Supplementary information, Data S1

Materials and Methods

Animals

All animal experiments were performed in compliance with the guidelines of the Animal Ethical Committee of the Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. Pregnant ICR outbred mice were identified by checking the existence of vaginal plug after mating. Developmental stage is counted as 0.5 days post-coitum (dpc) since the appearance of vaginal plug. Pregnant mice were killed by cervical dislocation at nominal day 6.5, 7.0 and 7.5 dpc. And embryos in decidua were removed from the implantation site. Embryos were then incubated in Dulbecco's modified Eagle's medium. The removal of decidua and Reichert's membrane were performed using surgeon tweezers. The developmental stage and morphology of embryos were reconfirmed using Theiler stage methodology. The acquired embryos were then dissected into essential parts feasible for the following experiments.

Embryo transcript preparation, gene expression analyses, cDNA sequencing library construction and sequencing

cDNA pool of each sample were acquired following Smart-seq2 protocol [11] for limited cell number samples. PCR primers for evaluating gene expression in mouse embryos were designed to test the abundance ratio of related transcripts. Quantitative PCR (qPCR) reaction was performed using SYBR green qPCR master mix (Sigma) in Eppendorf instrument with the manufacturer's instruction. The gene expression levels were normalized according to the housekeeping gene of Gapdh. cDNA sequencing libraries were generated using the Nextera XT DNA Library Preparation Kit (FC-131-1024) for the Illumina platform. Paired-end 125 bp sequencing was performed on a Hiseq2500 platform at the Berry Genomics Co., Ltd.

Embryo ChIP sample preparation and library construction

Essential parts of embryos were collected and pooled separately to reach a scale of 10,000 cells per analysis. To specify, 20 E6.5 embryos were dissected into extraembryonic and embryonic parts, and marked as E6.5 ExE and E6.5 Epi; 10 E7.0

embryos were dissected into extraembryonic, anterior and posterior parts, and marked as E7.0 ExE, E7.0 A and E7.0 P; 5 E7.5 embryos were dissected into extraembryonic, anterior, posterior and anterior mesoderm parts, and marked as E7.5 ExE, E7.5 A, E7.5 P and E7.5 AM. ChIP experiments were carried out by a modified published protocol [3]. Briefly, dissected embryos were fixed in 1% formaldehyde solution and quenched by 0.125 M glycine, then rinsed by DPBS twice before frozen in liquid nitrogen. After thawing on ice, fixed tissues were lysed with Solution I (10 mM HEPES, pH 7.9, 0.5 % NP40, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT) and SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.0) for 15 min each, fragmented to a size range of 200-500 bp by using Bioruptor Pico (Diagenode, Belgium). Solubilized fragmented chromatin was immunoprecipitated with antibody against H3K27me3 (Millipore 07-449). Antibody-chromatin complexes were pulled down using protein G beads (Dynabeads 10004D), washed several times and then eluted. Reverse crosslinking was performed subsequently under 65 °C for at least 4 h. Chromatin mixture solution was treated with RNase A and Proteinase K to remove residual RNAs and proteins. Finally, fragmented DNA was extracted with phenol-chloroform and precipitated with ethanol.

ChIP DNA was finally solved in nuclease free water and quantified using Qubit. Sequencing libraries were generated by using NEBNext Ultra DNA library preparation kit (NEB E7370).

Reduced Representation Bisulfite Sequencing

Samples were collected as described previously. The acquired embryo part was subjected to 200 μL lysis buffer (20 mM Tris-HCl, pH 8.0, 2 mM EDTA, 20 mM KCl, 0.3% (v/v) Triton X-100, 1 mg/mL proteinase K) and vortex for 5 s, quick spin down for 2 s. Then lysis the embryo part at 56 °C for at least 2 h. The resolved chromatin will be treated with RNase A to remove residual RNAs at 37 °C for 1 h. Genomic DNA from embryo parts would be acquired after phenol-chloroform extraction and ethanol precipitation. Further manipulation of genomic DNA concerning reduced representation bisulfite sequencing was performed as described previously [6].

Embryo explant culture

E7.0 embryos are dissected into anterior, posterior and extraembryonic parts and

cultured in N2B27 medium in 24-well cell culture cluster plate. The 24-well plate was pre-coated with fetal bovine serum (FBS) and washed twice with PBS (Gibco) before use. Chemical inhibitor of PRC2 complex, EPZ005687 (Selleck, S7004), were added when the explants attached to the bottom surface of the plate. Extracting RNA from explant samples at relevant time points. Further analysis will be performed as required.

Sequencing data quality control

All of the de-multiplexed sequencing reads were first cleaned to remove any artificial sequences, such as sequencing adapters introduced during the experimental process, and reads with more than 10% low quality bases were also discarded.

RNA-Seq data processing

For the RNA-Seq data, the annotation of the transcriptome was defined by combining the UCSC mm10 RefSeq genes and the GENCODE vM5 lncRNAs. Then the cleaned reads were aligned to the mm10 reference genome using TopHat (version 2.0.9) [12] with the default parameters. HTSeq [13] were used to count the number of

reads mapped in each annotated gene based on the mapping results. These results were further used for the calculating of differentially expressed genes using DESeq2 [14] package in bio-conductor. Furthermore, for the quantification of RNA expression, the gene expression levels (RPKM) of each sample were calculated based on the number of reads for each gene counted by HTSeq:

$$\mathrm{RPKM}_{gen,s} = \frac{\mathrm{Counts}_{\mathrm{gen,s}}}{10^9 \times \mathrm{Length}_{\mathrm{gen}} \times \sum_{gen} \mathrm{Counts}_{\mathrm{gen,s}}}$$

where gen for genes and s for samples.

Then genes with variation above 1 among all samples were selected and the ComBat function in sva package of Bioconductor [15] were used to remove the system error introduced by different sequencing batches. The normalized RPKM after removing batch effects were used for the following analyses, including clustering and comparison with ChIP-Seq results.

ChIP-Seq data processing

The ChIP-seq analysis were performed using the AQUAS histone ChIP-seq processing pipeline (https://github.com/kundajelab/TF_chipseq_pipeline.):

python chipseq.py -type histone --subsample-chip 30M --subsample-ctl 30M -species mm10 -fastq1_1 file1_R1 -fastq1_2 file1_R2 -fastq2_1 file2_R1 -fastq2_2 file2_R2 -ctl_fastq_1 control_R1-ctl_fastq_2 control_R2 -out_dir out_dir

In details, this pipeline first mapped fastq reads into bam files using bwa(0.7.13). Then the duplicated reads and low-quality reads were filtered, and the clean reads were converted into tagalign files. The peaks were then called based on the tagalign files using MACS2(2.1.0) with IDR criteria for pseudo replicates of each sample as well as pooled pseudo replicates. Finally, the signal tracks were generated using this pipeline and we used the fold-change track for further analysis.

The DNA methylation data processing

The trimmed reads were mapped to mm10 mouse genome following our modified bisulfite sequencing pipeline [6]. PCR duplicates are removed using samtools and DNA methylation level of every cytosine was calculated based on our previous publications [2]. This results are further used for differential methylation regions (DMRs) detection using RADMeth [16]. The pairwise comparisons of consecutive samples on a single

CpG level are calculated using a b-binomial model and the b difference distribution. Differential CpGs without neighbors are then embedded into a 100 bp region surrounding each CpG. Next, differential methylation analysis was repeated on the region level using a random effects model. Only regions significant at a Benjamini-Hochberg q-value below 0.05, an absolute methylation difference above 0.15, above 100 bps and containing at least 5 reads were considered as differentially methylated.

References

11 Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols* 2014; **9**:171-181.

12 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013; **14**:R36.

13 Anders S, Pyl PT, Huber W. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015; **31**:166-169.

14 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion

for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**:550.

15 Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012; **28**:882-883.

16 Dolzhenko E, Smith AD. Using beta-binomial regression for high-precision differential methylation analysis in multifactor whole-genome bisulfite sequencing experiments. *Bmc Bioinformatics* 2014; **15**.