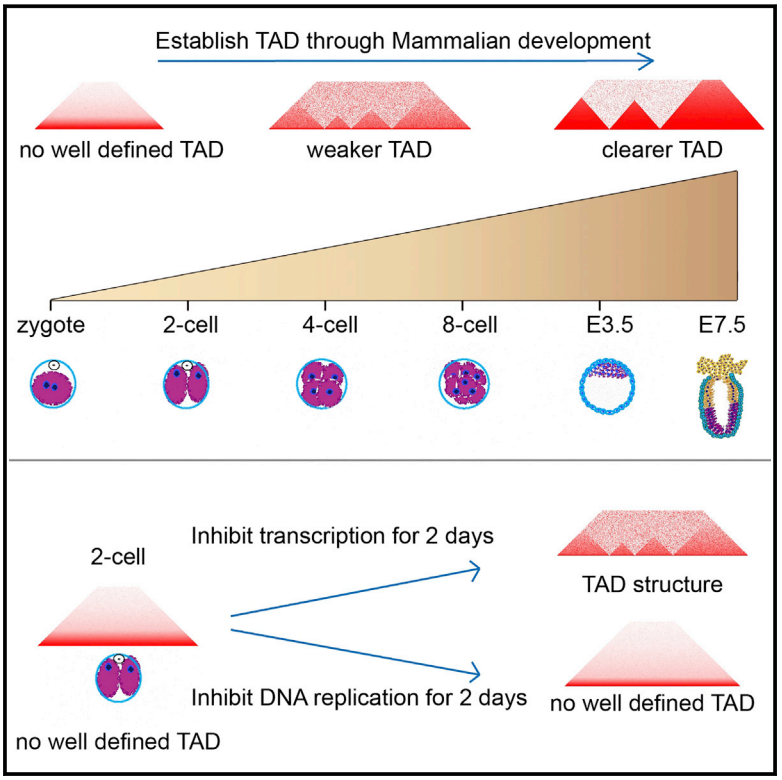
Resource

Cell

3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis

Graphical Abstract



Highlights

- Extra-long-range interactions along the genome are more frequent in sperm
- 3D chromatin structures undergo global reprogramming during embryo development
- Unmethylated CpGs and DNA demethylation are associated with chromatin structures
- TAD establishment requires DNA replication but not zygotic genome activation

Authors

Yuwen Ke, Yanan Xu, Xuepeng Chen, ..., Xiaocui Xu, Xingxu Huang, Jiang Liu

Correspondence

huangxx@shanghaitech.edu.cn (X.H.), liuj@big.ac.cn (J.L.)

In Brief

A comparative resource lays out the distinct conformational chromatin dynamics for mammalian sperm, oocytes, and early embryos.

Ke et al., 2017, Cell *170*, 367–381
July 13, 2017
http://dx.doi.org/10.1016/j.cell.2017.06.029

CellPress

Resource

Cell

3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis

Yuwen Ke,[1,3,7] Yanan Xu,[2,7] Xuepeng Chen,[1,3,7] Songjie Feng,[2,3,4,7] Zhenbo Liu,[1] Yaoyu Sun,[1,3] Xuelong Yao,[1,3] Fangzhen Li,[1] Wei Zhu,[1,3] Lei Gao,[1] Haojie Chen,[5] Zhenhai Du,[6] Wei Xie,[6] Xiaocui Xu,[1,3] Xingxu Huang,[2,*] and Jiang Liu[1,3,8,*]

[1]CAS Key Laboratory of Genome Sciences and Information, Collaborative Innovation Center of Genetics and Development, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China

[2]School of Life Science and Technology, ShanghaiTech University, Shanghai, China

[3]University of Chinese Academy of Sciences, Beijing, China

[4]Institute of Health Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

[5]Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

[6]MOE Key Laboratory of Bioinformatics, Center for Stem Cell Biology and Regenerative Medicine, THU-PKU Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing, China

[7]These authors contributed equally

[8]Lead Contact

*Correspondence: huangxx@shanghaitech.edu.cn (X.H.), liuj@big.ac.cn (J.L.)

http://dx.doi.org/10.1016/j.cell.2017.06.029

SUMMARY

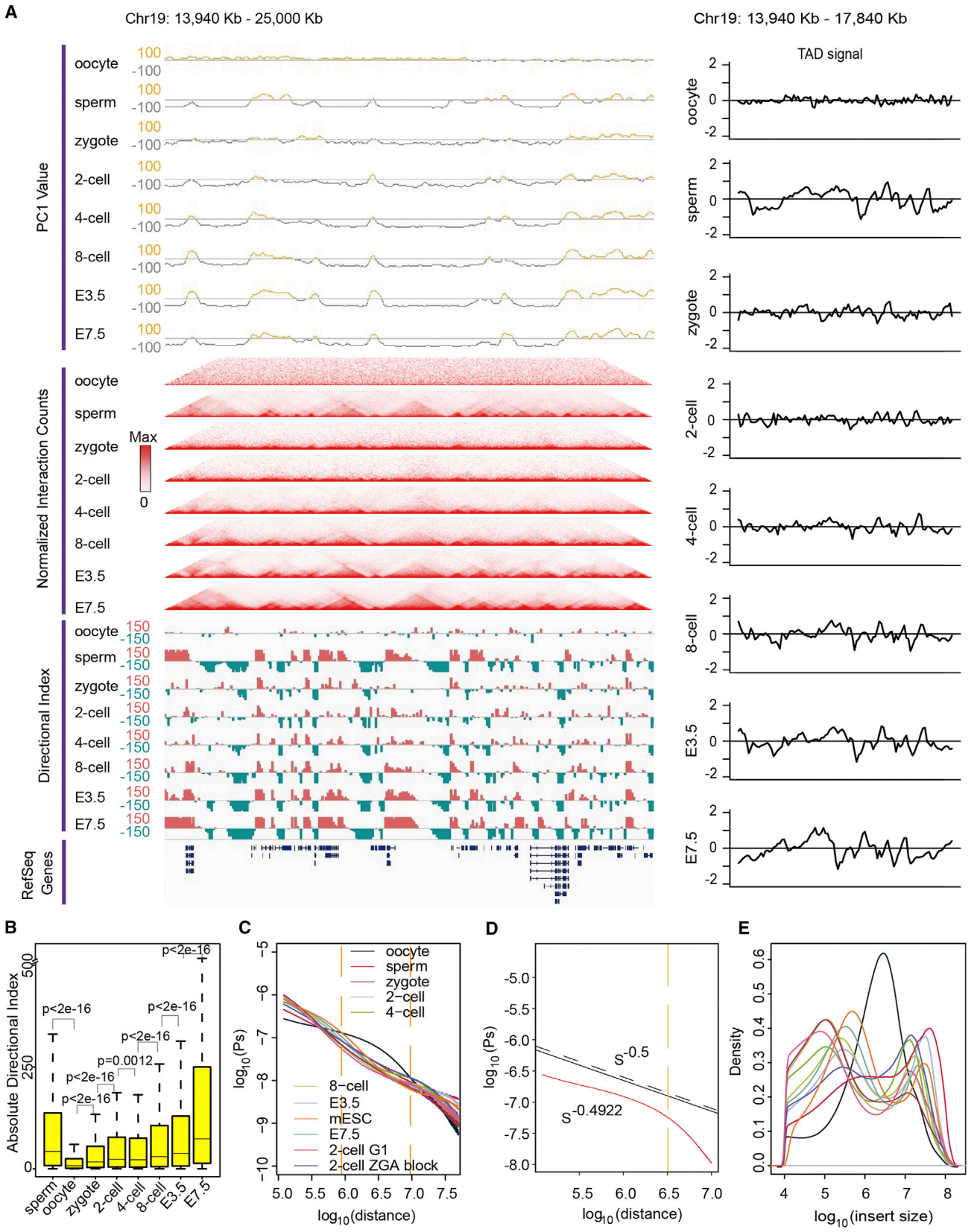
High-order chromatin structure plays important roles in gene expression regulation. Knowledge of the dynamics of 3D chromatin structures during mammalian embryo development remains limited. We report the 3D chromatin architecture of mouse gametes and early embryos using an optimized Hi-C method with low-cell samples. We find that mature oocytes at the metaphase II stage do not have topologically associated domains (TADs). In sperm, extra-long-range interactions (>4 Mb) and interchromosomal interactions occur frequently. The high-order structures of both the paternal and maternal genomes in zygotes and two-cell embryos are obscure but are gradually re-established through development. The establishment of the TAD structure requires DNA replication but not zygotic genome activation. Furthermore, unmethylated CpGs are enriched in A compartment, and methylation levels are decreased to a greater extent in A compartment than in B compartment in embryos. In summary, the global reprogramming of chromatin architecture occurs during early mammalian development.

INTRODUCTION

Chromosomes are packaged in a multi-hierarchical structure to fit into the crowded nucleus (Fullwood et al., 2009; Lieberman-Aiden et al., 2009). Recent studies have revealed that the topologically associated domain (TAD) appears to be a chromosomal secondary structure that is associated with histone modifications (Dixon et al., 2012; Sexton et al., 2012), gene expression (Dixon et al., 2015; Javierre et al., 2016; Le Dily et al., 2014; Nora et al., 2012), lamina association (Dixon et al., 2012; Kind et al., 2015), and DNA replication (Dixon et al., 2012; Pope et al., 2014). TADs are conserved across species. During cell differentiation, TADs remain relatively stable, but the frequency of interactions within them may change (Dixon et al., 2012, 2015; Krijger et al., 2016). TADs form the basis of the higher-level structures called "A" and "B" compartments. The "A" compartment region tends to be active, and the "B" compartment region tends to be inactive (Lieberman-Aiden et al., 2009).

In mammals, mature oocytes and mature sperm are transcriptionally quiescent. Mature oocytes pause at the metaphase II stage (MII stage) (Fulka et al., 1998). The majority of the sperm genome is occupied by protamine. The sperm nucleus is 10-fold smaller than that of the somatic cell (Johnson et al., 2011; Ooi and Henikoff, 2007). Nevertheless, it is not fully understood how the high-order chromatin structures fit the unique nuclear characteristics of mature sperm and oocytes in mammals.

Epigenetic reprogramming is essential for the precise gene regulation during early embryogenesis (Hackett and Surani, 2013). Genome-wide reprogramming of DNA methylomes, histone modifications, and chromatin accessibility occurs during embryo development in mammals (Dahl et al., 2016; Liu et al., 2016; Lu et al., 2016; Wang et al., 2014; Wu et al., 2016; Zhang et al., 2016). A recent paper has shown that TAD structures are almost unstructured in early *Drosophila* embryos before zygotic genome activation (ZGA), but TADs are re-established after ZGA (Hug et al., 2017). In mammals, it has been reported that the chromatin compartment of the maternal genome is much weaker than that of the paternal genome in zygotes (Flyamer et al., 2017). However, little is known about the reprogramming of 3D chromatin structures in mammalian embryos from the two-cell stage onward.



(legend on next page)

RESULTS

The Chromatin Architecture of Mature Oocytes at Metaphase

To study the chromatin architecture of mammalian gametes and early embryos, we optimized in situ Hi-C experiments by using a small number of cells (Lieberman-Aiden et al., 2009; Rao et al., 2014). We generated at least two replicates for mouse sperm, MII phase oocytes, zygotes at the PN4 stage, and early embryos at the two-cell, four-cell, eight-cell, embryonic day (E)3.5, and E7.5 stages. After filtering potentially artificial reads, we obtained an average of 170 million *cis*-paired reads for all stages (Table S1). After normalization, we validated the reproducibility of the Hi-C data according to normalized matrix Pearson correlation coefficient, interaction frequency, and A/B compartment PC1 (first principle component) values (Figures S1A–S1C).

We then analyzed chromatin conformation in mature gametes and early embryos at 40 kb resolution and plotted chromatin interaction frequency by using heatmaps, PC1 values, directional indices (DIs), and TAD signals (Figure 1A). The interaction maps of sperm display triangular patterns of regional enrichment (Figures 1A and S2A). In contrast, no triangular patterns are observed in any chromosomes of the mature oocytes at MII phase (Figures 1A and S2A). Our data show that both the absolute DI value (Figures 1B and S2B) and the absolute insulation score (Figure S2C) are lowest in oocytes, and these two values gradually increase through embryo development. Using the directional index method (Dixon et al., 2012) and the insulation score method (Giorgetti et al., 2016) to identify TADs, we did not find TADs in mature MII oocytes. The 3D structure of the MII phase oocyte genome is similar to that of mitotic cells, as reported previously (Naumova et al., 2013). This result is consistent with the fact that the chromosomes of mature oocytes are arrested at metaphase II (Fulka et al., 1998).

Next, we calculated the average intrachromosomal contact probability (Lieberman-Aiden et al., 2009) across all stages. Contact probability decreases monotonically from 10^5 to 10^8 bp for all samples (Figure 1C). The contact probability curves of sperm and embryos at different developmental stages are similar, but different from those of oocytes (Figure 1C). The power law exponent for the contact probability curve of MII oocytes is 0.4922 (Figure 1D), consistent with a previous report for mitotic cells (Naumova et al., 2013). We also plotted the density of paired loci at different insert sizes across all stages. The density pattern of MII oocytes differs from those of early embryos and sperm (Figure 1E). We observed two peaks in early embryos and one peak (at 3.1 million bases) in MII oocytes (Figure 1E). These results indicate that the chromosomes of mature MII oocytes do not form the TAD structure.

The Chromatin Architecture of Sperm

Previous studies have shown that TADs are present in mouse sperm (Battulin et al., 2015; Jung et al., 2017). However, because of limited sequencing depth, the resolution of the sperm structures was low. Here, we performed deep sequencing for mouse sperm. We observed one main insert size distribution peak in sperm (Figure 1E). This peak is ~41 Mb, which is longer than the insert sizes in other samples (Figure 1E). Consistently, empirical cumulative distribution function (ecdf) analysis shows that the distance reaching 80% of total interactions is the greatest in sperm (Figure S3A). These data suggest that the genomic distances of chromatin interactions in sperm tend to be longer than in early embryos. This is consistent with results showing that more interactions are far away from the diagonal line in sperm than in other samples (Figures 2A and S2A). To confirm this point, we calculated the relative ratio of the paired loci separated by genomic distance > 2 Mb versus those ≤ 2 Mb across different samples. Our data demonstrate that this ratio is the largest in sperm (Figure 2B). Cell cycle affects the high-order chromatin structure, with extra-long range interactions in the G1 phase being more frequent than those in other cell-cycle stages (Naumova et al., 2013). To exclude the effect of cell cycle, we wanted to compare the Hi-C data of sperm with that of somatic cells at the G1 phase. Because no Hi-C data for mouse cells synchronized at the G1 phase were available, we compared human sperm data with HeLa cell data at the G1 phase (Naumova et al., 2013). The extra-long-range interactions (>2 Mb) in human sperm are still more frequent than those in HeLa cells at the G1 phase (Figure S3B). These data indicate that there are more relatively longer-distance chromatin interactions in sperm than in somatic cells.

The interphase chromosomes can be subdivided into megabase-sized TADs, within which, loci prefer to contact each other (Dixon et al., 2012; Nora et al., 2012; Phillips-Cremins et al., 2013; Sexton et al., 2012). In sperm, a significant proportion of interactions are separated by genomic sequences >2 Mb, suggesting that the interactions outside TADs may be common. Our data show that the interactions within TADs only account for ~30% of the total interactions in sperm (Figure 2C). Furthermore, the majority of interactions outside TADs are likely to use TAD as the unit in the sperm genome (Figure 2A, blue box). Recently, it has been found that interactions between TADs (inter-TAD) occur frequently (Franke et al., 2016). Notably, more than 50% of the interactions in sperm, the largest proportion in all stages, are inter-TAD interactions (Figure 2C). Next, we

Figure 1. Overall Chromosome Structure

(A) Normalized Hi-C interaction frequencies displayed as a two-dimensional heatmap at 40 kb resolution overlaid on DI, A/B compartment-indicated PC1 value, and TAD signal.

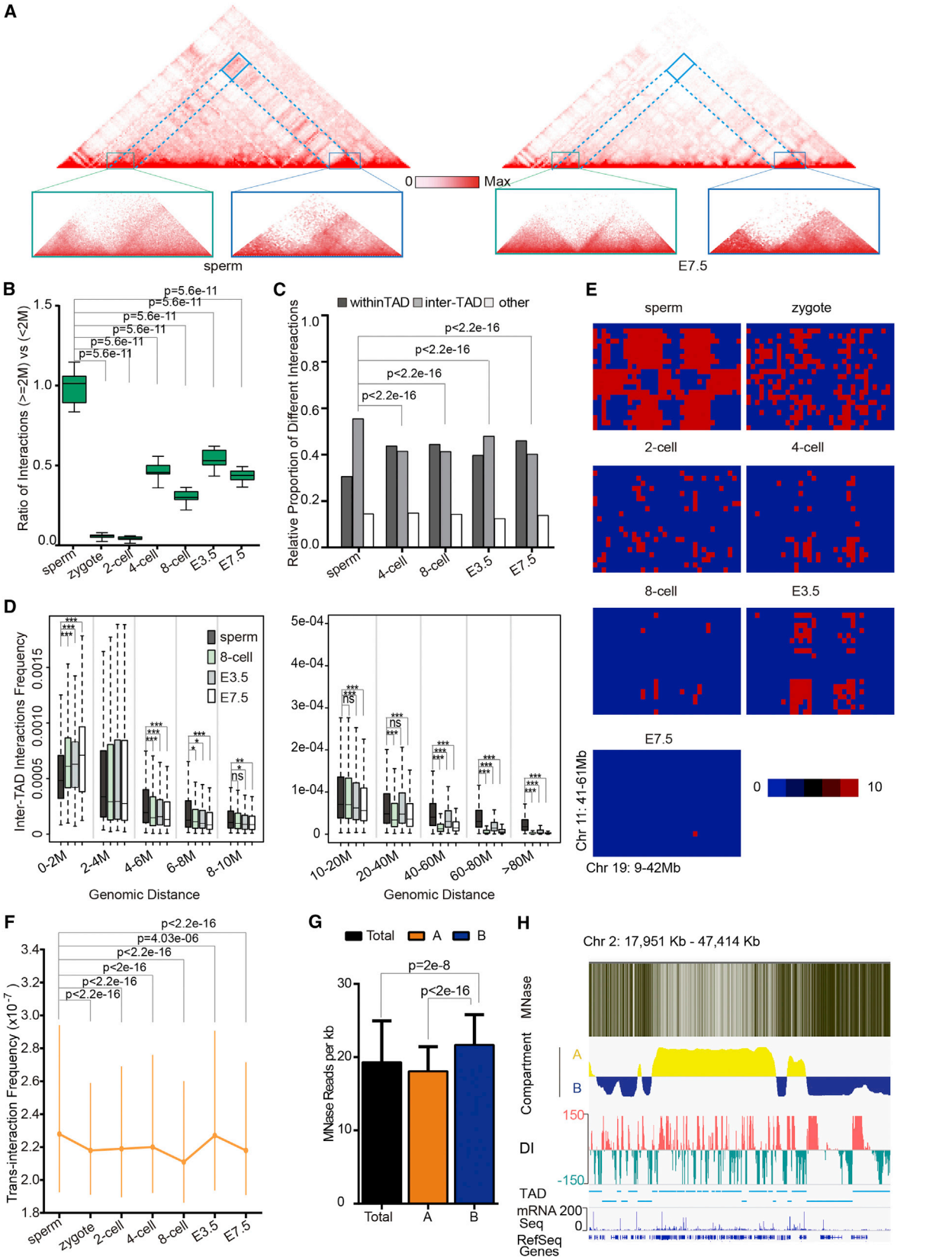
(B) Boxplot of absolute DI across all stages. P value was calculated by Kruskal-Wallis test followed by Dunn's multiple comparison test.

(C) The average contact probability across the genome decreases as a function of genomic distance. "2-cell G1" refers to two-cell embryos blocked at G1 phase by Aphidicolin treatment. "2-cell ZGA block" refers to two-cell embryos with zygotic activation blocked by α-Amanitin treatment.

(D) Power law exponents for the contact probability curve of MII oocytes.

(E) Relative density of paired loci with different insert sizes.

See also Figures S1 and S2 and Table S1.



(legend on next page)

examined the relative frequencies of inter-TAD interactions separated by different genomic distances. For TADs remaining stable in sperm, eight-cell, E3.5, and E7.5 embryos, the average interaction frequency between two TADs separated by $\leq$2 Mb of sequence in sperm is lower than that in E3.5 and E7.5 embryos (Figure 2D). When the distance is >4 Mb, the interaction frequency in sperm is higher than the frequency in early embryos, and the differences became more evident as the distance increased (Figure 2D). A similar trend is also observed for the relative proportion of all inter-TAD interactions versus the total interactions separated by different genomic distances (Figure S3C). These results indicate that the proportion of extra-long-range (>4 Mb) inter-TAD interactions in sperm is higher than that in early embryos.

We also analyzed the interchromosomal interactions in sperm (see the STAR Methods). The interchromosomal interactions are considerably more obvious in sperm than in E7.5 embryos (Figure S3D). Some regions from different chromosomes clearly interacted in sperm, but did not interact in E7.5 embryos (Figure 2E). We also calculated the relative frequencies of inter-chromosomal interactions across all stages (see the STAR Methods). The frequency of inter-chromosomal interactions between two *trans*-bins is highest in sperm (Figure 2F). Similarly, the relatively total inter-chromosomal interactions are also highest in sperm (Figure S3E).

The sperm genome is highly compressed because the majority of the genome is occupied by protamine (Carone et al., 2014). In mice, ~2%–4% of the sperm genome consists of histone-bound regions (Carone et al., 2014). The regions covered by nucleosomes have been mapped using micrococcal nuclease (MNase) sequencing (Carone et al., 2014). In the somatic tissues, regions with a high density of MNase-digested reads are usually associated with a low level of gene expression. Mammalian genomes can also be partitioned into A/B compartments (Dixon et al., 2012, 2015). The sperm genome was partitioned into A and B compartments and the MNase-digested reads were calculated in each compartment. Our data illustrate that MNase reads are enriched in B compartment regions (Figures 2G, 2H, and S3F) and genes in B compartments show lower expression than those in A compartments (Figure S3G), consistent with previous reports (Dixon et al., 2012, 2015).

In summary, sperm possess unique characteristics including a high frequency of interchromosomal interactions and extra-long-range intrachromosomal interactions, which are associated with the tight structure of sperm chromosome.

The Global Reprogramming of TADs during Early Embryogenesis

TAD domains are usually stable in different cell types and are highly conserved across species (Dixon et al., 2012; Ho et al., 2014). We were interested in the dynamics of TADs during early embryogenesis. We calculated the relative TAD signal variance across all stages. The relative TAD signal variance in zygotes is small, but it gradually increases through embryonic development (Figures 3A and S4A), suggesting that TAD structures become more evident during embryogenesis. Consistently, the triangular structures of TADs in the heatmaps at the zygote and two-cell stages are not as clear as those in other embryo stages (Figure 1A). We further calculated the proportion of the interactions less than 40, 80, 120, 160, or 200 kb versus total *cis*-interactions. The proportion of interactions within these distances in zygotes is much higher than that in gametes or embryos at the four-cell stage and onward (Figure 3B). The same results were also observed for two-cell embryos. These results show that relatively long-range interactions are limited in zygotes and two-cell embryos, indicating that the high-order structure of zygotes and two-cell embryos is obscure.

The DI value, which is associated with the calling of TADs, indicates the degree of confidence of a given bin having more interactions with upstream or downstream regions. DI values in the starting or ending bin of TAD at the zygotic and two-cell stages were very small, suggesting that the boundaries of the possible TADs are obscure at zygotic and two-cell stages (Figure S4B). Furthermore, the relative TAD intensities (see the STAR Methods) are very low in zygotes, but gradually increase throughout development (Figure S4C). Taken together, these data suggest that the TAD structures in zygotes and two-cell embryos are obscure.

Our data show that the TAD structures of sperm and embryos at eight-cell, E3.5, and E7.5 stages are clear. The number of TADs at these stages is ~2,000, similar to a previous report in mouse ES cells (Dixon et al., 2012). Furthermore, many TADs also undergo re-organization from four-cell embryos to E7.5 embryos (Figure 3C).

Clear TAD structures can be observed in somatic cells at G1 and S phases, but not at metaphase (Naumova et al., 2013).

Figure 2. Chromatin Architecture of Sperm

(A) Normalized Hi-C heatmaps of chromosome 19 at 40 kb resolution for sperm and E7.5 embryos. Blue rectangles represent extra-long distance inter-TAD interactions.

(B) The ratio of total interactions separated by genomic distance $\geq$2 Mb versus those <2 Mb. We used the chromosome as the unit for this analysis. P value was calculated by Wilcoxon test.

(C) The proportion of different kinds of interactions versus total interactions. P value was calculated by chi-square test.

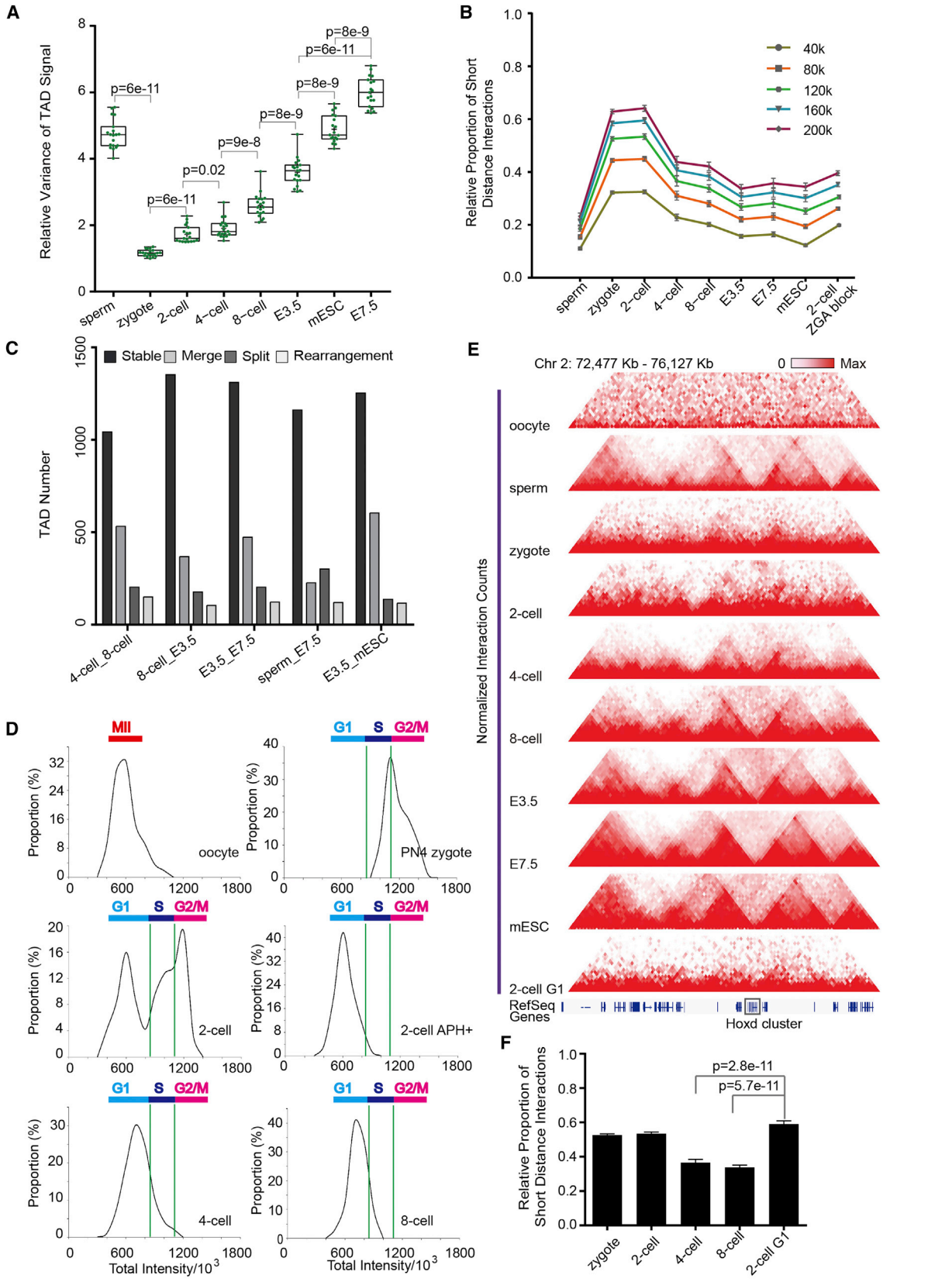
(D) Relative inter-TAD interaction frequency separated by a certain genomic distance. The relative interaction frequency was calculated using the interaction between two TADs divided by total *cis*-interactions at a specific stage. Statistical analysis was performed by using a one-sided Wilcoxon test. ***Indicates that interaction frequency at sperm is significantly higher than that at other stages (P value < 0.001). *p value < 0.05. ns, non-significant (p value > 0.05).

(E) Heatmap of trans-interactions between two regions from two chromosomes.

(F) The relative frequency of inter-chromosomal interactions between two bins. P value was calculated by Wilcoxon test.

(G) MNase reads are enriched in the B compartment in sperm. MNase reads per kb were calculated. "A" refers to the A compartment, and "B" refers to the B compartment. The p value was calculated by Wilcoxon test. Data are represented as mean ± SEM.

(H) The graph represents the MNase read intensity for a region in sperm chromosome 2, overlaid on the A/B compartments, DI, TAD, and mRNA-seq tracks. See also Figure S3.



(legend on next page)

Therefore, we investigated whether the obscure structures of early embryos are the result of cells at metaphase. First, we determined how many cells were in metaphase under a microscope. Less than 1% of cells are in metaphase for all the different samples examined (Figures S4D and S4E). These data indicate that the obscure structures of early embryos are not due to cells being in metaphase. We further checked the cell-cycle distribution at different embryo stages by quantifying DNA in cells using propidium iodide (PI) staining (Figure 3D). Our data illustrate that two-cell embryos are a mixture of cells at G1, S, and G2 phases (Figure 3D). We were interested in whether the dynamics among G1, S, and G2 phases cause the TAD structures of two-cell embryos to be difficult to detect. In this regard, we synchronized cells at the G1 phase using Aphidicolin, which can block DNA replication and arrest cells at G1 (Rosner et al., 2013). Aphidicolin-treated two-cell embryos were cultured for ~2 days after the control embryos had developed to the eight-cell stage. Heatmaps show that the high-order structure of the Aphidicolin-treated embryos is still obscure (Figure 3E). Consistently, the proportion of short-distance interactions in the treated two-cell embryos is significantly higher (p value $<2 \times 10^{-16}$, Wilcoxon test) than those in four-cell and later stage embryos (Figure 3F). These data demonstrate that obscure structures are still present in two-cell embryos at the G1 phase, suggesting that the weak structures observed in two-cell embryos are not caused by cell-cycle dynamics. Most cells in four-cell and eight-cell embryos are in G1 phase. The patterns of cell-cycle phases between four-cell and eight-cell embryos are similar (Figure 3D), although the high-order structures are different (Figure 3C). Taken together, these data indicate that the differences in high-order structures at different embryonic stages are not caused by cell-cycle variation. Our data also indicate that blocking DNA replication in two-cell embryos inhibits the establishment of high-order chromosome structures.

In summary, our data demonstrate that the TAD structures of zygotes and two-cell embryos are obscure, but the high-order chromatin structures are gradually re-established during embryo development.

The Association between Chromatin Structures and Embryonic Development

Compartment status is known to be associated with gene expression. Therefore, we investigated the switching of A/B compartment status during embryogenesis. As the compartment structures of zygotes and two-cell embryos are very weak (Figure S5A), we only examined the compartments in sperm and embryos at four-cell stages and onward. We found rearrangement of A and B compartments during embryogenesis (Figures 4A and S5B). Genes in the A compartments tend to show higher expression at different stages (Figure S5C), consistent with a previous report (Dixon et al., 2015). The expression of genes changing compartment status from A to B tends to decrease, while expression of genes changing compartment status from B to A tends to increase (Figure 4B). A set of 703 genes that showed concordance between changes in gene expression and A/B compartment switching (Table S2). Notably, the A/B compartments of many genes are stage-specific, and these genes are enriched for reproduction or embryonic process ontologies (Figure 4C; Table S3). Genes with A compartment status in sperm switching to B compartment status in E7.5 embryos are enriched in the category of male sex characteristics, such as *HOOK1* (Figure S5D), which is essential for producing sperm with normal morphology (Moreno et al., 2006). In addition, a number of developmental genes change their A/B compartment status during embryo development. For example, *Foxd3*, a pluripotency factor (Krishnakumar et al., 2016), is located in the B compartment in sperm, but changed to the A compartment in embryos (Figure S5D).

Hox genes play crucial roles in development. High-order chromatin structure is associated with the expression of *Hox* gene clusters (Narendra et al., 2015; Noordermeer et al., 2011). Herein, we checked the structures of *Hox* clusters in gametes and early embryos. The genomic length of *Hoxa*, *Hoxc*, and *Hoxd* clusters is ~100 kb, while the length of the *Hoxb* cluster is ~200 kb in mice. *Hoxb 13* and *Hoxb 9* are separated by a region of ~100 kb with no coding sequence (Figure 4D). The heatmap of sperm (5 kb resolution) shows that *Hoxb 13* and the rest of the *Hoxb* cluster belong to two different sub-TADs, and *Hoxb 13* has limited interactions with the rest of the *Hoxb* cluster (Figure 4D). In contrast, in E7.5 embryos, there is a special interaction structure between *Hoxb 13* and the rest of the *Hoxb* cluster and strong interactions were observed (Figure 4D). To validate this result, we performed circularized chromosome conformation capture sequencing (4C-seq) at two loci (Figure 4D; Table S4) with low-cell number samples and generated a relatively lower resolution interaction map. 4C-seq data of the *Hoxb 13* locus ("a" locus) confirmed that *Hoxb 13* has strong interactions with the rest of the *Hoxb* cluster in E7.5 embryos, but not in sperm (Figure 4E). Consistently, the *Hoxb 7* locus ("b" locus) had strong interactions with the *Hoxb 13*-containing sub-TAD in E7.5 embryos, but not in sperm (Figure 4F). Taken together, our data suggest that the structure of the *Hoxb* cluster differs between sperm and E7.5 embryos. The interactions between *Hoxb 13* and the rest of the *Hoxb*

Figure 3. Chromatin Dynamics at TAD Level during Embryogenesis

(A) Relative variance of TAD signal across all stages. P value is from Wilcoxon test.

(B) The ratio of paired loci separated by a genomic distance less than 40, 80, 120, 160, or 200 kb versus the total paired loci. The ratio of zygote and two-cell embryos is significantly higher (p value $<2 \times 10^{-16}$, Wilcoxon test) than in other stages for each distance cutoff.

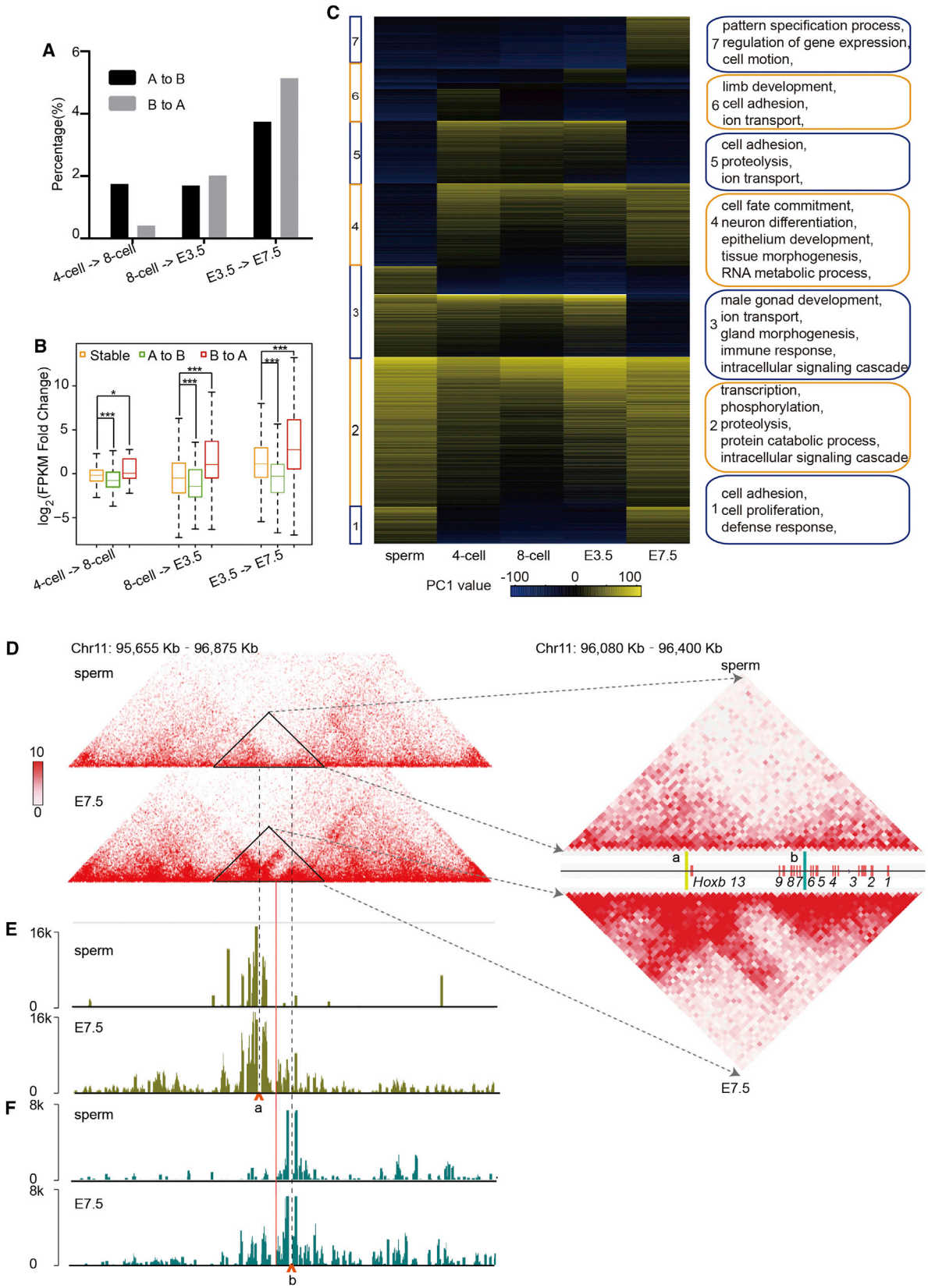
(C) The number of TADs that are stable, merged, split, and boundary-shifted between two stages.

(D) The distribution of cell cycle phases at different embryonic stages.

(E) Interaction frequency heatmap at 40 kb resolution at different stages. Two-cell G1 refers to Aphidicolin-treated two-cell embryos. The *Hoxd* cluster is located in the TAD boundary.

(F) The ratio of paired loci separated by a genomic distance <120 kb versus the total paired loci. The ratio of the Aphidicolin-treated two-cell embryos is significantly higher than that in four-cell and eight-cell embryos (Wilcoxon test). Data are represented as mean ± SEM.

See also Figure S4.



(legend on next page)

cluster may be required for spatiotemporal regulation of the *Hoxb* cluster in E7.5 embryos.

High-Order Chromatin Structures Are Associated with CpG Methylation and DNase I Hypersensitive Sites during Development

Global reprogramming of the DNA methylome, histone modification, and chromatin accessibility takes place during early mammalian embryogenesis. We were interested in the association between the reprogramming of these epigenetic modifications and 3D chromatin structure. Until now, an association between DNA methylation and chromatin structure has not been reported. Instead of analyzing the association between methylated CpGs and compartment status, we compared unmethylated CpGs with compartment status. Our data show that unmethylated CpGs (methylation level <0.25) are enriched in A compartment regions (Figures 5A and 5B). The unmethylated CpGs in A compartment regions are located in both promoter and genic regions (Figure 5C). Most of the unmethylated promoters are located in the A compartment (Figure 5D). Approximately three out of five promoters are in the unmethylated state in the A compartment, but only approximately one out of five promoters are in the unmethylated state in the B compartment in E7.5 embryos (Figure 5E). Unmethylated promoters are associated with better gene expression. Consistently, genes in A compartment regions tend to have better expression.

Genome-wide demethylation occurs during early mammalian embryogenesis. We further found that the decreased levels of CpG methylation in A compartments (74%) are higher than those in B compartments (66%) (Figure 5F). It is possible that the relatively loose structure of A compartments provides easier access for the DNA demethylation machinery in embryos.

Open chromatin structure can be accessed more readily by transcriptional factors, which are associated with long-range chromatin interactions. Recently, it has been shown that chromosomes lack DNase I hypersensitive sites (DHS) in zygotes and two-cell embryos, but DHSs are gradually established during embryo development (Lu et al., 2016), consistent with the gradual establishment of 3D chromatin structures during development. Combining our Hi-C data with DHS maps, we found that DHSs are enriched in A compartments (Figure 5G), consistent with the previous report. Furthermore, we found that the newly established DHSs in a specific stage are enriched in A compartments (Figure 5H).

Epigenetically modified H3K4me3 peaks are enriched in the A compartment. A similar result was observed in sperm and embryos (Figure S6). Broad H3K4me3 domains are present in oocytes and zygotes, but then disappear in late two-cell embryos (Dahl et al., 2016; Liu et al., 2016; Zhang et al., 2016). However, we did not find an association between broad H3K4me3 domains and compartment status. H3K27me3 modification are enriched in B compartments in preimplantation embryos, but are enriched in A compartments in ESC and E7.5 embryos (Figure S6A). We further analyzed the association between TAD boundaries and different histone modifications. As there are no clear TAD structures in zygotes and two-cell embryos, we focused our analyses on sperm and embryos at the four-cell stage and onward stage. Consistent with previous reports (Jung et al., 2017), our data show that H3K4me3, H3K27Ac, H3K9Ac, H3K4me1, CTCF, and SMC1 are enriched in the TAD boundaries, whereas H3K9me3, H3K36me3, and H3K27me3 are depleted in TAD boundaries (Figure S6B).

Allelic-Specific Chromatin Structure

MII oocytes and sperm differ dramatically in morphology and chromatin architecture. We were interested in whether the structure of the maternal genome also differs from that of the paternal genome after fertilization. Hi-C data were generated from embryos produced by the cross between female C57BL/6J strain and male DBA2/2J strain mice. We tracked the maternal genome and paternal genome by using SNPs between the two strains. The correlations according to interaction frequency are comparable between the paternal and maternal genomes at different stages (Figure S7A). Heatmaps illustrate that the high-order structures of both paternal and maternal genomes in zygotes and two-cell embryos are weak (Figure 6A). Consistent with these results, the proportion of short-range interactions also support that both maternal and paternal genomes have weak structures in zygotes and two-cell embryos, but they are re-established through development (Figure 6B).

We also examined compartment strength in maternal and paternal genomes. Compartment strength in both maternal and paternal genomes is weaker in zygotes than in E7.5 embryos, consistent with previous results (Figure S5A). In addition, the compartment strength in the maternal genome is much weaker than in the paternal genome in zygotes (Figure S7B), consistent with a recent publication (Flyamer et al., 2017). Our data also show that the A/B compartment status of some regions differ between the paternal and maternal genomes in early embryos (Figure S7C). Although the A/B compartment status of most imprinted regions is similar, the region encompassing the imprinted gene *DLK1* differs between the paternal and maternal

Figure 4. The Reprogramming of Chromatin Structure Is Associated with Early Development

(A) The proportion of the genome that changes A/B compartment status.

(B) Comparison of expression for genes with A/B compartment switching during development. ***p value <0.001, *p value <0.05 (Wilcoxon test).

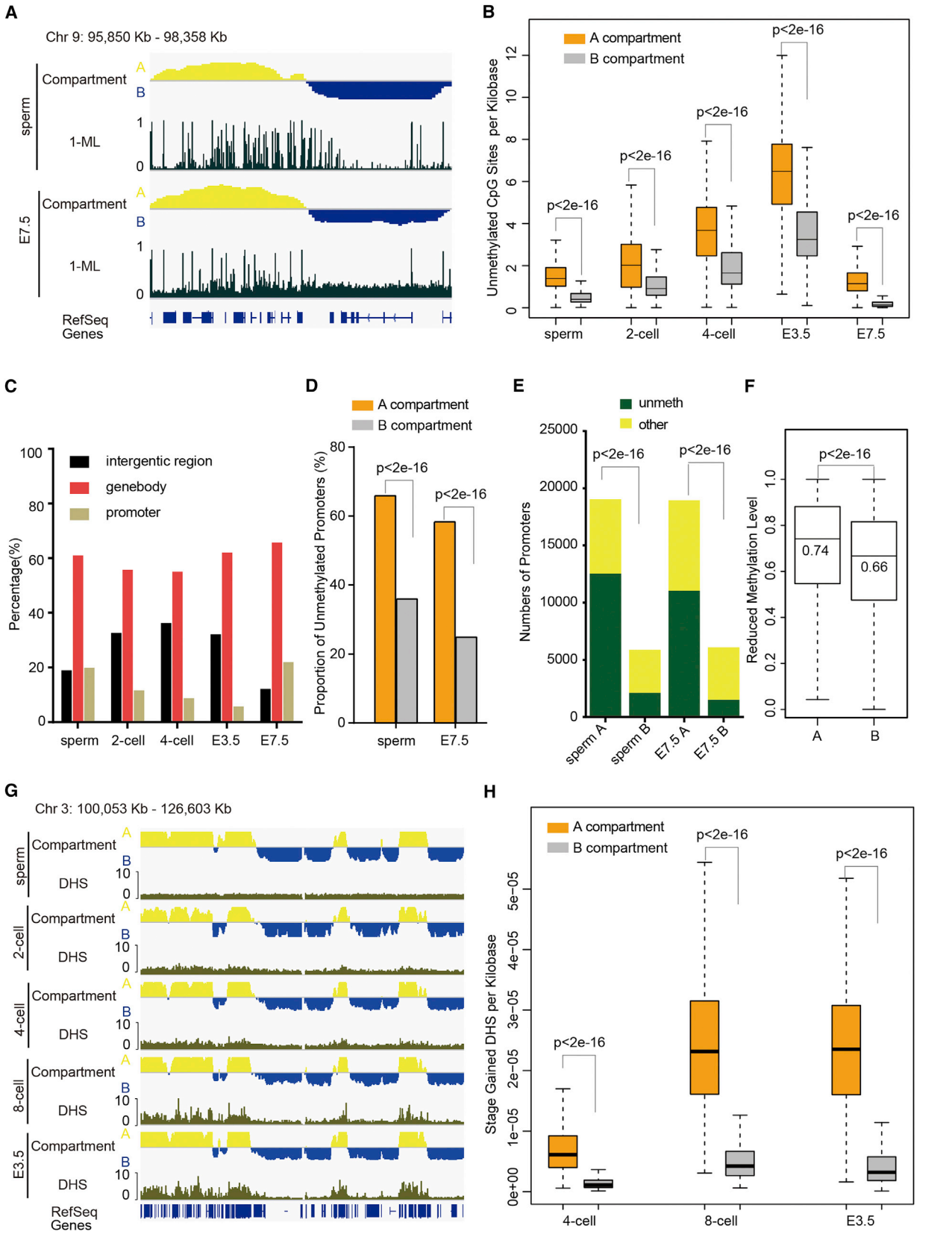
(C) GO enrichment for genes that show A/B compartment switching. PC1 values are sorted in the heatmap and clustered according to their dynamic pattern.

(D) The structural differences of *Hoxb* clusters between sperm and E7.5 embryos. The heatmap is displayed at 5 kb resolution in Juicebox (Durand et al., 2016). The heatmaps of both sperm and E7.5 embryos were generated with the same number of total contacts. Yellow (a) and cyan lines (b) (right panel) indicate the two viewpoints.

(E) 4C-seq results for the loci near *Hoxb 13*. Data were normalized between sperm and E7.5 embryos. The arrow refers to the view point. The red line indicates the TAD boundary in E7.5 embryos. We used the same number of total contacts for both sperm and E7.5 in the analysis.

(F) 4C-seq results for the loci near *Hoxb 7* are shown. We performed two replicates, each showing similar results. The same number of total contacts were analyzed for sperm and E7.5 embryos.

See also Figure S5 and Tables S2, S3, and S4.



(legend on next page)

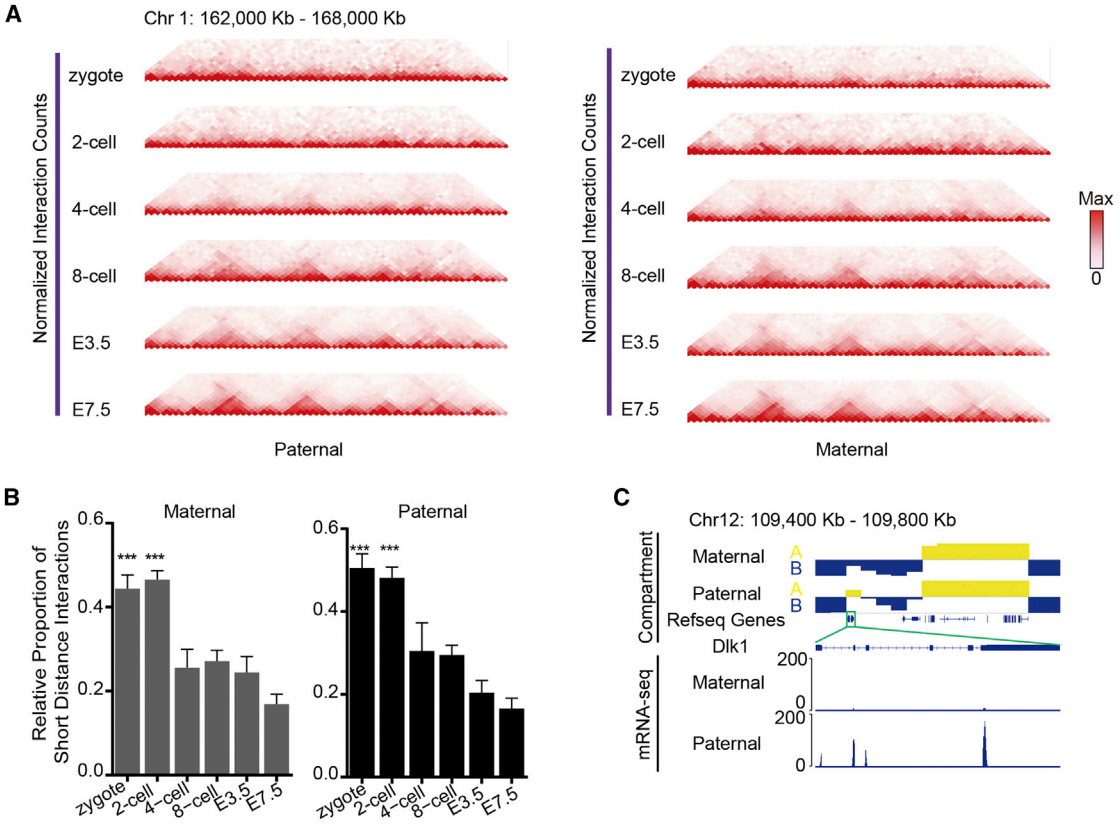


Figure 6. Allele-Specific Chromatin Structure

(A) Interaction frequency heatmap at 400 kb resolution for maternal and paternal genomes for zygote, two-cell, four-cell, eight-cell, E3.5, and E7.5 embryos.

(B) The ratio of paired loci separated by a genomic distance <120 kb versus the total paired loci for paternal and maternal genomes at different stages. ***Indicates that the proportion of short distance interactions in zygotes and two-cell embryos is significantly higher than that in four-cell, eight-cell, E3.5, and E7.5 embryos (Wilcoxon test). Data are represented as mean ± SEM.

(C) A/B compartment status and gene expression differ between maternal and paternal loci of the imprinted gene, DLK1, at E7.5 stage embryos.

See also Figure S7.

loci at the E7.5 stage (Figure 6C). Consistent with these data, *DLK1* expression differs between paternal and maternal copies (Figure 6C).

The Establishment of High-Order Structure Is Independent of Zygotic Genome Activation

Zygotic genome activation (ZGA) is an important biological process during early embryogenesis. We further investigated the effect of ZGA on the establishment of high-order chromatin structure. α-Amanitin can inhibit ZGA, but not DNA replication in two-cell embryos (Bolton et al., 1984), which we confirmed in mRNA-seq analysis (Figures 7A and 7B) and bromodeoxyuridine (BrdU) immunostaining (Figure S7D). α-Amanitin treatment blocks mouse embryos at the two-cell stage. Embryos remain alive, but are unable to undergo cell division. We collected embryos after ~2 days of treatment, when the untreated embryos developed into the eight-cell stage. Heatmaps illustrate that the high-order genome structures in the ZGA-blocked two-cell

Figure 5. Chromatin Structure Correlates with CpG Methylation and DHS

(A) The graph represents the A/B compartment status for a region in sperm and E7.5 embryos overlaid on CpG methylation. y axis using "1-ML" highlights the regions that are unmethylated. ML, methylation level.

(B) Unmethylated CpGs (ML <0.25) enriched in A compartment at different stages. P value was calculated by Wilcoxon test.

(C) Distribution of unmethylated CpGs in different genomic regions.

(D) Proportion of promoters that are unmethylated (ML <0.25) in A compartments and B compartments respectively. P value was calculated by chi-square test.

(E) Number of unmethylated promoters in A compartments and B compartments respectively, showing that promoters located in A compartments tend to be unmethylated. P value was calculated by chi-square test.

(F) The decreased methylation level of CpGs ($ML_{sperm} - ML_{E3.5}$) in A compartments and B compartments. P value was calculated by Wilcoxon test.

(G) The graph represents the A/B compartment status for a region overlaid on DHSs.

(H) Newly established DHSs enriched in A compartments during embryogenesis. P value was calculated by Wilcoxon test.

See also Figure S6.

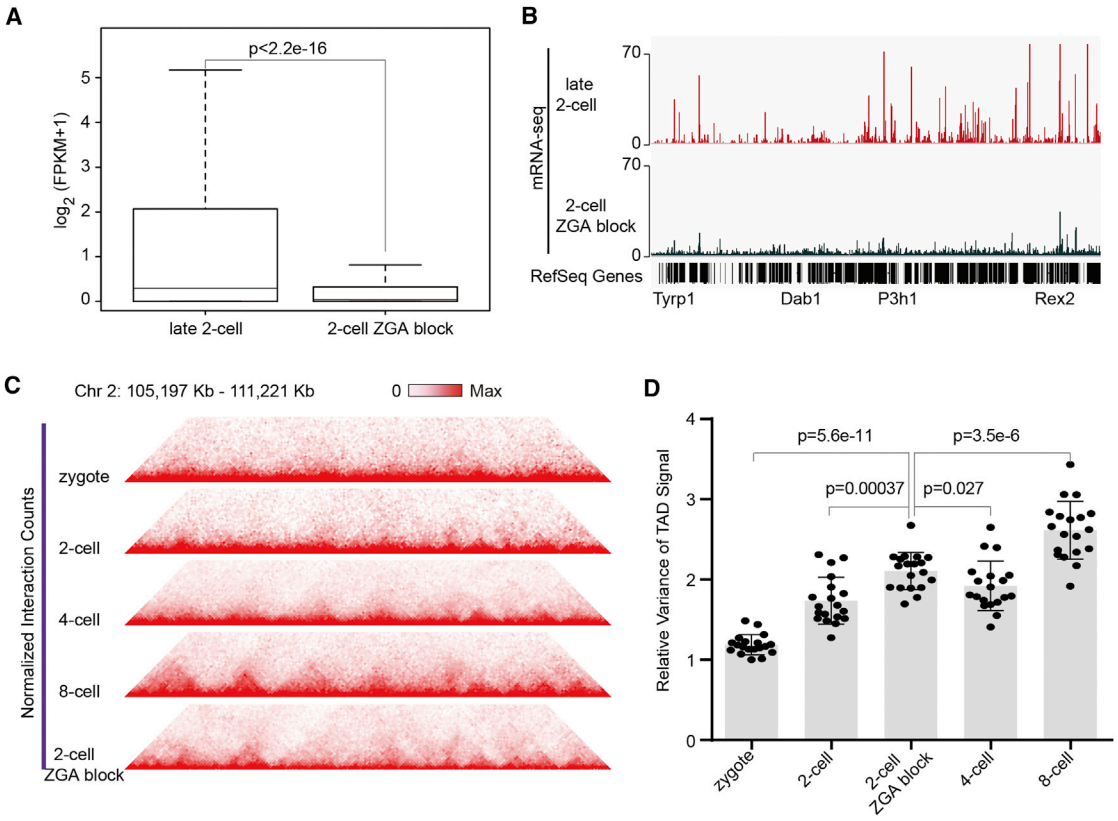


Figure 7. ZGA Is Not Required for the Establishment of High-Order Chromatin Structure

(A) Global expression difference between two-cell late-stage embryos and two-cell embryos treated with α-Amanitin. Statistical significance was determined by Wilcoxon test.

(B) Gene expression was inhibited after treatment with α-Amanitin.

(C) Interaction frequency heatmap at 40 kb resolution for zygote, two-cell, four-cell, eight-cell embryos, and two-cell ZGA-blocked embryos.

(D) The relative TAD signal variance of two-cell embryos treated with α-Amanitin compares with that of early embryos. Statistical significance was determined by Wilcoxon test. Data are represented as mean ± SEM.

See also Figure S7.

embryos are stronger than in the untreated two-cell embryos (Figure 7C). Consistent with this result, the relative variances of TAD signals in the ZGA-blocked embryos are also higher than in the untreated 2-cell embryos (Figure 7D). These data suggest that ZGA is not necessary for the establishment of high-order chromatin structure in early mouse embryos. Our result is consistent with previous work in *Drosophila* (Hug et al., 2017). In contrast, blocking DNA replication with Aphidicolin inhibits the establishment of TAD structures (Figures 3E and 3F). Taken together, these data suggest that the establishment of high-order chromatin structures is independent of ZGA but requires DNA replication in two-cell embryos.

DISCUSSION

TADs are relatively stable across various cell types. Interestingly, a recent study showed that *Drosophila* genome is mostly unstructured prior to ZGA, and TAD structures emerge during ZGA (Hug et al., 2017). Meanwhile, another group also investigated the chromatin structure of zygotes in mouse (Flyamer et al., 2017), they revealed that the compartment strength of maternal genome is much weaker than that of paternal genome in the zygotes, which is also confirmed by our data. Because Flyamer et al. (2017) did not generate the Hi-C data for mouse embryos at two-cell onward stages, they did not reveal that the TAD structure of zygotes was mostly unstructured compared to those of late embryos. In our study, we have generated the Hi-C data for the embryos at mouse zygotic, two-cell, four-cell, eight-cell, blastocyst, and E7.5 stages. By comparing the Hi-C data across different developmental stages, we reveal that the TAD structures of mouse zygotes and two-cell embryos are mostly unstructured too, and the high-order chromatin structure is gradually re-established during embryogenesis.

Hug et al. (2017) have found that the establishment of chromatin architecture is independent of ZGA in *Drosophila*. Consistently, we also find that blocking ZGA cannot inhibit the establishment of chromatin architecture in mouse embryos. These results suggest that a conserved mechanism may be involved in the establishment process. Furthermore, our data show that DNA replication, but not ZGA, is required for the establishment

of global chromatin structure in two-cell embryos. Perhaps, some DNA binding proteins required for high-order structure can easily access to DNA during the replication process. It is also possible that DNA replication in two-cell embryos can facilitate the replacement of some histone variances/modifications in chromosomes, which may be important for the establishment of the high-order chromatin structures. It is unknown why TAD structure of zygotes and two-cell embryos are obscure. It is known that zygotes and two-cell embryos have large nucleolar bodies. It is possible that these large nucleolar bodies are associated with the weak chromosomal structures.

We have presented high-resolution genome-wide maps of chromatin interactions in mature mammalian gametes and early embryos. Recent data have shown that germinal vesicle oocytes have TADs (Flyamer et al., 2017). Here, we show that mature MII oocytes do not have TADs. The extra-long-range interactions in sperm tend to be more frequent than those in somatic cells.

In summary, we have shown the unique 3D chromatin structure of mature gametes and the global reprogramming of chromatin structure during embryo development. Our data provide a valuable resource for the study of mammalian embryonic development.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Collection of sperm, oocytes, and early embryos
  - In situ Hi-C library generation using low cell numbers
  - 4C-seq library generation using a low number of cells
  - Profiling cell cycle phase by using propidium iodide immunofluorescence staining
  - BrdU staining
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Hi-C data processing, mapping, and ICE normalization
  - Pearson correlation coefficients analysis
  - TAD signal calculation and relative TAD signal variance
  - Insulation score analysis
  - Averaged contact probability P(s)
  - Density of paired loci at different insert sizes
  - TAD calling, TAD merge, and split between gametes and embryos
  - Inter-TAD interactions analysis
  - Relative inter-chromosomal interaction frequency
  - A/B Compartment PC1 value and status switch
  - Sperm MNase data processing and MNase fragment score
  - Identification of concordant genes with A/B compartment switch
  - Relative TAD intensity (RTI)
  - Domain boundary ChIP-seq signal enrichment analysis
  - Methylation and DHS enrichment analysis
  - 4C-seq bioinformatics
  - RNA-seq data processing
  - Hi-C allele-specific analysis
  - Compartment strength
  - Statistical methods
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2017.06.029.

AUTHOR CONTRIBUTIONS

X.H. and J.L. conceived the study. Y.K., X.C., and Y.X. facilitated its designs. Y.K., Y.X., S.F., Z.L., X.Y., Y.S., and L.G. collected samples. Y.K. and X.C. performed Hi-C library construction. Z.L. performed the mRNA library construction. X.C., F.L., W.Z., and X.X. performed the bioinformatics analyses. Y.K., Y.X., X.C., X.H., and J.L. interpreted the data. Y.K., X.C., X.H., and J.L. wrote the paper with the assistance of other authors.

ACKNOWLEDGMENTS

We thank Li Lan for the language editing and the sequencing facility in BIG, CAS. This work was supported by grants from CAS funding (XDB13040200), the National Key R&D Program (2016YFA0500903), the 973 Program of China (2015CB856200), and from the National Natural Science Foundation of China (31630040, 31425015, 31561130152, 91519306, 81171902, and 81402112).

Received: March 8, 2017
Revised: April 19, 2017
Accepted: June 19, 2017
Published: July 13, 2017

REFERENCES

Battulin, N., Fishman, V.S., Mazur, A.M., Pomaznoy, M., Khabarova, A.A., Afonnikov, D.A., Prokhortchouk, E.B., and Serov, O.L. (2015). Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach. Genome Biol. *16*, 77.

Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes Dev. *23*, 2484–2489.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Bolton, V.N., Oades, P.J., and Johnson, M.H. (1984). The relationship between cleavage, DNA replication, and gene expression in the mouse 2-cell embryo. J. Embryol. Exp. Morphol. *79*, 139–163.

Carone, B.R., Hung, J.H., Hainer, S.J., Chou, M.T., Carone, D.M., Weng, Z., Fazzio, T.G., and Rando, O.J. (2014). High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. Dev. Cell *30*, 11–22.

Dahl, J.A., Jung, I., Aanes, H., Greggains, G.D., Manaf, A., Lerdrup, M., Li, G., Kuan, S., Li, B., Lee, A.Y., et al. (2016). Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. Nature *537*, 548–552.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–498.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331–336.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst. *3*, 99–101.

Flyamer, I.M., Gassler, J., Imakaev, M., Brandão, H.B., Ulianov, S.V., Abdennur, N., Razin, S.V., Mirny, L.A., and Tachibana-Konwalski, K. (2017). Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature *544*, 110–114.

Franke, M., Ibrahim, D.M., Andrey, G., Schwarzer, W., Heinrich, V., Schöpflin, R., Kraft, K., Kempfer, R., Jerković, I., Chan, W.L., et al. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. Nature *538*, 265–269.

Fulka, J., Jr., First, N.L., and Moor, R.M. (1998). Nuclear and cytoplasmic determinants involved in the regulation of mammalian oocyte maturation. Mol. Hum. Reprod. *4*, 41–49.

Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H., et al. (2009). An oestrogen-receptor-alpha-bound human chromatin interactome. Nature *462*, 58–64.

Giorgetti, L., Lajoie, B.R., Carter, A.C., Attia, M., Zhan, Y., Xu, J., Chen, C.J., Kaplan, N., Chang, H.Y., Heard, E., and Dekker, J. (2016). Structural organization of the inactive X chromosome in the mouse. Nature *535*, 575–579.

Hackett, J.A., and Surani, M.A. (2013). Beyond DNA: programming and inheritance of parental methylomes. Cell *153*, 737–739.

Hammoud, S.S., Low, D.H., Yi, C., Carrell, D.T., Guccione, E., and Cairns, B.R. (2014). Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. Cell Stem Cell *15*, 239–253.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Ho, J.W., Jung, Y.L., Liu, T., Alver, B.H., Lee, S., Ikegami, K., Sohn, K.A., Minoda, A., Tolstorukov, M.Y., Appert, A., et al. (2014). Comparative analysis of metazoan chromatin organization. Nature *512*, 449–452.

Hug, C.B., Grimaldi, A.G., Kruse, K., and Vaquerizas, J.M. (2017). Chromatin architecture emerges during zygotic genome activation independent of transcription. Cell *169*, 216–228.

Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat. Methods *9*, 999–1003.

Javierre, B.M., Burren, O.S., Wilder, S.P., Kreuzhuber, R., Hill, S.M., Sewitz, S., Cairns, J., Wingett, S.W., Varnai, C., Thiecke, M.J., et al. (2016). Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. Cell *167*, 1369–1384.

Johnson, G.D., Lalancette, C., Linnemann, A.K., Leduc, F., Boissonneault, G., and Krawetz, S.A. (2011). The sperm nucleus: chromatin, RNA, and the nuclear matrix. Reproduction *141*, 21–36.

Jung, Y.H., Sauria, M.E., Lyu, X., Cheema, M.S., Ausio, J., Taylor, J., and Corces, V.G. (2017). Chromatin states in mouse sperm correlate with embryonic and adult regulatory landscapes. Cell Rep. *18*, 1366–1382.

Kind, J., Pagie, L., de Vries, S.S., Nahidiazar, L., Dey, S.S., Bienko, M., Zhan, Y., Lajoie, B., de Graaf, C.A., Amendola, M., et al. (2015). Genome-wide maps of nuclear lamina interactions in single human cells. Cell *163*, 134–147.

Krijger, P.H., Di Stefano, B., de Wit, E., Limone, F., van Oevelen, C., de Laat, W., and Graf, T. (2016). Cell-of-origin-specific 3D genome structure acquired during somatic cell reprogramming. Cell Stem Cell *18*, 597–610.

Krishnakumar, R., Chen, A.F., Pantovich, M.G., Danial, M., Parchem, R.J., Labosky, P.A., and Blelloch, R. (2016). FOXD3 regulates pluripotent stem cell potential by simultaneously initiating and repressing enhancer activity. Cell Stem Cell *18*, 104–117.

Krueger, F., and Andrews, S.R. (2016). SNPsplit: Allele-specific splitting of alignments between genomes with known SNP genotypes. F1000Res. *5*, 1479.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Le Dily, F., Baù, D., Pohl, A., Vicent, G.P., Serra, F., Soronellas, D., Castellano, G., Wright, R.H., Ballare, C., Filion, G., et al. (2014). Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. Genes Dev. *28*, 2151–2162.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289–293.

Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., Chen, J., Zhao, Y., Gao, H., Wang, H., et al. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature *537*, 558–562.

Lu, F., Liu, Y., Inoue, A., Suzuki, T., Zhao, K., and Zhang, Y. (2016). Establishing chromatin regulatory landscape during mouse preimplantation development. Cell *165*, 1375–1388.

Moreno, R.D., Palomino, J., and Schatten, G. (2006). Assembly of spermatid acrosome depends on microtubule organization during mammalian spermiogenesis. Dev. Biol. *293*, 218–227.

Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science *347*, 1017–1021.

Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Lajoie, B.R., Mirny, L.A., and Dekker, J. (2013). Organization of the mitotic chromosome. Science *342*, 948–953.

Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W., and Duboule, D. (2011). The dynamic architecture of Hox gene clusters. Science *334*, 222–225.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385.

Ooi, S.L., and Henikoff, S. (2007). Germline histone dynamics and epigenetics. Curr. Opin. Cell Biol. *19*, 257–265.

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics *20*, 289–290.

Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell *153*, 1281–1295.

Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K., et al. (2014). Topologically associating domains are stable units of replication-timing regulation. Nature *515*, 402–405.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*(W1), W160–W165.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

Rosner, M., Schipany, K., and Hengstschläger, M. (2013). Merging high-quality biochemical fractionation with a refined flow cytometry approach to monitor nucleocytoplasmic protein expression throughout the unperturbed mammalian cell cycle. Nat. Protoc. *8*, 602–626.

Saldanha, A.J. (2004). Java Treeview–extensible visualization of microarray data. Bioinformatics *20*, 3246–3248.

Servant, N., Lajoie, B.R., Nora, E.P., Giorgetti, L., Chen, C.J., Heard, E., Dekker, J., and Barillot, E. (2012). HiTC: exploration of high-throughput 'C' experiments. Bioinformatics *28*, 2843–2844.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458–472.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., and Ren, B. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116–120.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. *7*, 562–578.

Wang, L., Zhang, J., Duan, J., Gao, X., Zhu, W., Lu, X., Yang, L., Zhang, J., Li, G., Ci, W., et al. (2014). Programming and inheritance of parental DNA methylomes in mammals. Cell *157*, 979–991.

Wingett, S., Ewels, P., Furlan-Magaril, M., Nagano, T., Schoenfelder, S., Fraser, P., and Andrews, S. (2015). HiCUP: pipeline for mapping and processing Hi-C data. F1000Res. *4*, 1310.

Wu, J., Huang, B., Chen, H., Yin, Q., Liu, Y., Xiang, Y., Zhang, B., Liu, B., Wang, Q., Xia, W., et al. (2016). The landscape of accessible chromatin in mammalian preimplantation embryos. Nature *534*, 652–657.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., Ming, J., Wu, X., Zhang, Y., Xu, Q., et al. (2016). Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature *537*, 553–557.

Zheng, H., Huang, B., Zhang, B., Xiang, Y., Du, Z., Xu, Q., Li, Y., Wang, Q., Ma, J., Peng, X., et al. (2016). Resetting epigenetic memory by reprogramming of histone modifications in mammals. Mol. Cell *63*, 1066–1079.

Zhou, X., Lowdon, R.F., Li, D., Lawson, H.A., Madden, P.A., Costello, J.F., and Wang, T. (2013). Exploring long-range genome interactions using the WashU Epigenome Browser. Nat Methods. *10*, 375–376.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rat Anti-BrdU	Abcam	cat# ab6326; RRID: AB_305426
Donkey Anti-Rat Alexa Fluor 488	Invitrogen	cat# A-21208; RRID: AB_141709
Chemicals, Peptides, and Recombinant Proteins		
α-Amanitin	Sigma-Aldrich	cat#A2263
Aphidicolin	Sigma-Aldrich	cat#89458
Biotin-14-dATP	Life Technologies	cat#19524016
MboI	New England Biolabs	cat#R0147L
DpnII	New England Biolabs	cat#R0543L
HindIII	New England Biolabs	cat#R3104L
DNA Polymerase I Klenow Fragment	New England Biolabs	cat#M0210L
T4 DNA Ligase	New England Biolabs	cat#M0202L
Proteinase K solution	Thermo Fisher	cat#AM2548
Dynabeads MyOne Streptavidin C1	Life Technologies	cat#65002
AMPure XP beads	Beckman Coulter	cat#A63881
Critical Commercial Assays		
NEBNext Ultra II End Repair/dA-Tailing Module	New England Biolabs	cat#E7546L
NEBNext Ultra II End Ligation Module	New England Biolabs	cat#E7595L
Q5 High-Fidelity 2X Master Mix	New England Biolabs	cat#E0492L
hypotonic lysis buffer	Amresco	cat#,M334
Deposited Data		
Hi-C data from staged embryos	this paper	GSA: PRJCA000241
4C-seq data for sperm and E7.5 embryos	this paper	GSA: PRJCA000241
RNA-seq data from E3.5, E7.5 embryos, α-Amanitin treated 2-cell and control 2-cell	this paper	GSA: PRJCA000241
Human sperm Dataset	this paper	GSA: PRJCA000241
Human HeLa G1 Dataset	Naumova et al., 2013	ArrayExpress: E-MTAB-1948
Mouse Sperm MNase Dataset	Carone et al., 2014	GEO: GSE58101
Mouse Sperm RNaseq Dataset	Hammoud et al., 2014	GEO: GSE49622
mESC Hi-C Dataset	Dixon et al., 2012	GEO: GSE35156
Mouse early embryo (zygote,2-cell,4-cell,8-cell)Smart-seq2 Dataset	Wu et al., 2016	GEO: GSE66390
Mouse early embryo H3K27ac Dataset	Dahl et al., 2016	GEO: GSE72784
Mouse early embryo H3K4me3 Dataset	Zhang et al., 2016	GEO: GSE71434
Mouse early embryo H3K27me3 Dataset	Zheng et al., 2016	GEO: GSE76687
Mouse early embryo H3K4me3 and H3K27me3 Dataset	Liu et al., 2016	GEO: GSE73952
Mouse early embryo DNase-seq Dataset	Lu et al., 2016	GEO: GSE76642
mESC H3K9me3 Dataset	Bilodeau et al., 2009	GEO: GSE18371
mESC H3K4me3 and H3K27ac Dataset	Shen et al., 2012	GEO: GSE29184
mouse early embryo Methylation Dataset	Wang et al., 2014	GEO: GSE56697
Mouse Sperm Histone Modification and CTCF Dataset	Jung et al., 2017	GEO: GSE79230
Experimental Models: Organisms/Strains		
C57BL/6	Beijing Vital River Laboratory Animal Technology Co., Ltd.	Stock number 213

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
DBA/2J	SLAC Laboratory Animal Co., Ltd.	N/A
Oligonucleotides		
Primers for 4C-seq see Table S4	this paper; Synthesized by Sangon Biotech (Shanghai)	N/A
Software and Algorithms		
Bowtie2 (v2.2.6)	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
MACS2 (v2.1.0)	Zhang et al., 2008	https://github.com/taoliu/MACS/
HiCUP (v.0.5.7)	Wingett et al., 2015	http://www.bioinformatics.babraham.ac.uk/projects/hicup/
HiTC	Servant et al., 2012	http://www.bioconductor.org/packages/release/bioc/html/HiTC.html
Bedtools	Quinlan and Hall, 2010	http://bedtools.readthedocs.io/en/latest/
Samtools	Li et al., 2009	https://github.com/samtools/samtools
STAR(v 2.5.2b)	Dobin et al., 2013	https://github.com/alexdobin/STAR/releases
Cuffinks (v2.2.1)	Trapnell et al., 2012	http://cole-trapnell-lab.github.io/cufflinks/install/
SNPsplit (v0.3.2)	Krueger and Andrews, 2016	http://www.bioinformatics.babraham.ac.uk/projects/SNPsplit/
Juicer/Juicebox	Durand et al., 2016	https://github.com/theaidenlab/juicer
Deeptools(2.5.0.1)	Ramírez et al., 2016	http://deeptools.ie-freiburg.mpg.de/
C-world (Hi-C analysis software)	Job Dekker lab	https://github.com/dekkerlab/cworld-dekker
R	R Core Team	http://www.R-project.org/
Trimmomatic v0.32	Bolger et al., 2014	http://www.usadellab.org/cms/?page=trimmomatic
R package: ape	Paradis et al., 2004	https://cran.r-project.org/web/packages/ape/index.html
Java TreeView	Saldanha, 2004	http://jtreeview.sourceforge.net/
HOMER	Heinz et al., 2010	http://homer.ucsd.edu/homer/interactions/index.html
Picard Tools	DePristo et al., 2011	https://broadinstitute.github.io/picard/
WashU Eepigenome Bbrowser	Zhou et al., 2013	http://epigenomegateway.wustl.edu/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jiang Liu (liuj@big.ac.cn).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All animal maintenance and experimental procedures were carried out according to the guidelines of the Institutional Animal Care and Use Committee (IACUC) of the Beijing Institute of Genomics, CAS, Beijing, China.

All mice used in this study were housed under a 12 hr light/dark cycle. Littermates of the same sex were randomly assigned to experimental groups. Health status was normal for all mice. Mouse strains, C57BL/6J and DBA/2J, were used for this study.

METHOD DETAILS

Collection of sperm, oocytes, and early embryos

Mouse oocytes and sperm were collected from C57BL/6J mice. Zygotes, 2-cell, 4-cell, 8-cell, E3.5, and E7.5 embryos were from the cross of female C57BL/6J and male DBA/2J mice. C57BL/6J female mice were intraperitoneally injected with pregnant mare serum gonadotropin (PMSG, 5 IU), followed by human chorionic gonadotropin (hCG, 5 IU) 48 hr later. Metaphase II oocytes were isolated from the ampulla 12 hr after hCG injection and collected in hyaluronidase-containing M2 medium (Millipore) drops to

eliminate cumulus cell contaminants. To avoid somatic contaminants, oocytes were washed via progressive dilution (8 times) through sequential drops of CO_2-buffered, amino-acid supplemented KSOM (Millipore). Each oocyte was confirmed to be free of somatic contaminants under a microscope. The polar body of each oocyte was removed by pipetting through a fine pulled-glass needle. Superovulated C57BL/6J female mice were mated with DBA/2J male mice after hCG injection. PN4 stage zygotes were collected 26–27 hr after hCG injection. 2-cell embryos were collected 30–32 hr after hCG injection. All abnormal embryos were discarded. Embryos were carefully washed with KSOM to deplete somatic contaminants. The second polar bodies for zygotes and 2-cell embryos were also removed. To inhibit zygotic genome activation, zygotes (PN4) were cultured in KSOM supplemented with α-Amanitin for about 38–40 hr. 4-cell and 8-cell embryos were collected from the embryos cultured from 2-cell embryos in KSOM. Cavitated E3.5 blastocysts were collected by flushing the uteri of naturally mated mice with M2 medium (Sigma). Embryos were then treated with rabbit anti-mouse serum (Sigma) for 30 min in a CO_2 incubator at 37°C. Embryos were washed with KSOM and then added to standard guinea pig complement (Cedarlane) in KSOM for 30 min in a CO_2 incubator at 37°C. Finally, embryos were treated with acid Tyrode's solution (Sigma) and pipetted through a fine pulled-glass needle to remove the zona pellucida. Isolated blastocysts were serially washed to remove contaminants. E7.5 embryos were isolated through mechanical dissection of the decidua from the uterine lining of mated mice. Samples were again serially washed, and peripheral trophectodermal tissues were dissected away using fine glass capillaries. We collected the cauda epididymis from 8-week-old male C57BL/6J mice. The cauda epididymis was cut into pieces and placed in 4 mL of BSA-supplemented human tubule fluid (Millipore) in a 6-well plate (Corning). Sperm were incubated in HTF for 30 min in a CO_2 incubator at 37°C to reduce somatic contaminants. Then, the supernatant with swimming sperm was transferred to a new well. The incubation-transfer was repeated 4–5 times. Sample quality was assessed by microscopy before snap freezing. The purity of sperm was over 99%. Samples for RNA isolation were collected in Buffer RLT (RNeasy Micro Kit, QIAGEN).

In situ Hi-C library generation using low cell numbers

The generation of Hi-C libraries with a low number of cells was optimized according to a previous protocol (Rao et al., 2014). Samples were fixed with 100 μL of freshly made 1% formaldehyde solution and incubated at room temperature for 10 min. To quench the reaction, 2.5 M glycine solution was added to a final concentration of 0.2 M. Samples were incubated at room temperature for 5 min and then centrifuged for 5 min at 3,000 *g* at 4°C. Supernatants were discarded and samples were lysed in 100 μL of ice-cold Hi-C lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2% Igepal CA630) with a protease inhibitor cocktail on ice for 15 min. Samples were then centrifuged at 3000 *g* for 5 min and the supernatants carefully discarded. Pelleted nuclei were washed once with 100 μL of 1x NEBuffer 2. Discard the supernatant and remain 4.5 μL, and add 0.5 μL of 5% sodium dodecyl sulfate (SDS). Tubes were gently tapped to mix the pellet and were incubated at 62°C for 5 min. After incubating, 14.5 μL of water and 2.5 μL of 10% Triton X-100 were added to quench the SDS. Tubes were gently tapped to mix well, avoiding excessive foaming and then incubated at 37°C for 15 min. 2.5 μL of 10X NEBuffer 2 and 10 U of MboI restriction enzyme (NEB, R0147) were added and chromatin was digested at 37°C for 5 hr. Samples were incubated at 62°C for 20 min to inactivate MboI and then cooled to room temperature. To fill in the restriction fragment overhangs and mark the DNA ends with biotin, 5 μL of fill-in master mix (3.75 μL of 0.4 mM biotin-14-dATP, 0.45 μL of 10 mM dCTP/ dGTP/ dTTP mix, 0.8 μL of 5 U/μL DNA polymerase I, large) was added. Samples were mixed by pipetting and incubated at 37°C for 30 min. Ligation master mix (66.3 μL of water, 12 μL of 10X NEB T4 DNA ligase buffer, 10 μL of 10% Triton X-100, 1.2 μL of 10 mg/ml bovine serum albumin, 1 μL of 400 U/μL T4 DNA ligase) was added and samples were incubated at 16°C for 18 hr. Nuclei were pelleted by centrifugation for 5 min at 3000 *g* and were washed with 100 μL of 10 mM Tris buffer, pH 8.0. Pellets were then resuspended in 50 μL of 10 mM Tris buffer with 2 μL of 20 mg/ml proteinase K and incubated at 65°C for 18 hr. Proteinase K was inactivated by incubation at 75°C for 30 min. To make biotinylated DNA suitable for high-throughput sequencing using Illumina sequencers, DNA samples were sheared to a length of ~400 bp. Fifty 50 ng of carrier RNA was added and sheared DNA was transferred to a fresh PCR tube, then treated with the End Repair/dA-Tailing Module (NEB, E7442L) and Ligation Module (NEB, E7445L) following the operation manual. Samples were prepared for biotin pull-down by washing with 50 μL of 10 mg/ml Dynabeads MyOne Streptavidin T1 beads (Life technologies, 65602) in 100 μL of 1X Tween Washing Buffer (1X TWB: 5 mM Tris-HCl, pH 7.5; 0.5 mM EDTA; 1 M NaCl; 0.05% Tween 20). Samples were separated on a magnet and the solution was discarded. Ligation products were mixed with 83.5 μL of 2X Binding Buffer (2X BB: 10 mM Tris-HCl, pH 7.5; 1 mM EDTA; 2 M NaCl) and centrifuged at 1,0000 *g*, 1 min. Beads were resuspended with supernatant and incubated at room temperature for 30 min with rotation to bind the biotinylated DNA to the streptavidin beads. Samples were again separated on a magnet and the solution was discarded. Beads were washed by adding 200 μL of 1X TWB and tubes were heated on a Thermomixer at 55°C for 2 min with mixing. Beads were repelleted using a magnet and the supernatant was discarded. This wash step was repeated twice. Beads were then resuspended in 100 μL of 10 mM Tris buffer and transferred to a new tube. Beads were repelleted and the buffer discarded. Beads were resuspended in 20 μL of 10 mM Tris buffer and boiled at 98°C for 10 min. Beads were again repelleted and the supernatant were retained. The Hi-C library was amplified for 10 cycles of PCR with Q5 master mix (NEB, M0492L) following the operation manual. PCR products were confirmed by analyzing 1 μL of product using the FlashGel™ System (Lonza, 57063). PCR was continued with additional cycles until bright DNA bands were seen. A bottle of Agencourt AMPure XP beads (Beckman Coulter, A63881) was warmed to room temperature and gently shaken to resuspend the magnetic beads. One hundred μL of beads was added to 200 μL of diluted PCR product (0.5X volumes). Samples were mixed by pipetting and

incubated at room temperature for 10 min. Beads were pelleted on a magnet and the clear solution was transferred to a new tube. Another 30 μL of beads was added to the clear solution (0.65X volume), mixed by pipetting, and incubated at room temperature for 10 min. Keeping the beads on the magnet, samples were washed twice with 200 μL of 70% ethanol (freshly made) without mixing. Ethanol was then completely removed. Beads were left on the magnet for 5 min to allow the remaining ethanol to evaporate. DNA was eluted by adding 20 μL of ddH_2O, mixing by pipetting, and incubating at room temperature for 5 min. After separating on a magnet, the solution was transferred to a new tube. DNA was then quantified and sequenced using an Illumina sequencing platform.

4C-seq library generation using a low number of cells

To validate the Hi-C results, we generated 4C libraries with a low number of cells. We used about 50,000 cells to generate the libraries. We carried out two rounds of digestion and ligation. The primary digestion enzyme was *Hind*III. Reversal crosslinked DNA was purified with Agencourt AMPure XP beads and then digested with the second restriction endonuclease, *Dpn*II, in a 25 μL reaction volume (2.5 μL of 10x *Dpn*II restriction buffer; 2.5 U of *Dpn*II) at 37°C for 5 hr. Reactions were inactivated at 65°C for 30 min. The second ligation was then performed in a 150 μL reaction volume (15 μL of 10x ligation buffer; 0.5 μL of 400 U/ μL T4 DNA ligase) and incubated at 16°C overnight. Six μL of 20 mg/ml proteinase K was added and samples were incubated at 65°C for 5 hr, then 75°C for 30 min to inactivate the proteinase K. DNA was purified with Agencourt AMPure XP beads. Primer sequences are listed in Table S4.

Profiling cell cycle phase by using propidium iodide immunofluorescence staining

Mouse embryos at different stages were fixed with 4% paraformaldehyde in PBS for 15 min and permeabilized with 0.5% Triton X-100 in PBS for 15 min at room temperature. DNA was stained with propidium iodide (PI) and then analyzed with an Arrayscan VTI 700 instrument (Thermo Scientific). The fluorescence intensity of PI was quantified by HCS Studio (Thermo Scientific) software to indicate the relative DNA quantity of cells, which was used to estimate the cell cycle phase.

BrdU staining

Late zygotes were treated with α-Amanitin or Aphidicolin. A 100-μM 5-bromo-deoxyuridine (BrdU) solution was added within one hour after the first cell division. For the detection of BrdU-labeled DNA, embryos were fixed with 4% paraformaldehyde for 15 min and permeabilized for 15 min in 0.5% Triton X-100 in PBS. Embryos were then treated with 2 M HCl for 15 min at room temperature and blocked with 1% BSA, 0.05% Tween 20 in PBS for 1 hr. Anti-BrdU antibodies (Abcam) were detected by Alexa Fluor-488-conjugated secondary antibody. DNA was stained with Hoechst and embryos were mounted in 50% glycerol in PBS.

QUANTIFICATION AND STATISTICAL ANALYSIS

Hi-C data processing, mapping, and ICE normalization

For Hi-C data, we first trimmed the raw data to remove adaptor sequences and low-quality reads (Bolger et al., 2014). Then, we mapped the paired-end Hi-C reads to the mouse genome (mm10) using HiCUP (Wingett et al., 2015) (http://www.bioinformatics.babraham.ac.uk/projects/hicup/) pipeline v.0.5.7 with parameters (–longest 800–shortest 150). After HiCUP mapping, we discarded the same fragment reads, re-ligation reads, and continuous reads. Unique mapped reads (mapping quality $q > 10$) (Li et al., 2009) separated by a genomic distance greater than 10 kb were used for subsequent analysis. PCR duplication was removed by Picard Tools (DePristo et al., 2011). We generated the raw contact matrices at the binning resolution of 40 kb, 100 kb, and 250 kb. The ICE (Imakaev et al., 2012) normalization was applied to remove bias in the raw matrix using the ICE implemented in HiTC with parameters (–sparse. filter 0.02) (Servant et al., 2012). To compare the interaction frequencies among different development stages, interaction matrices should be normalized to remove the effect of sequencing depths at different stages. Thus, each row sum of interaction I_{ij} (*interacton between locus i and locus j*) in the normalized matrix was scaled to 1,0000.

Pearson correlation coefficients analysis

To validate the reproducibility of our data, we calculated the Pearson correlation coefficients (PCC) between two libraries as follows: The set of all possible interactions, I_{ij}, for two libraries, A and B, were correlated by comparing each point's interaction profiles in normalized interaction matrix, I_A, from library A with the same point, I_B, from library B. Because the interaction matrix is highly skewed toward proximal interactions, we restricted the interaction to a maximum distance of 4 Mb between points i and j. We used R functions to calculate the PCC between the two vectors of all points in I_A and I_B. R package *ape* was used for sample clustering with 1,000 bootstraps (Paradis et al., 2004) (Figure S1).

TAD signal calculation and relative TAD signal variance

To calculate the dynamics of TAD structures, we calculated TAD signal. The TAD signal value can indicate the strength of the TAD structure. We used intra-chromosomal maps at 40 kb resolution for each stage. The TAD signal was calculated as the log2 ratio of the number of corrected upstream-to-downstream interactions within a 2 Mb region. If a region contained less than 10 counts within a 2 Mb distance, the region was filtered. TAD signal variance was only calculated for the loci without filtered-out upstream and downstream regions within a 2 Mb region. Comparisons between different stages were performed only for the bins shared by all of the

different stages. Relative TAD signal variance was calculated relative to the minimum TAD signal variance in zygotes. The Wilcoxon test was applied to test statistical significance.

To determine the effect of sequencing depth on the TAD signal variance, we performed a downsampling analysis by a downsampling factor of 0.75. Variance due to sampling noise was confirmed to be proportional to 1/reads.

Insulation score analysis

Insulation score can be used to calculate TADs (Giorgetti et al., 2016). We used the public code to calculate the insulation score on Github (matrix2insulation.pl; https://github.com/dekkerlab/giorgetti-nature-2016). To calculate the insulation score of each bin in the 40 kb, the average number of interactions that occurred across each bin was calculated. This can be visualized by sliding a 480 kb × 480 kb square along the matrix diagonal. The IQR mean signal within the square was then assigned to the 40 kb diagonal bin. This procedure was then repeated for all 480 kb diagonal bins. The insulation score was normalized relative to all of the insulation scores across each chromosome by calculating the log2 ratio of each bin's insulation score versus the mean of all insulation scores. Valleys/minima along the normalized insulation score vector represent the loci of reduced Hi-C interactions that occur across the bin. These valleys/minima are interpreted as TAD boundaries or areas of high local insulation. We compared the insulation score at different stages. We used the Kruskal-Wallis test followed by Dunn's multiple comparison test to test statistical significance.

Averaged contact probability P(s)

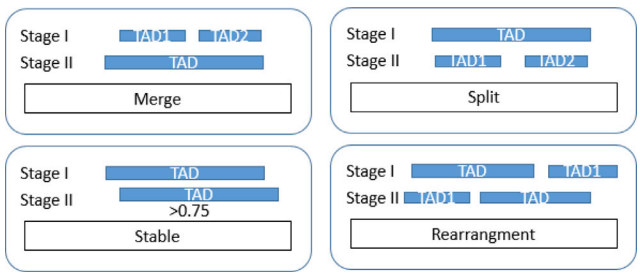
To calculate averaged contact probability (Ps), we first divided the genome into 40 kb bins. For each distance (separated by 40 kb, 80 kb, 120 kb …), we used the observed interaction frequency to calculate the expected value by LOWESS fitting at the respective distance (Naumova et al., 2013). Averaged contact probability decreases with genome distance increases (Figure 1C, log-log axis). For oocytes, we also calculated the slope of the P(s) linear fitting curve from 500 kb to 7 Mb, P(s) $\sim s^{-0.492}$, as shown in Figure 1C.

Density of paired loci at different insert sizes

We examined the relative abundance of paired loci at different insert sizes (Figure 1E). In probability theory, a probability density function (PDF) or density of a continuous random variable, is a function that describes the relative likelihood for this random variable to take on a given value. We assigned the unique paired reads to restriction enzyme fragments and plotted the distribution of the $\log_{10}$ of insert size between fragments.

TAD calling, TAD merge, and split between gametes and embryos

We used the DI (Directional Index) value to calculate TADs. First, we calculated the DI value for each bin based on the ICE-normalized and depth-normalized matrix. Then the DI value was taken as the input of the Hidden Markov Model (HMM) model to call TADs. The numbers of identified TADs across five samples were 1,810, 1,972, 2,004, 2,015, and 2,188 for sperm, 4-cell, 8-cell, E3.5, and E7.5, respectively. For the pairwise TAD reprogramming analysis, we used bedtools (Quinlan and Hall, 2010) to calculate the proportion of overlapped TADs between two adjacent stages. If two or more TADs in a stage fused into one TAD in another stage, we called this change a 'merge' (See the illustrated figure below). If one TAD in a stage divided into 2 or more TADs in another stage, we called this change a 'split'. Besides the TAD 'merge' and 'split', if the overlapped proportion of the TAD in two developmental stages was above 0.75, the TAD between these 2 stages was considered stable. The remaining situations were defined as TAD rearrangement or boundary shift.



Inter-TAD interactions analysis

The analyses for inter-TAD interactions were performed between sperm, 8-cell, E3.5, and E7.5 embryos, because TAD number and TAD size were comparable in these stages. Inter-TAD interactions refer to the interactions between TADs from the same

chromosome. We only compared the interactions between those TADs that maintained stability across sperm, 8-cell, E3.5, and E7.5 embryos. The relative frequency of inter-TAD interactions was calculated from the number of interactions between 2 TADs divided by the total number of cis-interactions in the corresponding chromosome. We plotted the relative inter-TAD interaction frequency at different genomic distances (Figure 2D). We also calculated the proportion of interactions for inter-TADs separated by a certain genomic distance within each chromosome (Figure S3C). Statistical significance was tested by performing the Wilcoxon test.

Relative inter-chromosomal interaction frequency

The relative inter-chromosomal (or trans-interaction) frequency between 2 trans bins was defined as the number of interactions between those 2-bins divided by the total number of interactions (cis-interactions plus trans-interactions) in one stage. We used a 1,000 kb resolution for trans-interaction analysis. The formula is below:

$$F_{Sn\ (Bi,\ Bj)} = C_{Sn(Bi,\ Bj)} / T_{Sn(\text{total cis} + \text{total trans})}$$

F refers to the relative trans-interaction frequency. Sn refers to stage n. (Bi, Bj) refers to interactions between bin i and bin j. C refers to the normalized interaction count. T refers to the total number of interactions (cis-interactions plus trans-interactions) in one stage.

We then calculated the average frequency of trans-interactions at different stages. The formula is below:

$$\text{Average} = \frac{\sum_{n=1}^{N} \sum_{j=1}^{Jn} \sum_{i=1}^{In} F_{Sn(Bi,\ Bj)}}{\sum_{n=1}^{N} (I_{Sn} \times J_{Sn})}.$$

We assigned the interactions between 2 trans-bins as the putative trans-interactions only if $F_{Sn\ (Bi,\ Bj)}$ is > 2 × Average. The same threshold was used across all stages. We plotted the distribution of the putative interactions in each stage (Figure 2F). Our results show that the average trans-interaction frequency in sperm is higher than in other stages after analysis by Wilcoxon test. Heatmaps were displayed using Java TreeView (Saldanha, 2004).

A/B Compartment PC1 value and status switch

We used HOMER (Heinz et al., 2010) software with parameters (–res 25,000, -superRes 1,00000) to obtain the PC1 value and A/B compartment status. We used a sliding window approach with a bin size of 100 kb and a step size of 25 kb to generate an observed/expected matrix. The observed interaction was calculated as the sum of all observed interactions of the 25 kb bins making up the larger 100 kb bin. Similarly, the expected interaction was calculated as the sum of the expected interactions of each of the 25 kb bins making up the larger 100 kb bin. This value was used to generate the observed/expected value. PC1 for each chromosome was used to identify regions of the genome as belonging to either the A or B compartment. The direction of the Eigen values is arbitrary, and therefore positive values were set to 'A' and negative values were set to 'B' based on their association with gene density. We identified regions with statistically significant variability in PC1 values across all developmental stages using ANOVA with p value < 0.05 as A/B compartment switched regions.

Sperm MNase data processing and MNase fragment score

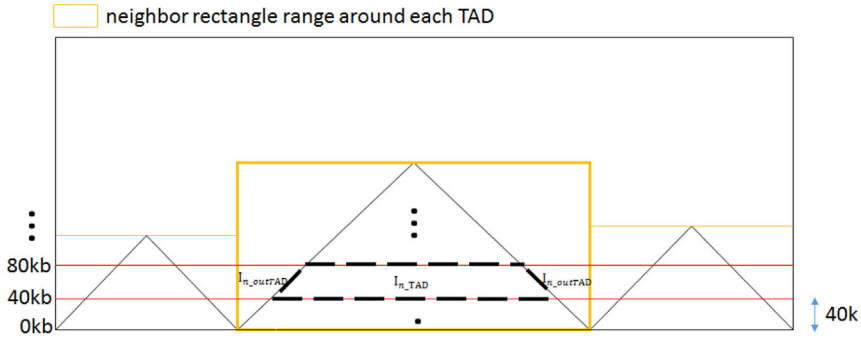
We investigated the association between nucleosomes and 3D structures in the sperm genome. We used published sperm MNase data (Carone et al., 2014) to obtain the nucleosome signal in sperm (GSE58101). Data were mapped to mm10 using bowtie2 (Langmead and Salzberg, 2012) with default parameters. Picard Tools (DePristo et al., 2011) were used to remove PCR duplication in MNase data. We only kept the reads with insert sizes between 135 bp and 165 bp, which is the typical size of the nucleosome. Next, we calculated the fragment score, which is associated with nucleosome signal. The MNase fragment score is analogous to RPKM (reads per kilobase per million). Briefly, for each restriction enzyme (MboI) fragment, we counted the reads from this fragment and normalized the count by using the RPKM method. Thus, this fragment score indicates the nucleosome signal. Then, we ranked restriction enzyme fragments based on their fragment score. In Figure S3F, 0%–10% indicates the fragments with top the 10% of fragment scores.

Identification of concordant genes with A/B compartment switch

To define concordance in gene expression and compartment status, we used a method similar to one previously described (Dixon et al., 2015). We calculated the covariance between the vector of the log2 of gene expression values and the vector of PC1 values for each gene across different stages. We used the calculated covariance as a metric to quantitatively define 'concordance'. To calculate a P value for the covariance for each gene, we compared these observed covariance values to a random background distribution. The background distribution was generated by randomly shuffling the vector of log2 of gene expression for each gene and then calculating the covariance between the random gene expression vector and the PC1 values. This was repeated 1,000 times for each gene, and a rank-based P value could then be calculated for the observed covariance values. Genes with a P value < 0.05 were defined as concordant genes.

Relative TAD intensity (RTI)

To compare the TAD (integrity) dynamics in gametes and early embryos, we calculated a relative TAD strength score (see the illustration in the figure below). In brief, we chose a neighbor rectangle range around each TAD where the TAD was embedded (the yellow rectangle in the picture below).



I_{n_TAD} : the frequency vector within TAD in the neighbor rectangle at distance n
I_{n_outTAD} : the frequency vector outside TAD in the neighbor rectangle at distance n

At a certain genomic distance, n, we have two frequency vectors. One is I_{n_TAD}, which is the frequency vector within the TAD in the neighbor rectangle at distance n (as shown by the dashed box in the picture). The other one is I_{n_outTAD}, which is the frequency vector outside the TAD in the neighbor rectangle at distance n (as labeled in the picture). The formula for the relative TAD intensity (RTI) is below:

$$RTI = median\left(\frac{median\left(I_{n_{TAD}}\right)}{median\left(I_{n_{outTAD}}\right)}\right).$$

A larger RTI indicates more interactions within the TAD compared to the neighbor region outside the TAD, indicating the TAD structure is clearer.

Domain boundary ChIP-seq signal enrichment analysis

Domain boundary profiles were identified by using the boundaries' centers as anchors with binning the ChIP-seq signals in 10 kb bins ± 500 kb up and downstream of the anchor. Results were plotted with deepTools2 (Ramírez et al., 2016). Boundaries were collapsed to the mean of the adjacent boundaries. All scores were sorted and the top 5% of scores were truncated. ChIP-seq signal was identified as the fold enrichment of the treatment signal versus the responding input sample (Zhang et al., 2008). Mouse sperm and embryo histone modification data are from public datasets (Bilodeau et al., 2009; Dahl et al., 2016; Jung et al., 2017; Liu et al., 2016; Shen et al., 2012; Zhang et al., 2016; Zheng et al., 2016).

Methylation and DHS enrichment analysis

Methylation data (GSE56697) (Wang et al., 2014) and DHS data (GSE76642) (Lu et al., 2016) were reanalyzed and overlapped with A/B compartments in early embryos using bedtools 2.25.0 (Quinlan and Hall, 2010). We found that unmethylated CpGs and stage-gained DHSs are enriched in A compartments.

4C-seq bioinformatics

To validate the interaction pattern observed in Hi-C data, 4C-seq analyses were performed at viewpoint 'a' near *Hoxb13* and viewpoint 'b' near *Hoxb7*. We discarded the reads that did not match the primary restriction enzyme sequence or the primer sequence. Primer sequences were trimmed from the reads and the reads were then mapped to the mouse reference genome (mm10) using bowtie2. We removed undigested and self-ligated fragment reads. To compare sperm and E7.5 embryo interactions, we used the same number of cis-reads in the analyses. To visualize the interactions, deepTools *bamCoverage* was used to generate tracks, which were displayed in the WashU Epigenome Browser (Zhou et al., 2013).

RNA-seq data processing

The RNA-seq data for oocytes, zygotes, 2-cell, 4-cell, and 8-cell embryos were obtained from a previous publication (Wu et al., 2016) (GSE66390). Sperm RNA-seq data was from GSE49622 (Hammoud et al., 2014). E3.5 and E7.5 embryo RNA-seq libraries were generated in this study by Smart-seq2. All RNA-seq data were mapped to the mm10 genome by STAR (version 2.5.2b) (Dobin et al., 2013) with default parameters. Gene expression FPKM was calculated by Cufflinks (version 2.2.1) using the refFat database from the UCSC genome browser. Sequencing depth was normalized. Maternal and paternal mRNA reads were identified by the SNPs between maternal or paternal copies. Smart-seq2 RNA-seq libraries were generated according to a previous publication (Wu et al., 2016). Cells were lysed in hypotonic lysis buffer (Amresco, M334), and polyadenylated mRNAs were captured by polyT primers. After 3–10 min of lysis at 72°C, Smart-seq2 reverse transcription reactions were performed. After pre-amplification and purification with AMPure XP beads, cDNAs were sheared with a Covaris instrument and were subjected to Illumina TruSeq library preparation. All libraries were sequenced on an Illumina platform according to the manufacturer's instructions.

Hi-C allele-specific analysis

Early embryos were obtained from crosses between female C57BL/6J mice and male DBA/2J mice. To achieve allele-specific chromosome conformation, we used SNPs that differ between the two strains to identify the reads from the maternal or paternal genome. We used SNP split software to distinguish paternal and maternal reads (Krueger and Andrews, 2016). For cis-interactions, we kept the read pairs with at least one-end and assigned them to the maternal genome or paternal genome. We further compared the differences in compartment status between maternal and paternal genomes for early embryos. We then used a 25 kb resolution with a 100 kb sliding window to calculate PC1 values for data in Figure 7C. A 100 kb resolution with a 400 kb sliding window was used for data in Figure S7C.

Compartment strength

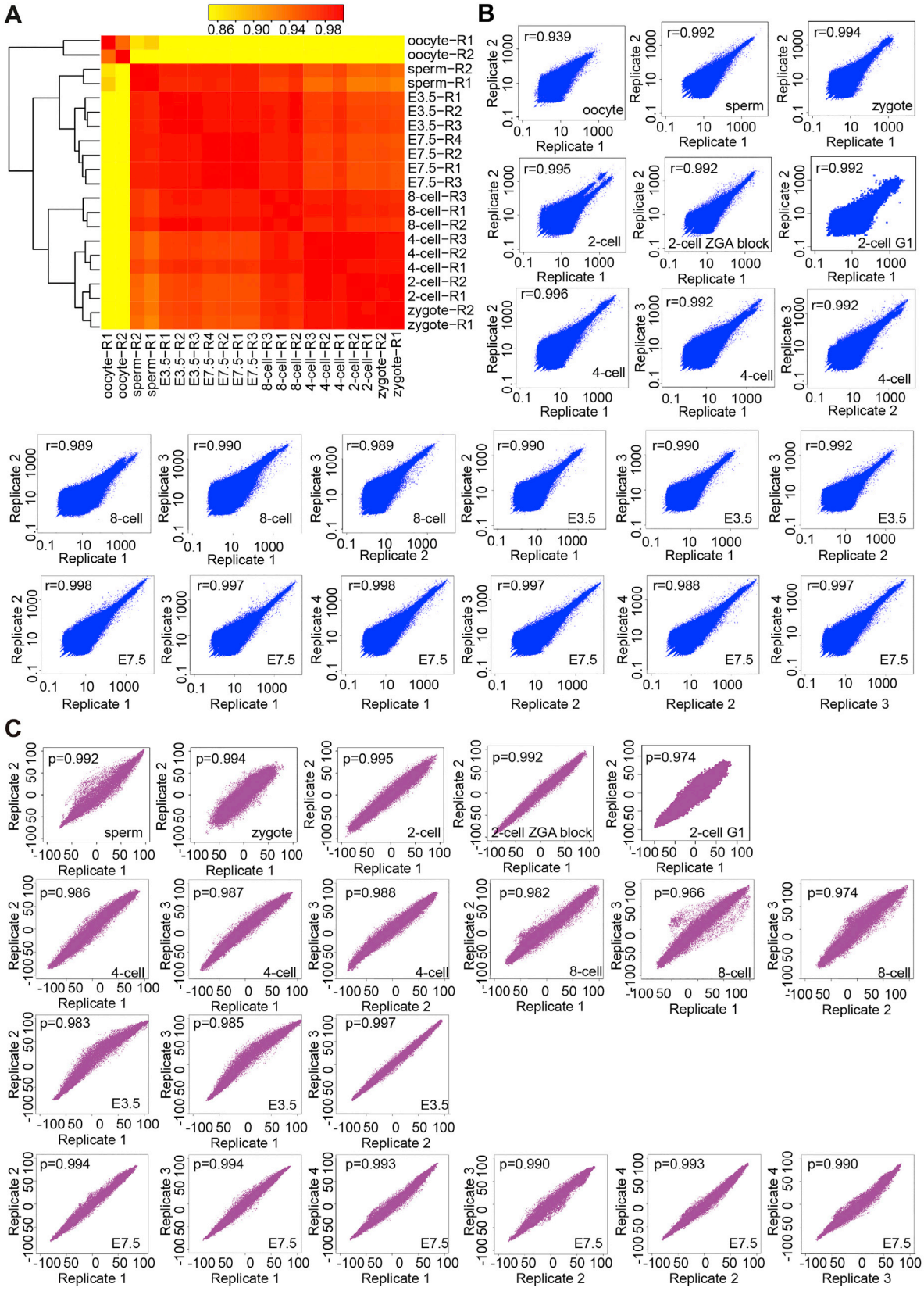
Compartment strength was calculated to compare the differences in compartment structures between different stages (Flyamer et al., 2017). To calculate the standard error, we created 100, 5 × 5 compartment enrichment matrices by bootstrapping. For each pixel of the 5 × 5 compartment enrichment map, we took all the observed-over-expected values that contributed to this pixel and took a random sample with the replacement of the same size as that of contributing values. We then proceeded downstream for each of the 100 reshuffled maps. To calculate the strength of compartment signal, we took the natural logarithm of $AA * BB / AB^2$. For bar plots, we chose the top 20% of eigenvector for A compartment, and the top 20% for B compartment.

Statistical methods

The Kruskal-Wallis test followed by Dunn's multiple comparison test was used to analyze data in Figure 1B and Figure S2C. We used the Chi-square test for data in Figure 2C, a one-sided Wilcoxon test for data in Figure 2D and Figure S3C, and the Kolmogorov-Smirnov (KS) test for data in Figures S2B and S3A. Data in the remaining figures were analyzed by Wilcoxon test.

DATA AND SOFTWARE AVAILABILITY

The accession number for the sequencing data reported in this paper is GSA: PRJCA000241.



(legend on next page)

Figure S1. Validation of Hi-C Data Quality, Related to Figure 1

(A) Heatmap shows the cluster classification for different Hi-C libraries according to their normalized interaction matrix Pearson correlation coefficient, restricted by distance shorter than 4 Mb.

(B) The correlation between Hi-C replicates for each stage according to normalized interaction frequency at 200 kb resolution. "r" indicates the Pearson correlation coefficient.

(C) The correlation between Hi-C replicates for each stage according to PC1 value at 40 kb resolution with 400 kb sliding window. "p" indicates the Pearson correlation coefficient.

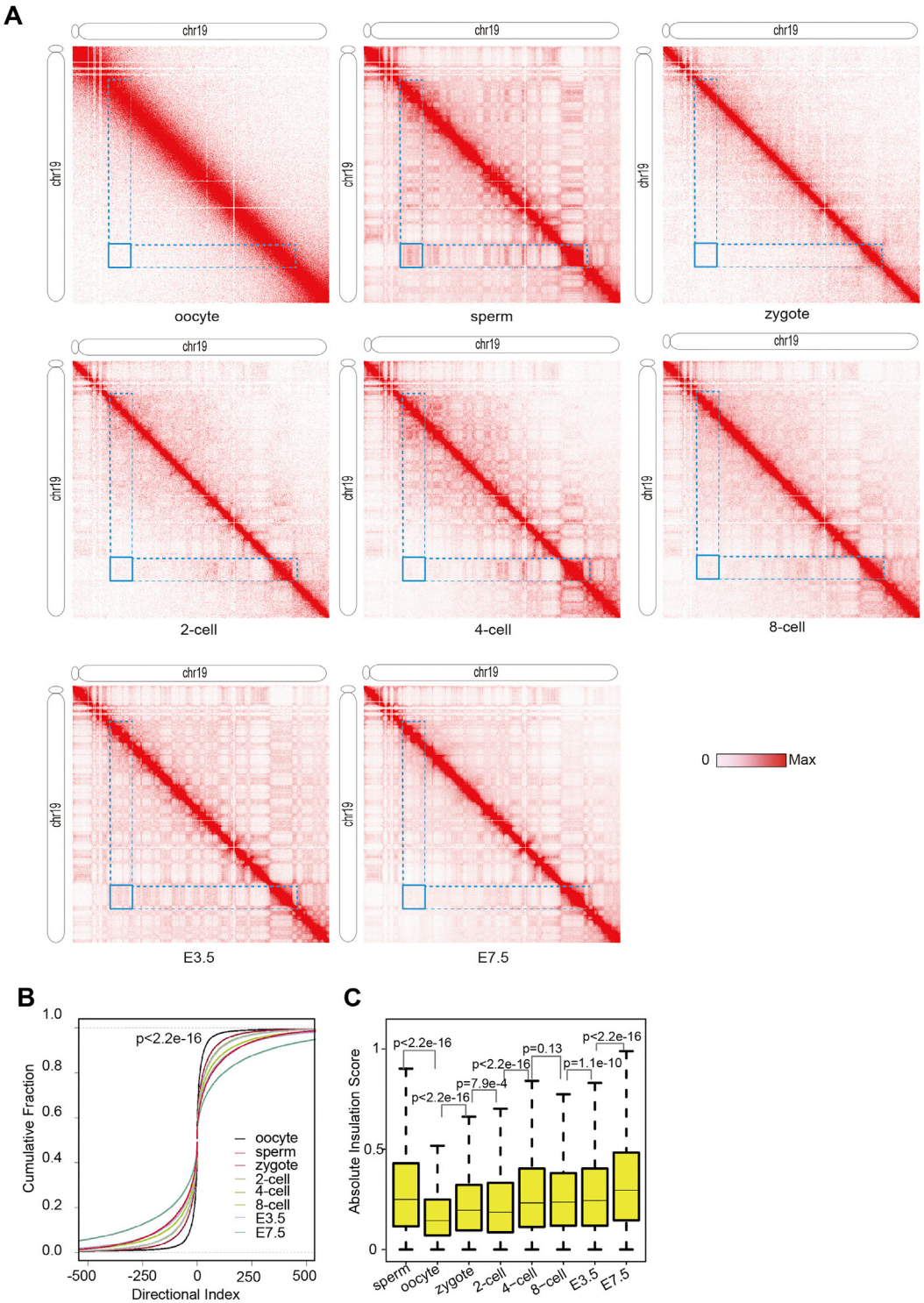
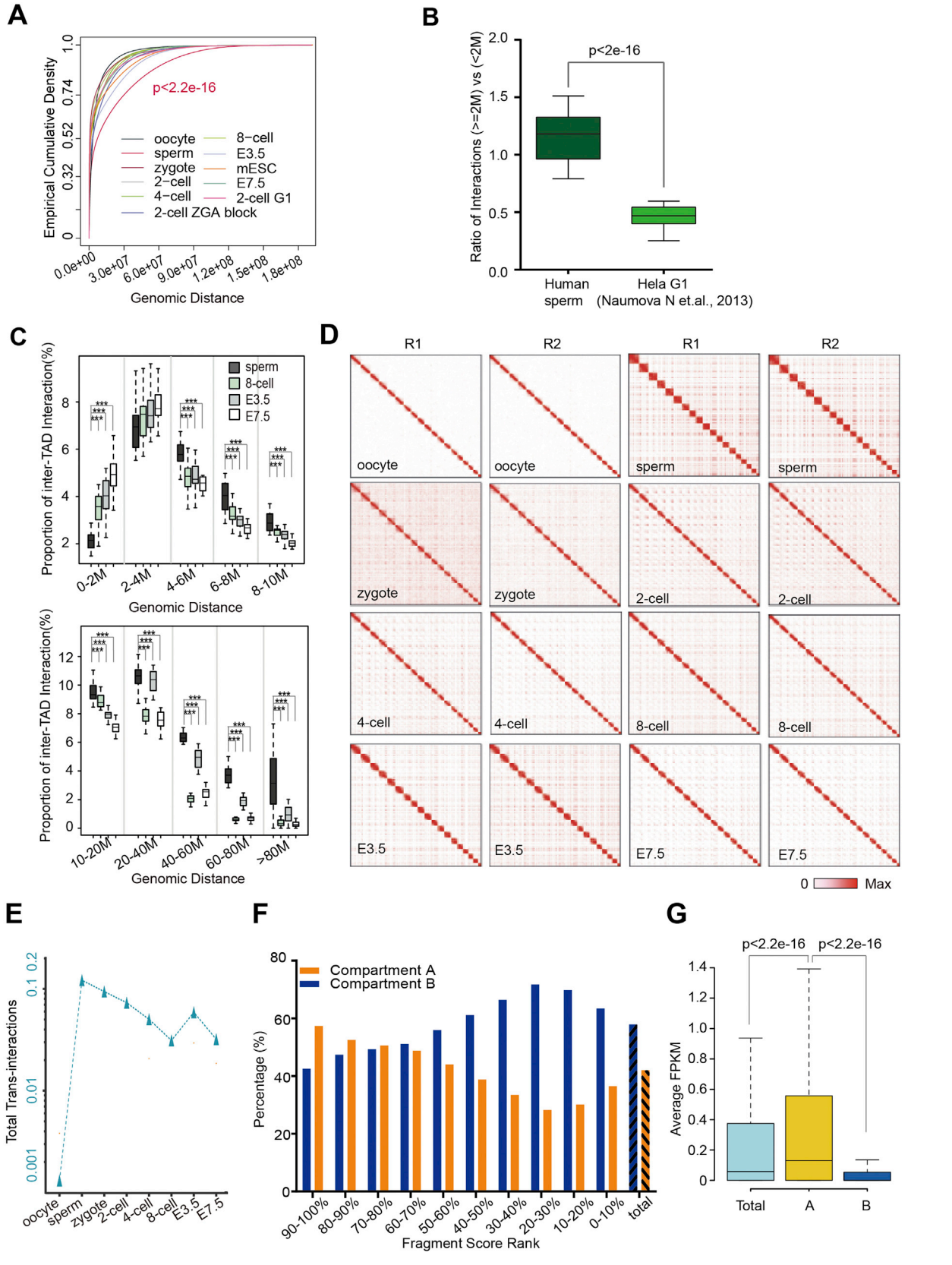


Figure S2. Oocyte Chromatin Structure, Related to Figure 1

(A) Normalized Hi-C interaction heatmaps at 40 kb resolution. Blue rectangles represent extra-long distance inter-TAD interactions, corresponding to Figure 2A. (B) The Empirical Cumulative Distribution plot for DI values across different stages. The bin whose upstream 50 bins or downstream 50 bins have no interactions, which caused gaps in the heatmap, has been filtered. Similar to Figure 1B, oocytes have lower DI values compared to other stages with the Kolmogorov-Smirnov (KS) test. (C) Boxplot of boundary insulation score across different stages. Statistical significance was determined by Kruskal-Wallis test, followed by Dunn's multiple comparison test.



(legend on next page)

Figure S3. Sperm Chromatin Structure, Related to Figure 2

(A) Empirical Cumulative Distribution Function analysis of paired loci interactions separated by different genomic distance. The y axis represents the accumulation percentage of the interactions. The x axis represents the genomic distance of the paired loci. Sperm is significantly different from other stages according to the Kolmogorov-Smirnov (KS) test. "2-cell G1" refers to Aphidicolin-treated 2-cell embryos. "2-cell ZGA block" refers to α-Amanitin-treated 2-cell embryos.

(B) The ratio of total interactions separated by genomic distance $\geq$ 2 Mb versus those < 2 Mb in human sperm and HeLa cells at G1 phase. We used the chromosome as the unit for this analysis. P value was calculated by Wilcoxon test.

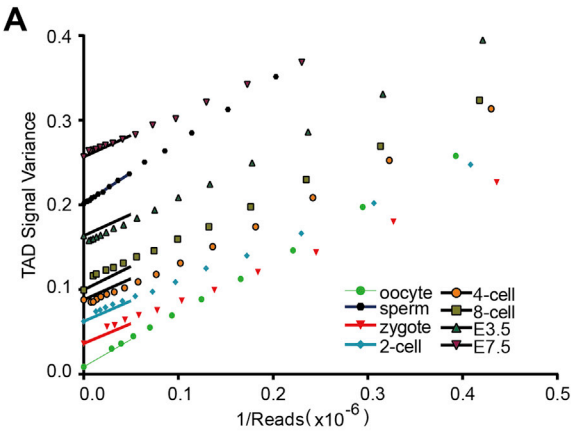
(C) Relative proportion was calculated from the amount of inter-TAD interactions separated by genomic distances within a certain window divided by the amount of total interactions. We used the chromosome as the unit for this analysis. A one-sided Wilcoxon test was used to test statistical significance. "***" refers to P value < 0.001.

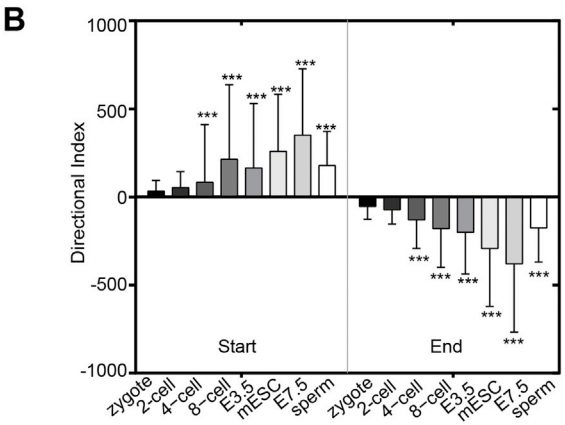
(D) Interaction frequency heatmaps show the interchromosomal interactions in all stages examined. Two replicates were analyzed for each stage. A two-dimensional heatmap indicates the normalized Hi-C interaction frequencies for 19 autosomes displayed at 1 Mb resolution.

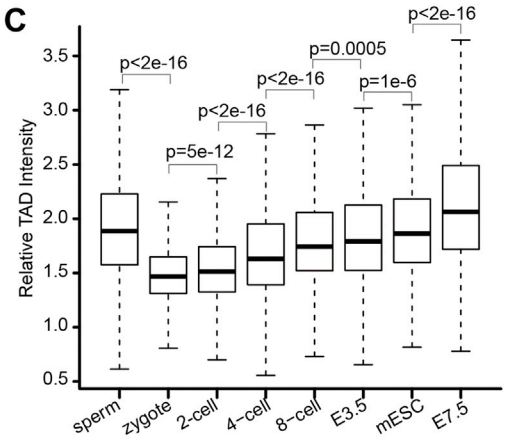
(E) The total inter-chromosomal interactions between two bins at different stages.

(F) Genomic fragments of high MNase read density tend to have compartment B status. Genomic fragments digested by MboI were grouped into deciles according to MNase read density.

(G) Expression comparison for genes in A or B compartments in sperm. Genes in A compartment tend to show higher expression. Statistical significance was determined by Wilcoxon test.

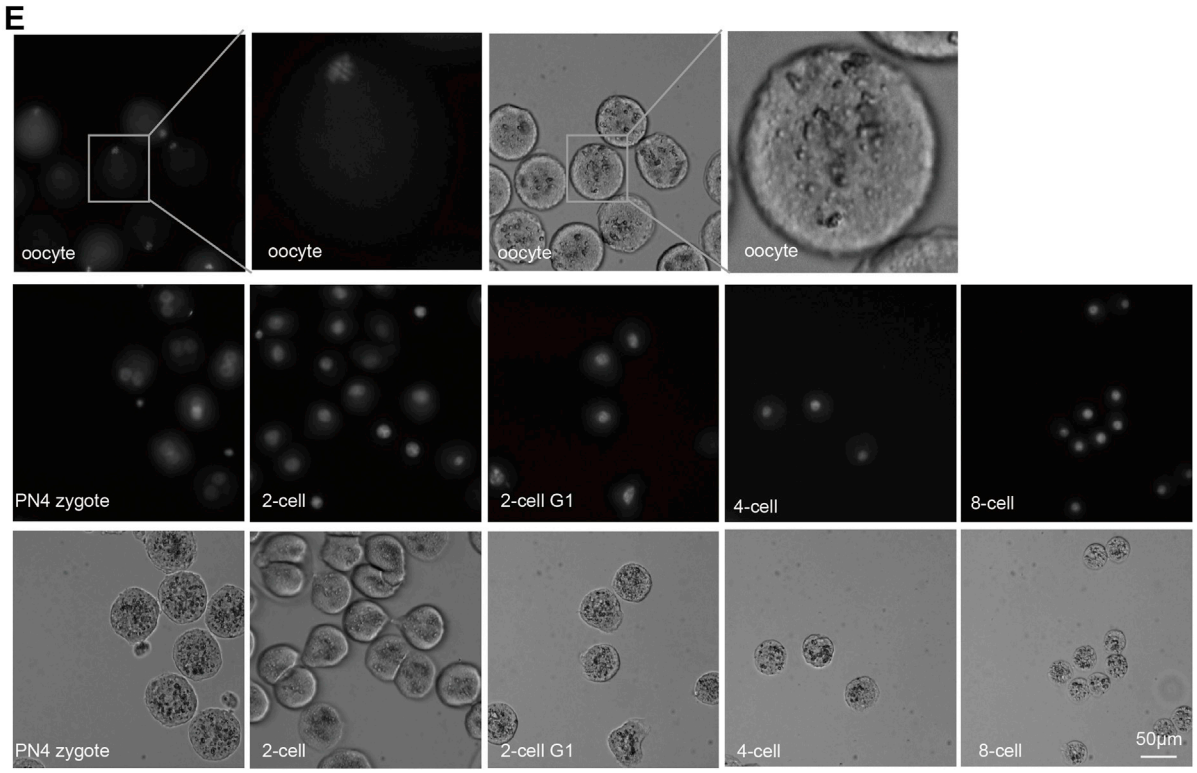






D

	cell number	Cells in M phase
oocyte	109	109
PN4 zygote	96	0
2-cell	94	1
4-cell	93	0
8-cell	90	0
2-cell G1	120	0



(legend on next page)

Figure S4. TAD Reprogramming in Early Embryos, Related to Figure 3

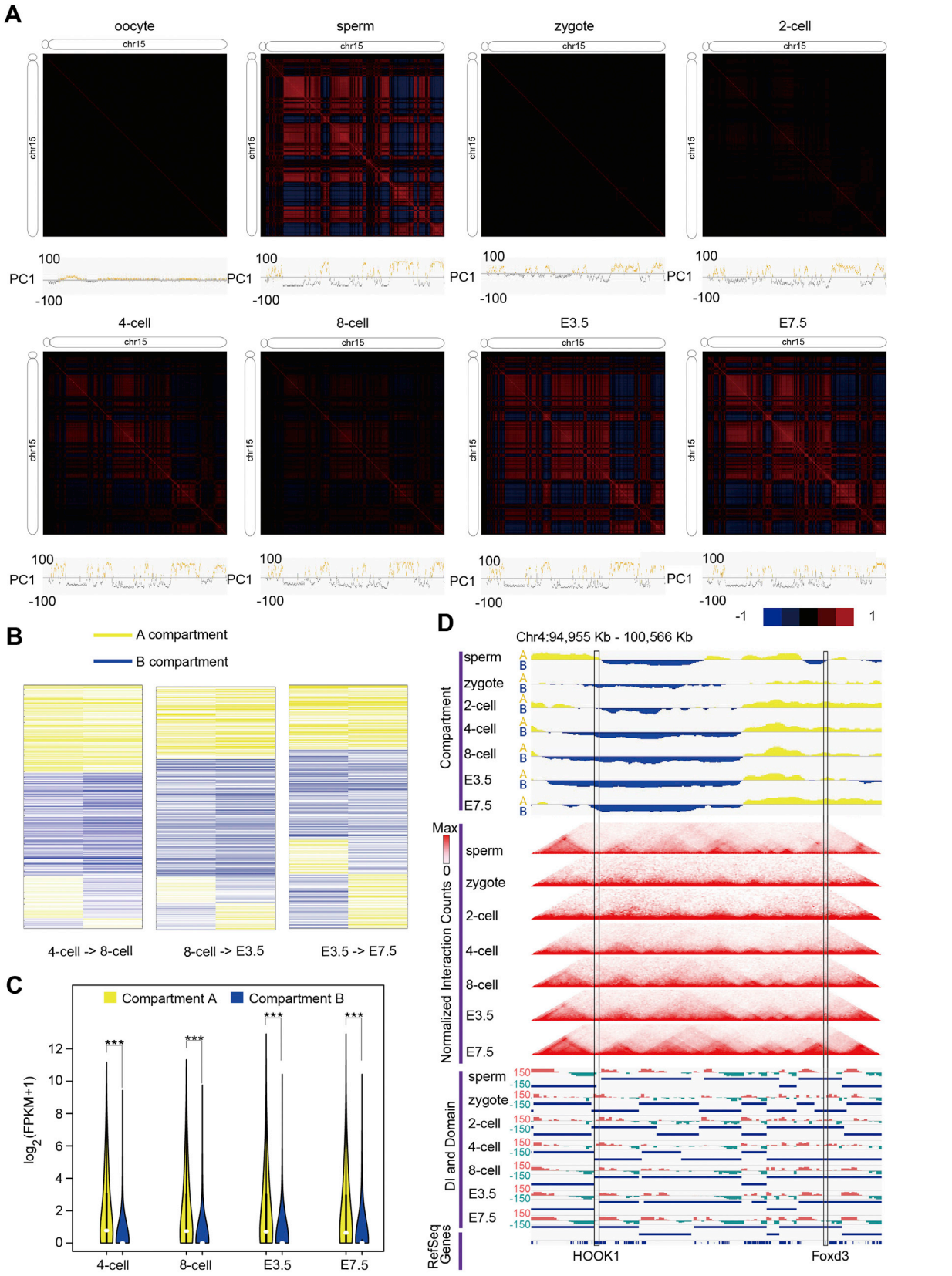
(A) Downsampling analysis for variance of the TAD signal for different developmental stages. Variance of the TAD signal has two contributing components: the variance of the underlying TAD signal and the variance due to sampling noise. Variance due to sampling noise is expected to be proportional to 1/reads. Our plot confirms this linear dependence and allows us to extrapolate and estimate the variance of the true TAD signal (for the case of very high coverage). Each line represents the indicated dataset. The far left side of each line represents the estimation of the TAD signal using all reads in an experiment. Reads were then downsampled by a factor of 0.75 at each progressively larger number of 1/reads. Oocyte extrapolation gives an estimation of the variance of the TAD signal approximately 10 times lower than in E7.5 embryos.

(B) The average values of DI at the TAD boundary bins across all stages. "***" indicates that the DI value in the specific stage is significantly higher than that in zygotes and 2-cell embryos, respectively ($p < 0.001$). P value was calculated by Wilcoxon test. Data are represented as mean ± SEM.

(C) Relative TAD intensity across all stages. P value was calculated by Wilcoxon test.

(D) The number of cells at metaphase at different developmental stages.

(E) PI staining and bright field images of cells at different stages.



(legend on next page)

Figure S5. Compartment Status in Gametes and Early Embryos, Related to Figure 4

(A) Correlation heatmaps for chromosome 15 at 40 kb resolution.

(B) A/B compartment status switching between two consecutive stages.

(C) Expression comparison for genes with A or B compartment status in 4-cell, 8-cell, E3.5, and E7.5 embryos. Genes in A compartments tend to show higher expression. Statistical significance was calculated by Wilcoxon test. "***" refers to P value < 0.001.

(D) Male gamete-related gene *HOOK1* shows A compartment status in sperm and B compartment status in early embryos. Developmental gene *Foxd3* shows B compartment status in sperm and A compartment status in early embryos.

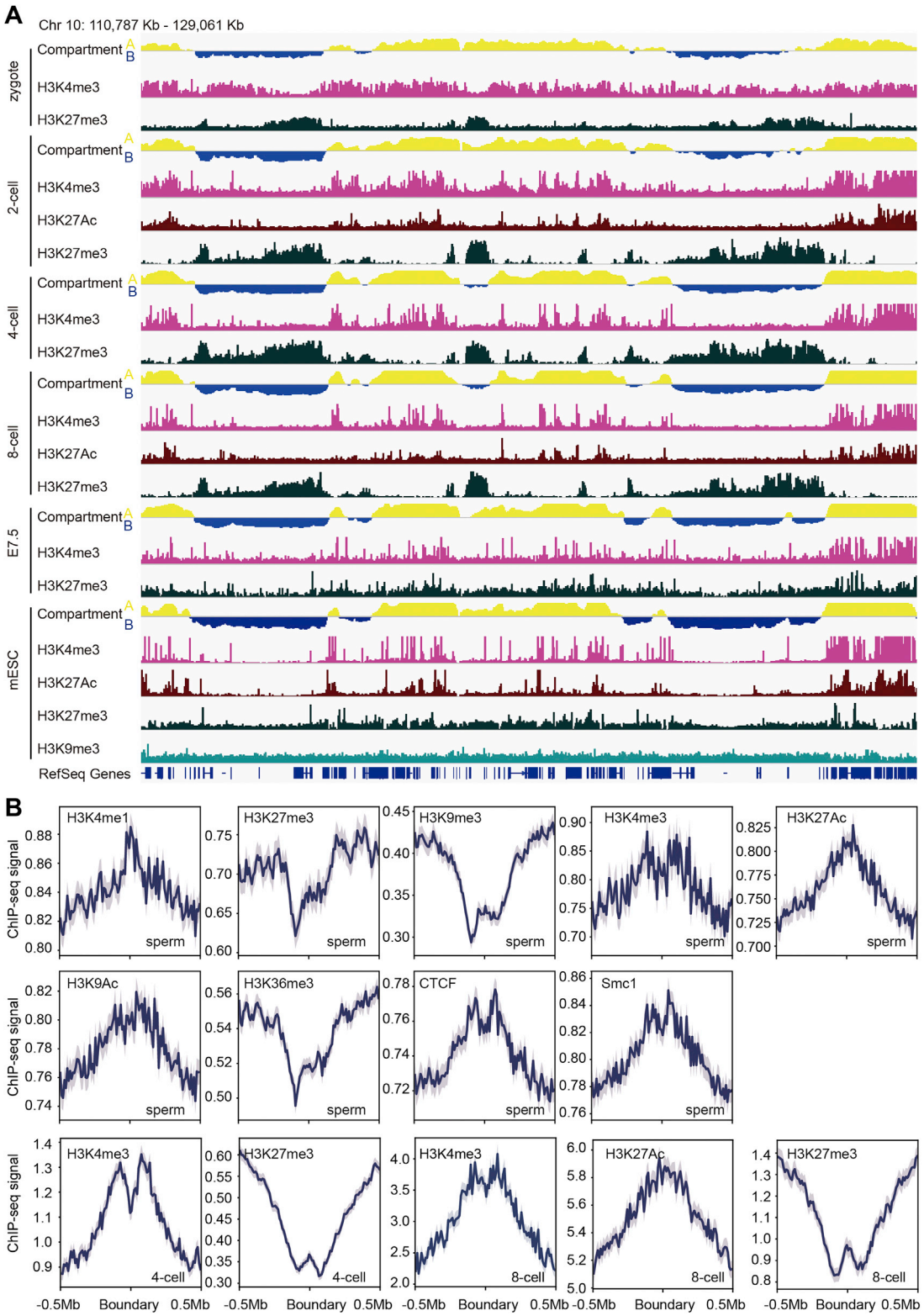
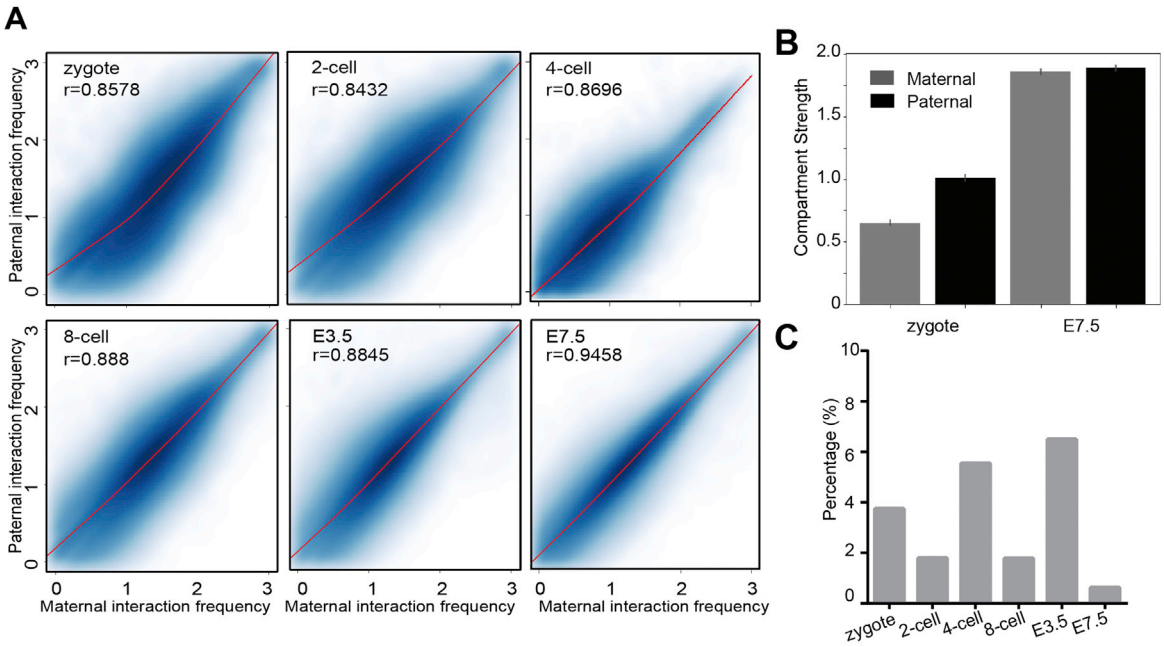
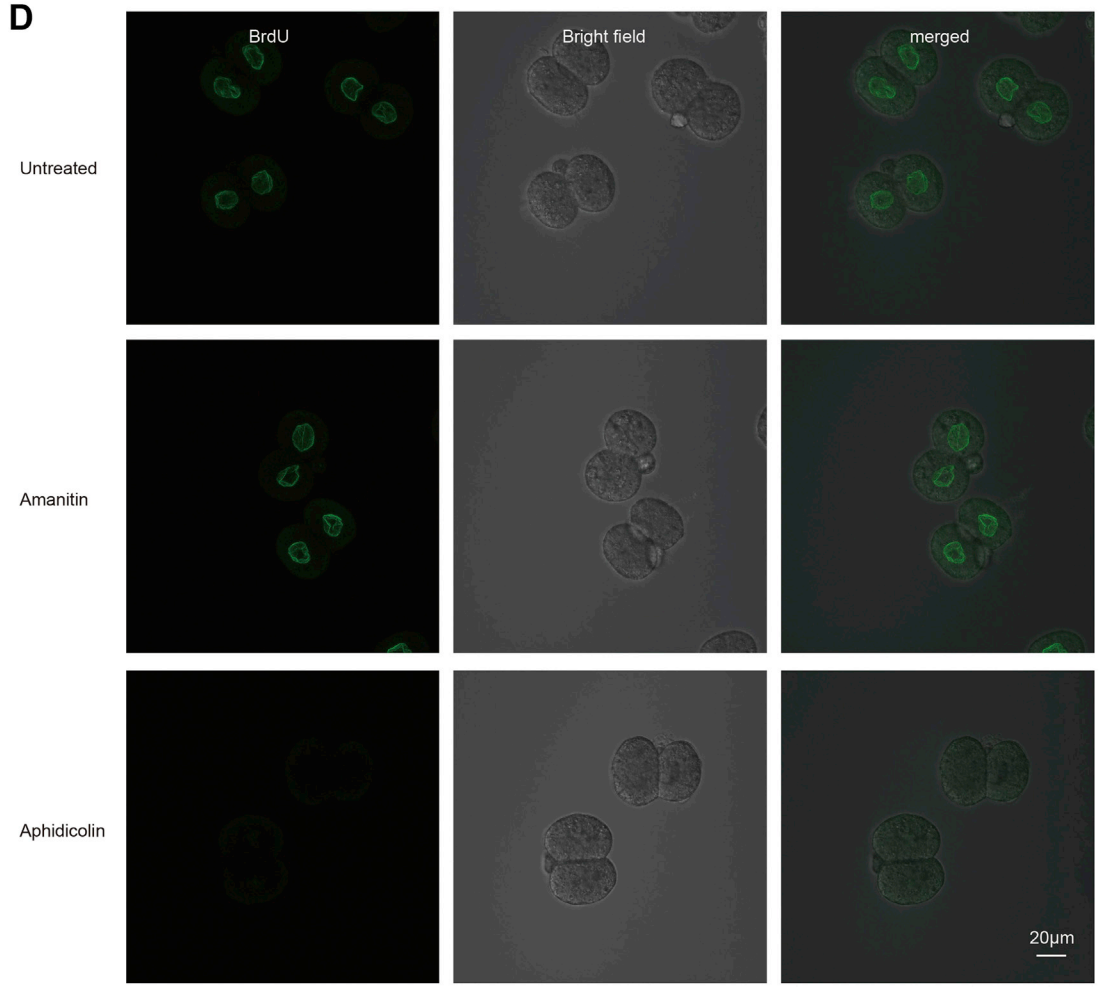


Figure S6. Association between Hi-C Data with Histone Modifications, Related to Figure 5

(A) Graph represents compartment status of a region overlayed by histone modification at different stages.

(B) Distribution plot of different histone modifications at TAD boundaries in sperm, 4-cell, and 8-cell embryos.





(legend on next page)

Figure S7. Allele-Specific Chromatin Structure and BrdU Staining Indicates DNA Replication in α-Amanitin-Treated Two-Cell Embryos, Related to Figures 6 and 7

(A) The correlation according to interaction frequency at 400 kb resolution between the paternal and maternal genomes for different stages.

(B) Compartment strength between the paternal and maternal genomes in zygotes and E7.5 embryos. Data are represented as mean ± SEM.

(C) The comparison of A/B compartment status between the maternal and paternal genome for early embryos.

(D) α-Amanitin or Aphidicolin was added in late stage zygotes and BrdU was added within one hour after the first cell division. Then, 2-cell embryos were cultured for about 12 hr after adding BrdU.