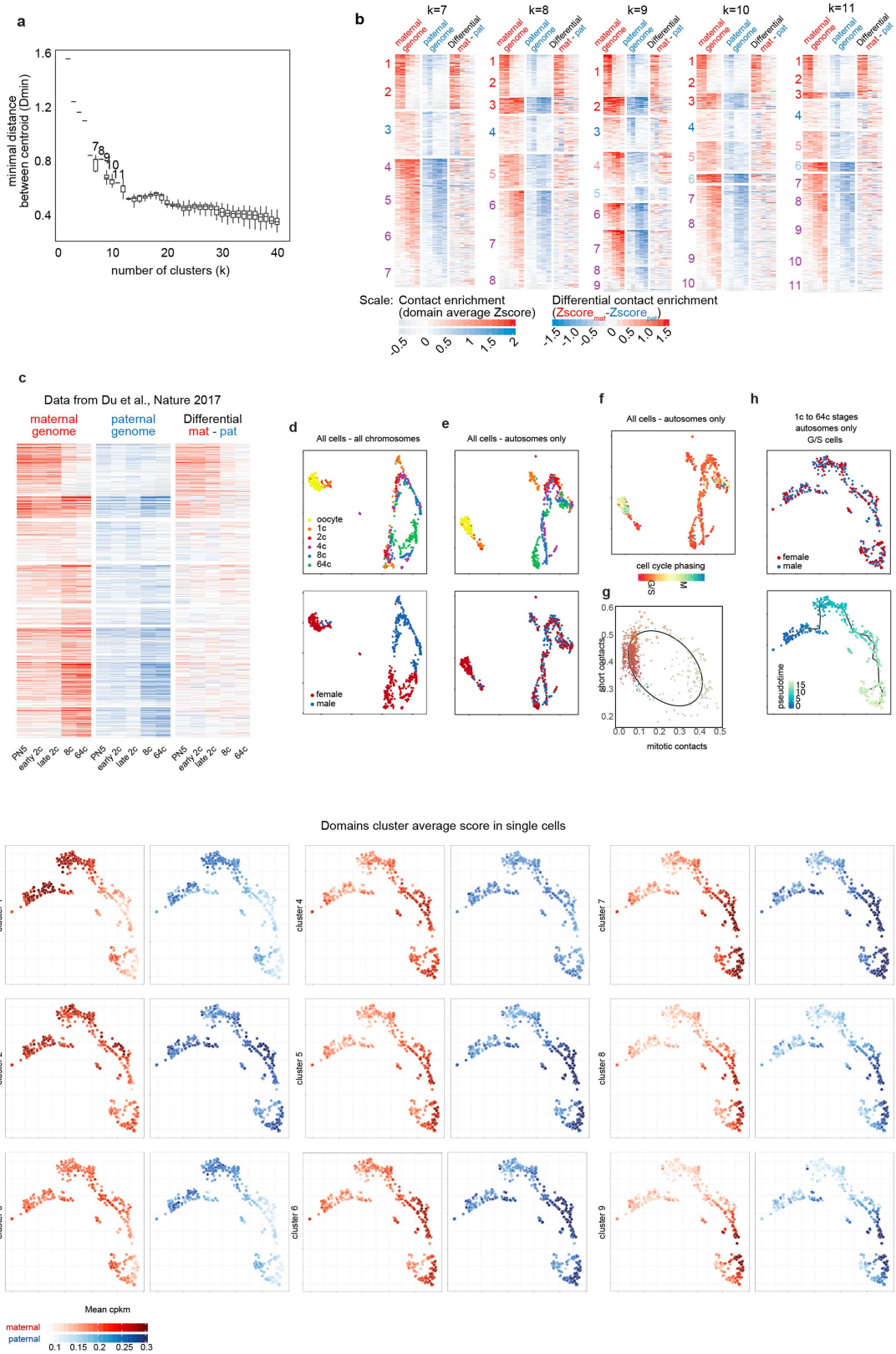
**Extended Data Fig. 1 | Single-cell HiC approach to studying chromosome organization in mouse preimplantation embryos.** **a**, Distribution of total contact versus *trans* ratio per single blastomere, according to developmental stage with given thresholds for exclusion. **b**, Fraction of maternal contacts on the X chromosome versus contacts on the Y chromosome. The colour of each dot indicates the fraction of reads that cover the maternal genome. Red rectangles highlight female diploid cells, blue rectangles highlight male cells and black rectangles highlight haploid cells (that is, oocytes or polar bodies). Cells outside these frames were excluded. **c**, Percentage of short-range

(25 kb–2 Mb) versus long-range or mitotic contacts (2–12 Mb) per single cell, coloured by developmental stages. **d**, Subset of the single cells at eight-cell stage, either in G1/S or G2 phase (top) or going towards mitosis (bottom), and their corresponding pseudo-bulk HiC heat maps. **e**, Table for the number of single blastomeres per stage of development that passed quality control, and the selected number after cell cycle phasing that were used to produce the subsequent analysis and heat maps. **f**, Bar plot of domain numbers for each developmental stage.



**Extended Data Fig. 2 | HiC view and DNA FISH for two independent genomic loci.** **a, d, g**, HiC contact maps for different genomic locations (as indicated), from the 1-cell to 64-cell stage. **b, c, e, f, h, i**, Analysis of the genomic locations for boundary formation (red and green probes in bottom of **a, d** and **g**) by 3D DNA FISH in two-cell-stage to eight-cell-stage embryos and embryonic stem cells (ESC), with insets of signal for the two independent pools (**b, e, h**). The total number of combined signals (red plus green) is reported in the box plot in the adjacent panels (**c, f, i**). DNA is stained by DAPI (blue). Scale bar, 10  $\mu$ m. **c,**

Box plot ( $\pm 1.5 \times$  interquartile range, 25th and 75th percentiles and median value) distribution of Pearson's correlation coefficient for red and green signals (in pools 1 and 2) of DNA FISH analysis. **a–c**, Chromosome 13 (region 90 Mb–92 Mb). **d–f**, X chromosome (region 104 Mb–105 Mb). **g–i**, Chromosome 13 (region 14 Mb–15 Mb). All experiments are performed in biological replicates,  $n$  is the combined signal number, centre lines denote the median coefficient. Statistical significance ( $P < 0.001$ ) was assessed using Wilcoxon's rank sum test (two-sided).

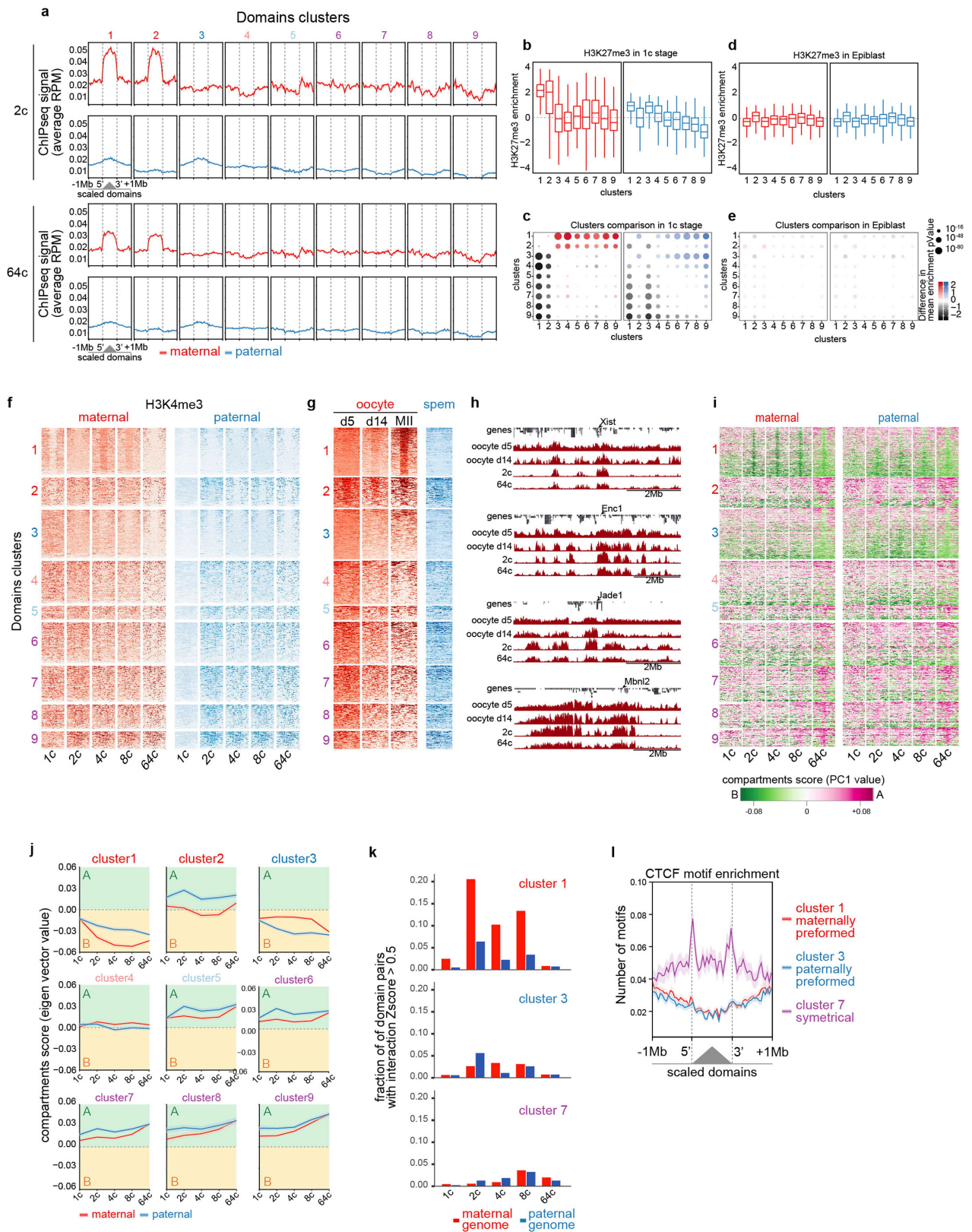


Extended Data Fig. 3 | See next page for caption.

# Article

**Extended Data Fig. 3 | Dynamics of domains in single cells.** **a**, Distribution of the minimal distance between cluster centroids ( $D_{\min}$ ) for a predefined number of clusters ( $k$ ) ranging from 2 to 40. Clustering was performed 100 times for each value of  $k$ . The optimal number of clusters is the highest value of  $k$  before the value  $D_{\min}$  becomes stagnant. **b**, Heat maps representing the result of clustering for different values of  $k$ . The same main categories are found for  $k > 8$ . The contact enrichment colour scale corresponds to the maternal (red) and paternal (blue) heat maps; the differential contact enrichment scale corresponds to the differential (maternal – paternal) heat maps. **c**, Heat maps showing domain enrichment in the bulk HiC data from GSE82185, with the same order as our clustering in Fig. 1d and showing similar dynamics. **d**, Single-cell

projection by UMAP from the quantification of domain contacts on each allele, using all cells and all chromosomes, coloured by stage (top) or by sex (bottom).  $n = 669$  single cells. **e**, As in **d** but excluding domains on the X chromosome. **f**, As in **e** but coloured by cell cycle phasing. **g**, Cell cycle phasing based on short-range versus mitotic contacts, with the same colour scale as in **f**. **h**, Single-cell projections after excluding oocytes, all cells in pre-M and M phase and domains on the X chromosome, as in Fig. 1f, coloured by sex (top) or by pseudotime overlaid with the inferred trajectory (bottom).  $n = 470$  single cells. **i**, As in **h**, coloured by mean count per kb per million (CPKM) on each allele, for the nine clusters identified in Fig. 1d.



Extended Data Fig. 4 | See next page for caption.

# Article

## Extended Data Fig. 4 | Chromatin changes and compartment formation over preimplantation.

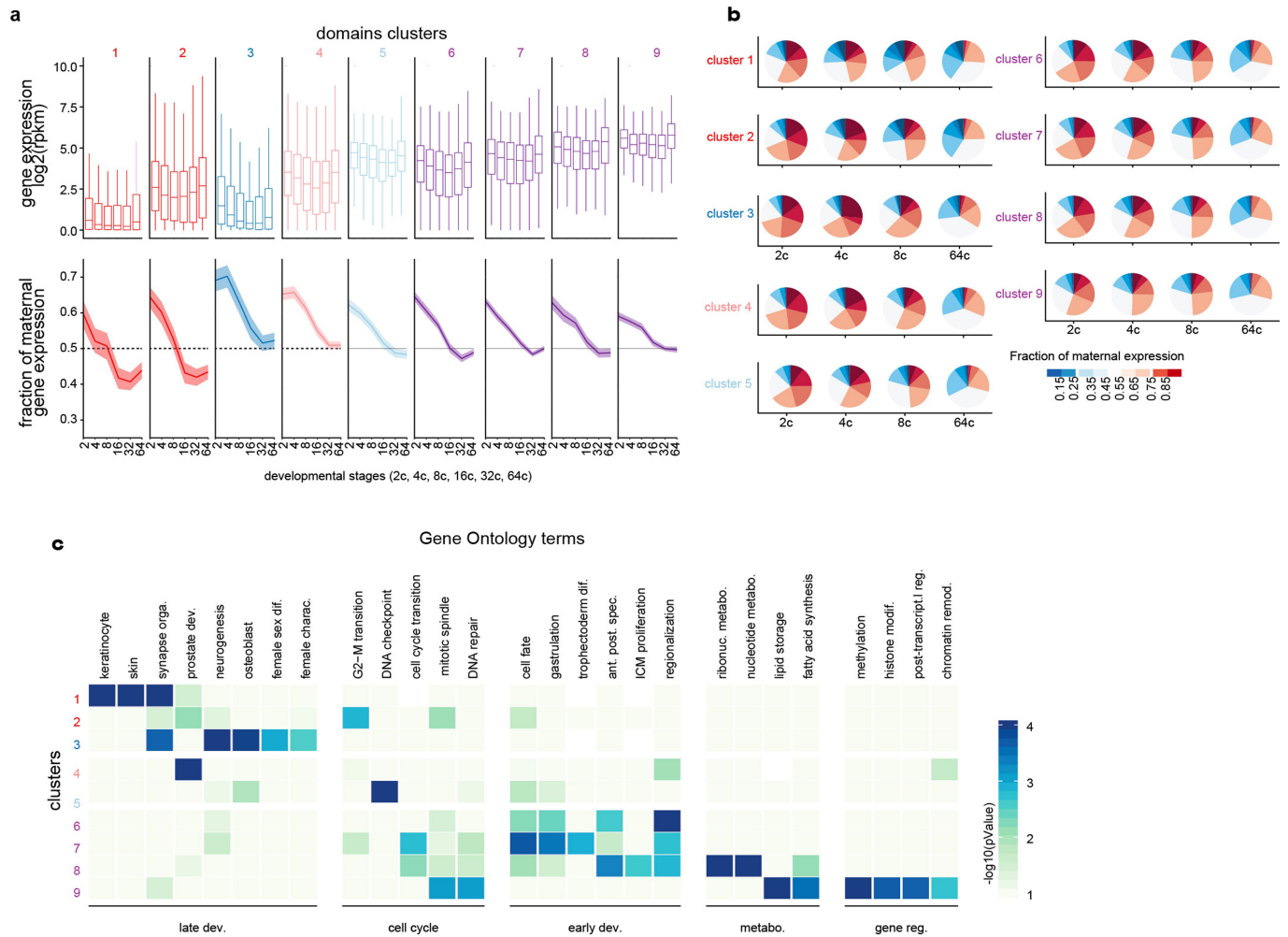
**a**, Average profile of H3K27me3 ChIP-seq signal at the domains for each parental allele at the 2-cell and 64-cell stages in clusters 1 to 9.  $n = 375, 238, 387, 338, 110, 327, 287, 194$  and  $141$  for each cluster from 1 to 9).

**b**, Distributions of H3K27me3 domain enrichment per cluster, on the maternal (red) and paternal (blue) genomes at the one-cell stage. Box plots represent  $\pm 1.5 \times$  interquartile range, 25th and 75th percentile and median value.  $n$  values are the same as in **a**. **c**, Statistical comparison, two-by-two, between each distribution in **b**.  $P$  values are calculated using a Wilcoxon test (two-sided, not paired).  $n$  values are the same as in **a**. **d, e**, As in **b, c** for H3K27me3 ChIP-seq data from epiblasts. **f**, Heat maps of H3K4me3 ChIP-seq signal at domains of each

cluster  $\pm 1$  Mb, with parental origin. **g**, Heat maps of H3K27me3 ChIP-seq signal at domains of each cluster  $\pm 1$  Mb in oocytes (post-natal day 5 or day 14; or ovulatory oocytes (MII)). **h**, Snapshots of H3K27me3 ChIP-seq signal covering 6 Mb at transiently imprinted loci (*Xist*, *Enc1*, *Jade1* and *Mbnl2*) for different stages of oogenesis, or the maternal allele in the 2-cell and 64-cell stages.

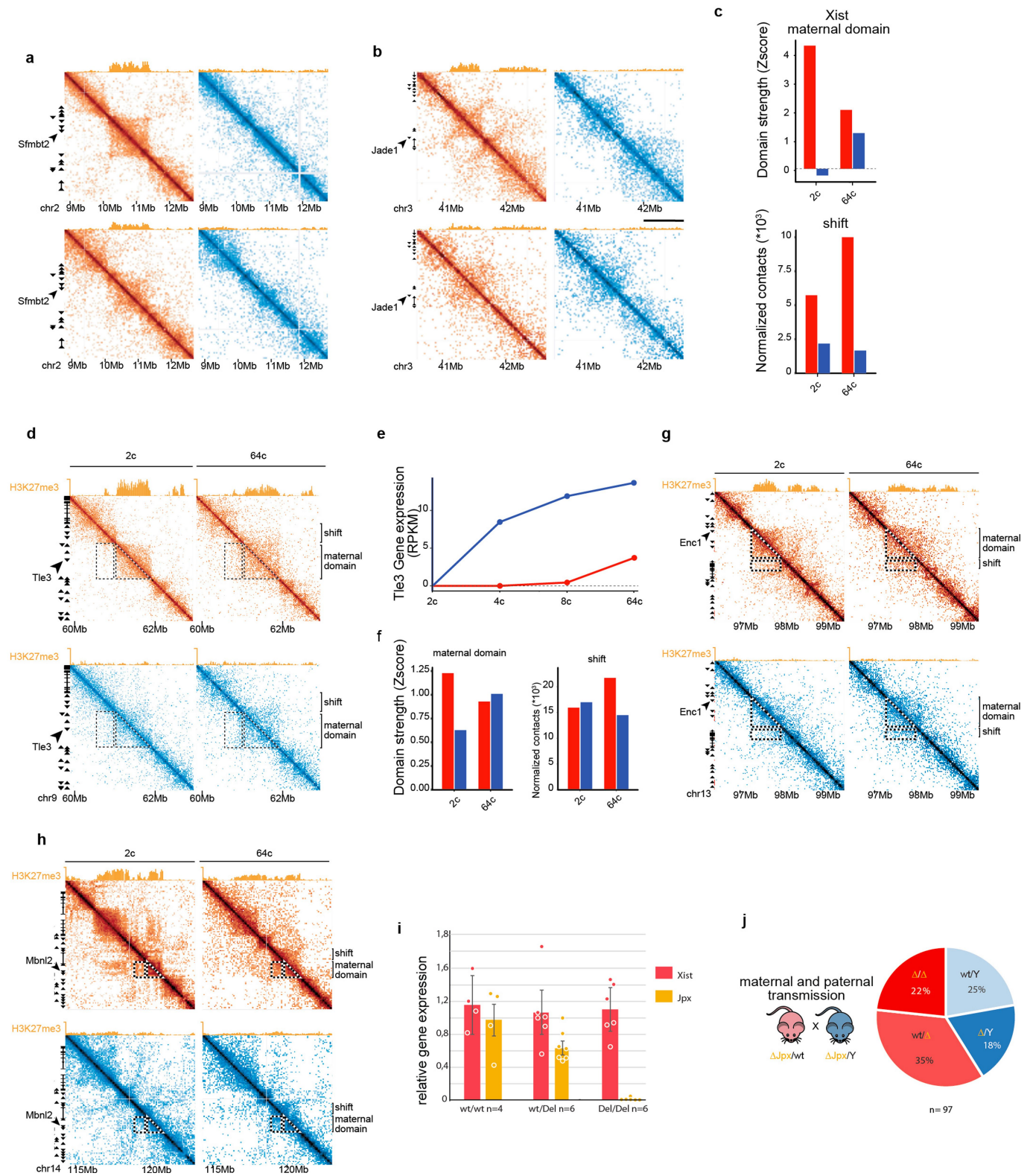
**i**, Compartment scores at domains of clusters 1–9, according to parental origin. **j**, Dynamics of the compartment scores for each cluster. Lines represent the mean, and shading represents the 95% confidence interval of the mean.  $n$  values are the same as in **a**. **k**, Bar plot of long-range interactions per stage, corresponding to the average heat map in Fig. 2f. **l**, CTCF-motif enrichment around domains.





**Extended Data Fig. 5 | Gene expression and functional annotation of domain clusters. a.** Distribution of gene expression (top;  $n = 797, 353, 612, 621, 268, 699, 562, 278$  and  $193$  genes for clusters 1 to 9) and fraction of maternal expression (maternal/(maternal + paternal)), bottom;  $n = 232, 249, 256, 502, 258, 664, 497, 179$  and  $269$  genes for which an allelic ratio could be calculated for clusters 1 to

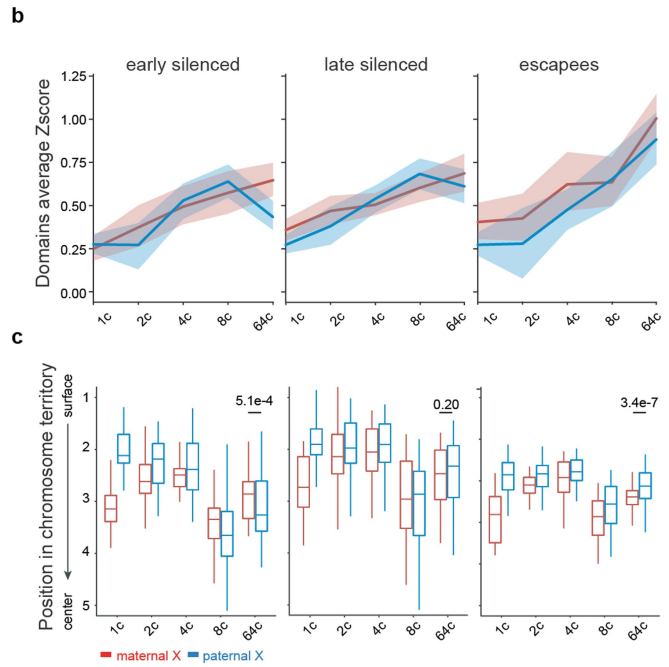
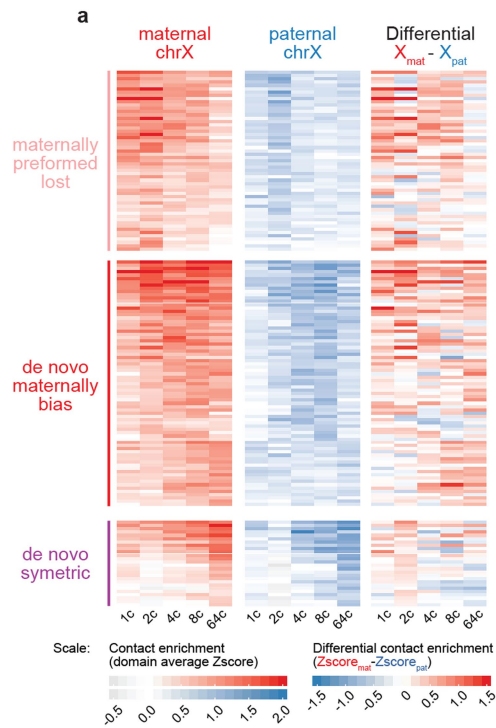
9, respectively) for genes present within domains of the different clusters. **b.** Pie charts for allelic expression bias from the 2-cell to the 64-cell stage for genes within clusters 1 to 9. **c.**  $P$  value (hypergeometric test) of Gene Ontology term enrichment in genes within each domain cluster.



Extended Data Fig. 6 | See next page for caption.

**Extended Data Fig. 6 | Structural tuning at maternal early domains during preimplantation.** **a**, Snapshots of HiC matrices and H3K27me3 ChIP-seq signal, showing the parental differences between the 2-cell and 64-cell stages for maternal (red) and paternal (blue) genomes at chromosome 2 (9–13.5 Mb) containing *Sfmbt2*. **b**, As in **a**, for chromosome 3 (40–43 Mb) containing *Jade1*. **c**, Quantification of contacts within the region presented in Fig. 3b. **d**, Snapshot of HiC matrices and H3K27me3 ChIP-seq signal, showing the parental differences between the 2-cell and 64-cell stages for maternal (red) and paternal (blue) genomes at chromosome 9 (60–62.5 Mb) containing *Tle3*. **e**, Gene-expression dynamic for *Tle3* for maternal (red) and paternal (blue) alleles. **f**, Quantification of contacts within the region shown in **d**. **g**, Snapshots of HiC matrices and H3K27me3 ChIP-seq signal, showing the parental differences between the 2-cell and 64-cell stages for maternal (red) and

paternal (blue) genomes at chromosome 13 (96–100 Mb) containing *Enc1*. **h**, As in **g**, for chromosome 14 (115–122 Mb) containing *Mbnl2*. **i**, Relative gene expression for *Xist* (in red) or *Jpx* (in yellow) in mouse embryonic fibroblasts derived from embryos issued from crossing  $\Delta Jpx$ /wild-type female mice with wild-type/Y or  $\Delta Jpx$ /Y male mice. The three genotypes analysed are indicated, as well as the number of independently derived mouse embryonic fibroblast cultures from independent single embryos ( $n=4, 6$  and  $6$  for wild-type/wild-type, wild-type/ $\Delta Jpx$  and  $\Delta Jpx/\Delta Jpx$  genotypes, respectively). Bar plot represents the mean of each independent expression value (for each embryo), error bars represent the s.d. and each dot represents an individual embryo value. **j**, Pie chart distribution of the genotypes obtained after mating  $\Delta Jpx$ /wild-type female mice with  $\Delta Jpx$ /Y male mice.  $n=104$  pups.



**Extended Data Fig. 7 | Analysis of X-linked gene position within the X chromosome as development progresses. a**, Clustering of X chromosome domain dynamics for contact enrichment (average Z-score). Domain number,  $n = 55, 75$  and  $26$  domains for each cluster. **b**, Structural changes. Lines represent the mean, and shading represents the 95% confidence interval of the

mean. **c**, Radial positions of X-linked genes, classified as early silenced, late-silenced and escapees as in a previous study<sup>20</sup>.  $n$  values are as in Fig. 4b. Box plot represents  $\pm 1.5 \times$  interquartile range, 25th and 75th percentiles and median value. Statistical difference was assessed using Wilcoxon's rank sum test (two-sided).

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

*Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.*

Data analysis

programs used:  
 R 3.4.2  
 Bioconductor 3.6  
 HiC-Pro 2.11.0  
 3dNetMod 1.0.10.06.17  
 HicExplorer 2.1.1  
 monocle 3  
 HiTC 1.26.0  
 Dip-C  
 Hickit  
 Trimgalore 0.4.4  
 picard 2.18.2  
 DeepTools 3.0  
 Subread 1.28.1  
 DESeq2 1.18.1  
 Codes developed for this study are openly available on our github repository : <https://github.com/heard-lab> .

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data generated in this study are available on GEO database under the number GSE129029

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences     Behavioural & social sciences     Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was not predetermined. Embryos were collected at many different development stages from independent females. Number (of cell and / or embryos are indicated for each appropriate section. We used sample size commonly used and accepted for each type of experiments (DNA FISH and single cell HiC) to allow for basic statistical inference while using an justifiable number of mice.
Data exclusions	no data were excluded from the analysis except for some single cell HiC when genomic coverage was inappropriate (extra chromosomal content, only maternal content). We also removed single cell HiC data with insufficient coverage (see Materials and Methods as well as extended figure 1
Replication	experiments were repeated independantly n=3 or more
Randomization	No randomization was used as animals were utilized to produced embryos at specific developmental time. However samples were treated, to when possible in parallel, and were all analyzed in parallel
Blinding	No experimental blinding was used as embryos embryo was done at specific developmental time post mating of animals.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	mouse embryonic stem cells mESCs E14 were used in this study as control (GSM1366337)
Authentication	as the cells used in this study have been recurrently used by the authors in previous studies they have not been authenticates
Mycoplasma contamination	ESCs were tested negative for mycoplasma
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	cells used are not in the ICLAC database

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	We used adult animals (from 6 weeks to 3 months for females, and 8 weeks to 1 year for males) for producing preimplantation embryos (first 4 days post mating). C57Bl/6J females (purchased from Charles river) were mated onto CASTEi/J males (purchase from Jackson Laboratory) for the HiC data and B6D2F1 females were mated onto B6D2F1 males for DNA FISH data. For CRISPRCas9 engineered animals we used B6D2F1 donor embryos transferred into pseudopregnat OF1 females. After founder identification the line was backcross on B6D2F1 animals
Wild animals	this study does not contain any wild animals
Field-collected samples	this study does not contain animals collected from the fields.
Ethics oversight	Animal care and use for this study were performed in accordance with the recommendations of the European community (2010/63/UE). All experimental protocols were approved by the ethics committee of Institut Curie CEEA-IC118 under the number APAFIS#8812-2017020611033784v2 given by national authority in compliance with the international guidelines.

Note that full information on the approval of the study protocol must also be provided in the manuscript.