

Coordination of Engineered Factors with TET1/2

Promotes Early-Stage Epigenetic Modification

during Somatic Cell Reprogramming

Gengzhen Zhu, Yujing Li, Fei Zhu, Tao Wang, Wensong Jin, Wei Mu, Wei Lin, Weiqi Tan, Wenqi Li, R. Craig Street, Siying Peng, Jian Zhang, Yue Feng, Stephen T. Warren, Qinmiao Sun, Peng Jin, and Dahua Chen

Inventory of Supplemental Information

Supplemental Figures:

Figure S1. Establishment of a rapid and highly efficient cell reprogramming system by modified factors. (Related to Figure 1)

Figure S2. OySyNyK-induced iPSCs exhibit complete pluripotency. (Related to Figure 2)

Figure S3. TET1/2-mediated 5hmC modification plays important roles in rapid iPSC induction by OySyNyK factors. (Related to Figure 3)

Figure S4. SOX2 and NANOG interacts with Tet1 and Tet2. (Related to Figure 4)

Supplemental Experimental Procedures: A detailed description of all the experimental procedures used in this work.

Supplemental References

FIGURE S1

A



B

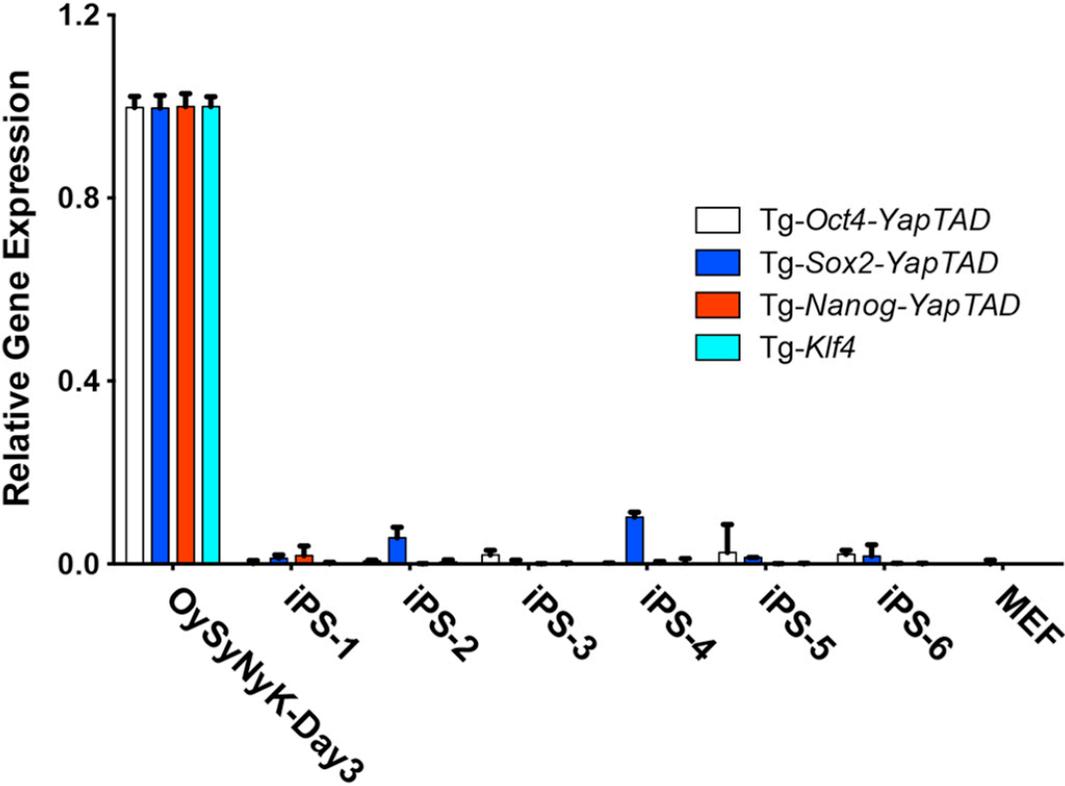


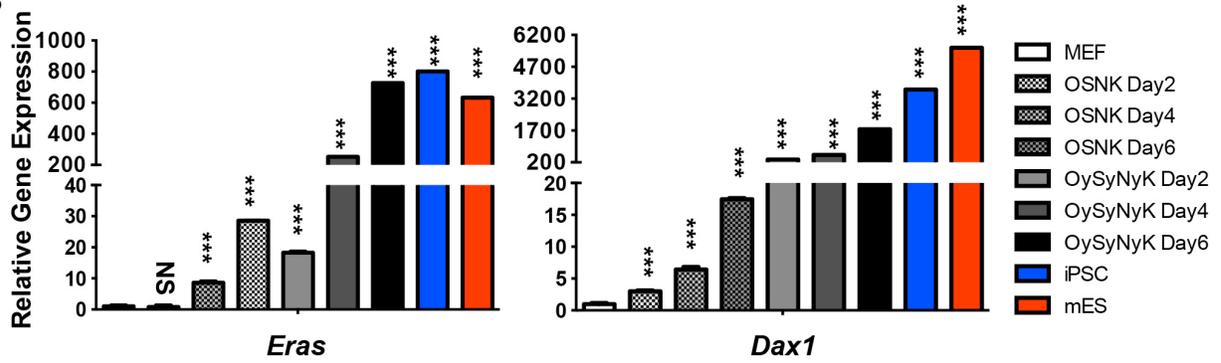
FIGURE S2

A

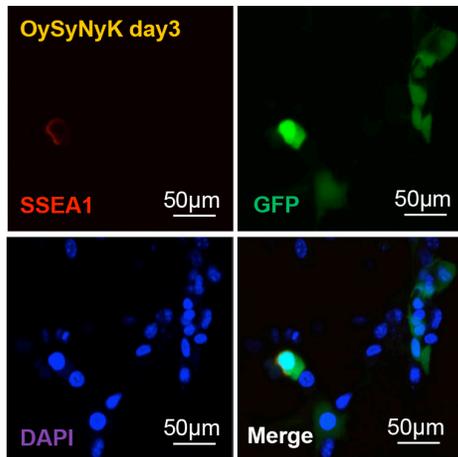
Summary of chimera formation and germline transmission from injections of OySyNyK-iPS cells into blastocysts.

Blastocysts injected	Mice born	Chimeras	Chimerism		Germline Transmission
			>50%	<50%	
77	29	2	1	1	1

B



C



D

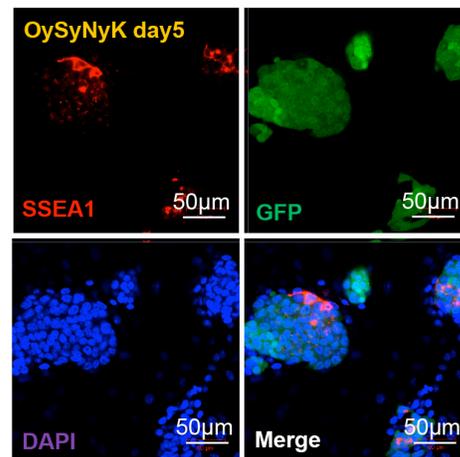


FIGURE S3

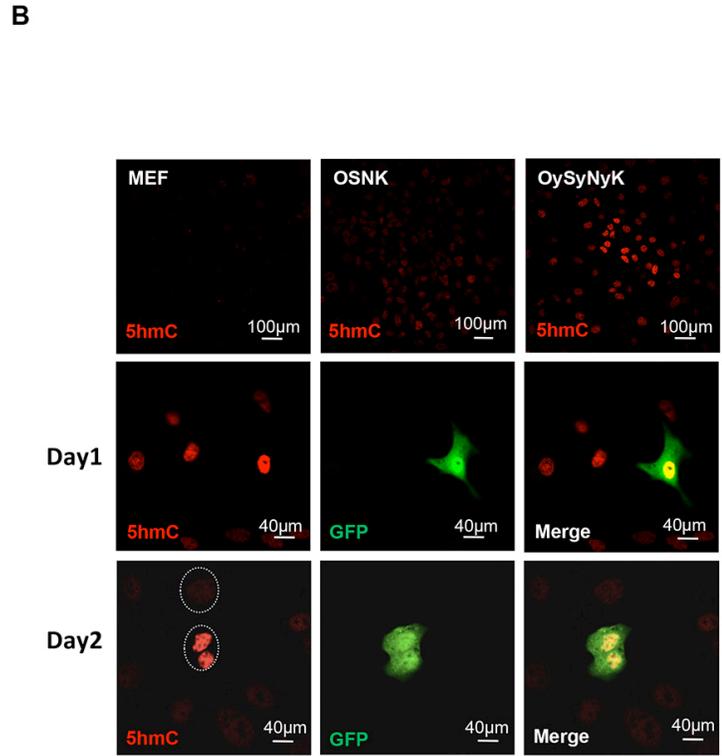
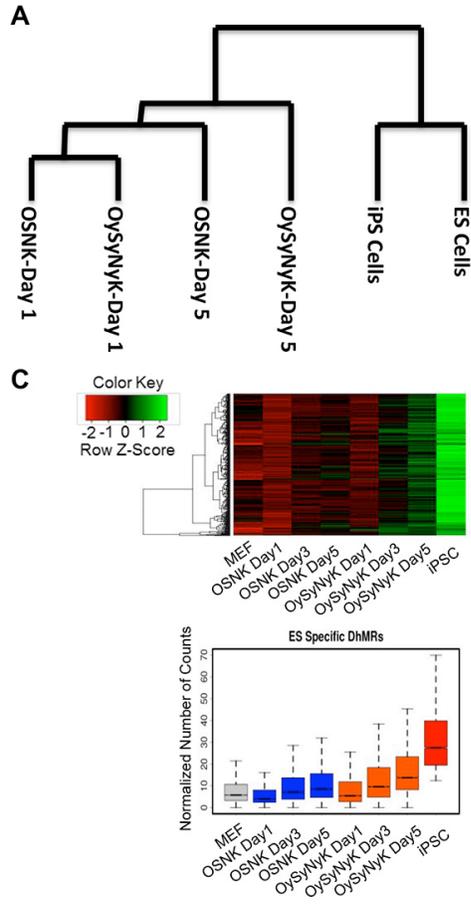
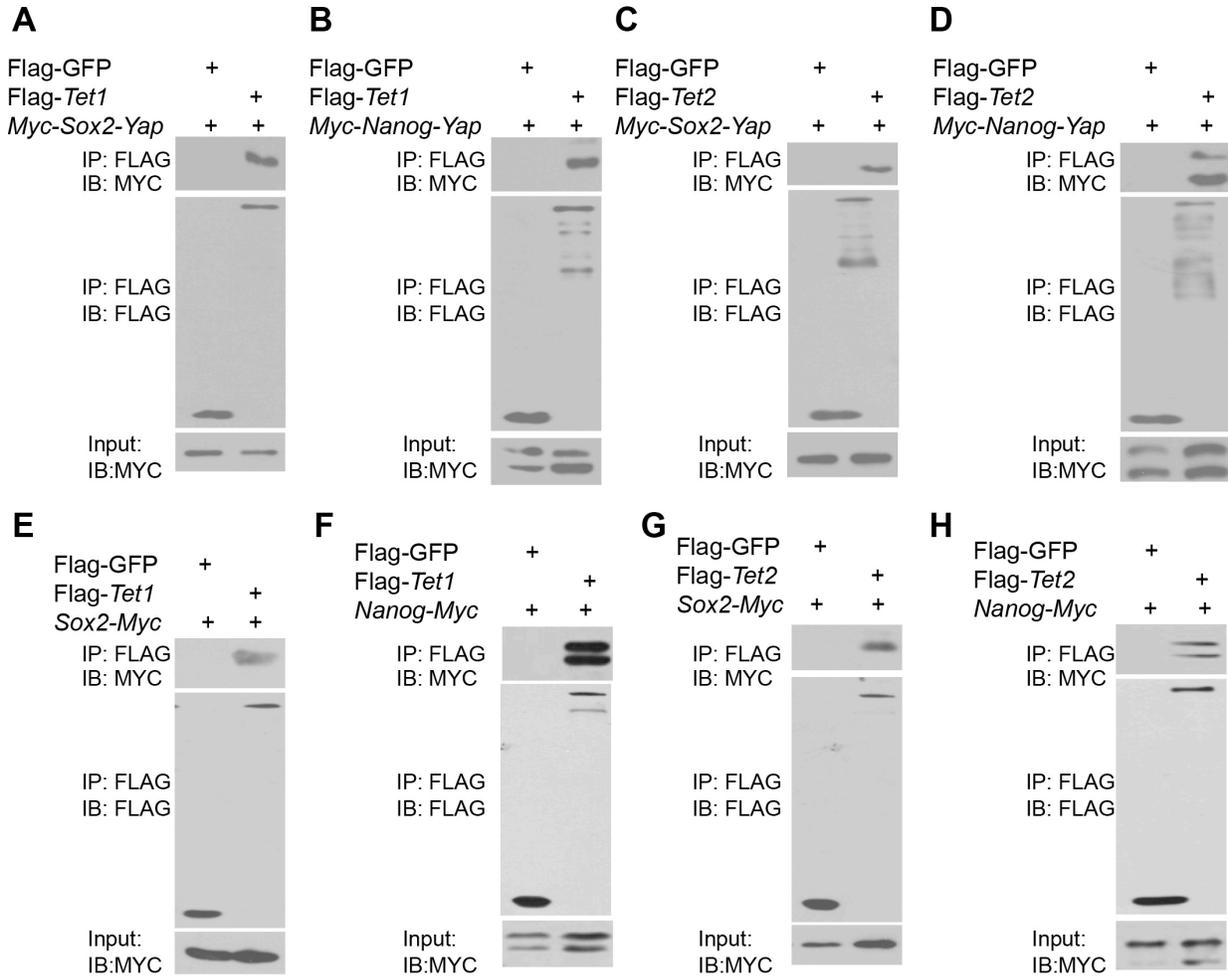


FIGURE S4



SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure 1. Establishment of a rapid and highly efficient cell reprogramming system by modified factors.

(A) Comparison of the efficiency of the OySyNyK- and the OSNK-induced iPSC cell colonies was shown by AP staining at the day 7.

(B) qRT-PCR analysis showing that the exogenous genes were silenced in multiple OySyNyK-induced iPSC lines, whereas the exogenous genes were highly expressed in OySyNyK-transduced MEFs at day 3. Data are representative of at least three independent experiments, (mean and s.d. of triplicate assays). OySyNyK-transduced MEFs at day 3 were used as control groups. Groups were compared using Student's t test. ***, $P < 0.001$ versus control groups.

Supplemental Figure 2. OySyNyK-induced iPSCs exhibit complete pluripotency.

(A) Summary of chimera formation and germline transmission from injections of OySyNyK-iPS cells into blastocysts.

(B) qRT-PCR analysis showing the expression levels of endogenous pluripotency markers *Eras* (left) and *Dax1* (right) in OSNK- and OySyNyK-transduced MEFs at different time points. Both iPSC and ES cells were included for the comparison. OSNK- and OySyNyK-transduced MEFs groups were compared to MEF groups respectively. ***, $P < 0.001$ versus control groups.

(C and D) Immunostaining assays showing expression of endogenous SSEA1 protein in in OSNK- and OySyNyK-transduced MEFs at day 3 (C) and day 5 (D).

Supplemental Figure 3. TET1/2-mediated 5hmC modification plays important roles in rapid iPSC induction by OySyNyK factors.

(A) Cluster of samples analyzed by gene expression profiling. Gene expression profile of OySyNyK-transduced MEF at day 5 is more similar to those from iPSC and ES cells.

(B) Immunostaining of 5hmC in MEFs, OSNK- and OySyNyK-transduced MEFs at day1 and day 2.

(3) Top: Heatmap of the top 500 ESC-specific DhMRs. Compared with the OSNK reprogramming method, the intermediate stages of cells reprogrammed by the OySyNyK method are more similar to iPSCs. Green represents more 5hmC modification and red represents less 5hmC modification. Bottom: Box plots of hydroxymethylation levels in ESC-specific DhMRs among MEFs, cells reprogrammed by the OSNK method at different stages, cells reprogrammed by the OySyNyK method, and iPSCs.

Supplemental Figure 4. SOX2 and NANOG interact with TET1 and TET2.

(A-D) HEK293T cells were transfected with various expression plasmids as indicated using the calcium phosphate precipitation method. Thirty-six hours post-transfection, cells were lysed for immunoprecipitation assays and further western blot assays. These assays showed physical interaction of the modified SOX2 and NANOG with TET1/2 proteins.

(E-H) HEK293T cells were transfected with various expression plasmids as indicated using the calcium phosphate precipitation method. Thirty-six hours post-transfection, cells were lysed for

immunoprecipitation assays and further western blot assays. These assays showed physical interaction of native SOX2 and NANOG with TET1/2 proteins.

SUPPLEMENTAL COMPLETE EXPERIMENTAL PROCEDURES

Plasmid construction

pMXs-*Oct4* (O), pMXs-*Sox2* (S), pMXs-*Nanog* (N), and pMXs-*Klf4* (K) retroviral vectors were obtained from Addgene. Coding sequences of mouse *Oct4*, *Sox2*, *Nanog* were fused directly with the TAD domain of mouse *Yap* (amino acids 275–489) in the C-terminal to generate *Oct4-Yap^{TAD}* (Oy), *Sox2-Yap^{TAD}* (Sy), and *Nanog-Yap^{TAD}* (Ny). The modified factors were then cloned into a pMXs-retroviral vector.

A pGL4.2-Basic-6XCR4 reporter plasmid was constructed as follows. One copy of CR4, which contains OCT4, SOX2 and KLF4 binding sites, was amplified from the mouse (C57BL/6) genome by PCR, and the 6XCR4 fragment was then subcloned into a pGL4.2-[*luc2*/Puro] vector (Promega) to generate the pGL4.2-Basic-6XCR4 vector. pRL-CMV vector driving the expression of Renilla luciferase under the CMV promoter was used as transfection control.

Cell cultures

Culture medium for HEK293T cells consisted of high-glucose Dulbecco's Modified Eagle's Medium (DMEM; Hyclone) supplemented with 10% heat-inactivated fetal bovine serum (Hyclone) and antibiotics (100 U/ml penicillin and 100 µg/ml streptomycin; Gibco). For MEF cells, the above medium was supplemented with 0.055 mM β-mercaptoethanol (Gibco), 2 mM L-GlutaMax (Gibco) and 0.1 mM non-essential amino acids (NEAA; Gibco), and for iPSC generation, 50 µg/ml vitamin C (Sigma) and 1000 U/ml LIF (Millipore) were further added.

iPSCs were maintained on mitomycin C-treated MEF feeder cells in knockout DMEM medium supplemented with 15% knockout serum replacement, 100 U/ml penicillin, 100 µg/ml streptomycin (Gibco), 1mM sodium pyruvate, 0.055 mM β-mercaptoethanol (Gibco), 2 mM L-GlutaMax (Gibco), 0.1 mM NEAA (Gibco), and 1000 U/ml LIF (Millipore). CGR8 mES cells were cultured in plates coated with 0.1% gelatin without feeder layers with GMEM medium (Sigma) supplemented with 15% fetal bovine serum (Millipore), 100 U/ml penicillin, 100 µg/ml streptomycin (Gibco), 1mM sodium pyruvate, 0.055 mM β-mercaptoethanol (Gibco), 2 mM L-GlutaMax (Gibco), 0.1 mM NEAA (Gibco), and 1000 U/ml LIF(Millipore)

Antibodies

The antibodies used in this work include anti-TET1 (09-872, Millipore), anti-TET2 (Ab94580, Abcam), anti-OCT4 (Ab19857, Abcam), anti-SOX2 (Ab5603, Abcam), anti-NANOG (A300-397A, Bethyl Laboratories), anti-GFP (M2004L, Abmart), anti-FLAG (F7425, Sigma), anti-MYC (MB2562, MBL) and anti-5hmC (39791, Active Motif) antibodies.

Luciferase reporter assay

HEK293T cells were transfected in a 24-well plate with various expression plasmids together with the pGL4.2-Basic-6XCR4 vector (100 ng/well) and renilla (2.5 ng/well) using the calcium phosphate precipitation method. At 36 h post-transfection, cells were lysed for measurement of luciferase activity. Firefly luciferase activities were normalized to the renilla activity.

Retroviral production and mouse iPSC induction

HEK293T cells were transfected with pMXs retroviral vector and packaging plasmid Ecopac (1:1) by the calcium phosphate precipitation method. At 12 h post-transfection, the medium was replaced, and at 48 h post-transfection, the virus supernatant was collected and filtered through a 0.45 μm polyvinyl difluoride (PVDF) filter. OCT4-GFP MEFs were seeded at 5×10^4 cells per well in a 12-well plate at 1 day before infection. The four different virus-containing supernatants (O/S/N/K, or Oy/Sy/Ny/K) were added at a ratio of 1:1:1:1 together with 8 $\mu\text{g}/\text{ml}$ polybrene. After incubation for 24 h, the virus supernatant was replaced with iPSC generation medium (this time point was referred to as day 0). The medium was replaced every day for iPSC generation. For OySyNyK-iPSC generation, at approximately 24 ± 6 h later, *Oct4-GFP* expression was observed and *Oct4-GFP*-positive iPSC colonies first appeared at day 3. iPSC colonies were then counted and picked up at days 6 or 7. For OSNK-iPSC generation, *Oct4-GFP*-positive iPSC colonies first appeared at days 9–10, which were counted and picked up at days 14–16.

Immunofluorescence and AP staining

Cells cultured on glass slides were fixed with 4% paraformaldehyde for 15 min, and then treated with 2M HCl for 1 h followed by 100mM Tris-HCl PH8.5 for 15min. The treated cells were further permeabilized and blocked with 5% bovine serum albumin (BSA) and 0.2% Triton X-100 for 30 min. The cells were then incubated with primary antibodies at 4°C overnight, followed by a secondary antibody at room temperature for 1 h. The antibodies used were: mouse anti-GFP (1:1000; Abmart) and rabbit anti-5hmC antibodies (1:500; Active Motif). Imaging of the cells

was performed using a Zeiss LSM 710 META laser scanning confocal system. For AP staining, cells were stained using an Alkaline Phosphatase Detection kit (SCR004; Millipore) according to the manufacturer's recommendations. OySyNyK-iPSCs were stained at days 7 or 8, whereas OSNK-iPSCs were stained at days 15 or 16.

Co-immunoprecipitation

For co-immunoprecipitation assays, cells were collected at 30 h after transfection and lysed for 50 min on ice in lysis buffer (20 mM Tris, pH 7.4, 150 mM NaCl, 0.5% Triton X-100, 1 mM EDTA, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride, 10 µg/ml aprotinin, 10 µg/ml leupeptin, 1 mM sodium fluoride, 1 mM sodium orthovanadate, and 25 mM β-glycerophosphate), sonicated briefly, and then centrifuged at 4°C at maximal speed for 10 min. The supernatants were then incubated with anti-Flag (Sigma) agarose beads for 4 h. Immuno-blotting assays were carried out by standard procedures. For endogenous IP, the lysates of CGR8 were incubated with anti-SOX2 or anti-NANOG antibody or control IgG for overnight at 4 °C, followed by further incubation with protein A/G bead (Pierce) for 2–4 h. The following immuno-blotting assays were carried out to detect the target proteins as indicated.

Genomic DNA dot blotting

Genomic DNA isolated from iPSC colonies generated using regular OSNK or SyKOyNy methods were serially diluted to ensure equal loading of samples, followed by denaturation at 99°C for 10 min, and then placed on ice immediately before loading onto a Hybond positively

charged nylon transfer membrane. The membrane loaded with genomic DNA was incubated at 80°C for 1 h for crosslinking, and then blocked with 5% BSA for 1 h at room temperature. Hybridization was carried out with a rabbit polyclonal antibody against 5hmC (1:10,000; Active Motif) overnight at 4°C.

Blastocyst injection

To produce chimeras, OySyNyK-iPSCs were injected into blastocysts from ICR mice, followed by implantation into pseudo-pregnant ICR mice. Chimeric mice were bred with wild-type ICR mice to determine the germline transmission ability of OySyNyK-iPSCs.

Bisulfite genomic sequencing

Genomic DNA (0.5 µg) was bisulfite-treated with a CpGenome™ Turbo Bisulfite Modification Kit (Millipore) according to the manufacturer's recommendations. Promoter sequences of OCT4 and NANOG were amplified by PCR, and then cloned into a pEASY-T3 vector (TransGen). Ten clones were selected randomly to be sequenced and analyzed.

qRT-PCR analysis

Total RNA was isolated using TRIZOL (Invitrogen), and 2 µg total RNA was used to synthesize cDNA using a TransScript First-Strand cDNA Synthesis SuperMix (TransGen). qRT-PCR was performed using SYBR Premix Ex Taq™ (Takara) and analyzed with a Bio-rad iQ5. The primers used in this study are listed in the Supplemental information.

Flow cytometric analysis and cell sorting

Cells were trypsinized, washed with PBS, and then filtered through a 300µm mesh. To calculate the efficiency of iPSC induction, the proportion of GFP-positive cells was analyzed by a BD FACSCalibur. GFP-positive cells were also sorted by a MoFlo XDP Cell Sorter.

***Tet1* and *Tet2* knockdown**

Knockdown of *Tet1* and *Tet2* was achieved by a pLKO.1-based short hairpin RNA lentiviral vector. A sequence targeting no specific gene was used as a control. HEK293T cells were transfected with lentiviral vectors together with packaging plasmids pMDLg, pMD.1G and REV by the calcium phosphate precipitation method. After incubation for 12 h, the medium was changed. The supernatant was collected and filtered through a 0.45 µm PVDF filter (Millipore) at 48 h post-transfection. A total of 4×10^5 *Oct4-GFP* MEFs were plated in a 6-cm dish at 1 d before viral infection. Virus supernatant was added to MEFs together with 8 µg/ml polybrene, followed by incubation for 24 h. MEFs were passaged twice before transduction with reprogramming factors.

Chromatin immunoprecipitation

The OySyNyK- or the OSNK-induced intermediate cells at day 1 were collected and used for ChIP assays. The cells were crosslinked with 1% formaldehyde for 10 min at room temperature, followed by neutralization with 125 mM glycine for 5 min. After washing with ice-cold PBS

twice, the crosslinked cells were collected for lysis on ice for 10 min with a SDS lysis buffer containing 1% SDS, 50 mM Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, and a protease inhibitor cocktail (Roche). The chromatin was then sonicated to ~500 bp fragments. Immunoprecipitations without beads were carried out using the sonicated chromatin and antibodies against Flag, TET1, TET2, NANOG and SOX2 at 4°C overnight, followed by a further 2 h of incubation with 30 µl Dynabeads protein G (Invitrogen) or 60µl Protein A Agarose/Salmon Sperm DNA (Millipore). After serial washing with buffers (low salt, high salt, LiCl, and TE buffers), the immunoprecipitated chromatin was eluted with SDS elution buffer containing 1% SDS and 0.1 M NaHCO₃, pH 8.0. The immunoprecipitated chromatin was reversed and digested with RNase and protease K, and the ChIP-DNA was extracted using phenol, chloroform and isoamyl alcohol at a ratio of 25:24:1 and precipitated with ethanol.

Capture of 5hmC-containing genomic DNA fragments

The 5hmC capture was performed as described previously (Song et al., 2011). Genomic DNA sonicated to fragments shorter than 500 bp was labeled with azide UDPG catalyzed by β-glucosyl transferase for 1 h at 37°C. The labeled DNA fragments were purified using a PCR purification kit (Qiagen), and then eluted in nuclease-free water. The labeled DNA fragments were biotinylated with a biotin-bisulfide linker at 37°C for 2 h. After purification with the PCR purification kit and elution in 100 µl nuclease-free water, the biotinylated DNA was incubated with Dynabeads MyOne Streptavidin C1 (Cat#: 650.02; Invitrogen) in binding buffer containing 10 mM Tris-HCl, pH 7.5, 1 M NaCl, 1 mM EDTA, and 0.5% Tween 20 at room temperature for

15 min. After thorough washing with binding buffer, the captured DNA was eluted with 50 mM fresh DTT at room temperature for 2 h. A Bio-spin 6 Tris column (Cat#: 732-6227; Bio-Rad) was used to remove the DTT from the elution buffer, followed by purification using a Qiagen MinElute PCR Purification Kit, and then the captured DNA was eluted in 10 μ l elution buffer.

Preparation of a captured DNA library

The captured DNA fragments were used for generation of libraries following the Illumina protocol for “Preparing Samples for ChIP Sequencing of DNA” (Part# 111257047 Rev. A). A total of 25 ng input genomic DNA, 5hmC-captured DNA or ChIP-DNA was used to initiate the protocol. DNA fragments (~150–300 bp) were selected using AMPure beads after the adaptor ligation step. PCR-amplified DNA libraries were quantified on an Agilent 2100 Bioanalyzer and diluted to 6–8 pM for cluster generation and sequencing. We performed 38-cycle single-end sequencing using Version 4 Cluster Generation and Sequencing Kits (Part #15002739 and #15005236, respectively) and the version 7.0 protocol. Image processing and sequence extraction were performed using the standard Illumina Pipeline.

5hmC profiling

Images acquired from the Illumina HiSeq2000 were processed using the default program by Illumina. The raw sequences were aligned using Bowtie 0.12.6 for the mouse genome UCSC mm9 as the reference genome with no more than two mismatches within the first 30 bp. A custom pipeline was used to remove duplicated genomic matches. ESC-specific DhMRs across

the genome were called using a model-based analysis of ChIP-Seq (MACS) with ESC 5hmC profiles as the treatment and the MEF 5hmC profile as the control. Parameters used for MACS peak calling were: effective genome size: 1.87e+09; tag size: 50; bandwidth: 200; P-value cutoff: 1.00e-05. The association of DhMRs with genomic features was acquired by overlapping with known genomic features. Genomic features such as UCSC RefSeq genes and CpG islands were obtained from UCSC tables for mm9. The overlap with a given genomic feature was called if overlapping by ≥ 1 bp.

All sequence related statistical analysis and data processing were analyzed using R (<http://www.r-project.org/>). Gene ontology analysis was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7.

List of Primers used in this study

qRT-PCR for mouse *Tet1*:

Forward primer: 5'-ACACAGTGGTGCTAATGCAG-3'

Reverse primer: 5'-AGCATGAACGGGAGAATCGG-3'

qRT-PCR for mouse *Tet2*:

Forward primer: 5'-AGAGAAGACAATCGAGAAGTCGG-3'

Reverse primer: 5'- CCTTCCGTACTCCCAAACATCAT-3'

qRT-PCR for mouse *Tet3*:

Forward primer: 5'- TGCGATTGTGTCGAACAAATAGT-3'

Reverse primer: 5'- TCCATACCGATCCTCCATGAG-3'

qRT-PCR for endogenous *Oct4*:

Forward primer: 5'-TAGGTGAGCCGTCTTTCCAC-3'

Reverse primer: 5'-GCTTAGCCAGGTTTCGAGGAT-3'

qRT-PCR for endogenous *Sox2*:

Forward primer: 5'-AGGGCTGGGAGAAAGAAGAG-3'

Reverse primer: 5'-CCGCGATTGTTGTGATTAGT-3'

qRT-PCR for endogenous *Nanog*:

Forward primer: 5'-ATCCCTTCCCTCGCCATCAC-3'

Reverse primer: 5'-GGCATTGATGAGGCGTTCC-3'

qRT-PCR for *Dax1*:

Forward primer: 5'- TGCTGCGGTCCAGGCCATCAAGAG -3'

Reverse primer: 5'- GGGCACTGTTCAGTTCAGCGGATC -3'

qRT-PCR for *Eras*:

Forward primer: 5'- TGCCTACAAAGTCTAGCATCTTG -3'

Reverse primer: 5'- CTTTTACCAACACCACTTGCAC -3'

qRT-PCR for exogenous *Oct4*:

Forward primer: 5'-GGGTGGACCATCCTCTAGAC-3'

Reverse primer: 5'-CCAGGTTTCGAGAATCCAC-3'

qRT-PCR for exogenous *Sox2*:

Forward primer: 5'-GGGTGGACCATCCTCTAGAC-3'

Reverse primer: 5'-GGGCTGTTCTTCTGGTTG-3'

qRT-PCR for exogenous *Nanog*:

Forward primer: 5'-GGGTGGACCATCCTCTAGAC-3'

Reverse primer: 5'-GGCATTGATGAGGCGTTCC-3'

qRT-PCR for exogenous *Klf4*:

Forward primer: 5'-GGGTGGACCATCCTCTAGAC-3'

Reverse primer: 5'-GCTGGACGCAGTGTCTTCTC-3'

qRT-PCR for *Gapdh*:

Forward primer: 5'-AGTCAAGGCCGAGAATGGGAAG-3'

Reverse primer: 5'-AAGCAGTTGGTGGTGCAGGATG-3'

qRT-PCR for ChIP assay *Tet1* promoter

Forward primer: 5'-TTGGTTTGGTTTGGTTGGTCCAGG -3'

Reverse primer: 5'-GGCAATGGCTGCTGCACTTTCTTA -3'

qRT-PCR for ChIP assay *Tet2* promoter

Forward primer: 5'-GGAGAAGAATGCAACTCCTCTTCAG -3'

Reverse primer: 5'-ACGGGAAGTTGGGATTCACAGAGA -3'

qRT-PCR for ChIP assay *Oct4* promoter

Forward primer: 5'-GGAAGTGGGTGTGGGGAGGTTGTA -3'

Reverse primer: 5'-AGCAGATTAAGGAAGGGCTAGGACGAGAG -3'

qRT-PCR for ChIP assay *Nanog* promoter

Forward primer: 5'-GCATAAACCTTGATATTTTGAACGGCCTATT -3'

Reverse primer: 5'- ACAGATGGACTAAAGCCCCTAAGTAGAAATCAT -3'

DNA Methylation analysis of the promoter of the *Nanog*:

Forward primer: 5'-GATTTTGTAGGTGGGATTAATTGTGAATTT-3'

Reverse primer: 5'-ACCAAAAAAACCACACTCATATCAATATA-3'

DNA Methylation analysis of the promoter of the *Oct4*:

Forward primer: 5'-ATGGGTTGAAATATTGGGTTTATTTA-3'

Reverse primer: 5'-CCACCCTCTAACCTTAACCTCTAAC-3'

ShRNA for Knockdown of *Tet1*:

5'-TTTCAACTCCGACGTAAATAT-3'

ShRNA for Knockdown of *Tet2*:

5'-ATGCAGTATTTCCCGAATAAT-3'

SUPPLEMENTAL REFERENCES

Song, C.X., Szulwach, K.E., Fu, Y., Dai, Q., Yi, C., Li, X., Li, Y., Chen, C.H., Zhang, W., Jian, X., *et al.* (2011). Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol* 29, 68-72.