Supporting Information

Hayashi et al. 10.1073/pnas.1502855112

SI Discussion

Based on our results (summarized in Table S2), the role of each residue used in this study is classified as follows.

K137, T141, N145, and R147 on NANOG are critical for DNA binding and mESC self-renewal. K137 and R147 are also important in maintaining protein lifetime in mESCs. These critical residues are located on the reading helix H3, and form specific contacts with DNA (Fig. 1*D*). Although N145 and R147 are highly conserved in various HDs and support the general DNA recognition of HD to the canonical consensus TAAT sequence (1), K137 and T141 are highly specific to mammalian NANOG HD and could support the recognition of specific DNA by NANOG HD (2). Together, these residues may have a fundamental role in DNA recognition of NANOG HD.

T100, Y119, and Q138 could not be analyzed as purified protein:DNA complexes, but were shown to be important for mESC self-renewal. Y119 was also important in maintaining protein lifetime in mESCs. These residues are highly conserved among various HD. T100 is situated in the N-terminal arm of the HD, which has nonbonded contacts to the target DNA in the minor groove (3) (Fig. S1*D*). Y119 is situated near the start of H2 helix and contacts phosphates in the DNA backbone (Fig. S1*C*). This interaction is conserved in various HD structures (3), and is observed in the hNANOG HD complex. Q138 supports H3 helix interaction with the major groove of DNA through phosphate contacts (3). Together, these residues may be critical in maintaining the integrity of NANOG HD–DNA complex and preventing degradation.

F102 is important in maintaining protein lifetime in mESCs and mESC self-renewal, although substitution of this residue for an alanine did not significantly affect DNA-binding affinity. This residue is conserved in other HD proteins. The aromatic F102 makes nonbonded backbone contacts with phosphates in the DNA in the hNANOG HD-DNA complex structure, and also contributes to the hydrophobic core of the protein that involves packing of the three helices, consistent with that described for other HDs (4, 5) (Fig. S1*D*).

Q124, M125, Y136, K140, and K151 alanine substitutions weaken hNANOG DNA-binding activity. In addition, Y136, K140, and K151 were weaker in maintaining protein lifetime in mESCs. Residues K140 and K151 are highly conserved among various HD proteins and form hydrogen bonds with phosphates in the DNA in hNANOG HD–DNA complex. In addition, K140 was mediates the interaction with a specific DNA in PDX1 HD (6). All of the other residues described (Q124, M125, and Y136) are not conserved in the HD family, and thus are specific to NANOG (2) (Fig. S1*C*). Interestingly, M125 and Y136 constitute one of the most frequently occurring covarying pairs among various HDs and contact the same phosphate (7). Thus, there might be functional redundancy in these residues, which would not be evident in single mutation to alanine.

Alanine mutants of Q144 and M148 showed no or little effects in our experiments. These results were unexpected because these residues are conserved among many HDs (1) and were shown to be important in DNA recognition and specificity in *Drosophila* Ftz and Antp proteins (3). However, our results are consistent with previous reports, where alanine substitution of these residues maintained DNA-binding activity for a glutamine to alanine mutant in *Drosophila* Engrailed protein (8, 9) or a methionine to alanine mutant in human HOXC9 protein (10). Because substitution of other amino acids than alanine in these homologous positions greatly affected the DNA-binding activity in these studies, the substitution of Q144 or M148 to residues other than alanine in hNANOG HD may have stronger effects on DNA-binding activity and cellular functions.

The L122 residue is of particular interest, and its uniqueness is described in this study. Structurally, the 122 position in HD is situated in the N-terminal of the H2 helix, and it is highly diversified among HDs (2) (Fig. 1*C*). Leucine in this position is coincidentally found in mammalian OCT4 proteins and the DLX protein family. Interestingly, some other HDs, such as *Drosophila* Bicoid and vnd/NK2 proteins, bear an alanine in the equivalent position, whereas *Drosophila* Ant and Ftz and vertebrate HOXA9 and HOXB1 proteins bear an arginine. This arginine forms a salt bridge with a DNA phosphate and is required for efficient target site recognition observed in *Drosophila* Ant and Ftz structures in complexes with DNA (3, 11). From our structural analysis, we could rationalize the enhanced affinity of the L122A mutant by noting that it is proximal to Y136 in helix H3. We speculate that an alanine substitution may affect the conformation of Y136 to form more stable interactions and contribute overall to stabilizing helix 3 and foster a reading conformation for the major groove of DNA. Additionally, this mutation might stabilize an unknown protein:protein interface.

SI Materials and Methods

Protein and DNA Production and Purification. The region encoding hNANOG HD (aa 94–162, UniprotKB entry: Q9H9S0) with an N-terminal tobacco etch virus (TEV) protease-cleavable site was introduced into pET-46 Ek/Ligation Independent Cloning vector (Novagen). As a result, the vector contained an N-terminal 6-His-tag and TEV recognition site MAHHHHHHVDDDDKM-SENLYFQ/S. The construct was subsequently transformed into *Escherichia coli* BL21Star (DE3) cells (Invitrogen) and cultured in Luria-Bertani medium supplemented with 50 μg/mL ampicillin at 37 °C. Protein expression was induced with 0.2 mM isfopropyl-β-D-thiogalactopyranoside at OD_{600} ~0.6–0.8, and cells were grown at 16 °C for 16–18 h. Cells were pelleted by centrifugation at 3,500 × *g* for 15 min at 4 °C and resuspended with cold lysis buffer A [20 mM Tris HCl pH 8, 300 mM NaCl, 10% (vol/vol) glycerol, 40 mM Imidazole, 5 mM β-mercaptoethanol, 1 mM CHAPS]. The protein was purified using Ni-nitrilotriacetate beads (Qiagen) by elution with buffer A supplemented with 300 mM imidazole. The purified protein was combined with the 12-bp duplex region of the *OCT4* (POU5F1) promoter (5′-GGCCCATTCAAG-3′/ 3′-CCGGGTAAGTTC-5′) and incubated on ice for at least 30 min. The complex was isolated by size-exclusion chromatography using a GE Healthcare HiLoad 16/600 Superdex 75 prep grade column using 20 mM Tris·HCl pH 8.0, 5 mM DTT, 200 mM NaCl as the running buffer. The peak fractions were pooled and concentrated to 30 mg/mL by centrifugal ultrafiltration (Millipore) for crystallization trials. After size-exclusion chromatography, the presence of DNA in the protein sample was confirmed by agarose gel electrophoresis. Native PAGE confirmed that the complex eluted as a monomer. The hNANOG:DNA complex stability was checked by differential scanning fluorimetry.

For BLI experiments (see below), the 6-His tag was cleaved by His-tagged TEV protease (at 5:1 protein to protease ratio) in buffer B [20 mM Tris HCl, 20% (vol/vol) glycerol, 300 mM NaCl, 1 mM DTT] overnight at 4 °C. The following day, the concentration of DTT was reduced to <0.5 mM by dilution with buffer A and the sample was purified by Ni-NTA Agarose gel filtration (Qiagen). The flow-through was concentrated and loaded onto a Superdex 75 (16/60) column equilibrated with buffer C (20 mM

Tris HCl pH 8, 200 mM NaCl, 5 mM DTT). The eluted fractions were analyzed by SDS/PAGE to assess purity. The purified protein was distributed into 1-mL aliquots, flash-frozen in liquid nitrogen and stored at −80 °C.

Crystallization. The hNANOG:DNA complex was crystallized using the nanodroplet vapor diffusion method (12) with standard Joint Center for Structural Genomics (JCSG) crystallization protocols (13, 14). Sitting drops composed of 100-nL protein solution mixed with 100-nL crystallization solution in a sitting-drop format were equilibrated against a 50-μL reservoir at 293 K for 153 d before harvest. The crystallization reagent consisted of 0.2 M calcium acetate, 12% (vol/vol) 2-propanol and 0.1 M sodium cacodylate pH 5.66. Glycerol was added to a final concentration of 15% (vol/vol) as a cryoprotectant. Initial screening for diffraction was carried out at the Stanford Synchrotron Radiation Lightsource (SSRL, Stanford Linear Accelerator Center National Accelerator Laboratory, Menlo Park, CA) using the SSRL Automated Mounting sytem (15). The diffraction data were indexed in hexagonal space group $P6_522$.

X-Ray Data Collection, Structure Determination, and Refinement. A native dataset was collected at the SSRL on beamline 11–1 using the BLU-ICE data collection environment (16). The dataset was collected at 100 K using a Pilatus 6M PADdetector (Dectris). The data were integrated with MOSFLM (17) and scaled with SCALA from the CCP4 suite (18). Molecular replacement using the crystal structure of the mNANOG HD (PDB ID code 2VI6) as the search model was performed using PHASER (19) from the PHENIX software suite (20). Model building and crystallographic refinement were performed using COOT (21) and PHENIX. The refinement protocol included TLS (translation/libration/screw) refinement with six TLS groups, one for each of the NANOG HD molecules and the DNA chains in the asymmetric unit. Data and refinement statistics are summarized in Table S1 (22–24).

Validation and Deposition. The quality of the crystal structure was analyzed using the JCSG Quality Control server (smb.slac.stanford.edu/jcsg/QC). This server verifies the stereochemical quality of the model using AutoDepInputTool (25), MolProbity (26), and PHENIX, agreement between the atomic model and the data using RESOLVE (27), the protein sequence using CLUSTALW (28), the ADP distribution by using PHENIX, and differences in $R_{crystal}/R_{free}$, expected $R_{free}/R_{crystal}$, and various other items, including atom occupancies, consistency of noncrystallographic symmetry pairs, ligand interactions, special positions, and so forth, by using in-house scripts to analyze the refinement log file and PDB header. Protein quaternary structure analysis was performed using the PDBePISA server (pdbe.org/pisa/) (29). Fig. 1 *A*, *C*, and *D*, and Fig. S1 *C* and *D* were prepared using PyMOL (www.pymol.org/; DeLano Scientific) and Fig. 1*B* was adapted from an analysis using PDBsum (www.ebi.ac.uk/pdbsum).

Site-Directed Mutagenesis for Recombinant NANOG HD Mutants. The pET46 hNANOG HD vector obtained as described above was specifically mutated using KOD Hot Start DNA polymerase PCR (EMD Millipore) and DPN-I (New England BioLabs) digestion following standard protocol. The DNA oligos used for each mutation are listed in Table S3. Successful mutagenesis reaction was verified by sequencing. Recombinant hNANOG HD mutant proteins were expressed and purified following the same protocol described for the WT hNANOG HD.

Bio-Layer Interferometry. BLI experiments were performed on the Octet RED system (Fortebio) at 25 °C. The streptavidin (SA) sensors were dipped in the assay buffer [20 mM Tris·HCl pH 8, 300 mM NaCl, 5 mM $MgCl_2$, 5% (vol/vol) glycerol, and 5 mM DTT, supplemented with 0.05% Tween-20 and 0.1 mg/mL bovine serum albumin (BSA)] for 30 min to 1 h before use. In this optimized assay buffer, nonspecific background signal (given by the protein binding to the reference sensors) was significantly reduced. The *OCT4* promoter DNA duplex was purchased with a C6 biotin group attached to the 5′ end of the forward strand (IDT DNA) and loaded on the streptavidin biosensor. To avoid carryover resulting from slower dissociation rates, each sensor was assigned a single-point concentration of the protein. A seven-point concentration series were assayed for each protein, NANOG HD WT, and mutants, as described in Table S2 (20 μM to 0.5 μM range). Wells with assay buffer only were used as reference wells. Each protein concentration was also assigned a single reference sensor, where assay buffer instead of the protein was loaded on the streptavidin biosensor. The double reference sensors/wells were then subtracted from the raw data and processed with the Octet Data Analysis Software. The binding assay protocol included the following steps: wash for 60 s (to establish a stable base line before starting association-dissociation cycles of the proteins), immobilization of the biotin-conjugated *OCT4* promoter 12-mer (20 μg/mL) for 150 s, baseline for 150 s, association for 100 s or 150 s, and dissociation for 300 s. The last three steps were repeated for all of the protein concentrations for two more cycles. After each cycle, biosensors were regenerated with 1 M $MgCl_2$ for two steps of 15 s. The binding assays were performed at least in triplicate with three different protein batches. Dissociation (k_{diss}) and association rate constants (k_{ass}) were determined with the Octet Data Analysis Software, as a result of a global fit considering the entire step times, and assuming a 1:1 binding model.

Differential Scanning Fluorimetry. Recombinant NANOG HD WT and mutant protein thermal unfolding was monitored using 2.5× Sypro Orange (Invitrogen). Recombinant proteins were diluted to a final concentration of 20 μM in assay buffer [20 mM Tris HCl pH 8, 200 mM NaCl, 10% (vol/vol) glycerol, 5 mM DTT containing 2.5× Sypro Orange dye diluted from 5,000× stock] and incubated with 40 μM *OCT4* promoter 12-bp duplex for 30 min at 4 °C. Samples were aliquoted to a final volume of 100 μL in a 96-well PCR plate (Agilent Technologies) and sealed with optical quality sealing (Applied Biosystems). Each well was repeated in triplicate. Thermal unfolding was carried out using ViiA 7 Real-Time PCR System (Life Technologies) by heating samples from 25 °C to 95 °C with 3 °C increments per minute. The fluorescent intensity was measured using excitation/emission wavelengths of 470 nm/558 nm. Thermal denaturation graphs were plotted using Prism 6 (GraphPad). Fluorescence intensity data were normalized to the baseline and plotted as percent of unfolded protein in function of temperature. Melting curves were fitted using nonliner Boltzmann sigmoidal equation and T_m values (V50) were calculated using the Boltzmann equation in Prism 6.

Mouse Cell Culture. For mESCs, the RF8 line (a gift from R. V. Farese Jr., Harvard University, Boston) (30) were cultured in conventional culture medium ("Normal" culture condition) consisting of Knockout-DMEM (Gibco) supplemented with 15% (vol/vol) FCS (Gibco), nonessential amino acid (Gibco), Glutamax (Gibco), β-mercaptoethanol (Gibco), and 1,000 U/mL LIF (Millipore). The medium was changed every day. The cells were passaged every 3–5 d. For differentiation assays, the cells were seeded at 5,000 cells/cm^2 onto gelatin-coated dishes in the same medium only without LIF or the same medium only ("–LIF" condition) without LIF and added 5 μM RA (Sigma R2625; "+RA" condition).

For EpiSCs, the EpiSC-5 line (a gift from Paul Teser, Case Western Reserve University, Cleveland) (31) was cultured in N2B27 medium (StemCells or Clontech) supplemented with basic FGF (10 ng/mL; Millipore) and Activin A (10 ng/mL; R&D Systems) on fibronectin (1 μg/cm^2; Sigma F1141)-coated dishes. The medium was changed every day. The cells were passaged every 3–5 d. For reprogramming assays, the cells were seeded at 5,000 cells/cm^2 onto fibronectin-coated six-well plates in the

same medium. The next day, the medium was changed to "2i+LIF" medium made of N2B27 medium supplemented with 1,000 U/mL mouse recombinant LIF (Millipore) and 2i (mixture of G3K3β and MEK inhibitors; Millipore). The medium was changed every day for 5 d.

Constitutive Expression of Full-Length NANOG Mutants by pPyCAG Vectors in Mammalian Cells. Each mNANOG mutant was constructed as follows. pPyCAG-Nanog-IP plasmids (Addgene) were used as template in site-mutagenesis. Site mutagenesis was performed using GeneArt site-mutagenesis kit (Invitrogen) with KOD Xtreme PCR polymerase (Toyobo). Thermal cycle conditions are 37 °C for 20 min, 98 °C for 2 min, 20 cycles of 94 °C for 10 s, 60 °C for 15 s, and 68 °C for 10 min. After recombining, PCR products are transformed and cultured in Oneshot Max-Efficiency DH5α strain (Invitrogen) onto a LB/ampicillin agar plate. Then, each colony are picked up and cultured in LB/ampicillin medium. Then, mNANOG inserts were verified by sequencing with CAG-F and IRES-R primer sets. pPyCAG-EGFP-IP plasmids (a gift of Hitoshi Niwa, RIKEN, Kobe, Japan, through RIKEN DNA bank) were used as a control. Primer sequences for each mutant are shown in Table S3. Primer sequences for adding N-terminal 3×FLAG tag are 5-AAA GCA GGC TCT GAC ATG GAC TAC AAA GAC CAT GAC GGT GAT TAT AAG GAT CAC GAT ATC GAC TAC -3 and 5- ACC AGG AAG ACC CAC ACT CTT GTC ATC GTC ATC CTT GTA GTC GAT ATC GTG ATC CTT ATA ATC ACC-3.

Electroporation using an Amaxa Nucleofection device (Lonza; A-23 program) was performed to transfect each plasmid (4 μg) into mESCs or mEpiSCs (2×10^6 cells) with Mouse ES Cell Nucleofector kit (Lonza). Puromycin selection (2 μg/mL) was started 2 d after the electroporation and was continued throughout experiments.

PiggyBac Vector Transfection and Excision. PiggyBac transposon-based plasmid vectors were constructed with pHULK backbone (DNA2.0). (N-term) 3×FLAG-tagged hNANOG-2A-mCherry cassette was inserted under EF1α promoter. Site mutagenesis was performed using GeneArt site-mutagenesis kit (Invitrogen) with KOD Xtreme PCR polymerase (Toyobo). Primer sequences for each mutant are shown in Table S3. Electroporation using an Amaxa Nucleofection device (Lonza; A-23 program) was performed to transfect each plasmid (4 μg) into mEpiSCs (2×10^6 cells) with Mouse ES Cell Nucleofector kit (Lonza). Puromycin selection (2 μg/mL) was started 2 d after the electroporation and was continued throughout experiments. Reprogramming into ground-state was performed in 2i+LIF medium for 10 d. Then, the cells were passaged onto gelatin-coated dishes in the same medium. After 2 passages, pPyCAG-EGFP-IB plasmids (a gift of Hitoshi Niwa, RIKEN, Kobe, Japan, through RIKEN DNA bank) was transfected using electroporation. Blasticidin selection (10 μg/mL) was started 2 d after the electroporation and was continued throughout experiments. To delete transgenes, 4 μg of Excision Only PiggyBac Transposase Expression Vector (System Biosciences) was transfected using electroporation. Five days after the electroporation, GFP^+ and $mCherry^-$ cells were sorted with BD FACSAria III. The sorted cells were cultured in the same conditions for several passages until they were used in the chimeric mouse formation.

Embryo Manipulation for Chimeric Mouse Formation Assay. Mouse 8-cell/morula-stage embryos collected in M2 medium (Millipore) from the oviduct and the uterus of C57BL/6 mice 2.5 d postcoitum were transferred into KSOM-AA medium (Millipore) and were cultured for 24 h before blastocyst injection. For micromanipulation, dissociated PSCs were suspended in PSC culture medium. A piezo-driven micromanipulator (Prime Tech) was used to drill zona pellucida and trophectoderm under the microscope, and 10 PSCs were introduced into blastocyst cavities near the inner cell mass. After blastocyst injection, embryos underwent follow-up culture for 1–2 h, after which they were transferred into the uteri of pseudopregnant recipient ICR mice (2.5 d postcoitum). All of the procedures were performed at Gladstone Transgenic Core Laboratory.

Human iPS Cell Culture. For HiPSCs, the HiPS-WTc11 line (GM25256 at Coriell Institute) was generated by electroporation of episomal plasmid vectors carrying OCT4, SOX2, KLF4, L-MYC, LIN28, and shRNA for TP53 (32). This HiPSC line was originated from dermal fibroblast of a 30-y-old Japanese male and was verified to maintain self-renewal, pluripotent, integration-free, and normal karyotype for long-term culture. This line was cultured in mTeSR1 medium (Stemcell Technologies) on recombinant Laminin-511-E8 fragment (iMatrix-511; 0.25 μg/cm^2; Iwai North America)-coated dishes. The medium was changed every day. The cells were passaged every 4–6 d at 1:5 ~1:20 split ratio with Accutase (Millipore). ROCK Inhibitor, 10 μg/mL of Y27632 (Wako) was added to the medium only at passaging day. For reprogramming assays, the cells were seeded in 2i+LIF medium [including 10 ng/mL human recombinant LIF (Peprotech) instead of mouse LIF]. For single-cell survival assay, the cells were dissociated with Trypsin-EDTA (0.05%; Life Technologies) with pipetting and then defined trypsin inhibitor (Life Technologies) was added. Then, the cells were seeded at 1,000 cells per well of a six-well plate in the medium without ROCK inhibitor.

RNA Extraction, Reverse Transcription, and qPCR. Total RNA was purified with RNeasy Plus Mini kit (Qiagen), which includes the steps to eliminate genomic DNA. One microgram of total RNA was used for reverse-transcription reaction with SuperScript III First-Strand Synthesis System (Invitrogen). qPCR was performed with Taqman Gene Expression Master Mix or Power SYBR green master mix and analyzed with the 7900 real-time PCR system (all from Applied Biosystems). The expression levels of each gene are normalized with the amount of Gapdh expression in the same samples. All of the procedures were according to the manufacturer's recommendations. Taqman probe numbers and primer sequences used for qPCR are shown below; *Gapdh* (Mm99999915_g1), *Oct4* (*Pou5f1*; Mm00658129_gH), *Sox2* (Mm00488369_s1), *Klf4* (Mm00516104_m1), *Fgf5* (Mm00434989_m1), *Rex1* (*Zfp42*; Mm03053975_g1), *T* (*Brachyury*; Mm01318252_m1), *Dax1* (*Nr0b1*; Mm00431729_m1), *Esrrb* (*Nr3b2*; Mm00442411_m1), total *mNanog* (Mm02384862_g1), endogenous *mNanog* (Mm02019550_s1), *Blimp1* (*Prdm1*; Mm00476128_m1), *Stella* (*Dppa3*; Mm01184198_g1) and *3×FLAG – mNanog* (FW: 5′-GACTACAAAGACCATGACGGTGAT-3′, RV: ACTGCA-GGCATTGATGAGGCGTTC-3′), *GAPDH* (Hs02758991_g1), *OCT4* (Hs04260367_gH), endogenous *NANOG* (Hs04399610_g1), *TFCP2L1* (Hs00232708_m1), *KLF2* (Hs00360439_g1), *KLF4* (Hs00358836_m1), *KLF5* (Hs00156145_m1), *REX1* (Hs01938187_s1), *PRDM14* (Hs01119056_m1), *STELLA* (Hs01931905_g1), *FGF5* (Hs03676587_s1), *DAX1* (Hs03043658_m1), and *ESRRB* (Hs01584024_m1).

RNA-Seq. RNA-seq libraries were constructed according the manufacturers guidelines, using the TruSeq Stranded total RNA Sample Prep Kit (Illumina) at Gladstone Genomics Core Laboratory. These libraries were sequenced with Illumina TruSeq v3 cluster generation and sequencing chemistry for paired-end 100-bp reads on an Illumina HiSeq2500 (Illumina) at the University of California, San Francisco genomic core facilities.

Reads were mapped to assembly NCBI37/mm9 of the mouse genome using Bowtie2 and constructed spliced alignments using Tophat2 with default settings. Following read mapping, we selected fragments (read-pairs) where at least one of mate-pairs had a quality score of >10, aligned with no gaps, with three base

mismatches or less. Any read pairs with an insert size less than 150 bp or greater than 1 Mb, or on different chromosomes, were excluded from subsequent analyses. The amount of each gene expression was determined with log normalized fragments per kilobase of transcript per million reads sequenced-adjusted counts (counts adjusted by the gene length and sequencing depth). Heatmap was drawn using the "heatmap.2" function in R script.

Western Blotting. To determine protein expression in mESCs overexpressing mNANOG mutants in RA-added conditions, the cells were collected at day 5 after RA addition. To determine protein half-life of mNANOG in mESCs, the protein translation was inhibited by treating cells with 100 μg/mL cycloheximide (Sigma) and the cells were collected at different time points (0, 0.5, 1, 2, 4, and 8 h). To determine protein expression in EpiSCs overexpressing mNANOG mutants, the cells were collected at day 5 cultured in 2i+ LIF medium. These cells were lysed with 2× Laemmli sample buffer (Bio-Rad) with β-mercaptoethanol directly. Cell lysates were separated by electrophoresis on 4–20% (wt/vol) SDS-polyacrylamide gel (456-1094, Bio-Rad) and transferred to a nitrocellulose membrane with the iBlot transfer system (IB3010-01, Invitrogen). The membrane was blocked with Odyssey Blocking Buffer (Li-COR Biosciences) and incubated with primary antibody diluted in the blocking buffer in SNAP i.d. protein detection system (Millipore). After washing with TBS with 0.1% Tween 20 buffer, the membrane was incubated with IRDye 680LT- or IRDye 800CW-conjugated donkey secondary antibody (Li-COR Biosciences) diluted in the blocking buffer. Signals were detected with Odyssey Imaging Systems (Li-COR Biosciences). Primary antibodies used for Western blotting were mouse monoclonal anti-OCT3/4 (Santa Cruz Biotechnology; C-10 clone, 1 μg/mL,), goat anti-mNANOG (Reprocell; 0.4 μg/mL), mouse monoclonal anti-SOX2 (5 μg/mL; R&D Systems), mouse monoclonal anti-ESRRB (Perseus Proteomics; 5 μg/mL), goat anti-BRACHYURY (R&D Systems; 1 μg/mL), mouse monoclonal anti-FLAG antibody (Sigma; M2 clone, 5 μg/mL), and chicken anti-GAPDH antibody (0.2 μg/mL; Millipore). Each protein amount was calculated from the band intensity with ImageJ software.

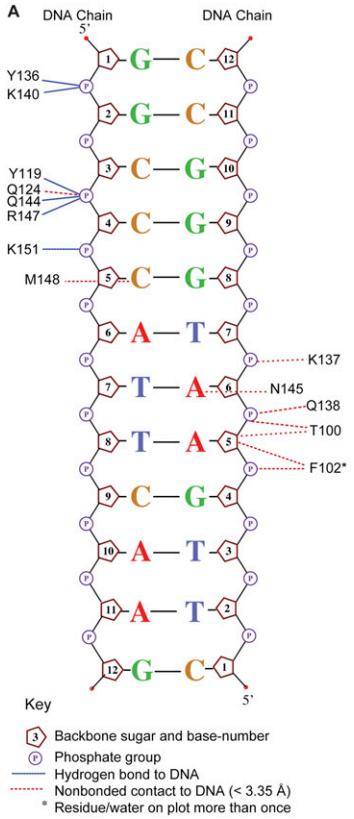
AP Staining and Immunocytochemistry. For AP staining, the cells were fixed with PBS containing 4% (vol/vol) paraformaldehyde for 2 min at room temperature. Staining was performed using the Alkaline Phophatase detection kit (Millipore), according to the manufacturer's recommendations.

For Immunocytochemistry, the cells were fixed with PBS containing 4% (vol/vol) paraformaldehyde for 10 min at room temperature. The cells were permeabilized with PBS containing 0.1% Triton X-100 for 10 min at room temperature, washed with PBS twice, treated with 1% BSA for blocking, and incubated with primary antibodies for overnight at 4 °C. The primary antibodies used were mouse monoclonal anti-OCT4 (1 μg/mL, C-10 clone; Santa Cruz Biotechnology). Then, the cells were washed with PBS for 3 times, incubated with secondary antibodies for 1 h at room temperature. The secondary antibodies used were Alexa488-conjugated donkey anti-goat IgG (1 μg/mL; Invitrogen). Then, the cells were sashes with PBS three times and were stained with DAPI contained in the Vectashield set (Vector Laboratories). The images were taken with BZ-9000 (BIOREVO) fluorescence microscope (Keyence) and analyzed with BZ Image Analysys software (Keyence) to count $Oct4^+$ cell ratio automatically.

Flow Cytometry. Flow cytometry analysis was performed according to standard procedures. Anti-CD31 antibody conjugated with Allophycocyanin (1 μL per sample; BD Biosciences) were incubated with cell samples for 30 min. Samples were analyzed with MACSQuant VYB flow cytometer (Miltenyl Biotec) and FlowJo software.

Statistics. Results are shown as mean ± SEM from an appropriate number of samples indicated in the figure legends. Dunnett's test was used to determine statistical significance in Fig. 3*D* using R statistical language. Student's *t* test (one-tailed, paired) was used to determine statistical significance in Figure S6 using Microsoft Excel "TTEST" function.

1. Chirgadze YN, Sivozhelezov VS, Polozov RV, Stepanenko VA, Ivanov VV (2012) Recognition rules for binding of homeodomains to operator DNA. *J Biomol Struct Dyn* 29(4):715–731.
2. Jauch R, Ng CK, Saikatendu KS, Stevens RC, Kolatkar PR (2008) Crystal structure and DNA binding of the homeodomain of the stem cell transcription factor Nanog. *J Mol Biol* 376(3):758–770.
3. Gehring WJ, et al. (1994) Homeodomain-DNA recognition. *Cell* 78(2):211–223.
4. Baird-Titus JM, et al. (2006) The solution structure of the native K50 Bicoid homeodomain bound to the consensus TAATCC DNA-binding site. *J Mol Biol* 356(5):1137–1151.
5. Rajasekaran M, Chen C (2012) Structural effect of the L16Q, K50E, and R53P mutations on homeodomain of pituitary homeobox protein 2. *Int J Biol Macromol* 51(3): 305–313.
6. Babin V, Wang D, Rose RB, Sagui C (2013) Binding polymorphism in the DNA bound state of the Pdx1 homeodomain. *PLOS Comput Biol* 9(8):e1003160.
7. Clarke ND (1995) Covariation of residues in the homeodomain sequence family. *Protein Sci* 4(11):2269–2278.
8. Grant RA, Rould MA, Klemm JD, Pabo CO (2000) Exploring the role of glutamine 50 in the homeodomain-DNA interface: Crystal structure of engrailed (Gln50 → ala) complex at 2.0 A. *Biochemistry* 39(28):8187–8192.
9. Ades SE, Sauer RT (1994) Differential DNA-binding specificity of the engrailed homeodomain: The role of residue 50. *Biochemistry* 33(31):9187–9194.
10. Zhou B, Liu C, Xu Z, Zhu G (2012) Structural basis for homeodomain recognition by the cell-cycle regulator Geminin. *Proc Natl Acad Sci USA* 109(23):8931–8936.
11. Furukubo-Tokunaga K, et al. (1992) In vivo analysis of the helix-turn-helix motif of the fushi tarazu homeo domain of *Drosophila melanogaster*. *Genes Dev* 6(6):1082–1096.
12. Santarsiero BD, et al. (2002) An approach to rapid protein crystallization using nanodroplets. *J Appl Cryst* 35(Pt 2):278–281.
13. Elsliger MA, et al. (2010) The JCSG high-throughput structural biology pipeline. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 66(Pt 10):1137–1142.
14. Lesley SA, et al. (2002) Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline. *Proc Natl Acad Sci USA* 99(18):11664–11669.
15. Cohen AE, Ellis PJ, Miller MD, Deacon AM, Phizackerley RP (2002) An automated system to mount cryo-cooled protein crystals on a synchrotron beam line, using compact sample cassettes and a small-scale robot. *J Appl Cryst* 35(6):720–726.
16. McPhillips TM, et al. (2002) Blu-ice and the distributed control system: Software for data acquisition and instrument control at macromolecular crystallography beamlines. *J Synchrotron Radiat* 9(Pt 6):401–406.
17. Leslie AGW, Powell HR (2007) Processing diffraction data with MOSFLM. *Evolving Methods for Macromolecular Crystallography* 245:41–51.
18. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760–763.
19. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
20. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
21. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
22. Diederichs K, Karplus PA (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* 4(4):269–275.
23. Weiss MS, Hilgenfeld R (1997) On the use of the merging R factor as a quality indicator for X-ray data. *J Appl Cryst* 30(Pt 2):203–205.
24. Weiss MS, Metzner HJ, Hilgenfeld R (1998) Two non-proline *cis* peptide bonds may be important for factor XIII function. *FEBS Lett* 423(3):291–296.
25. Yang H, et al. (2004) Automated and accurate deposition of structures solved by X-ray diffraction to the Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* 60(Pt 10):1833–1839.
26. Davis IW, et al. (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35(web server issue):W375–W383.
27. Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr* 56(Pt 8):965–972.
28. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21): 2947–2948.
29. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797.
30. Meiner VL, et al. (1996) Disruption of the acyl-CoA:cholesterol acyltransferase gene in mice: evidence suggesting multiple cholesterol esterification enzymes in mammals. *Proc Natl Acad Sci USA* 93(24):14041–14046.
31. Tesar PJ, et al. (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448(7150):196–199.
32. Okita K, et al. (2011) A more efficient method to generate integration-free human iPS cells. *Nat Methods* 8(5):409–412.



B

Residue	HSDC	ASA	BSA	BAP	$\Delta^{I}G$
T100		184.6	1.84	/	0.03
Y119	H	152.4	38.5	///	0.21
L120		20.68	8.03	////	-0.1
L122		95.43	12.7	//	0.2
M125		10.35	8.52	//////////	0.17
Y136		78.09	22.2	///	0.01
K140	H	101	61.7	///////	-0.6
Q144	H	77.3	60	///////	-0.3
N145		86.3	6.86	/	-0.1
R147	H	26.34	21.8	//////////	-2
M148		88.11	56.8	///////	0.96
K151	H	135.9	44.5	////	-1.2

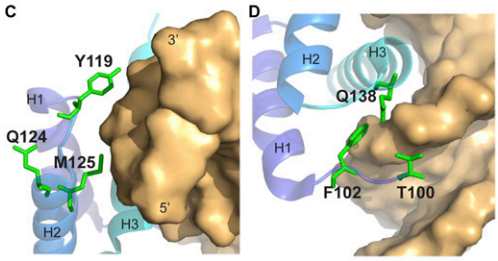


Fig. S1. (*A*) Schematic of hNANOG HD residues and their potential interaction with DNA in the crystal structure, modified from the data of NUCPLOT. (*B*) Residues which have potential interaction with DNA in the crystal structure, modified from the data of PDBePISA servers. ASA, accessible surface area (Å^2); BAP, buried area percentage (one bar per 10%); BSA, buried surface area (Å^2); HSDC, residues making a hydrogen/disulphide bond, salt bridge, or covalent link; $\Delta^{I}G$, solvation energy effect (kcal/mol). (*C*) Residues in hNANOG HD (Y119, Q124, and M125) are represented in green. (*D*) Residues in hNANOG HD (T100, F102, and Q138) are represented in green.

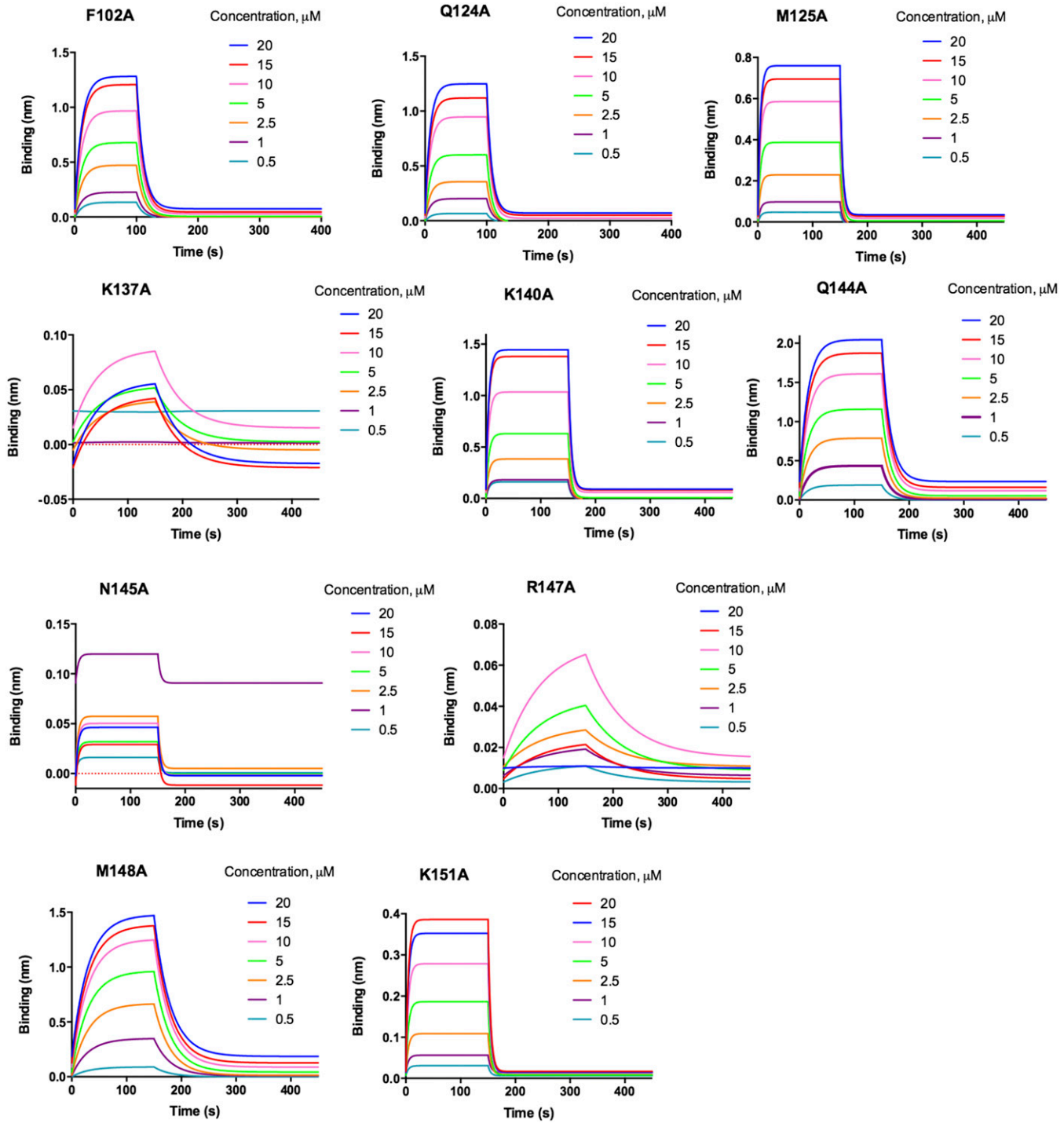


Fig. S2. Octet BLI analysis of hNANOG HD mutants on biotinylated *OCT4* promoter DNA (continued from Fig. 2). Dose–response curves of F102A, Q124A, M125A, K137A, K140A, Q144A, N145A, R147A, M148A, and K151A showing processed binding response (nm) to biotinylated *OCT4* promoter DNA of a range of hNANOG HD protein concentrations. Note: All of the hNANOG HD mutants showed a 1:1 type binding interaction, which could be fitted with a single exponential model. Some proteins (L122A, K140A, Q144A, and M148A) deviated from this model showing a heterogeneous profile at the highest concentrations assayed (20 and 15 μM). Such behavior was attributed to nonspecific protein interaction to the *OCT4* promoter DNA rather than a more complex binding event. The binding data were easily fit as a 1:1 model when association times were reduced.

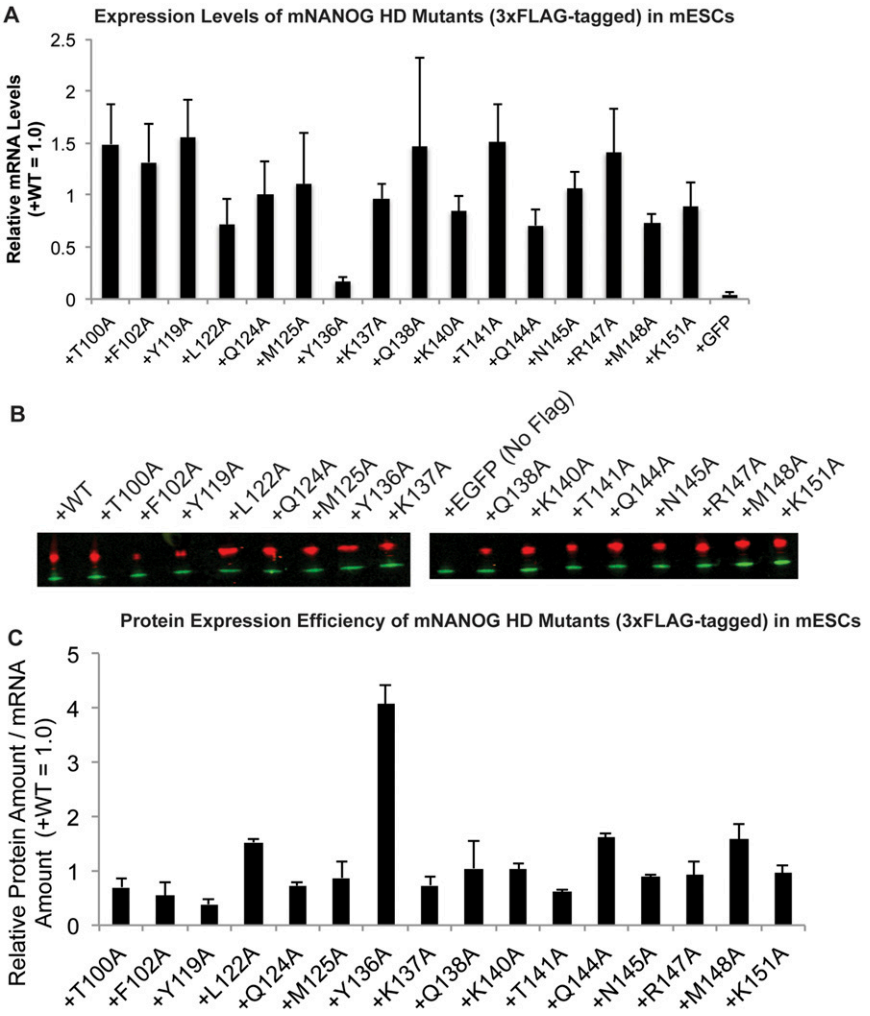


Fig. S3. Expression levels of 3×FLAG-tagged mNANOG WT and mutants in mESCs. (*A*) mRNA expression of 3×FLAG-mNANOG normalized with Gapdh expression in mESCs overexpressing 3×FLAG-tagged mNANOG WT or HD mutants or GFP detected by RT-qPCR analysis. The values are mean + SEM. $n = 3$. (*B*) Protein expression of 3×FLAG protein (red) and GAPDH (Green) in the same samples used for qRT-PCR in *A* detected by Western blotting analysis. (*C*) The efficiency of exogenous mNANOG protein expression calculated from protein amounts of 3×FLAG-mNANOG divided by their mRNA expression levels. The values are mean + SEM. $n = 3$.

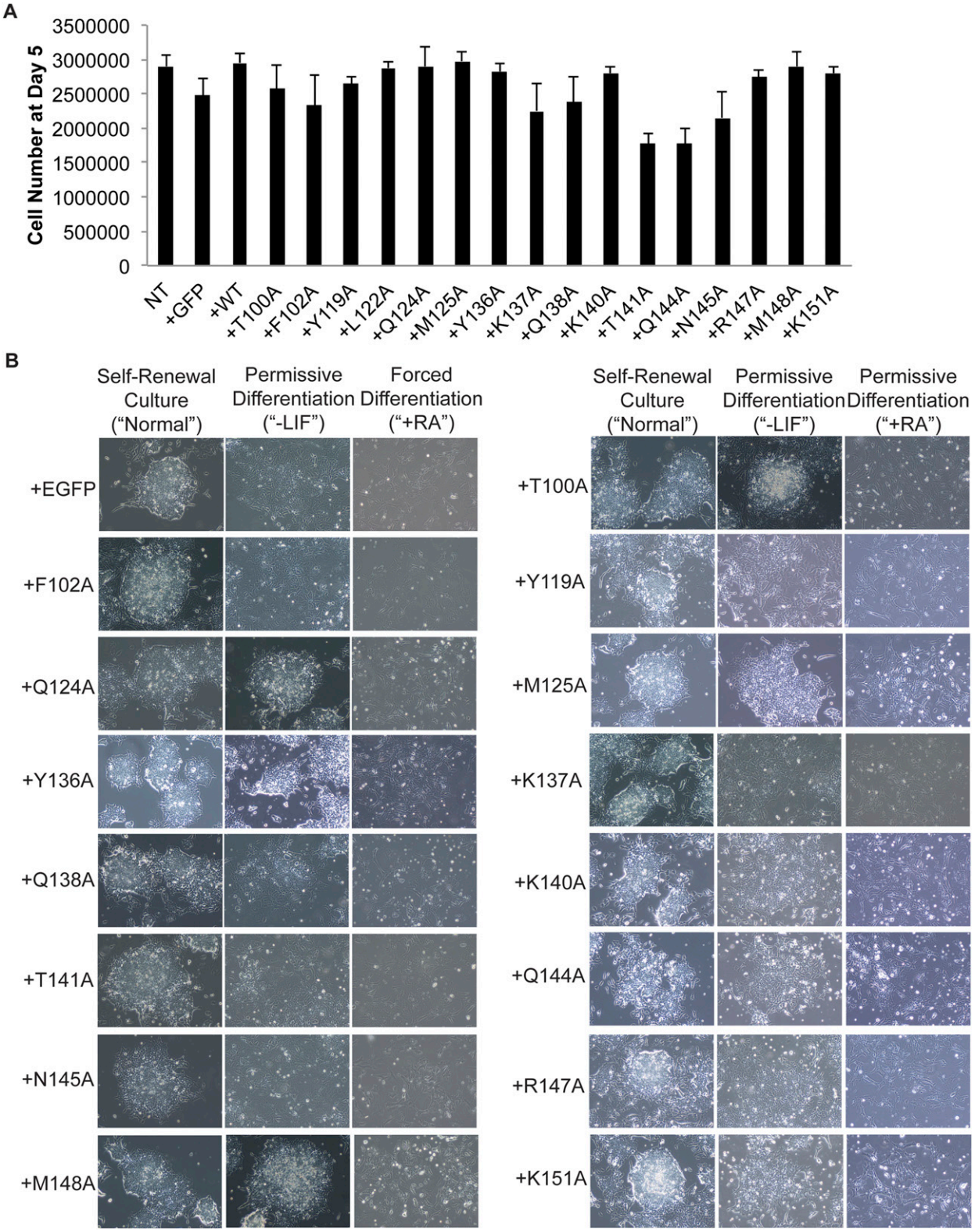


Fig. S4. The effects of mNANOG mutants on mESC morphologies and proliferation. (*A*) Cell growth of mESCs overexpressing mNANOG mutants in Normal culture conditions. Cell numbers were counted 5 d after seeding at 50,000 cells per well of a six-well plate. The values are mean + SEM. $n = 4$. (*B*) Typical colony morphologies of mESCs overexpressing each mNANOG mutants (T100A, F102A, Y119A, Q124A, M125A, Y136A, K137, Q138A, K140A, T141A, Q144A, N145A, R147A, M148, or K151A) in Normal, −LIF, or +RA culture conditions (continued from Fig. 4*A*). The images were taken 5 d after seeding. Note: when these mESCs transfectantants were cultured with LIF (i.e., Normal culture condition), they grew stably (Fig. S4*A*) and showed undifferentiated cell morphologies (this panel and Fig. 4*A*) and did not decrease their gene expression of self-renewal markers [i.e., *Rex1* (*Zfp42*), endogenous *Nanog*, *Oct4 (Pou5f1)*, *Sox2*, *Dax1* (*Nr0b1*), or *Esrrb* (*Nr3b2* or *Errβ*)] (Fig. 4*B* and Fig. S5*A*). These results indicated that these mutants did not disrupt mESC self-renewal in the Normal culture condition.

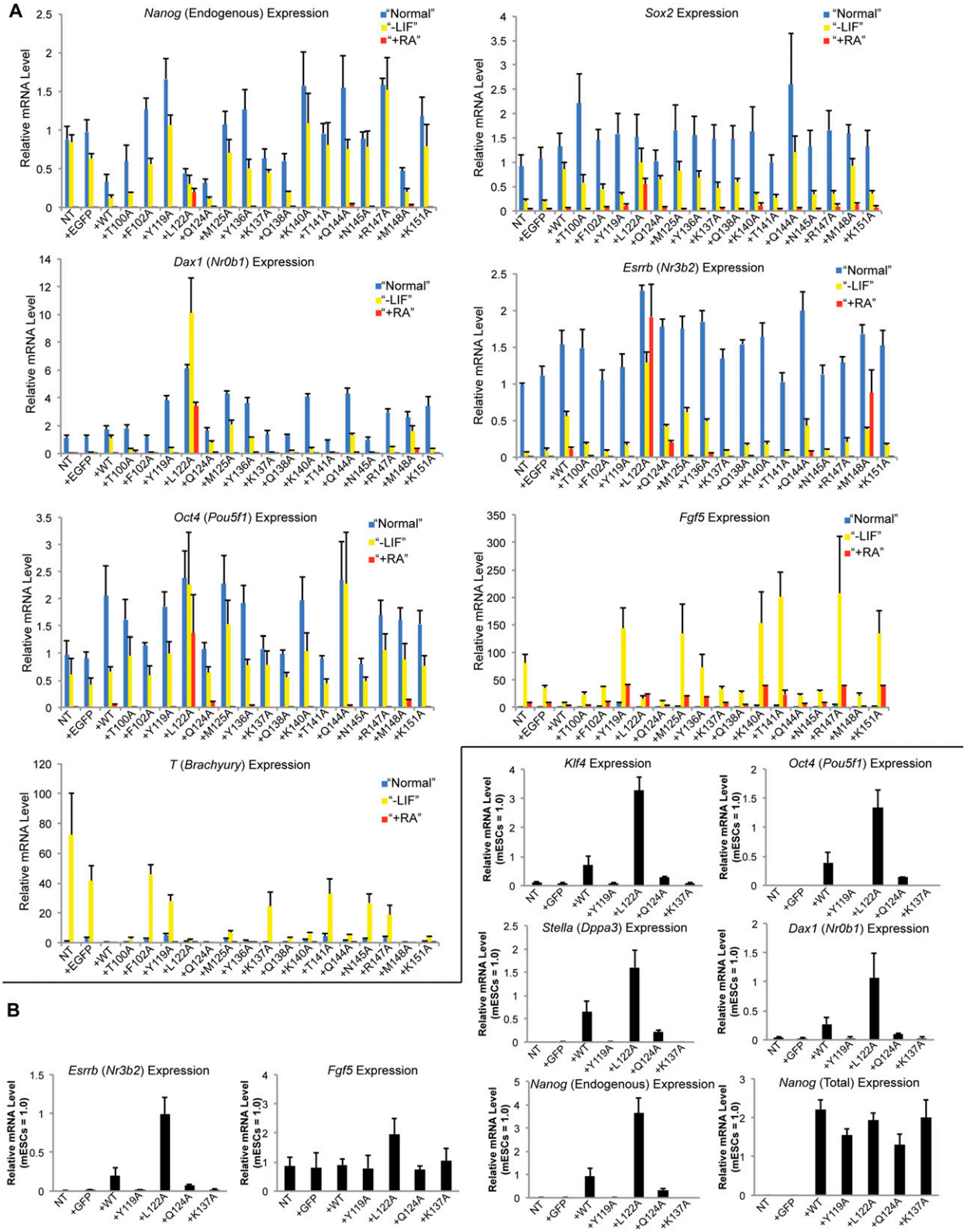


Fig. S5. Gene expression of pluripotency markers in mESCs and EpiSCs. (*A*) Gene expression of endogenous *Nanog*, *Sox2*, *Dax1 (Nr0b1)*, *Esrrb* (*Nr3b2*), *Fgf5*, *T*, or *Oct4* (*Pou5f1*) detected by RT-qPCR in mESC transfectants overexpressing each mNANOG mutant in Normal, −LIF, or +RA culture conditions (continued from Fig. 4*B*). The amount of an undifferentiated mESC sample was set as 1.0. $n = 4$, values are mean + SEM. (*B*) Gene expression of total *Nanog,* endogenous *Nanog*, *Oct4*, *Dax1,* Esrrb, *Klf4*, *Stella*, and *Fgf5* detected by RT-qPCR in EpiSC transfectants overexpressing each mNANOG mutant cultured in 2i+LIF medium for 5 d (continued from Fig. 5C). The amount of an undifferentiated mESC sample was set as 1.0. $n = 4$, values are mean + SEM. Note: total *mNanog* expression was similar in all of the mNANOG transfectants, confirming that the amount of *mNanog* expression was not the cause of the phenotypical differences among the mNANOG transfectants.

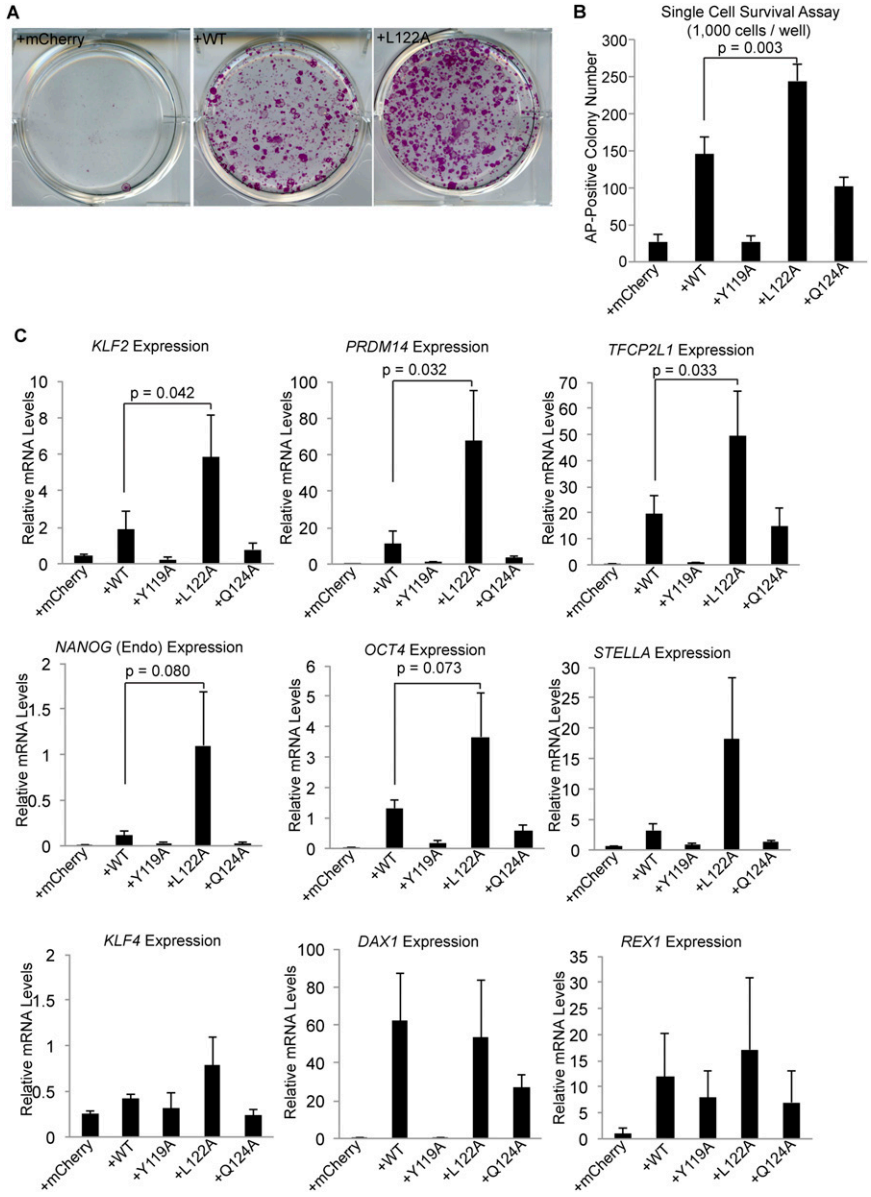


Fig. S6. The effect of L122A on putative characteristics of ground-state pluripotency in human iPSCs. (*A*) Colony forming activity from a single cell of dissociated human iPSCs treated with Trypsin/EDTA. AP staining (shown in red) was performed on mCherry, WT, or L122A-overexpressing human iPSCs seeded at 1,000 cells per well and cultured in LIF2i medium for 12 d. The images were taken from a whole well of six-well plate. (*B*) AP^+ colony numbers counted in a well of a six-well plate. $n = 4$, values are mean +SEM. *P* values of Student's *t* test between WT and L122A samples are shown. (*C*) Gene expression of pluripotency markers in human iPSC transfectants cultured in LIF2i medium for 5 d. The amount of an undifferentiated human iPSC sample cultured in mTeSR1 medium was set as 1.0. $n = 6$, values are mean + SEM. *P* values (less than 0.1) of Student's *t* test between WT and L122A samples are shown.

Table S1. Summary of crystal parameters, data collection, and refinement statistics for hNANOG HD- *OCT4* (*Pou5f1*) promoter DNA complex (PDB ID code 4RBO)

Data collection and refinement statistics	4RBO
Space group	P 6_5 2 2
Unit cell parameters (Å)	a = 215.6, b = 215.6, c = 42.3
Data collection	
Wavelength (Å)	1.0
Resolution range (Å)	29.9–3.30 (3.39–3.30)
No. of observations	30,402
No. of unique reflections	8,834
Completeness (%)	97.0 (94.3)
Mean $I/\sigma(I)$	9.4 (1.9)
R_{merge} on I^* (%)	8.3 (55.2)
R_{meas} on $I^\dagger$ (%)	9.9 (65.8)
$R_{p.i.m.}$ on $I^\ddagger$ (%)	5.2 (35.0)
Model and refinement statistics	
Resolution range (Å)	29.9–3.3
No. of reflections (total)[§]	8,823
No. of reflections (test)	417
Completeness (% total)	96.1
Cutoff criteria	$\|F\|>0$
R_{cryst}[¶]	23.5
R_{free}[¶]	26.8
Stereochemical parameters	
Restraints RMSD bond length (Å)	0.003
Restraints RMSD bond angle (°)	0.8
Average isotropic $B^{\|\|}$ value, protein (Å^2)	174.1
Average isotropic $B^{\|\|}$ value, DNA (Å^2)	205.9
Coordinate error (maximum-likelihood based)	0.48 (Å)
Protein residues/atoms	109/1,873
Nucleic acid residues/atoms	48/1,518
Ramachandran plot: residues (%) in favored/allowed	97.1%/2.9%
Matthews' coefficient (an estimated solvent content)	~4.5 Å^3 (~73%)

Numbers in parentheses are for the outer shell.

*$R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - <I(hkl)>|/\Sigma_{hkl}\ \Sigma_i(hkl)$.

†$R_{meas} = \Sigma_{hkl}[N/(N-1)]^{1/2}\Sigma_i|I_i(hkl) - <I(hkl)>|/\Sigma_{hkl}\Sigma_i I_i(hkl)$ (22).

‡$R_{p.i.m}$ (precision-indicating R_{merge}) $= \Sigma_{hkl}[(1/(N-1)]\ 1/2\ \Sigma_i|I_i\ (hkl) - <I(hkl)>|/\Sigma_{hkl}\Sigma_i\ Ii(hkl)$ (23, 24).

§Typically, the number of unique reflections used in refinement is slightly less than the total number that were integrated and scaled. Reflections are excluded because of systematic absences, negative intensities, and rounding errors in the resolution limits and unit-cell parameters.

¶$R_{cryst} = \Sigma_{hkl}||F_{obs}| - |F_{calc}||/\Sigma_{hkl}|F_{obs}|$, where F_{calc} and F_{obs} are the calculated and observed structure-factor amplitudes, respectively. R_{free} is the same as R_{cryst} but for a small percentage of the total reflections chosen at random and omitted from refinement.

‖This value represents the total B that includes TLS and residual B components.

Table S2. Summary of alanine scanning experiments in this study

NANOG HD species	Purified protein	Binding affinity (K_D, μM)	k_{on} ($\times 10^4$ $mol^{-1} \cdot s^{-1}$)	k_{off} (s^{-1})	Full R^2 (fitted according a 1:1 binding model)	Purified protein-DNA complex stability (T_m, °C)	Protein half-life in mESCs (h)	mESC self-renewal without LIF	mESC self-renewal with RA
WT	Yes	5.9 ± 1.4	1.71 ± 0.546	0.195 ± 0.0079	0.988	61.06	3.1 ± 0.5	+	–
T100A	No	NA	NA	NA	NA	NA	ND	–	–
F102A	Yes	7.1 ± 1.5	1.81 ± 0.413	0.101 ± 0.0039	0.972	58.24	1.5 ± 0.2	–	–
Y119A	No	NA	NA	NA	NA	NA	1.5 ± 0.3	–	–
L122A	Yes	1.4 ± 0.32	1.87 ± 0.435	0.170 ± 0.0065	0.987	64.56	11.5 ± 0.7	+	+/–
Q124A	Yes	10 ± 2.3	0.68 ± 0.349	0.121 ± 0.0046	0.978	58.53	3.0 ± 0.8	+	–
M125A	Yes	9.1 ± 2.0	2.29 ± 0.934	0.228 ± 0.013	0.976	55.71	2.7 ± 0.3	+	–
Y136A	Yes	17 ± 6.6	2.13 ± 0.814	0.260 ± 0.016	0.978	57.47	1.7 ± 0.4	+	–
K137A	Yes	∞	NA	NA	NA	NA	1.4 ± 0.4	–	
Q138A	No	NA	NA	NA	NA	NA	ND	–	–
K140A	Yes	13 ± 2.9	8.12 ± 0.817	0.355 ± 0.023	0.980	55.76	1.8 ± 0.4	–	–
T141A	Yes	∞	NA	NA	NA	NA	2.0 ± 0.4	–	–
Q144A	Yes	6.3 ± 1.3	1.26 ± 0.376	0.102 ± 0.0043	0.967	55.62	2.8 ± 0.4	+	–
N145A	Yes	∞	NA	NA	NA	NA	3.5 ± 0.3	–	–
R147A	Yes	∞	NA	NA	NA	NA	1.2 ± 0.1	–	–
M148A	Yes	4.2 ± 0.85	1.08 ± 0.194	0.058 ± 0.0017	0.972	64.41	3.5 ± 0.4	+	–
K151A	Yes	13 ± 0.84	4.55 ± 0.827	0.210 ± 0.028	0.976	55.15	2.1 ± 0.3	–	–

Each symbol and abbreviation indicates as follows: NA, not applicable; ND, not determined; +, positive; +/–, weakly positive; –, negative.

Table S3. Primer sets used for mutagenesis of hNANOG HD recombinant proteins and full-length mNANOG expression vectors

Name	Primers (5′-3′)
T100A hNANOG	AAACAGAAGACCAGA**GCG**GTGTTCTCTTCCACCCAGCTG
T100A mNANOG	CCAGGAAGCAGAAGAT**GCG**GGCTGTGTTCTCTCAGGCCCA
F102A hNANOG	CAGAAGACCAGAACTGTG**GCG**TCTTCCACCCAGCTGTGT
F102A mNANOG	AGAAGATGCGGACTGTG**GCC**TCTCAGGCCCAGCTGTGTGC
Y119A hNANOG	AGATTTCAGAGACAGAAA**GCG**CTCAGCCTCCAGCAGATGCAA
Y119A mNANOG	AGGTTTCAGAAGCAGAAG**GCC**CTCAGCCTCCAGCAGATG
Y119A hNANOG (PiggyBac)	TTTCAGAGACAGAAA**GCC**CTCAGCCTCCAGCAG
L122A hNANOG	CAGAAATACCTCAGC**GCG**CAGCAGATGCAAGAACTCTCC
L122A mNANOG	AAGCAGAAGTACCTCAGC**GCC**CAGCAGATGCAAGAACTC
L122A hNANOG (PiggyBac)	CAGAAATACCTCAGC**GCC**CAGCAGATGCAAGAA
Q124A hNANOG	CAGAAATACCTCAGCCTCCAG**GCG**ATGCAAGAACTCTCCAAC
Q124A mNANOG	AAGTACCTCAGCCTCCAG**GCG**ATGCAAGAACTCTCCTCC
Q124A hNANOG (PiggyBac)	TACCTCAGCCTCCAG**GCG**ATGCAAGAACTCTCC
M125A hNANOG	TACCTCAGCCTCCAGCAG**GCG**CAAGAACTCTCCAACATCCTG
M125A mNANOG	TACCTCAGCCTCCAGCAG**GCG**CAAGAACTCTCCTCCATT
Y136A hNANOG	CTCTCCAACATCCTGAACCTCAGC**GCG**AAACAGGTGAAGACCTGGTTCCAGAACCAG
Y136A mNANOG	TCCATTCTGAACCTGAGC**GCT**AAGCAGGTTAAGACCTGG
K137A hNANOG	CTGAACCTCAGCTAC**GCG**CAGGTGAAGACCTGGTTC
K137A mNANOG	CCATTCTGAACCTGAGCTAT**GCG**CAGGTTAAGACCTGGTT
Q138A hNANOG	CTGAACCTCAGCTACAAA**GCG**GTGAAGACCTGGTTCCAG
Q138A mNANOG	TCTGAACCTGAGCTATAAG**GCG**GTTAAGACCTGGTTTCAA
K140A hNANOG	CTCAGCTACAAACAGGTG**GCG**ACCTGGTTCCAGAACCAGAGA
K140A mNANOG	CTGAGCTATAAGCAGGTT**GCG**ACCTGGTTTCAAAACCAA
T141A hNANOG	CTCAGCTACAAACAGGTGAAG**GCG**TGGTTCCAGAACCAGAGA
T141A mNANOG	CAGGTTAAGACCTGGTTT**GCA**AACCAAAGGATGAAGTGC
Q144A hNANOG	CAGGTGAAGACCTGGTTC**GCG**AACCAGAGAATGAAATCTAAGAGG
Q144A mNANOG	CAGGTTAAGACCTGGTTT**GCA**AACCAAAGGATGAAGTGC
N145A hNANOG	GTGAAGACCTGGTTCCAG**GCG**CAGAGAATGAAATCTAAGAGGTGG
N145A mNANOG	AGGTTAAGACCTGGTTTCAA**GCC**AAAGGATGAAGTGCAA
R147A hNANOG	ACCTGGTTCCAGAACCAG**GCG**ATGAAATCTAAGAGGTGGCAG
R147A mNANOG	ACCTGGTTTCAAAACCAA**GCG**ATGAAGTGCAAGCGGTGG
M148A hNANOG	GTGAAGACCTGGTTCCAGAACCAGAGA**GCG**AAATCTAAGAGGTGGCAG
M148A mNANOG	TGGTTTCAAAACCAAAGG**GCG**AAGTGCAAGCGGTGGCAG
K151A hNANOG	CAGAACCAGAGAATGAAATCT**GCG**AGGTGGCAGAAAAACAACTGG
K151A mNANOG	AACCAAAGGATGAAGTGC**GCG**CGGTGGCAGAAAAACCAG

Because reverse primers are just reverse complementary to the forward primer, only forward primer sequences are shown. Mutant bases are shown in boldface.