Identification of a putative transactivation domain in human Nanog

Jong-Hyun Oh[1], Hyun-Jin Do[2],
Heung-Mo Yang[2], Shin-Yong Moon[3],
Kwang-Yul Cha[1,2], Hyung-Min Chung[1,2]
and Jae-Hwan Kim[1,2,4]

[1]Graduate School of Life Science and Biotechnology
[2]Cell and Gene Therapy Research Institute
Pochon CHA University
[3]Stem Cell Research Center
Seoul 135-907, Korea
[4]Corresponding author: Tel, 82-2-3468-3195;
Fax, 82-2-3468-3373; E-mail, jaehwan_k@cha.ac.kr

Abstract

Nanog is a newly identified divergent homeodomain protein that directs the infinite propagation and sustains the pluripotency of embryonic stem cells. It has been reported that murine Nanog has two potent transactivation domains in N-terminal and C-terminal regions. Human Nanog (hNanog) polypeptide shares about 58% and 87% identity to the open reading frame and homeodomain of murine Nanog, respectively. However, the functional domains and molecular mechanisms of hNanog are poorly understood. In this study, for the first time, we presented that only C-terminus of hNanog contains a potent transactivation domain. Based on the amino acid sequences of homeobox domain, we roughly divided hNanog open reading frame into the three regions such as N-terminal, homeodomain and C-terminal regions and constructed either the fusion proteins between hNanog individual and Gal4 DNA binding domain or the context of native hNanog protein. Reporter assays by using reporter plasmid containing Gal4 or Nanog binding site revealed that the only C-terminal region exhibited the significant fold induction of transactivation. However, interestingly, there was no significant activation through N-terminal region unlike mu-rine Nanog, suggesting that C-terminal region may have more critical roles in the transcriptional activation of target genes. Taken together, the finding of a putative transactivation domain in hNanog may contribute to the further understanding of molecular mechanism on the regulation of downstream genes involved in self-renewal and pluripotency of human stem cells.

Key words: homeodomain; nanog; pluripotency; stem cell; transactivation domain; transcription factor

Introduction

The capability for self-renewal and the pluripotency of embryonic stem (ES) cells seem to be under the control of multiple transcriptional factors (Pesce and Scholer, 2001; Pan et al., 2002) in the presence of leukemia inhibitory factor (LIF). However, LIF is not sufficient to prevent differentiation and maintain the pluripotency of human ES cells (Niwa et al., 2000; Reubinoff et al., 2000; Cavaleri and Scholer, 2002). Recently, Nanog, a member of the homeobox family of DNA binding transcription factors, was identified as a possible candidate for the LIF/Stat3-independent factor for the pluripotency of ES cells. It does not only block endodermal differentiation but also maintain the pluripotency of ES cells even in the absence of LIF (Chambers et al., 2003; Mitsui et al., 2003). Nanog-deficient murine ES cells and embryos underwent extraembryonic endoderm differentiation (Mitsui et al., 2003). Gata4 and gata6, endoderm transcription factors, were upregulated in Nanog-null cells, suggesting that the self-renewal mechanism of Nanog may be maintained by transcriptional repression of downstream genes critical for differentiation (Mitsui et al., 2003). In the mean while, Nanog may also exert to maintain the pluripotency by activating ES cell specific gene(s), such as rex1 through binding to the consensus sequences in the enhancer region of rex1 (Ben-Shushan et al., 1998; Mitsui et al., 2003). However, little is known how Nanog acts as a transcriptional regulator on the downstream genes.

Recently, it has been reported that two distinct transactivational domains were identified in murine Nanog (Pan and Pei, 2003). Murine Nanog was divided into three regions based on the position of homeodomain. N-terminal region to the homeodomain is rich in Ser, Thr and Pro residues, a frequent characteristic of transcriptional activation domain (Rambaldi et al., 1994). C-terminal region to the homeodomain contains novel Trp (W) repeat motif of unknown

structure and function. They reported that both N- and C-terminal regions to the homeodomain in the murine Nanog have active transcriptional activities and that the C-terminal region unexpectedly seems to be more active than the N-terminal region (Pan and Pei, 2003). hNanog shares about 57% overall and 87% identity within the homeodomain to the murine Nanog. In addition, the Ser, Thr and Pro rich residues in the N-terminus and the W repeats in the C-terminus are conserved in both species, suggesting that the conserved motifs may also play critical roles in the transcriptional activity of hNanog. To examine the transcriptional activities of each region containing the conserved motifs, we constructed the fusion proteins between the hNanog individual domain and either Gal4 or Nanog homeobox DNA binding domain. In this study, for the first time, we report that only C-terminal region of hNanog possesses the potent transcriptional activity. Surprisingly, our study indicates that N-terminal region of hNanog does not exhibit the significant transcriptional activity, unlike murine Nanog.

Materials and Methods

Plasmid constructs

A full length hNanog isoform lacking Lys (aa168) residue (designated as hNanog 1a, unpublished data) was obtained by RT-PCR from NTERA-2 cl.D1 cDNA and cloned into a mammalian expression vector pCMV-Tag (Stratagene, La Jolla, CA). hNanog 1a was roughly divided into three regions based on the position of the homeodomain, as reported in murine Nanog (Pan and Pei, 2003). N-terminal (aa1-94), homeodomain (aa95-154) and C-terminal (aa155-304) regions of hNanog 1a were amplified by PCR using the full-length hNanog 1a cDNA as a template. To construct the fusion proteins between Gal4 DNA binding domain (aa1-147) and hNanog individual domains, the fragments of N-, C-terminal and homeodomain regions were inserted into multiple cloning site of pBIND vector (Promega, Madison, WI), and then fused in-frame to the C-terminus of Gal4 DNA binding domain. To construct the fusion proteins between Nanog homeobox DNA binding and hNanog individual domains, the expression vectors for N- or C-terminal regions fused to homeodomain as well as homeobox itself of hNanog 1a were prepared by inserting each region into the pcDNA3.1, and then a SV40 nuclear localization signal (NLS, PKKKRKV) and a Flag epitope tag were added to the N-terminus of each region. All PCR reactions were performed by using *pfu* DNA polymerase (Core Bio System, Seoul) and 15 pM each of the forward and reverse primers under the following conditions: 1 cycle at 95°C for 2 min followed by 35 cycles at 95°C for 30 sec, 55°C for 30 sec, 72°C for 2 min and a final extension for 10 min at 72°C. A minimal promoter construct, pGL3ti (Jonk *et al.,* 1998), was prepared from pGL3-basic (Promega) by inserting the adenovirus major late promoter TATA box and the mouse terminal de-

oxynucleotidyl transferase gene initiator sequence in the multi cloning site. Reporter construct bearing Nanog binding site was prepared by inserting oligonucleotides containing one copy of the Nanog binding site (ACCCTTCGCCGATTAAGTACTTAA; Pan and Pei, 2005) in the upstream of the minimal promoter within pGL3ti vector. The cloned PCR products and reporter plasmids were verified by sequencing.

Cell culture

Human embryonic carcinoma (hEC; NCCIT and NTERA-2 cl.D1), COS1, HeLa and 293T (American Type Cell Collection) cells were grown in Dulbecco's modified Eagle's medium (Life Technologies; Paisley, UK) supplemented with 10% fetal bovine serum (Hyclone, Logan, UT), and 50 IU penicillin and 50 µg/ml streptomycin.

Reporter assays

Cells were cotransfected with either 1 µg pG5*luc* vector (Promega) containing the 5 copies of Gal4 binding sites and 2 µg each Gal4 fusion plasmid or 1 µg Nanog binding site reporter plasmid and 2 µg each NLS/Flag-tagged expression vector by using ExGen500 *in vitro* transfection reagent (Fermentas, Hanover, MD) in 60-mm tissue culture dish. The luciferase activities from the reporter plasmids were normalized to the β-galactosidase activity of the cotransfected internal control plasmid pcDNA 3.1/hygro/ *LacZ* (Invitrogen). Cells were harvested 48 h after transfection and luciferase activities were measured by using firefly luciferase assay system (Promega). Transfection and reporter assay experiments were carried out in duplicates and independently repeated at least three times.

Western blotting

COS1 and 293T cells were transfected with each Gal4 fusion or NLS/Flag-tagged expression vector containing hNanog individual domain by using ExGen500 *in vitro* transfection reagent under the manufacture's protocol and harvested 48 h after transfection for Western blot analysis. Briefly, the harvested cells were washed with PBS and lysed with 200 µl of Pro-PREP protein extraction solution (iNtRON Biotech., Seoul) for 20 min on ice or -20°C. Whole cell lysates were collected by the centrifugation at 13,000 rpm for 25 min at 4°C and electrophoresed through 12% SDS-PAGE. The SDS-PAGE gels were transferred to PVDF membrane (Amersham Biosciences, Piscataway, NJ). The transferred membranes were blocked in 5% non-fat milk and incubated with anti-Gal4 monoclonal antibody (1:500, Santa Cruz, Santa Cruz, CA) or anti-Flag monoclonal antibody (1:10,000, Sigma, St. Louis, MO), followed by the incubation with alkaline phosphate-conjugated anti- mouse secondary antibody (1:1,000, Santa Cruz). The membranes were developed by Nitrotetrazolium blue chloride/5-Bromo-4-chloro-3-indolyl phosphate (Sigma) substrates.

Results and Discussion

To analyze the transcriptional activity of hNanog, we cloned full length hNanog isoform lacking Lys (aa168) residue, designated as hNanog 1a (Unpublished data), from hEC cells and the open reading frame was roughly divided into three regions as studied in murine Nanog (Pan and Pei, 2003): N-terminal (aa1-94), homeodomain (aa95-154), and C-terminal (aa155-304) regions based on the amino acid sequences of homeodomain and the individual domains were fused to the yeast transcription factor, Gal4 DNA binding domain (aa1-147) lacking the capability of transcriptional activation by itself (Figure 1A). The expression of the fusion protein of Gal4 DNA binding domain and hNanog individual domain plasmid constructs was confirmed by Western blot analysis with the anti-Gal4 DNA binding domain monoclonal antibody (Figure 1B). Western blot analysis revealed that the additional

fragment with larger molecular weight than we predicted was observed in Gal4/N-terminal fusion protein as in murine Nanog (Pan and Pei, 2003). To examine the transactivational potentials of hNanog individual regions in endogenously Nanog-expressing cells, NCCIT cells were transiently cotransfected with each fusion protein expression plasmid construct and reporter vector, pG5*luc*, containing 5 copies of Gal4 DNA binding sites in the upstream of the minimal promoter. As reported on murine Nanog (Pan and Pei, 2003), C-terminal region of hNanog drove a significant transcriptional activation, while N-terminal region did not exhibit the significant activation unlike murine N-terminal one (Figure 1C). We also performed the same Gal4 fusion protein/reporter assay in Nanog-nonexpressing cell lines, HeLa (Figure 1D), 293T and COS1 cells (data not shown), and observed the consistent transactivational potentials as in hEC cells, NCCIT.
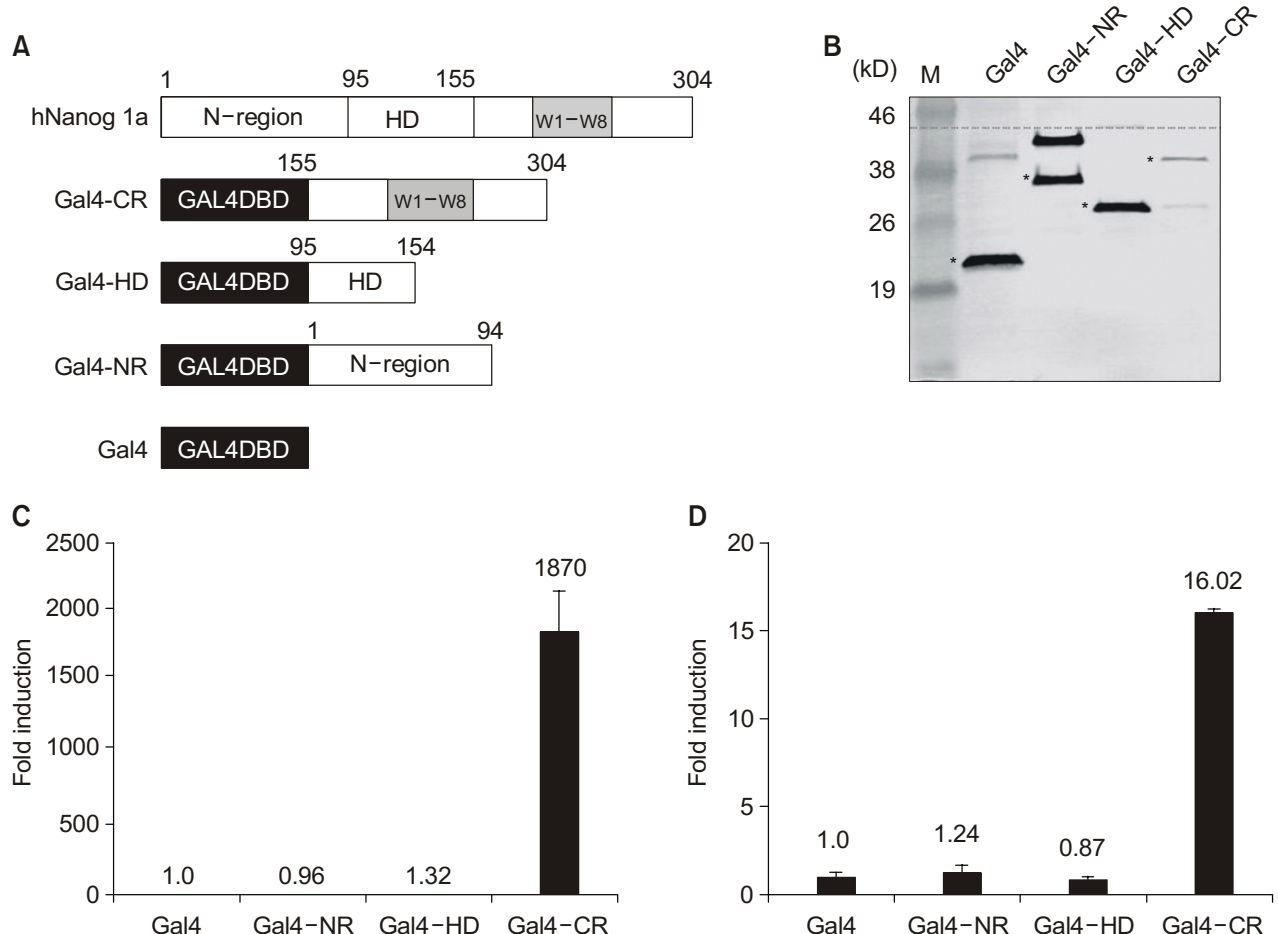

Figure 1. Identification of a putative transactivation domain in hNanog 1a. (A) Schematic representation of the C-terminal region (CR), N-terminal region (NR) or homeodomain (HD) fused to Gal4 DNA binding domain. (B) Western blot analysis of COS1 cells transfected with Gal4 DNA binding domain fusion constructs. The asterisks indicate the predicted position based on its calculated molecular weight. (C, D) Transactivation activities of fusion constructs in NCCIT, human embryonic carcinoma cells (C) and HeLa cells (D). The transcriptional activity of each construct was calculated relative to the expression of Luciferase reporter gene observed with pBIND, which was set to 1. Data shown are means and S.D. for duplicate measurements from one representative transfection.
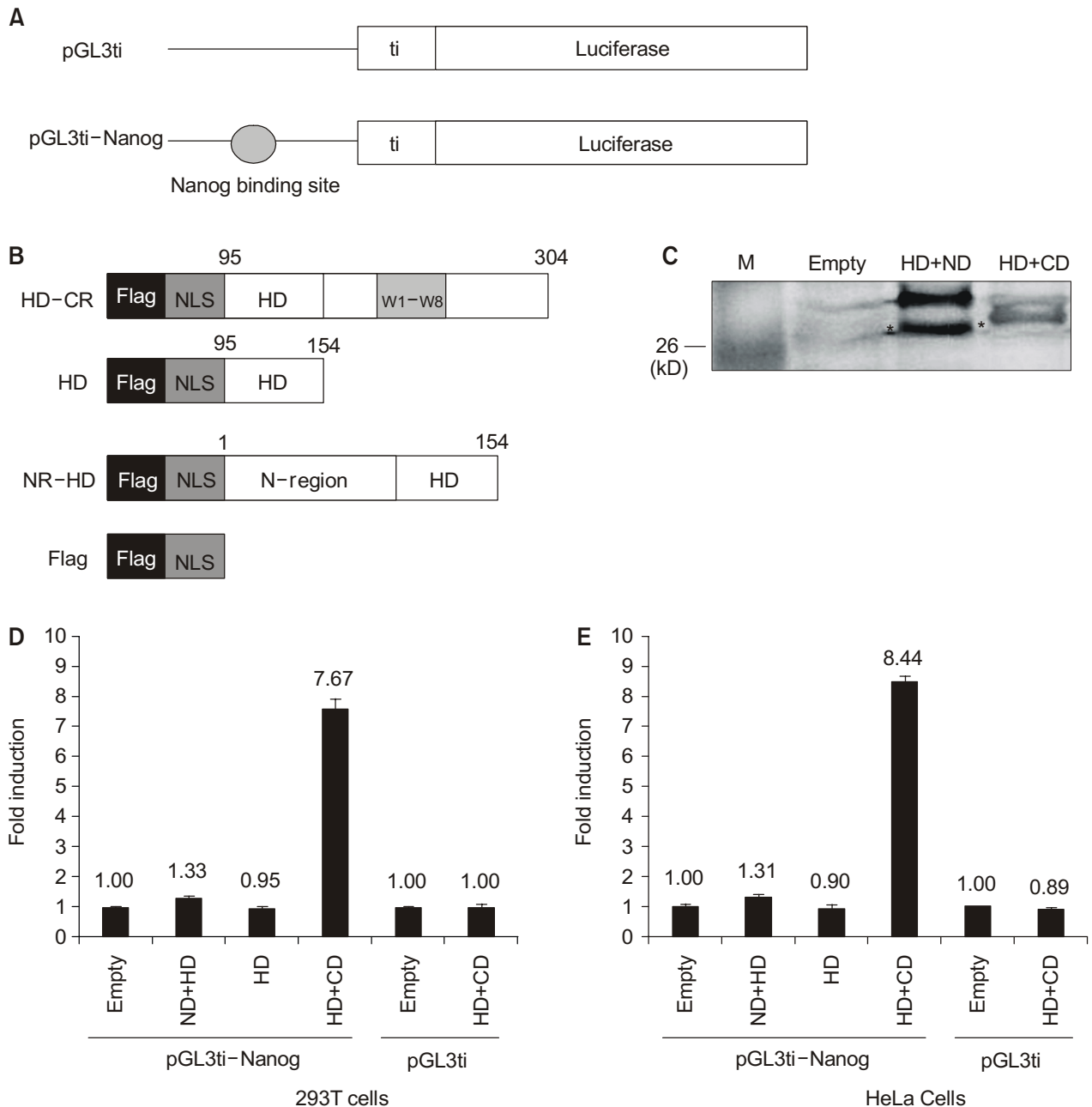
Figure 2. Transactivational activity of native hNanog 1a domains through Nanog binding site. (A) Schematic representation of pGL3ti and pGL3ti-Nanog binding site reporter gene constructs. (B) Schematic representation of the homeodomain (HD), N-terminal fused to HD (NR-HD), C-terminal regions fused to HD (HD-CR) inserted with a NLS and a Flag epitope. (C) Western blot analysis in 293T cells transfected with expression vectors for N-terminal + HD (NR-HD), HD + C-terminal (HD-CR) regions. The asterisks indicate the predicted position based on its calculated molecular weight. (D, E) Transactivation activities of each construct in 293T (D) and HeLa cells (E). Data are presented as fold activation relative to cells transfected with either pGL3ti or pGL3ti-Nanog binding site reporter gene construct. Data shown are means and S.D. for duplicate measurements from one representative transfection.

We next evaluated the transactivation properties of hNanog in the context of native hNanog protein instead of Gal4-fusion protein. For this study, NLS/Flag tagged expression vectors for homeodomain and N- or C-terminal regions fused to homeodomain of hNanog were prepared and the expression of fusion proteins was confirmed by the Western blot analysis with the anti-Flag monoclonal antibody (Figure 2C). Western blot analysis also revealed the additional fragment with larger molecular weight than we pre-

dicted was observed in hNanog N-terminal region fused to homeodomain, as shown in Gal4/N-terminal fusion protein (Figure 1B). To examine the transactivational potentials of native hNanog individual regions, the expression vectors were cotransfected with pGL3ti reporter plasmid containing the one copy of Nanog binding site in 293T and HeLa cells (Figure 2D and 2E). As a control, the reporter construct without the Nanog binding site was used. Reporter assay revealed that C-terminal region of hNanog drove the significant transcriptional activation, but N-terminal region did not exhibit the significant activation of the reporter, as shown in Figure 1.

This study presented, for the first time that hNanog possesses the potent transactivaion activity only in the C-terminal region. In the recent studies on murine Nanog (Pan and Pei, 2003; 2005), they reported that both N- and C-terminal regions to the homeodomain have transcriptional activities, suggesting that the W repeats and the last part of 58 amino acid residues in C-terminal, and Ser and Thr residues in N-terminal regions may play a critical role for transactivation activity. Human Nanog shares about 45% identity to the N- and C-terminal regions of murine Nanog. Even though the Ser and Thr rich region in N-terminus and the W repeats in C-terminus are conserved in both species, our data suggest that in hNanog, the C-terminal region containing eight W repeats motif at every 5th position of amino acid sequence 196-236 seems to play more critical roles in the transcriptional activation on the downstream target genes. Although N-terminal region in hNanog didn't show the significant transactivational activity, Western blot analysis (Figure 1B and 2C) suggests that some of the Ser or/and Thr residues in N-terminal region may undergo post-translational modifications such as phosphorylation, possibly affecting the transactivational activity of the C-terminal region in hNanog. However, more functional studies on the individual domains are required to identify the unknown regulation mechanism on transactivational activity of hNanog.

Recent study (Hart *et al.*, 2004) on the identification of hNanog2 strengthens the importance of C-terminal region in hNanog for transactivational activity. Human *Nanog2* cloned from hES cells has the similar structure to hNanog but possesses a much longer first intron and encodes a protein of only 233 amino acids with the partial truncation of N-terminal region. Although the possibility for post-translational modifications in partial N-terminal region for transactivational activity can not be excluded, the existence of the N-terminal truncated hNanog 2 suggests that C-terminal region may be evolutionarily more critical for transcriptional activation, even though the hNanog2 is still functionally unknown. Moreover, N-terminal region may provide a structurally important motif for the transcriptional activation of hNanog by tight regulation through phosphorylation or other modifications.

In conclusion, we identified a potent transactivation domain in hNanog and might provide a new tool to understand the molecular mechanism on the regulation of downstream target genes involved in the self-renewal and pluripotency of human stem cells.

References

Ben-Shushan E, Thompson JR, Gudas LJ, Bergman Y. Rex-1, a gene encoding a transcription factor expressed in the early embryo, is regulated via Oct-3/4 and Oct-6 binding to an octamer site and a novel protein, Rox-1, binding to an adjacent site. Mol Cell Biol 1998;18:1866-78

Cavaleri F, Scholer HR. Nanog: a new recruit to the embryonic stem cell orchestra. Cell 2002;113:551-2

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. Functional expression cloning of nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 2003;113:643-55

Hart AH, Hartley L, Ibrahim M, Robb L. Identification, cloning and expression analysis of the pluripotency promoting nanog genes in mouse and human. Dev Dyn 2004;230:187-98

Jonk LJ, Itoh S, Heldin CH, ten Dijke P, Kruijer W. Identification and functional characterization of a Smad binding element (SBE) in the JunB promoter that acts as a transforming growth factor-β, activin, and bone morphogenetic protein-inducible enhancer. J Biol Chem 1998;273:21145-52.

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. The homeoprotein nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 2003;113: 631-42

Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet 2000;24:372-6

Pan G, Pei D. The stem cell pluripotency factor nanog activates transcription with two unusually potent subdomains at its C-terminus. J Biol Chem 2005;280:1401-7

Pan GJ, Chang ZY, Scholer HR, Pei D. Stem cell pluripotency and transcription factor Oct4. Cell Res 2002;12:321-9

Pan GJ, Pei DQ. Identification of two distinct transactivation domains in the pluripotency sustaining factor nanog. Cell Res 2003;13:499-502

Pesce M, Scholer HR. Oct-4: gatekeeper in the beginnings of mammalian development. Stem Cells 2001;19:271-8

Rambaldi I, Kovacs EN, Featherstone MS. A proline-rich transcriptional activation domain in murine HOXD-4 (HOX-4.2). Nucleic Acids Res 1994;22:376-82

Reubinoff BE, Pera MF, Fong CY, Trounson A, Bongso A. Embryonic stem cell lines from human blastocysts: somatic differentiation *in vitro*. Nat Biotechnol 2000;18:399-404