SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Preparation of mouse preimplantation embryos and oocytes

All animal studies were performed in accordance with guidelines of the Institutional Animal Care and Use Committee at Harvard Medical School. MII-stage oocytes were collected from 8 week-old superovulated B6D2F1/J (BDF1) females by injecting 7.5 I.U. of PMSG (Millipore) and hCG (Millipore). For *in vitro* fertilization (IVF), MII oocytes were transferred into HTF medium supplemented with 10 mg/ml bovine serum albumin (BSA; Sigma-Aldrich) and inseminated with activated spermatozoa obtained from the caudal epididymides of adult JF1/MsJ male mice (Jackson laboratory, 003720). Spermatozoa capacitation was attained by 1 hour incubation in HTF medium. Six hours post-insemination (hpi), fertilized oocytes were transferred and cultured in KSOM in a humidified atmosphere of 5% CO_2/95% air at 37.8 ℃. The 1-cell, 2-cell, 4-cell, 8-cell, and morula stage embryos were harvested at 12, 30, 46, 60, 78 hpi, respectively. The numbers of the embryos collected for DNase-seq are 100, 100, 50, 25, and 10, respectively. For preparation of the 1-cell stage zygotes, BDF1 sperm was used. For preparation of oocytes, the ovaries were harvested from 14 day-old BDF1 female mice, punctured with a 30-gauge needle; 150 oocytes were collected for DNase-seq.

Removal of polar bodies from embryos

Just before sampling the preimplantation embryos for DNase-seq, the polar bodies were removed to avoid genomic contamination. The embryos were briefly treated with the Acid Tyrode's solution (Sigma-Aldrich) supplemented with 0.5% polyvinylpyrrolidone (PVP, Irvine Scientific, 90123) and 50 mM NaCl to gently remove zona pellucida. After washing with M2 media (Millipore), the embryos were incubated in Trypsin/EDTA (Life technologies) supplemented with 0.5% PVP and 50 mM NaCl for 2-3 min on a multi-test glass slide (MP Biomedicals, 6040805). Gentle pipetting with a narrow-bore glass pipette enabled dissociation of polar bodies from the embryos. The dissociated embryos were then washed with M2 media followed by washing with PBS containing 0.2% BSA and transferred into an Eppendorf LoBind 1.5 ml tube (Eppendorf).

Oct4 knockdown by siRNA injection into 1-cell embryos

MII oocytes harvested from 8-week-old BDF1 females were inseminated with BDF1 sperm. Fertilized oocytes at 7 hpi were injected with 2 μM of a custom designed siRNA against Oct4 (siOct4#1, 5'-AAGGAUGUGGUUCGAGUAUGG-3') (Hay et al., 2004), a Silencer Select siRNA specific for Oct4 (Life technologies, siOct4#2, s71991) or 2 μM ON-TARGETplus non-targeting control siRNA #1 (Control siRNA) (Dharmacon, #D-001810-01-05) using a Piezo impact-driven micromanipulator. The 8-cell stage embryos were collected at 60 hpi for DNase-seq. Twenty-five of the 8-cell stage embryos per group were collected at 60 hpi for DNase-seq.

Nfya knockdown by siRNA injection into GV-stage oocytes

Fully grown GV-stage oocytes were obtained from 8- to 12-week-old BDF1 mice 44-48 h after injection with 7.5 I.U PMSG. The ovaries were removed and transferred to M2 media. The ovarian follicles were punctured with a 27-gauge needle, and the cumulus cells were gently removed from the cumulus–oocyte complexes using a narrow-bore glass pipette. The oocytes were then transferred into α-MEM (Life technologies, 12571-063) supplemented with 5% FBS, 10 ng/ml EGF, and 0.2 mM 3-isobutyl-1-methylxanthine (IBMX; Sigma–Aldrich).

One hour after collection, GV oocytes were injected with 2 μM of Silencer Select siRNA specific for Nfya (Life technologies, s70571 for siNfya#1, s70570 for siNfya#2) or 2 μM Control siRNA using a Piezo impact-driven micromanipulator (Prime Tech Ltd., Ibaraki, Japan). Two hours after injection, oocytes were transferred to IBMX-free α-MEM to induce meiotic maturation. After maturation for 16 hours, the MII-stage oocytes were transferred into HTF media supplemented with 10 mg/ml BSA, and inseminated with BDF1 sperm. Fertilized oocytes were transferred into KSOM. The 2-cell stage embryos were collected at 30 hpi for DNase-seq (70-90 embryos) and RNA-seq (10-16 embryos) after removing polar bodies.

Physical isolation of maternal and paternal pronuclei at PN3 and PN5 stages

MII oocytes were collected from 8-12 week-old BDF1 females by superovulation. Fertilized oocytes were obtained by *in vitro* fertilization with BDF1 sperm. For PN3-stage pronuclei isolation, zygotes were transferred into M2 media containing 10 μM cytochalasin B (Sigma-Aldrich) at 5 hpi. Zona pellucidae were cut by a Piezo impact-driven micromanipulator and the pronuclei were isolated from the zygotes. At 7.5 hpi (PN3-stage), they were washed with 0.2% BSA/PBS, transferred into Eppendorf LoBind 1.5 ml tubes, and treated with DNase I. The numbers of PN3 pronuclei collected for DNase-seq were 140 each. For PN5-stage pronuclei isolation, zygotes were transferred into M2 media containing 10 μM cytochalasin B at 10 hpi. The isolated pronuclei were washed with 0.2% BSA/PBS, transferred into Eppendorf LoBind 1.5 ml tubes, and treated with DNase I at 12 hpi (PN5-stage). The numbers of PN5 pronuclei collected for DNase-seq were 170 each. We confirmed, by BrdU incorporation assay, that 7.5 hpi and 12 hpi were in the middle of S-phase and in G2-phase, respectively (data not shown). The parental pronuclei were distinguished by (1) the distance from the second polar body and (2) the size of the pronucleus. We selected PN3-stage as the earliest time point for pronuclei isolation because 5 hpi was the earliest time when the parental pronuclei were clear enough to be distinguished under microscopy.

Whole mount immunostaining

Embryos were fixed in either PBS containing 3.7% paraformaldehyde (PFA) (for the Nfya antibody) or PBS containing 3.7% PFA and 0.2% Triton X-100 (for the Oct4 antibody) for 20 min, and then washed with PBS containing 10 mg/ml BSA (PBS/BSA). After permeabilization with 0.5% Triton X-100 for 15 min (for the samples fixed by PFA without 0.2% Triton), embryos were washed in PBS/BSA and treated with primary antibodies overnight at 4 ℃. The primary antibodies used are rabbit anti-Nfya (1:1600, Thermo Fisher Scientific, PA5-28990) and goat anti-Oct4 (1:500, Santa Cruz Biotechnology, sc8628). After washing with PBS/BSA for 30 min, samples were incubated with Alexa Flour 568 donkey anti-rabbit IgG (1:250, Life technologies) or Alexa Flour 568 donkey anti-goat IgG (1:250, Life technologies) for 1 h at room temperature. The embryos were then mounted on a glass slide in Vectashield anti-bleaching solution with DAPI (Vector Laboratories, Burlingame, CA). Fluorescence was imaged under a laser-scanning confocal microscope with a spinning disk (CSU-10, Yokogawa) and an EM-CCD camera (ImagEM, Hamamatsu). All images were acquired and analyzed using the Axiovision software (Carl Zeiss).

The fluorescent signal intensity was quantified with Axiovision software. Briefly, the signal intensity within nuclei was determined, and the cytoplasmic signal was subtracted as background signal. The average signal of the blastomeres was defined as the signal intensity of each embryo. The average of signal intensity in control embryos was set as 1.0.

Low-input DNase-seq

Fresh cells were collected and resuspended in 36 μl lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Triton X-100) and incubated on ice for around 5 minutes. DNase I (10 U/μl, Roche) was added to final concentration of 80 U/ml (for 1-cell embryo and oocyte samples) or 40 U/ml (for all the other samples in this study) and incubated at 37 ℃ for exactly 5 minutes. The reaction was stopped by adding 80 μl Stop Buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 0.15% SDS, 10 mM EDTA) containing 2 μl Proteinase K (20 mg/ml, Life technologies). Then 20 ng circular carrier DNA (pure plasmid DNA without mammalian sequence which was purified again with 0.5x Beckman SPRI beads to remove trace amount of small fragments) was added. The mixture was incubated at 50 ℃ for 1 hour. DNA was purified by extraction with phenol-chloroform and precipitated by ethanol in the presence of linear acrylamide (Life technologies) overnight. Precipitated DNA was resuspended in 55.5 μl TE (2.5 mM Tris, pH 7.6, 0.05 mM EDTA) and the entire volume is used in sequencing library construction. Sequencing library was prepared using NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs) with the exception that adaptor ligation step was done with 0.03 μM adaptor in the ligation reaction for 30 minutes at 20 ℃ and PCR amplification was done using Kapa hifi hotstart readymix (Kapa Biosystems) for 8-cycles. The PCR product was purified with 1.3 volume SPRIselect beads (Beckman Coulter) and then size selected with 0.7 volume plus 0.7 volume SPRIselect beads and eluted in 24 μl TE. The number of cycles for the second PCR amplification was measured by qPCR using 1 ul of the 1:1,000 diluted eluent. Then the remaining 23 μl sample was amplified

with Kapa hifi hotstart readymix (7 to 11 cycles for the samples analyzed here). The PCR product was again purified and size selected using SPRIselect beads as the steps after the first PCR amplification. The libraries were sequenced on a Hiseq2500 with single-end 100 bp reads (Illumina).

For qPCR validation of DHSs, samples (32 4-cell embryos or 10 morula embryos) were treated as described in liDNase-seq until resuspending the ethanol precipitated DNA in 15 μl TE. DNase I minus control samples with the same number of embryos was treated the same way except lysis buffer was used instead of DNase I diluent. Then the DNA was used for multiplexed amplification with the primers listed in Table S6 for 14 cycles in 50 μl volume with SsoFast EvaGreen Supermix (Bio-Rad) in the presence of 4% DMSO. Then 0.2 ul of the pre-amplified product was used for quantification with technical replicates. The relative amount of intact DNA at DHSs amplifiable after DNase I treatment was quantified against DNase I free control samples. The primers used here are included in Table S6.

RNA-seq

The embryos were directly lysed and used for cDNA synthesis using SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (Clontech). After cDNA amplification, the samples were fragmented to average size of 150-200 bp using Covaris sonicator (Covaris). Sequencing libraries were prepared with the fragmented DNA using NEBNext Ultra DNA Library Prep Kit for Illumina according to manufacturer's instruction (New England Biolabs). The libraries were sequenced on a Hiseq2500 with single-end 50 bp reads (Illumina).

Genomic datasets

In this study, 32 liDNase-seq datasets were generated including: two 1-cell embryo libraries, two 2-cell embryo libraries, two 4-cell embryo libraries, two 8-cell embryo libraries, two morula embryo libraries, two mouse ESC libraries, two 14 day growing oocyte library, two sperm library, two maternal PN3 pronuclei library, two paternal PN3 pronuclei library, two maternal PN5 pronuclei library, two paternal PN5 pronuclei library, two siRNA control 2-cell embryo libraries, two Nfya KD (siNfya#1 and siNfya#2) 2-cell embryo libraries, two siRNA control 8-cell embryo libraries and two Oct4 KD 8-cell (siOct4#1 and siOct4#2) embryo libraries. Five RNA-seq datasets were generated including two siRNA control 2-cell embryo libraries, two Nfya KD (siNfya#1 and siNfya#2) 2-cell embryo libraries and one 1-cell embryo library. The detailed reads information for these dataset can be found in Table S1. In addition to the liDNase-seq and RNA-seq libraries prepared in this study, we downloaded the DNase-seq dataset of ESCs from mouse ENCODE project (Vierstra et al., 2014), the RNA-seq datasets of mouse early embryos (Xue et al., 2013) and the Oct4 ChIP-seq dataset of mouse ESCs (Whyte et al., 2013).

liDNase-seq data analysis

Reads of liDNase-seq data were mapped to the mouse genome (mm9) using Bowtie v0.12.9. We use '-m 1' parameter for unique hits mapping. The reads with mapping quality (MAPQ) ≤ 10 or redundant reads that mapped to the same location with the same orientation were removed from further analysis in each library.

The DHS peaks in liDNase-seq data were identified by Hotspot program with FDR <= 0.01 (John et al., 2011). The tag density at each DHS was calculated by normalizing the number of reads in the DHS to the total number of reads in the library (possibility of a tag located on a base-pair per million reads). For identifying DHS peaks in preimplantation embryos, we used an additional stringent cutoff requirement of RPM > 1 for the peak in both replicates. The Pearson correlation coefficient (r) of tag densities at genome-wide DHS between two replicates was calculated to show the correlation between different liDNase-seq libraries. To compare the 30-cell and 100-cell liDNase-seq data with the ENCODE DNase-seq data, the peaks identified in the ENCODE data were extended ±1 kb from the summit of the peak if the peak size was <2 kb and overlapping peaks were merged. Then the number of reads in each DHS for the 30-cell and 100-cell liDNase-seq libraries was counted. The Pearson correlation coefficient (r) of tag densities at genome-wide DHS was calculated. The RefSeq gene assembly (mm9) from the UCSC Genome Browser database was used as genomic feature distribution analysis. The regions ±0.5 kb around the TSS were defined as promoters in this study except for the allele specific DHS analysis. Gene ontology enrichment

analysis and motif analysis: to predict the function of dynamic DHS during preimplantation embryo development, gene ontology analysis were performed using GREAT (McLean et al., 2010). Gene ontology analysis for 1-cell promoter DHS related genes, 8-cell stage primed promoter genes and morula stage primed promoter genes were performed with DAVID (Huang et al., 2008). Enrichment of known motifs was analyzed using the Homer tool (http://homer.salk.edu/homer/motif/).

Allele specific DHS analysis

To identify SNPs between the BDF1 (B6 x DBA/2J) genome and the JF1 genome, we obtained 5.1 million SNPs between B6 and DBA/2J from the Sanger Institute (http://www.sanger.ac.uk/science/data/mouse-genomes-project), and 12.5 million SNPs between B6 and JF1 from (ftp://molossinus.lab.nig.ac.jp/pub/msmdb/For_Seq_Analysis/list_of_variations/). Comparing these two SNPs lists resulted in a total of 11.1 million SNPs between the BDF1 genome and the JF1 genome. DNase-seq reads of the BDF1/JF1 hybrid samples were mapped to reference genome (B6) allowing up to 3 mismatch per read using Bowtie v0.12.9. Read covering SNPs were assigned to its parental origins using the SNP information. Mouse imprinting genes data were obtained from mouse book (http://www.mousebook.org/mousebook-catalogs/imprinting-resource). To validate this analysis pipeline, we processed the DNase-seq reads from 8-cell control knockdown embryos, which were derived from BDF1/BDF1 background. Using this data, we obtained 1,552,511 and 5,503 reads containing BDF1 and JF1 SNPs, respectively. The ratio between BDF1 and JF1 reads was 282.1. In contrast, when we processed the DNase-seq reads from BDF1/JF1 hybrid 8-cell embryos, the ratio between BDF1 and JF1 reads was 1.5. Thus, our pipeline allows detection of the JF1 allele with a low false discover rate (about 0.5%). To include most of the known imprinting control regions, we used the regions ±1.5 kb around the TSS as promoters for imprinting genes. Among 150 imprinting genes, 109 of them contain at least one SNP between the BDF1 and the JF1 genome in the promoter region.

RNA-seq data analysis

Reads of RNA-seq were mapped to the mouse genome (mm9) with TopHat v2.0.6. All programs were run with default parameters unless otherwise specified. Uniquely mapped reads were subsequently assembled into transcripts guided by the reference annotation (UCSC gene models) with Cufflinks v2.2.1. Expression level of each gene was quantified with normalized FPKM (fragments per kilobase of exon per million mapped fragments). The 2-cell activated genes were identified by comparing the 2-cell control RNA-seq with the 1-cell RNA-seq data using five-fold change as a cut-off.

Statistical analyses and data visualization

Statistical analyses were implemented with R (http://www.r-project.org/). Pearson's r coefficient was calculated using the cor function with default parameters. Independent 2 group Wilcoxon rank sum test were used to compare distributions using the wilcox.test function in R. The Figure 1C were generated using *heatmapr* function in Cistrome (Wang et al., 2014). Position-wise coverage of the genome by sequencing reads was determined by normalizing to the total unique mapped reads in the library using macs2 v2.1.0 (Zhang et al., 2008) and visualized as custom tracks in the UCSC genome browser.

SUPPLEMENTAL REFERENCES

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protocols *4*, 44-57.
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. Nature genetics *43*, 264-268.
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. Nature biotechnology *28*, 495-501.
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.
Wang, Q., Huang, J., Sun, H., Liu, J., Wang, J., Wang, Q., Qin, Q., Mei, S., Zhao, C., Yang, X., *et al.* (2014). CR Cistrome: a ChIP-Seq database for chromatin regulators and histone modification linkages in human and mouse. Nucleic acids research *42*, D450-458.
Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307-319.
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.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology *9*, R137.