Open
LETTER TO THE EDITOR

Cell Research (2014) :1-5.

www.nature.com/cr

H3K4me3 epigenomic landscape derived from ChIP-Seq of 1 000 mouse early embryonic cells

Cell Research advance online publication 2 September 2014; doi:10.1038/cr.2014.119

Dear Editor,

Epigenetic regulation is crucial to the establishment and maintenance of the identity of a cell. Recent studies suggest that transcription is implemented amongst a mixture of various histone modifications [1]. It has also been recognized that to interrogate function of genetic information, comprehensively systematic profiling of the epigenome in multiple cell stages and types is required [2]. Chromatin immunoprecipitation (ChIP) has become one of the most critical assays to investigate the complex DNA-protein interactions [3]. Combined with profiling technologies such as microarrays (ChIP-on-chip) or high-throughput sequencing (ChIP-Seq), this assay becomes a great tool to study the epigenetic regulatory networks in cells [4-6]. However, the ChIP process produces limited amount of DNA due to the low yield of antibody pull-down, DNA damage during fragmentation and cleavage of DNA-protein complex, and complicated downstream analysis [7]. The conventional approaches have to consume a considerable amount of samples, typically 10^6 - 10^7 cells, to overcome this low-yield issue and obtain reliable results [5]. This limitation also restricts ChIP applications from precious primary tissue samples such as early embryonic cells or rare tumor stem cells.

ChIP-Seq, compared with ChIP-on-Chip, deeply sequences the target DNA fragments and generates highly comprehensive data with higher resolution, fewer artifacts, greater coverage and larger dynamic range [6]. Although recent application of automated microfluidic ChIP (AutoChIP) was successfully performed using 2 000 cells through locus-specific analysis by qPCR [8], such assays do not achieve the comprehensiveness afforded by DNA sequencing approaches. Recently, several approaches have been developed to perform ChIP-Seq using as low as 10 000 or even only 5 000 cells [7, 9-11, 14]. However, all of these methods rely on ChIP reactions in tens of microliters and preamplification of ChIP product before sequencing library preparation, either through linear amplification (by *in vitro* transcription) or exponential amplification (by PCR), both of which potentially introduce significant bias. Adli *et al.* [7] reported a modified protocol to realize the ChIP-Seq using 10 000 cells by revising the random primers used in amplification to reduce the primer self-annealing, with an optimized PCR condition to cover the GC-rich regions. Ng *et al.* [14] developed another protocol to perform ChIP-Seq of H3K4me3 modification using 10 000 mouse primordial germ cells, requiring pre-amplification before the sequencing library preparation. Sachs *et al.* [15] reported a chromatin immunoprecipitation study with low number of cells without pre-amplification, however, it needs at least 50 000 cells as starting material. Here we present a new method that implements a microfluidic device to facilitate the ChIP process, providing a technology to obtain the high-quality ChIP-Seq data from merely 1 000 mammalian cells, with no need of pre-amplification. The whole ChIP process has been greatly shortened to 8 h. Through this method, we have accomplished, for the first time, a rapid, semi-automated, and highly sensitive ChIP assay to investigate the genome-wide landscape of histone modification H3K4me3 using 1 000 mouse epiblast cells at E6.5, and found that the H3K4me3 landscape of post-implantation epiblast is more similar to that of the mEpiSCs than that of mESCs.

Microfluidic devices are the ideal reaction systems for handling small number of cells. We fabricated our PDMS-based device to treat four samples in parallel on a single chip (Figure 1A). Each reaction pipeline could accept no more than 1 200 cells. We first performed the whole ChIP assay on a microfluidic device, including trapping the magnetic beads that pre-incubated with antibody, binding of the sonicated chromatin fragments with H3K4me3 to the beads, and washing away the chromatin fragments without H3K4me3 modification. Alternatively, the micro-device can be used for fragmentation of chromatins through micrococcal nuclease (MNase) treatment on chip, after the cell permeabilization. After the enrichment we collected the DNA out of the device and finished the sequencing library construction in microcentrifuge tubes. We eliminated preamplification of the ChIP product. Instead, we combined the end-repair,

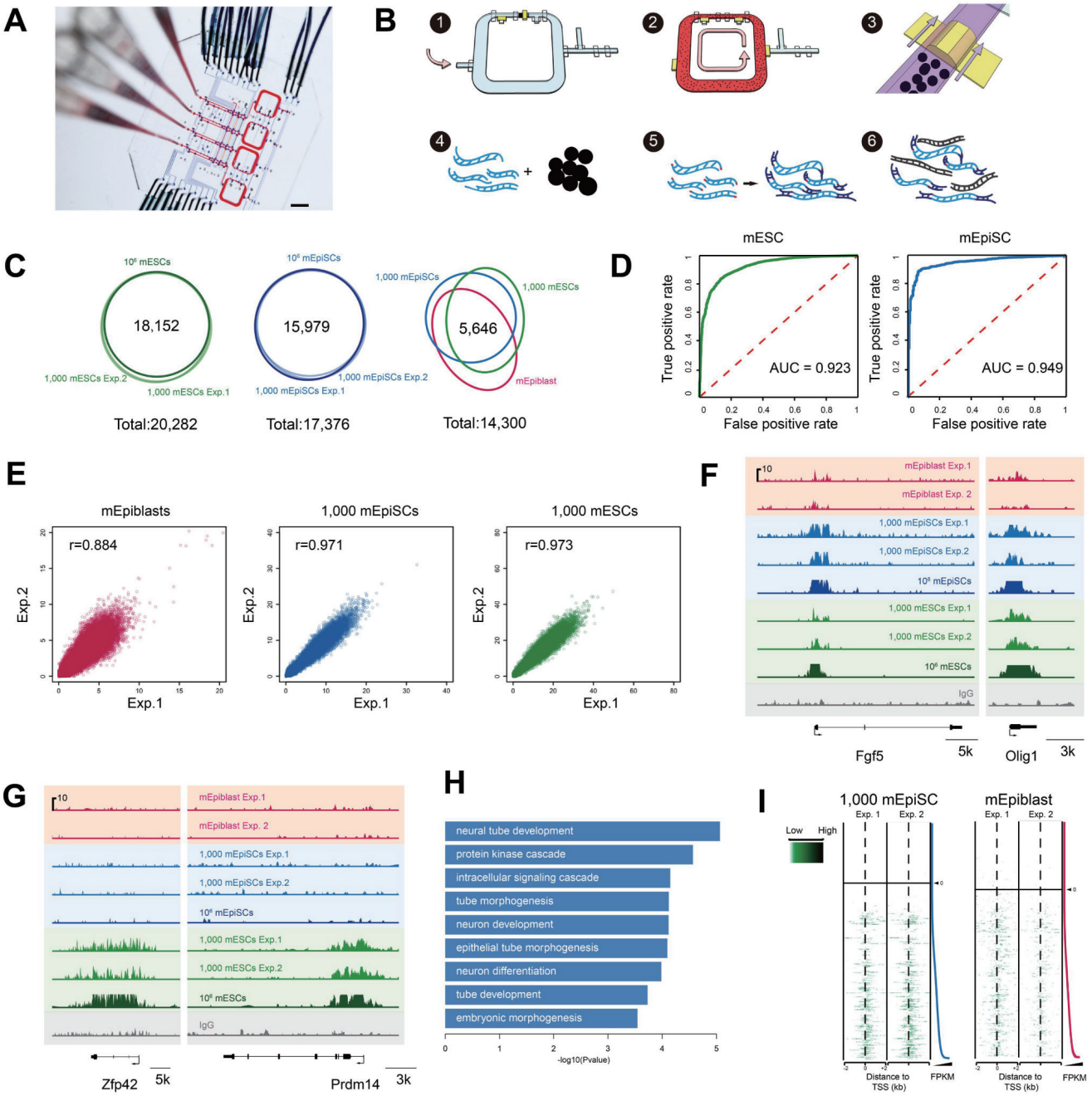


Figure 1 Microfluidic device and the ChIP-Seq from 1 000 mEpiSCs, 1 000 mESCs and epiblast cells of the E6.5 mouse embryos. **(A)** Optical micrograph of a 4-plex microfluidic device with control lines and sample inputs. Scale bar: 5 mm. **(B)** Key operational steps of a ChIP-Seq flow pipeline. Step 1: load the chromatin fragments to fill dead-end flow-channels; Step 2: mixing and immunoprecipitation; Step 3: trap the antibody-functionalized beads (Ab-beads) to form a column; Step 4: release the DNA from the chromatin-antibody-bead complex; Step 5: end-repair, adenylation, and ligation; Step 6: amplification for sequencing. **(C)** The Venn diagram of enriched TSS regions of mEpiSCs, mESCs, and E6.5 epiblast cells. **(D)** The receiver operating characteristic (ROC) curve representing the true positive and false positive rates for the 1 000-cell experiment of mEpiSCs and mESCs. The standards are the enriched TSS regions called from the ChIP-Seq experiments using one million cells. **(E)** The correlation of the enrichment of H3K4me3 markers around TSS regions of epiblast cells of E6.5 mouse embryos, mEpiSCs, and mESCs. Each point represents an individual gene. **(F)** Representative loci of the ChIP-Seq of the H3K4me3 markers for three cell types, E6.5 epiblast cells, mEpiSCs, and mESCs, showing the peaks shared by all these three types of pluripotent cells. **(G)** Representative loci of the ChIP-Seq of the H3K4me3 markers specifically enriched in mESCs. **(H)** Gene ontology terms enriched in epiblast cells of E6.5 mouse embryos and mEpiSCs, compared to mESCs. **(I)** Heatmap showing read distribution around TSS regions of different transcripts ranked by the FPKM.

adenylation, and ligation steps in a one-tube reaction and then used carrier DNA to facilitate high-efficiency purification of the ligated DNA product. Then we used PCR to amplify these ligated DNAs to get nanogram amount of DNA for sequencing. Conventionally, the experimental process is tedious, taking at least two days to complete the sample treatment [3]. However, our new protocol can greatly accelerate the whole process; the complete experimental procedure, including microfluidic-based cell permeabilization, chromatin fragmentation, antibody pull-down, and many washing steps, can be completed within 8 h.

We compared two approache for chromatin fragmentation, ultrasonic shearing and MNase digestion, and mainly used the former approach in our experiments since the result was more robust. The fragmentation step could be performed off-chip using the probe-free sonicator. 10 μl formaldehyde-crosslinked cell suspension was fragmented by the ultrasonic. Then the fragmented chromatin suspension was concentrated from 10 μl to 1 μl by evaporation.

Since the starting material is very limited and the reaction yield is intrinsically low, even losing a small fraction of the target DNA [8] would potentially compromise sequencing efforts. We hence carefully designed the microfluidic channels to sequentially perform the necessary reactions without losing target DNA. We implement the dead-end filling method [12] to transfer the chromatin fragments into a ring-shape chamber to react with antibody-coated beads (Figure 1B). Dead-end filling was practical because PDMS was gas permeable; the air in the ring-chambers was expelled and replaced by liquid within a few minutes. Integrated 3-valve peristaltic pumps circulate the liquid in the ring chambers, facilitating the immunoprecipitation. This step is the most challenging practice in traditional ChIP assays. The challenges come from the incomplete crosslinking, nonspecific adsorption, low-efficiency chromatin fragmentation, and low specificity of binding between the antibodies and histone. However, with reduced reaction volume and active mixing, the efficiency of the whole process has been greatly improved. The semi-automated microfluidic control improved the precision and synchronization among all reaction pipelines in the same chip, ensuring the high reproducibility of the assay. We found that the amount of the beads used in an individual assay was also a critical factor. Excess beads elevated the background signals while insufficient beads would decrease the enrichment efficiency.

The chromatin-antibody-bead complex was then flushed out of the chip and collected by microcentrifuge tubes. The complex was incubated with proteinase K at 68 ℃ for 2 h to release the DNA, which was then purified by phenol-chloroform-isoamyalcohol extraction. Precipitated DNA was re-suspended in 10 μl RNase-free H_2O, and analyzed by qPCR using the specific primers (Supplemental methods) and by high-throughput sequencing.

To quantitatively assess the sensitivity and accuracy of this new protocol, we performed ChIP-Seq of H3K4me3 for mEpiSCs and mESCs from both a bulk amount (10^6) of cells and from 1 000 cells and compared the results. We found that for mEpiSCs, the ChIP-seq of the two 1 000-cell samples recovered 16 351 and 16 245 out of the 16 929 enriched transcription start site (TSS) regions with H3K4me3 peaks from the bulk sample, and for mESCs, the two 1 000-cell samples recovered 18 256 and 18 206 out of the 18 367 enriched TSS regions from the bulk sample, demonstrating the high sensitivity (on average 96.3% for mEpiSCs and 99.3% for mESCs) of our method. Moreover, 98.3% and 98.4% of the 16 636 and 16 516 enriched TSS regions from the 1 000 mEpiSCs overlapped with those from bulk sample, and 91.0% and 91.7% of the 19 917 and 19 733 enriched TSS regions from the 1 000 mESCs overlapped with those from bulk sample (Figure 1C and Supplementary information, Table S1), exhibiting the high accuracy of our method. The mean correlation coefficient is 0.94 when comparing the H3K4me3 profile of 1 000 mEpiSCs and 10^6 mEpiSCs, and 0.76 of 1 000 mESCs and 10^6 mESCs (Supplementary information, Figure S1J), proving the general agreement of the profile based on 1 000 cells of our method and that of standard ChIP-Seq. We plot the receiver operating characteristic (ROC) curves using the *P*-value ranked peaks from the 1 000-cell experiments against the positive TSS from the bulk samples (Figure 1D and Supplementary information, Figure S1G), and then calculated the area under curve (AUC) of mEpiSC and mESC as 0.949 and 0.923, indicating our method a good classifier to distinguish positive TSSs from negative ones. These results clearly prove that our method is able to recover a majority of H3K4me3 peaks from as low as 1 000 mammalian cells with a very low false positive rate.

Next we tested the robustness of our 1 000-cell ChIP-Seq method. Besides mEpiSCs and mESCs, we also performed ChIP-Seq on two biological replicates of the H3K4me3 marker from the 1 000 epiblast cells of E6.5 mouse embryos. We found that most of the TSS regions that have enrichment peaks are overlapped between these duplicates (Figure 1E). The correlation coefficients (r) of these replicates are 0.884 (epiblast cells of E6.5 mouse embryos), 0.971 (mEpiSCs), and 0.973 (mESCs), which are comparable to the previous report that required 10 000 cells [7].

Since pluripotent mEpiSC cells were derived from epiblast cells of E6.5 mouse embryos, we asked whether the H3K4me3 landscape of mEpiSCs *in vitro* was similar to that of E6.5 epiblast *in vivo*. We found that, in general, the H3K4me3 pattern of mEpiSCs was very similar to that of epiblast from E6.5 embryos. Among the top 10 000 *P*-value ranked peaks around the TSS regions, 7 000 in mEpiSCs overlapped with those in E6.5 epiblast (Figure 1C). When we compared the mESCs and the E6.5 epiblast cells, the overlapped peak number was 6 680. These proved that the H3K4me3 epigenetic landscape of epiblast cells of E6.5 mouse embryos is more similar to that of mEpiSCs than to that of mESCs as expected.

Since H3K4me3 was an active marker for gene expression, we analyzed the transcriptome of mEpiSCs and E6.5 epiblast cells by RNA-Seq. We found that the gene expression profiles of mEpiSCs and E6.5 epiblast cells were very similar as well ($r = 0.940$) for transcripts with FPKM ≥ 0.1 in at least one of the samples. We compared, in-depth, the ChIP-Seq result at some important gene loci for early embryonic development and found high similarity among epiblast cells of E6.5 mouse embryos, mEpiSCs, and mESCs (Figure 1F). At the same time, we also found the specific gene loci only enriched in mESCs (Figure 1G). Gene ontology terms showed that, compared with mESCs, both epiblast cells of E6.5 mouse embryos and mEpiSCs enriched for the ectodermal differentiation-related characteristics such as neural tube development and neuronal differentiation (Figure 1H). Furthermore, the RNA expression level of the genes clearly correlated with the enrichment of H3K4me3 around their TSS regions both in mEpiSCs and E6.5 epiblast cells (Figure 1I and Supplementary information, Figure S1M). This correlation verifies the previous assumption that EpiSCs are a reliable *in vitro* model for post-implantation epiblast cells [13].

In summary, we have developed a highly sensitive ChIP-Seq method by combining microfluidic chip-based chromatin immunoprecipitation with one-tube carrier sequencing library preparation. The integrated device is able to finish the major steps of ChIP, including concentration of the cells from tens of microliters to nanoliters, fixation and permeabilization of the cells, fragmentation of chromatins, binding of the target chromatin fragments onto the beads, as well as the elution of enriched chromatin fragments. Subsequently, without any preamplification, the purified DNA fragments were converted into a sequencing library by a one-tube reaction containing end-repair, adenylation, and ligation followed by carrier PCR. We have demonstrated that this microfluidic-assisted ChIP-Seq method works robustly for as low as 1 000 mammalian cells. We have shown that the quality of H3K4me3 profile acquired by our method from 1 000 cells is comparable to that of traditional approach using bulk materials. Moreover, our method is highly reproducible with the correlation coefficient of the two biological replicates of E6.5 epiblast cells, mEpiSCs and mESCs as high as 0.884, 0.971, and 0.973, respectively. Finally we have demonstrated that the H3K4me3 epigenetic landscape of mEpiSCs is very similar to that of epiblast cells from E6.5 mouse embryos, validating that mEpiSC is an appropriate *in vitro* model to study the epigenetic regulation of post-implantation epiblast cells *in vivo*. Our method will permit thorough analysis of the epigenomic landscape of early embryos or other situations in which only a very limited amount of materials are available.

Acknowledgments

We thank Prof Azim Surani for kindly giving us the mEpiSC cell line, and thank Tao Chen, Dr Yongfan Men, Zhilong Yu, Zitian Chen, Dr Liang Zhao, Haiwei Qiu, Lu Yang and Dr Yun Zhang for discussion and experimental help. This work was supported by the Ministry of Science and Technology of China (2011CB809106, 2012CB966704) and the National Natural Science Foundation of China (21222501, 21327808, 31271543, 91313302). AMS was supported by the Whitaker International Biomedical Engineering Fellowship.

Jie Shen[1, 2, 3, *], Dongqing Jiang[1, 3, *], Yusi Fu[1, 3, *], Xinglong Wu[1, 3, *], Hongshan Guo[1, 3], Binxiao Feng[1], Yuhong Pang[1, 3], Aaron M Streets[1, 2], Fuchou Tang[1, 3], Yanyi Huang[1, 2]

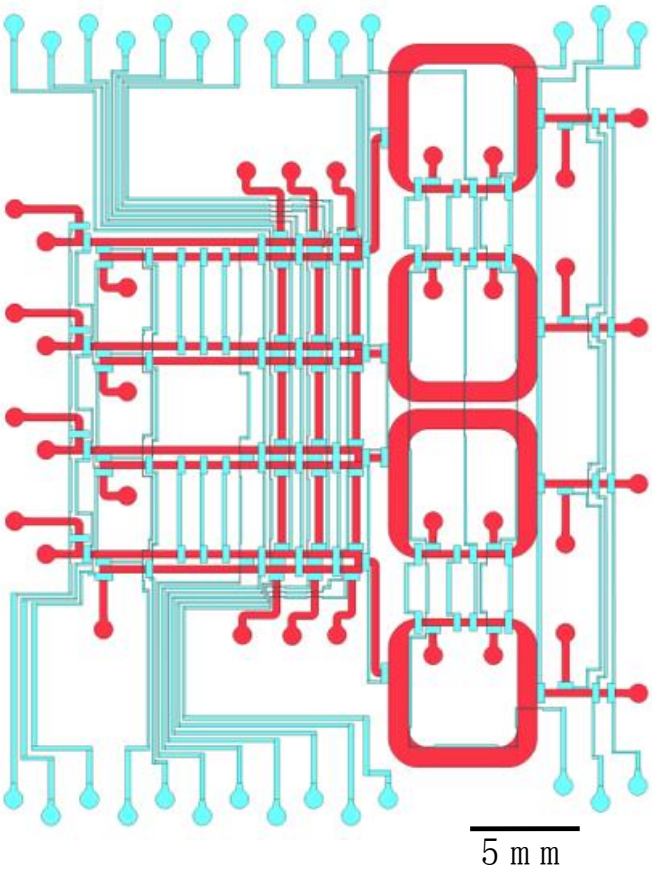
[1]*Biodynamic Optical Imaging Center (BIOPIC),* [2]*College of Engineering,* [3]*School of Life Sciences, Peking University, Beijing 100871, China*
*These four authors contributed equally to this work.
Correspondence: Yanyi Huang[a], Fuchou Tang[b]
[a]E-mail: yanyi@pku.edu.cn
[b]E-mail: tangfuchou@pku.edu.cn

References

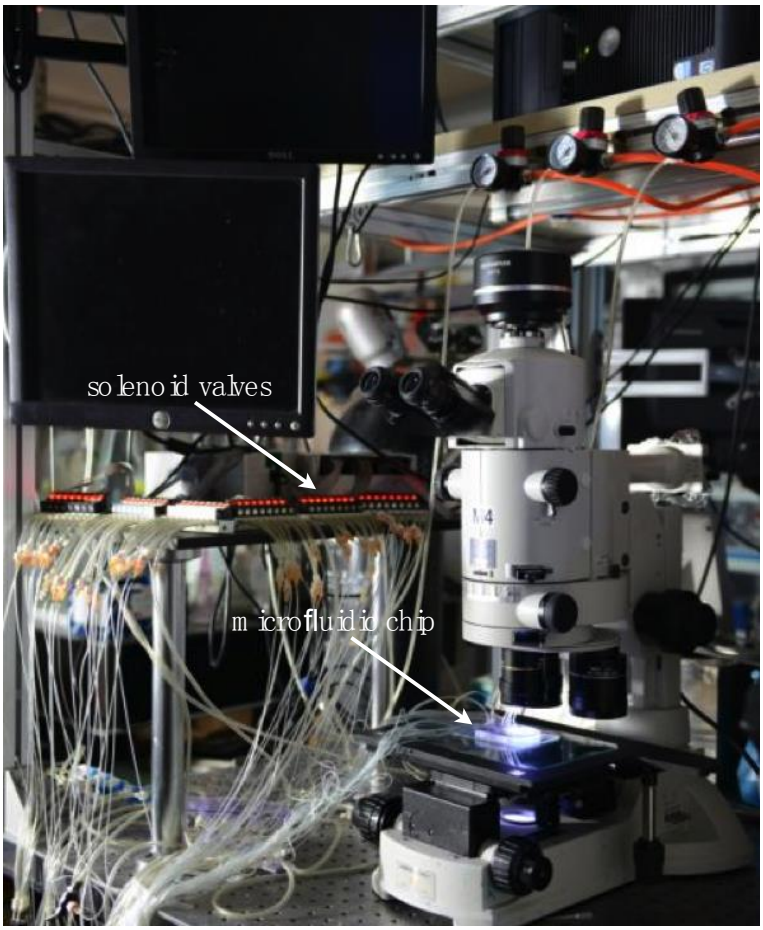
1 Berger SL. *Nature* 2007; **447**:407-412.
2 Shen Y, Yue F, McCleary DF, *et al*. *Nature* 2012; **488**:116-120.
3 O'Geen H, Nicolet CM, Blahnik K, *et al*. *Biotechniques* 2006; **41**:577-580.
4 Park PJ. *Nat Rev Genet* 2009; **10**:669-680.
5 Hawkins RD, Hon GC, Ren B. *Nat Rev Genet* 2010; **11**:476-486.
6 Furey TS. *Nat Rev Genet* 2012; **13**:840-852.
7 Adli M, Zhu J, Bernstein BE. *Nat Methods* 2010; **7**:615-618.
8 Wu AR, Hiatt JB, Lu R, *et al*. *Lab Chip* 2009; **9**:1365-1370.
9 Shankaranarayanan P, Mendoza-Parra M-A, Walia M, *et al. Nat Methods* 2011; **8**:565-567.
10 Gilfillan GD, Hughes T, Sheng Y, *et al. BMC Genomics* 2012; **13**:645.
11 Shankaranarayanan P, Mendoza-Parra M-A, van Gool W, *et al. Nat Prot* 2012; **7**:328-339.
12 Marcy Y, Ouverney C, Bik EM, *et al. Proc Natl Acad Sci USA* 2007; **104**:11889-11894.
13 Tesar PJ, Chenoweth JG, Brook FA, *et al. Nature* 2007; **448**:196-199.

14 Sachs M, Onodera C, Blaschke K, *et al*. *Cell Rep* 2013; **3**:1777-1784.
15 Ng JH, Kumar V, Muratani M, *et al*. *Dev Cell* 2013; **24**:324-333.

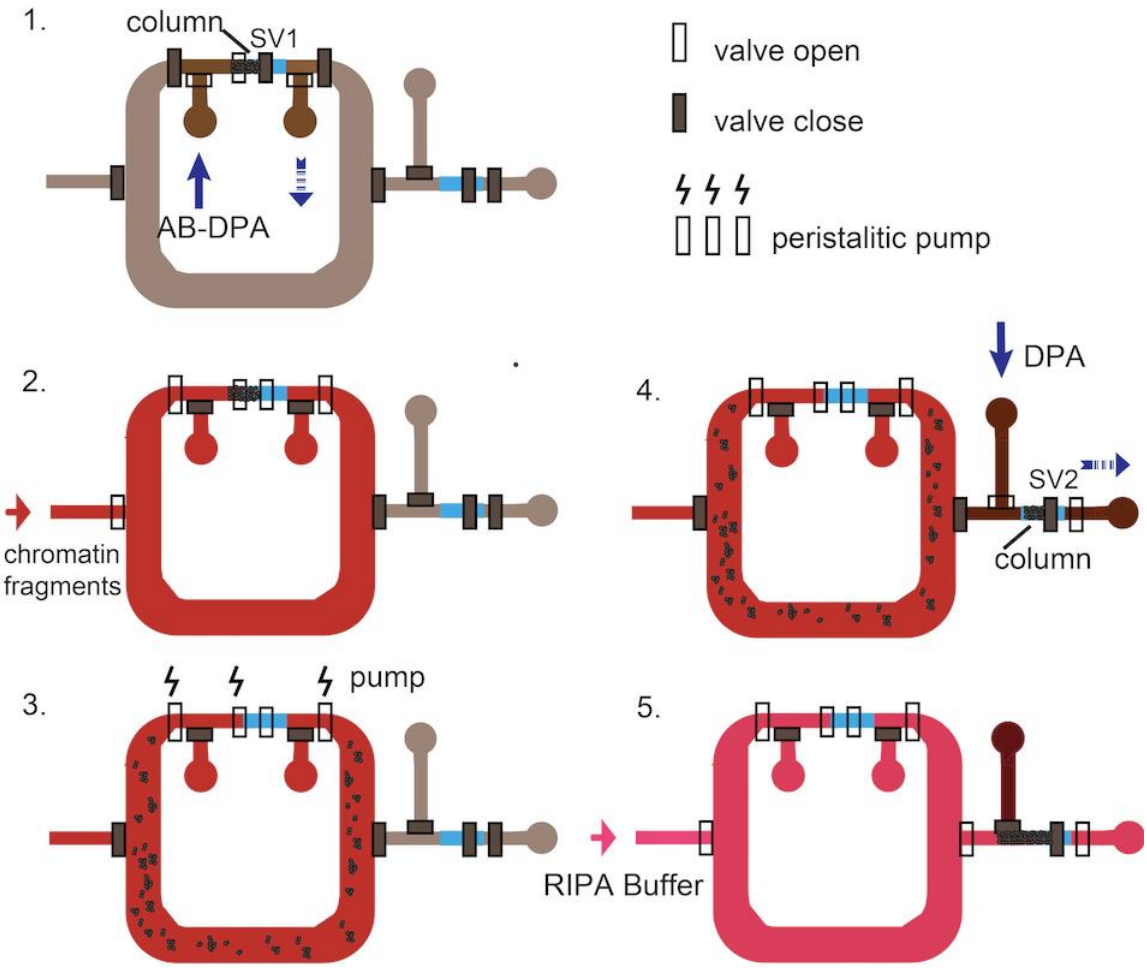
(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)



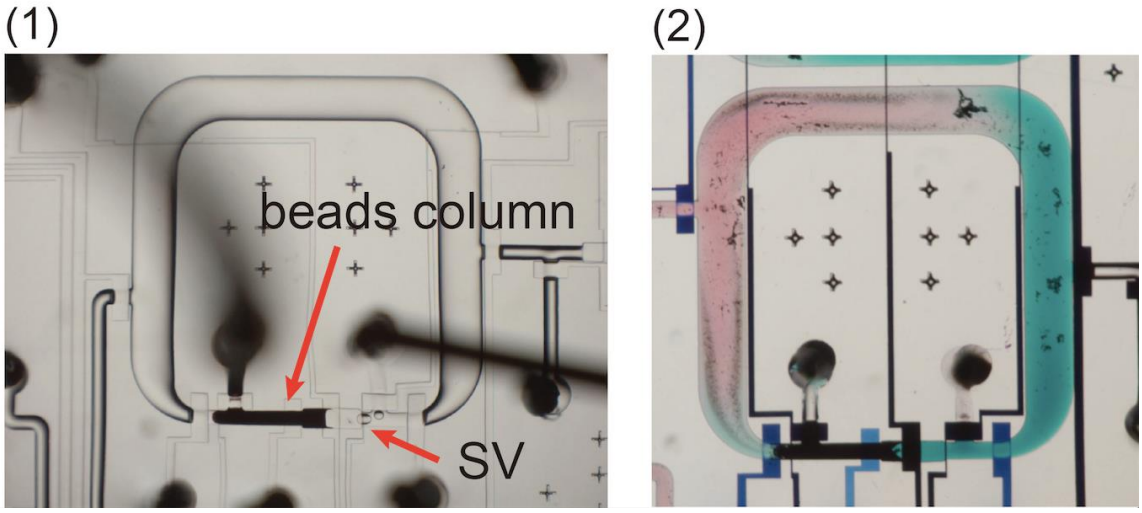
Supplementary information, Figure S1A The design of a microfluidic device with 4 identical independent reaction pipelines. The fluidic channels are shown as red in the figure, whereas the control channels cyan.



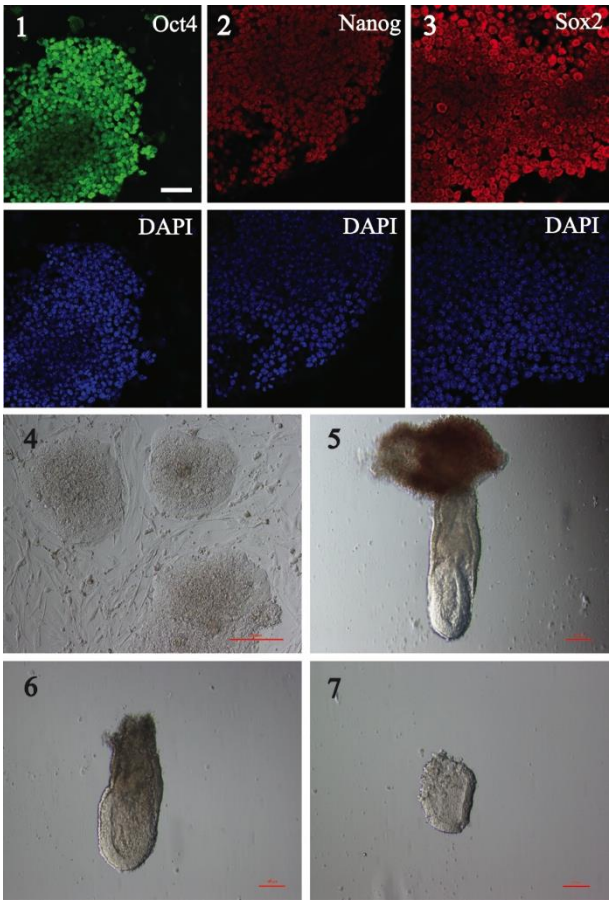
Supplementary information, Figure S1B The experimental setup. All integrated pneumatic valves in the chips were driven by a series of computer-controlled solenoid valves through home-developed LabVIEW programs. The air pressure for actuating integrated valves was 0.1 MPa, and the pressure for driving liquid sample was 0.01–0.02 MPa. The chip can be directly observed under a stereomicroscope (AZ100, Nikon).



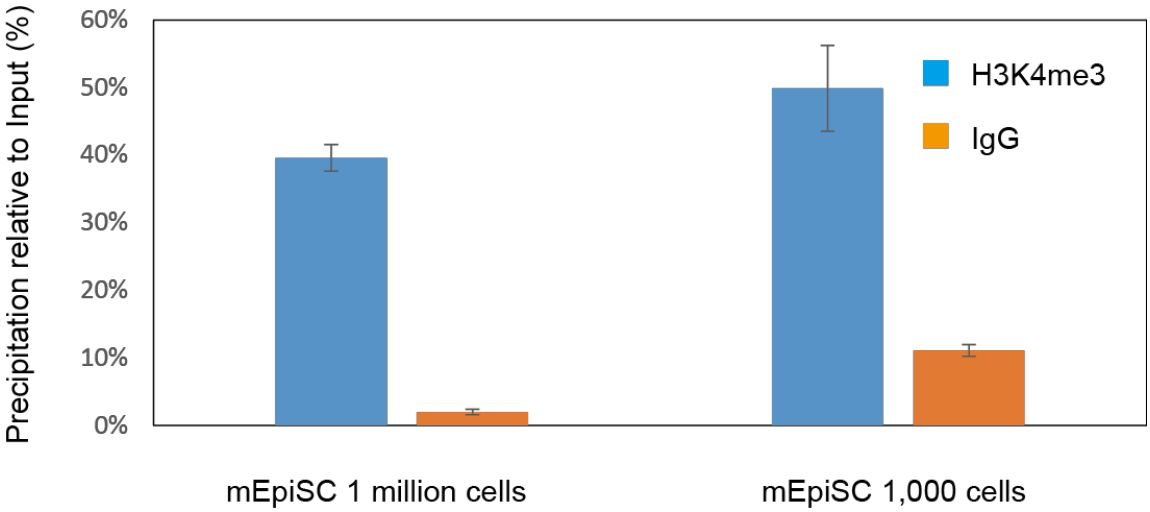
Supplementary information, Figure S1C Demonstrations of ChIP process on a microfluidic chip. (1) The antibody-functionalized beads (AB-DPA) are trapped by SV1 to form a column. (2) The chromatin fragments are loaded from the inlet to dead-end fill the flow-channels. (3) The mixing of chromatin fragments and Ab-beads is facilitated by peristaltic pump for 2 h at 0°C. (4) A short bead-column is formed at SV2. (5) The Ab-beads with captured chromatin fragments are trapped by the bead-column at SV2 and washed by RIPA buffer. Finally the beads are eluted off the device and collected by microcentrifuge tubes.



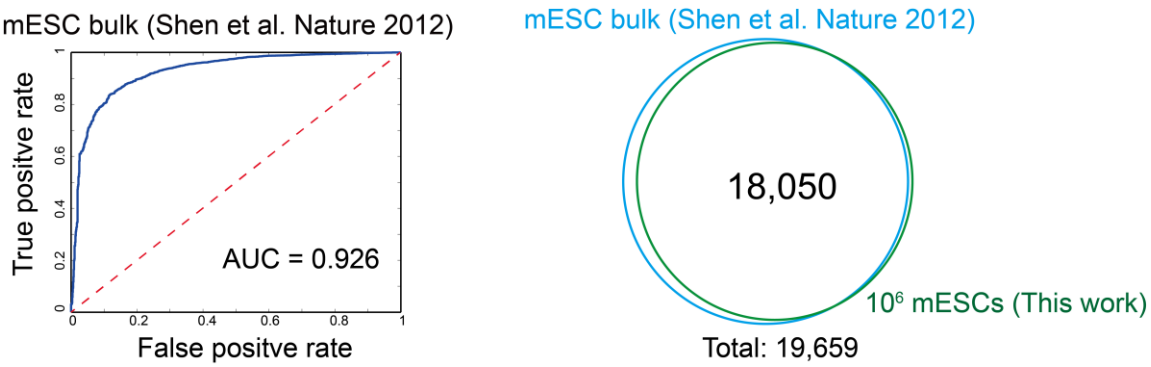
Supplementary information, Figure S1D The microphotographs demonstrating the operation on-chip. The controlled channels are filled with black dye. (1) The micrographs of the formation of bead-columns. We fabricated sieve valves (SVs) at the specific locations in the microfluidic devices to trap the beads and form columns. In a ring-shape channel, a column ~ 1,000 μm long is formed in the upstream side of the SV. (2) We can see two reagents, indicated with different dyes, and the beads (the black dots) are loaded in a loop-channel.



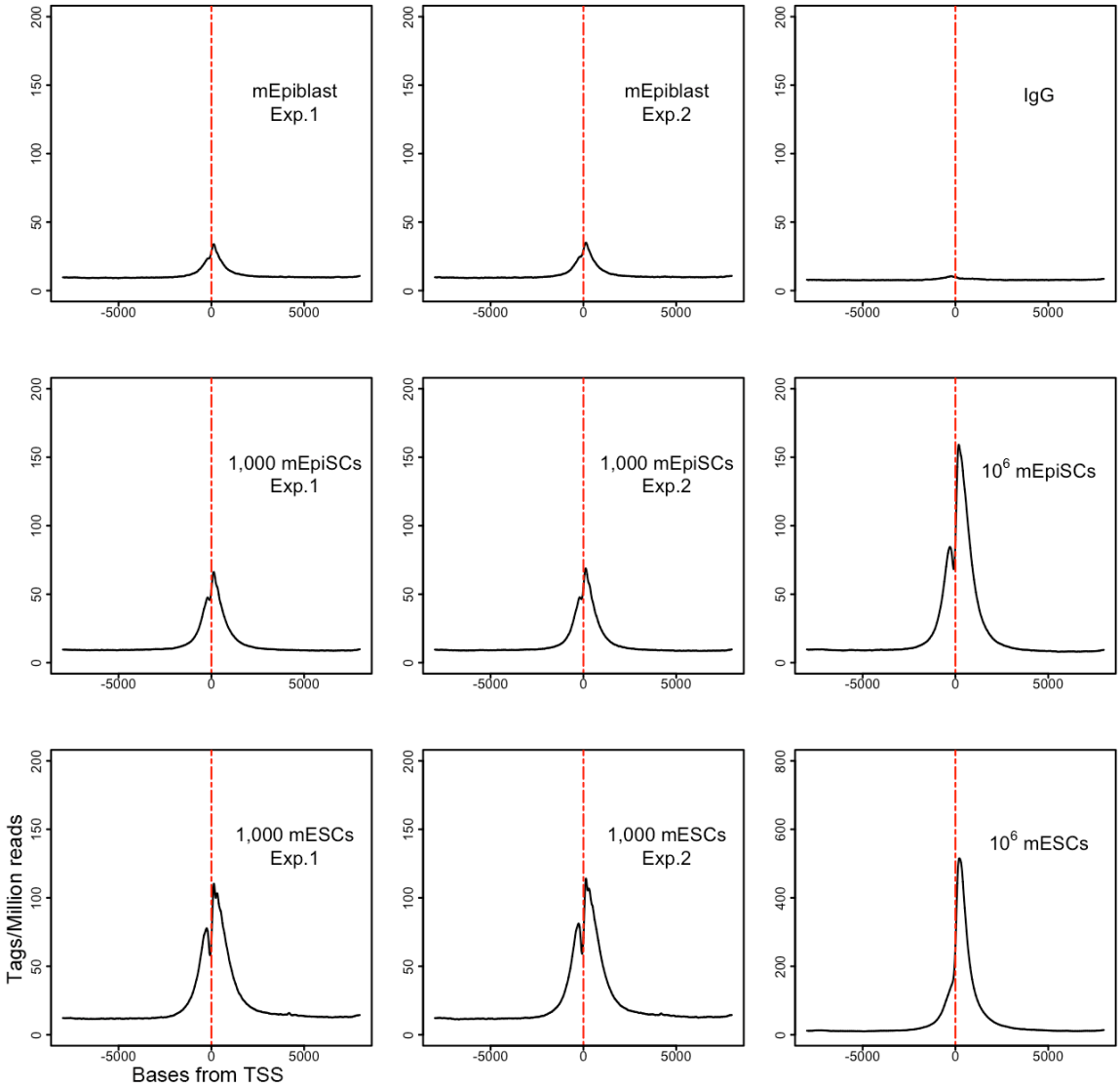
Supplementary information, Figure S1E (1)-(3) The Immunostaining of mouse epiblast stem cells (mEpiSCs) for pluripotency markers, Pou5f1 (also known as Oct4), Nanog, and Sox2, and DAPI (scale bar 50 µm).(4) The phase contrast microphotograph of mEpiSCs colonies (scale bar 200 µm). (5) A 6.5 dpc mouse embryo with ectoplacental cone (scale bar 100 µm). (6) The 6.5 dpc mouse embryo (scale bar 100 µm). (7) The epiblast part of the 6.5 dpc mouse embryo (the visceral endoderm was removed) (scale bar 100 µm).



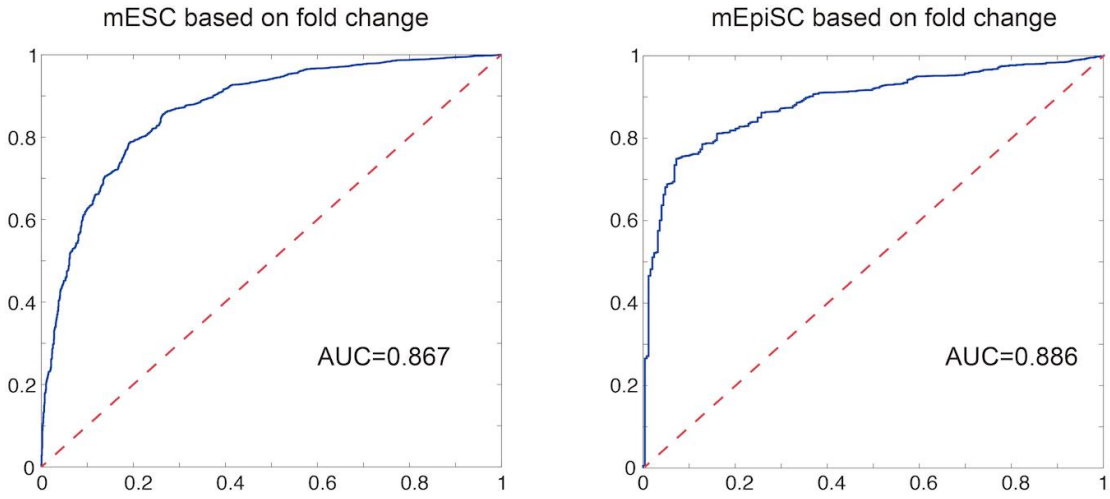
Supplementary information, Figure S1F ChIP enrichment of mEpiSCs for H3K4me3 with different amount of cells. Enrichment is calculated from the quantitative PCR Ct values. The primers used in testing K4 enrichment are 5'-CGCATGCAGGACCTGAACTT-3' (forward) and 5'-CGCTGCCGAGTA GGGTAGGATA-3' (reverse), which bind to H3K4me3 specific gene site Olig1.



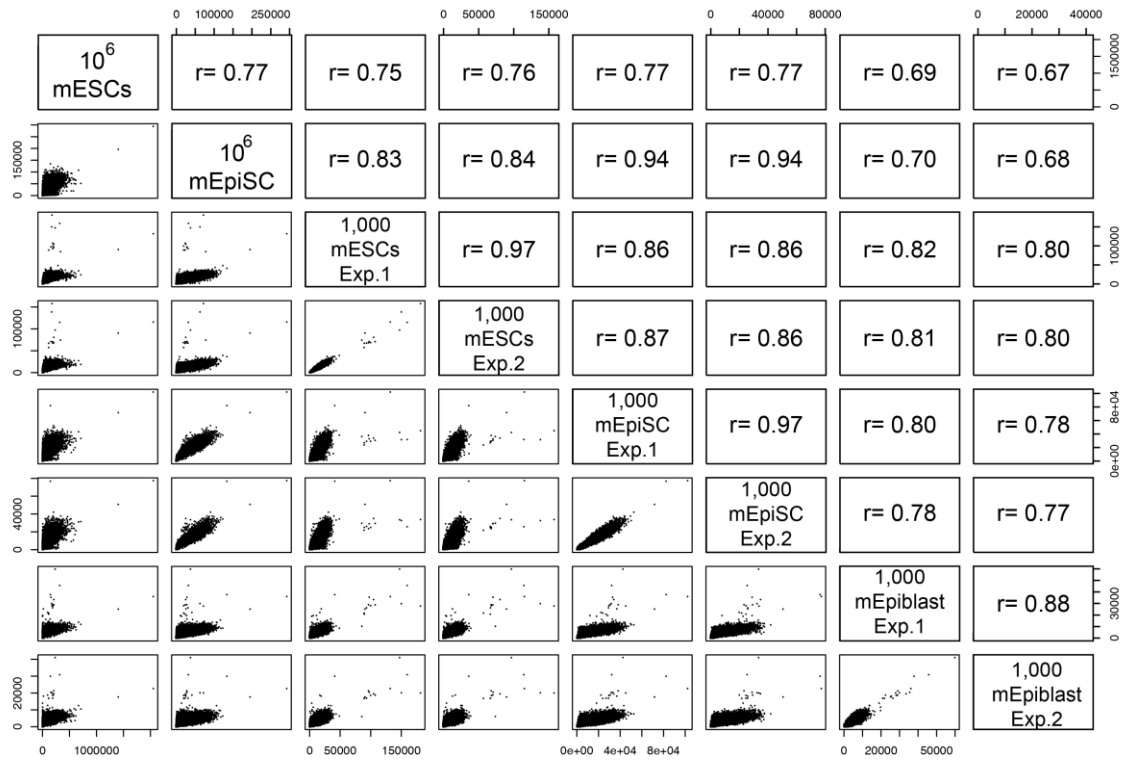
Supplementary information, Figure S1G Our mESC bulk data compared to Shen et al [10]. The similarities between our work and theirs are high when comparing with both ROC curve and Venn diagram.



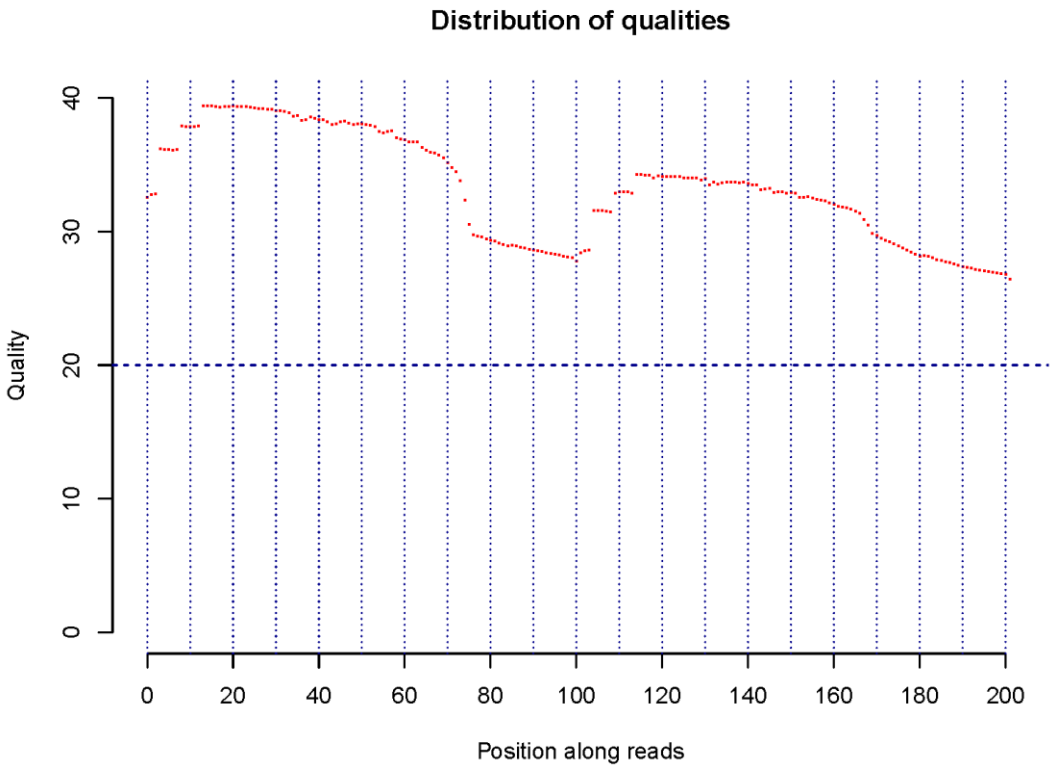
Supplementary information, Figure S1H Depth distribution of bases 2 kb upstream and 2 kb downstream of Transcription Start Site (TSS) with H3K4me3 marker. All the depths of bases of the same position relative to TSS are added up to view the distribution. The summit is slightly shifted downstream from TSS, similar to the previous findings (Mikkelsen, T. S. *et al. Nature* 2007; **448**, 553–560).



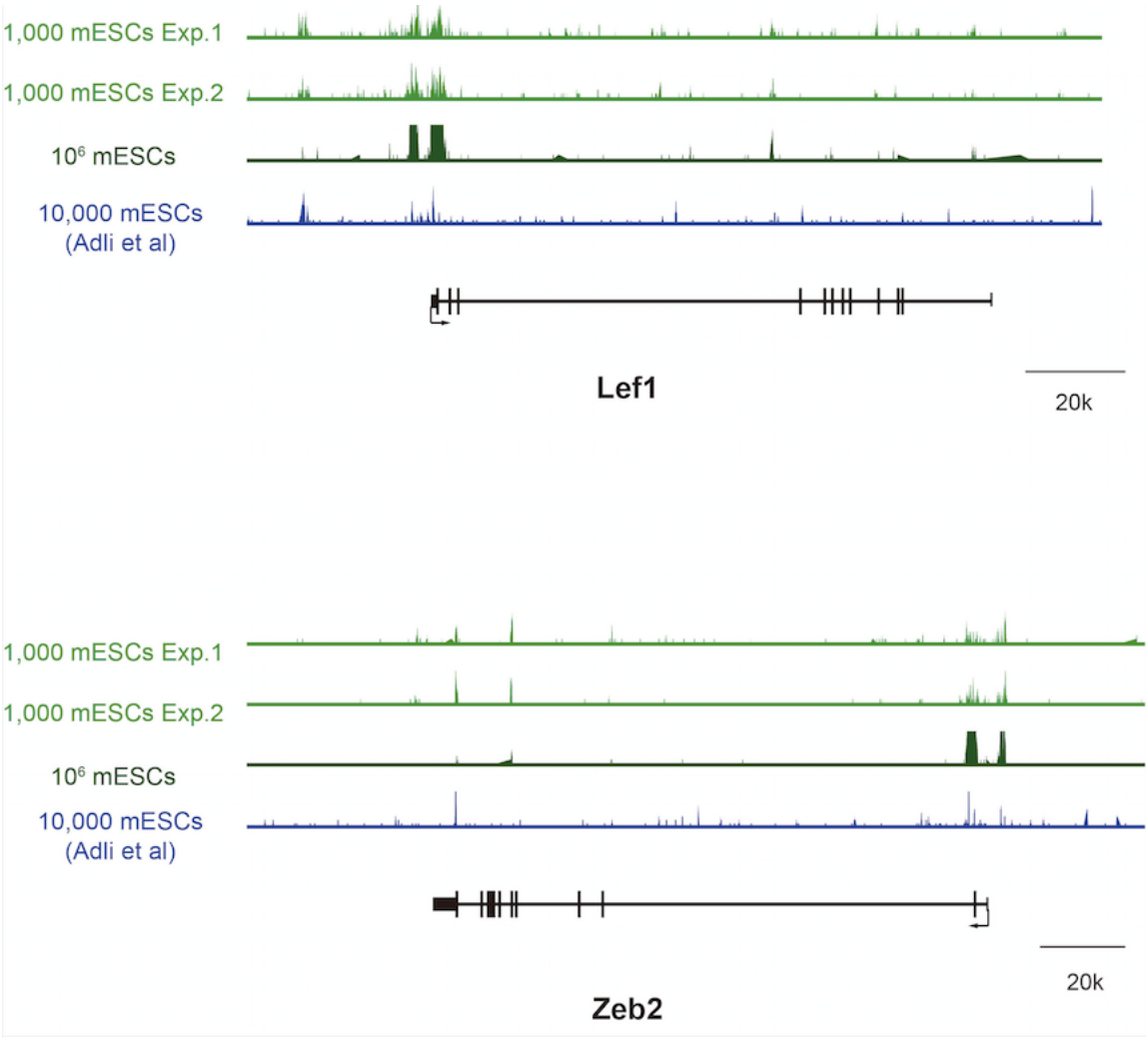
Supplementary information, Figure S1I The ROC curve based on fold change. The positive peaks called by MACS from the data of 1,000 mESCs and mEpiSCs are ranked by fold change, and are compared to the positive peaks called from the 1,000,000-cell sample.



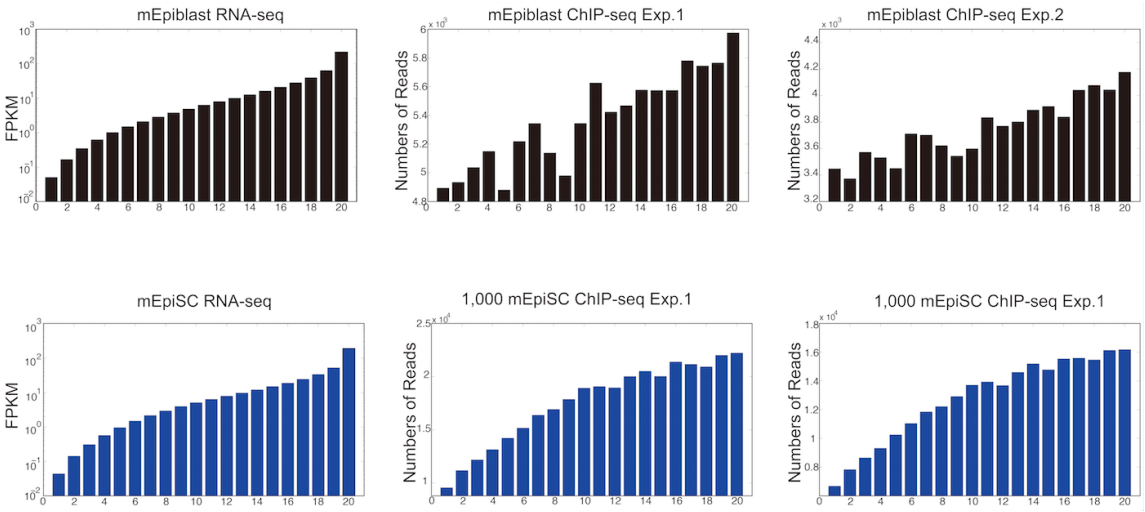
Supplementary information, Figure S1J The correlation map among different samples. The r represents the Pearson correlation coefficient. This figure shows the high similarity between bulk amount of cells samples and 1,000 cell samples as well as between biological replicates.



Supplementary information, Figure S1K Distributions of qualities of ChIP-Seq data. Quality is based on the Phred Quality Score, which $Q=-10\log_{10}P$. Shown here is the mean quality of different bases position of all the reads sequenced in 1,000 mEpiblast cells.



Supplementary information, Figure S1L Comparison of the ChIP-Seq of the H3K4me3 marker in mESCs in Adli's results [11] and ours. The Y axis for the figures is reads number covered on each base, the limit is 10 reads.



Supplementary information, Figure S1M The correlation between the expression level of genes and the enrichment of H3K4me4 marker around their TSS regions. Transcripts with FPKM>0.01 is sub-grouped into 20 groups and the mean expression level and the mean number of reads within H3K4m3 peaks at TSS regions of the genes of each group is calculated. The spearman correlation coefficient is 0.991 and 0.990 for mEpiSCs experiment; 0.917 and 0.944 for the mEpiblast experiment based on this subgrouping. It shows a clear statically correlation between the expression level with the H3K4me4 enrichment.

1. Methods and Material

1.1. Microfluidic device fabrication

All microfluidic devices were fabricated using multilayer soft lithography [1]. Devices are composed of three layers of polydimethylsiloxane (PDMS, RTV615 kit, GE), bonded to a cleaned glass slide (7x10 cm). Two separate master molds, one for the fluidic layer and the other for the control layer, were fabricated by photolithography. The silicon wafers were treated with hexamethyldisilazane (HMDS, Alfa Aesar, USA) vapor for 3 min at 25°C before being coated with photoresist. The hybrid master mold of the fluid layer was fabricated through a multi-step photolithography to form the molds with different thickness. The flow channels were fabricated by spin-coating positive photoresist (50XT, AZ Electronic Materials) to a thickness of 50 µm. After photolithographic etching, the patterned positive photoresist was re-flowed on a hot plate ramped from 35°C to 220°C at 6°C/h, to obtain rounded channel profiles, with a peak height of 45 to 50 µm. Subsequently, the chambers for magnetic bead trapping were patterned by 15µm-thick negative photoresist (SU-8 2025, MicroChem, Newton, MA). Finally the hybrid master mold was baked at 160°C for 1 h to fully crosslink the SU8. The mold of the control layer had 15 µm thick features made by P4620 positive photoresist (AZ Electronic Materials). Before the fabrication of PDMS chips, both molds were treated with trimethylchlorosilane (TMCS, Sinopharm, China) vapor for 5 min at 25°C. The fluidic layer was made by pouring PDMS (5: 1, elastomer to crosslinker ratio) onto its mold to a thickness of 5 to 7 mm. The control layer of the chip was made by spin-coating PDMS (20: 1, elastomer to crosslinker ratio) onto the mold at 1400 rpm for 60 s. Then the fluid and control layers were baked at 80°C for 20 min and 30 min, respectively. After being peeled off from its mold and hole-punched, the fluidic layer was aligned over the control layer, and then bonded at 80°C for 45 min. The bonded layers were peeled off from the control mold, hole-punched, then placed on a slide glass with a thin, cured PDMS layer (10: 1, elastomer to crosslinker ratio). Finally, the whole chip was incubated at 80°C for at least 6 h. The complete chip design is shown in Supplementary information Figure S1.

1.2. Cell culture

We acquired the mouse epi-stem cells (mEpiSCs) from Azim Surani's lab at Gordon Institute, University of Cambridge. This mEpiSC cell line was established from the epiblast of an E6.5 pregastrula-stage mouse embryo, with the detailed information described in publication [2].

mEpiSCs were grown to confluence in DMEM/F12 media (480 ml, GIBCO) supplemented with N-2 supplement (2.5 ml, 100x, Invitrogen), B-27 supplement (5 ml, 100x, Invitrogen), BSA (25 mg, Sigma), L-Glutamine (5 ml, 200 mM, Sigma), NEAA (5 ml, 100x, GIBCO), ß-ME (1 ml, 50 mM, Sigma), Penicillin/Streptomycin (5 ml, 10000 U/ml & 10 mg/ml, Sigma), Activin A (200 µl, final concentration: 20 ng/ml, Sigma) and bFGF (600 µl, final concentration: 12ng/ml, R&D systems) at 37°C and 5% CO_2 incubator conditions. Cells were passaged every two days.

Mouse Embryonic Stem cell (mESCs) were grown to confluence in DMEM/F12 media (480 ml, GIBCO) supplemented with FCS (120 ml, GIBCO), L-Glutamine (5 ml, 200 mM, Sigma), ß-ME (1 ml, 50 mM, Sigma), Penicillin/Streptomycin (5 ml, 10000 U/ml & 10 mg/ml, Sigma) Sodium Pyruvate (6 ml, Sigma), Nucleotide (6 ml, sigma), Nonessential amino acid (6 ml, gibco), Sodium bicarbonate (9.6 ml, sigma) and LIF (60 µl, final concentration: 1000 units/ml, Millipore) at 37°C and 5% CO_2 incubator conditions. Cells were passaged every two days.

1.3. Cell preparation

mEpiSCs were grown to 70% confluence, then washed in 1x PBS, digested by 0.1% Collagenase type IV (Gibco) and further diluted in culture medium to stop digestion. The mEpiSC clones were cut off by glass pipette and collected. The cell suspension was thoroughly pipetted to achieve a single cell suspension, further washed with 1x PBS and then pelleted at 1,000 rpm for 3 min. We then resuspended the cells with 500 µl 1x PBS, and crosslinked histone with DNAs by adding 13.5 µl formaldehyde (1% vol/vol final concentration), vortexing gently and keeping the reaction at room temperature for 8 min. Then we added 57 µl 1.25 mM Glycine and incubated the cell suspension at room temperature for 5 min to stop the crosslinking. The fixed cells were subsequently spun down at 6,000 rpm for 3 min and washed twice by 500 µl 1x DPBS. In the last washing process, we picked 2 µl of the suspension from the tube and counted the cell number, then centrifuged and aspirated the supernatant. According to the cell density, we resuspended the cells in an appropriate volume of DPBS to a final concentration of 200 cells/ µl and picked 5 µl suspension, which contained about 1,000 cells, for further on-chip processing.

mESCs were grown to 70% confluence, then washed in 1x PBS, digested by 0.5% trypsin (Invitrogen) and further diluted in culture medium to stop digestion. The cell suspension was thoroughly mixed to achieve a single cell suspension, further washed with 1x PBS and

then pelleted at 1,000 rpm for 3 min. We then resuspended cells and crosslinked histone with DNAs the same way as we did on mEpiSCs.

The mouse early embryonic cells were obtained from pregnant female mice on day 6.5 of pregnancy (E6.5). We first isolated embryos from the uterus of a pregnant mouse and subsequently dissected the embryo from maternal decidua. The Reichert's membrane and visceral endoderm were removed carefully by glass pipette carefully. We digested the cells with 0.25% trypsin for about 10 minutes, then added FBS (10% in PBS) for about 5 minutes to stop digestion and subsequently broke up the large clumps to a mostly single-cell suspension by mouth pipette apparatus. Finally, we used formaldehyde to perform the crosslinking of chromatins as well as we did on mEpiSCs.

1.4. Magnetic beads-antibody complex preparation

41 µl stock solution of Dynabeads-protein A (DPA, Invitrogen) was put into a 200 µl tube and then placed on a magnet rack for 3 min. We washed the DPA twice by removing the supernatant, adding 100 µl of RIPA Buffer (140 mM NaCl, 10 mM Tris-HCl (pH7.5), 1% Triton X-100, 0.1% SDS, 1 mM EDTA, 0.5 mM EGTA), vortexing, and placing the tube back to the magnet rack for 3 min. Supernatant was removed and 41 µl of RIPA Buffer was added to re-suspend DPA. 10 µl of resuspended DPA was diluted by 90 µl of RIPA Buffer in a new 200 µl tube on ice. We then added 4.8 µl H3K4me3 (Cell Signaling Technology) antibody, and 2.4 µl rabbit IgG (Santa Cruz Biotechnology) or GFP antibody into the tube. The tubes were stuck on a rotator running at 20 rpm at 4°C for 2 h.

1.5. Microfluidic ChIP process

Both ChIP and washing processes were performed in a microfluidic device (Supplementary information Figure S2). Before each experiment, the inside surface of the device was treated with a 0.2% pluronic solution (0.2% w/v Pluronic F127 (Sigma) in DPBS) for 30 min, followed by air drying. We performed four individual reactions in parallel with a single device. The cell suspension was loaded into the device through four independent inlets. We performed the chromatin fragmentation off chip by ultrasonic shearing in the probe sonicator Covaris S2, with fragment size of 150-800 bases. 10 µl Formaldehyde-cross-linked cells suspension was sonicated and collected. The fragmented chromatin suspension was concentrated from 10 µl by evaporation to 1 µl. Then the concentrated chromatin suspension was loaded into the device followed by beads-column trapping.

The antibody-functionalized beads (Ab-beads) pre-incubated with antibody slurry were loaded into the microfluidic device, and then trapped by the sieve valves SV1 to form a column about 1200 μm long. The chromatin fragments were loaded from the inlet in order to dead-end fill the flow-channels. With outlets keeping closed, the rings can be ''dead-end filled'' by RIPA ChIP Buffer (140 mM NaCl, 10 mM Tris-HCl (pH7.5), 1% Triton X-100, 0.1% SDS, 1 mM EDTA, 0.5 mM EGTA, 1:100 protease inhibitior (Roche)) with all the cellular material, reducing loss of soluble DNAs [3]. One of the rings, as designed for reference sample or input, was not loaded with beads. The mixing of chromatin fragments and Ab-beads is facilitated by peristaltic pump for 2 h at 0°C. After mixing, a short bead-column was formed at SV2. The Ab-beads with captured chromatin fragments were trapped by the bead-column at SV2 and washed by 50 μl RIPA Buffer for 10 min (Supplementary information Figure S3). Finally the beads were eluted off the device (Supplementary information Figure S3) and collected by centrifuge tubes. The microphotographs that demonstrate the operation on-chip were shown as Supplementary information Figure S4.

1.6. Isolation of DNA

After IP, the chromatin-antibody-bead complexes were collected through outlets by flushing 100 μl RIPA Buffer from inlet, respectively. The tubes were slightly centrifuged to avoid bubbles, and then placed on a magnet rack on ice for 3 min followed by removing supernatant. To reduce the nonspecific adhesion, we washed the chromatin-antibody-bead complex with 100 μl ice-cold RIPA Buffer twice. To wash the chromatin-antibody-bead complexes, we added 100 μl ice-cold RIPA Buffer followed by gently manual agitation and put them on magnet rack for 3 min to remove supernatant. The beads were then resuspended using 100 μl TE Buffer (10 mM Tris-HCl (pH8.0), 10 mM EDTA), transferred into a new 1.5ml tube on ice. The tubes were placed on a magnet holder on ice for 3 min for beads capture with supernatant being removed. DNA was then released by adding 300 μl Complete Elution Buffer (50 mM NaCl, 20 mM Tris-HCl (pH7.5), 5 mM EDTA, 1% SDS, 200 μg/ml protease K) to the captured beads and incubated on a thermomixer (Eppendorf) at 68°C, 1300 rpm for 2 h. The input DNA was dissolved in 100 μl RIPA Buffer and then treated by adding 200 μl Elution Buffer (50 mM NaCl, 20 mM Tris-HCl(pH7.5), 5 mM EDTA) and 3 μl 10 ng/nl proteinase K (Sigma) at 68°C for 2 h.

After incubation on a thermomixer, we placed samples on a magnet rack, and transferred supernatant to 1.5 ml tubes while keeping the input untouched. In the following steps, the treatment of samples was tantamount to that of input. We added 300 μl phenol-

chloroform-isoamyalcohol (25:24:1), vortexed and centrifuged the tubes at 15,000 g for 5 min, and collected 290 µl supernatant. We repeated the extraction step with chloroform-isoamyalcohol (24:1) and collected 270 µl upper liquid into a new tube. We then added 44 µl 3M NaAc (pH5.2) and 1 µl glycogen(10 ng/nl) into the supernatant with agitation, followed by adding 1 ml 100% alcohol at -20°C. We immediately mixed the content and incubated tubes at -80°C for 1 h. Tubes were taken from -80°C and melted at room temperature, and centrifuged at 13,200 rpm at 4°C for 15 min. After removing the supernatant, we added 1 ml 75% alcohol at -20°C to wash the sediment, and then centrifuged tubes at 13,200 rpm at 4°C for 10 min. The precipitated glycogen-DNA complex was dried at 50°C for 20 min. We finally dissolved the ChIP DNA with 10 µl H_2O. Dahl and Collas published an optimized protocol for ChIP [4]. Compared to that edition, we reduced the washing volume from tens of microlittles to less than one micolitter for further reducing the loss of chromatin fragments with H3K4me3 marker.

The solution was used for further analysis including quantitative PCR (qPCR) or deep sequencing.

1.7. Real-time qPCR

We validated this protocol by analyzing the ChIP samples by qPCR. We designed a positive primer Olig1 that yield short amplicons (66 bp). Which the forward sequence is 'cgcatgcaggacctgaactt', and the reverse sequence is 'cgctgccgagtagggtaggata'.

For real-time qPCR cycling, SYBR green (Takara) was used as a probe in a Rox reference dyell (Takara) according to following thermal cycling protocol: 40 cycles of 95°C for 5 s, 60°C for 34 s. Detection was enabled at 60°C. Post-run melting curve analysis proved that primer dimers were excluded.

1.8. High-throughput sequencing

DNA sequencing was carried out using an Illumina HiSeq2000/2500 sequencer. The reagents for library preparation were provided by Illumina and were used according to the manufacturer's protocol with appropriate modifications as described below. We purified DNA from the ChIP assay by AMPure XP beads and eluted the DNA by 20 µl resuspension Buffer. The DNA suspension was concentrated from 20 µl to 5 µl by evaporation. Then, 4 µl End Repair Mix and 1 µl resuspension Buffer were added to the 5 µl cleaned up DNA suspension before incubating at 30°C for 30 minutes and inactivating at 75°C for 15

minutes for end repairing. Adenylation of 3' ends is performed by adding 8.4 µl thawed A-Tailing Mix and 1.6 µl resuspension Buffer, incubating at 37°C for 30 minutes. We ligated adapters by adding 1.7 µl of DNA ligase Mix, 3.1 µl resuspension Buffer and 0.2 µl Adapter Index. 30°C for 10 minutes is the duration for adapter ligation. We added 1 µl DNA which contains 100 ng of 5 kb sized plasmid DNA as carrier, and then purified DNA twice by AMPure XP Beads. Enrichment of DNA was achieved by PCR for 15 cycles. Polyacrylamide gel electrophoresis followed another purification by AMPure XP Beads and we separated DNA with the length of 250-700 bp. We extracted DNA from PAGE gel and purified it by a QIAGEN gel extraction kit.

1.9. Analysis of the ChIP-Seq data

The raw images were converted to sequence data in fastq format using the RTA v1.9 and CASAVA v1.8.2, which generated 2-3 Gb data for each sample. We aimed to keep the high quality data and discarded the reads that 1) with too much uncertain sequencing base which called as N'; 2) with low quality (50% of reads with Phred quality value ≤5); and 3) contain adaptor sequences. The filtered ChIP-Seq data was then aligned to the reference genome (mm9, UCSC) using bowtie2 [5] with the default setting. The alignment SAM format results were converted to BAM format via samtools12, simultaneously PCR duplication alignment was removed.

The resulted SAM file was converted to coordinated-sorted BAM file with samtools [6], simultaneously the PCR duplication reads were removed. The filtered BAM file was then fed to MACS [7] to call peaks and generate a single wig file for each sample. We used 10^{-5} p-value cut-off for the 10^6 cells experiments and 10^{-3} for the 1,000 cells experiments as the peaks are less enriched. We kept all the reads while using MACS as the PCR duplicates have already been deleted and we used 180 bp as our inferred average length of the reads.

1.10. Analysis of the RNA-Seq data

The RNA-Seq data was sequenced using Illumina HiSeq 2000 to generate 100 bp pair-ended reads. Either read of a pair that failed to pass the data processing filtering above would be discarded. The filtered data was aligned to mm9 with Tophat [8] using the default setting, then the output BAM file was assembled and converted to FPKM of each transcript using Cufflinks with recommended options [9]. The gene structure annotation data was downloaded from UCSC Genome Bioinformatics. After converting the annotation data to GTF format and merging the redundancy transcripts, we got 28,056 transcripts with unique

annotated structure. The data was used to extract TSS positions and was regarded as an input of the reference transcript annotation for Cufflinks to calculate the FPKM of each transcript.

References:

1. Unger MA, Chou HP, Thorsen T, Scherer A, Quake SR. Monolithic microfabricated valves and pumps by multilayer soft lithography. *Science* 2000; **288**: 113-116.

2. Bao S, Tang F, Li X, *et al.* Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* 2009; **461**:1292-1295.

3. Marcy Y, Ouverney C, Bik EM, *et al.* Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *Proceedings of the National Academy of Sciences of the United States of America* 2007; **104**: 11889-11894.

4. Dahl JA, Collas P. A rapid micro chromatin immunoprecipitation assay (uChIP). *Nature Protocols* 2008; **3**: 1032-1045.

5. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature methods* 2012; **9**: 357-359.

6. Li H, Handsaker B, Wysoker A, *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078-2079.

7. Zhang Y, Liu T, Meyer CA, *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008; **9**: R137.

8. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 2009; **25**: 1105-1111.

9. Roberts A, Trapnell C, Donaghey J, *et al.* Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* 2011; **12**: R22.

10. Shen Y, Yue F, McCleary, DF, *et al*. A map of the cis-regulatory sequences in the mouse genome. *Nature* 2012; **488**: 116–120.

11. Adli M, Zhu J. Bernstein BE. Genome-wide chromatin maps derived from limited numbers of hematopoietic progenitors. *Nature Methods* 2010; **7**: 615–618.

Table S1. Summary of the sequencing data of ChIp-Seq

Sample Name	mEpibalst Exp.1	mEpibalst Exp.2	1,000 mEpiSC Exp.1	1,000 mEpiSC Exp.2	1,000 mESC Exp.1	1,000 mESC Exp.2	1,000,000 mEpiSC	1,000,000 mESC
Mapping(bowtie2)								
Reads(pair)	8757625	7138086	14151254	15394284	5080693	4427769	10097245	18139324
unique_mapped reads(pair)	6503200	4762866	10848961	11332733	4093318	3534133	7695187	14311799
multi_mapped reads (pair)	1241371	837539	2227463	2433190	674313	580873	2019169	1603059
discordantly reads(pair)	100419	71627	284437	235530	99633	73140	99627	83496
unique_mapped reads(mate)	134917	82324	124985	122642	52260	42996	87424	101432
multi_mapped reads(mate)	81728	56818	159162	144453	49087	38862	97055	58403
failed_mapped reads(mate)	1608625	2792966	1296639	2518567	325511	397388	382045	4122105
uniq mapping rate	74.26%	66.72%	76.66%	73.62%	80.57%	79.82%	76.21%	78.90%
overall mapping rate	90.82%	80.44%	95.42%	91.82%	96.80%	95.51%	98.11%	88.64%
PCRduplicat_rate	16.58%	22.03%	21.73%	44.68%	4.44%	5.43%	6.97%	29.97%
Peak Calling(MACS)								
peaks	25637	23344	18979	19366	66998	61285	16535	24786
positive TSS	12719	11668	16636	16516	19917	19733	16929	18367