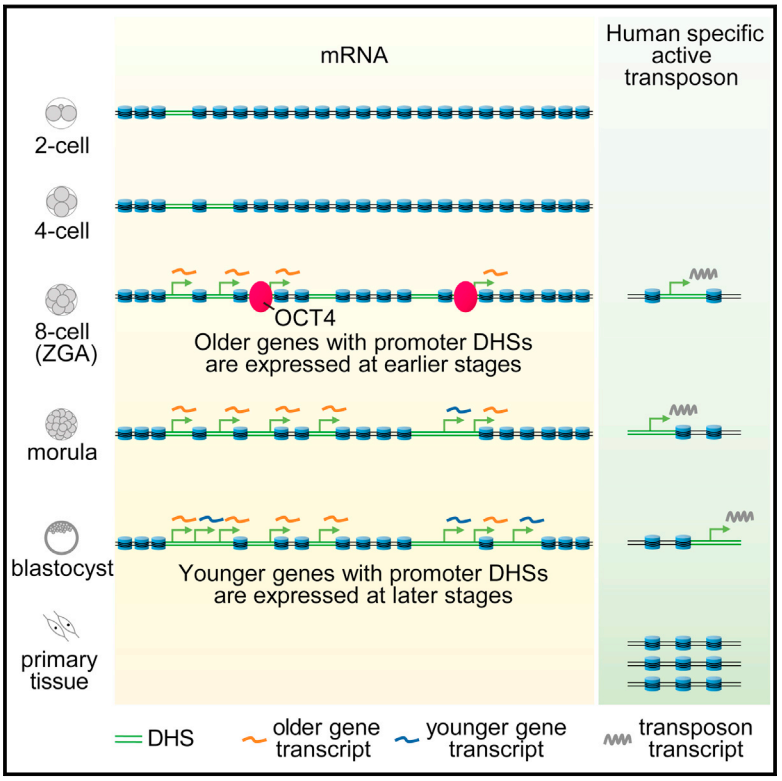
Resource

Cell

Chromatin Accessibility Landscape in Human Early Embryos and Its Association with Evolution

Graphical Abstract



Highlights

- The DHS landscape is gradually established during human embryo development
- OCT4 contributes to zygotic genome activation in humans, but not in mice
- Younger genes establish DHS at later stages, and older genes show the opposite trend
- Human transposons SVA/HERVK harbor DHSs and are specifically expressed in embryos

Authors

Lei Gao, Keliang Wu, Zhenbo Liu, ..., Jianqiao Liu, Zi-Jiang Chen, Jiang Liu

Correspondence

ljq88gz@163.com (J.L.), chenzijiang@hotmail.com (Z.-J.C.), liuj@big.ac.cn (J.L.)

In Brief

The dynamic landscape of open chromatin during early human embryogenesis provides a rich resource and insights into a zygotic genome activation.

Gao et al., 2018, Cell *173*, 248–259
March 22, 2018
https://doi.org/10.1016/j.cell.2018.02.028

CellPress

Cell

Chromatin Accessibility Landscape in Human Early Embryos and Its Association with Evolution

Lei Gao,[1,7] Keliang Wu,[2,7] Zhenbo Liu,[1] Xuelong Yao,[1,3] Shenli Yuan,[1,3] Wenrong Tao,[2] Lizhi Yi,[1,3] Guanling Yu,[2] Zhenzhen Hou,[2] Dongdong Fan,[4] Yong Tian,[4] Jianqiao Liu,[5,*] Zi-Jiang Chen,[2,*] and Jiang Liu[1,6,8,*]

[1]CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, 100101 Beijing, China

[2]Center for Reproductive Medicine, Shandong University, National Research Center for Assisted Reproductive Technology and Reproductive Genetics, Key Laboratory of Reproductive Endocrinology, Shandong University, Ministry of Education, Jinan, 250001 Shandong, China

[3]CAS Center for Excellence in Animal Evolution and Genetics, University of Chinese Academy of Sciences, 100049 Beijing, China

[4]Key Laboratory of RNA Biology of CAS, University of Chinese Academy of Sciences, Institute of Biophysics, Chinese Academy of Sciences, 100101 Beijing, China

[5]Center for Reproductive Medicine, Third Affiliated Hospital, Guangzhou Medical University, 510150 Guangzhou, China

[6]CAS Center for Excellence in Molecular Cell Science, University of Chinese Academy of Sciences, 100049 Beijing, China

[7]These authors contributed equally

[8]Lead Contact

*Correspondence: ljq88gz@163.com (J.L.), chenzijiang@hotmail.com (Z.-J.C.), liuj@big.ac.cn (J.L.)

https://doi.org/10.1016/j.cell.2018.02.028

SUMMARY

The dynamics of the chromatin regulatory landscape during human early embryogenesis remains unknown. Using DNase I hypersensitive site (DHS) sequencing, we report that the chromatin accessibility landscape is gradually established during human early embryogenesis. Interestingly, the DHSs with OCT4 binding motifs are enriched at the timing of zygotic genome activation (ZGA) in humans, but not in mice. Consistently, OCT4 contributes to ZGA in humans, but not in mice. We further find that lower CpG promoters usually establish DHSs at later stages. Similarly, younger genes tend to establish promoter DHSs and are expressed at later embryonic stages, while older genes exhibit these features at earlier stages. Moreover, our data show that human active transposons SVA and HERV-K harbor DHSs and are highly expressed in early embryos, but not in differentiated tissues. In summary, our data provide an evolutionary developmental view for understanding the regulation of gene and transposon expression.

INTRODUCTION

Epigenetic information plays crucial roles in the transition from totipotent zygotes to pluripotent primitive ectoderm cells in inner cell mass (ICM) in mammals. Totipotency and pluripotency are two quite distinct epigenetic states with different potentials in development. The zygotes are totipotent, as they can develop into a whole organism. The zygotes lack the capacity for self-renewal and can only undergo cleavage division. Pluripotent cells are established from totipotent blastomeres within the ICM of blastocysts. ICM cells acquire the properties of normal cell division and can self-renew when they are properly cultured *in vitro* (Surani et al., 2007). During mammalian development from the zygotes to blastocysts, significant epigenetic reprogramming happens (Ci and Liu, 2015; Dahl et al., 2016; Liu et al., 2016; Lu et al., 2016; Wu et al., 2016; Zhang et al., 2016). Proper epigenetic reprogramming is essential for embryonic development. Recent studies have shown that genome-wide abnormal reprogramming of DNA methylome, which is associated with aberrant embryonic development, occurs in human early embryos (Li et al., 2017).

Epigenetic and genetic mechanisms coordinate to regulate gene expression during development and cell-fate determination. Gene activation usually requires a *cis*-regulatory DNA sequence to be accessed by *trans* factors (Thurman et al., 2012). DNase I hypersensitivity assay (Jin et al., 2015; Stergachis et al., 2013) and assay for transposase accessible chromatin (Buenrostro et al., 2015; Cusanovich et al., 2015) are two commonly used methods to measure chromatin accessibility. Recent studies have shown that significant reprogramming of chromatin accessibility landscape occurs during embryogenesis in mice (Lu et al., 2016; Wu et al., 2016). Oct4 and Nfya contribute to the establishment of chromatin-accessibility patterns in mouse early embryos (Lu et al., 2016). The overall process of embryogenesis is similar between humans and mice, but the embryo stage of zygotic genome activation (ZGA) differs. Mouse major ZGA occurs at the two-cell stage, while human major ZGA occurs at the eight-cell stage (Niakan et al., 2012). As the dynamics of chromatin accessibility have not been investigated in human early embryos, it will be of interest to know the conservation and difference of chromatin accessibility between humans and mice.

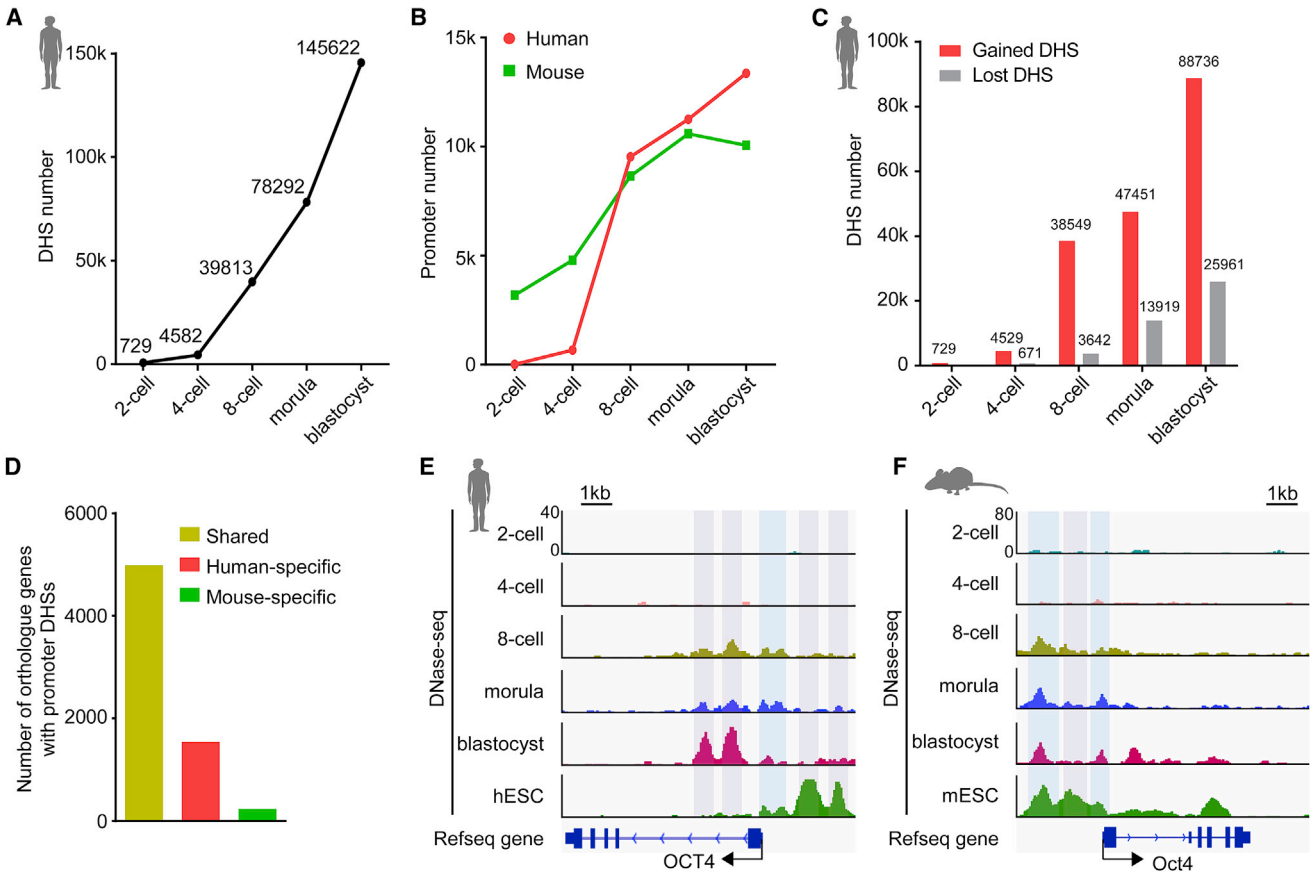


Figure 1. DHSs Are Dynamic in Human Early Embryos

(A) DHS number at each embryo stage in humans.

(B) The number of promoters harboring DHSs at each embryo stage in humans and mice.

(C) The number of gained or lost DHSs at each embryo stage in humans. The number of gained or lost DHSs is indicated above each bar.

(D) The numbers of the orthologous genes with shared or species-specific promoter DHSs between humans and mice at the blastocyst stage.

(E and F) Genome browser view of DNase-seq signal around *OCT4* promoter in human embryos, H1 hESC (ENCODE) (E) and mouse embryos and E14 mESC (ENCODE) (F).

See also Figures S1 and S2 and Tables S1, S2, and S3.

RESULTS

The Dynamics of DHSs during Human Early Embryogenesis

To investigate the chromatin regulatory landscape in human early embryos, we collected the embryos at the two-cell, four-cell, eight-cell, morula, and blastocyst stages. All human embryos were derived from *in vitro* fertilization. Embryos were collected after obtaining written informed consent signed from donor couples. We only used embryos with high morphology quality. 50–100 cells were used to perform DNase sequencing (DNase-seq) with two biological replicates for each embryo stage (Table S1). The DNase-seq results between the biological replicates are highly correlated (Figures S1A and S1B). We detected DNase I hypersensitive sites (DHSs) using a hotspot algorithm (John et al., 2011). Our data show that a large proportion of DHSs are overlapped between human blastocysts and the H1 human embryonic stem cell line (hESC) (Figure S1C). Furthermore, 230 DHSs in human blastocysts are human enhancers, which have been experimentally validated in mouse embryos (Figure S1D and Table S2) (Visel et al., 2007). These results indicate that our DNase-seq results are reliable.

Next, we were interested in the global dynamics of DHSs during early embryogenesis. Our data show that the numbers of DHSs gradually increase from the two-cell to the blastocyst stage, suggesting a progressive increase of chromatin accessibility during human embryo development (Figure 1A). The overall trend of DHS dynamics in humans is similar to that in mice (Lu et al., 2016), which is also confirmed by our DHS data in mouse embryos (Figure S1E and Table S3). Then, we examined the genomic distribution of DHSs at each stage. Our data show that a large part of human DHSs are in the intergenic regions (Figures S2A–S2C).

Mouse ZGA occurs at the two-cell stage, whereas human ZGA occurs at the eight-cell stage. We were interested in whether there is a relationship between DHS and the timing of ZGA in mammals. Unexpectedly, the proportion of DHSs locating in promoter regions of protein-coding genes at the two-cell and

four-cell stages dramatically differs between human and mouse (Figure S2D). Consistently, there are only 17 and 665 promoters harboring DHSs in human two-cell and four-cell embryos, respectively, while there are 3,190 and 4,791 promoters harboring DHSs in mouse two-cell and four-cell embryos, respectively (Figure 1B). The number of promoters harboring DHSs becomes comparable at the eight-cell stage in humans (9,533) and mice (8,647). Then, we wanted to know the chromatin accessibility of ZGA gene promoters in human and mouse embryos. By comparing the transcriptomes of two-cell and eight-cell human embryos, we defined 2,863 genes that are significantly upregulated in eight-cell embryos (fold-change > 3) as ZGA genes. In mice, we defined 2,947 ZGA genes by comparing the transcriptomes of the zygotes and late two-cell embryos. Our data show that promoter DHSs at the two-cell stage in mice are enriched in ZGA genes ($p < 2.2e^{-16}$, chi-square test), and 32% (920 of 2,947) of ZGA genes harbor promoter DHSs at the two-cell stage. In contrast, only 8% (236 of 2,863) of ZGA genes harbor promoter DHSs at the four-cell stage in humans. Upon the eight-cell stage in humans, 81% of ZGA genes (2311 of 2,863) harbor promoter DHSs. Taken together, mouse ZGA occurs at the two-cell stage, when 3,190 promoters have DHSs, while human ZGA occurs until the eight-cell stage, when 9,533 promoters have gained DHSs. These data suggest that the timing of ZGA in mammals is associated with the number of the promoters with DHSs.

To evaluate the detail of DHS dynamics during human embryogenesis, we analyzed DHSs gained or lost at each stage. Our data show that the number of DHSs gained between two consecutive stages is bigger than that of DHSs lost (Figure 1C). Both the gained and lost DHSs predominantly locate at distal regulatory regions (Figure S2E). We also find that most promoters of orthologous genes between humans and mice have similar chromatin-accessibility patterns (Figures 1D, S2F, and S2G [left panels]). However, some promoters show species-specific DHS patterns in humans and mice (Figures S2F and S2G [middle and right panels]). Interestingly, the DHSs around the *OCT4* gene locate at the intragenic region in human embryos, but at the promoter region in mouse embryos (Figures 1E and 1F). The data also show that the *OCT4* DHSs in hESCs locating at promoter regions differ from those in human early embryos (Figure 1E), suggesting that the regulation of *OCT4* in hESCs differs from that in human blastocysts.

DHSs Are Associated with Gene Expression

DHS formation is crucial for the activation of gene expression (Bell et al., 2011; Thurman et al., 2012). Consistently, the genes with promoter DHSs show higher expression than those without promoter DHSs (Figure 2A). Furthermore, DHS signal of promoters is positively correlated with gene expression level in embryos from the eight-cell stage onward in both humans and mice (Figures S3A–S3D). Moreover, the genes with higher expression tend to have higher DNase-seq signals around transcriptional end sites (TESs) in the embryos from the eight-cell stage onward (Figures S3A–S3D). We find that most genes with promoter DHSs are expressed at each developmental stage. In addition, some genes (1.5%–5%) are not expressed at the stage when these genes gain promoter DHSs, but they start transcription at a later stage (Figures S3E and S3F), suggesting that these promoters are primed to be active. Our results are consistent with the previous report in mouse embryos (Lu et al., 2016).

Establishment of DHSs Is Associated with Embryonic Development

To gain insight into the potential function of DHS dynamics during human early embryogenesis, we performed gene ontology (GO) analyses between chromatin accessibility and embryo development. Our data show that genes with promoter DHSs gained at the four-cell and eight-cell stages are enriched in RNA and DNA metabolic processes. Genes with promoter DHSs gained at the morula and blastocyst stages are enriched in cell development, cell migration, and cell commitment (Figure 2B and Table S4). *HOX* gene clusters play crucial functions in body patterning during animal development (Ferrier and Holland, 2001). Our data illustrate that no DHS is observed in all *HOX* clusters at the two-cell and four-cell stages, but the DHS signals become evident through embryo development in both humans and mice (Figures 2C, S4A, and S4B).

Many distal DHSs function as enhancers (Thurman et al., 2012). We annotated distal DHSs to their nearest genes. Genes with distal DHSs gained at the eight-cell, morula, and blastocyst stages are all enriched in nuclear-transcribed mRNA catabolic processes. Genes with distal DHSs gained at the morula stage are also enriched in endodermal cell-fate determination, and those at the blastocyst stage are enriched in blastocyst development (Figure S4C). These results suggest that the dynamics of distal DHSs are associated with embryo development.

OCT4 Contributes to ZGA in Humans

DHSs provide permissive binding sites for transcription factors, and DHS formation is crucial for the activation of gene expression. Oct4 is an important transcription factor that plays an important role during ZGA in zebrafish (Leichsenring et al., 2013). But it remains unknown whether OCT4 also plays a role during human ZGA. Enrichment analyses of transcription-factor binding motifs in the DHSs show that OCT4 binding motifs are enriched at the time of human ZGA (eight-cell stage) (Figure S5A). After injecting OCT4 small interfering RNAs (siRNAs) into human zygotes, we found that 713 (25%) ZGA genes are downregulated in knockdown (KD) eight-cell embryos (Figures 3A and 3B). 85% of the downregulated ZGA genes have established promoter DHSs at the eight-cell stage, suggesting that establishing DHSs in these genes is associated with the transcriptional regulation of OCT4 during human ZGA. Further GO analysis shows that the downregulated genes in OCT4 KD embryos are enriched in transcriptional processes, such as transcription elongation from RNA polymerase II and regulation of transcription DNA template (Figures 3A and 3C), which agrees with the requirement of OCT4 for genome activation. In contrast, Oct4 binding motifs are not enriched in the DHSs at the time of mouse ZGA (two-cell stage) (Figure S5A). Consistently, only 1.4% (40 of 2,947) of ZGA genes in mouse are downregulated in Oct4 KD embryos at the four-cell stage (Figures 3B and S5B). This limited effect

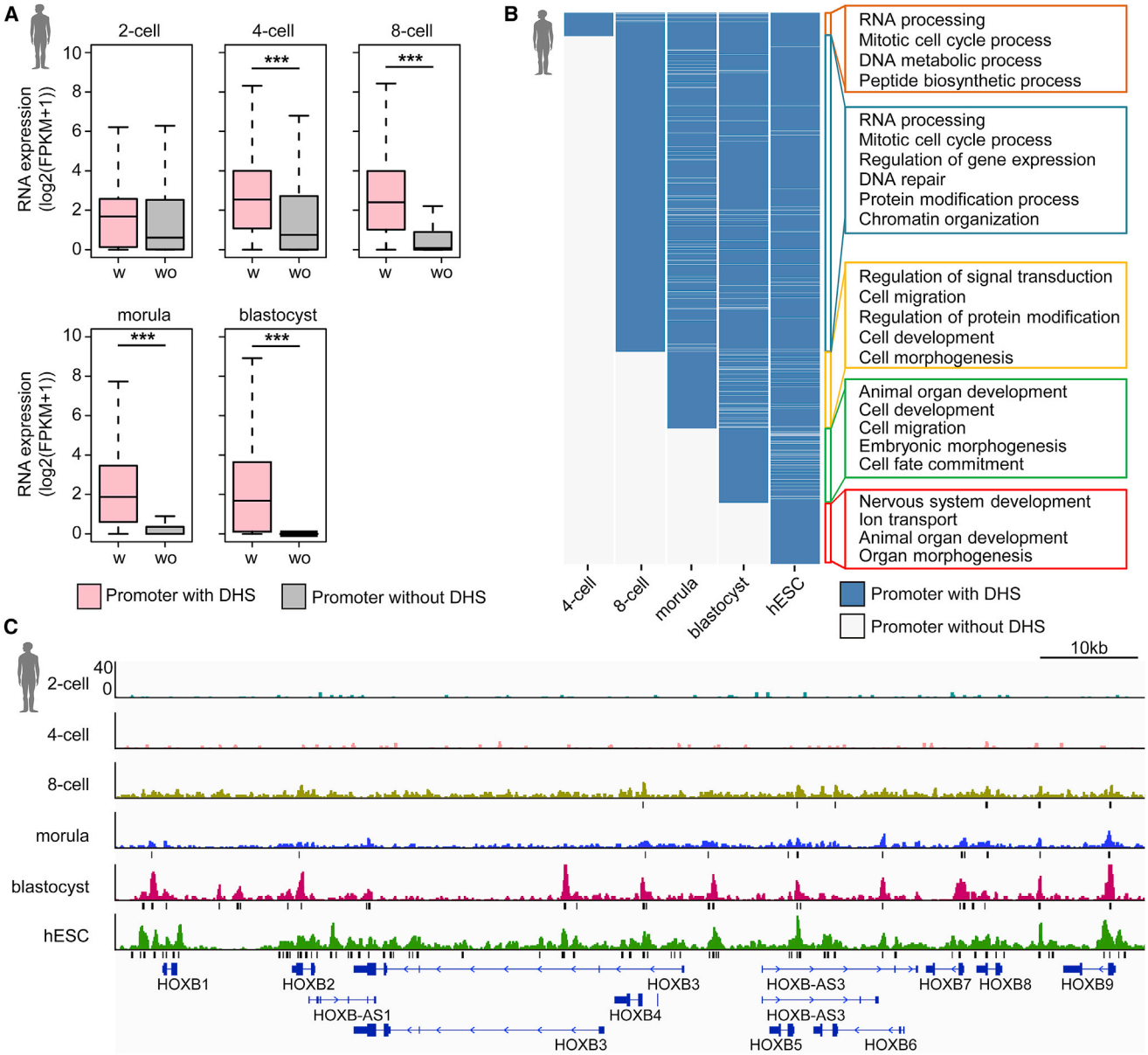


Figure 2. Chromatin Accessibility Is Associated with Embryo Development

(A) The expression level of the genes with promoter DHSs (w) or without promoter DHSs (wo) at each stage in humans. Boxes and whiskers represent the 25th/75th percentiles and 1.5 × the interquartile range, respectively. *** represents $p < 0.001$ (Wilcoxon rank-sum test).

(B) GO enrichment of the genes with promoter DHSs gained at each stage. 15,541 protein-coding genes with promoter DHSs in human embryos and H1 hESC are analyzed. The p values of the listed biological processes are <1e−5 (Fisher's exact test).

(C) Genome-browser view of DHS patterns around the *HOXB* cluster in human early embryos. Black bars indicate DHSs.

See also Figure S3 and S4 and Table S4.

of Oct4 during mouse ZGA in our results agrees with the previous report (Le Bin et al., 2014).

Among the potential targets of OCT4, SOX2 is downregulated in OCT4 KD embryos at the eight-cell stage (Figures 3A and 3D). As SOX2 binding motifs harboring DHSs are also enriched in human eight-cell embryos (Figure S5A), we investigated the impact of human SOX2 on gene regulation. Because Sox2 is not expressed in mouse early embryos, we did not examine the role of Sox2 in mice. Our data show that 200 ZGA genes are downregulated in human SOX2 KD embryos at the eight-cell stage (Figures 3E and S5C). Among these genes, 161 genes overlap with the downregulated genes in OCT4 KD embryos (Figure 3E), supporting that SOX2 is the downstream target of OCT4. Taken together, our data suggest that OCT4 contributes to human ZGA, which is associated with DHS establishment in human early embryos.

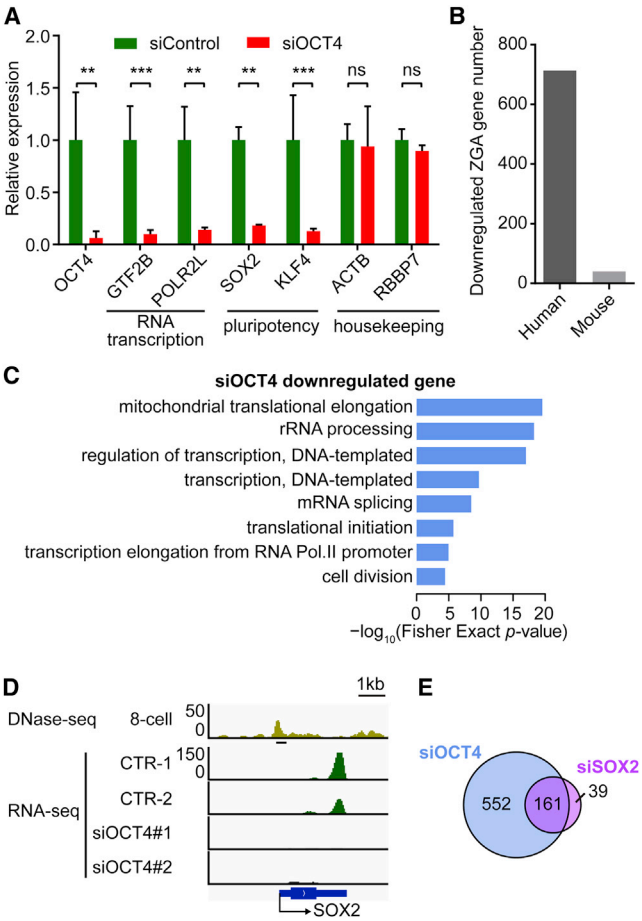


Figure 3. OCT4 Contributes to Human ZGA

(A) Relative expression of the genes in control or OCT4 KD human embryos at the eight-cell stage. RNA sequencing (RNA-seq) data were from two replicates for both control and OCT4 KD. Two replicates used two independent siRNAs for OCT4. Error bars represent standard errors. ** represents $p < 0.01$; *** represents $p < 0.001$ (t test).

(B) The numbers of downregulated ZGA genes in OCT4 KD embryos at the human eight-cell stage and mouse four-cell stage.

(C) GO analysis of downregulated ZGA genes in OCT4 KD human embryos.

(D) Genome-browser view of promoter DHS patterns and RNA expression level of *SOX2*.

(E) The number of the overlapped genes that are downregulated in OCT4 and SOX2 KD embryos. RNA-seq data were from two replicates by using two independent siRNAs for SOX2.

See also Figure S5.

DHSs Are Associated with DNA Methylation

DNA methylome reprogramming occurs in mammalian embryos (Guo et al., 2014; Smith et al., 2014; Smith et al., 2012; Wang et al., 2014), so we investigated the relationship between DNA methylation patterns and chromatin-accessibility landscapes in early embryos. First, we found that promoters with DHSs tend to be hypomethylated compared to those without DHSs in both humans (Figure 4A) and mice (Figure 4B). Then, we found an inverse correlation between DHS signal and DNA methylation level in both promoters and distal regulatory elements (Figures 4C, 4D, S6A, and S6B). In general, the methylation levels of the regions with DHSs are lower than those without DHSs (Figures S6C and S6D). These data suggest that open chromatin regions tend to be hypomethylated.

Global DNA demethylation occurs in mammalian embryogenesis (Wang et al., 2014). Next, we asked whether there is a correlation between DNA demethylation and chromatin accessibility. Our data show that the demethylation levels of DHS regions are higher than those of non-DHS regions at each embryo stage (Figure 4E). In addition, almost all demethylated promoters do not harbor DHSs until the morula and blastocyst stages (Figure 4F).

DHSs with Low CpG Densities Tend to Open at Later Embryonic Stages

CpG density is reversely correlated with DNA methylation (Lister et al., 2009; Weber et al., 2007). Here, we find that among the promoters with gained DHSs at different stages, those with higher demethylation levels typically have lower CpG densities (Figure 5A). To understand the potential roles of the genes with highly demethylated promoters, we performed GO enrichment analysis. Genes with gained promoter DHSs having lower CpG density and higher demethylation levels (Figure 5A [pink]) are enriched in the categories of immune response and reproductive pathways at the blastocyst stage (Figure 5B and Table S5).

Next, we asked whether promoter CpG density influences DHS establishment during development. Our data show that the CpG densities of the promoters establishing DHSs at earlier stages are higher than those at later stages in both humans (Figure 5C) and mice (Figure 5D). This result suggests that higher-CpG promoters tend to establish DHSs at earlier stages, while lower-CpG promoters tend to establish DHSs at later stages.

DHSs in Human Active Transposons

Transposons account for approximately 40% of mammalian genomes and play an important role in human evolution (Cordaux and Batzer, 2009). In agreement with the global dynamics of DHSs in early embryos, the number of DHSs locating in transposons also gradually increases during embryo development (Figure 6A). Our data show that the number of transposons with DHSs in the blastocysts is comparable to that of those in differentiated tissues (Figure 6A). We further analyzed the distribution of DHSs locating in different types of transposons. Our data show that the long interspersed nuclear element (LINE) transposons harboring DHSs account for a higher proportion in humans than in mice (Figures S7A and S7B).

In the human genome, only a small percentage of transposon elements are currently active and mobile. Human active transposons include several classes, such as human endogenous retroviruses (HERV), SINE-VNTR-Alus (SVAs), and Alus (Friedli and Trono, 2015; Mills et al., 2007). HERV-Ks are active in the presence of an intact long terminal repeat (LTR) containing coding sequences of retroviral proteins (Subramanian et al., 2011), which are shown to be transcribed during human embryogenesis (Grow et al., 2015). We were interested in the relationship between DHSs and the expression of human active transposons. Our data show that the HERV-Ks harboring DHSs are enriched in early embryos compared to differentiated tissues (Figure 6B). In addition, the HERV-Ks with intact human-specific LTR5

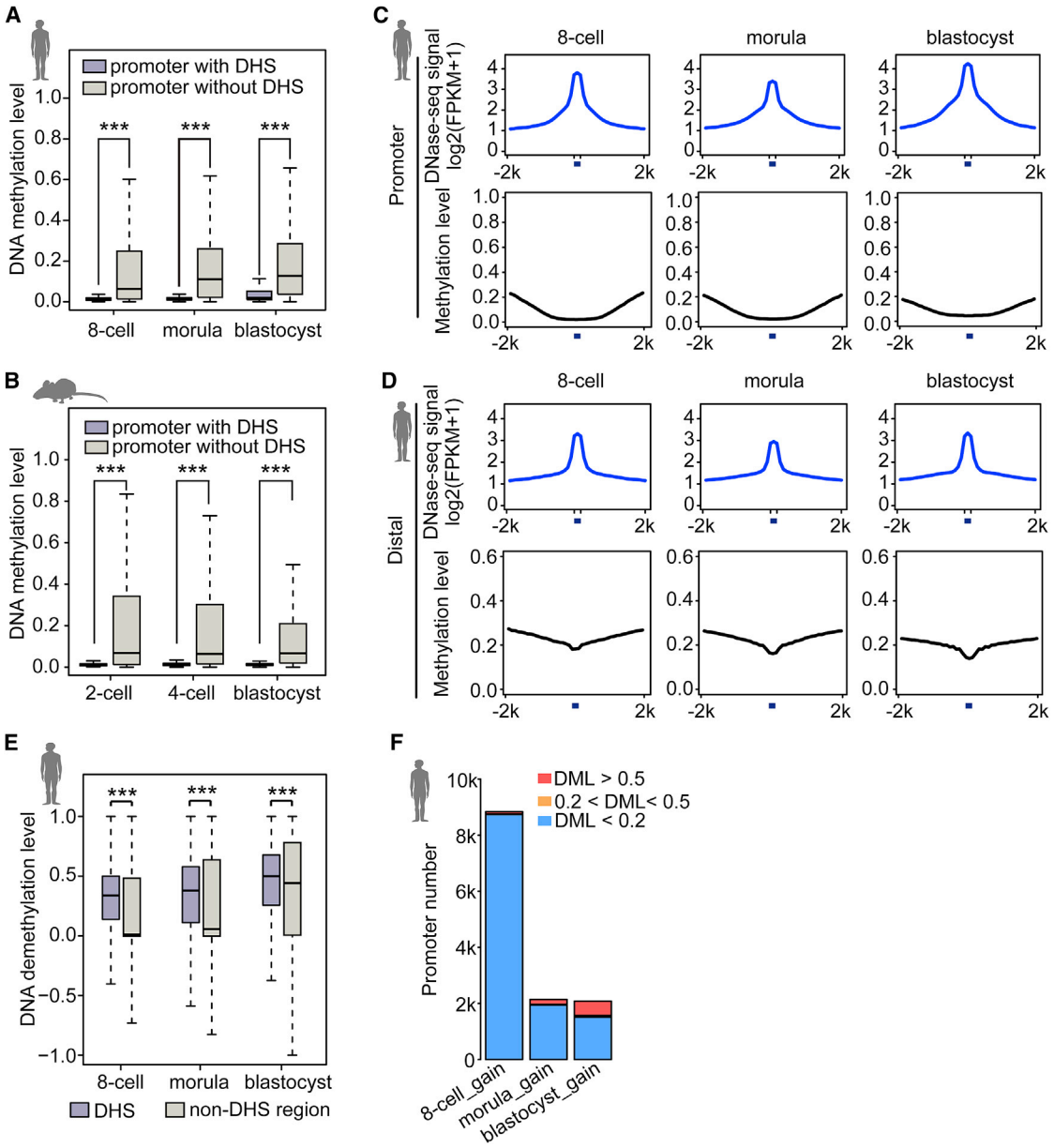


Figure 4. DHS Is Associated with DNA Methylation

(A and B) Boxplots showing DNA methylation levels (ML) of gene promoters with or without DHSs in human (A) and mouse (B) embryos. Boxes and whiskers represent the 25th/75th percentiles and 1.5 × the interquartile range, respectively. *** represents $p < 0.001$ (Wilcoxon rank-sum test).

(C and D) DNase-seq signals and DNA methylation levels around DHSs in promoter (C) and distal (D) regions in human early embryos. Blue bars indicate DHSs. Upstream and downstream 2-kb regions of DHSs are divided into 20 bins, respectively. The DHS region is divided into 3 bins. The average DNase-seq signal and DNA methylation level are calculated in each bin.

(E) Distribution plots of demethylation levels of DHSs and non-DHS regions in human early embryos. Demethylation level = (ML_{sperm} – ML_{embryo}). Boxes and whiskers represent the 25th/75th percentiles and 1.5 × the interquartile range, respectively. *** represents $p < 0.001$ (Wilcoxon rank-sum test).

(F) Bar plots showing the numbers of gained DHS promoters with different demethylation levels at a specific stage. DNA demethylation level, DML.

See also Figure S6.

elements (LTR5_Hs) show good expression in early embryos, but not in the differentiated tissues (Figure 6C). Our data further show that the HERV-Ks with DHSs have higher expression than those without DHSs (Figure S7C). Figure 6D shows that a HERV-K harboring DHSs in LTR5_Hs regions is expressed in early embryos. Expression of HERV-Hs is a hallmark of naive pluripotency in humans (Santoni et al., 2012). Our data confirm that the expression of HERV-Hs can be only observed in early embryos, but not in differentiated tissues (Figure 6E). Importantly, the HERV-Hs (flanked by LTR7 elements) harboring DHSs are

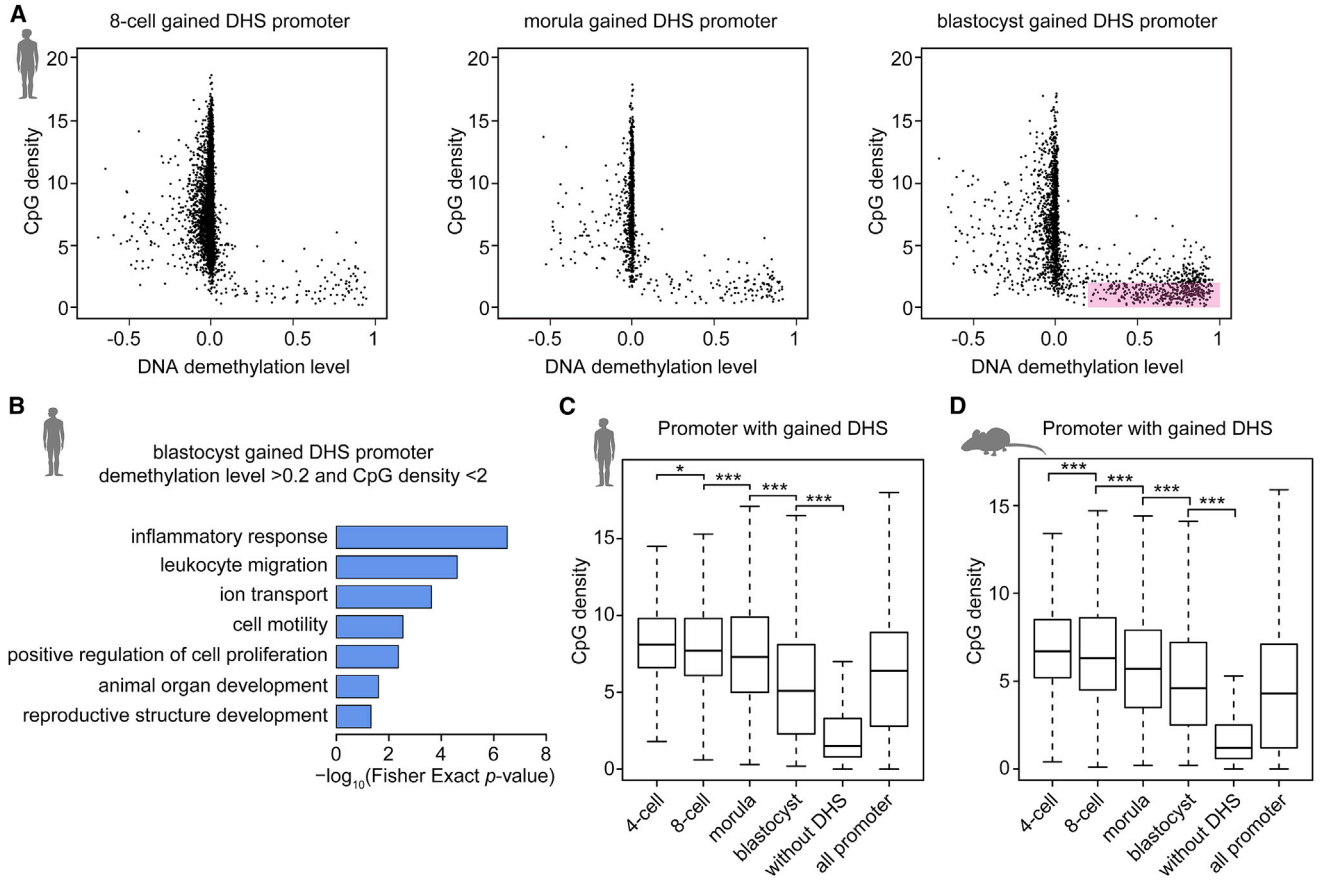


Figure 5. Low-CpG Promoters Tend to Establish DHSs at Later Embryo Stages

(A) Plot distribution of DNA demethylation levels versus CpG densities for the promoters with gained DHSs at human eight-cell, morula, and blastocyst stages, respectively.

(B) GO enrichment of the genes with the promoters in the pink region in (A). The promoters in the pink region are demethylated (DML > 0.2) and have low CpG density (<2 CpG/100 bp).

(C and D) Boxplots of the CpG density of the promoters with gained DHSs at different stages in humans (C) and mice (D). Boxes and whiskers represent the 25th/75th percentiles and 1.5 × the interquartile range, respectively. Wilcoxon rank-sum test is used for the statistical test. * represents $p < 0.05$; *** represents $p < 0.001$.

See also Table S5.

also enriched in early embryos compared to differentiated tissues (Figure 6F). In addition, the HERV-Hs with DHSs have higher expression than those without DHSs (Figure S7D).

SVAs are hominid-specific non-LTR transposons. SVAs are active in the human genome (Burns and Boeke, 2012) and are shown to be expressed in human early embryos (Guo et al., 2014). The SVAs harboring DHSs are also highly enriched in human embryos (Figure 6G). Consistently, high expression of SVAs is observed in early embryos, but not in the differentiated tissues (Figure 6H). Moreover, the SVAs with DHSs have higher expression than those without DHSs (Figure S7E).

Alus are non-LTR transposons. It has been reported that a portion of AluS subfamily and most of AluY subfamily are active and play roles in some human diseases (Bennett et al., 2008). Similar to SVAs, Alus also usually have open chromatin states in early embryos, but not in differentiated tissues (Figure S7F). Consistently, the number of expressed Alus is higher in early embryos than in differentiated tissues (Figure S7G).

DNA methylation plays important roles in the silence of transposons (Friedli and Trono, 2015). Consistently, the transposons with DHSs are hypomethylated compared to those without DHSs in both human and mouse embryos (Figures 6I and 6J).

In summary, human active transposons have higher expression in early embryos than in differentiated tissues, and the expression of active transposons is associated with chromatin accessibility. Expression of active transposons may lead to mobility of transposons, which can generate mutations in the human genome. Mutations occurring in early embryos can have a good chance to pass to germlines and offspring. Therefore, the open chromatin states of transposons in early embryos may play important roles in human evolution.

DHSs and Gene Age

Previous studies have used phylostratigraphic approaches to classify gene ages and divided human and mouse genes into 20 groups (P1–P20) respectively (Domazet-Loso and Tautz, 2008;

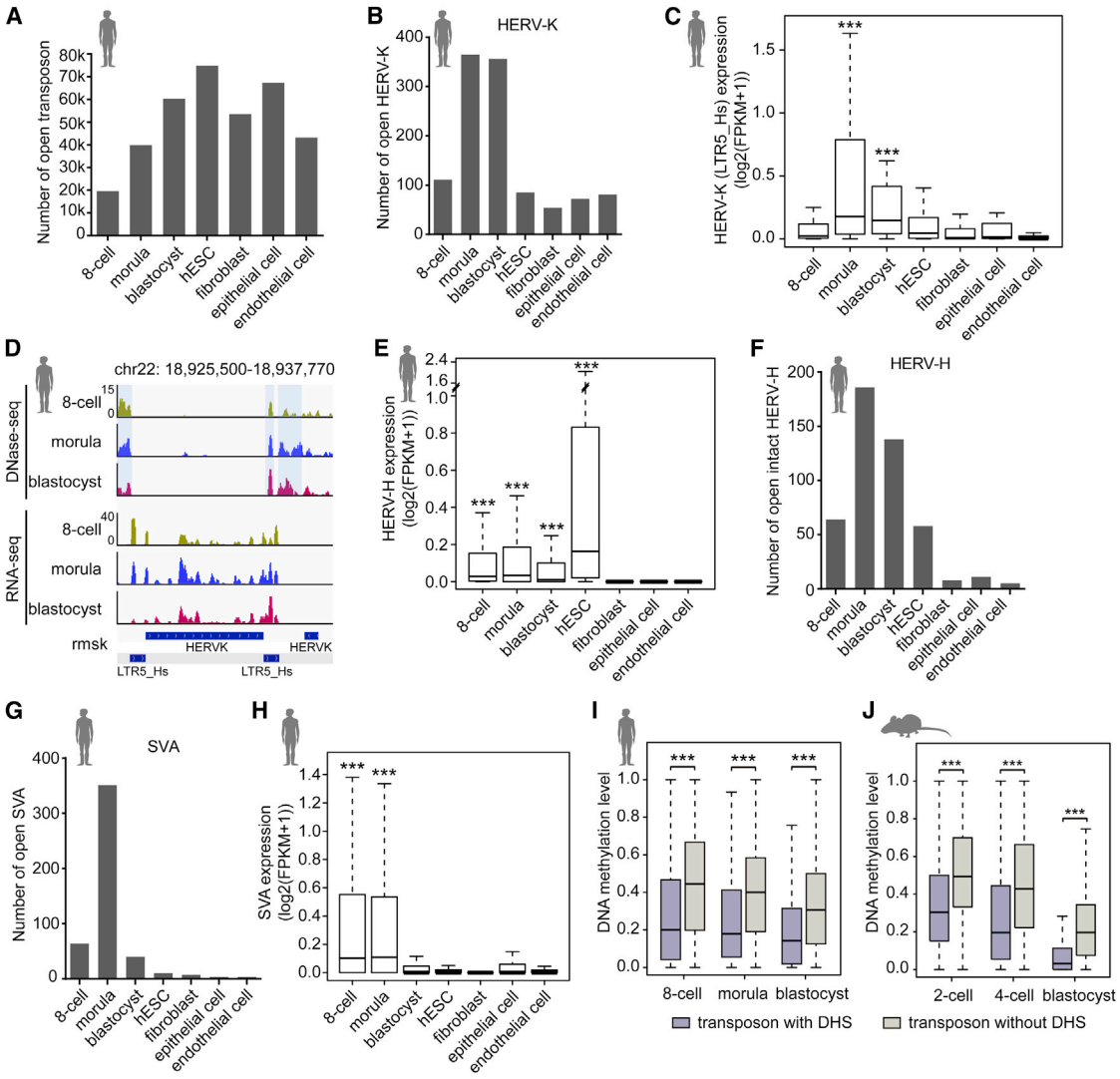


Figure 6. DHSs in Transposon Elements

(A) The number of the transposons harboring DHSs in human embryos, H1 hESC, and the differentiated cells. H1 hESC and primary cell data are from ENCODE.

(B) The number of HERV-Ks with DHSs in human embryos and the differentiated cells.

(C) RNA expression level of intact HERV-K flanked by LTR5_Hs in human embryos and differentiated cells. Boxes and whiskers represent the 25th/75th percentiles and 1.5 × the interquartile range, respectively. *** represents that transposon expression in a specific stage is significant higher than that in fibroblast, epithelial, and endothelial cells ($p < 0.001$, Wilcoxon rank-sum test). The same statistical test is used in (E) and (H).

(D) Genome-browser view of the DNase-seq and RNA-seq signals of a HERV-K flanked by LTR5_Hs. DHS peaks are shaded.

(E) RNA expression level of intact HERV-H (flanked by LTR7) in human early embryos and differentiated cells.

(F) The numbers of intact HERV-Hs with DHSs in embryos, H1 hESC, and the differentiated cells.

(G) The numbers of SVAs with DHSs in human embryos, H1 hESC, and differentiated cells.

(H) RNA expression level of SVAs in human embryos and the differentiated cells.

(I and J) Boxplots of DNA methylation levels of transposons with and without DHSs in human (I) and mouse (J) embryos. Boxes and whiskers represent the 25th/75th percentiles and 1.5 × the interquartile range, respectively. *** represents $p < 0.001$(Wilcoxon test).

See also Figure S7.

Domazet-Lošo and Tautz, 2010; Neme and Tautz, 2013). P1–P10 genes originate before the split of vertebrates and are considered relatively older genes, and P10–P20 genes are considered relatively younger genes. Here, we investigated the association between gene age and chromatin accessibility. Our data show that the relative proportion of older genes with promoter DHSs gained at an earlier stage (versus all genes with promoter DHSs gained at the corresponding stage) is higher than that at a later stage (Figure 7A). In contrast, the relative proportion of younger genes with the promoter DHSs gained at a later stage is higher than that at an earlier stage in both humans and mice (Figures 7B–7D). Taken together, older gene promoters tend to establish

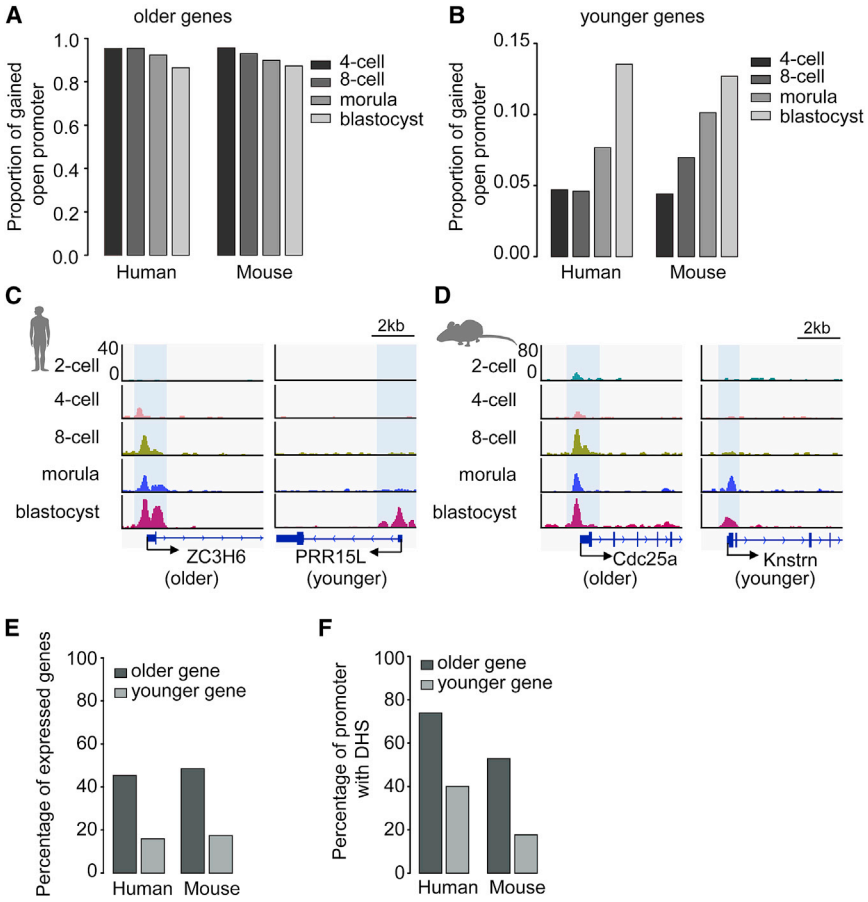


Figure 7. Association between Gene Age and Chromatin Accessibility

(A) The relative proportion of the older genes with promoter DHSs gained at each stage (versus all genes with promoter DHSs gained at the corresponding stage) in humans and mice.

(B) The relative proportion of the younger genes with promoter DHSs gained at different stages in humans and mice.

(C and D) Genome-browser view of the DNase-seq signal of representative older and younger genes in human (C) and mouse (D) embryos.

(E) The proportion of the expressed older genes (fragment per kilobase per million mapped reads [FPKM] > 1) versus all older genes in the blastocyst and the proportion of the expressed younger genes versus all younger genes.

(F) The proportion of the older genes with promoter DHSs versus all older genes in the blastocysts and the proportion of the younger genes with promoter DHSs versus all younger genes.

See also Table S6.

DHSs at earlier developmental stages, while younger gene promoters tend to establish DHSs at later developmental stages.

As older genes, compared to younger genes, tend to have DHSs at earlier developmental stages, it will be interesting to know whether older genes tend to be expressed at earlier stages. We focused our analysis at the blastocyst stage. We calculated the relative proportion of the expressed older/younger genes versus all older/younger genes. Our data show that the proportion of the expressed older genes is higher than that of the expressed younger genes in the blastocysts (Figure 7E), suggesting that older genes tend to be expressed at earlier stages. Consistently, the proportion of older genes harboring promoter DHSs (versus all older genes) is higher than that of younger genes harboring promoter DHSs in blastocysts (Figure 7F). Taken together, older genes tending to be expressed at earlier stages is associated with older genes establishing DHSs at earlier stages.

To understand the potential functions of younger genes harboring DHSs in early embryos, we performed GO enrichment. Our results show that the younger genes with the promoter DHSs gained at the morula and blastocyst stages are enriched in adaptive immune systems (Table S6). This result is consistent with the fact that some primate-specific new genes are only expressed in the placenta to fit the challenge of the immune response between mother and fetus (Than et al., 2009).

DISCUSSION

The timing of ZGA differs between humans and mice. ZGA occurs at the two-cell stage in mice but at the eight-cell stage in humans (Rossant and Tam, 2017). Our data show that 3,190 promoters have DHSs at the two-cell stage in mouse. In contrast, promoter DHSs are seldom established at the two-cell and four-cell stages in humans, and promoter DHSs are dramatically established upon the eight-cell stage. Perhaps ZGA in mammals requires a sufficient number of promoter DHSs. Oct4 plays an important role in zebrafish ZGA (Leichsenring et al., 2013). We find that OCT4 binding motifs in DHSs are enriched at the stage of human ZGA, but not of mouse ZGA. Consistently, OCT4 contributes to human ZGA, but not mouse ZGA. Our data are consistent with a recent report showing that knockout *OCT4* leads to the defect of human blastocysts, but not of mouse blastocysts (Fogarty et al., 2017). OCT4 binding motifs are not enriched until the eight-cell stage, and OCT4 significantly contributes to human ZGA, which may be another reason why human ZGA occurs at the eight-cell stage.

Transposon activity plays important roles in evolution (Cordaux and Batzer, 2009). The mobility of transposons is also involved in some human diseases (Burns and Boeke, 2012). Active transposons with DHSs have higher expression than those without DHSs, suggesting that DHS patterns mark the potential activity of transposons in human early embryos. Interestingly, human active transposon HERV-Ks and SVAs harbor DHSs and show good expression in early embryos, but not in the differentiated tissues. The expression of active transposons may lead to the activation of transposons, which can produce mutations in the human genome. Mutations in early embryos have a greater opportunity to pass into the germlines and

offspring. Thus, the chromatin accessibility pattern of active transposons in embryos may have an important impact on human evolution. Mutations in early embryos may also pass into many different organs, possibly causing some human diseases.

Our data indicate that older genes tend to establish DHSs and show expression at earlier embryo stages, while younger genes tend to establish DHSs and show expression at later embryo stages. These suggest that younger genes usually are not required for earlier cleavage stages. Through embryo development, younger genes are required when cells start differentiation in the morula and blastocysts. For example, our data illustrate that younger genes with promoter DHSs gained at later stages are enriched in the immune system. It is known that some primate-specific new genes are only expressed in the placenta to fit the challenge of the immune response between mother and fetus (Than et al., 2009). Thus, the establishment of DHSs at different embryonic stages for different-aged genes is a reasonable strategy for mammalian development.

Taken together, our data provide a very valuable resource for future studies in human embryo development. We also provide new insights into the activity of human active transposons and epigenetic regulation on mammalian embryogenesis from an evolutionary developmental perspective.

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
  - Human subjects
  - Animals
- METHOD DETAILS
  - Experimental Design
  - Human embryos collection
  - Mouse embryos collection
  - Gene knockdown by siRNA injection
  - DNase-seq
  - RNA-seq library preparation
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - DNase-seq data analysis
  - Genomic annotation
  - Gained or lost DHSs at a specific stage
  - RNA-seq data analysis
  - DNA methylation data analysis
  - Transposon analysis
  - Gene Ontology (GO) analysis
  - Gene Age
  - Human and mouse ortholog analysis
  - Statistical analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and six tables and can be found with this article online at https://doi.org/10.1016/j.cell.2018.02.028.

ACKNOWLEDGMENTS

This work was supported by the grants from the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB13040200), The National Key R&D Program (2015CB856200), National Natural Science Foundation of China (91731312, 31425015, 31601052, 31561130152, and 81601256), and CAS (QYZDY-SSW-SMC016). We thank Li Lan for the language editing.

AUTHOR CONTRIBUTIONS

Jianqiao Liu, Z.-J.C., and Jiang Liu conceived the study. L.G. and K.W. facilitated its designs. K.W., W.T., and G.Y. collected human embryos. L.G., Z.L., X.Y., and S.Y. collected mouse embryos. K.W. and Z.H. performed siRNA microinjection in human embryos. D.F. and Y.T. performed siRNA injection in mouse embryos. L.G. performed DNase-seq library construction. Z.L. performed RNA-seq library construction. L.G., X.Y., and L.Y. performed the bioinformatics analyses. L.G., K.W., Z.-J.C., and Jiang Liu interpreted the data. L.G., K.W., Z.-J.C., and Jiang Liu wrote the paper with the assistance of other authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 5, 2017
Revised: November 15, 2017
Accepted: February 9, 2018
Published: March 8, 2018

REFERENCES

Baczkowski, T., Kurzawa, R., and Głabowski, W. (2004). Methods of embryo scoring in in vitro fertilization. Reprod. Biol. *4*, 5–22.

Bao, W., Kojima, K.K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob. DNA *6*, 11.

Bell, O., Tiwari, V.K., Thomä, N.H., and Schübeler, D. (2011). Determinants and dynamics of genome accessibility. Nat. Rev. Genet. *12*, 554–564.

Bennett, E.A., Keller, H., Mills, R.E., Schmidt, S., Moran, J.V., Weichenrieder, O., and Devine, S.E. (2008). Active Alu retrotransposons in the human genome. Genome Res. *18*, 1875–1883.

Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R., et al. (2010). The NIH Roadmap Epigenomics Mapping Consortium. Nat. Biotechnol. *28*, 1045–1048.

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

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and Greenleaf, W.J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. Nature *523*, 486–490.

Burns, K.H., and Boeke, J.D. (2012). Human transposon tectonics. Cell *149*, 740–752.

Ci, W., and Liu, J. (2015). Programming and inheritance of parental DNA methylomes in vertebrates. Physiology (Bethesda) *30*, 63–68.

Consortium, E.P.; ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Cordaux, R., and Batzer, M.A. (2009). The impact of retrotransposons on human genome evolution. Nat. Rev. Genet. *10*, 691–703.

Cusanovich, D.A., Daza, R., Adey, A., Pliner, H.A., Christiansen, L., Gunderson, K.L., Steemers, F.J., Trapnell, C., and Shendure, J. (2015). Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. Science *348*, 910–914.

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.

Dang, Y., Yan, L., Hu, B., Fan, X., Ren, Y., Li, R., Lian, Y., Yan, J., Li, Q., Zhang, Y., et al. (2016). Tracing the expression of circular RNAs in human pre-implantation embryos. Genome Biol. *17*, 130.

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.

Domazet-Loso, T., and Tautz, D. (2008). An ancient evolutionary origin of genes associated with human genetic diseases. Mol. Biol. Evol. *25*, 2699–2707.

Domazet-Lošo, T., and Tautz, D. (2010). A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. Nature *468*, 815–818.

Ferrier, D.E., and Holland, P.W. (2001). Ancient origin of the Hox gene cluster. Nat. Rev. Genet. *2*, 33–38.

Fogarty, N.M.E., McCarthy, A., Snijders, K.E., Powell, B.E., Kubikova, N., Blakeley, P., Lea, R., Elder, K., Wamaitha, S.E., Kim, D., et al. (2017). Genome editing reveals a role for OCT4 in human embryogenesis. Nature *550*, 67–73.

Friedli, M., and Trono, D. (2015). The developmental control of transposable elements and the evolution of higher species. Annu. Rev. Cell Dev. Biol. *31*, 429–451.

Grow, E.J., Flynn, R.A., Chavez, S.L., Bayless, N.L., Wossidlo, M., Wesche, D.J., Martin, L., Ware, C.B., Blish, C.A., Chang, H.Y., et al. (2015). Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature *522*, 221–225.

Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., Yan, J., Ren, X., Lin, S., Li, J., et al. (2014). The DNA methylation landscape of human early embryos. Nature *511*, 606–610.

Guo, F., Li, L., Li, J., Wu, X., Hu, B., Zhu, P., Wen, L., and Tang, F. (2017). Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. Cell Res. *27*, 967–988.

Hay, D.C., Sutherland, L., Clark, J., and Burdon, T. (2004). Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. Stem Cells *22*, 225–235.

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.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. *37*, 1–13.

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

Ishibashi, R., Kozuki, S., Kamakura, S., Sumimoto, H., and Toyoshima, F. (2016). c-Rel Regulates Inscuteable Gene Expression during Mouse Embryonic Stem Cell Differentiation. J. Biol. Chem. *291*, 3333–3345.

Jin, W., Tang, Q., Wan, M., Cui, K., Zhang, Y., Ren, G., Ni, B., Sklar, J., Przytycka, T.M., Childs, R., et al. (2015). Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. Nature *528*, 142–146.

John, S., Sabo, P.J., Thurman, R.E., Sung, M.H., Biddie, S.C., Johnson, T.A., Hager, G.L., and Stamatoyannopoulos, J.A. (2011). Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. Nat. Genet. *43*, 264–268.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat. Methods *12*, 357–360.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571–1572.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Le Bin, G.C., Muñoz-Descalzo, S., Kurowski, A., Leitch, H., Lou, X., Mansfield, W., Etienne-Dumeau, C., Grabole, N., Mulas, C., Niwa, H., et al. (2014). Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. Development *141*, 1001–1010.

Leichsenring, M., Maes, J., Mössner, R., Driever, W., and Onichtchouk, D. (2013). Pou5f1 transcription factor controls zygotic gene activation in vertebrates. Science *341*, 1005–1009.

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.

Li, G., Yu, Y., Fan, Y., Li, C., Xu, X., Duan, J., Li, R., Kang, X., Ma, X., Chen, X., et al. (2017). Genome wide abnormal DNA methylome of human blastocyst in assisted reproductive technology. J. Genet. Genomics *44*, 475–481.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.

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.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

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.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Mills, R.E., Bennett, E.A., Iskow, R.C., and Devine, S.E. (2007). Which transposable elements are active in the human genome? Trends Genet. *23*, 183–191.

Muller, J., Szklarczyk, D., Julien, P., Letunic, I., Roth, A., Kuhn, M., Powell, S., von Mering, C., Doerks, T., Jensen, L.J., and Bork, P. (2010). eggNOG v2.0: extending the evolutionary genealogy of genes with enhanced non-supervised orthologous groups, species and functional annotations. Nucleic Acids Res. *38*, D190–D195.

Neme, R., and Tautz, D. (2013). Phylogenetic patterns of emergence of new genes support a model of frequent de novo evolution. BMC Genomics *14*, 117.

Niakan, K.K., Han, J., Pedersen, R.A., Simon, C., and Pera, R.A. (2012). Human pre-implantation embryo development. Development *139*, 829–841.

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.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

Rossant, J., and Tam, P.P.L. (2017). New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. Cell Stem Cell *20*, 18–28.

Roy, T.K., Bradley, C.K., Bowman, M.C., and McArthur, S.J. (2014). Single-embryo transfer of vitrified-warmed blastocysts yields equivalent live-birth rates and improved neonatal outcomes compared with fresh transfers. Fertil. Steril. *101*, 1294–1301.

Santoni, F.A., Guerra, J., and Luban, J. (2012). HERV-H RNA is abundant in human embryonic stem cells and a precise marker for pluripotency. Retrovirology *9*, 111.

Shirasawa, H., Kumagai, J., Sato, W., Kumazawa, Y., Sato, N., and Terada, Y. (2013). Retrieval and in vitro maturation of human oocytes from ovaries

removed during surgery for endometrial carcinoma: a novel strategy for human oocyte research. J. Assist. Reprod. Genet. *30*, 1227–1230.

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature *484*, 339–344.

Smith, Z.D., Chan, M.M., Humm, K.C., Karnik, R., Mekhoubad, S., Regev, A., Eggan, K., and Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo. Nature *511*, 611–615.

Stergachis, A.B., Neph, S., Reynolds, A., Humbert, R., Miller, B., Paige, S.L., Vernot, B., Cheng, J.B., Thurman, R.E., Sandstrom, R., et al. (2013). Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. Cell *154*, 888–903.

Subramanian, R.P., Wildschutte, J.H., Russo, C., and Coffin, J.M. (2011). Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. Retrovirology *8*, 90.

Surani, M.A., Hayashi, K., and Hajkova, P. (2007). Genetic and epigenetic regulators of pluripotency. Cell *128*, 747–762.

Than, N.G., Romero, R., Goodman, M., Weckle, A., Xing, J., Dong, Z., Xu, Y., Tarquini, F., Szilagyi, A., Gal, P., et al. (2009). A primate subfamily of galectins expressed at the maternal-fetal interface that promote immune cell death. Proc. Natl. Acad. Sci. USA *106*, 9731–9736.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75–82.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Vierstra, J., Rynes, E., Sandstrom, R., Zhang, M., Canfield, T., Hansen, R.S., Stehling-Sun, S., Sabo, P.J., Byron, R., Humbert, R., et al. (2014). Mouse regulatory DNA landscapes reveal global principles of cis-regulatory evolution. Science *346*, 1007–1012.

Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser–a database of tissue-specific human enhancers. Nucleic Acids Res. *35*, D88–D92.

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.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat. Genet. *39*, 457–466.

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.

Xue, Z., Huang, K., Cai, C., Cai, L., Jiang, C.Y., Feng, Y., Liu, Z., Zeng, Q., Cheng, L., Sun, Y.E., et al. (2013). Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature *500*, 593–597.

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1131–1139.

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.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
Human early embryos	Center for Reproductive Medicine, Shandong University, Jinan, China	N/A
Chemicals, Peptides, and Recombinant Proteins		
DNase I recombinant, RNase-free	Roche	Cat#04716728001
carrier RNA	Tiagen	Cat#RT416-02
phenol-chloroform	Amresco	Cat#0883
SPRIselect beads	Beckman Coulter	Cat#B23318
Critical Commercial Assays		
NEBNext Ultra II DNA Library Prep Kit for Illumina	New England Biolabs	Cat#E7645S
SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing	Takara	Cat#634888
Deposited Data		
Human early embryos DNase-seq dataset	This study	GSA: CRA000297
Mouse early embryos DNase-seq dataset	This study	GSA: CRA000297
Human siRNA knockdown 8-cell embryo RNA-seq dataset	This study	GSA: CRA000297
Mouse siRNA knock-down 4-cell embryo RNA-seq dataset	This study	GSA: CRA000297
Human H1 hESC DNase-seq dataset	Bernstein et al., 2010	GEO: GSE18927
Human fibroblast DNase-seq dataset	Consortium, 2012	GEO: GSM2308477
Human epithelial cell DNase-seq dataset	Bernstein et al., 2010	GEO: GSM753973
Human endothelial cell DNase-seq dataset	Thurman et al., 2012	GEO: GSM736627
Mouse fibroblast DNase-seq dataset	Vierstra et al., 2014	GEO: GSM1014199
Mouse kidney DNase-seq dataset	Consortium, 2012	GEO: GSM2195802
Mouse heart DNase-seq dataset	Vierstra et al., 2014	GEO: GSM1014166
Human preimplantation embryos RNA-seq dataset	Yan et al., 2013	GEO: GSE36552
Mouse preimplantation embryos RNA-seq dataset	Xue et al., 2013	GEO: GSE44183
Human H1 hESC RNA-seq dataset	Bernstein et al., 2010	GEO: GSM915328
Human fibroblast RNA-seq dataset	Consortium, 2012	GEO: GSM2072585/GSM2072586
Human epithelial cell RNA-seq dataset	Consortium, 2012	GEO: GSM2072478
Human endothelial cell RNA-seq dataset	Consortium, 2012	GEO: GSM2072324/GSM2072325
Mouse blastocyst RNA-seq dataset	Wu et al., 2016	GEO: GSE66582
Mouse zygote and 2-cell embryos RNA-seq dataset for ZGA gene analysis	Wu et al., 2016	GEO: GSE66582
Human preimplantation embryos total RNA-seq dataset	Dang et al., 2016	GEO: GSE71318
Human early embryos and sperm WGBS dataset	Li et al., 2017	GSA: CRA000114
Mouse early embryos and sperm WGBS dataset	Wang et al., 2014	GEO: GSE56697
VISTA enhancer database	Visel et al., 2007	https://enhancer.lbl.gov/
Repbase22.05	Bao et al., 2015	http://www.girinst.org/server/RepBase/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Mouse: C57BL/6	Beijing Vital River Laboratory Animal Technology, Beijing, China	Cat#213
Oligonucleotides		
Negative control siRNA: UUCUCCGAACGUGUCACGUdTdT	Synthesized by Shanghai GenePharma	N/A
Human OCT4 siRNA #1: AAGGAUGUGGUCCGAGUGUGGdTdT	Hay et al., 2004, synthesized by Shanghai GenePharma	N/A
Human OCT4 siRNA #2	Santa Cruze	Cat# sc-43980
Human SOX2 siRNA #1: CCUGUGGUUACCUCUUCCUdTdT	This study, synthesized by Bioneer	N/A
Human SOX2 siRNA #2: GCAGCUGAAAUUUAGGACdTdT	This study, synthesized by Bioneer	N/A
Mouse OCT4 siRNA #1: AAGGAUGUGGUUCGAGUAUGGdTdT	Lu et al., 2016, synthesized by Shanghai GenePharma	N/A
Mouse OCT4 siRNA #2: GGAGUCCCAGGACAUGAAAdTdT	Ishibashi et al., 2016, synthesized by Shanghai GenePharma	N/A
P5 primer: AATGATACGGCGACCACCGA	Synthesized by Sangon Biotech (Shanghai)	N/A
P7 primer: CAAGCAGAAGACGGCATACGA	Synthesized by Sangon Biotech (Shanghai)	N/A
Software and Algorithms		
Trimmomatic (v0.33)	Bolger et al., 2014	http://www.usadellab.org/cms/index.php?page=trimmomatic
Bowtie (v1.2.0)	Langmead et al., 2009	http://bowtie-bio.sourceforge.net/index.shtml
Samtools (v1.3.1)	Li et al., 2009	http://www.htslib.org/doc/samtools-1.3.1.html
Bedtools (v2.26.0)	Quinlan and Hall, 2010	http://bedtools.readthedocs.io/en/latest/#
Picard	DePristo et al., 2011	https://broadinstitute.github.io/picard/
hotspot	John et al., 2011	http://www.uwencode.org/proj/hotspot/
HOMER	Heinz et al., 2010	http://homer.ucsd.edu/homer/interactions/index.html
R	R Core Team	http://www.R-project.org/
IGV	Thorvaldsdóttir et al., 2013	http://software.broadinstitute.org/software/igv/
deepTools2	Ramírez et al., 2016	http://deeptools.ie-freiburg.mpg.de/
hisat2 (v2.0.4)	Kim et al., 2015	https://ccb.jhu.edu/software/hisat2/index.shtml
cufflinks (v2.2.1)	Trapnell et al., 2010	http://cole-trapnell-lab.github.io/cufflinks/
DESeq2	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
Bismark (v0.16.0)	Krueger and Andrews, 2011	http://www.bioinformatics.babraham.ac.uk/projects/bismark/
bamUtil (v1.0.9)	Goncalo Abecasis's lab	https://genome.sph.umich.edu/wiki/BamUtil
DAVID	Huang da et al., 2009a, 2009b	https://david.ncifcrf.gov/
GREAT	McLean et al., 2010	http://great.stanford.edu/public/html/

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

Human subjects

The study of human embryos was approved by the Reproductive Study Ethics Committee in Center for Reproductive Medicine, Shandong University, Jinan, China. The human embryos used in this study were obtained from 82 donor couples after obtaining the written informed consents signed by them. The women are 25-38 year-old with tubal-factor infertility and their husbands have normal semen. The embryos with high quality were randomly assigned to experimental groups without examination about the gender of embryos due to the limited cell number, and collected using standard clinical protocols. The number of embryos at each developmental stage used for DNase-seq in this study is shown in Table S1. One injected 8-cell stage embryo for each siRNA, including 2 human OCT4 siRNAs and 2 human SOX2 siRNAs, and two control siRNA injected 8-cell stage embryos were employed for RNA-seq experiments. No statistical methods were used to predetermine sample size.

Animals

Wild-type 6-8 weeks old female and male C57BL/6 mice were used in this study. All the mice were in healthy status, which were drug or test naive. The husbandry and experimental procedures were carried out according to the guidelines of the Institutional Animal Care and Use Committee of the Beijing Institute of Genomics and the Institutional Committee of Institute of Biophysics, Chinese Academy of Sciences. The mice used in this study were housed under a 12 hr light/dark cycle in specific pathogen-free conditions. Age and sex matched littermates between 6 to 8 weeks of age were randomly assigned to experimental groups, which were not involved in previous procedures. The mouse early embryos with high morphological quality were randomly chosen for DNase-seq and RNA-seq experiments. The number of mouse embryos at each developmental stage used for DNase-seq experiments in this study is shown in Table S3. Two injected 4-cell stage embryos for each mouse Oct4 siRNA and four control siRNA injected embryos were used for RNA-seq analysis.

METHOD DETAILS

Experimental Design

No statistical methods were used to predetermine sample size. No data were excluded. The samples were not randomized for the experiments. The investigators were not blinded to allocation during data analysis. Experiments were done with two biological replicates.

Human embryos collection

Human 2-cell, 4-cell, 8-cell, morula and morphological AA grade blastocyst embryos were vitrified at appropriate time after *in vitro* fertilization (IVF) as described below.

All of the 2-cell and 4-cell embryos were from immature MI oocytes that were clinically useless and donated by patients from IVF treatments after signing informed consent by the donor couples. MI oocytes were kept in *in vitro* maturation (IVM) medium at 37°C in an atmosphere with 5% CO_2 for 18-24 hr (Shirasawa et al., 2013). The IVM medium consists of M199 medium (GIBCO, 11-150-059) with 20% Systemic Serum Substitute (Irvine Scientific, 99193) and 75 mIU/mL of recombinant follicle stimulating hormone (Merck Serono). These *in vitro* matured oocytes were fertilized by donated sperm from sperm bank which was used just for research, then cultured in G1.5 medium (Vitrolife) in a humidified atmosphere at 37°C with 6% CO_2 in air. The 2-cell and 4-cell embryos were collected and vitrified around 27 hr and 48 hr after routine fertilization, respectively. The embryo vitrification was carried out as described (Roy et al., 2014). Briefly, the embryos were incubated in Vitrification Solution 1 consisting of 8% ethylene glycol and 8% dimethyl sulfoxide in Cryobase (10 mM HEPES-buffered media containing 20 mg/mL human serum albumin and 0.01 mg/mL gentamicin) at room temperature for 11 min. After initial shrinkage, embryos with original volume were transferred into Vitrification Solution 2 (16% ethylene glycol, 16% dimethyl sulfoxide and 0.68 M trehalose in Cryobase) for 1-1.5 min. Embryos were finally transferred onto Cryotop strip in an extremely small volume of solution (< 0.1 μL) and plunged into liquid nitrogen. After addition of the protective cover, the Cryotop was transferred into liquid nitrogen for storage.

The embryos at 8-cell, morula and blastocyst stages were from donated frozen embryos with signed informed consent. These embryos were vitrified at 3 day, 4 day and 5 day after routine *in vitro* fertilization, respectively.

The vitrified embryos on Cryotop strip were thawed rapidly by taking them from the liquid nitrogen after removal of the protective cover, then immersed in 2.5 mL of 37°C Warming Solution 1 (1 M trehalose in Cryobase) for 1 min on a heated stage. Embryos were then transferred to 0.5 mL of Warming Solution 2 (0.5 M trehalose in Cryobase) for 3 min, and then placed into 0.5 mL Cryobase for 5 min followed by fresh 0.5 mL Cryobase for 1 min. Embryos were finally transferred to G1.5 or G2 medium (Vitrolife) to evaluate embryo quality.

The embryos with high quality were selected for this study according to the following morphological criteria (Baczkowski et al., 2004): 2-cell embryos with 2 symetric blastomeres of equal size and with negligible fragmentation; 4-cell embryos with 4 blastomeres of equal size and with negligible fragmentation; 8-cell embryos with 8 blastomeres of equal size and without cytoplasmic fragments; morulae are compacted with more than 16 cells; blastocysts with clearly visible blastocyst cavity, well developed inner cell mass with many tightly packed cells, smooth trophoectoderm and thin zona.

The thawed embryos with high quality were picked randomly for experimental groups. Zona pellucida was removed by mechanical dissection with a glass needle. Embryos were washed several times by gentle pipetting with a narrow-bore glass pipette to remove the attached cumulus or polar bodies. After final washing with 0.1% BSA/PBS for 3 times, the embryos were transferred into micro-centrifuge tubes (Sorenson BioScience, 39640T) for DNase-seq experiments. The number of embryos at each developmental stage for DNase-seq was shown in Table S1. Two biological replicates were done for each embryonic stage.

Mouse embryos collection

Mouse embryos were collected from the cross of C57BL/6 mice (Beijing Vital River Laboratory Animal Technology, Beijing, China). To induce ovulation, female mice were injected intraperitoneally with pregnant mare serum gonadotropin (PMSG, 5 IU), followed by human chorionic gonadotropin (hCG, 5 IU) injection 48 hr later. Mouse 2-cell, 4-cell, 8-cell, morula and blastocyst stage embryos were collected at 30 hr, 52 hr, 68 hr, 80 hr and 3.5 day post hCG administration, respectively. After washing several times with M2 medium (Sigma, M7167) to remove contaminants, embryos were treated with acid Tyrode's solution (Sigma, T1788) to remove zona pellucida. Polar bodies were dissociated by gentle pipetting with a narrow-bore glass pipette. Finally, embryos were washed with 0.1% BSA/PBS, and then transferred into 1.7 mL tubes for DNase-seq experiments. The number of embryos at each developmental stage for DNase-seq was shown in Table S3. Two biological replicates were used for each embryonic stage.

Gene knockdown by siRNA injection

To investigate the function of OCT4 and SOX2 in human ZGA, the following siRNAs targeting human *OCT4* or *SOX2* were used for microinjection: human *OCT4* siRNA #1 (AAGGAUGUGGUCCGAGUGUGGdTdT) (Hay et al., 2004); human *OCT4* siRNA #2 (Santa Cruze, sc-43980); *SOX2* siRNA #1 (CCUGUGGUUACCUCUUCCUdTdT); *SOX2* siRNA #2 (GCAGCUGAAAUUUAGGACdTdT); negative control siRNA (UUCUCCGAACGUGUCACGUdTdT). For this assay, the donated IVM oocytes were used. They were *in vitro* fertilized and cultured in G1.5 medium as described above. 10 μM siRNA solution was loaded into injection pipette and injected into the zygote before pronuclear fading using Eppendorf PiezoXpert and Eppendorf CellTram vario microinjector. The injected embryos were cultured in G1.5 medium (Vitrolife) in a humidified atmosphere at 37°C with 6% CO_2 in air. The injected embryos with normal morphology were harvested at 8-cell stage at 3 day after fertilization. One 8-cell embryo was used for each RNA-seq library preparation assay, including one injected embryo for each human OCT4 or SOX2 siRNA and two control siRNA injected embryos for two control replicates.

To examine the role of Oct4 in mouse ZGA, the *in vitro* fertilized zygotes from C57BL/6 mouse were injected with 5 μM siRNA at 7 hr post insemination. The following siRNAs were used: mouse *Oct4* siRNA #1 (AAGGAUGUGGUUCGAGUAUGGdTdT) (Lu et al., 2016); mouse *Oct4* siRNA #2 (GGAGUCCCAGGACAUGAAAdTdT) (Ishibashi et al., 2016) and negative control siRNA (UUCUCCGAACGU GUCACGUdTdT). The siRNA solution was injected into the cytoplasm of zygotes using Eppendorf FemtoJet 4i microinjector. The siRNA injected embryos were cultured in EmbryoMax® KSOM Medium (1X) with 1/2 Amino Acids (Millipore, MR-106-D) in 20 μL droplets covered with mineral oil (SIGMA, M5904) at 37°C under 5% CO_2 in air, and harvested at 4-cell stage. Two siRNA injected 4-cell embryos with normal morphology were pooled for a RNA-seq library preparation assay. Two mouse *Oct4* siRNA #1 injected embryos, two mouse *Oct4* siRNA #2 injected embryos and four negative control siRNA injected embryos for two control replicates were used.

DNase-seq

DNase-seq experiment was carried out as described (Jin et al., 2015; Lu et al., 2016). Briefly, embryos were lysed in 36 μL cold lysis buffer (10mM Tris-HCl pH 7.5, 10mM NaCl, 3mM $MgCl_2$, 0.1% Triton X-100) and kept on ice for 20-30 min. 4 μL diluted DNaseI (Roche, 04716728001) was add to final concentration of 70 U/mL and incubated at 37°C for exactly 5 min. Reaction was stopped by adding 80 μL of stop buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 0.15% SDS, 10 mM EDTA) containing 40 μg Proteinase K (QIAGEN, 19133) followed by incubation at 55°C for 1 hr. After adding 30 ng carrier RNA (Tiagen, RT416-02), the DNA was extracted by phenol-chloroform (Amresco, 0883) and precipitated by ethanol with 20 μg glycogen (Thermo Fisher, R0551) and 1/10 volume 3M NaOAc (Thermo Fisher, R1181) at −20°C overnight. Following centrifugation at max speed for 15 min, DNA precipitation was washed with 800 μL ice-cold 70% ethanol, then dissolved in 50 μL TE (2.5mM Tris-HCl pH 7.5, 0.05mM EDTA) after air-dry. NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, E7645S) was used for library construction according to manufactory's instruction. DNA was end repaired and A-tailed by adding 7 μL NEBNext Ultra II End Prep Reaction Buffer and 3 μL NEBNext Ultra II End Prep Enzyme Mix. Samples were incubated in a thermal cycler at 20°C for 30min, 65°C for 30min, and finally cooled to 4°C. Adaptor ligation was performed by adding 30 μL NEBNext Ultra II Ligation Master Mix, 1 μL NEBNext Ligation Enhancer, 0.5 μL 200mM ATP and 2.5 μL Y-shapped Illumina Multiplexing Adaptors (1.5 μM). Samples were thoroughly mixed and incubated at 20°C for 25min. After adaptor ligation, 1.3 volume SPRIselect beads (Beckman Coulter, B23318) were used to purify DNA and 8 cycles PCR amplification was performed with NEBNext Ultra II Q5 Master Mix. 150-400 bp DNA fragments were selected with 0.65 volume plus 0.65 volume

SPRIselect beads, then eluted in 25 μL TE. 1 μL eluent was diluted into 1 mL nuclease-free water, 8.8 μL dilution was used for qPCR (20 μL reaction) with SsoFast Evagreen Supermix (Bio-Rad, 172-5200) on Bio-Rad CFX96 real-time PCR instrument. P5 primer (AATGATACGGCGACCACCGA) and P7 primer (CAAGCAGAAGACGGCATACGA) were used for qPCR. The Ct value at threshold of 300 was recorded and the number of cycles for second PCR amplification was calculated by Ct_{300} value minus 9. After second PCR amplification, PCR product was purified with 1.3 volume SPRIselect beads. The libraries were sequenced on Hiseq X10 with paired-end 150 bp (Illumina).

RNA-seq library preparation

The siRNA injected human or mouse embryos were lysed directly and used for cDNA synthesis using SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (Takara, 634888). Briefly, the sample volume of human or mouse embryos was adjusted to 9.5 μL with nuclease-free water. After adding 1 μL 10X Reaction Buffer (0.95 μL 10X Lysis Buffer, 0.05 μL RNase Inhibitor), samples were incubated at room temperature for 5 min, then placed on ice. 2 μL 3′ SMART-Seq CDS Primer II A (12 μM) was added. Following incubation at 72°C for 3 min, samples were placed on ice for 2 min. cDNA synthesis reaction was set up by adding 4 μL 5X Ultra Low First-Strand Buffer, 1 μL SMART-Seq v4 Oligonucleotide (48 μM), 0.5 μL RNase Inhibitor (40 U/μL) and 2 μL SMARTScribed Reverse Transcriptase. The reaction was performed in a thermal cycler with following program: 42°C for 90min, 70°C for 10min, 4°C forever. The first-strand cDNA product was amplified by adding 25 μL 2X SeqAmp PCR Buffer, 1 μL PCR Primer II A (12 μM), 1 μL SeqAmp DNA Polymerase and 3 μL nuclease-free water. 16 rounds PCR amplification was employed with following program: 95°C for 1 min; 98°C for 10 s, 65°C for 30 s and 68°C for 3 min, repeat these 3 steps for 15 times; 72°C for 10 min; 4°C forever. The amplified cDNA was purified using 1 volume SPRIselect beads, then fragmented to 200-400 bp by Covaris sonicator (Covaris). Sequencing libraries were prepared with NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, E7645S) as described above. In order to obtain adequate amount of DNA for sequencing, the cycle of PCR amplification was determined according to the amount of 1 μL amplified DNA, which was evaluated using FlashGel System (Lonza, 57063). The libraries were sequenced on Hiseq X10 with paired-end 150bp (Illumina).

QUANTIFICATION AND STATISTICAL ANALYSIS

DNase-seq data analysis

Low quality reads of DNase-seq data were removed. Then, the remaining reads were cropped to 100 bp by Trimmomatic v0.33 (Bolger et al., 2014). Paired reads1 and unpaired reads were used for mapping. Human libraries were aligned to human genome hg19 and mouse libraries were aligned to mm10 by Bowtie v1.2.0 (Langmead et al., 2009) with parameter "-m 1." Low mapping quality (MAPQ < 10) and PCR duplicated reads were removed by Samtools and Picard (DePristo et al., 2011; Li et al., 2009). DHSs were called by hotspot algorithm with FDR < 0.01 (John et al., 2011). DHSs covered by less than 8 reads were filtered out.

The Fragment Per Kilobase per Million mapped reads (FPKM) value for each genome-wide non-overlapped 2kb window was calculated as tag density in DNase-seq data. The Pearson correlation coefficient (r) of tag densities between two replicates was used to evaluate the reproducibility. The alignment data of two replicates at each stage were merged, and DHS peaks were called as described above for downstream analysis.

DNase-seq signal tracks visualized in Integrative Genomics Viewer (IGV) were generated by bamCoverage in Deeptools2 suite with parameters "–noralizeUsingRPKM–extendReads 150" (Ramírez et al., 2016; Robinson et al., 2011; Thorvaldsdóttir et al., 2013).

Published DNase-seq data were downloaded from indicated GEO datasets: human H1 hESC from GSE18927 (Bernstein et al., 2010); human fibroblast from GSM2308477 (Consortium, 2012); human epithelial cell from GSM753973 (Bernstein et al., 2010); human endothelial cell from GSM736627 (Thurman et al., 2012); mouse fibroblast, kidney and heart from GSM1014199, GSM2195802 and GSM1014166, respectively (Consortium, 2012; Vierstra et al., 2014). The DHSs were called through method as aforesaid without cropping reads.

Genomic annotation

The hg19 and mm10 refGene files from UCSC Table Browser were used for genome annotations. Only protein-coding genes were selected for analysis. TSS ± 500bp was defined as promoter region. If a DHS is overlapped with a promoter region, the DHS was considered as a promoter DHS. If a DHS is overlapped with exon but not overlapped with promoter, the DHS was considered as an exon DHS. If a DHS is overlapped with intron but not overlapped with promoter and exon, the DHS was considered as an intron DHS. The other DHSs were classed into intergenic DHSs. The non-promoter DHSs were defined as distal DHSs. The Bedtools v2.26.0 was used for the analysis (Quinlan and Hall, 2010).

To check whether the DHSs in human embryos were experimentally validated enhancers, the experimental validated human enhancer data (1792 enhancers) were downloaded from VISTA database (Visel et al., 2007).

The script findMotifsGenome.pl in HOMER software was used to analyze transcription factors binding motifs enrichment in DHSs (Heinz et al., 2010).

Gained or lost DHSs at a specific stage

To investigate the dynamics of chromatin accessibility, the DHSs at one stage that do not overlap with the DHSs at the previous stage were defined as the gained DHS at this stage. The DHSs at one stage that do not overlap with the DHSs at the next stage were defined as the lost DHSs at this stage.

RNA-seq data analysis

Low quality reads of RNA-seq and adaptor sequences were removed by Trimmomatic v0.33. Then human data were aligned to hg19 and mouse data were aligned to mm10 by hisat v2.0.4 with parameter "–dta-cufflinks" (Kim et al., 2015). The FPKM value of each gene was calculated by Cufflinks v2.2.1 (Trapnell et al., 2010). The reference transcript annotation of protein-coding genes was downloaded from refGene in UCSC Table Browser. For protein-coding genes with multiple transcripts, the longest transcripts were kept.

The R package DESeq2 were used for gene differential expression analysis (Love et al., 2014). To figure out downregulated genes in siRNA knockdown embryos, 2 sets of control siRNA RNA-seq data, 1 set of siRNA #1 data and 1 set of siRNA #2 data were used for analysis. FDR < 0.05, fold change > 2 were used as cutoff for downregulated genes.

RNA-seq raw data of human and mouse preimplantation embryos were downloaded from GSE36552 and GSE44183, respectively (Xue et al., 2013; Yan et al., 2013). RNA-seq raw data of human H1 hESC, fibroblast, epithelial cell and endothelial cell were downloaded from GSM915328, GSM2072585/GSM2072586, GSM2072478 and GSM2072324/GSM2072325, respectively (Bernstein et al., 2010; Consortium, 2012). RNA-seq raw data of mouse blastocyst were downloaded from GSE66582 (Wu et al., 2016).

The ZGA genes were defined as those significantly upregulated in human 8-cell embryos (FPKM > 1 and fold change > 3 comparing with 2-cell embryo) or mouse 2-cell embryos (FPKM > 1 and fold change > 3 comparing with zygote). For mouse ZGA gene analysis, the RNA-seq raw data of mouse zygote and late 2-cell embryos were used which were downloaded from GSE66582 (Wu et al., 2016).

To analyze the expression of transposon in human early embryos, the total RNA-seq raw data downloaded from GSE71318 were analyzed as described above (Dang et al., 2016). Middle blastocyst RNA-seq data were used for blastocyst. The reference repeat annotation was downloaded from rmsk in UCSC Table Browser. The FPKM value of each repeat or transposon was calculated by Cufflinks.

RNA-seq tracks visualized in IGV were generated by bamCoverage program in Deeptools2 with parameter "–noralizeUsingRPKM."

To analyze correlation between gene expression level and DNase-seq signal, protein coding genes were divided into 5 groups according to RNA FPKM values (0-1, 1-5, 5-10, 10-50, > 50). The genic region and upstream/downstream 10kb regions of each gene were divided into 5 bins, respectively. The DNase-seq signal of each bin in a gene group was measured by the average DNase-seq FPKM.

DNA methylation data analysis

The DNA methylation data of human early embryos and sperm were downloaded from CRA000114 in PRJCA000248 (Li et al., 2017), which had been deposited in the Genome Sequence Archive of Beijing Institute of Genomics, Chinese Academy of Sciences. The DNA methylation data of mouse early embryos and sperm were downloaded from GSE56697 (Wang et al., 2014). Low quality reads were removed by Trimmomatic v0.33. The filtered reads were aligned to reference genomes by Bismark with default parameters (Krueger and Andrews, 2011). PCR duplicates were removed by Picard and overlapped parts of pair-end reads were trimmed from one end by bamUtil. The methylation level of each CpG site was calculated by a custom script.

The DNA methylation level of each promoter (TSS ± 500 bp) or transposon element was measured as the average DNA methylation level of all CpG sites in the region.

To analyze the correlation between DNase-seq signal and DNA methylation level around DHS, each DHS was divided into 3 bins, and the upstream or downstream 2kb regions were both divided into 20 bins. The average DNase-seq FPKM value and DNA methylation level were calculated in each bin.

To investigate DNA demethylation level of DHSs in each stage embryo, DNA demethylation level of each DHS were calculated by the difference value of DNA methylation level between sperm and corresponding stage embryo.

The CpG number was counted in each assigned region by a custom script. The CpG density was defined as the average number of CpG site per 100 bp for this region.

Transposon analysis

The annotation of DNA, LTR, SINE, LINE, RC and Others in the rmsk file in UCSC Table Browser were chosen as transposon annotation. Transposon overlapping with DHS was considered as transposon with DHS (open transposon). HERV-Ks contain HERVK, LTR5_Hs, LTR5A and LTR5B in ERVK class. The HERV-K flanked by LTR5_Hs was selected as intact HERV-K with LTR5_Hs. The HERV-H flanked by LTR7 was selected as intact HERV-H. Alu elements showing > 268 bp length and > 92% identity to their consensus sequences downloaded from Repbase22.05 were referred to Alu, including 36880 AluY and 11471 AluS (Bao et al., 2015; Bennett et al., 2008). The annotation files for these transposons were generated from rmsk annotation file in UCSC Table Browser with custom script. RNA expression levels of these transposons were calculated by Cufflinks.

Gene Ontology (GO) analysis

GO analysis was performed by DAVID tool (Huang da et al., 2009a, 2009b). For the distal DHSs, the nearest genes within 100 kb were chosen for GO analysis by GREAT program (McLean et al., 2010).

Gene Age

To investigate the association between gene age and chromatin accessibility, the human and mouse gene age data were downloaded from published study (Neme and Tautz, 2013). Human and mouse genes are classified into 20 groups (P1-P20) by using phylostratigraphic approaches, respectively. P1 - P10 genes are originated before the split of vertebrates and are considered as relatively older genes; and P10 - P20 genes are considered as relatively younger genes.

Human and mouse ortholog analysis

To analyze the conservation of chromatin accessibility patterns between human and mouse, the human and mouse ortholog gene groups were downloaded from website: (http://eggnogdb.embl.de/download/eggnog_4.5/data/meNOG/) (Muller et al., 2010). We chose the ortholog gene pairs with only one ortholog gene in both human and mouse, the ortholog gene pairs with multiple ortholog genes in either species were discarded.

Statistical analysis

R was used for statistics analysis. Pearson's correlation coefficient was calculated using the cor.test function with default parameters to evaluate reproducibility of replicates in Figure S1A. Wilcoxon rank sum test was used for RNA expression comparisons (Figures 2A, 6C, 6E, 6H, and S7C-S7E), DNA methylation comparisons (Figures 4A, 4B, 4E, 6I, 6J, S6C, and S6D) and CpG density comparisons (Figures 5C and 5D) using the wilcox.test function in R. * represents $p < 0.05$, ** represents $p < 0.01$, and *** represents $p < 0.001$. t test was used for RNA differential expression analysis between siRNA KD embryos (2 embryos for each gene) and NC siRNA injected embryos (2 embryos for control knockdown) in Figures 3A, S5B, and S5C. In t test, ** represents $p < 0.01$, *** represents $p < 0.001$, and ns represents no significance ($p > 0.05$). The chi-square.test function was used to calculate the enrichment of promoter DHSs in mouse ZGA genes. Fisher Exact test was used for GO enrichment analysis by DAVID tool in Figures 2B, 3C, and 5B and Tables S4, S5, and S6. The number of embryos or biological replicates used in this study can be found in Tables S1 and S3 and Method Details section. No statistical methods were used to predetermine sample size.

DATA AND SOFTWARE AVAILABILITY

The accession number for the sequencing data reported in this paper is GSA: CRA000297. These data have been deposited in the Genome Sequence Archive under project PRJCA000484.

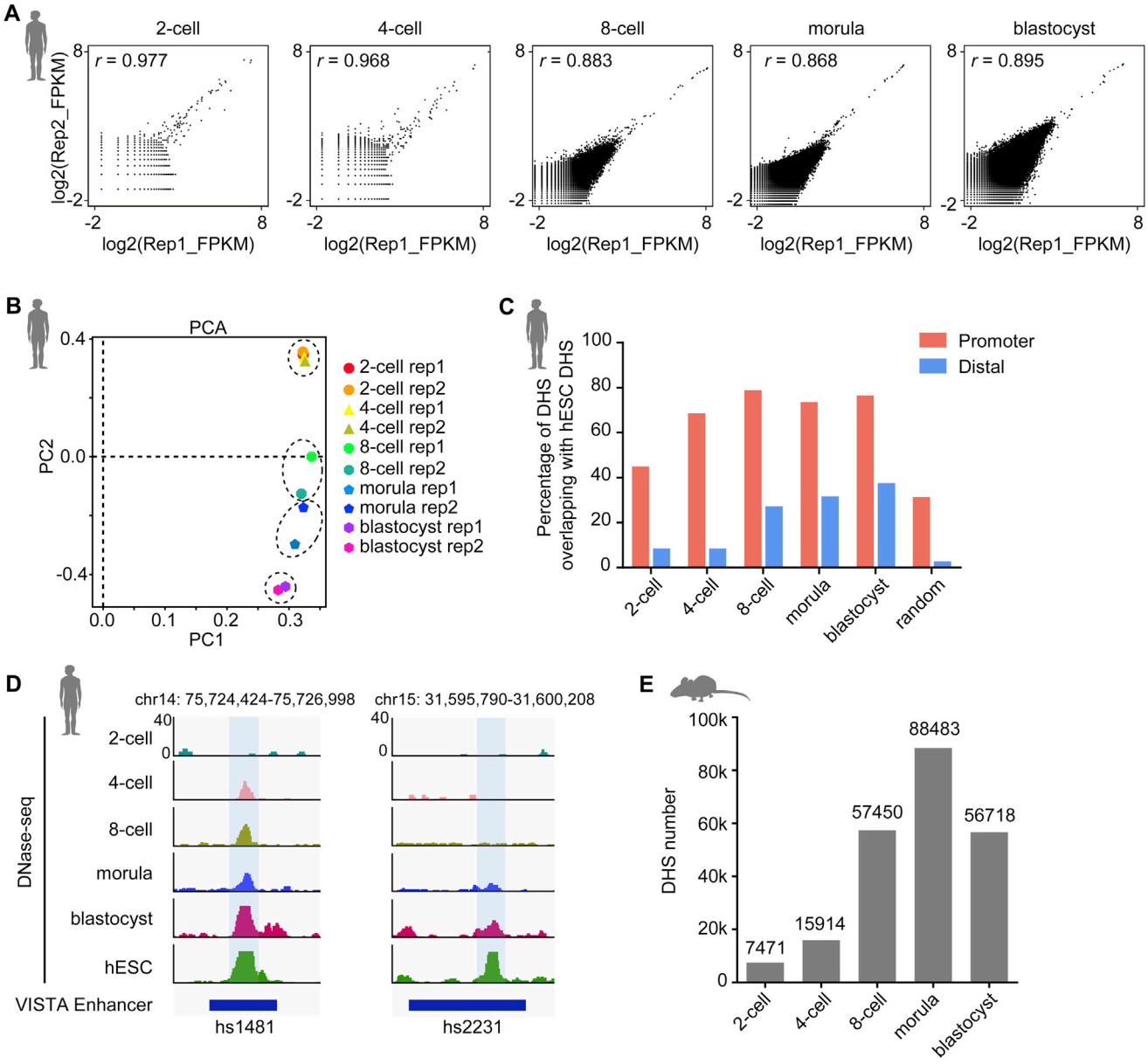


Figure S1. Validation of DNase-Seq Data Quality, Related to Figure 1

(A) The correlation between biological replicates at each stage. DNase-seq Fragment Per Kilobase per Million reads (FPKM) is calculated for each non-overlapped 2kb window in whole genome. "*r*" indicates the Pearson correlation coefficients.

(B) Principal component analysis (PCA) of DNase-seq data for all samples.

(C) Percentage of the promoter or distal DHSs in human early embryos overlapping with the DHSs in H1 hESC (ENCODE). DHSs within TSS ± 500bp were referred as promoter DHSs. Distal DHSs represent non-promoter DHSs. A set of random regions was added.

(D) Genome browser view of two representative loci in human early embryos. The two loci were experimentally validated enhancers (VISTA database). DHSs are shaded.

(E) DHS number at each stage in mouse early embryos. The DHS number at the blastocyst stage is smaller than that at the morula stage, which is consistent with the previous observation (Guo et al., 2017).

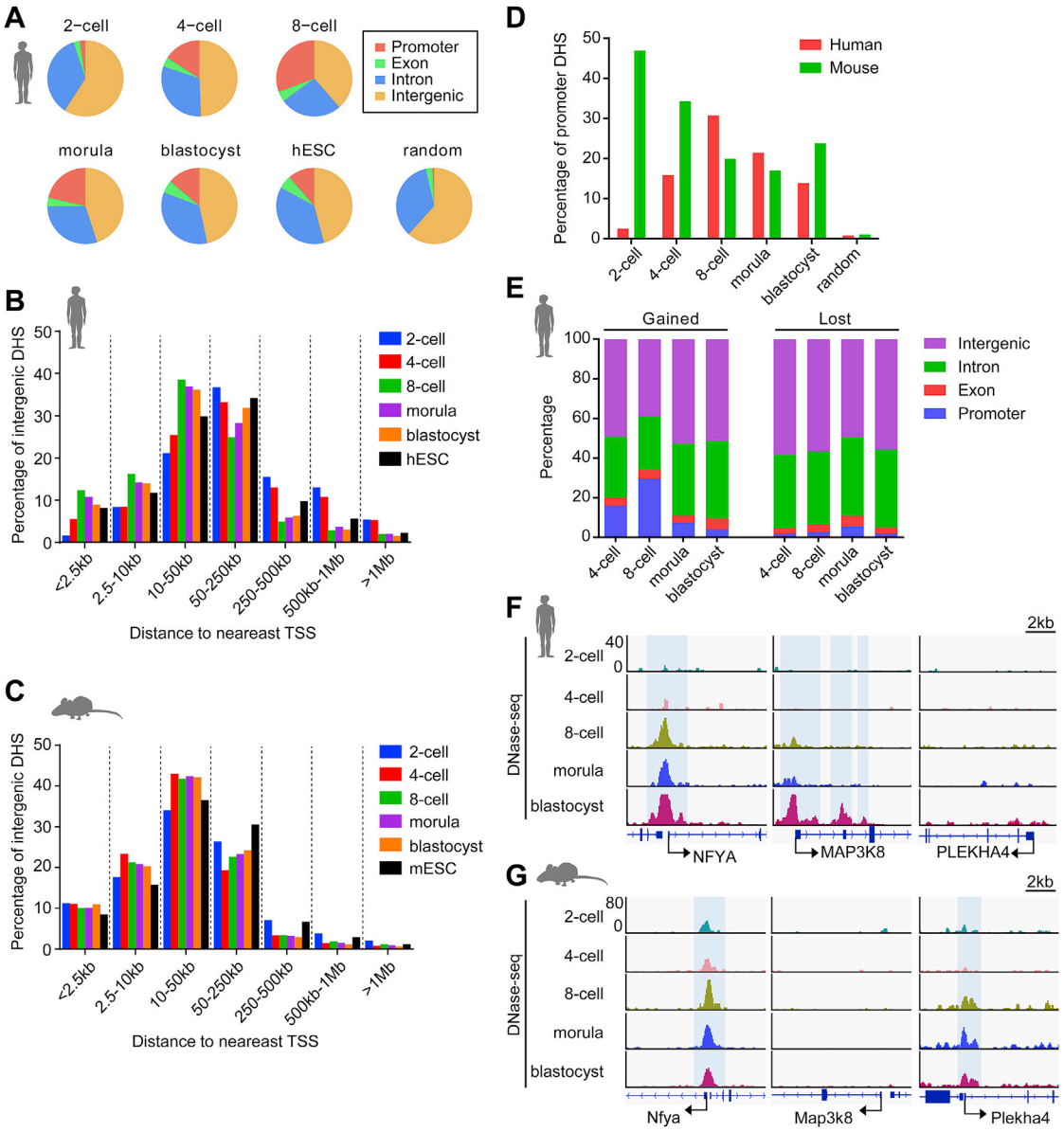


Figure S2. DHS Dynamics in Human Early Embryos, Related to Figure 1.

(A) Pie graphics showing genomic distribution of DHSs in human early embryos and H1 hESC (ENCODE). A set of random regions matching the length and number of DHSs in human blastocyst is plotted.

(B and C) The percentage of the intergenic DHSs within different distance windows to the nearest TSS in human (B) and mouse (C) embryos.

(D) The percentage of promoter DHSs (versus all DHSs) at each stage. A set of random regions matching the length and number of DHSs in human or mouse blastocyst is plotted.

(E) Bar plots showing genomic distribution of gained or lost DHSs at each embryo stage in human.

(F and G) Genome browser view of the representative DHS patterns around three orthologous genes in human (F) and mouse (G) embryos. DHSs are shaded.

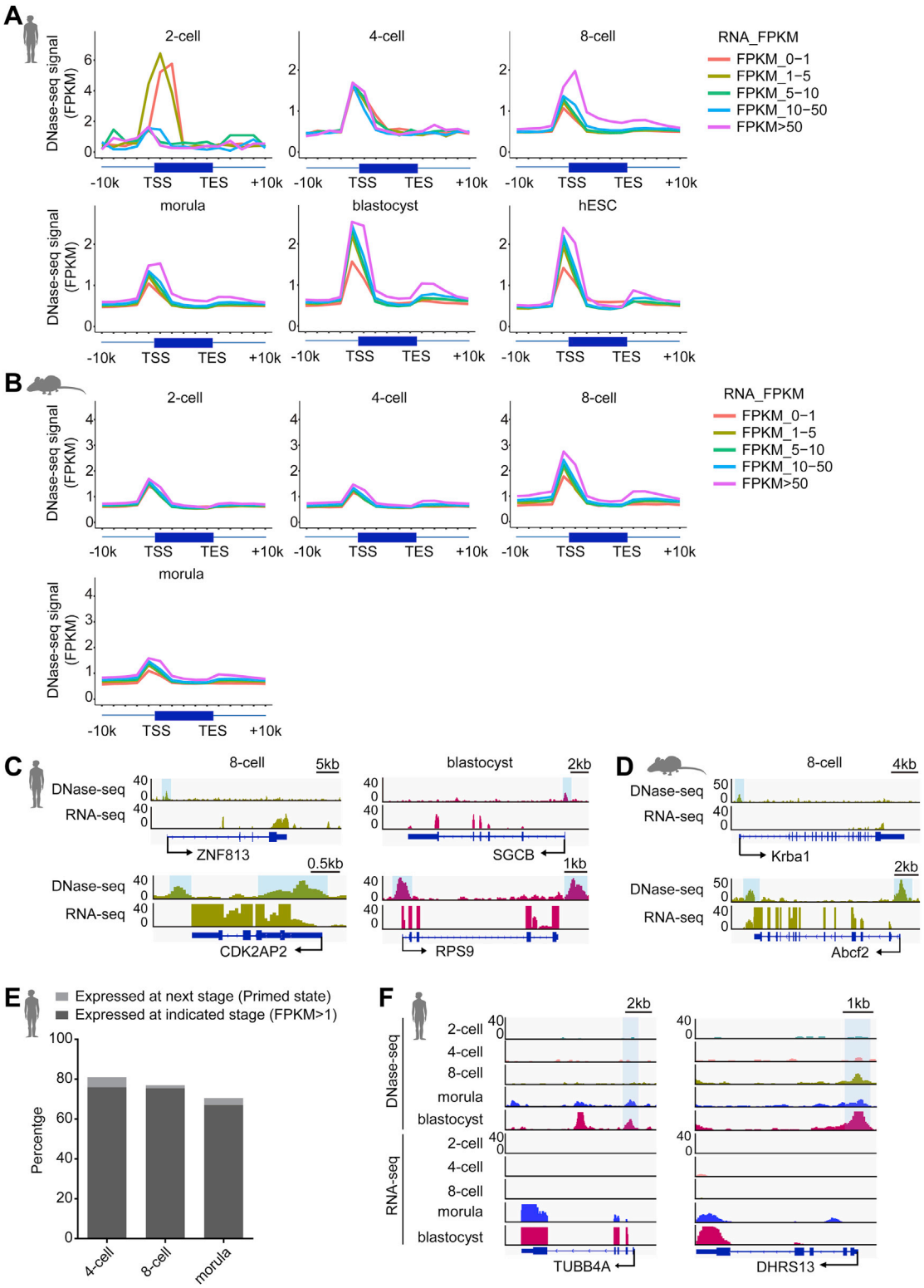


Figure S3. Gene Expression Is Associated with DNase-Seq Signal around Gene TSS and TES, Related to Figure 2

(A and B) DNase-seq signal around the genes with promoter DHSs at each developmental stage in human (A) and mouse (B). The genes are classified into 5 groups according to their expression levels. Upstream 10kb of TSS, downstream 10kb of TES and gene body are divided into 5 bins respectively. Average DNase-seq FPKM is calculated for each bin.

(legend continued on next page)

(C and D) Genome browser view showing DNase-seq signal around TSS/TES and RNA-seq signal of the representative genes in human (C) and mouse (D) embryos. DHSs are shaded. Genes with low expression level (1 < FPKM < 5) correspond to weak DNase-seq signal around TSS and TES. Genes with high expression level (FPKM > 50) correspond to strong DNase-seq signal around TSS and TES.

(E) The percentage of the genes with promoter DHSs that are expressed at the next stage (primed state) or at the indicated stages.

(F) Genome browser view of DNase-seq and RNA-seq signal of the representative genes with promoter DHSs that are expressed at the corresponding stage (left panel) or are primed to be expressed (right panel) in human embryos. DHSs are shaded.

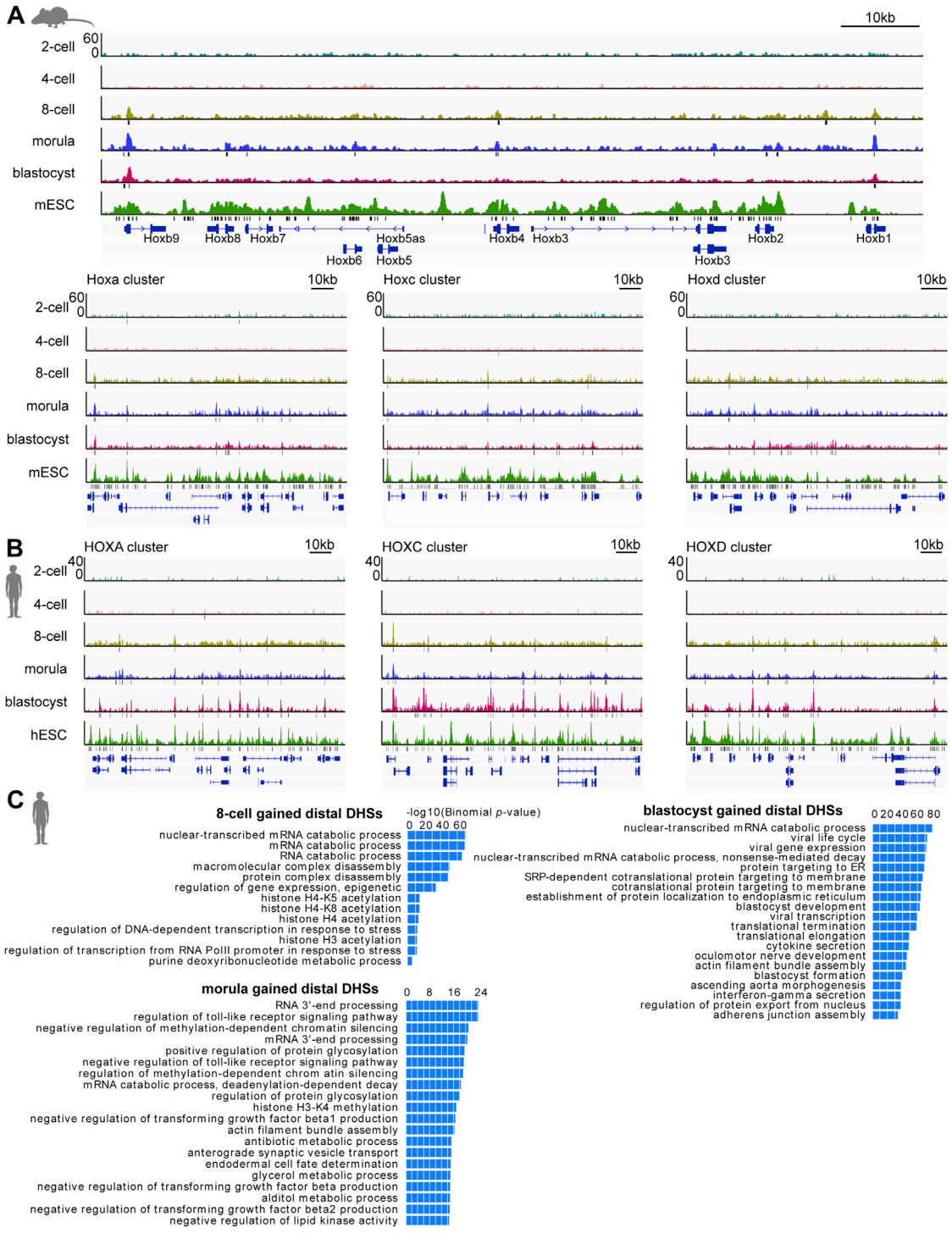


Figure S4. DHS Pattern of *Hox* Cluster in Mouse Embryos and GO Analysis of Gained Distal DHSs in Human Early Embryos, Related to Figure 2

(A) Genome browser view of DHS patterns around *Hox* clusters in mouse early embryos. DHSs are labeled with bars below.

(B) Genome browser view of DHS patterns around *HOXA*, *HOXC* and *HOXD* clusters in human early embryos. DHSs are labeled with black bars below.

(C) GO analysis of gained distal DHSs associated genes at each stage. GREAT online tool was used to analyze the nearest genes within 100kb around distal DHSs.

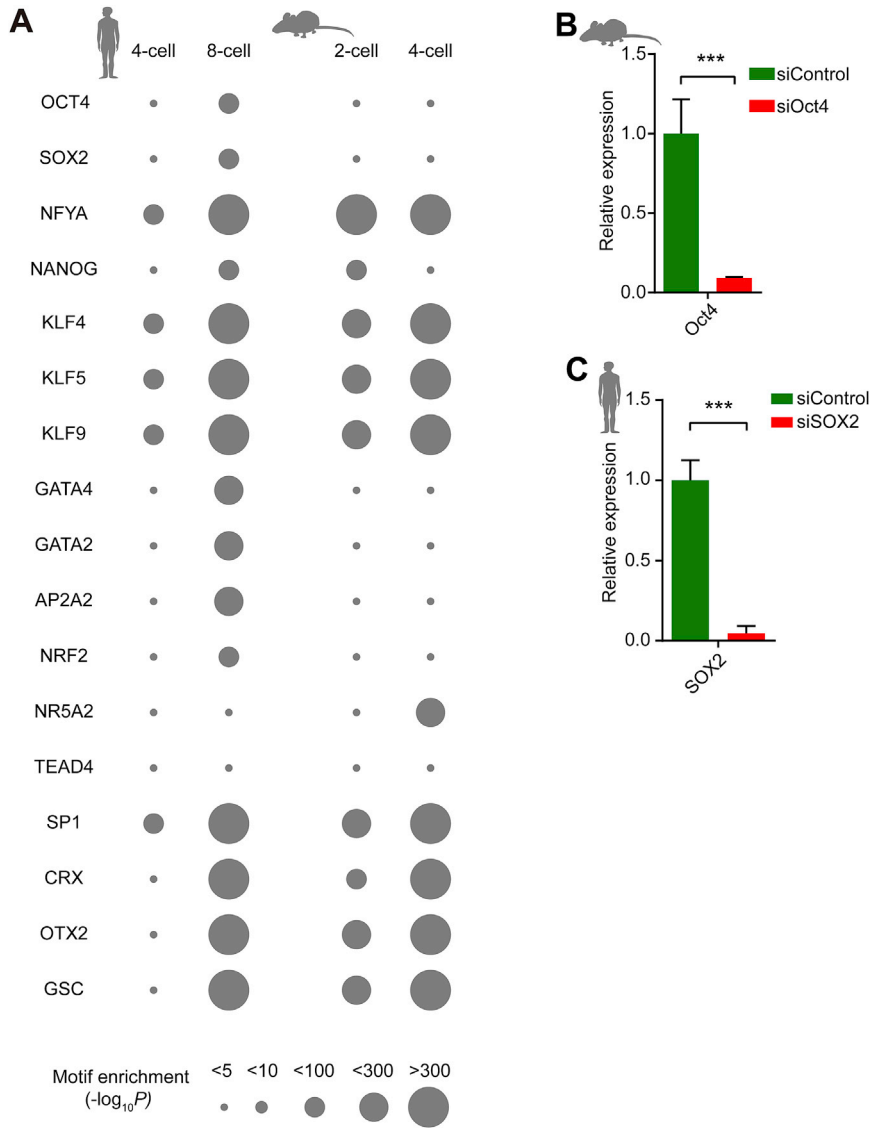


Figure S5. Motif Dynamics, Related to Figure 3

(A) Enrichment scores of transcription factor binding motifs for the DHSs in human and mouse embryos.

(B) Relative expression level of Oct4 in mouse 4-cell embryos after control and Oct4 knockdown. RNA-seq data were from two replicates for both control and Oct4 KD. Two replicates used two independent siRNAs for Oct4. Error bars denote the standard errors. *** represents $p < 0.001$ (t test).

(C) Relative expression level of SOX2 in human 8-cell embryos after control and SOX2 knockdown. Error bars denote the standard errors. *** represents $p < 0.001$ (t test).

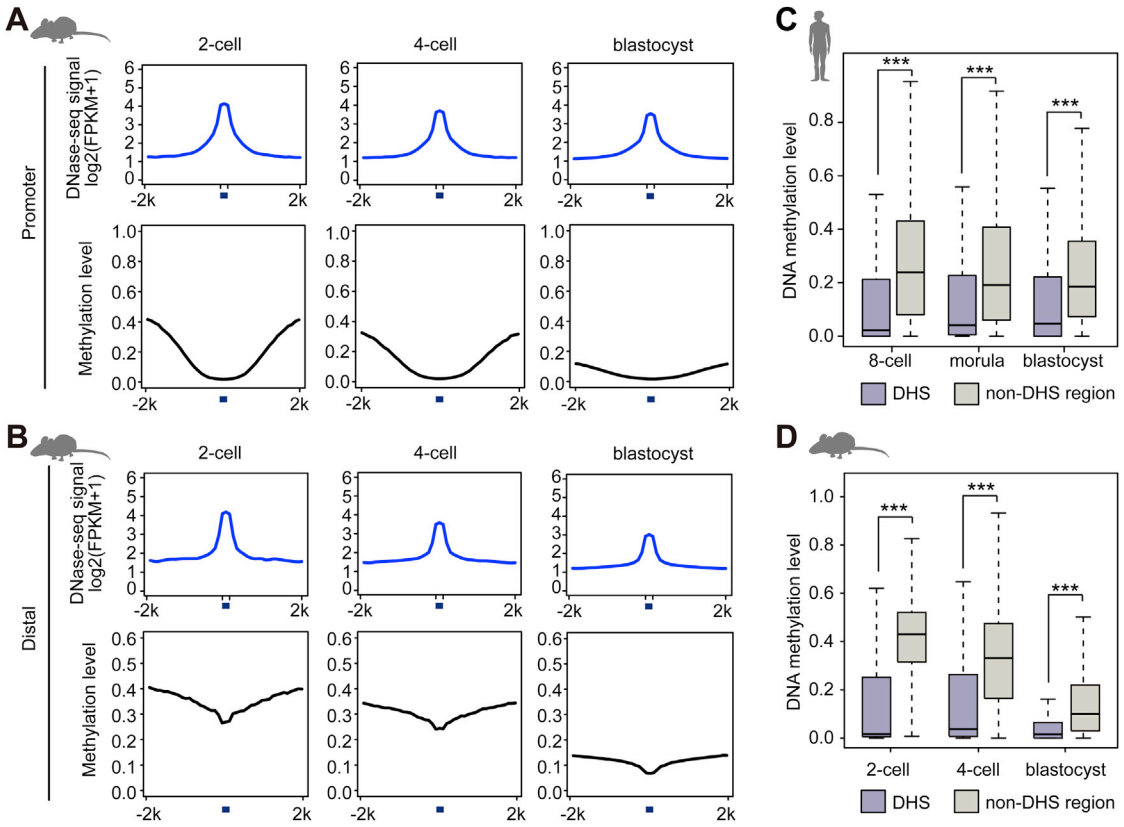


Figure S6. DHSs Are Associated with DNA Methylation in Mouse Early Embryo, Related to Figure 4

(A and B) DNase-seq signal and DNA methylation level around DHSs in promoter (A) and distal (B) regions at each stage in mouse embryos. Blue bars represent the positions of DHSs. The upstream 2kb region and downstream 2kb region are divided into 20 bins respectively. DHS region is divided into 3 bins. The average DNase-seq signal and DNA methylation level are calculated in each bin.

(C and D) The DNA methylation levels of DHS region and non-DHS region in human (C) and mouse (D) early embryos. Boxes and whiskers represent the 25th/75th percentiles and 1.5 × the interquartile range, respectively. Wilcoxon test is used for statistical test. *** represents $p < 0.001$.

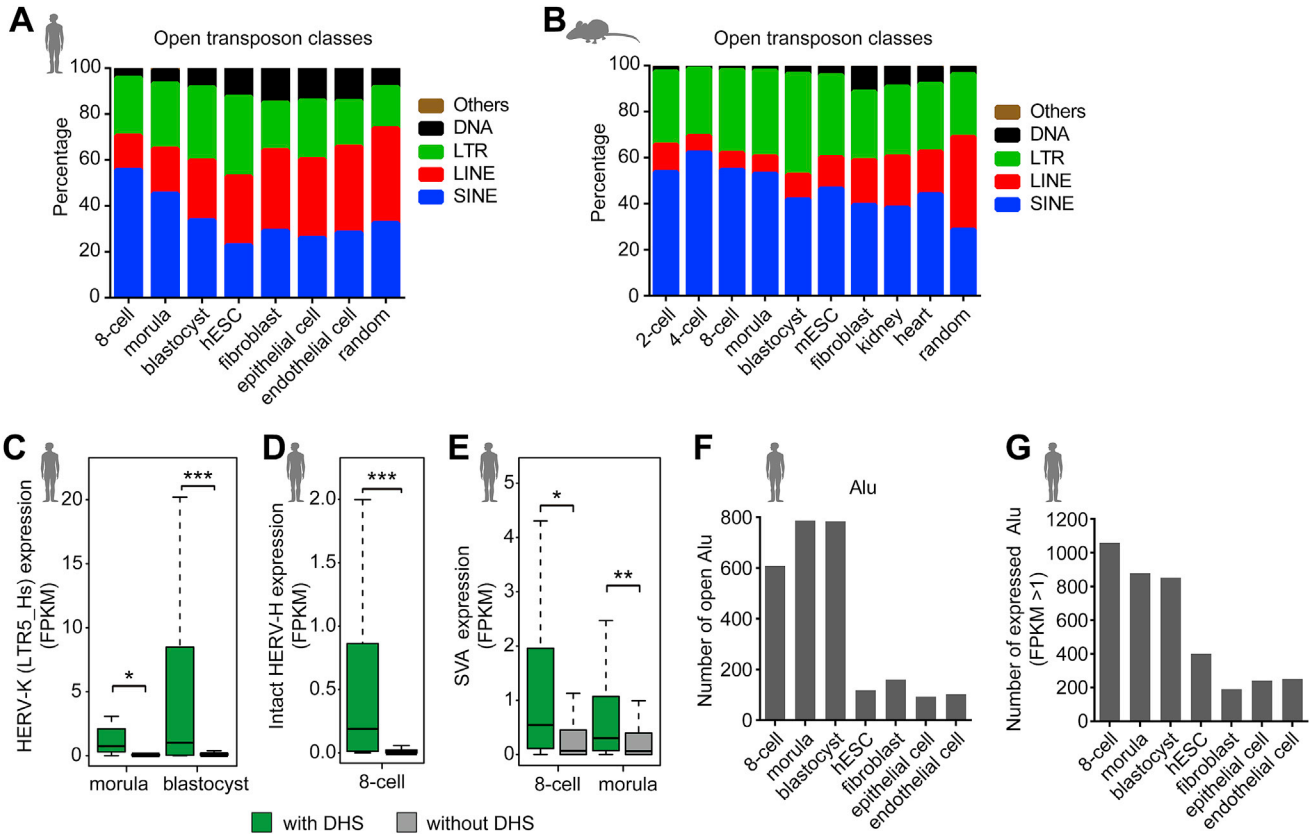


Figure S7. The DHSs in Transposon Elements, Related to Figure 6

(A) Bar plots showing the percentages of different classes of transposons harboring DHSs in human embryos, H1 hESC and differentiated cells. Open transposons refer to transposons with DHSs.

(B) Bar plots showing the percentages of different classes of transposons harboring DHSs in mouse embryos, E14 mESC and the differentiated tissues.

(C–E) RNA expression level of the transposons with DHSs and without DHSs at the indicated stages in human. HERV-K flanked by LTR5_Hs (C), intact HERV-H flanked by LTR7 (D) and SVA (E) were analyzed. Boxes and whiskers represent the 25th/75th percentiles and 1.5 × the interquartile range, respectively. Wilcoxon test is used for statistical test. *** represents $p < 0.001$, * represents $p < 0.05$.

(F) The number of the Alu harboring DHSs in human embryos, H1 hESC and the differentiated cells.

(G) The number of the expressed Alu (FPKM > 1) in human embryos, H1 hESC and the differentiated cells.