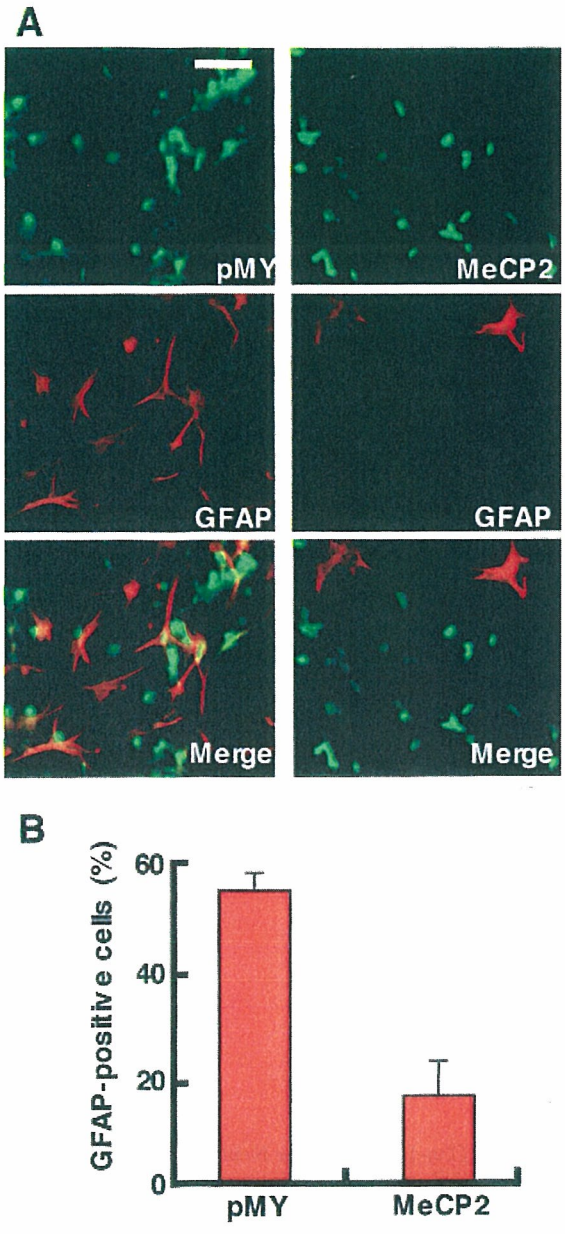

Fig. 3. MeCP2 suppresses GFAP expression. **A:** Four-day-expanded E14.5 neuroepithelial cells were infected with recombinant retrovirus engineered to express only GFP (pMY), or MeCP2 together with GFP (MeCP2), and cultured with LIF (50 ng/ml) for 4 days to induce astrocyte differentiation. The cells were then stained with antibodies against GFP (green) and GFAP (red). **B:** The percentage of GFAP-positive cells in each GFP-positive cell population after 4 days' culture was determined (mean ± SD). Scale bar = 50 μm.

from differentiating into GFAP-positive cells (16% in Fig. 3B). In addition, there was no significant difference in the number of dead cells between control and MeCP2-expressing virus-infected cells (Supplemental Fig. 2), indicating that the reduction in the number of GFAP-positive cells caused by MeCP2 expression was not due to the specific killing of astrocytes. Furthermore, the proliferation rate of control and MeCP2-expressing virus-infected cells was similar when judged by bromodeoxyuiridine (BrdU) uptake (data not shown). Neither selective cell death nor inhibition of proliferation, therefore, appears to account for the reduction of astrocyte differentiation. These observations suggest that MeCP2 functions as a suppressor of astrocytic gene expression.

MeCP2 Binds to the Hypermethylated GFAP Exon 1

MeCP2 is known to suppress gene expression by binding to methylated CpG dinucleotides via a DNA-binding motif called the *methyl-cytosine binding domain*. Accordingly, if MeCP2 inhibits GFAP gene expression, the gene should possess a methylated region. As mentioned above, CpG sites in the GFAP promoter region between −1,622 bp and −1,423 bp, including the STAT3-binding site, are hypomethylated both in neuroepithelial cells at late gestation and in differentiated cell populations containing many neurons. In an effort to identify highly methylated regions of the GFAP gene other than the STAT3 site-containing region, we found that the exon 1 region (+20 to +449 bp) was heavily methylated in E14.5 neuroepithelial cells as well as in neuron-containing differentiated cells (Fig. 4A). These results raised the possibility that exon 1 could be a binding target for MeCP2 in neurons.

To examine whether MeCP2 binds to hypermethylated exon 1 of the GFAP gene in neurons, we performed a ChIP assay using anti-MeCP2 antibody followed by PCR to detect coimmunoprecipitated exon 1 in undifferentiated neuroepithelial cells and neuron-containing differentiated cells. As shown in Figure 4B, the level of MeCP2 binding to exon 1 was significant in neuron-containing differentiated cells but was much lower in undifferentiated neuroepithelial cells. We also confirmed that MeCP2 could bind to hypermethylated exon 1 in undifferentiated neuroepithelial cells when expressed ectopically (data not shown). It seems likely from these experiments that the expression of MeCP2 in neurons contributes to the silencing of the GFAP gene as a consequence of MeCP2 binding to the hypermethylated exon1 region of the gene.

MeCP2 Binds to the Hypermethylated Proximal Region of the S100β Gene To Suppress Its Expression

S100β is a soluble calcium-binding protein synthesized in astrocytes and, like GFAP, is known as a typical marker for astrocytes (Burette et al., 1998). The results decsribed above showed that MeCP2 has the potential to suppress an astrocytic marker, GFAP expression; however, they did not reveal whether the MeCP2-mediated suppression is confined only to the GFAP gene or might extend to other astrocyte-specific genes such as S100β. Neuroepithelial cells derived from E14.5 telencephalon were infected with control and MeCP2-expressing retro-

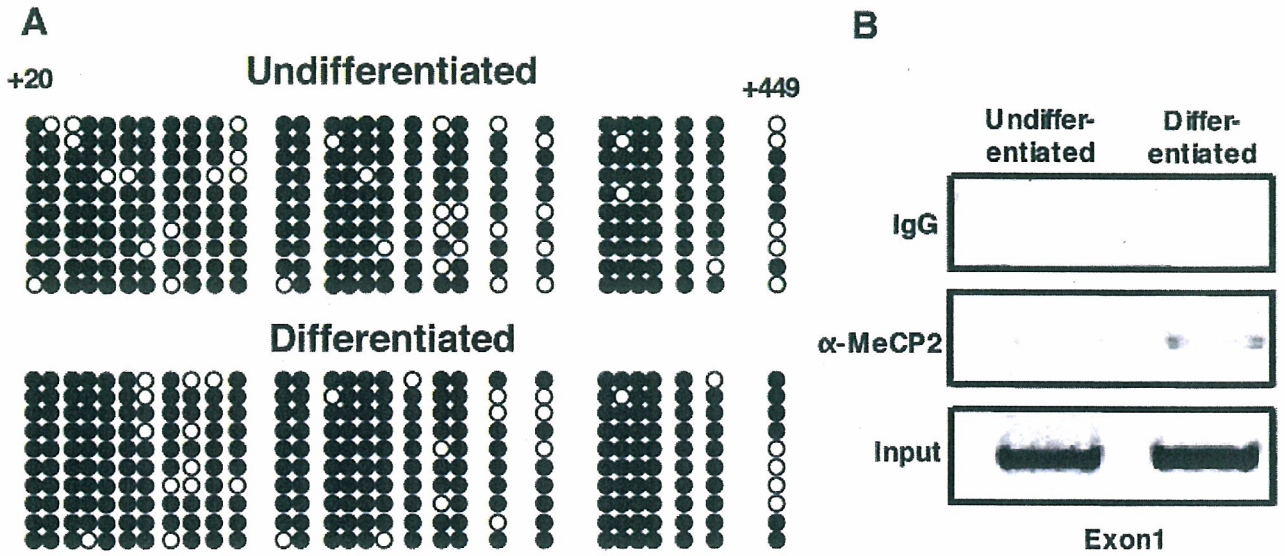

Fig. 4. MeCP2 binds to the hypermethylated exon 1 of GFAP gene in neurons. **A:** Four-day-expanded E14.5 neuroepithelial cells (undifferentiated) were differentiated in medium containing 0.5% FBS for 4 days (differentiated). The methylation status of CpG sites within the GFAP gene exon 1 region (+20 to +449 bp) in undifferentiated neuroepithelial cells (upper) and differentiated cells (lower) was analyzed by bisulfite sequencing. White and black circles indicate unmethylated and methylated CpG sites, respectively. **B:** ChIP assays were performed using anti-MeCP2 antibody and PCR primers to detect a DNA fragment spanning −18 and +510 bp in the exon1 region of GFAP gene in undifferentiated neuroepithelial cells (left lanes) and differentiated cells (right lanes).

viruses and cultured for 4 days under conditions that favor astrocyte differentiation. We found that 59% and 3% of the virus-infected cells were S100β positive in the control and MeCP2-expressing populations, respectively (Fig. 5A,B). These indicate that MeCP2 can dramatically inhibit S100β expression as well as GFAP expression.

To address the question of whether the suppression of S100β gene expression by MeCP2 occurs through a mechanism analogous to that of the GFAP gene, we next analyzed the methylation status of CpG sites around the transcriptional initiation site, including exon 1, of the S100β gene in undifferentiated neuroepithelial cells and neuron-containing differentiated cells. There are four CpG sites, at −322, −207, −64, and +7 bp relative to the transcriptional start site. As shown in Figure 5C, although the methylation frequency of the CpG dinucleotide at −322 bp was very low, it was high at the other three CpG sites in both undifferentiated and differentiated cells. These three CpG sites may thus be targets for MeCP2 binding to repress gene expression. To examine whether MeCP2 can associate with these methylated CpG sites, ChIP analysis was performed with an anti-MeCP2 antibody followed by PCR to detect a coimmunoprecipitated DNA fragment spanning −234 and +84 bp in the S100β gene. As shown in Figure 5D, a strong association between MeCP2 and the hypermethylated region of the S100β gene was observed in cells expressing MeCP2. These data reveal that MeCP2 can bind to the highly methylated region of the S100β gene as well as to that of the GFAP gene to suppress expression of both genes. Collectively, these results further support the argument that MeCP2 restricts astrocytic lineage-specific gene expression in neurons expressing higher amounts of MeCP2.

MBD1 and MeCP2 Have a Redundant Function Regarding the Suppression of Astrocytic Gene Expression

MeCP2 belongs to a protein family composed of at least five members, all of which share a functional methyl-CpG-binding domain, except for MBD3, which is incapable of binding to methylated DNA (Hendrich and Bird, 1998). The other methyl-CpG binding proteins may thus have a functional redundancy with MeCP2 in inhibiting astrocytic gene expression. In particular, it has been reported that MBD1 is expressed in the brain and plays an important role in adult neurogenesis and hippocampal function in mouse (Zhao et al., 2003). Therefore, we decided to determine whether MBD1 can suppress GFAP expression. To this end, we infected E14.5 neuroepithelial cells with control and MBD1-expressing retroviruses and cultured them with LIF for 4 days to induce astrocyte differentiation. As shown in Figure 6A, MBD1 expression resulted in the inhibition of GFAP-positive astrocyte differentiation of neuroepithelial cells. The control retrovirus-infected cells became GFAP positive astrocytes at a frequency of 54% (Fig. 6A), whereas only 24% of MBD1-infected cells were GFAP positive (Fig. 6B). These data imply that MeCP2 and MBD1, both of which are highly expressed

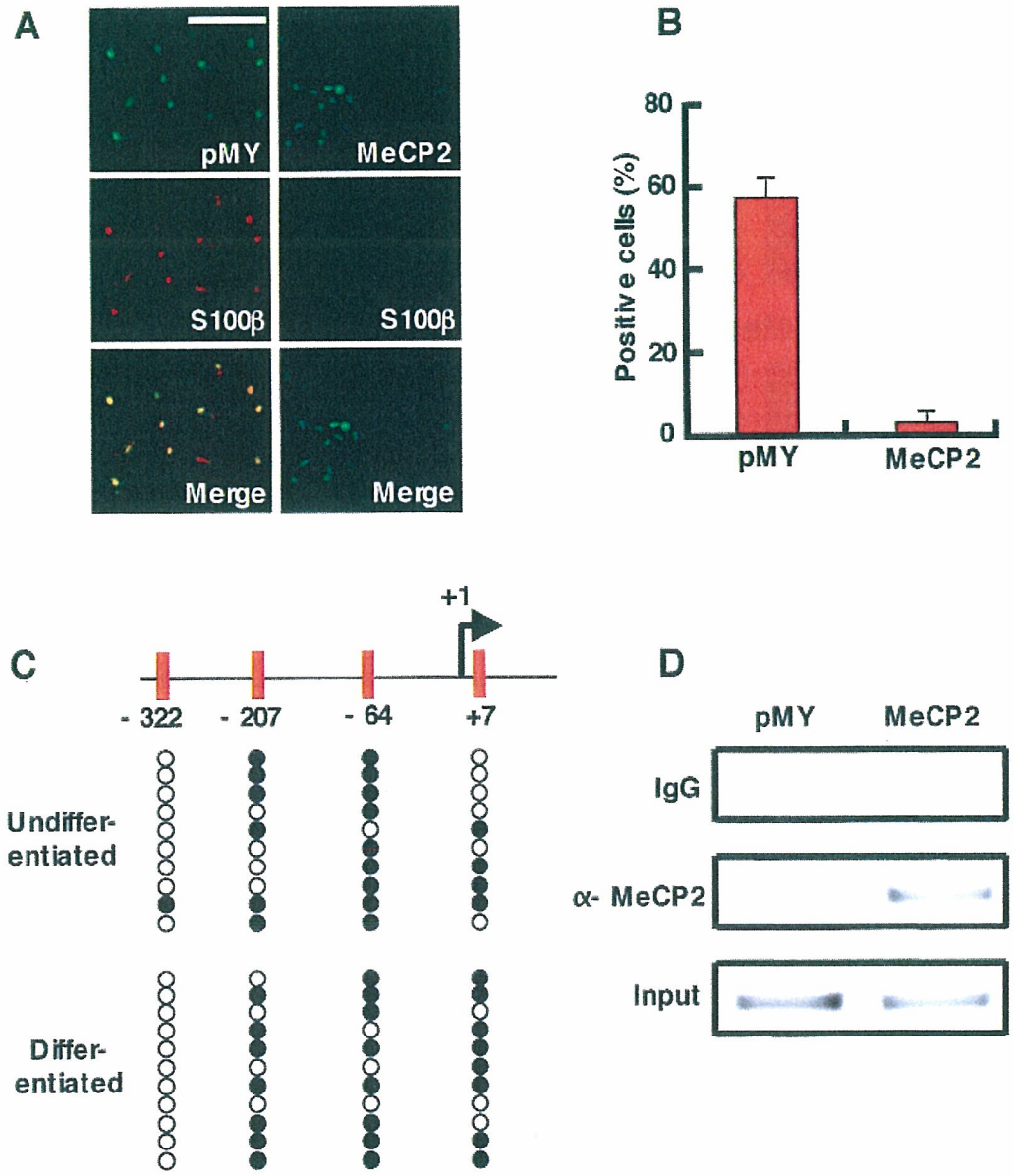

Fig. 5. MeCP2 binds to the hypermethylated proximal region of the S100β gene to suppress its expression. **A:** Neuroepithelial cells were infected with recombinant retrovirus engineered to express only GFP (pMY), or MeCP2 together with GFP (MeCP2), and cultured with LIF (50 ng/ml) for 4 days to induce astrocyte differentiation. The cells were then stained with antibodies against GFP (green) and S100β (red). **B:** The percentage of S100β-positive cells within the GFP-positive cell population after 4 days' culture was determined (mean ± SD). **C:** The methylation status of four CpG sites (−322, −207, −64, and +7 bp) around the S100β transcriptional initiation site (+1 bp) including exon 1, in undifferentiated neuroepithelial cells and differentiated cells, was analyzed by bisulfite sequencing. White and black circles indicate unmethylated and methylated CpG sites, respectively. Nucleotide position are designated base on GeneBank accession number L22144. **D:** ChIP assays were performed using anti-MeCP2 antibody and PCR primers to detect a DNA fragment spanning −234 and +84 bp in the S100β gene in control (pMY) and MeCP2-expressing virus-infected neuroepithelial cells. Scale bar = 50 μm.

in neurons, are functionally redundant and suppress astrocytic gene expression in neurons.

DISCUSSION

It is becoming apparent that interplay between cell-extrinsic cues and cell-intrinsic programs has important roles in the fate specification of NSCs. We have previously reported that the change in DNA methylation of a cell-type-specific gene promoter is a critical determinant for the switch from neurogenesis to astrocytogenesis of NSCs during development (Takizawa et al., 2001). We propose here that not only DNA methylation

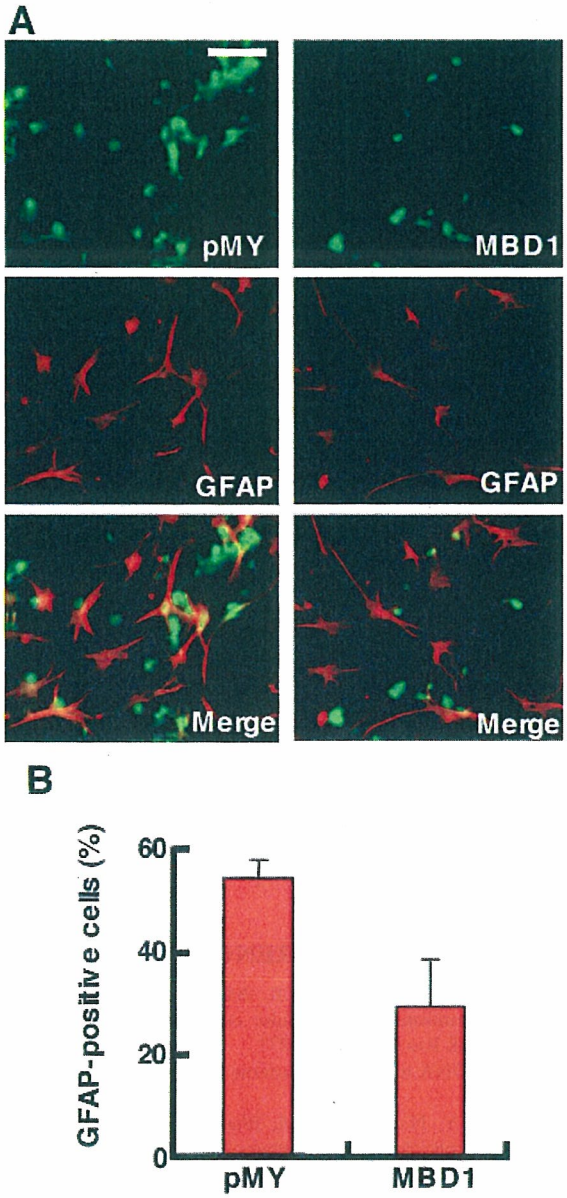

Fig. 6. MBD1 inhibits GFAP expression in neuroepithelial cells. **A:** Neuroepithelial cells were infected with recombinant retrovirus engineered to express only GFP (pMY), or MBD1 together with GFP (MBD1), and the cells were cultured with LIF (50 ng/ml) for 4 days to induce astrocyte differentiation. The cells were then stained with antibodies against GFP (green) and GFAP (red). **B:** The percentage of GFAP-positive cells within the GFP-positive cell population after 4 days' culture was determined (mean ± SD).

but also the expression of MBDs is involved in the maintenance of neuronal identity.

Neurons differentiated from neuroepithelial cells at midgestation do not express an astrocyte-specific gene GFAP, even when stimulated with the astrocyte-inducing cytokine LIF. However, this is not surprising, in that the hypermethylated status of the STAT3-binding site in the GFAP gene promoter is maintained (Takizawa et al., 2001). As gestation proceeds, neuroepithelial cells gradually lose methylation at the STAT3-binding site and gain the potential to differentiate into GFAP-expressing astrocytes. Nevertheless, these neuroepithelial cells can still give rise to neurons, and these neurons do not express GFAP even if they are stimulated by the STAT3-activating cytokine LIF. These findings suggested to us that a mechanism other than STAT3 site methylation must exist to suppress astrocytic genes in neurons.

DNA methylation-mediated gene silencing occurs mainly through the following two mechanisms. First, CpG methylation within a transcription factor recognition sequence can interfere directly with binding of the transcription factor to the sequence (Watt and Molloy, 1988). Second, and more generally, methylation-mediated gene silencing can be executed through the action of MBD family members such as MeCP2 and MBD1, which preferentially bind to methylated CpG (Hendrich and Bird, 1998). The MBDs themselves are transcriptional repressors whose association with other corepressor proteins and histone-modifying enzymes leads to repressive chromatin remodeling and gene silencing (Hendrich and Bird, 1998). Among the MBDs, at least MeCP2, MBD1, MBD2, and MBD3 are known to be expressed predominantly in neurons in the CNS (Coy et al., 1999; Jung et al., 2002; Shahbazian et al., 2002; Kishi and Macklis, 2004). We found in the present study that exon 1 of the GFAP gene is heavily methylated even in neuroepithelial cells at late gestation and in neurons differentiated from these neuroepithelial cells. Likewise, the proximal promoter region including exon 1 of another astrocytic marker gene, S100β, is highly methylated in both types of cells. We further demonstrated that MeCP2 actually binds to these highly methylated regions and suppresses expression of both GFAP and S100β genes. Similarly, MBD1 also inhibited GFAP expression in neuroepithelial cells, which normally respond to LIF stimulation and express the gene. Collectively, these observations suggest that MBDs participate in the suppression of astrocytic gene expression in neurons expressing abundant amounts of MBDs. In this context, it is intriguing that neurons can be divided into two groups: neurons with a hypermethylated STAT3-binding site in the GFAP gene promoter and those without it. However, the functional difference between these two neuronal types is currently unknown.

In this report, we have suggested the involvement of MBDs, including MeCP2, in the suppression of GFAP gene expression in neurons. Thus, a dysfunction of MeCP2 or other MBDs in the brain might affect GFAP expression. Indeed, it has been reported that many glial gene transcripts, including GFAP, related to known neuropathological mechanisms are up-regulated in the brains of RTT patients (Colantuoni et al., 2001). In addition, GFAP is the major intermediate filament of mature astrocytes, and its specific expression in this cell type suggests an important function (Messing and Brenner, 2003). While the absence of GFAP has relatively

minor effects on development, its overexpression leads to an increase in lethality in transgenic mice (Messing et al., 1998). Furthermore, GFAP-coding mutations are known to be responsible for most cases of Alexander disease, a devastating neurodegenerative disorder (Messing et al., 2001). Therefore, the precise control of GFAP expression level modulated by MBDs may be necessary for normal development and function of the CNS. In contrast to reports on the human case of RTT, Tudor et al. (2002) have reported that the results of numerous microarray analyses revealed no dramatic changes in transcription even in MeCP2 mutant mice displaying overt disease symptoms; statistical analyses of the data indicated that even relatively subtle changes in transcription would have been detected if present. We also examined whether GFAP is expressed in neurons prepared from MeCP2-deficient mice at postnatal day 2 and found no obvious GFAP expression in neurons as assessed by immunocytochemistry (data not shown). These observations might reflect either a difference in sensitivity of humans and mice to the loss of MeCP2 (note that male mutant mice are viable, whereas it is thought that many, although not all, human males with MeCP2 mutations die perinatally; Hagberg et al., 1983; Dotti et al., 2002), or functional compensation by another methyl-binding protein, such as MBD1, that has similar and possibly redundant functions, as suggested in this work. Whatever the reason, we need to await further studies to unravel the cause of the discrepancy between MeCP2 mutant mice and human RTT patients. Another group has shown recently that MeCP2 is expressed in embryonic nonneuronal cells, including astrocytes, as judged by immunocytochemistry and Western blotting (Nagai et al., 2005). They further indicated that knockdown of MeCP2 in astrocytes reduces their proliferation. Conversely, we could not detect MeCP2 expression in astrocytes in our experiments. Although the reason for this inconsistency is unclear, it may be attributable to differences in sensitivity between the two antibodies used to detect MeCP2.

Kishi and Macklis (2004) have reported that MeCP2 is not involved in fate determination of neuroepithelial cells as judged from both in vivo and in vitro analysis of MeCP2 mutant mice. They showed that MeCP2 is abundantly expressed in neurons but that expression is very low, if any exists, in undifferentiated neuroepithelial cells, so that the absence of MeCP2 might not affect the fate decision of the cells. Neuroepithelial cells at late gestation express astrocytic genes in response to astrocyte-inducing cytokines, a property that makes them suitable to examine whether MeCP2 suppresses astrocytic gene expression. In the present study, we demonstrated that astrocytic gene expression was inhibited when E14.5 neuroepithelial cells were transduced by MeCP2-expressing retroviruses. This result suggests that MeCP2 is involved in the suppression of astrocytic gene expression in cells such as neurons, which abundantly express MeCP2. It does not, however, necessarily mean that MeCP2 physiologically controls fate specification of neuroepithelial cells.

We have argued in this paper that MBDs inhibit astrocyte-specific gene expression in neurons, but other interpretations remain possible. It has been suggested that neurogenin1 (Ngn1), a proneural basic helix-loop-helix transcriptional factor, can inhibit transcription of GFAP by sequestering the CBP-Smad1 complex away from astrocyte-specific genes and by inhibiting the activation of STATs, which are necessary for astrocytogenesis (Sun et al., 2001). We have also detected Ngn1 expression in mature neurons in the adult mouse brain (J.K. and K.N., unpublished observations). Therefore, it is possible that proneural gene products such as Ngn1 function as transcriptional suppressors for astrocytic genes, in concert with MBDs. In any case, the present study suggests that the binding of MBDs to highly methylated regions of their target genes may be a key molecular mechanism restricting the differentiation plasticity of neurons.

ACKNOWLEDGMENTS

We thank Dr. X. Zhao (University of New Mexico) for rat MBD1 cDNA, Dr. T. Kitamura (Tokyo University) for pMY vector and Plat-E cells, and Drs. Y. Bessho and T. Matsui for valuable discussions. We also thank Dr. I. Smith for helpful comments and critical reading of the manuscript. We are very grateful to N. Ueda for her excellent secretarial assistance and to N. Namihira for technical help.

REFERENCES

Abematsu M, Smith I, Nakashima K. 2006. Mechanisms of neural stem cell fate determination: extracellular cues and intracellular programs. Curr Stem Cell Res Ther 1:267–277.

Bird A. 2002. DNA methylation patterns and epigenetic memory. Genes Dev 16:6–21.

Bonni A, Sun Y, Nadal-Vicens M, Bhatt A, Frank DA, Rozovsky I, Stahl N, Yancopoulos GD, Greenberg ME. 1997. Regulation of gliogenesis in the central nervous system by the JAK-STAT signaling pathway. Science 278:477–483.

Bugga L, Gadient RA, Kwan K, Stewart CL, Patterson PH. 1998. Analysis of neuronal and glial phenotypes in brains of mice deficient in leukemia inhibitory factor. J Neurobiol 36:509–524.

Burette A, Jalenques I, Romand R. 1998. Developmental distribution of astrocytic proteins in the rat cochlear nucleus. Brain Res Dev Brain Res 107:179–189.

Clark SJ, Harrison J, Paul CL, Frommer M. 1994. High sensitivity mapping of methylated cytosines. Nucleic Acids Res 22:2990–2997.

Colantuoni C, Jeon OH, Hyder K, Chenchik A, Khimani AH, Narayanan V, Hoffman EP, Kaufmann WE, Naidu S, Pevsner J. 2001. Gene expression profiling in postmortem Rett Syndrome brain: differential gene expression and patient classification. Neurobiol Dis 8:847–865.

Coy JF, Sedlacek Z, Bachner D, Delius H, Poustka A. 1999. A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3′-untranslated region of the methyl-CpG-binding protein 2 gene (MeCP2) suggests a regulatory role in gene expression. Hum Mol Genet 8:1253–1262.

Doetsch F. 2003. A niche for adult neural stem cells. Curr Opin Genet Dev 13:543–550.

Dotti MT, Orrico A, De Stefano N, Battisti C, Sicurelli F, Severi S, Lam CW, Galli L, Sorrentino V, Federico A. 2002. A Rett syndrome MECP2 mutation that causes mental retardation in men. Neurology 58:226–230.

Edlund T, Jessell TM. 1999. Progression from extrinsic to intrinsic signaling in cell fate specification: a view from the nervous system. Cell 96: 211–224.

Endres M, Meisel A, Biniszkiewicz D, Namura S, Prass K, Ruscher K, Lipski A, Jaenisch R, Moskowitz MA, Dirnagl U. 2000. DNA methyltransferase contributes to delayed ischemic brain injury. J Neurosci 20: 3175–3181.

Hagberg B, Aicardi J, Dias K, Ramos O. 1983. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. Ann Neurol 14:471–479.

He F, Ge W, Martinowich K, Becker-Catania S, Coskun V, Zhu W, Wu H, Castro D, Guillemot F, Fan G, de Vellis J, Sun YE. 2005. A positive autoregulatory loop of Jak-STAT signaling controls the onset of astrogliogenesis. Nat Neurosci 8:616–625.

Hendrich B, Bird A. 1998. Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol 18:6538–6547.

Hsieh J, Gage FH. 2004. Epigenetic control of neural stem cell fate. Curr Opin Genet Dev 14:461–469.

Issa JP, Ottaviano YL, Celano P, Hamilton SR, Davidson NE, Baylin SB. 1994. Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. Nat Genet 7:536–540.

Jones PA, Baylin SB. 2002. The fundamental role of epigenetic events in cancer. Nat Rev Genet 3:415–428.

Jung BP, Zhang G, Ho W, Francis J, Eubanks JH. 2002. Transient forebrain ischemia alters the mRNA expression of methyl DNA-binding factors in the adult rat hippocampus. Neuroscience 115:515–524.

Kishi N, Macklis JD. 2004. MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. Mol Cell Neurosci 27:306–321.

Koblar SA, Turnley AM, Classon BJ, Reid KL, Ware CB, Cheema SS, Murphy M, Bartlett PF. 1998. Neural precursor differentiation into astrocytes requires signaling through the leukemia inhibitory factor receptor. Proc Natl Acad Sci U S A 95:3178–3181.

Kriaucionis S, Bird A. 2003. DNA methylation and Rett syndrome. Hum Mol Genet 12:R221–R227.

Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69:915–926.

Messing A, Brenner M. 2003. GFAP: functional implications gleaned from studies of genetically engineered mice. Glia 43:87–90.

Messing A, Head MW, Galles K, Galbreath EJ, Goldman JE, Brenner M. 1998. Fatal encephalopathy with astrocyte inclusions in GFAP transgenic mice. Am J Pathol 152:391–398.

Messing A, Goldman JE, Johnson AB, Brenner M. 2001. Alexander disease: new insights from genetics. J Neuropathol Exp Neurol 60:563–573.

Morita S, Kojima T, Kitamura T. 2000. Plat-E: an efficient and stable system for transient packaging of retroviruses. Gene Ther 7:1063–1066.

Nagai K, Miyake K, Kubota T. 2005. A transcriptional repressor MeCP2 causing Rett syndrome is expressed in embryonic non-neuronal cells and controls their growth. Brain Res Dev Brain Res 157:103–106.

Nakashima K, Wiese S, Yanagisawa M, Arakawa H, Kimura N, Hisatsune T, Yoshida K, Kishimoto T, Sendtner M, Taga T. 1999a. Developmental requirement of gp130 signaling in neuronal survival and astrocyte differentiation. J Neurosci 19:5429–5434.

Nakashima K, Yanagisawa M, Arakawa H, Kimura N, Hisatsune T, Kawabata M, Miyazono K, Taga T. 1999b. Synergistic signaling in fetal brain by STAT3-Smad1 complex bridged by p300. Science 284:479–482.

Qian X, Shen Q, Goderie SK, He W, Capela A, Davis AA, Temple S. 2000. Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. Neuron 28:69–80.

Rajan P, McKay RD. 1998. Multiple routes to astrocytic differentiation in the CNS. J Neurosci 18:3620–3629.

Robertson KD, Wolffe AP. 2000. DNA methylation in health and disease. Nat Rev Genet 1:11–19.

Sauvageot CM, Stiles CD. 2002. Molecular mechanisms controlling cortical gliogenesis. Curr Opin Neurobiol 12:244–249.

Shahbazian MD, Antalffy B, Armstrong DL, Zoghbi HY. 2002. Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. Hum Mol Genet 11:115–124.

Sun Y, Nadal-Vicens M, Misono S, Lin MZ, Zubiaga A, Hua X, Fan G, Greenberg ME. 2001. Neurogenin promotes neurogenesis and inhibits glial differentiation by independent mechanisms. Cell 104:365–376.

Takizawa T, Nakashima K, Namihira M, Ochiai W, Uemura A, Yanagisawa M, Fujita N, Nakao M, Taga T. 2001. DNA methylation is a critical cell-intrinsic determinant of astrocyte differentiation in the fetal brain. Dev Cell 1:749–758.

Tudor M, Akbarian S, Chen RZ, Jaenisch R. 2002. Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. Proc Natl Acad Sci U S A 99:15536–15541.

Watt F, Molloy PL. 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev 2:1136–1143.

Zhao X, Ueba T, Christie BR, Barkho B, McConnell MJ, Nakashima K, Lein ES, Eadie BD, Willhoite AR, Muotri AR, Summers RG, Chun J, Lee KF, Gage FH. 2003. Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. Proc Natl Acad Sci U S A 100:6777–6782.

Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells

Atsushi Suzuki*†‡, Ángel Raya*†§, Yasuhiko Kawakami*, Masanobu Morita*†, Takaaki Matsui*¶, Kinichi Nakashima‖**, Fred H. Gage†‖, Concepción Rodríguez-Esteban*, and Juan Carlos Izpisúa Belmonte*†§††

*Gene Expression Laboratory, †Stem Cell Research Center, and ‖Laboratory of Genetics, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037; §Institució Catalana de Recerca i Estudis Avançats and Center of Regenerative Medicine, Dr. Aiguader 80, 08003 Barcelona, Spain; and **Laboratory of Molecular Neuroscience, Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma 630-0101, Japan

Edited by Irving L. Weissman, Stanford University School of Medicine, Stanford, CA, and approved May 19, 2006 (received for review August 11, 2005)

ES cells represent a valuable model for investigating early embryo development and hold promise for future regenerative medicine strategies. The self-renewal of pluripotent mouse ES cells has been shown to require extrinsic stimulation by the bone morphogenetic protein (BMP) and leukemia inhibitory factor signaling pathways and the expression of the transcription factors Oct4 and Nanog. However, the network of interactions among extrinsic and intrinsic determinants of ES cell pluripotency is currently poorly understood. Here, we show that *Nanog* expression is up-regulated in mouse ES cells by the binding of T (Brachyury) and STAT3 to an enhancer element in the mouse *Nanog* gene. We further show that Nanog blocks BMP-induced mesoderm differentiation of ES cells by physically interacting with Smad1 and interfering with the recruitment of coactivators to the active Smad transcriptional complexes. Taken together, our findings illustrate the existence of ES cell-specific regulatory networks that underlie the maintenance of ES cell pluripotency and provide mechanistic insights into the role of Nanog in this process.

pluripotency | T (Brachyury) | self-renewal | mesoderm differentiation | leukemia inhibitory factor

Mouse ES cells are self-renewing, pluripotent cell lines derived from preimplantation embryos (1, 2). Strict culture conditions must be followed to maintain the self-renewal of pluripotent mouse ES cells. Two extrinsic culture requirements, a feeder layer of fibroblasts and the addition of FBS, have been identified as necessary to sustain proliferation of undifferentiated mouse ES cells and their activities pinpointed to specific molecules (reviewed in ref. 3). Thus, self-renewal of mouse ES cells can be sustained in feeder-free conditions by supplementing the culture media with the cytokine leukemia inhibitory factor (LIF) (4, 5). In the absence of LIF, ES cell colonies flatten and form epithelium-like sheets (4, 5). More recently, the self-renewal promoting activity of animal serum has been identified as being mediated by ligands of specific families of the TGF-β superfamily, including the bone morphogenetic protein (BMP) family members BMP2 and BMP4 and the growth and differentiation factor (GDF) family member GDF6 (6). In the absence of BMP/GDF signals, LIF is not sufficient to prevent the neural differentiation of ES cells, whereas the absence of both BMP/GDF and LIF stimulation results in a flattened cell phenotype similar to that observed during LIF withdrawal (6).

The intracellular signaling cascades initiated by both LIF and BMP/GDF that sustain self-renewal of mouse ES cells have been worked out to a significant degree of detail (reviewed in ref. 3). In summary, binding of LIF to its cognate LIF receptor results in the recruitment of gp130 and the formation of a ternary complex that catalyzes the tyrosine phosphorylation, dimerization, and nuclear translocation of the downstream signal transducer STAT3. BMP/GDF, in turn, promotes ES cell self-renewal by inducing the expression of members of the inhibitor of differentiation (Id) family of negative transcriptional modulators, most likely mediated by activation of the TGF-β downstream signal transducer Smad1 (6).

In addition to extrinsic requirements, the pluripotency of mouse ES cells has been shown to depend on intrinsic determinants, such as the expression of the POU transcription factor Oct4 (7) and the divergent homeodomain-containing factor Nanog (8, 9). Both factors are absolutely required for ES cells to maintain their pluripotent identity. Thus, the lack (7) or down-regulation (10) of *Oct4* expression induces trophoectoderm differentiation, whereas ES cells lacking *Nanog* function differentiate to endoderm lineages (8). The relationships among extrinsic and intrinsic determinants of ES cell identity are only recently beginning to be understood. The maintenance of pluripotent ES cell self-renewal by Oct4 requires functional LIF/STAT3 and BMP/GDF/Id signaling cascades (6, 10), but the function of LIF/STAT3 does not seem to be the maintenance of *Oct4* expression (10). Overexpression of *Nanog*, in turn, circumvents the necessity of either LIF or BMP/GDF stimulation (6, 9), although synergism between Nanog function and LIF/STAT3 signaling has been noted (9).

During the investigation of the early events of ES cell specification toward mesoderm lineages, we have recently shown that ES cell cultures normally contain a population of cells expressing *T* (*Brachyury*) [T encodes one of the earliest markers of mesoderm differentiation (11, 12)] that we have termed early mesoderm-specified (EM) progenitors (13). Interestingly, in the presence of LIF, the mesoderm-specification of EM progenitors is reverted to generate fully pluripotent ES cells by a mechanism involving Nanog and T, prompting the possibility that T regulates *Nanog* expression in EM progenitors (13). Here, we show that T and STAT3 coordinately bind to a regulatory element in the mouse *Nanog* promoter, resulting in increased *Nanog* expression in EM progenitors. Furthermore, we provide evidence from gain- and loss-of-function experiments demonstrating that Nanog prevents the progression of BMP-induced mesoderm differentiation of ES cells by directly binding to Smad1 and interfering with the recruitment of coactivators, thus blocking the transcriptional activation of downstream targets, including that of *T*.

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: LIF, leukemia inhibitory factor; BMP, bone morphogenetic protein; GDF, growth and differentiation factor; EM, early mesoderm-specified.

‡Present address: Research Unit for Organ Regeneration, Center for Developmental Biology, 2-2-3 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan.

¶Present address: Gene Regulation Research, Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma 630-0101, Japan.

††To whom correspondence should be addressed at: Gene Expression Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037. E-mail: belmonte@salk.edu.

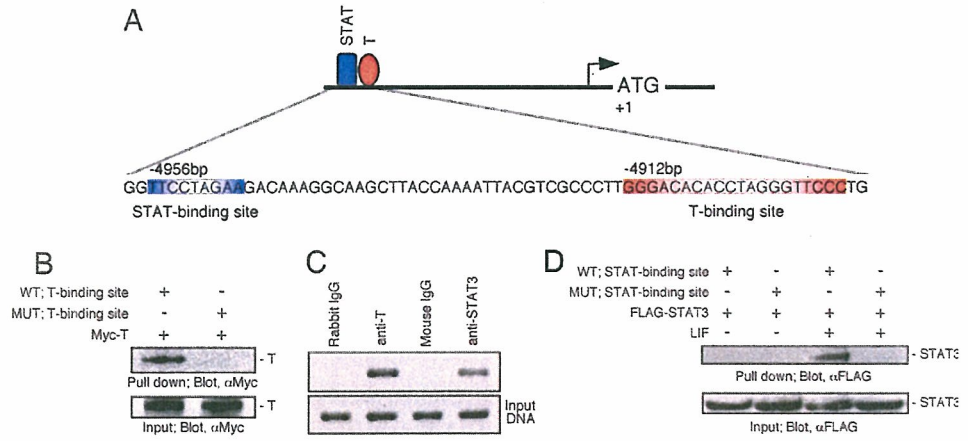

Fig. 1. Presence of STAT- and T-binding sites in the mouse *Nanog* promoter. (*A*) Schematic representation of the 5′ upstream regulatory region of the mouse *Nanog* gene. Putative STAT- and T-binding sites are indicated. (*B*) T binds to the putative T-binding site in the *Nanog* regulatory region, as shown by pull-down assays. WT and mutated (MUT) versions of double-strand oligonucleotides representing the putative T-binding site were used as probes. Input lysates were also blotted with anti-Myc antibody. (*C*) Chromatin immunoprecipitation (ChIP) assays for the putative T- and STAT-binding sites in the *Nanog* regulatory region demonstrate specific binding of T and STAT3 to the regulatory region. (*Lower*) A PCR amplification of input DNA before immunoprecipitation. (*D*) LIF-dependent binding of STAT3 to the putative STAT-binding site in the *Nanog* regulatory region. WT and mutated versions of double-strand oligonucleotides for the putative STAT-binding site were used as probes. Input lysates were also blotted with anti-FLAG antibody.

Results

T- and STAT-Binding Sites in the Mouse *Nanog* Gene. To gain insights into the regulation of *Nanog* expression by T, we analyzed the mouse *Nanog* gene in search of regulatory sequences. At 4.91 kb upstream of the translation start site of *Nanog*, we identified a 20-bp sequence forming an imperfect palindrome that shares homology with the proposed binding site for T (14) (Fig. 1*A*). We tested the ability of T to bind to oligonucleotides representing this sequence, but not to mutated versions thereof, by performing *in vitro* pull-down assays of biotin-labeled oligonucleotides incubated with lysates of NIH 3T3 cells expressing Myc-tagged T (Fig. 1*B*). We also investigated the ability of endogenous T to bind the region of interest in the *Nanog* promoter *in vivo* by chromatin immunoprecipitation (ChIP) assays of ES cells with a T-specific antibody (Fig. 1*C*). Our search for putative regulatory elements in the *Nanog* promoter also identified a predicted STAT-binding site 44 bp upstream of the T-binding site (Fig. 1*A*). We tested the ability of STAT3 to bind to this site *in vitro* (Fig. 1*D*) and *in vivo* (Fig. 1*C*) using experimental approaches similar to the ones used to characterize the T-binding site. These results uncover the presence of functional binding sites for T and STAT3 in the mouse *Nanog* promoter.

Functional Characterization of the *Nanog* EM Enhancer. We next analyzed the significance of the T- and STAT3-binding sites in the *Nanog* promoter for the biology of EM progenitors. We generated two constructs driving the expression of luciferase, one comprising 5.2 kb of the *Nanog* genomic sequence upstream of the translation start (−5203Nanog-Luc, which included both STAT3- and T-binding sites) and the other lacking the 5′-most 1 kb (and thus both STAT3- and T-binding sites, −4191Nanog-Luc). Transient transfection of ES cells with either reporter construct resulted in a similar, ≈40-fold transcriptional induction (compared with a promoterless luciferase construct) when ES cells were cultured in medium containing 1,000 units/ml LIF (Fig. 2*A*), a condition in which EM progenitors are generated at very low frequency (13). These results indicate that the regulatory elements responsible for the constitutive expression of *Nanog* in ES cells are located in the first 4.2 kb of the mouse *Nanog* gene upstream of the translation start, consistent with the recent identification of functional Oct4/Sox2-binding sites in the proximal mouse *Nanog* promoter (15, 16).

Importantly, the transcriptional activity of the −5203Nanog-Luc was increased by ≈4-fold with respect to that of the −4191Nanog-Luc in ES cells adapted to grow in medium supplemented with 400 units/ml LIF (Fig. 2*A*), in which the EM progenitor population represents ≈20% of the culture (13) (see also Fig. 3). These findings suggest that the up-regulation of *Nanog* expression in EM progenitors is controlled by regulatory sequences located between base pairs −5,203 and −4,192 upstream of the translation start of the mouse *Nanog* gene, a region containing the functional STAT3- and T-binding sites. To address whether this region could function as a transcriptional enhancer, we cloned it into a luciferase reporter

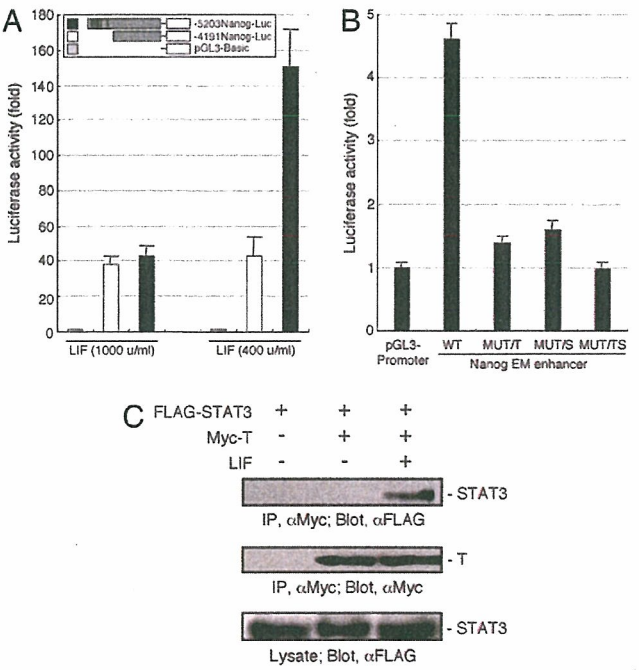

Fig. 2. Regulation of *Nanog* expression by LIF/STAT3 signaling and T. (*A* and *B*) Analysis of transcriptional activities of the *Nanog* regulatory region by luciferase reporter assay in mouse ES cells (1,000 or 400 units/ml LIF). (*A*) Both −5203Nanog-Luc and −4191Nanog-Luc showed a similar activation with 1,000 units/ml LIF, whereas −5203Nanog-Luc activity was further increased in cultures maintained with 400 units/ml LIF. (*B*) Both the T- and STAT3-binding sites were required for activation of *Nanog* EM enhancer activity in ES cells cultured with 400 units/ml LIF. WT indicates base pairs −5203 to −4192. MUT/T, MUT/S, and MUT/TS indicate the mutation in T-, STAT3-, or both T- and STAT3-binding sites, respectively, in the *Nanog* EM enhancer. Bars show mean ± SD ($n = 4$). (*C*) T and STAT3 physically interact inside cells. T and STAT3 were coimmunoprecipitated when STAT3 was activated by LIF. IP, immunoprecipitation.

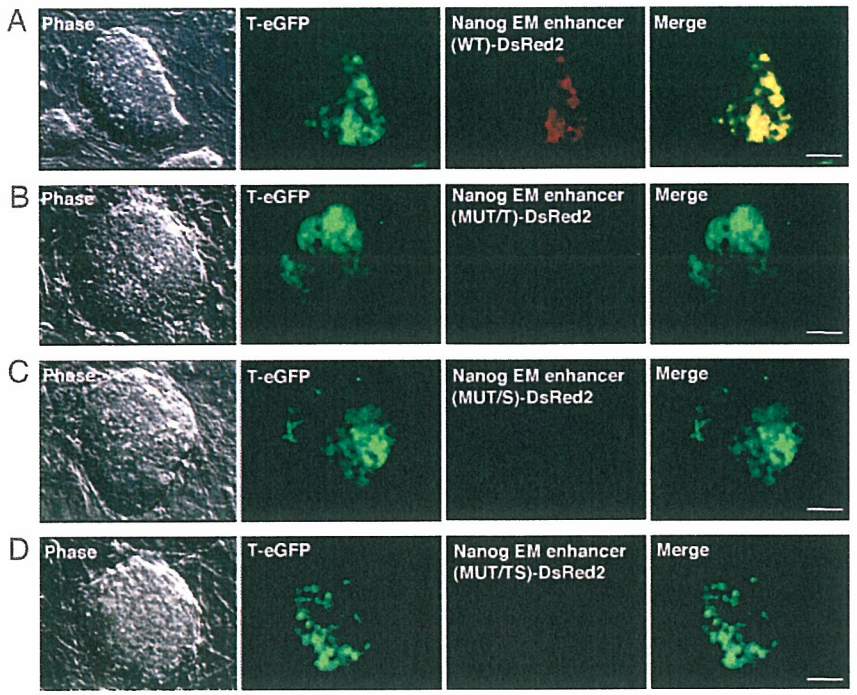

Fig. 3. The *Nanog* EM enhancer is active in EM progenitors. Fluorescent images of *T-EGFP* and *Nanog* EM enhancer DsRed2 expression in ES cell colonies formed in culture with 400 units/ml LIF. The coexpression of EGFP with DsRed2 in WT (*A*), but not when T- and/or STAT3-binding sites are mutated (*B–D*), indicates that the activity of the *Nanog* EM enhancer in EM progenitor cells is regulated by T and STAT3. (Scale bar: 10 μm.)

construct driven by a minimal promoter. This construct increased transcription levels by ≈4.5-fold when transiently transfected into ES cells cultured with 400 units/ml LIF (Fig. 2*B*). Moreover, the enhancer activity of this region was lost when either or both the STAT3- and the T-binding sites were mutated (Fig. 2*B*). These results demonstrate the existence of an enhancer of *Nanog* expression located between base pairs −5,203 and −4,192 in the mouse *Nanog* gene that is active in conditions that promote the appearance of EM progenitors and to which we refer as the *Nanog* EM enhancer.

Because the binding sites for STAT3 and T are located in close proximity to one another in the *Nanog* EM enhancer and because both T box transcription factors (17–19) and STAT3 (20, 21) have been described as physically interacting with other transcription factors for the regulation of specific promoters, we decided to analyze whether T and STAT3 could interact inside the cell. We tested this possibility in NIH 3T3 cells by cotransfecting expression vectors encoding tagged versions of STAT3 and T (FLAG-STAT3 and Myc-T) and carrying out immunoprecipitation assays. Interestingly, we found an association of T with STAT3 only when nuclear translocation of STAT3 was activated by stimulation with LIF (Fig. 2*C*).

The *Nanog* EM Enhancer Is Active in EM Progenitors. We next investigated whether the activity of the *Nanog* EM enhancer was restricted to EM progenitors. For this purpose, we used transgenic reporter ES cells that express EGFP under the regulatory sequences of the mouse *T* gene (*T-EGFP* ES cells). The maintenance of these cells in culture medium supplemented with 400 units/ml LIF increased the size of the EM progenitor population, which reached a plateau of ≈20% of the cells after 15 passages and was readily detected by the activity of the *T-EGFP* reporter (13) (see also Fig. 3). To visualize the activity of the *Nanog* EM enhancer in specific cells, we used it to drive the expression of a red fluorescent protein (DsRed2). *T-EGFP* ES cells stably expressing this second reporter showed activity of the *Nanog* EM enhancer only in EM progenitors, as evaluated by the colocalization of EGFP and DsRed2 signals in these cells (Fig. 3*A*). Consistent with the results of the luciferase reporter assays, mutation of either or both STAT3- and T-binding sites in the *Nanog* EM enhancer abrogated the activity of this reporter in EM progenitors (Fig. 3 *B–D*). Our results thus far demonstrate that the up-regulation of *Nanog* expression in EM progenitors depends on the binding of STAT3 and T to specific sites in the EM enhancer in the mouse *Nanog* gene.

Nanog Blocks Mesoderm Induction by BMPs. Because BMPs are potent inducers of mesoderm differentiation in the context of embryo development (22–24), as well as in mouse ES cells (25–28), we next tested whether the generation of EM progenitors was modified by increasing or decreasing BMP signaling in cultures of ES cells. After three passages in medium containing 400 units/ml LIF, the size of the EM progenitor population reached ≈6% in cultures of *T-EGFP* ES cells (Fig. 4*A*). This percentage almost doubled when cells were incubated in the presence of recombinant BMP2, BMP4, or BMP7 and was reduced by half upon incubation with noggin (Fig. 4*A*), a secreted factor that blocks BMP signaling (29, 30). We then tested whether BMP signaling also regulated the maintenance of EM progenitors. When pure populations of EM progenitors were plated in culture medium containing 400 units/ml LIF, ≈75% of the resulting cells underwent a transition to ES cells, whereas the remaining ≈25% maintained EM progenitor identity (Fig. 4*B*). When the cultures were supplemented with BMPs, the maintenance of EM progenitors increased by ≈2-fold, whereas it was decreased by half upon incubation with noggin (Fig. 4*B*). Interestingly, overexpression of *Nanog* in EM progenitors resulted in a decrease in their maintenance similar to that induced by noggin (Fig. 4*B*). These results indicate that the generation and maintenance of EM progenitors depends, at least in part, on the differentiation-promoting activity of BMPs and suggest that Nanog's ability to reduce the numbers of EM progenitors may depend on the blockade of BMP signaling.

Signaling by BMPs is intracellularly transduced by receptor-regulated Smads (Smad1, Smad5, and Smad8) and the comediator Smad4 and is antagonized by inhibitory Smads [Smad6 and Smad7 (reviewed in ref. 31)]. To characterize the mechanism by which Nanog blocks BMP signaling, we first analyzed the effects of *Nanog*

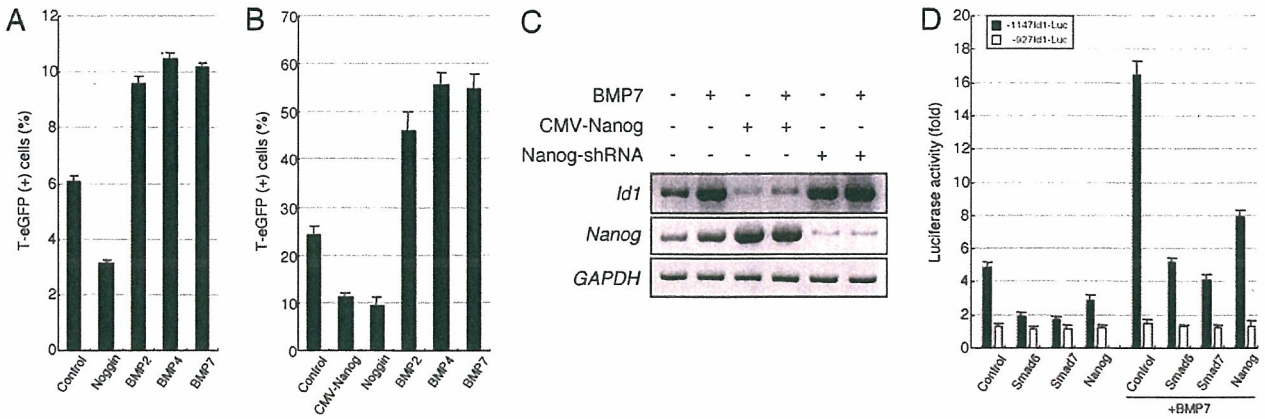

Fig. 4. BMP signaling promotes the mesoderm specification of ES cells. (*A*) Flow-cytometric analysis of T-positive cells in *T-EGFP* ES cells cultured for three passages with 400 units/ml LIF under conditions of inhibition (noggin) or activation (BMP2, BMP4, and BMP7) of BMP signaling. Bar shows mean ± SD ($n = 4$). (*B*) Flow-cytometric analysis of T-positive cells produced from purified T-positive cells cultured with 400 units/ml LIF under conditions of inhibition or activation of BMP signaling. Inhibition of endogenous BMP signaling by noggin decreased the percentage of T-positive cells at a similar level of Nanog overexpression, whereas BMP activation increased the percentage of T-positive cells. Bar shows mean ± SD ($n = 4$). (*C*) RT-PCR analyses in ES cells cultured with 400 units/ml LIF with or without addition of BMP7 show that BMP-dependent *Id1* expression is negatively regulated by overexpressing *Nanog* and enhanced when Nanog function is down-regulated. (*D*) A reporter construct of −1147Id1-Luc containing the Smad-binding sites, but not −927Id1-Luc, was activated in a BMP-dependent manner in ES cells cultured with 400 units/ml LIF. Nanog and inhibitory Smads (Smad6 and Smad7) down-regulated −1147Id1-Luc activity in a similar manner. Bars show mean ± SD ($n = 4$).

overexpression and down-regulation in the BMP-induced expression of *Id1*, a well characterized transcriptional target of BMP signaling (32). Gain of Nanog function in ES cells induced decreased basal levels of *Id1* expression and greatly impaired the ability of BMP signaling to up-regulate *Id1* expression (Fig. 4*C*). Conversely, down-regulating Nanog function with *Nanog*-specific short hairpin RNAs (shRNAs) resulted in increased levels of *Id1* expression (Fig. 4*C*). Next, we used *Id1*-luciferase reporter constructs containing (−1147Id1-Luc) or lacking (−927Id1-Luc) the Smad-binding sites (33) and analyzed their activity in ES cells. Transient transfection of these reporters in ES cells resulted in a ≈4.5-fold activation of the −1147Id1-Luc reporter when compared with the −927Id1-Luc reporter (Fig. 4*D*), indicating the existence of a significant level of endogenous BMP signaling associated with our culture conditions (see *Discussion*). Addition of BMP to the culture medium resulted in a strong up-regulation of the −1147Id1-Luc reporter as compared with the −927Id1-Luc reporter (Fig. 4*D*). That the activation of the −1147Id1-Luc reporter was due to BMP signaling was further confirmed by the fact that cotransfection of ES cells with cDNAs encoding inhibitory Smads drastically reduced the transcriptional activity of the reporter induced by endogenous or exogenous BMPs (Fig. 4*D*). Interestingly, *Nanog* overexpression in ES cells closely mimicked the effect of inhibitory Smads (Fig. 4*D*), suggesting that Nanog may block BMP signaling by interfering with the formation of activated Smad complexes.

Nanog Binds to Smad1. Inhibitory Smads negatively regulate BMP signaling by binding to activated receptor-regulated Smads, hence limiting their availability to form transcriptionally active complexes with Smad4 and/or other nuclear cofactors (reviewed in ref. 31). To address whether Nanog blocked BMP signaling by a similar mechanism, we first analyzed its ability to interact with the receptor-regulated Smad1 inside the cell. Coimmunoprecipitation assays in NIH 3T3 cells revealed that Nanog was indeed able to bind Smad1 only when the latter was activated by cotransfection of a constitutively active ALK3. Similarly, the interaction of endogenous Nanog and Smad1 could be detected in ES cells and was enhanced by BMP stimulation (Fig. 6 *A* and *B*, which is published as supporting information on the PNAS web site).

Next, we mapped the interaction domain of Smad1 with Nanog. The different Smads contain two conserved domains, the N-terminal Mad homology (MH) 1 and the C-terminal MH2 domain, separated by a poorly conserved linker. The interaction of receptor-regulated Smads with Smad4 and other transcription factors and cofactors, as well as with inhibitory Smads, occurs through the MH2 domain (reviewed in ref. 31). In cells cotransfected with Nanog and expression constructs encoding the individual MH1, MH1 plus linker, or MH2 domains of Smad1, interaction with Nanog was found exclusively with the MH2 domain (Fig. 6*C*). These results are consistent with a negative role of Nanog in BMP signaling by interfering with the interaction of receptor-activated Smads with Smad4 and/or additional nuclear factors.

The paralogous transcriptional coactivators cAMP responsive element-binding protein (CREB) binding protein (CBP) and p300 are nuclear cofactors important for TGF-β signaling, including that of BMPs, which interact with the MH2 domain of receptor-regulated Smads and Smad4 (reviewed in ref. 31). To gain further insights into the mechanism of Nanog-mediated down-regulation of BMP signaling, we tested whether Nanog interfered with the recruitment of p300 to the complexes of activated Smads. For this purpose, Myc-tagged Smad1, hemagglutinin (HA)-tagged p300, and constitutively active ALK3 were expressed in NIH 3T3 cells with or without HA-tagged Nanog. Immunoprecipitations of cell lysates were performed with anti-Myc antibodies followed by Western blotting using anti-HA antibodies. In the absence of Nanog, Smad1 efficiently coimmunoprecipitated p300. In the presence of coexpressed Nanog, the amount of p300 bound to Smad1 decreased in a Nanog dose-dependent manner (Fig. 6*D*). The functional significance of these findings was further verified by the fact that overexpression of p300 completely rescued the down-regulation in the transcriptional activity of the *Id* promoter induced by Nanog (Fig. 5*A*). These results indicate that Nanog negatively regulates BMP signaling by interfering with the recruitment of the coactivator p300 to the Smad transcriptional complex.

Nanog Blocks the Induction of *T* Expression by BMPs. Finally, the finding that the expression of *Xbra*, the homologue of *T* in *Xenopus*, is regulated by TGF-β signals (34) prompted us to investigate whether *T* could be a transcriptional target of BMP signaling in ES cells, and, if so, whether Nanog could directly block the induction of *T* by BMPs. In a preliminary analysis, we identified a BMP-responsive element in the ≈1.2-kb region upstream of the transla-

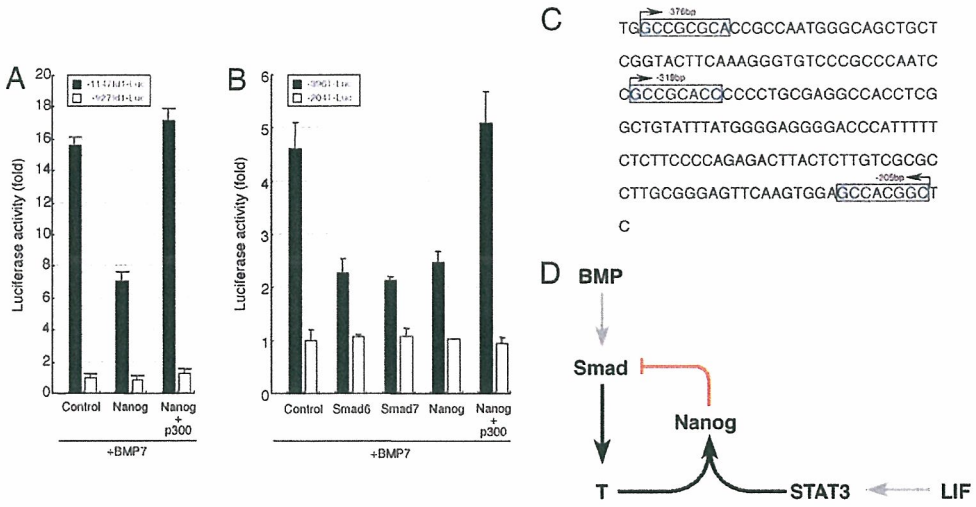

Fig. 5. Nanog down-regulates the expression of BMP targets. (*A*) Overexpression of p300 rescues the down-regulation of −1147Id1-Luc activity induced by *Nanog*. Bars show mean ± SD (n = 4). (*B*) The −396T-Luc, but not the −204T-Luc, reporter construct is activated in a BMP-dependent manner in ES cells cultured with 400 units/ml LIF. Nanog and inhibitory Smads (Smad6 and Smad7) down-regulate −396T-Luc activity in a similar manner. The down-regulation of −396T-Luc activity induced by Nanog is rescued by overexpression of p300. Bars show mean ± SD (n = 4). (*C*) Sequence of the 5′ upstream regulatory region of the mouse *T* gene. Three putative BMP-responsive Smad-binding sites are indicated by boxes. (*D*) Schematic representation of the negative feedback mechanism by which Nanog blocks BMP-induced *T* expression in the presence of LIF/STAT3 signaling. Black arrows depict positive direct transcriptional regulation, gray arrows depict positive posttranslational regulation, and red lines represent inhibitory regulation. See *Results* for details.

tion initiation site of the mouse *T* promoter (data not shown). We then generated a series of luciferase reporter constructs covering this region. We transfected these constructs into ES cells cultured in medium containing 400 units/ml LIF and supplemented with BMP7 and further mapped the BMP-responsive element to a region located between base pairs −396 and −204 of the mouse *T* gene (Fig. 5*C*). Under these conditions, the activity of the −396T-Luc reporter was ≈5-fold that of the −204T-Luc reporter and decreased by half upon coexpression of inhibitory Smads or Nanog (Fig. 5*B*). Interestingly, the down-regulation of the −396T-Luc activity induced by Nanog could be completely rescued by the coexpression of p300 (Fig. 5*B*). The analysis of this region in the mouse *T* promoter detected three motifs with homology to the reported consensus of BMP-responsive Smad-binding sites (35). These results indicate that *T* is a direct transcriptional target of BMP signaling and that Nanog down-regulates *T* expression by inhibiting BMP signaling at the level of the formation of active Smad/p300 complexes.

Discussion

Differentiation-Promoting Activity of BMPs. The results from our analyses indicate that the generation of EM progenitors from ES cells depends on the direct mesoderm-inducing ability of BMP stimulation (Figs. 4 and 5). This finding is consistent with reported roles of BMP signaling during embryo development (22–24) and with previous studies of ES cell differentiation *in vitro* (25–28). However, the mesoderm-differentiating activity of BMPs seems to be at odds with their role in maintaining the self-renewal of pluripotent ES cells (6). Indeed, BMP signaling seems to have contrasting effects on the maintenance of ES cell pluripotency. On the one hand, BMPs are necessary to prevent ES cell differentiation toward neural fates (6, 26, 36). On the other hand, signaling by BMPs induces a loss of ES cell pluripotency by promoting their differentiation toward nonneural fates such as mesoderm-derived lineages (refs. 25 and 27 and this work). These opposing effects of BMPs can be partially explained by differences in the experimental conditions used in those studies. Thus, in the absence of LIF, low concentrations of BMPs (≈0.25–10 ng/ml) promote mesoderm differentiation (25–27) at the expense of neural fates (26). In the presence of LIF, however, similarly low concentrations of BMPs prevent neural differentiation of ES cells (6, 36) and maintain their pluripotency with no signs of mesoderm differentiation (6). Consistent with this notion, we did not detect increased generation of EM progenitors with BMP concentrations <100 ng/ml in the presence of LIF (data not shown). Thus, LIF seems to render ES cells refractory to the mesoderm-inducing activity of BMPs. Our studies demonstrate that this resistance depends, at least in part, on a negative feedback mechanism mediated by Nanog and T.

A Negative Feedback That Blocks Mesoderm Specification in ES Cells. Our results also provide mechanistic insights into the relationships among ES cell pluripotency determinants. Thus, we characterize a negative feedback mechanism that prevents mesoderm specification of ES cells in the presence of LIF (Fig. 5*D*). In this mechanism, mesoderm differentiation of ES cells is initiated by BMP signaling. Possible sources of BMP activity in our culture conditions include FCS (25), the fibroblast feeder layer, and/or ES cells themselves (6). Consistent with this, we detect a significant activation of the −1147Id1-Luc reporter even in the absence of exogenous BMP supplements (Fig. 4*C*). As a direct consequence of BMP signaling, ES cells undergoing mesoderm specification activate the expression of *T* (Fig. 5*B*). In the presence of LIF, activated STAT3 interacts with T and binds to the *Nanog* EM enhancer, thus resulting in the up-regulation of *Nanog* expression in these cells. Increased levels of Nanog, in turn, directly block the mesoderm-differentiation activity of BMPs, thereby limiting the progression of mesoderm specification and down-regulating *T* expression, ultimately regenerating pluripotent ES cells from EM progenitors.

However, it is clear that the functions of LIF and Nanog in the maintenance of ES cell pluripotency are not restricted to participating in the negative feedback mechanism characterized here. Thus, the complete lack of Nanog function promotes differentiation of ES cells to endoderm lineages (8), indicating the existence of roles of Nanog other than that of preventing mesoderm differentiation. Indeed, the up-regulation of *Nanog* expression by T and STAT3 takes place only in EM progenitors (Fig. 3*A*), whereas the constitutive expression of *Nanog* in ES cells is regulated by more proximal regions of the *Nanog* promoter (Fig. 2*A*).

Mouse ES cells, consistent with their developmental origin in the embryo epiblast, have the ability to give rise to derivatives of all three primary germ layers. However, unlike cells in the epiblast, in which pluripotency is very transient, mouse ES cells can be maintained in culture indefinitely in a pluripotent state. The mechanism(s) whereby the adaptation to culture conditions releases epiblast cells from the loss of pluripotency remain an outstanding question in the biology of ES cells. Our results provide insights into

the mechanisms by which the determinants of ES pluripotency interact to actively prevent lineage specification of ES cells.

Materials and Methods

ES Cell Culture. Mouse ES cells (J1 line, ref. 37) were maintained on mouse embryonic fibroblasts in standard media containing 15% FBS (HyClone) and LIF (1,000 or 400 units/ml, ESGRO; Chemicon). In some cases, BMP2, BMP4, BMP7 (all 200 ng/ml), or noggin (100 ng/ml) was supplemented in the cultures [BMP7 and noggin were kind gifts from S. Choe (The Salk Institute for Biological Studies); BMP2 and BMP4 were from R & D Systems]. In each passage, 3×10^5 cells were plated in a 6-cm dish. For clonal analyses, *T*-positive and -negative cells were separately isolated by using FACS (MoFlo; Cytomation, Fort Collins, CO) and plated at a density of 1×10^4 cells per 6-cm dish. Before analyses, trypsinized cells were allowed to attach to gelatin-coated plates for 45 min, thus removing >95% of feeder cells and recovering >95% of ES cells.

Fluorescent Reporter Constructs. The first 24 bp of the *T* coding sequence in a mouse genomic bacterial artificial chromosome (BAC) (RP23-456E5, BACPAC Resources, Oakland, CA) were replaced with an EGFP-Neo cassette by homologous recombination in bacteria (38). The bacterial strain DY380 was kindly provided by N. G. Copeland (National Cancer Institute, Frederick, MD). Transgenic ES cell lines expressing this construct (*T-EGFP*) were obtained by electroporation and G418 selection. DNA fragments of the mouse *Nanog* gene were obtained from a mouse genomic BAC clone (RP23-406B15). The *Nanog* EM enhancer region (base pairs −5203 to −4192) and its mutated versions, in which the T- and/or STAT-binding sites were mutated (T, GG-GAC*GCGCCT*GAATCCTAT; STAT, *CCAA*TAGAA; italics indicate mutated nucleotides), were inserted into a c-fos minimal promoter–DsRed2 vector. Mouse ES cells were transfected with these reporter constructs by using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions.

Supporting Information. The immunoprecipitation, Western blotting, RT-PCR, and chromatin immunoprecipitation procedures and the expression constructs used are in *Supporting Text*, which is published as supporting information on the PNAS web site.

We thank Robert Benezra, Senyon Choe, Neal G. Copeland, Richard Eckner, Kohei Miyazono, and Shinya Yamanaka for sharing reagents; Dirk Buscher, Chris Kintner, Isao Oishi, Junichiro Sonoda, and Ayumu Tashiro for helpful suggestions; Harley Pineda, Timothy Chapman, and Henry Juguilon for excellent technical assistance; and May-Fun Schwarz for help in the preparation of the manuscript. A.S. was partially supported by Japan Society for the Promotion of Science (JSPS) Research Fellowships for Young Scientists; A.S., T.M., and K.N. are partially supported by JSPS Postdoctoral Fellowships for Research Abroad; and A.R. and C.R.-E. were partially supported by Fundación Inbiomed. Funding for mouse ES cell work in J.C.I.B.'s laboratory was from the G. Harold and Leila Y. Mathers Charitable Foundation and the National Institutes of Health.

1. Martin, G. R. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 7634–7638.
2. Evans, M. J. & Kaufman, M. H. (1981) *Nature* **292,** 154–156.
3. Chambers, I. & Smith, A. (2004) *Oncogene* **23,** 7150–7160.
4. Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M. & Rogers, D. (1988) *Nature* **336,** 688–690.
5. Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A. & Gough, N. M. (1988) *Nature* **336,** 684–687.
6. Ying, Q. L., Nichols, J., Chambers, I. & Smith, A. (2003) *Cell* **115,** 281–292.
7. Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H. & Smith, A. (1998) *Cell* **95,** 379–391.
8. Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. & Yamanaka, S. (2003) *Cell* **113,** 631–642.
9. Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S. & Smith, A. (2003) *Cell* **113,** 643–655.
10. Niwa, H., Miyazaki, J. & Smith, A. G. (2000) *Nat. Genet.* **24,** 372–376.
11. Herrmann, B. G., Labeit, S., Poustka, A., King, T. R. & Lehrach, H. (1990) *Nature* **343,** 617–622.
12. Wilkinson, D. G., Bhatt, S. & Herrmann, B. G. (1990) *Nature* **343,** 657–659.
13. Suzuki, A., Raya, A., Kawakami, Y., Morita, M., Matsui, T., Nakashima, K., Gage, F. H., Rodriguez-Esteban, C. & Belmonte, J. C. (2006) *Nat. Clin. Pract. Cardiovasc. Med.* **3,** S114–S122.
14. Kispert, A. & Herrmann, B. G. (1993) *EMBO J.* **12,** 3211–3220.
15. Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S. Y., Suemori, H., Nakatsuji, N. & Tada, T. (2005) *Mol. Cell. Biol.* **25,** 2475–2485.
16. Rodda, D. J., Chew, J. L., Lim, L. H., Loh, Y. H., Wang, B., Ng, H. H. & Robson, P. (2005) *J. Biol. Chem.* **280,** 24731–24737.
17. Hiroi, Y., Kudoh, S., Monzen, K., Ikeda, Y., Yazaki, Y., Nagai, R. & Komuro, I. (2001) *Nat. Genet.* **28,** 276–280.
18. Stennard, F. A., Costa, M. W., Elliott, D. A., Rankin, S., Haast, S. J., Lai, D., McDonald, L. P., Niederreither, K., Dolle, P., Bruneau, B. G., *et al.* (2003) *Dev. Biol.* **262,** 206–224.
19. Garg, V., Kathiriya, I. S., Barnes, R., Schluterman, M. K., King, I. N., Butler, C. A., Rothrock, C. R., Eapen, R. S., Hirayama-Yamada, K., Joo, K., *et al.* (2003) *Nature* **424,** 443–447.
20. Zhu, M., John, S., Berg, M. & Leonard, W. J. (1999) *Cell* **96,** 121–130.
21. Collum, R. G., Brutsaert, S., Lee, G. & Schindler, C. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 10120–10125.
22. Dale, L., Howes, G., Price, B. M. & Smith, J. C. (1992) *Development (Cambridge, U.K.)* **115,** 573–585.
23. Jones, C. M., Lyons, K. M., Lapan, P. M., Wright, C. V. & Hogan, B. L. (1992) *Development (Cambridge, U.K.)* **115,** 639–647.
24. Winnier, G., Blessing, M., Labosky, P. A. & Hogan, B. L. (1995) *Genes Dev.* **9,** 2105–2116.
25. Johansson, B. M. & Wiles, M. V. (1995) *Mol. Cell. Biol.* **15,** 141–151.
26. Finley, M. F., Devata, S. & Huettner, J. E. (1999) *J. Neurobiol.* **40,** 271–287.
27. Czyz, J. & Wobus, A. (2001) *Differentiation* **68,** 167–174.
28. Yuasa, S., Itabashi, Y., Koshimizu, U., Tanaka, T., Sugimura, K., Kinoshita, M., Hattori, F., Fukami, S. I., Shimazaki, T., Okano, H., *et al.* (2005) *Nat. Biotechnol.* **23,** 607–611.
29. Holley, S. A., Neul, J. L., Attisano, L., Wrana, J. L., Sasai, Y., O'Connor, M. B., De Robertis, E. M. & Ferguson, E. L. (1996) *Cell* **86,** 607–617.
30. Zimmerman, L. B., De Jesus-Escobar, J. M. & Harland, R. M. (1996) *Cell* **86,** 599–606.
31. Shi, Y. & Massague, J. (2003) *Cell* **113,** 685–700.
32. Hollnagel, A., Oehlmann, V., Heymer, J., Ruther, U. & Nordheim, A. (1999) *J. Biol. Chem.* **274,** 19838–19845.
33. Nakashima, K., Takizawa, T., Ochiai, W., Yanagisawa, M., Hisatsune, T., Nakafuku, M., Miyazono, K., Kishimoto, T., Kageyama, R. & Taga, T. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 5868–5873.
34. Latinkic, B. V., Umbhauer, M., Neal, K. A., Lerchner, W., Smith, J. C. & Cunliffe, V. (1997) *Genes Dev.* **11,** 3265–3276.
35. Kim, J., Johnson, K., Chen, H. J., Carroll, S. & Laughon, A. (1997) *Nature* **388,** 304–308.
36. Tropepe, V., Hitoshi, S., Sirard, C., Mak, T. W., Rossant, J. & van der Kooy, D. (2001) *Neuron* **30,** 65–78.
37. Li, E., Bestor, T. H. & Jaenisch, R. (1992) *Cell* **69,** 915–926.
38. Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G. & Court, D. L. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 5978–5983.

www.nature.com/clinicalpractice/cardio

Maintenance of embryonic stem cell pluripotency by Nanog-mediated reversal of mesoderm specification

Atsushi Suzuki, Ángel Raya, Yasuhiko Kawakami, Masanobu Morita, Takaaki Matsui, Kinichi Nakashima, Fred H Gage, Concepción Rodríguez-Esteban and Juan Carlos Izpisúa Belmonte*

SUMMARY

Embryonic stem cells (ESCs) can be propagated indefinitely in culture, while retaining the ability to differentiate into any cell type in the organism. The molecular and cellular mechanisms underlying ESC pluripotency are, however, poorly understood. We characterize a population of early mesoderm-specified (EM) progenitors that is generated from mouse ESCs by bone morphogenetic protein stimulation. We further show that pluripotent ESCs are actively regenerated from EM progenitors by the action of the divergent homeodomain-containing protein Nanog, which, in turn, is upregulated in EM progenitors by the combined action of leukemia inhibitory factor and the early mesoderm transcription factor T/Brachyury. These findings uncover specific roles of leukemia inhibitory factor, Nanog, and bone morphogenetic protein in the self-renewal of ESCs and provide novel insights into the cellular bases of ESC pluripotency.

KEYWORDS embryonic stem cell, mesoderm specification, Nanog, pluripotency, regenerative medicine

A Suzuki and T Matsui were Research Associates, Á Raya was a Senior Research Associate, Y Kawakami is a Staff Scientist, M Morita is a Research Associate, C Rodríguez-Esteban is a Senior Research Associate, and JC Izpisúa Belmonte is a Professor in the Gene Expression Laboratory; and K Nakashima was a Research Associate and FH Gage is a Professor in the Laboratory of Genetics, all at the Salk Institute for Biological Studies, La Jolla, CA, USA. K Nakashima is also a Professor at the Laboratory of Molecular Neuroscience, Nara Institute of Science and Technology, Ikoma, Japan. A Suzuki is a Research Scientist at the Research Unit for Organ Regeneration, Center for Developmental Biology, Kobe, Japan. Á Raya is an ICREA Research Professor, and JC Izpisúa Belmonte is Scientific Director at the Center of Regenerative Medicine, in Barcelona, Spain. T Matsui is an Assistant Professor at the Graduate School of Biological Sciences, Nara Institute of Science and Technology, Takayama, Ikoma, Japan.

Correspondence
*Gene Expression Laboratory, Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
belmonte@salk.edu

Received 7 September 2005 **Accepted** 7 November 2005
www.nature.com/clinicalpractice
doi:10.1038/ncpcardio0442

INTRODUCTION

Mouse embryonic stem cells (ESCs) are permanent cell lines derived from preimplantation embryos[1,2] that display the peculiarities of combining unlimited self-renewal and pluripotency abilities while retaining a normal KARYOTYPE (see review[3]). Arguably, most of the recent interest in ESC research results from the successful derivation of human pluripotent cell lines,[4,5] which has created new prospects for future cell replacement therapies. However, many basic questions about the biology of these promising cells need to be answered if their potential is to be realized. Human ESCs share with their mouse homonyms the peculiarities of self-renewal and pluripotency. The molecular mechanisms by which self-renewal and pluripotency are maintained in human and mouse ESCs, however, appear to differ (see review[6]). Moreover, the cellular bases of pluripotency of either mouse or human ESCs are largely unknown.

In this paper, we analyze the dynamics of expression of *Brachyury* (*T*), which encodes the transcription factor T/Brachyury, one of the earliest markers of mesoderm differentiation[7,8] in cultures of mouse ESCs. We show that mouse ESCs cultured on feeders in the presence of LEUKEMIA INHIBITORY FACTOR[9,10] (LIF) and serum contain mesoderm-specified progenitors, the number of which is dependent on the amount of LIF in the culture medium. We also show that, in the presence of LIF, the specification of these cells to a mesoderm fate can be reversed, so that they give rise to fully pluripotent ESCs. We further demonstrate that the reversal of mesoderm specification is regulated by NANOG.[11,12] Specifically, we identify a negative feedback mechanism by which, in the presence of LIF, *Nanog* expression is upregulated by T, the expression of which is, in turn, repressed by Nanog. Our findings indicate that ESC pluripotency is actively maintained by specific molecular mechanisms that prevent and/or reverse cell differentiation.

METHODS

Embryonic stem cell culture

Mouse ESC lines J1[13] and SAT1, 2, 6, and 11 (derived in our laboratory from C57BL/6×129/TerSv blastocysts) were maintained on mouse embryonic fibroblasts in standard media containing 15% fetal bovine serum (HyClone, Logan, UT, USA) and LIF (1,000 or 400 U/ml, ESGRO; Chemicon, Temecula, CA). In each passage, 3×10^5 cells were plated in a 6-cm dish. For clonal analyses, T^+ and T^- cells were separately isolated using fluorescence-activated cell sorting (FACS) (MoFlo; Dako, Fort Collins, CO, USA) and plated at a density of 1×10^4 per 6-cm dish. Attached cells were found after 8 h and traced over the following 2 days. Before FACS analysis, trypsinized cells were allowed to attach to gelatin-coated plates for 45 min, thus removing more than 95% of feeder cells and recovering more than 95% of ESCs.

For differentiation assays, ESCs were aggregated in hanging drops (5×10^3 cells each) for 3 days without LIF, and the embryoid bodies thus formed were cultured without LIF on gelatin-coated dishes for 5 days. For immunocytochemical analyses of myosin or albumin, embryoid bodies were cultured for 12 days, and for Tuj1 analyses, embryoid bodies were cultured for 2 days in medium supplemented with 1 μM all-*trans* retinoic acid (Sigma-Aldrich Corp., St Louis, MO, USA).

The human ESC line HUES7 (a gift from D Melton) was maintained essentially as described elsewhere,[14] except that LIF was omitted from the culture medium and Plasbumin-25 (Bayer Corp., Elkhart, IN, USA) was substituted for Plasmanate. In some cases, a supplement of LIF (400 U/ml) was added.

Fluorescent reporter constructs

The first 24 bp of the *T* coding sequence in a mouse genomic bacterial artificial chromosome (BAC) (RPCI23–456E5; BACPAC Resources Center, Oakland, CA, USA) were replaced with ENHANCED GREEN FLUORESCENT PROTEIN (eGFP)-Neo or DsRed2-Neo cassettes by homologous recombination in bacteria.[15] The bacterial strain DY380 was kindly provided by NG Copeland. Transgenic ESC lines expressing these constructs (T-eGFP or T-DsRed2) were obtained by electroporation and G418 selection.

Reverse transcriptase PCR analysis

Total RNA isolation, cDNA synthesis, and semi-quantitative reverse transcriptase polymerase chain reaction (RT-PCR) analysis were performed as described elsewhere.[16,17] Primer sequences and PCR conditions are available on request.

Immunostaining

Cells were fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) for 30 min at 25 °C, followed by washing with PBS + 0.1% Tween-20 (PBT), blocking with 5% goat or donkey serum/PBT, and incubation at 4 °C overnight with the following primary antibodies: anti-Oct4 (1:200, Santa Cruz Biotechnology, Santa Cruz, CA, USA), anti-GFP (1:200, Molecular Probes, or 1:100, Santa Cruz), anti-DsRed (1:200, Clontech, Mountain View, CA, USA), anti-Tuj1 (1:200, Sigma-Aldrich), anti-myosin (1:50, Developmental Studies Hybridoma Bank, University of Iowa, clone MF20), or anti-albumin (1:300, Biogenesis, Kingston, NH, USA). After washing with PBT, cells were incubated with Alexa 488-conjugated (1:200, Molecular Probes, Invitrogen, Carlsbad, CA, USA) and/or Cy3-conjugated (1:200; Jackson ImmunoResearch Laboratories, West Grove, PA, USA) secondary antibodies specific to the appropriate species for 2 h at 25 °C. After incubation with 4',6-diamidino-2-phenylindole and washing with PBT, specific immunoreactivities were imaged by confocal microscopy (Radiance 2100, Bio-Rad, Hercules, CA, USA).

Nanog–short hairpin RNA

Nanog–short hairpin RNAs (Nanog–shRNAs) driven by a mouse U6 promoter (a gift from G Tiscornia) were amplified by PCR and inserted into a vector containing a puromycin-resistance cassette. The sequence and design of the Nanog–shRNAs are available on request.

Western blotting analysis

Antibodies against STAT3 (1:1,500; Cell Signaling Technology, Beverly, MA, USA), tyrosine-phosphorylated STAT3 (1:1,500; Cell Signaling Technology), Nanog (1:600; a gift from S Yamanaka), T (1:1,000; Orbigen, San Diego, CA, USA), and β-actin (1:5,000; Abcam, Cambridge, MA, USA) were used for western blotting analysis.

Expression constructs

The entire coding sequences of mouse T, dnT, Nanog, and human Nanog1 (hNanog1) were subcloned into pTRE2pur (Clontech) or a vector containing a cytomegalovirus (CMV) promoter and puromycin-resistance cassette. Transfection

GLOSSARY

KARYOTYPE
Characterization of the chromosomal complement of the cell, including number, form, and size of the chromosomes

LEUKEMIA INHIBITORY FACTOR
A glycoprotein cytokine of pleiotropic actions, most notably to preserve the pluripotency of mouse embryonic stem cells in culture

NANOG
A divergent homeodomain-containing transcription factor identified as an absolute requirement of mouse embryonic stem cell pluripotency

ENHANCED GREEN FLUORESCENT PROTEIN
A spontaneously fluorescent protein isolated from the jellyfish *Aequoria victoria*, frequently used as a reporter for imaging of living cells

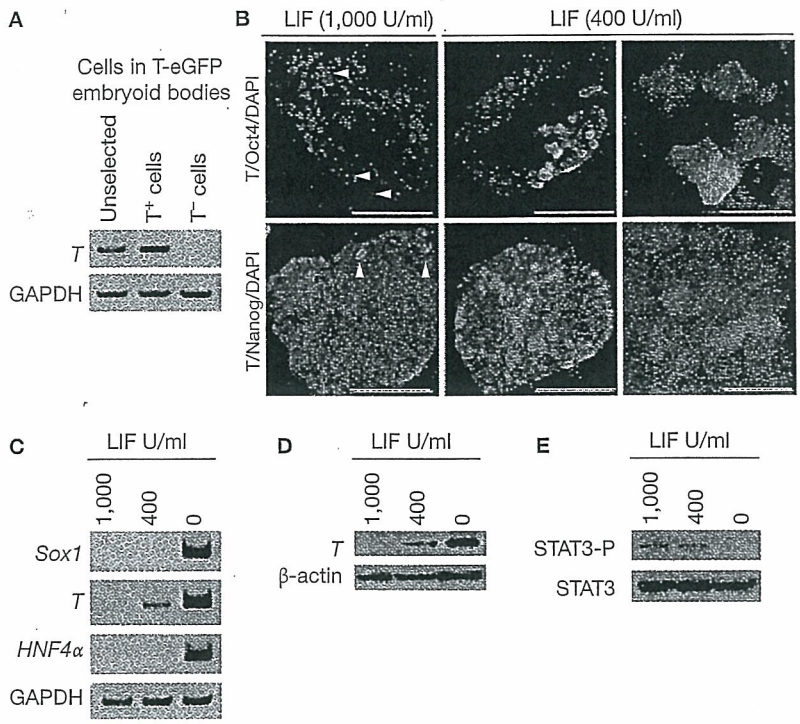

Figure 1 Leukemia inhibitory factor regulates the percentage of *T*-expressing cells in mouse embryonic stem cell cultures. (**A**) Reverse transcription polymerase chain reaction analysis of *T* expression in *T-eGFP* embryoid bodies (unselected cells), and in T^+ or T^- cells isolated from them. *T-eGFP* expression recapitulated endogenous *T* expression. (**B**) Fluorescent images of *T* (eGFP)/Oct4 and *T* (DsRed2)/*Nanog* (eGFP) expression in mouse embryonic stem cell colonies after 25 passages in medium containing leukemia inhibitory factor 1,000 or 400 U/ml. 4',6-Diamidino-2-phenylindole staining identifies individual nuclei in each field. T^+ cells were observed in colonies formed in culture with leukemia inhibitory factor not only at 400 U/ml (right and center panels), but also at 1,000 U/ml (left panels, arrowheads). Scale bars, 50 μm (left and center panels) and 10 μm (right panels). (**C**) Embryonic stem cells cultured with leukemia inhibitory factor 400 U/ml expressed the mesoderm marker *T*, but not ectoderm or endoderm markers. Markers of all three lineages were expressed in embryonic stem cells cultured in the absence of leukemia inhibitory factor (0). Reverse transcriptase polymerase chain reaction analyses of ectoderm (*Sox1*), mesoderm (*T*), and endoderm (*HNF4α*) early marker genes in *T-eGFP* embryonic stem cells obtained after 20 passages in medium containing leukemia inhibitory factor 1,000 or 400 U/ml, or after 4 passages in medium without the factor. (**D,E**) Western blotting analysis of (D) T and (E) tyrosine-phosphorylated STAT3 (STAT3-P) in *T-eGFP* embryonic stem cells cultured as in (C). The levels of β-actin (D) and STAT3 (E) are shown as controls. GAPDH, glyceraldehyde-3-phosphate dehydrogenase; LIF, leukemia inhibitory factor; STAT3-P, tyrosine-phosphorylated STAT3; T-eGFP, transgenic embryonic stem cell line expressing enhanced green fluorescent protein under the regulatory regions of the mouse *T* gene.

of mouse ESCs was carried out with lipofectamine 2000 and transfection of human ESCs with ExGen 500 (Fermentas, Hanover, MD, USA), following the manufacturer's instructions.

RESULTS

Mouse embryonic stem cells contain a population of *T*-expressing cells

To characterize the early steps of mouse ESC differentiation toward mesoderm lineages, we generated transgenic ESC lines expressing eGFP under the regulatory sequences of *T*, one of the earliest markers of mesoderm differentiation.[7,8] The expression of eGFP in nine independent *T-eGFP* ESC lines faithfully recapitulated that of endogenous *T*, as assayed by the presence of *T* transcripts in *T-eGFP*-positive (T^+) cells sorted from embryoid bodies differentiated *in vitro*, and by their absence in *T-eGFP*-negative (T^-) cells (Figure 1A and data not shown). Interestingly, colonies of undifferentiated *T-eGFP* ESCs grown under standard culture conditions (on a fibroblast feeder layer in culture medium containing serum and LIF 1,000 U/ml) contained T^+ cells. These cells were found in small numbers (1–3 cells per colony) in colonies of otherwise undifferentiated morphology (6.7%, $n=120$, Figure 1B), and

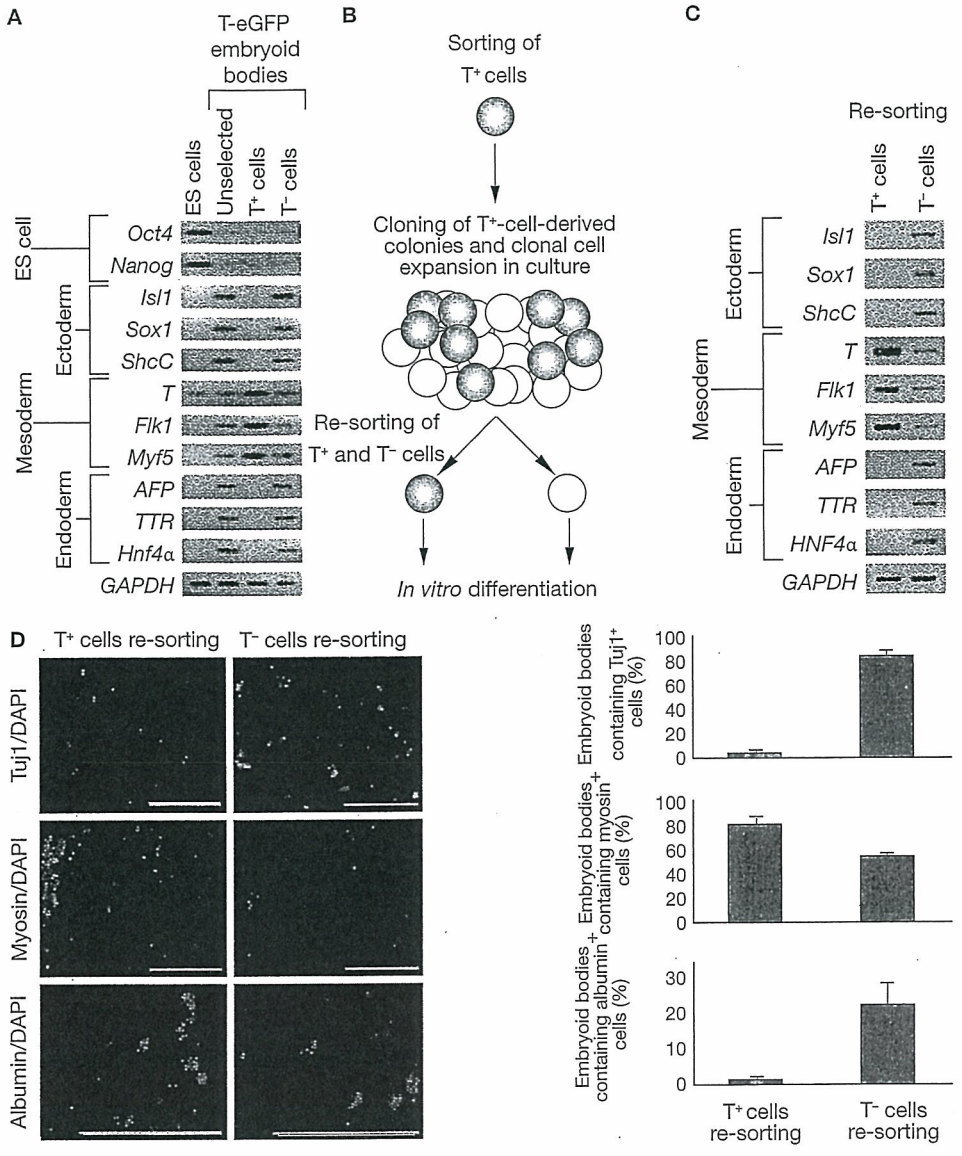

Figure 2 Dedifferentiation of lineage-specified mesoderm progenitors into pluripotent stem cells. (**A**) Reverse transcription polymerase chain reaction analysis of embryonic stem cell, ectoderm, mesoderm, and endoderm marker gene expression in *T-eGFP* embryonic stem cells obtained after 20 passages in medium containing leukemia inhibitory factor 400 U/ml, and in embryoid bodies derived from unselected *T-eGFP* embryonic stem cells, and from purified T^+ or T^- cells. (**B**) Experimental procedure for characterizing T^+ and T^- cells derived from a single T^+ cell (early mesoderm-specified progenitor cell). A single T^+ cell forms a mosaic colony in clonal density cultures with leukemia inhibitory factor 400 U/ml. Each mosaic colony was independently picked up and expanded. Then, T^+ and T^- cells originally derived from a single T^+ cell were re-sorted by fluorescence-activated cell sorter to analyze their differentiation potential *in vitro*. (**C**) Reverse transcriptase polymerase chain reaction analysis for ectoderm, mesoderm, and endoderm marker gene expression in embryoid bodies derived from re-sorted T^+ or T^- cells. (**D**) Immunocytochemical analysis of Tuj1 (neuronal lineage), myosin (muscle lineage), and albumin (hepatic lineage) production in embryoid bodies derived from re-sorted T^+ or T^- cells. Scale bars, 100 μm. Graphs show the percentage of embryoid bodies containing antigen-positive cells. Bar shows mean ± SD ($n = 3$; 30 embryoid bodies were examined in each dish).
DAPI, 4',6-diamidino-2-phenylindole; ES cells, embryonic stem cells; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; T-eGFP, transgenic embryonic stem cell line expressing enhanced green fluorescent protein under the regulatory regions of the mouse *T* gene.

no colonies formed exclusively by T^+ cells were ever detected under these culture conditions. Flow cytometric analyses of cultures of *T-eGFP* ESCs revealed a $0.59 \pm 0.05\%$ of T^+ cells ($n=4$), a fraction that remained virtually invariable after more than 50 passages (data not shown). These findings are consistent with the widely known (although frequently overlooked) fact that mouse ESCs, even when maintained under optimal culture conditions, exhibit some degree of spontaneous differentiation. Indeed, this very fact has been recognized as a hallmark of a 'good' ESC line.[3] Our observation is also consistent with the identification of *T* as a transcript selectively enriched in undifferentiated mouse ESCs.[11,18]

Since the addition of LIF is critical for the maintenance of undifferentiated ESCs, we next investigated whether the percentage of T^+ cells in our cultures depended on the amount of LIF in the culture medium. *T-eGFP* ESCs cultured in the absence of exogenous LIF differentiated extensively over time and could not be grown over five passages. Under these conditions, the percentage of T^+ cells rose exponentially and markers of ectoderm, mesoderm, and endoderm lineages were expressed (Figure 1C and data not shown), indicating that the majority of ESCs underwent spontaneous differentiation. In contrast, reduction of exogenous LIF to 400 U/ml resulted in colonies of normal ESC morphology (Figure 1B) that could be maintained in culture for over 50 passages with no signs of differentiation or crisis. This finding is consistent with the range of exogenous LIF concentrations reported to sustain self-renewal of pluripotent mouse ESCs,[10,19] with the activation of STAT3 phosphorylation under these conditions (Figure 1E), and with the fact that ESCs cultured in medium containing LIF 400 U/ml for extended periods of time, when injected into blastocysts, result in degrees of chimerism and contribution to the germline not different from ESCs maintained in LIF 1,000 U/ml (our unpublished observations).

Despite their apparently undifferentiated morphology, colonies of *T-eGFP* ESCs adapted to grow in medium supplemented with LIF 400 U/ml contained large numbers of T^+ cells (Figure 1B). The adaptation process was gradual, so that the percentage of T^+ cells increased over time when the cells were switched to culture medium containing LIF 400 U/ml and reached a plateau of $21.2 \pm 2.0\%$ at 15 passages ($n=4$), after which the fraction of T^+ cells remained constant (data not shown). Also, the expression of *T* at both mRNA (Figure 1C) and protein (Figure 1D) levels was increased in cultures supplemented with LIF 400 U/ml when compared with *T-eGFP* ESCs grown in medium containing 1,000 U/ml. To rule out the possibility that the increased number of T^+ cells under these conditions represented a peculiarity of the ESC line (J1) used in these experiments, we established four independent ESC lines from a different genetic background (C57BL/6 × 129/TerSv) that were used to generate additional *T-eGFP* transgenic lines. With small variations, all four lines displayed similar dynamics of accumulation of T^+ cells when cultured in medium containing LIF 400 U/ml (data not shown).

Other than the increased numbers of T^+ cells, we could not detect any differences in cultures maintained with LIF 400 versus 1,000 U/ml. Thus, the increase in *T* expression was not accompanied by upregulation of other transcripts involved in mesoderm differentiation, and no markers of ectoderm or endoderm differentiation were detected in ESC cultures supplemented with LIF 400 U/ml (Figure 1C). Moreover, T^+ cells generated under these conditions showed proliferation rates similar to those of T^- cells when plated at high density (0.2×10^6 cells yielded $4.18 \pm 0.27 \times 10^6$ and $4.00 \pm 0.31 \times 10^6$ cells, respectively, after 7 days in culture, $n=3$), coexpressed the pluripotency-associated markers *Oct4*, *Nanog*, and *Rex1* (Figure 1B), and stained positive for alkaline phosphatase (not shown).

Taken together, our results indicate that, in cultures of mouse ESCs, a population of T^+ cells exists, the size of which is controlled by the amount of LIF present in the culture medium. Importantly, lowering the concentration of LIF to 400 U/ml did not appear to be detrimental for the self-renewal of mouse ESCs, even though T^+ cells formed up to 20% of the cells in these culture conditions. For these reasons, and since the size of the T^+ fraction was more amenable to analysis, we continued our studies in culture medium containing LIF 400 U/ml (unless otherwise stated).

T^+ embryonic stem cells are mesoderm-specified progenitors

To characterize the identity of T^+ cells, we first analyzed their ability to generate differentiated progeny. For this purpose, we performed *in vitro* differentiation assays of bulk *T-eGFP* ESCs as well as of sorted populations of T^+ and T^- cells. In these assays, unsorted ESCs and T^- cells behaved

similarly, giving rise to differentiated cells that expressed markers of ectoderm, mesoderm, and endoderm fates (Figure 2A). In contrast, embryoid bodies formed from T^+ cells differentiated exclusively into cells expressing markers of mesoderm lineages (Figure 2A), indicating that the T^+ cells present in our cultures of *T-eGFP* ESCs were lineage-specified. Thus, we termed this population of T^+ cells 'early mesoderm-specified' (EM) progenitors.

Next, we analyzed the differentiation potential of T^- cells generated from EM progenitors (Figure 2B). *In vitro* differentiation assays revealed that T^- cells derived from EM progenitor cells were able to give rise to cells of ectoderm, mesoderm, and endoderm lineages (Figure 2C). Moreover, T^- cells readily generated beating cardiomyocytes that stained positive for myosin, albumin-positive cells after prolonged periods under differentiation-promoting conditions, and Tuj1-positive cells after treatment with retinoic acid (Figure 2D). Our results show that T^+ cells represent a population of EM progenitors that, in the presence of LIF, are able to recover the self-renewal abilities and pluripotency characteristics of mouse ESCs.

Nanog regulates the transition of early-mesoderm-specified progenitors to embryonic stem cells

In our characterization of the transcriptional profile of EM progenitors, we did not detect changes in the expression of the pluripotency-associated markers *Oct4* and *Rex1* or in the levels of *Gbx2*, *Fgf5*, or *Lif* expression (see below). However, we detected a clear upregulation in the expression of the pluripotency-associated marker *Nanog*, when compared with that of ES T^- cells (Figure 3A), raising the possibility that Nanog function is mechanistically linked to the transition of ESCs to EM progenitors, or vice versa. To investigate these possibilities, we used a *Nanog* expression transgene driven by a strong constitutive promoter, which resulted in sustained overexpression of *Nanog* transcripts in *T-eGFP* ESCs (Figure 3B). In these conditions, the expression of *T* was downregulated (Figure 3B), and the transition of EM progenitors to ESCs was facilitated, as evaluated by the ~3-fold reduction in the percentage of T^+ cells ($7.8 \pm 1.1\%$, $n=4$; Figure 3F), when compared with that of mock-transfected *T-eGFP* ESCs ($23.3 \pm 2.0\%$, $n=4$; Figure 3E).

For the converse experiment, downregulation of Nanog function, we assayed the efficiency of shRNAs to induce partial *Nanog* silencing, since complete loss of Nanog function is incompatible with ESC pluripotency.[11] Introduction of *Nanog–shRNA* into *T-eGFP* ESCs resulted in a marked decrease in the levels of Nanog, as evaluated by RT-PCR (Figure 3D) and immunoblotting with specific antibodies (Figure 3C). In these conditions, cell colonies of undifferentiated ESC morphology formed, and no signs of endoderm differentiation were apparent (data not shown). However, the number of T^+ cells generated from EM progenitors expressing *Nanog–shRNA* more than doubled that of mock-transfected *T-eGFP* ESCs maintained under similar culture conditions ($60.4 \pm 5.1\%$ versus $23.3 \pm 2.0\%$, respectively, both $n=4$; compare Figure 3G and 3E). Together with the results of our gain-of-function experiments, these findings demonstrate that Nanog controls the dedifferentiation of EM progenitors to ESCs.

Positive regulation of *Nanog* expression by T and LIF/STAT3

We next investigated the mechanism by which *Nanog* expression is upregulated in EM progenitors. Based on our observations that *Nanog* expression is increased in cell populations with high levels of *T* expression (Figure 3A), we first asked whether T could induce *Nanog* expression. For this purpose, we used a doxycycline-inducible conditional expression system. ESCs expressing the tetracycline-inducible transcriptional activator (rtTA)[20] were transfected with a construct containing the full-length *T* cDNA driven by a tetracycline-responsive promoter (T-rtTA ESCs). In the absence of doxycycline, T-rtTA ESCs cultured in medium supplemented with LIF 1,000 U/ml formed colonies of undifferentiated ESCs, with a morphology indistinguishable from that of parental untransfected ESCs. Under these conditions, *T* expression in T-rtTA ESCs was negligible, demonstrating the tight regulation of the inducible system (Figure 3H). Upon addition of doxycycline to cultures of T-rtTA ESCs, *T* expression was induced progressively and became detectable by RT-PCR after 12 h of doxycycline induction (Figure 3H). T-rtTA ESCs cultured in the presence of doxycycline started to acquire the characteristic flattened morphology of differentiated colonies 48–72 h after doxycycline induction (not shown). Importantly, *Nanog* expression increased in T-rtTA ESCs after 12 h of doxycycline induction, paralleling that of *T* (Figure 3H). These results indicate that the expression of *Nanog* can be regulated

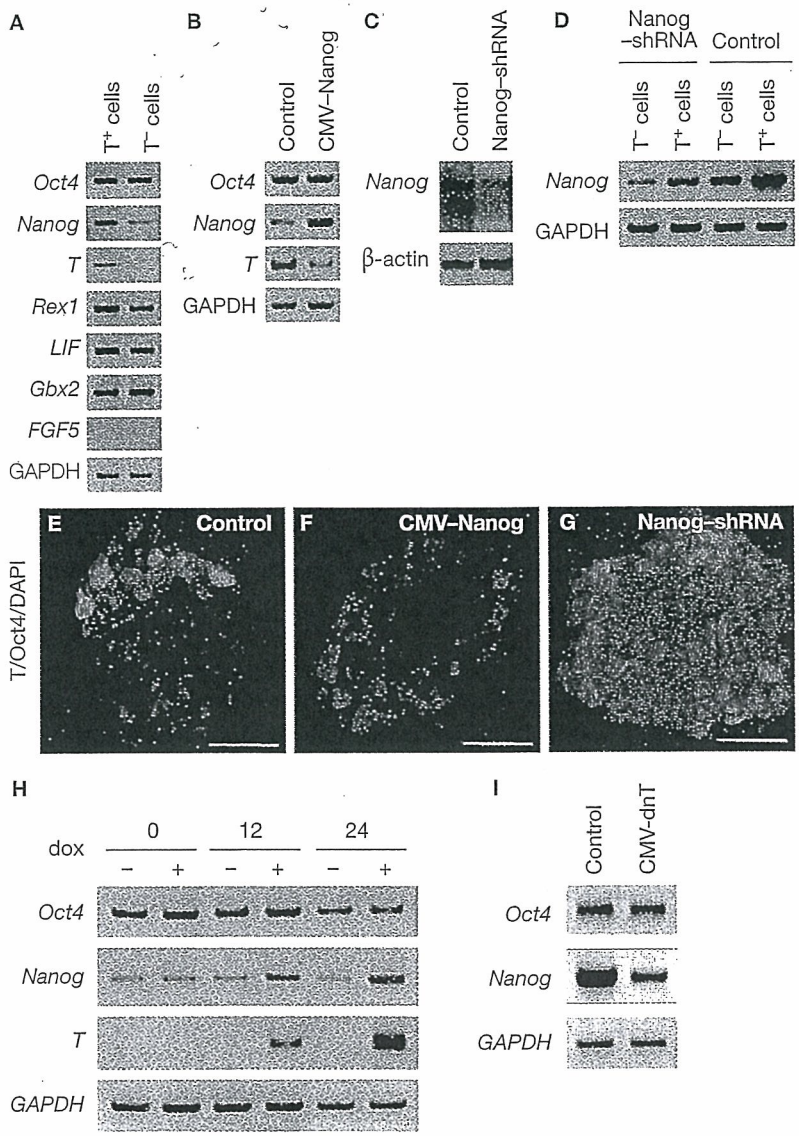

Figure 3 Nanog is sufficient and necessary for dedifferentiation of early mesoderm-specified progenitors into pluripotent stem cells. (**A**) RT-PCR analysis of gene expression in T^+ or T^- cells isolated from *T-eGFP* ES cells. T^+ cells had a higher level of *Nanog* expression. (**B**) RT-PCR analysis of *Oct4*, *Nanog*, and *T* expression in *T-eGFP* ES cells carrying control or CMV–*Nanog* constructs. Overexpression of *Nanog* is associated with downregulation of *T* expression. (**C**) Western blotting analysis of Nanog in *T-eGFP* ES cells carrying control or *Nanog–shRNA* constructs. *Nanog–shRNA* efficiently downregulates the level of Nanog protein in ES cells. (**D**) RT-PCR analysis of *Nanog* expression in T^+ or T^- cells isolated from *T-eGFP* ES cells carrying control or *Nanog–shRNA* constructs. (**E–G**) Fluorescent images of *T* (eGFP)/Oct4 expression in representative colonies formed from single T^+ cells carrying control (E), CMV–*Nanog* (F), or *Nanog–shRNA* (G) constructs. *T-eGFP* ES cells were transfected with each construct, then T^+ cells were isolated, selected with puromycin, and cultured for 7 days with leukemia inhibitory factor (LIF) 400 U/ml. 4',6-Diamidino-2-phenylindole staining identified individual nuclei in each field. Scale bar, 20 μm. (**H**) RT-PCR analysis of *Oct4*, *Nanog*, and *T* expression in doxycycline-dependent *T*-inducible ES cells cultured with LIF 1,000 U/ml. *T* expression was induced by doxycycline, and *Nanog* upregulation was associated with an increased level of *T*. Numbers at the top indicate induction hours with doxycycline. (**I**) RT-PCR analysis of *Oct4* and *Nanog* expression in *T-eGFP* ES cells carrying control or CMV–*dnT* constructs. dnT downregulated *Nanog* expression without affecting *Oct4* expression in ES cells cultured with LIF 400 U/ml.

CMV, cytomegalovirus; DAPI, 4',6-diamidino-2-phenylindole; dox, doxycycline; eGFP, enhanced green fluorescent protein; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; RT, reverse transcription; shRNA, short hairpin RNA; T-eGFP, transgenic ES cell line expressing eGFP under the regulatory regions of the mouse *T* gene.

by T and suggest that the upregulation of *Nanog* expression found in EM progenitors may depend on their increased levels of *T* expression.

To test this possibility, we analyzed the consequences of blocking the function of T in ESCs in the expression of *Nanog* and the generation of EM progenitors. For this purpose, we used a truncated version of T previously shown to function as a dominant-negative (dnT).[21] *T-eGFP* ESCs stably expressing *dnT* formed colonies with undifferentiated ESC morphology in culture medium containing LIF 400 U/ml (data not shown). The level of *Nanog* expression in these cells, however, was downregulated in comparison with that of mock-transfected *T-eGFP* cells under similar culture conditions (Figure 3I). Moreover, the blockade of T function resulted in an impaired transition of EM progenitors to ESCs (46.7 ± 2.6% of T^+ cells versus 22.1 ± 1.5% in control conditions, both $n=4$), consistent with the reduced levels of *Nanog* expression found in these conditions. Thus, our results from gain-of-function and loss-of-function experiments identify a negative feedback mechanism by which increased *T* expression in EM progenitors upregulates the expression of *Nanog*, which, in turn, downregulates *T* expression and promotes the regeneration of an ESC phenotype.

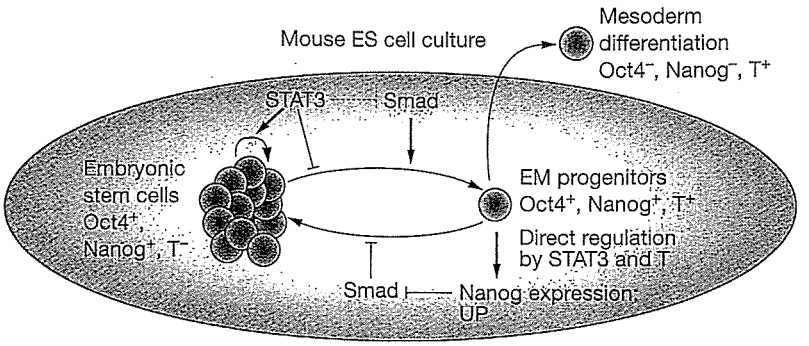

Figure 4 Maintenance of embryonic stem cell pluripotency by a negative feedback involving Nanog and T. Schematic diagram outlining the genetic relationships identified in this study and the proposed mechanism of Nanog action for the maintenance of pluripotent mouse embryonic stem cells. Embryonic stem cells that initiate mesoderm differentiation express the early mesoderm marker *T*. When leukemia inhibitory factor (LIF) is present, T upregulates the expression of *Nanog*, which, in turn, provides negative feedback to downregulate *T* expression and eventually leads to the regeneration of embryonic stem cells from early mesoderm-specified progenitors. In the absence of LIF signaling, T is not sufficient to upregulate *Nanog* expression, and mesoderm differentiation proceeds. The relevance of this negative feedback mechanism for maintaining the pluripotency of embryonic stem cells is evident in long-term cultures. Thus, the size of the early mesoderm-specified population cannot be maintained in embryonic stem cells in which either Nanog or T function is experimentally downregulated. In such conditions, the cultures progressively accumulate early mesoderm-specified progenitors and eventually lose pluripotency.
EM, early mesoderm-specified; ES cell, embryonic stem cell.

DISCUSSION

Our studies identify a population of *T*-expressing cells normally present in cultures of mouse ESCs; we have termed these T-expressing cells 'early mesoderm-specified (EM) progenitors'. The expression of *T* is generally considered to be a marker of early mesodermal fate during embryo development, as well as during ESC differentiation. However, the significance of *T* expression in ESC cultures is not straightforward, for T itself does not confer mesoderm identity and is dispensable for mesoderm formation.[22] Nevertheless, we show that EM progenitors are specified to mesoderm fates, as evidenced by their differentiation to mesoderm, but not endoderm or ectoderm, lineages, upon LIF withdrawal *in vitro* (Figure 2A,C,D). It is formally possible that EM progenitors could also give rise to endoderm lineages *in vitro* under specific culture conditions (serum-free medium supplemented with >30 ng/ml of activin A or hepatocyte-differentiation medium[23]), although we have not explored this possibility. In the presence of LIF, however, EM progenitors are phenotypically indistinguishable from ESCs, as the two populations coexist in colonies of undifferentiated morphology and both maintain pluripotency over extended periods of time in culture. Moreover, EM progenitors maintain the expression of pluripotency-associated markers such as *Oct4*, *Nanog*, and *Rex1* and have high levels of alkaline phosphatase activity (Figures 1B, 3A, and data not shown). Indeed, in the presence of LIF, EM progenitors and ESCs interchange their identities at a rate that depends, precisely, on the amount of LIF. In this sense, the regeneration of a pluripotent ESC phenotype from EM progenitors is reminiscent of the reversion to ESCs of early primitive ectoderm-like (EPL) cells.[24] Unlike EPL cells, however, EM progenitors display expression levels of *Rex1* and *Gbx2* comparable to those of ESCs and do not express *Fgf5* (Figure 3A).

Our results also show that, in the presence of LIF, *Nanog* overexpression is sufficient to accelerate the transition of EM progenitors to ESCs (Figure 3F). More importantly, downregulation of Nanog function results in impaired transition of EM progenitors to ESCs (Figure 3G). Thus, we identify Nanog as a critical component of a negative feedback mechanism that blocks the progression of ESC differentiation toward mesoderm