Functional Analysis of Sox15 and Sox2

Masayoshi Maruyama

Nara Institute of Science and Technology, Center for Genetic Information and Education, Laboratory of Animal Molecular Technology

(Ph.D. Professor Shinya Yamanaka)

2005.01.31

バイオサイエンス研究科　博士論文要旨

所属 （主指導教員）	遺伝子教育研究センター　動物分子工学 （山中　伸弥　教授）		
氏名	丸山　昌良	提出	平成 17 年　1 月　6 日
題目	Functional analysis of Sox15 and Sox2		

<Background>

Embryonic stem (ES) cells are derived from inner cell mass of blastocysts and proliferate infinitely while maintaining pluripotency, the ability to differentiate into all kinds of cells existed in mammalian body. The rapid growth and pluripotency made human ES cells attractive sources for cell therapy to patients with degenerative diseases. We have shown that a translational regulator NAT1 is essential for pluripotency. NAT1 is thought to regulate translation of specific proteins involved in pluripotency. One of the candidates is Sox2, a Sry-related transcription factor. Sox2 has been shown to play essential roles in the maintenance of pluripotency. In addition to Sox2, we found that another Sry-related factor Sox15 is also enriched in mouse ES cells.

<Purpose>

In the current study, I tried to answer the following three questions:

(1) Is NAT1 involved in pluripotency in non-mammalian organisms? As the first step to answer this question, I tried to identify NAT1 orthologs in lower organisms.

(2) What is the function of Sox15? To this end, I disrupted the mouse Sox15 gene by homologous recombination and compared phenotypes with those of Sox2-null mutants.

(3) What are the downstream target genes of Sox2? I was especially interested in whether Sox proteins are involved in the expression control of the ES cell specific homeoprotein Nanog, since disruption of *Nanog* or *Sox2* both resulted in loss of pluripotency.

<Results and Discussion >

(1) Is NAT1 involved in pluripotency in non-mammalian organisms?

In order to identify NAT orthologs in lower organisms, I searched genomic databases with Blast. Putative NAT1 orthologs were found in *Drosophila melanogaster*, *Halocynthia roretzi*, zebrafish, frog and chicken, but not in yeast or C. elegans. Yeast two hybrid analyses showed that the putative Drosophila ortholog binds to eukaryotic translation initiation (eIF) 4A, but not to eIF4E, as is the case with mammalian NAT1. Surprisingly, all of the NAT1 orthologs utilize non-AUG translation initiation codons. We have now studied roles of these NAT1 orthologs in development and cell differentiation.

(2) What is the function of Sox15?

Our laboratory has identified several genes that are highly enriched in mouse ES cells and early embryos. One of them encodes Sox15. RT-PCR confirmed that *Sox15* expression is highest in undifferentiated ES cells and repressed upon differentiation. We compared functions of Sox15 and Sox2. SELEX analyses showed that Sox15 binds to similar DNA sequences to those of Sox2. By using gel mobility shift assay and luciferase reporter assay, we found that Sox15 resembled Sox2 in that it binds the Fgf4 and Fbx15 enhancers and synergistically activates them with Oct3/4. To elucidate the physiological functions of Sox15, we established *Sox15*-/- mice and ES cells by gene targeting. *Sox15*-deficient mice were normal in gross appearance and fertile. *Sox15*-null ES cells a normal in morphology and the expression of *Fgf4* and *Fbx15*, two know targets of Sox2, suggesting that its function is largely compensated by Sox2. To identity target genes of Sox15, which is not compensated by Sox2, we performed DNA microarray analysis. Comparison between *Sox15*-heterozygous ES cells and *Sox15*-null cells showed that the expression of *Otx2*, *Ctgf* and *Ebaf* were significantly decreased, whereas *Hrc* was increased in *Sox15*-null ES cells. RT-PCR analyses confirmed these changes. The expression of these genes was normalized by introduction of Sox15 cDNA into null ES cells. Gel-mobility shift assay showed that Sox2 bound to the Sox binding site of *Hrc* more tightly than Sox15 *in vitro*. In great contrast, chromatin immunoprecipitation demonstrated that only Sox15 bound to the same site *in vivo*. These data demonstrated that Sox15 plays a specific role in gene expression in mouse ES cells, but the gross effect of its gene disruption is not noticeable in normal condition.

(3) What are the downstream target genes of Sox2?

Nanog is a key transcriptional factor underlying pluripotency in both ICM and ES cells. The null phenotypes of Nanog are nearly identical to those of Sox2. Previously, an upstream distal enhancer (DE) fragment was identified to be responsible for the specific expression of *Nanog* in ES cells. However, little is known about the cis-element and trans-factors required for the DE activity. In this study, I identified cis-elements in DE and a novel proximal enhancer (PE) of the mouse *Nanog* gene. Using the luciferase reporter assay, I showed both DE and PE are necessary for specific expression in ES cells. Gel mobility shift assay and chromatin immunoprecipitation show that DE was specifically bound by Sox2 only, whereas PE was bound by Oct3/4 and Sox2. Nanog expression was significantly reduced in delayed blastocysts that were deficient in *Sox2*. These results demonstrated that Nanog is synergistically regulated by Oct3/4 and Sox2, as reported previously with *FGF4*, *Fbx15* and *UTF1*.

INDEX

1. INTRODUCTION

Embryonic stem (ES) cells are derived from inner cell mass of blastocysts and proliferate infinitely while maintaining pluripotency, the ability to differentiate into all kinds of cells existed in mammalian body [1]. Mouse ES cells were first established in 1981 and led to the development of knockout mouse technology [2,3]. Pluripotenct stem cells were subsequently generated from human blastocysts in 1998 [4]. The rapid growth and pluripotency make human ES cells attractive sources for cell therapy to patients with degenerative diseases, such as diabetes and Parkinson's disease.

Molecular mechanisms underlying self-renewal of mouse ES cells and embryonic carcinoma (EC) cells have been targets of extensive basic researches [5,6]. In contrast to somatic stem cells in which self-renewal is established by asymmetrical cell division, self-renewal of ES and EC cells is maintained by symmetrical division that produces two identical pluripotent daughter cells. Several transcription factors have been identified to be essential for pluripotency, including Oct3/4 [7,8], Nanog [9,10], and Sox2 [11].

In addition, we have shown that the translational regulator NAT1 is essential for pluripotency. NAT1 might be involved in translation control of the aforementioned pluripotency-associated transcription factors.

1.1 Analyses of evolutional conservation of NAT1

NAT1 [12] is a mammalian protein that shows ~30% identity with the eukaryotic translation initiation factor (eIF) 4G [13], which functions as an essential scaffold protein for various proteins involved in translation initiation, such as eIF4A, eIF4E and eIF3 [14,15]. NAT1 is similar to the C-terminal two thirds of eIF4G [16,17,18]. NAT1 binds to eIF4A and eIF3, but not to eIF4E [12,19]. It has been reported that NAT1 is involved in translation control of various transcription factors and signaling molecules.

By gene targeting strategy, we have shown that NAT1 is essential for early mouse development and ES cell differentiation [20]. NAT1-deficient embryos die due to failure of gastrulation. NAT1-deficient ES cells can not differentiate properly by various induction methods. These data indicate

that NAT1 may control translation of factors that are involved in cell differentiation, such as the Sox protein family.

In this study, I tried to study whether NAT1 plays similar roles in pluripotency in non-mammalian species. As a first step, I tried to isolate NAT1 orthologs from various species.

1.2 Functional analyses of Sox15

The Sox family proteins are transcription factors related to the mammalian testis-determining factor Sry, which is located on the Y chromosome of mouse and man [21,22]. Sox proteins are characterized with the highly conserved high mobility group (HMG) domains that consist of 79 amino acids and are involved in DNA recognition and binding [23]. The HMG domains are also involved in association with partner proteins [24,25].

According to the sequence similarity outside the HMG domain and the genomic organization, Sox proteins are grouped into A to H subgroups [26]. Analysis of expression patterns revealed that the expression of some Sox

genes is cell-specific and restricted to definite developmental stages, whereas other genes are ubiquitously expressed. Furthermore, the expressions of different Sox genes overlap in many cell types and tissues [27]. Gene knockouts in mouse and the identification of human mutations have demonstrated essential roles of Sox family members in developmental process, including sex differentiation, chondrogenesis, gliogenesis, B-cell development, and lens development [27,28,29].

Sox2, which belongs to group B, was isolated as a Sox protein specific for EC cell. Northern blot analyses showed the highest expression of Sox2 in mouse EC and ES cells. These data indicated that Sox2 is the sole Sox protein expressed in mouse EC and ES cells. Sox2 forms a complex with the POU family transcription factor Oct3/4 on the enhancer of the fibroblast growth factor 4 gene, which is expressed in ICM and essential for the survival of post-implantation mouse embryo [30]. Similar complex formation was demonstrated on the enhancers of other target genes, such as UTF1 [31] and Fbx15 [32]. In addition, both Oct3/4 [33] and Sox2 [34] are

auto-regulated by the Oct3/4-Sox2 complex. Sox2-null embryos are lethal at peri-implantation stages, and Sox2-null ES cells could not be established [11], indicating that Sox2 and its association with Oct3/4 are essential for maintenance of pluripotency.

For better understanding of self-renewal of mouse ES cells, we analyzed expressed sequence tag (EST) databases with digital differential display and identified more than twenty genes that are highly enriched in early mouse embryos and ES cells[10,25,35]. One of them encodes Sox15 [36], suggesting that it is the second Sox protein expressed in ES cells. Sox15 is a single member of group G, but its HMG domain is closely related to the group B Sox proteins including Sox2 [37]. Its expression has been reported in muscle [38] and testis [39]. In the current study, I compared functions of Sox15 and Sox2 in mouse ES cells. In addition I generated Sox15-null mice to determine its physiological functions.

1.3 Transcriptional hierarchy among Sox2, Oct3/4 and Nanog

Fertilized eggs are totipotent in that they can produce complete animals if transferred to properly conditioned uterus. Totipotency is maintained in each blastomere at least untill 4-cell stage embryos. A single blastomere in 8-cell stage embryos can not produce complete animal, but chimera experiments have shown that it keeps the potential to differentiate into all embryonic and extraembryonic cells.

Subsequently totipotency is lost and cells become one of two lineages, namely trophectoderm or inner cell mass (ICM). Trophectoderm is an extraembryonic lineage that will differentiate cells in placenta. Cells in ICM are the founder of all cells in embryos and some cells in extraembryonic tissues, but they can not differentiate into trophectoderm in normal condition. Therefore, cells in ICM are pluripotent, but not totipotent. Cells in ICM then take one of two fates: they either differentiate into primitive endoderm or remain pluripotent as primitive ectoderm (also known as epiblast).

Totipotency is lost during gastrulation when epiblast produces all three

germ layers. In addition, epiblast also produces primodial germ cells (PGC), which are founders of sperm and oocytes. PGCs are not pluripotent, since they can only differentiate into either sperm or oocytes. However, totipotency is fully recovered upon fertilization of sperm and oocytes. Eggs, ICM, epiblast, PGC and germ cells can be considered as a circuit that transmits totipotency to the next generation.

Pluripotent cell lines can be isolated from this circuit. Embryonic stem (ES) cells are derived from ICM or epiblast. They proliferate infinitely while maintaining pluripotency. Similar pluripotent cells can be generated from PGC and are designated embryonic germ (EG) cells. Recently, pluripotent cells were also derived from neonate mouse testis and are designated multipotent germ stem (mGS) cells [40].

Two transcription factors, Oct3/4 and Sox2, are expressed in all cells in the totipotency circuit, but not in somatic cells. Two exceptions are that their expression is faint in sperm and that Sox2 is also expressed in neural stem cells. Oct3/4 and Sox2 are also specifically expressed in ES, EG and mGS cell

lines. Gene targeting experiments have demonstrated that the two transcription factors are essential for pluripotency in peri-implantation embryos and ES cells.

The homeobox protein Nanog is also indispensable for pluripotency in both epiblast and ES cells. Overexpression of Nanog renders ES cells independent of LIF for self-renewal. This is in great contrast to Oct3/4 and Sox2 that can not induce LIF-independent maintenance of ES cells.

Another important difference between Nanog and the other two transcription factors is their expression patters. The expression of Nanog is limited in pluripotent cells, including ICM, epiblast, ES cells, EG cells and mGS cells. Nanog is not expressed in earlier embryos or germ cells including PGC.

The third purpose of my study is to understand the molecular mechanism underlying tightly regulated expression of Nanog. I found that Oct3/4 and Sox2 play essential roles in the expression of Nanog. I also found that the binding sites of Oct3/4 and Sox2 in the mouse Nanog gene are highly and

specifically methylated in germ cells.

2. MATERIAL AND METHODS

Identification of NAT1 orthologs from Xenopus and zebrafish

Xenopus laevis and zebrafish EST clones similar in sequence to mammalian NAT1 were identified using the BlastN algorithm and purchased from IncytoGenomics. Nucleotide sequences of these clones were determined by primer walk. Primers used for zebrafish NAT1A were zf-S1, zf-S2, zf-AS1, zf-S3, zf-S4, zf-AS3, zf-AS4, zf-Ai-S1, zNAT1-ai-2086, zNAT1-ai-S3092, zNAT1-ai-S3301, and ai-S3838. Primers used for zebrafish NAT1B were zf-S1, zf-S2, zf-AS1, zf-S3, zf-S4, zf-AS3, zf-AS4, zf-AW-AS1, and zf-AW-S1. Primers used for *Xenopus* NAT1 were XNAT1-S1, XNAT1-S2, XNAT1-AS1, and XNAT1-AS2.

Identification of a chicken NAT1 ortholog

A chicken cDNA sequence similar to that of mammalian NAT1 has been deposited with the GenBank/EMBL/DDBJ database (Accession No. AF093110). However, the reported sequence has multiple ambiguous

nucleotides and lacks open reading frames. Based on this information, we designed five sets of primers to amplify from embryonic total RNA overlapping fragments covering the chicken NAT1 cDNA. Primer pairs used were chNAT1-U29 /chNAT1-L636, chNAT1-U482 /chNAT1-L1135, chNAT1-U878/chNAT1-L1948, chNAT1-U1955/chNAT1-L2829, and chNAT1-U2767/chNAT1-L3552. First-strand cDNA was synthesized with Revertra Ace (Toyobo, Japan). PCR was performed with *Taq* polymerase (Toyobo) and consisted of initial denaturation at 94°C for 5 sec; 35 cycles at 94°C for 2 sec, 55°C for 2 sec, and 72°C for 1 min; and final extension at 72°C for 5 min. PCR products were subcloned into a pCR2.1 vector (Invitrogen) and sequenced with 21M13 primer and M13 reverse primer.

Western blot analyses

Total cellular lysate was isolated from mouse RF8 ES cells and *Xenopus* oocytes with M-Per (Pierce). *Xenopus* oocytes, chicken extract, and zebrafish extract were kind gifts from Drs. Toshiaki Tanaka, Chio Oka, and Kunio

Inoue, respectively. Lysates were separated on 5% SDS–polyacrylamide gels. Western blot using an anti-NAT1 polyclonal antibody was performed.

Site-specific mutagenesis of mouse NAT1 cDNA

Mutant cDNAs containing single nucleotide mutations in the vicinity of the GUG start codon were generated by PCR using primers containing the corresponding mutation. Sense primers used were U15wt, U15-3C, U15+4C, U15+5T, and U15+6A. The antisense primer used was L4. The PCR products were subcloned into pCR2.1, which contains a T7 promoter for in vitro transcription.

In vitro transcription and translation

The mutated NAT1 cDNAs and the NAT1-luciferase chimeric cDNAs were transcribed, translated, and labeled with [35S]methionine in vitro with a TNT Quick Coupled Transcription/Translation System (Promega, Madison, WI, USA) as previously described. The resulting polypeptides were

separated using SDS–PAGE and analyzed with an imaging plate scanner (BAS-5000, Fuji Film, Japan).

Identification of a NAT1 ortholog from H. roretzi

Two *H. roretzi* EST clones (175H10 and 11B12) similar in sequence to human NAT1 were identified using the Blast server in Magest (http://www.genome.ad.jp/magest/). An *Xho*I fragment of 175H10 was subcloned into 11B12 to obtain a full-length cDNA.

Identification of an NAT1 ortholog from D. melanogaster

We searched for *Drosophila* EST clones similar in sequence to mammalian NAT1 using the *Drosophila* Genome Project Blast Search (http://www.fruitfly.org/cgi-bin/blast/run_blast.pl). We obtained four EST clones (LD199465, LD25657, GH12286, SD01173) from Research Genetics.

To generate full-length dNAT1 cDNA (pOT2-dNAT1), an *Xho*I fragment of LD199465 was subcloned into the corresponding site in GH12286.

Yeast two-hybrid system

Protein interactions of *Drosophila* NAT1 were analyzed using Matchmaker GAL4 Two-Hybrid System 3 (Clontech, Palo Alto, CA, USA). The open reading frame of dNAT1was amplified by PCR with primers dNAT1-F1 and dNAT1-R1 and cloned into pGBKT7. Since the precise translation codon of dNAT1 had not been determined, we amplified the entire open reading frame including a putative 5′ untranslated region. *Drosophila* eIF4A and eIF4E were amplified by PCR with primers d4A-F-*Eco*RI and d4A-R-*Bam*HI for eIF4A, and d4E-F-*Eco*RI and d4E-R-*Bam*HI for eIF4E. Amplified products were cloned into pGADT7. This plasmid containing mouse NAT1 was introduced into AH109 yeast, which were cultured on SD/−His/−Leu/−Trp plates or SD/−Leu/−Trp plates.

ES cell lines and culture conditions

In this study, RF8 and MG1.19 ES cells were used. RF8 ES cells which used for gene targeting, production of knockout mice or some assays were maintained in Dulbecco's modified eagle medium (DMEM, Nacalai tesque) containing 15% fetal bovine serum (FBS, selected batches for ES cells, Hyclone), 1 x 10^{-4} M non-essential amino acids (Invitrogen), 2 mM L-glutamine (Invitrogen), 1 x 10^{-4} M 2-mercaptoethanol (Invitrogen) and 50U penicillin/streptomycin (Invitrogen). Most of cases, RF8 ES cells were cultured on mitomycin C-treated STO feeder layer. To maintain the pluripotency of ES cells in feeder-free condition, ES cells were plated on gelatin-coated tissue culture plates, and LIF was supplemented into the medium. MG1.19 ES cells were grown on gelatin-coated tissue culture plates in the medium containing LIF. MG1.19 ES cells are stably express the large T antigen of murine polyoma virus. When pCAG-IP vectors containing the replication origin of polyoma virusare introduced into these cells, they can be replicated without being integrated into chromosomes.

Vector constructions

Most of vectors were modified for using Gateway Technology (GIBCOBRL). Insert fragment were amplified by PCR, subcloned pENTR-D-TOPO (Invitrogen) and verified by sequencing. Information of vector constructions and primer sequences used in this study were described in Appendix 1.

RT-PCR

Total RNA was isolated with Trizol (Invitrogen). First-strand cDNA was synthesized from total RNA with ReverTra Ace (Toyobo, Japan), and PCR was carried out with Ex Taq DNA polymerase (Toyobo).

Immunoblotting and Immunoprecipitation

In all of experiments, cell lysates were collected at subconfluent density. Cells were washed once with PBS, lysed with M-PER (PIERCE) and then incubated for 15 minutes at 4°C. Cells were scraped, transferred to microtubes and centrifuged at 15000 rpm for 5 min. The supernatants were transferred to new microtubes. The concentrations of total protein were measured by Bradford method. The cleared lysates were added with one-fifth volume of 5 x SDS sample buffer and boiled at 100°C for 5 min. 50μg

of proteins were separated by SDS-polyaclylamide gel electrophoresis and transferred to PVDF membranes (Millipore). Then membranes were incubated in blocking solution for 1 hr at room temperature with gently agitation. Primary antibodies were diluted at appropriate concentrations in fresh blocking buffer and membranes were placed in the primary antibody solution overnight at 4°C. Next day, membranes were washed three times for 10 min each with TBST at room temperature with agitation. Then membranes were incubated in secondary antibody solution for 1 hr at room temperature with agitation. All of secondary antibodies used for immunoblotting in this study were conjugated with horse radish peroxydase. After incubation, membranes were washed three times for 10 min each with TBST at room temperature with agitation. Detection was performed using enhanced chemiluminescence (ECL, Amersham) reagents.

Gel mobility shift assay

pCAG-IP expression vector cloned coding region of each genes were introduced into Cos7 cells with Fugene6 (Roche). Cell extracts were collected with extract

buffer (20mM HEPES[pH7.8], 450mM NaCl, 0.4mM EDTA, 0.5mM dithiothreitol, 25% glycerol, 0.5mM phenylmethlylsulfonyl fluoride) supplemented with protease inhibitor cocktail and then dialyzed to remove. F9 embryonic carcinoma cell extract was purchased from Funakoshi. The DNA probes used in this assay were complementary double-stranded DNA oligonucleotides. To generate complementary double stranded DNA, 1nmol sense strand, 1mM antisense strand, 3μl 10×annealing buffer (200 mM Tris-HCl [pH 7.5], 100 mM MgCl2, 500 mM NaCl), 7μl distilled water were combined in a tube and incubated at 85°C for 5min. The tube was cooled down at room temperature. Binding reactions were performed performed as described by Dent and Latchman[41], expect that 0.5% Tris-borate-EDTA buffer was used for electrophoresis, gel were prerun for 5 min, and poly(dG-dC) was used instead of poly(dI-dC).

Luciferase assay

0.8μg of each reporter constructs were transfected into RF8, MG1.19 ES cells, Cos7 cells or NIH3T3 cells seeded on 24-well culture plates along with 0.025 μg of pRL-TK (Promega) by using Lipofectamine 2000. Twenty-four

hours after transfection, cells were lysed and measured the luciferase activities by using Dual-reporter assay system (Promega) as a manufacturer's protocol.

SELEX

The coding region of the mouse Sox15 and Sox2 genes was introduced into pIH1119 to produce a fusion protein consisting of MBP (maltose binding protein) and Sox15 or Sox2. The fusion was induced in *E. coli* BL21AI (invitrogen) and was purified with amylose-beads (New England Biolads).SELEX was performed as follows. Double strand oligonucleotide was synthesized in reaction mixture containing 1μg SELEX-N20-Oligo, 1μg SELEX-N20RV primer, 2.5μl 10×Klenow buffer, 4μl 10mM dNTP, and 14.7μl distilled water. The mixture was denatured at 95 °C for 5min and annealed at 54°C for 5min. Klenow fragment (0.225U in 3.6μl of 1×Klenow buffer) was added to the mixture, which was then incubated at 25°C for 40min. The double-strand DNA was purified by phenol/chloroform extraction and ethanol precipitation and resuspended at 4°C for 30min in reaction mixture

containing of 20μl DNA, 30μl MBP-Sox15 or Sox2-bound beads, 20μl of 5×binding buffer (100mM HEPES-KOH[pH7.9], 1mM EDTA, 1M KCl, and 50% Glycerol), and 30μl of water. The beads were washed six times with 100μl lysis buffer (20mM Tris-HCl[pH7.4], 200mM NaCl, and 1mM EDTA) supplemented with 1mM DTT and 0.2mM PMSF. DNA was purified by phenol/chloroform extraction and ethanol precipitation and resuspended in 20μl distilled water, 5μl of this was used for PCR amplification in reaction buffer containing 5μl of 10×ExTaq buffer, 4μl of 2.5mM dNTP, 1μl of 30μM selex-N20FW primer, 1μl of 30μM selex-N20RV primer, 0.25μl of ExTaq polymerase, and 33.75μl of water. PCR program consisted of the initial denaturation at 95°C for 1min, 20cycles of 95°C for 15 sec, 54°C for 10 sec, and 72°C for 10 sec, and the final extension at 72°C for 5 min. Amplified products were purified by Micro-bio spin column (BioRad) and concentrated to 20μl ethanol precipitation, of which half was applied to another selection round. This procedure was repeated five times for enrichment. Amplified products in the final reaction were ligated into pCR2.1 vector (Invitrogen).

Fourty-five clones were randomly selected and sequenced.

Generation of anti-Sox15 and anti-Sox2 antibodies

The C terminal regions of Sox15 and Sox2, which are devoid of the HMG domains, were amplified by PCR. The primers used were anti-Sox15-s and anti-Sox15-as for Sox15 and anti-Sox2-s band anti-Sox2-as for Sox2. PCR products were cloned into pENTR/D-TOPO (Invitrogen) to construct pENTR-Sox15C and pENTER-Sox2C. These "entry" vectors were recombined with the "destination" vector pDEST17 for the expression of N-terminally histidine-tagged proteins in E. coli, The resulting expression vectors, pDEST17-Sox15C and pDEST17-Sox2C, were introduced into BL21-AI E. coli (Invirogen). Purification of histidine-tagged proteins was performed with Ni-nitrilotriacetic acid agarose (Qiagen) under denaturing conditions with 8M urea. After being dialyzed in 6M urea, the recombinant proteins were injected into New Zealand White rabbits to generate anti-Sox15 or anti-Sox2 serum.

Construction of Sox15 targeting vectors

To disrupt the mouse Sox15 gene, a cassette carrying the internal ribosome entry site (IRES) and a fusion of the b-galactosidase and neomycin resistance genes (bgeo) was inserted into the single exon of the gene, upstream of the HMG domain [42]. A 1.5-kbp 5' arm of the targeting vector was amplified by the Expand long-template PCR system (Roche) with primers sox15-5arm-s-NotI and sox15-5arm-as-SpeI. A 3.9-kbp 3' arm was amplified with primers sox15-3arm-s-BamHI and sox15-3arm-as-XhoI. The IRES-ßgeo cassette was ligated between the two PCR fragments. A DTA (diphtheria toxin A) cassette was placed downstream of the 3' arm. The resulting targeting vector was linearized with *SacII* digestion and introduced into RF8 ES cells by electroporation [43]. Genomic DNAs from G418-resistant colonies were screened for homologous recombination by Southern blot analyses.

Construction of Sox2 targeting vectors

To disrupt the mouse Sox2 gene, a cassette carrying the internal ribosome entry site (IRES) and a fusion of the b-galactosidase and neomycin resistance

genes (bgeo) was inserted into the Sox2 coding region. A 3.8-kbp 5' arm of the targeting vector was amplified by the Expand long-template PCR system (Roche) with primers sox2-5arm-s-NotI and sox2-5arm-as-NheI. A 1.9-kbp 3' arm was amplified with primers sox2-3arm-s-BamHI and sox2-3arm-as-XhoI. The resulting targeting vector was linearized with *NsiI* and introduced into RF8 ES cells by electroporation [42]. Genomic DNAs from G418-resistant colonies were screened for homologous recombination by Southern blot analyses.

Mice

All of mice used in this study were maintained in specific pathogen free area. To generate mice carrying Sox15 and Sox2 heterozygous or null mutation, Sox15 $^{ß\text{-}geo/+}$ or Sox2 $^{ß\text{-}geo/+}$ RF8 ES cells were injected to blastocysts (C57BL6) and then embryos were transferred to the uterus of pseudo pregnant female mice at 3.5 dpc. High percentage chimeric mice (judged by coat color) were mated with C57BL6 female mice. The genotypes of littermate were confirmed by southern blotting or PCR.

Genotyping of ES cells and mice for Sox15

For 5' recombination, genomic DNA was digested with *EcoRI*, separated on a 1% agarose gel, and transferred to nylon membrane. A 550-bp 5' probe was amplified with sox15-5'south-s and sox15-5'southern-as. Hybridization with this probe resulted in a 8.0-kbp band from the wild-type locus and an 4.3-kbp band from the targeted locus. For 3' recombination, genomic DNA was digested with *MluI*. An 870-bp 3' probe was amplified with sox15-3'south-s and sox15-3'south-as. Hybridization with this probe resulted in an 18.0-kbp bnd from the wild-type locus and an 8.8-kbp band from the targeted locus. After identifying ES cell clones that were correctly targeted, Genotypes of mice and ES cells were determined with three-primer PCR. A sense primer, ßgeo-screening1, was designed from ßgeo cassette to amplify the targeted locus. Another sense primer, sox15-3'-m-tail-s, was designed from HMG domain to amplify the wild-type locus. An antisense primer, sox15-3'-m-tail-as, was designed to amplify both the wild-type and targeted loci. Amplification with these three primers produced a 1036-bp band from

the wild-type locus and a 1371-bp band from the targeted locus.

Genotyping of ES cells and mice for Sox2

For 5' recombination, genomic DNA was digested with *EcoRI*, separated on a 1% agarose gel, and transferred to nylon membrane. A 527-bp 5' probe was amplified with sox2-5'south-s and sox2-5'southern-as. Hybridization with this probe resulted in a 15.7-kbp band from the wild-type locus and an 6.8-kbp band from the targeted locus. For 3' recombination, genomic DNA was digested with *EcoRI*. An 607-bp 3' probe was amplified with sox2-3'south-s and sox2-3'south-as. Hybridization with this probe resulted in an 15.7-kbp bnd from the wild-type locus and an 6.1-kbp band from the targeted locus. After identifying ES cell clones that were correctly targeted, Genotypes of mice and ES cells were determined with three-primer PCR. A sense primer, ßgeo-screening1, was designed from ßgeo cassette to amplify the targeted locus. Another sense primer, sox2-3'-m-tail-s, was designed from HMG domain to amplify the wild-type locus. An antisense primer, sox2-3'-m-tail-as, was designed to amplify both the wild-type and targeted

loci. Amplification with these three primers produced a 611-bp band from the wild-type locus and a 945-bp band from the targeted locus

Generation of Sox15-null and rescued ES cells

ES cells deficient in Sox15 were obtained by culturing the heterozygous ES cells with high concentration (6mg/ml) of G418 [44]. To obtained rescue cells, pCAG-IP-Sox15 was transferred into Sox15-deficient ES cells by electroporation. To identify clone expressing Sox15, we screened colonies resistant to 2μg/ml puromycine by northern blot and western blot analysis with anti-Sox15 antibody.

In situ hybridization

For hybridization, embryo was dissected and washed twice in cold PBS. After fixation for 5 min in 4% paraformaldehyde in PBS. Sample were washed twice in PBS containing 0.1% Tween 20 at RT, fixed again in 0.1% glutaraldehyde in PBT for 20 min at RT, and washed again five times in PBT. Thereafter, the sample were incubated for 5 min at RT in a 1:1 mixture of PBT and hybridization buffer (50 % formamid, 5×SSC [pH4.0], 0.1 %Tween

20, 50 μg/ml heparin) containing 10μg/ml t-RNA, and 100 μg/ml sheared denatured herring sperm DNA. Samples were prehybridized for 1 h in hybridization buffer at 70°C. After quickly adding 0.5 μg DIG-labeled cRNA probe in 100 μl hybridization buffer previously denatured at 95°C, hybridization was carried out overnight at 70°C. To remove unbound cRNA, three 30-min washing steps were performed at 65°C in SSC-FC buffer (2×SSC[pH4.5], 50% formamide, 0.1% Tween 20) followed by washing for 10 min at 70°C in hybridization buffer. Samples were cool down and washed again three times in TBST, blocked by incubating for 1 h in 10% heat inactivated sheep serum in TBST. After overnight incubation at 4°C with anti-DIG alkaline phosphatase antibody diluted to 1:2,000 in TBST with 1% sheep serum, unbound antibody was washed with TBST. Four 30 min washes of TBST at RT were followed by four 10 min washing steps in APB(100 mM Tris[pH 9.5], 100 mM NaCl, 50 mM MgCl2, 0.1% Tween 20) at RT. To visualize alkaline phosphatase activity, samples were stained in a solution of 4.5μl/ml NBT and 3.5μl/ml BCIP in APB and incubate in the dark

at RT for up to 48 h. To stop the reaction, samples were transferred to PBS.

DNA microarray

Total RNA from Sox15 heterozygous ES cells and homozygous ES cells were labeled with Cy3 and Cy5 hybridized to Mouse Development Microarray (Algilent) according to the manufacturer's protocol. The arrays were scanned with a Agilent G2565BA Microarray Scanner System(Agilent). Hybridization was repeated with different clones. Data were analyzed with GeneSprings (Silico Genetics).

Chromatin immunoprecipitation (CHIP) assays

Formaldehyde was added directly to the culture medium to a final concentration of 1 %(v/v) and the dishes were gently shaken on a shaker at room temperature for 8 min. Glycine was added to 125 mM, and the dishes were returned to the shaker for an additional 5 min. Dishes were washed three times with ice-cold PBS and harvested by scraping into 3 ml of cold PBS. Cells from dishes were collected by centrifugation at 2000 rpm for 5

min at 4 °C, and the supernatants were discarded. Cell pellets were resuspended in 10 ml of cold PBS plus 200μM phenylmethylsulfonylfluoride. Cells were collected by centrifugation as above and resuspended in 5 ml of an ice cold solution containing 5 mM HEPES, pH 8.0, 85 mM KCl, 0.5% (v/v) Nonidet P-40 plus protease inhibitors (200 μM phenylmethylsulfonyl fluoride, 1.4 μg/ml pepstatin, 1 μg/ml leupeptin). Samples were allowed to swell on ice for 10 min before homogenization with three strokes of a glass Dounce homogenizer to release nuclei. Nuclei were collected by centrifugation as above, and the supernatants were discarded. Nuclei were resuspended in 50 mM Tris, pH 7.6, 10 mM EDTA, 1% SDS plus protease inhibitors as above (50 μl/15-cm dish of cells). Samples were incubated on ice for 10 min, and immunoprecipitation dilution buffer A (0.01% SDS, 1.1% (v/v) Triton X-100, 1.2 mM EDTA, 16.7 mM Tris, pH 7.6, 167 mM NaCl) was added to bring the volume to 750 μl. Samples were sonicated 30 sec ON and 1 min OFF using a Bioruptor(Cosmo bio). Routinely, ten repetitions of the sonication were performed. Chromatin samples (500 μl) were first

precleared with normal mouse IgG (5 µl) in the presence of protein G-Sepharose () bead slurry (60 µl of a 50/50 slurry of beads in TBS (16.7 mM Tris, pH 7.6, 167 mM NaCl), 1 mg/ml bovine serum albumin, 200 µg/ml salmon sperm DNA). Samples were incubated for 2 h at 4 °C on a rotator, and beads were collected by centrifugation in a microcentrifuge at 2000 rpm. The unbound material (chromatin) was transferred to a new tube and 5 µg of each antibody were added. Samples were incubated overnight at 4 °C on a rotator, 60 µl of blocked protein G slurry were added, and incubation was continued on the rotator for an additional 2 h at 4 °C. Beads were collected by centrifugation in a microcentrifuge at 12000 rpm×1min. Beads were first washed twice with 500 µl of ice-cold buffer B (0.05% (w/v) SDS, 1% (v/v), Triton X-100, 20 mM Tris, pH 7.6, 2 mM EDTA, 150 mM NaCl), washed once sequentially with buffer D (0.05% (w/v) SDS, 1% (v/v) Triton X-100, 20mM Tris, pH 7.6, 2 mM EDTA, 500 mM NaCl), buffer 3 (0.25 M LiCl, 1.0% (v/v) Nonidet P-40, 1.0% deoxycholate, 10 mM Tris, pH 7.6, 1 mM EDTA), and buffer C (0.1% (v/v) Triton X-100, 150 mM NaCl, 20 mM Tris, pH 7.6, 2 mM

EDTA). Beads were transferred to a microtube, and bound material was eluted by incubating the beads with 75 µl of elution buffer (0.1 mM sodium bicarbonate, 1.0% (w/v) SDS) while vortexing for 10 min. This was repeated a total of four times. The four eluates were pooled, and NaCl and RNase were added to 300 mM and 10 µg, respectively, and samples were heated at 65°C for 6 h to reverse the Schiff's base linkage. DNA was collected by ethanol precipitation and diluted a total of 20 µl.

3. RESULTS

3.1 Identification of non-mammalian orthologs of NAT1

To isolate non-mammalian NAT1 orthologs, we searched the GenBank/EMBL/DDBJ databases using the BlastN algorithm. In the non-redundant nucleotide database, we identified a chicken cDNA sequence (accession number: AF093110) that was similar to mouse NAT1. However, the sequence contained several ambiguous nucleotides and did not contain any long open reading frames, suggesting that it was not an accurate representation. Using this sequence information, we designed five sets of primers that amplified overlapping fragments that covered the entire sequence. By assembling the sequences of these five PCR products, we obtained a cDNA fragment of 3329 nt containing a single open reading frame[2]. The nucleotide sequence of the chicken NAT1 cDNA was 81% identical with its mouse counterpart (Table 1).

By searching the EST database, we found one Xenopus (accession number: BE679183) and two zebrafish EST clones (accession numbers:

AI588745 & AW133753) that shared a high degree of similarity to the 5' portion of the mouse NAT1 sequence. We obtained these clones and sequenced the entire DNA fragment. The Xenopus NAT1 cDNA was 3831-nt long and 72% identical to the mouse sequence (Table 1). The two zebrafish clones were 66% identical to each other and are believed to have originated by gene duplication. One clone (AI588745) was 4030-nt long and 64% identical to mouse NAT1, whereas the other one (AW133753) was 4030-nt long and shared 65% identity with mouse NAT1. We designated them zebrafish NAT1A and NAT1B.

We then analyzed the non-mammalian NAT1 orthologs, which all contained single, long open reading frames. In all species, the GUG codon that functions as the translation initiation codon in mammals was identified near the 5' end of each of the presumptive coding regions, which were conserved among all sequences examined (Figure 1A). The predicted amino acid sequences determined by the open reading frames in the mouse and non-mammalian cDNAs were highly conserved, apart from the first 20 to 30

amino acids (Figure 1B). Beginning at the 5' end of the conserved region for each gene, the GUG triplet was the first possible non-AUG initiation codon. In contrast, the first AUG codon in the open reading frame was located approximately 700-nt downstream from the GUG codon. These data, taken together, indicated that the GUG triplet functions as the translation initiation codon in all species.

To demonstrate the GUG-mediated initiation in non-vertebrate species, we performed Western blot analyses. If the GUG codon was used as the start codon, the predicted molecular weights of the frog and the two zebrafish NAT1 orthologs were 101.8, 101.2, and 102.2 kDa, respectively. These are approximately equal to the molecular weight of human NAT1 (102.1 kDa). Chicken NAT1 was predicted to be slightly smaller (97.5 kDa). If the downstream AUG was used as an initiation codon, the protein product was predicted to be ~80 kDa. Since the predicted amino acid sequences of the non-mammalian NAT1s were very similar to that of their mouse counterpart (Table 2), we reasoned that a polyclonal □-NAT1 antibody generated against

mammalian NAT1 should be able to detect non-vertebrate orthologs. This was indeed the case; the antibody recognized frog and zebrafish NAT1 as bands of approximately the same size as mouse NAT1, and chicken NAT1 as a protein of slightly smaller size (Figure 2). No products as small as ~80 kDa were detected.

We next studied whether the GUG codons in the non-mammalian NAT1 orthologs are used as translation initiation codons *in vitro*. We located the cDNAs of mouse, chicken, frog and zebrafish NAT1 downstream to the T7 promoter. We then transcribed and translated these constructs *in vitro*. From all the constructs, translated products of predicted sizes were obtained (Figure 3). This result further supports the notion that the GUG codon is the start codon in the non-mammalian NAT1 orthologs.

3.2 A role of Sox15 in ES cells and development

Expression pattern of the Sox15 gene

To identify candidates of ES cell-specific genes, digital differential display was used to compare Gene representation between EST libraries derived from mouse ES cells and those from various somatic tissues. Several genes appeared exclusive in ES cell-derived libraries. One of them encodes Sox15. This is surprising because Sox2 had been considered to be the sole Sox factor family expressed in mouse ES cells.

Sox15 had been reported to be expressed in testis and muscle. Its expression in ES cells had not been studied. To determine whether Sox15 is indeed expressed predominantly in ES cells, RT-PCR assay was carried out with mSox15F1 and mSox15R1 primers. A 69-bp fragment was amplified in two independent ES cell lines, RF8 and MG1.19 (Figure 4A).

To confirm specific expression in ES cells at the protein level, polyclonal antibody against the C-terminal portion (amino acids 125 to 231) of mouse Sox15 was generated as described in Materials and Methods. Western blot

analysis with the antiserum detected a single band of 25 KDa in ES cells, but not in Cos7 cells (Figure 4B).

To further verify specific expression of Sox15 protein, ES cells were stained with anti-Sox15 or anti-Sox2 antibody. Figure 4C shows that ES cells are stained by Cy3-conjugated secondary antibody. In addition, anti-Sox15 staining was almost restricted to nucleus, in consistent with its predicted role as a transcription factor. These data showed that Sox15 gene is predominantly expressed in the nucleus of undifferentiated ES cells.

Interaction between Sox15 and Oct3/4

It is thought that Sox15 plays similar roles to Sox2, because Sox15 was expressed in undifferentiated ES cells. To examine whether Sox15 interacts to Oct3/4, myc-tagged Oct3/4 was introduced into MG1.19 mouse ES cells together with either HA-tagged Sox2 or HA-Sox15. Myc-Oct3/4 protein was precipitated with anti-myc antibody and examined whether HA-Sox proteins were co-precipitated (Figure 5). Western blot analysis with anti-HA antibody detected co-precipitation of both HA-Sox2 and HA-Sox15 with myc-Oct3/4.

When HA-Sox proteins alone were introduced into cells, they were not precipitated by anti-myc antibody. These data showed that both Sox15 and Sox2 interact to Oct3/4 in mouse ES cells.

DNA binding capacity of Sox15

I performed gel mobility-shift assay to study whether Sox15 binds to the enhancer of the mouse Fgf4 gene, a target gene of Sox2 (Figure 6). The enhancer contains the consensus octamer motif (ATTAGCAT) and Sox binding site (AACAAAG). Sox2 and Oct3/4 had been shown to bind synergistically on the Fgf4 enhancer. When ^{32}P-labled Fgf4 enhancer fragment was incubated with Cos7 cell extracts expressing Oct3/4, Sox2 or both, shifted bands corresponding to an Oct3/4 monomer, a Sox2 monomer and an Oct3/4-Sox2 complex were observed, respectively. When ^{32}P-labled Fgf4 probe was incubated with Cos7 cell extracts expressing Sox15, a shifted band was also observed. When Cos7 cell extracts expressing both Sox15 and Oct3/4 were used, the shifted band migrated more slowly than that with Oct3/4 alone did, demonstrating that Sox15 and Oct3/4 synergistically bind

to the Fgf4 promoter. The band corresponding to a Sox15 monomer is less intense than that corresponding to a Sox2 monomer, suggesting weaker affinity of Sox15 to the Fgf4 enhancer. In contrast, the bands corresponding to the Oct3/4-Sox15 dimer showed similar intensity to that corresponding to the Oct3/4-Sox2 dimer. This suggests that synergism with Oct3/4 compensates the weaker affinity of Sox15.

Next, enhancer of mouse Fbx15 gene was examined (Figure 7). Fbx15 enhancer contains the canonical Sox binding site and a motif (TTTATCAT) two nucleotides different from the octamer consensus. Oct3/4 can bind to this site only with the help of Sox2. I found that Oct3/4 alone or Sox15 alone barely binds to the Fbx15 enhancer, in consistent with the weaker affinity of Sox15 than Sox2. However, the Oct3/4-Sox15 complex can bind to the Fbx15 enhancer as effectively as the Oct3/4-Sox2 complex. These data support the notion that synergism with Oct3/4 compensates the weaker affinity of Sox15.

Transcriptional activation of FGF4 and Fbx15 by Sox15

To study whether Sox15 can activate the Fgf4 and Fbx15 enhancers

synergistically with Oct3/4, I performed luciferase reporter assays using reporter genes containing either the Fgf4 enhancer or the Fbx15 enhancer (Figure 8). These reporter genes were introduced into Cos7 cells together with expression vectors of Oct3/4 and/or Sox proteins. When both Sox15 and Oct3/4 were introduced, the Fgf4 enhancer was activated ~4 fold. Sox15 or Oct3/4 alone did not show such effect. However, the enhancement by Sox15 and Oct3/4 was weaker than that by Sox2 and Oct3/4, which showed ~7-fold enhancement. Similar result was obtained with the Fbx15 enhancer. These data indicated that Sox15 synergistically activates the Fgf4 and Fbx15 enhancers with Oct3/4, but with less potency than Sox2 does, in consistent with its weaker affinity to the enhancers.

Differential binding potency of Sox15-Sox2 chimeric proteins

I pursued why Sox15 showed lower affinity to the FGF4 and Fbx15 enhancers. To this end, I constructed expression vectors for chimeric proteins in which the HMG domains of Sox2 and Sox15 were exchanged. Sox2-15-2 contains the HMG domain of Sox15 in the background of Sox2. Sox15-2-15 is

the opposite combination.

These constructs were introduced into Cos7 cells and gel mobility shift assay was performed with these Cos7 cell extracts and either the Fgf4 enhancer or the Fbx15 enhancer (Figure 9). With both enhancers, we found that Sox2-15-2 showed weaker affinity than Sox2 did, whereas Sox15-2-15 showed stronger affinity than Sox15 did. When Oct3/4-expressing extracts were included in the reaction, the differences between these Sox protein became smaller. These data indicate that the weaker affinity of Sox15 to the Fgf4 and Fbx15 enhancer is, at least in part, attributable to the HMG domains.

Comparison of the binding sequences of Sox15 and Sox2

The finding that the Sox15 is less potent than Sox2 in binding to the Fgf4 and Fbx15 enhancers suggests that the recognition sequence of Sox15 might be different from that of Sox2. To test this possibility, SELEX analysis was performed (Figure 10A). Affinity purification columns with recombinant Maltose binding protein (MBP)-Sox15 or MBP-Sox2 were prepared.

Oligonucleotides with random sequences were purified on these affinity columns, amplified by PCR, and re-applied to the affinity column. This procedure was repeated five times and the final elution was subcloned and sequenced. This analysis showed that Sox2 binds to AACAATG, as expected. Sox15 also preferentially binds to nearly identical sequences. These data suggest that the lower potency of Sox15 is attributable to lower affinity to the binding sequences, but not to different preferences in recognition sequences.

To further characterize the DNA recognition sequences of Sox15 and Sox2 , I performed gel mobility shift assays with a series of probes in which each nucleotide of the consensus Sox binding sequence (AACAATG) was mautated to one of the other three nucleotides (Figure 10B). When incubated with these probes, Sox2 and Sox15 showed similar preferences. The analyses showed that A at position 4 and C at position 5 are critical, while positions 1, 2 and 7 are less strict required. These data confirm that Sox15 shares the same binding sequences with Sox2.

Identification of an ES cell-specific enhacer in the mouse Sox15 gene

To identity regulatory regions of the mouse Sox15 gene, we isolated overlapping DNA fragments from ~20 kbp flanking region and subcloned them in pGV-P vector in which the luciferase cDNA is driven by the minimum SV40 promoter. We found that a region containing the second exon showed a high enhancer activity in undifferentiated ES cells, but not in NIH3T3 cell (Figure 11). We then tested smaller fragments from this region and was able to narrow down the enhancer to a 40bp fragment in second exon (Figure 12). Computer-assisted analysis did not identify putative binding sites for known transcription factors in this sequence (Figure 13A). When a 32p-labeled Sox15 enhancer probe was incubated with undifferentiated ES cell extract, a shifted band with strong signal was observed (Figure 13B). The band was not observed with Cos7 extracts. These results demonstrate that unknown factor(s) expressed highly in undifferentiated ES cell binds to the Sox15 enhancer.

Analyses of Sox15 knock-out mice

To study functions of Sox15 in development, its mouse gene was inactivated by homologous recombination. The Sox15 gene consists of two exons. I constructed a targeting vector in which the HMG domain was replaced with a cassette consisting of IRES (internal ribosome entry site) and b-geo (a fusion of ß-galactocidase and the neomycin resistant gene) (Figure 14 A). The vector was introduced into RF8 ES cells by electroporation. Screening of 96 G418-resistant clones identified three positive clones by both PCR and Southern blot analyses (Figure 14B, C). One of the positive clones was injected into blastocysts of C57/BL6 mice and germ line transmission was obtained. Sox15-null mice were born with the expected Mendelian ratio (+/+ : +/- : -/- =30 : 57 :25). They were normal in gross appearance and fertile, as recently reported. This result suggests that Sox15 is dispensable for normal development, probably because its function is compensated by Sox2 or other Sox family transcription factors.

Function of Sox15 in ES cells

To study the function of Sox15 in ES cells, homologous mutant ES cells

were established by selecting heterozygous cells with high concentrations of G418. Among 48 colonies obtained with 2〜3mg/ml of G418, four were found homozygous for Sox15 deletion by PCR and Southern blot analyses. Northern and Western blot analyses demonstrated that Sox15 transcripts and proteins were absent in these clones (Figure 14D, E). Sox15-null ES cells were normal in morphology and proliferation when maintained undifferentiated on STO feeder cells (Figure 15A, B). They were also competent in differentiation after LIF removal and retinoic acid treatment. Northern blot analyses showed that the expression level of Sox/Oct target genes, such as Fgf4, UTF1 and Fbx15 were indistinguishable between wild-type, Sox15-heterozygous, and Sox15-null ES cells (Figure 15C). These data demonstrated that the function of Sox15 is largely compensated by Sox2 in ES cells.

Generation of Sox2/Sox15 double knockout mice

Sox15-/- homozygous mutant mice showed no gross developmental defects and fertile. This result is raised a possibility that Sox2 compensates Sox15

function. To test this hypothesis, we crossed Sox2-heterozygous mice and Sox15-homozygous mice. Sox2+/- Sox15-/- mice developed normaly and were born at the ratio expected from the Mendelian law (Figure 16).

Expression of Sox15 in mouse embryo

To study the expression pattern of Sox15 during development, I performed in situ hybridization (Figure 17). Sox15 RNA was first detected at the morula stage [2.5 days postcoitum (dpc)]. At the blastocyst stage (3.5 dpc), the Sox15 was detected specifically in ICM. Expression persists throughout epiblast and extraembryonic ectoderm of 6.5 dpc and 7.5 dpc embryos. However, Sox15 expression was lower than that of Sox2.

Identification of Sox15 target genes

To identity target genes of Sox15, I performed DNA microarray analyses with Agilent Mouse Development arrays that contain ~20,000 genes expressed in early embryos. Comparison between Sox15-heterozygous ES cells and Sox15-null cells showed that the expression of Otx2 [45], Ctgf [46] and Ebaf [47] was significantly decreased, whereas Hrc [48] was increased in

Sox15-null ES cells. RT-PCR analyses confirmed these changes (Figure 18B). The changes were rescued by introduction of Sox15 cDNA into Sox15-null ES cells. However, Fgf4 and Fbx15 were normally expressed in Sox15-null ES cells, in consistent with the result of northern blot analyses (Figure 18A). In addition, ES cell specific genes, such as Nanog and Oct3/4, were also normally expressed in Sox15-deficient cells.

Binding of Sox15 to the Hrc gene

To study whether Sox15 directly regulates the Hrc expression, I performed chromatin immunoprecipitation assay. I found five putative Sox binding sites (R1 ~ R5) around the mouse Hrc gene (Figure 19A). We found the R5 fragment was precipitated with anti-Sox15 antibody in wild-type ES cells, but not in Sox15-/- cells (Figure 19B). This site was not precipitated with anti-Sox2 antibody. I also performed gel mobility-shift assay with a probe containing R5 sequence (Figure 19C). When this probe was incubated with Cos7 cell extracts expressing Sox2 or Sox15, shifted bands corresponding to a Sox2 monomer or a Sox15 monomer were obtained, respectively. These data

showed that Sox2 can bind to the R5 sequence *in vitro*, but not *in vivo*.

3.3 Transcriptional hierarchy between Oct3/4, Sox2 and Nanog

Binding of Sox2 to the Nanog distal enhancer (DE)

We have previously demonstrated that that an ES-cell specific transcriptional regulatory region of the mouse Nanog gene is located at approximately 5kb upstream from the transcriptional initiation site (nucleotide -4737~ -4386, Figure 20). Computer-assisted analysis indicated that the enhancer contains putative STAT3, TCF, Nanog binding sites. Point mutation of STAT binding site (TTCCTAGAA) to TTCCTAGTC significantly reduced enhancer activity. This indicated that STAT may be involved in enhancer activity of Nanog.

I studied whether STAT3 directly binds to the putative STAT binding site

in Nanog enhancer. To this end, I performed gel mobility-shift assay with a probe containing the putative STAT binding site of DE (Figure 21A). To determine the position of bands corresponding to STAT3, APRE (STAT3 high affinity binding site) oligonucleotides was incubated with nuclear extracts of Cos7 cells expressing STAT3.

When the Nanog DE probe was incubated with undifferentiated ES cell extracts, an intense shifted band was observed. However, the position of this band was different from that corresponding to STAT3, as observed in the control lane.

The Nanog DE probe contains a putative TCF binding site in adjacent to the putative STAT binding site. I performed gel mobility shift assay with probes in which the putative STAT binding site, the putative TCF binding site, or both were mutated. The shifted band was abolished with any of the three mutations (Figure 21B).

I noticed that the shift band observed with the Nanog DE probe was similar to the band that was observed when Sox2 was incubated with the

FGF4 enhancer probe. To confirm that the shift was caused by the binding of Sox2, I performed super-shift assay with anti-Sox2 antibody (Figure 22). The antibody impaired the generation of the slow-migrating band. This indicates that Sox2 indeed binds to the Nanog DE sequence.

Binding of Sox2 to non-canonical binding site in DE

DE does not contain the canonical Sox binding site. However, a sequence AAGACAAAG, which overlaps the putative STAT and TCF binding sites, is similar to the canonical Sox binding sequence (AAACAAAG). The three aforementioned mutations change this sequence. These data indicate that Sox2 binds to the non-canonical sequence in DE. To test this possibility, I performed gel mobility-shift assay with probes in which the 5', middle or 3' part of the putative STAT binding site were mutated (Figure 23). We found that the Sox2 binding was abolished by the 3' end mutation, but not the 5'-end or middle part mutations. These data support the notion that Sox2 binds to the non-canonical sequence in DE.

Sox binding site in human Nanog enhancer

DE is conserved in the human Nanog gene. However, the Sox binding site in DE was only partially conserved in the human Nanog gene (Figure 24A). To study whether Sox2 can bind to the human sequence, I performed, gel mobility-shift assay (Figure 24B). However, no shifted band appeared with the human probe.

Interplay of Sox and POU binding site in proximal enhancer (PE)

Considering the similarity of Sox2-, Oct3/4- and Nanog-knockout mice phenotypes [9, 12, 14], Oct3/4 may play a role in Nanog expression. Thus, we searched for putative POU binding sites in the 5' flanking regulatory region from –5047 to+72. A candidate octamer motif, TTTTGCAT, was found at approximately 180 nt upstream from the transcription initiation site (Figure 25A). We also found a canonical Sox binding site, TACAATG, in adjacent to the octamer sequence. These sequences are also conserved in the human Nanog gene (Figure 25B).

To confirm that Oct3/4 and Sox2 can bind to these two cis-elements, I

performed gel mobility-shift assay (Figure 26A). When a ^{32}P-labled probe containing the two cis-elements was incubated with undifferentiated ES cell extract and Cos7 cell extracts expressing Oct3/4, Sox2 or both, shifted bands corresponding to an Oct3/4 monomer, a Sox2 monomer and an Oct3/4-Sox2 complex were observed, respectively.

To confirm that this band is reflected the simultaneous binding of Sox2 and Oct3/4 to the probe, I performed super-shift assay with anti-Sox2 antibody (Figure 26B). Sox2 or/and Oct3/4 was expressed in Cos7 cells, and the binding reaction, including the whole cell extract prepared from the transfected cells, was incubated with anti-Sox2 antibody before electrophoresis. As shown in Figure 26B, the Sox2 antibody impaired the generation of the slow-migrating band. This indicates that Sox2 indeed bound to the sequence. A similar result was obtained in the case of anti-Oct3/4 antibody (not shown). These data indicated that Oct3/4 and Sox2 synergistically bind to the proximal region, which we will refer the proximal enhancer (PE).

Sox2 and Oct3/4 binding *in vivo*

To confirm the binding of Sox2 and Oct3/4 to DE and PE *in vivo*, I performed chromatin immunoprecipitation assay. As shown in Figure 27, anti-Sox2 antibody precipitated the Sox binding site in DE and PE, whereas anti-Oct3/4 antibody precipitated PE. Other sites were not precipitated with both antibodies. This result confirms the *in vivo* binding of the two transcription factors to DE and PE.

Activation of the Nanog enhancers by Sox2 and Oct3/4

To evaluate the roles of Sox2 and Oct3/4 in ES cell specific expression of the Nanog gene, we constructed reporter vectors containing the Nanog promoter and enhancers (Figure 28). The wild-type reporter gene showed strong luciferase activity in undifferentiated ES cells, but not in differentiated ES cells or NIH3T3 cell. Mutation in on of the Sox2 or/and Oct3/4 binding sites in PE or DE decrease the enhancer activity. We all of these binding sites were mutated, the enhancer activity nearly abolished. These results indicated that Sox2 and Oct3/4 in both PE and DE play important roles in

ES cell-specific expression of the mouse Nanog gene.

We then tested whether Sox2 and Oct3/4 are sufficient to induce c the Nanog expression in differentiated cells. To this end, the reporter gene was introduced into Cos7 cells together with the expression vectors of Sox2 or/and Oct3/4 (Figure 29). However, the reporter was inactive in Cos7 cells, indicating that Sox2 and Oct3/4 alone are not enough to induce the expression Nanog.

Generation of Sox2/Nanog double heterozyous mice

To better understand the role of Sox2 in the expression of Nanog, we crossed Nanog- and Sox2- heterozygous mice (Figure 30A). Nanog/Sox2 double heterozygous mice were born at a slightly lower ratio than expected from the Mendelian law (WT : Nanog+/- : Sox2+/- : Nanog/Sox2+/- = 32 : 23 : 25 :19).

Nanog expression in Sox2 knockout blastocyst

To better understand the role of Sox2 in the Nanog expression, I performed whole-mount in situ hybridization (WISH) in blastocyst from Sox2 heterozygous intercrossed. Since maternal Sox2 transcripts still persist at

3.5 dpc, we treated mice with tamoxifen and obtained delayed blastcysts at 8.5 dpc (Figure 30B). WISH with the Sox2 probe showed that Sox2 is negative in approximately one forth of embryos, indicating that the maternal transcripts disappeared by this time. With the Nanog probe, I observed low or no signals in several embryos. I genotyped these embryos by PCR and found they were Sox2-/-. These data demonstrated that Sox2 plays an important role in the Nanog expression.

Binding of Naong to the regulatory region of the mouse Nanog gene

Expression of endogenous Nanog was reduced when exogenous Nanog was expressed in ES cells. This result suggests the possibility that Nanog negatively auto-regulate itself. Thus, I performed chromatin immunoprecipitation assay (Figure 27). Anti-Nanog antibody precipitated a fragment located between DE and PE. This indicates that the Nanog binding site existed in the regulatory region of its own. Further experiments are required to identify the exact Nanog binding sequence and confirm the negative auto-regulation.

4. DISCUSSION

4.1 Identification of non-mammalian orthologs of NAT1

Evolutionarily conserved non-AUG initiation in NAT1 orthologs

Previously, we reported the use of GUG-mediated translation initiation in human, mouse, and rabbit NAT1 [12,19]. EST analyses indicated that this was also the case in chimpanzee, swine, horse and rat NAT1. We have now identified NAT1 orthologs in non-mammalian vertebrates including chicken, *Xenopus*, and zebrafish; in the latter, two distinct NAT1 orthologs were identified that are likely to have arisen by gene duplication. All of the ortholog cDNAs contained a single open reading frame, and the GUG codon that functions as the initiator in mammals was conserved in all the non-mammalian vertebrates. Moreover, the first AUG codon in the open reading frame was located too far downstream to account for the observed molecular weight in all species, which was shown by Western blot and *in vitro* transcription/translation to be nearly identical to that of the mouse protein. We therefore concluded that the GUG triplet was used as the

translation initiation codon not only in mammals, but also in non-mammalian vertebrates.

We inspected other reported cases of non-AUG initiation from an evolutionary point of view. We found that TEF-1 homologs have been reported in human (TEF-3, TEF-4, TEF-5) [49] and in chicken (TEF-1, TEF-3, TEF-5) [50,51]. Human TEF-3 and TEF-5 and chicken TEF-1, TEF-3, and TEF-5 all use AUA or AUU triplets as initiation codons. In contrast, human TEF-4 has a longer 5' terminus and utilizes an AUG start codon. By searching GenBank EST databases, we identified zebrafish TEF-1 and TEF-5 cDNA fragments, as well as a cDNA fragment of a *Xenopus* TEF-1 ortholog. Sequence comparison revealed that zebrafish orthologs also use AUU as an initiation codon, while an AUG codon exists in place of AUU in *Xenopus* TEF-1. Thus, AUG, AUU, and AUA appear to be interchangeable as initiation codons in TEF homologs. The strict usage of GUG initiation codons in vertebrate NAT1 is therefore exceptional.

Why has non-AUG translation initiation been conserved through evolution

in NAT1? We found that the overexpression of NAT1 in ES cells led to spontaneous differentiation, and that NAT1 protein levels declined upon differentiation of ES cells due to post-transcriptional regulation (unpublished observation). These results indicate that the NAT1 protein level must be tightly regulated for proper cellular function. The GUG initiation codon may be important for the maintenance of the proper level of NAT1 protein.

4.2 Functional analyses of Sox15

Sox2 was previously thought to be the sole Sox protein expressed in mouse ES and EC cells [29]. However, we now report that Sox15 is also expressed in mouse ES cells. Our study demonstrated that Sox15 and Sox2 regulate different sets of genes *in vivo*, despite similar protein-protein interactions and DNA recognition *in vitro*.

Both gel mobility shift assays and SELEX analyses showed that Sox2 and Sox15 share the same DNA binding sequence. Sox15 can bind the enhancers

of Sox2 target genes including Fgf4 and Fbx15, albeit with decreased affinity. Like Sox2, Sox15 also binds DNA synergistically with Oct3/4, and similar binding was seen with Sox15/Oct3/4 complexes as with Sox2/Oct3/4 complexes. Taken together, these *in vitro* biochemical data indicate that Sox2 and Sox15 behave similarly with respect to DNA binding.

In stark contrast to the early embryonic lethality observed in Sox2 null animals, mouse embryos deficient in Sox15 implanted and developed completely normally in all aspects examined. The expression of Sox2 target genes including Fgf4, Fbx15 and Nanog was not affected by Sox15 deficiency. Chromatin immunoprecipitation assays showed that Sox15 does not bind to the Sox binding site of the mouse Nanog gene. Thus, despite similar *in vitro* DNA bind properties, Sox2 and Sox15 clearly regulate different target genes *in vivo*.

DNA microarray analyses revealed that expression of Otx2, Ctgf and Ebaf was decreased while expression of Hrc was increased in Sox15-null ES cells. Chromatin immunoprecipitation assays showed that the Sox binding site of

the Hrc gene is bound by Sox15, but not by Sox2 *in vivo*. In contrast, gel mobility shift assay showed that both Sox2 and Sox15 bind to Hrc. These data further highlight the differences between *in vivo* and *in vitro* experiments and demonstrate that Sox15 and Sox2 regulate different sets of target genes in mouse ES cells.

The mechanisms by which Sox proteins control gene expression *in vivo* are unclear. Sox15 and Sox2 recognize and bind similar DNA sequences, and the binding of both is enhanced by the presence of Oct3/4 binding. This is not wholly unexpected since the HMG domains, which bind to DNA and Oct3/4, are 78% identical between the two Sox proteins. In contrast, the identity outside the HMG domains is less than 30%. Some Sox proteins bind to other proteins through non-HMG domains [24]. For example, the non-HMG motif PLNLSSR is required for binding of Sox6 to the co-repressor CtBP2 [52]. Sox15 and Sox2 may bind to different transcription regulators through non-HMG domains and therefore regulate unique sets of genes.

During the preparation of this manuscript, the generation of Sox15-null

mice were reported[53]. Consistent with our data, these Sox15-deficient mice were normal in development, gross appearance and fertility. Histological analysis revealed normal myofiber ultrastructure and the presence of comparable amounts of satellite cells in the skeletal muscle compared to wild-type animals. The authors propose that other Sox family transcription factors likely compensated for any gross deficiencies resulting from Sox15 deletion.

Cultured Sox15-null myoblasts displayed a marked delay in differentiation *in vitro* [53]. Moreover, skeletal muscle regeneration in Sox15-null mice was attenuated after application of a crush injury. These results suggest a requirement for Sox15 in the myogenic program. Expression of the early myogenic regulated factors MyoD and Myf5 was altered in Sox15-deficient myoblasts with MyoD being downregulated and Myf5 expression increased. These results are consistent with our data and suggest a specific role for Sox15 that cannot be compensated by other Sox family members.

Our data demonstrated that Sox15 is dispensable for normal mouse

development, sex determination and fertility. This is surprising given the essential roles of other Sox family proteins in development [26,27]. Sox15 is the sole member of Sox family group G, and it is only found in mammals[38]. Other organisms including Fugu, Drosophila melanogaster and C. elegans do not have Sox15 orthologs. These data indicate that the Sox15 gene evolved relatively recently.

4.3 Roles of Sox2 and Oct3/4 in the Nanog expression

The third purpose of my study was to clarify the interplay between the three essential transcription factors, Sox2, Oct3/4 and Nanog [7,8,9,10,11]. Disruption of these factors all resulted in loss of pluripotent cells in ICM or epiblast. Thus, there may be interactions between the three factors. I found that the expression of Nanog is directly regulated by Sox2 and Oct3/4.

We previously identified a ~400 bp fragment located ~5 kb from the transcription initiation site is important for ES cell-specific expression of Nanog. Comparison between mouse and human genome sequences indicate

that the sequence is significantly conserved in the human Nanog gene. The enhancer contains a putative STAT binding site, thus I first hypothesized that STAT3 binds to this sequence.

However, I found that is not the case, and alternatively, Sox2 tightly binds to a non-canonical sequence in the enhancer. The binding of Sox2 was shown by both gel mobility shift assay and chromatin immuno-precipitation assay. Mutation of this sequence decreased the Nanog enhancer activity.

In addition, I found adjacent octamer motif and Sox binding site at ~150 bp upstream from the transcription initiation site. Gel mobility shift assay and chromatin immunoprecipitation showed that Oct3/4 and Sox2 bind to these sequences. Mutation of these sequences decreased the Nanog enhancer activity. Mutation of all the three sites abolished the enhancer activity.

In addition, I studied the expression of Nanog in embryos obtained from intercrosses of Sox2-heterozyous mice. Due to maternal expression, Sox2-null blastoysts contained significant amount of Sox2 transcripts. I therefore injected tamoxifen into mice and obtained delayed blastodysts at

7.5 dpc. In approximately one forth of these embryos, the expression of Nanog was severely impaired and these embryos turned out to be Sox2 homozyous by PCR analysis. These data, taken together, indicated that Sox2 and Oct3/4 plays important roles in the expression of Nanog.

Synergy between Sox2 and Oct3/4 was first documented in the regulation of Fgf4 [30]. Subsequently, similar mechanism was identified in UTF1, Fbx15, Sox2 itself and Oct3/4 itself [31,32,33,34]. In the case of the mouse osteopontin gene [54], Oct3/4 promotes the expression whereas Sox2 suppresses transcription. Thus, interaction of Sox2 and Oct3/4 is an important key to the transcription factor network in ES cells.

5. Acknowledgment

First, I sincerely and deeply appreciate to Ph.D. Professor Shinya Yamanaka for giving me the chance to study in institute. I learned lots of things from him. His help has been very important for my life. It was a great honor to study here under his teaching. I thank to Kaoru Mitsui, Masato Nakagawa, Koji Shimozaki, Keisuke Okita, Yumiko Sasaoka, Kazutoshi Takahashi, Yoshimi Tokuzawa, Mirei Murakami, and other members in Yamanaka lab for valuable discussion, Tomoko Ichisaka, Yukiko Ikeguchi, Junko Iida, Masako Shirasaka and Chihiro Takigawa for technical and administrative assistance, Dr. Hitoshi Niwa for MG1.19 ES cells and pCAG-IP. This work was supported in part by research grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan, The Uehara Memorial Foundation, The Naito Foundation, The Sumitomo Research Foundation, The Mitsubishi Foundation and Toray Science and Technology Grant (to S.Y.). This work was also supported in part by a Grant-in-Aid for 21st Century COE Research from the Ministry of Education, Culture, Sports, Science and

Technology.

At last, I thank to my parents to support my life.

6. REFERENCES

[1] A. G. Smith, EMBRYO-DERIVED STEM CELLS: Of Mice and Men, Annu Rev Cell Dev Biol 17 (2001) 435-62.

[2] M. J. Evans, M. H. Kaufman, Establishment in culture of pluripotential cells from mouse embryos, Nature 292 (1981) 154-6.

[3] G. R. Martin, Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells, Proc Natl Acad Sci U S A 78 (1981) 7634-8.

[4] J. A. Thomson, et al., Embryonic stem cell lines derived from human blastocysts, Science 282 (1998) 1145-7.

[5] I. Chambers, A. Smith, Self-renewal of teratocarcinoma and embryonic stem cells, Oncogene 23 (2004) 7150-60.

[6] M. Rao, Conserved and divergent paths that regulate self-renewal in mouse and human embryonic stem cells, Dev Biol 275 (2004) 269-286.

[7] J. Nichols, et al., Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4, Cell 95 (1998)

379-91.

[8] H. Niwa, J. Miyazaki, A. G. Smith, Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells, Nat Genet 24 (2000) 372-6.

[9] I. Chambers, et al., Functional expression cloning of nanog, a pluripotency sustaining factor in embryonic stem cells, Cell 113 (2003) 643-55.

[10] K. Mitsui, et al., The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells, Cell 113 (2003) 631-42.

[11] A. A. Avilion, et al., Multipotent cell lineages in early mouse development depend on SOX2 function, Genes Dev 17 (2003) 126-40.

[12] H. Imataka, H. S. Olsen, N. Sonenberg, A new translational regulator with homology to eukaryotic translation initiation factor 4G, EMBO J 16 (1997) 817-25.

[13]J. D. Shaughnessy, Jr, N. A. Jenkins, N. G. Copeland, cDNA cloning,

expression analysis, and chromosomal localization of a gene with high homology to wheat eIF-(iso)4F and mammalian eIF-4G, Genomics 39 (1997) 192-7.

[14] R. E. Thach, Cap recap: the involvement of eIF-4F in regulating gene expression, Cell (1992) 177–80.

[15] W.C.Merrick, J.W.B Hershey, The pathway and mechanism of eukaryotic protein synthesis. In Hershey,J.W.B., Mathews,M.B. and Sonenberg,N. (eds), *Translational Control*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1996) 31–70.

[16] M. W. Hentze, eIF4G: a multipurpose ribosome adapter? Science 275 (1997) 500–1.

[17] S. J. Morley, P. S. Curtis, V. M. Pain, eIF4G: translation's mystery factor begins to yield its secrets. RNA 3 (1997) 1085–104.

[18] A. Gingras, B. Brian, N. Sonenberg, eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. Annu. Rev. Biochem. 68 (1999) 913–63.

[19] S. Yamanaka, K. S. Poksay, K. S. Arnold, T. L. Innerarity, A novel translational repressor mRNA is edited extensively in livers containing tumors caused by the transgene expression of the apoB mRNA-editing enzyme. Genes Dev. 11 (1997) 321–333.

[20] S. Yamanaka, X. Y. Zhang, M. Maeda, K. Miura, S. Wang, R. V. Farese Jr, H. Iwao, T. L. Innerarity, Essential role of NAT1/p97/DAP5 in embryonic differentiation and the retinoic acid pathway, EMBO J 19 (2000) 5533-41

[21] A. H. Sinclair, et al., A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif, Nature 346 (1990) 240-4.

[22] J. Gubbay, et al., A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes, Nature 346 (1990) 245-50.

[23] V. R. Harley, R. Lovell-Badge, P. N. Goodfellow, Definition of a consensus DNA binding site for SRY, Nucleic Acids Res 22 (1994) 1500-1.

[24] M. Wilson, P. Koopman, Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators, Curr Opin Genet Dev 12 (2002) 441-6.

[25] Y. Kamachi, M. Uchikawa, H. Kondoh, Pairing SOX off: with partners in the regulation of embryonic development, Trends Genet 16 (2000) 182-7.

[26] J.Bowles, G Schepers, P Koopman, Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators, Dev Biol 277 (2000) 239-55

[27] M Wegner, From head to toes: the multiple facets of Sox proteins, Nucleic Acid Res 27 (1999) 1409-20

[28] G.E Shepers, R.D Teasdale, P Koopman, Twenty pairs of Sox: entext, homology, and nomenclature of the mouse and human Sxo transcription factor gene families, Dev Cell 2 (2002) 167-70

[29] L. H. Pevny, R. Lovell-Badge, Sox genes find their feet, Curr Opin Genet Dev 7 (1997) 338-44.

[30] H. Yuan, N. Corbi, C. Basilico, L. Dailey, Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3, Genes Dev 9 (1995) 2635-45.

[31] M. Nishimoto, A. Fukushima, A. Okuda, M. Muramatsu, The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2, Mol Cell Biol 19 (1999) 5453-65.

[32] Y. Tokuzawa, et al., Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development, Mol Cell Biol 23 (2003) 2699-708.

[33] S. Okumura-Nakanishi, M. Saito, H. Niwa, F. Ishikawa, Oct-3/4 and Sox2 regulate Oct3/4 gene in ES cells, J Biol Chem (2004).

[34] M. Tomioka, et al., Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex, Nucleic Acids Res 30 (2002) 3202-13.

[35] K. Takahashi, K. Mitsui, S. Yamanaka, Role of ERas in promoting

tumour-like properties in mouse embryonic stem cells, Nature 423 (2003) 541-5.

[36] M. van de Wetering, H. Clevers, Sox 15, a novel member of the murine Sox family of HMG box transcription factors, Nucleic Acids Res 21 (1993) 1669.

[37] P. Koopman, G. Schepers, S. Brenner, B. Venkatesh, Origin and diversity of the SOX transcription factor gene family: genome-wide analysis in Fugu rubripes, Gene 328 (2004) 177-86.

[38] F. Beranger, C. Mejean, B. Moniot, P. Berta, M. Vandromme, Muscle differentiation is antagonized by SOX15, a new member of the SOX protein family, J Biol Chem 275 (2000) 16103-9.

[39] M. A. Sarraj, H. P. Wilmore, P. J. McClive, A. H. Sinclair, Sox15 is up regulated in the embryonic mouse testis, Gene Expr Patterns 3 (2003) 413-7.

[40] M . Kanatsu-Shinohara, K. Inoue, J. Lee, M. Yoshimoto, N. Ogonuki, H. Miki, S. Baba, T. Kato, Y. Kazuki, S. Toyokuni, M. Toyoshima, O. Niwa,

M. Oshimura, T. Heike, T. Nakahata, F. Ishino, A. Ogura, T. Shinohara, Generation of pluripotent stem cells from neonatal mouse testis, Cell 29 (2004) 1001-12.

[41] C. L Dent, D.S Latchman, The DNA mobility shift assay, In D.S.Latchman (ed.), Transcription factors: a practical approach, Oxford University Press, New York, N.Y. (1993) 1-26.

[42] P. Mountford, et al., Dicistronic targeting constructs: reporters and modifiers of mammalian gene expression, Proc Natl Acad Sci U S A 91 (1994) 4303-7.

[43] V. L. Meiner, S Cases, H.M Myers, E. R Sande, S Bellosta, M Schambelan, R. E Pitas, J McGuire, J Herz, and R. V Farese, Jr, Disruption of the acyl-CoA: cholesterol acyltransfease gene in mice: evidence suggesting multiple cholesterol esterification enzymes in mammals, Proc Natl Acad Sci USA 93 (1996) 14041-6

[44] R. M. Mortensen, D. A. Conner, S. Chao, A. A. Geisterfer-Lowrance, J. G. Seidman, Production of homozygous mutant ES cells with a single

targeting construct, Mol Cell Biol 12 (1992) 2391-5.

[45] A. Simeone, D. Acampora, M. Gulisano, A. Stornaiuolo, E. Boncinelli, Nested expression domains of four homeobox genes in developing rostral brain, Nature 358 (1992) 687-90.

[46] R. P. Ryseck, H. Macdonald-Bravo, M. G. Mattei, R. Bravo, Structure, mapping, and expression of fisp-12, a growth factor-inducible gene encoding a secreted cysteine-rich protein, Cell Growth Differ 2 (1991) 225-33.

[47] C. Meno, et al., Two closely-related left-right asymmetrically expressed genes, lefty-1 and lefty-2: their distinct expression domains, chromosomal linkage and direct neuralizing activity in Xenopus embryos, Genes Cells 2 (1997) 513-24.

[48] A. G. Ridgeway, H. Petropoulos, A. Siu, J. K. Ball, I. S. Skerjanc, Cloning, tissue distribution, subcellular localization and overexpression of murine histidine-rich Ca2+ binding protein, FEBS Lett 456 (1999) 399-402

[49] P. Jacquemin, J. J. Hwang, J. A. Martial, P. Dolle, I. Davidson, A novel family of developmentally regulated mammalian transcription factors containing the TEA/ATTS DNA binding domain, J Biol Chem 271 (1996) 21775-85.

[50] A. F. Stewart, et al., Muscle-enriched TEF-1 isoforms bind M-CAT elements from muscle-specific promoters and differentially activate transcription, J Biol Chem 269 (1994) 3147-50.

[51] A. Azakie, S. B. Larkin, I. K. Farrance, G. Grenningloh, C. P. Ordahl, DTEF-1, a novel member of the transcription enhancer factor-1 (TEF-1) multigene family, J Biol Chem 271 (1996) 8260-5.

[52] R. P. Ryseck, H. Macdonald-Bravo, M. G. Mattei, R. Bravo, Structure, mapping, and expression of fisp-12, a growth factor-inducible gene encoding a secreted cysteine-rich protein, Cell Growth Differ 2 (1991) 225-33.

[53] H. J. Lee, et al., Sox15 is required for skeletal muscle regeneration, Mol Cell Biol 24 (2004) 8428-36

[54] V. Botquin, H. Hess, G. Fuhrmann, C. Anastassiadis, M. K. Gross, G. Vriend, H. R. Scholer, New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2, Genes Dev. 12 (1998) 2073-90.

Appemdix 1(Primer list)

Name	Sequence (5'-3')	Application
oct3/4gw-s	AAA AAG CAG GCT CCA CCT TCC CCA TGG CTG GAC ACC	Mouse Oct3/4 ORF cloning to pDONR201
oct3/4gw-as	AGA AAG CTG GGT TGA TCA ACA GCA TCA CTG AGC TTC	Mouse Oct3/4 ORF cloning to pDONR201
sox2-gw-s	AAA AAG CAG GCT TGT ATA ACA TGA TGG AGA CGG	Mouse Sox2 ORF cloning to pDONR201
sox2-gw-as	AGA AAG CTG GGT TTC ACA TGT GCG ACA GGG GCA GT	Mouse Sox2 ORF cloning to pDONR201
Sox15-gw-s	CAC CAT GGC GCT GAC CAG CTC CTC ACA A	Mouse Sox15 ORF cloning to pENTR/D-TOPO
Sox15-gw-as	TTA AAG GTG GGT TAC TGG CAT GGG	Mouse Sox15 ORF cloning to pENTR/D-TOPO
sox7-gw-S	CACCTCGGCCATGGCCTCGCTGCTGGG	Mouse Sox7 ORF cloning to pENTR/D-TOPO
sox7-gw-AS	CTCCATTCCTCCAGCTCTATGACACAC	Mouse Sox7 ORF cloning to pENTR/D-TOPO
Sox17-gw-s	CACCAGAGCCATGAGCAGCCCGGATG	Mouse Sox17 ORF cloning to pENTR/D-TOPO
Sox17-gw-as	CGTCAAATGTCGGGGTAGTTGCAATA	Mouse Sox17 ORF cloning to pENTR/D-TOPO
sox2HMG15-out-s	CCC GGC GTA AGA GCA AAA ACA CGC TCA TGA AGA AGG ATA A	C-terminal of Sox2 in Sox2-15-2 chimeric
sox2HMG15-out-as	ATG GGC CGC TTC ACC TTC TCC GGG CTG TTC TTC TGG TTG C	N-terminal of Sox2 in Sox2-15-2 chimeric
sox2HMG15-hmg-s	GCA ACC AGA AGA ACA GCC CGG AGA AGG TGA AGC GGC CCA T	HMG domain of Sox15 of Sox2-15-2 chimeric gene
sox2HMG15-hmg-as	TTA TCC TTC TTC ATG AGC GTG TTT TTG CTC TTA CGC CGG G	HMG domain of Sox15 of Sox2-15-2 chimeric gene
sox15HMG2-out-s	GGC CGC GGC GCA AAA CCA AGT CGA GCA CCG GGT CTG TCC C	C-terminal of Sox15 in Sox15-2-15 chimeric
sox15HMG2-out-as	ATG GGC CTC TTG ACG CGG TCC AGC GGA AGT CCC CCA GAC G	N-terminal of Sox15 in Sox15-2-15 chimeric
sox15HMG2-hmg-s	CGT CTG GGG GAC TTC CGC TGG ACC GCG TCA AGA GGC CCA T	HMG domain of Sox15 in Sox15-2-15 chimeric gene
sox15HMG2-hmg-as	GGG ACA GAC CCG GTG CTC GAC TTG GTT TTG CGC CGC GGC C	HMG domain of Sox15 in Sox15-2-15 chimeric gene
MyoD-s,gw	CAC CAT GGA GCT TCT ATC GCC GCC ACT C	Mouse MyoD ORF cloning to pENTR/D-TOPO
MyoD-as,gw	GCT GCA GTC GAT CTC TCA AAG CAC CTG A	Mouse MyoD ORF cloning to pENTR/D-TOPO
myogenin-s,gw	CAC CCT GAT GGA GCT GTA TGA GAC ATC C	Mouse Myogenin ORF cloning to pENTR/D-TOPO
myogenin-as,gw	CCA CTT AAA AGC CCC CTG CTA CAG AAG T	Mouse Myogenin ORF cloning to pENTR/D-TOPO
anti-sox15-s	CAC CTC GAG CAC CGG GTC TGT CCC C	C-terminal of Sox15 cloning to pENTR/D-TOPO
anti-sox15-as	TTA AAG GTG GGT TAC TGG CAT	C-terminal of Sox15 cloning to pENTR/D-TOPO
anti-sox2-s	CAC CCT CAT GAA GAA GGA TAA GTA C	C-terminal of Sox2 cloning to pENTR/D-TOPO
anti-sox2-as	TCA CAT GTG CGA CAG GGG CAG	C-terminal of Sox2 cloning to pENTR/D-TOPO
SELEX_N-20 FW	TAG GCA TGT GGA TCC GTC TGG C	SELEX
SELEX_N20 RV	ATC GAA GGT GGA TCC GGT ACG C	SELEX
SelexN20-Oligo	TAGGCATGTGGATCCGTCTGGCN20GCGTACCGGATCCACCTT	SELEX
fgf4, gel-s	TAGAAAACTCTTTGTTTGGATGCTAATGGGATACTTAAA	FGF4 probe using Gel mobilityshift assay
fgf4, gel-as	TTTAAGTATCCCATTAGCATCCAAACAAAGAGTTTTCTA	FGF4 probe using Gel mobilityshift assay
fgf4-gel-oct(m)-s	TAG AAA ACT CTT TGT TTG GGG TAA CCC GGG ATA CTT AAA	FGF4 probe(Oct mutation) using Gel mobilityshift assay
fgf4-gel-oct(m)-as	TTT AAG TAT CCC GGG TTA CCC CAA ACA AAG AGT TTT CTA	FGF4 probe(Oct mutation) using Gel mobilityshift assay
fgf4-gel-sox(m)-s	TAG AAA ACT ATT GGG TTG GAT GCT AAT GGG ATA CTT AAA	FGF4 probe(Sox mutation) using Gel mobilityshift assay
fgf4-gel-sox(m)-as	TTT AAG TAT CCC ATT AGC ATC CAA CCC AAT AGT TTT CTA	FGF4 probe(Sox mutation) using Gel mobilityshift assay
fgf4-gel-double(m)-s	TAG AAA ACT ATT GGG TTG GGG TAA CCC GGG ATA CTT AAA	FGF4 probe(Oct/Sox mutation) using Gel mobilityshift assay
fgf4-gel-double(m)-as	TTT AGT ATC CCG GGT TAC CCC AAC CCA ATA GTT TTC TA	FGF4 probe(Oct/Sox mutation) using Gel mobilityshift assay
fbx, gel, wt-s	CCA GAT GTG CTT TAT CAT AAC AAT GGA ATT CCT AGG GGC T	Fbx15 probe using Gel mobilityshift assay
fbx, gel, wt-as	AGC CCC TAG GAA TTC CAT TGT TAT GAT AAA GCA CAT CTG G	Fbx15 probe using Gel mobilityshift assay
fbx, gel, sox-m-s	CCA GAT GTG CTT TAT CAT AAC CAT GGA ATT CCT AGG GGC T	Fbx15 probe(Sox mutation) using Gel mobilityshift assay
fbx, gel, sox-m-as	AGC CCC TAG GAA TTC CAT GGT TAT GAT AAA GCA CAT CTG G	Fbx15 probe(Sox mutation) using Gel mobilityshift assay
fbx, gel, oct-m-s	CCA GAT GTG CTC CCT CAT AAC AAT GGA ATT CCT AGG GGC T	Fbx15 probe(Oct mutation) using Gel mobilityshift assay
fbx, gel, oct-m-as	AGCCCCTAGGAATTCCATTGTTATGAGGGAGCACATCTGG	Fbx15 probe(Oct mutation) using Gel mobilityshift assay
Fbx-wt-g-PCR(TT)-s	TAG GCA TGT GGA TCC GTC TGT TCC AGA TGT GCT TTA TCA TAA CAA TGG AAT TCC TAG GGG CTT TGT ACC GGA TCC ACC TTC GAT	Fbx15 probe using Gel mobilityshift assay

Fbx-wt-g-PCR(TT)-as	ATC GAA GGT GGA TCC GGT ACA AAG CCC CTA GGA ATT CCA TTG TTA TGA TAA AGC ACA TCT GGA ACA GAC GGA TCC ACA TGC CTA	Fbx15 probe using Gel mobilityshift assay
AACAATG-s(WT)	TAG GCA TGT GGA TCC GTC TGG CCC AGA ATC GAT TTA TCA TAA CAA TGC ATA TGC TAG GGG CTG CGT ACC GGA TCC ACC TTC GAT	Fbx15 probe using Gel mobilityshift assay
AACAATG-as(WT)	ATC CGT ACA CCT AGG CAG ACC GGG TCT TAG CTA AAT AGT ATT GTT ACG TAT ACG ATC CCC GAC GCA TGG CCT AGG TGG AAG CTA	Fbx15 probe using Gel mobilityshift assay
AACCATG-s(4C)	CGA TTT ATC ATA ACC ATG CA	Fbx15 probe(4C mutation) using Gel mobilityshift assay
AACCATG-as(4C)	TAT GCA TGG TTA TGA TAA AT	Fbx15 probe(4C mutation) using Gel mobilityshift assay
Fbx1C-s	CGA TTT ATC ATC ACA ATG CA	Fbx15 probe(1C mutation) using Gel mobilityshift assay
Fbx1C-as	TAT GCA TTG TGA TGA TAA AT	Fbx15 probe(1C mutation) using Gel mobilityshift assay
Fbx1G-s	CGA TTT ATC ATG ACA ATG CA	Fbx15 probe(1G mutation) using Gel mobilityshift assay
Fbx1G-as	TAT GCA TTG TCA TGA TAA AT	Fbx15 probe(1G mutation) using Gel mobilityshift assay
Fbx1T-s	CGA TTT ATC ATT ACA ATG CA	Fbx15 probe(1T mutation) using Gel mobilityshift assay
Fbx1T-as	TAT GCA TTG TAA TGA TAA AT	Fbx15 probe(1T mutation) using Gel mobilityshift assay
Fbx2C-s	CGA TTT ATC ATA CCA ATG CA	Fbx15 probe(2C mutation) using Gel mobilityshift assay
Fbx2C-as	TAT GCA TTG GTA TGA TAA AT	Fbx15 probe(2C mutation) using Gel mobilityshift assay
Fbx2G-s	CGA TTT ATC ATA GCA ATG CA	Fbx15 probe(2G mutation) using Gel mobilityshift assay
Fbx2G-as	TAT GCA TTG CTA TGA TAA AT	Fbx15 probe(2G mutation) using Gel mobilityshift assay
Fbx2T-s	CGA TTT ATC ATA TCA ATG CA	Fbx15 probe(2T mutation) using Gel mobilityshift assay
Fbx2T-as	TAT GCA TTG ATA TGA TAA AT	Fbx15 probe(2T mutation) using Gel mobilityshift assay
Fbx3A-s	CGA TTT ATC ATA AAA ATG CA	Fbx15 probe(3A mutation) using Gel mobilityshift assay
Fbx3A-as	TAT GCA TTT TTA TGA TAA AT	Fbx15 probe(3A mutation) using Gel mobilityshift assay
Fbx3G-s	CGA TTT ATC ATA AGA ATG CA	Fbx15 probe(3G mutation) using Gel mobilityshift assay
Fbx3G-as	TAT GCA TTC TTA TGA TAA AT	Fbx15 probe(3G mutation) using Gel mobilityshift assay
Fbx3T-s	CGA TTT ATC ATA ATA ATG CA	Fbx15 probe(3T mutation) using Gel mobilityshift assay
Fbx3T-as	TAT GCA TTA TTA TGA TAA AT	Fbx15 probe(3T mutation) using Gel mobilityshift assay
Fbx4G-s	CGA TTT ATC ATA ACG ATG CA	Fbx15 probe(4G mutation) using Gel mobilityshift assay
Fbx4G-as	TAT GCA TCG TTA TGA TAA AT	Fbx15 probe(4G mutation) using Gel mobilityshift assay
Fbx4T-s	CGA TTT ATC ATA ACT ATG CA	Fbx15 probe(4T mutation) using Gel mobilityshift assay
Fbx4T-as	TAT GCA TAG TTA TGA TAA AT	Fbx15 probe(4T mutation) using Gel mobilityshift assay
Fbx5C-s	CGA TTT ATC ATA ACA CTG CA	Fbx15 probe(5C mutation) using Gel mobilityshift assay
Fbx5C-as	TAT GCA GTG TTA TGA TAA AT	Fbx15 probe(5C mutation) using Gel mobilityshift assay
Fbx5G-s	CGA TTT ATC ATA ACA GTG CA	Fbx15 probe(5G mutation) using Gel mobilityshift assay
Fbx5G-as	TAT GCA CTG TTA TGA TAA AT	Fbx15 probe(5G mutation) using Gel mobilityshift assay
Fbx5T-s	CGA TTT ATC ATA ACA TTG CA	Fbx15 probe(5T mutation) using Gel mobilityshift assay
Fbx5T-as	TAT GCA ATG TTA TGA TAA AT	Fbx15 probe(5T mutation) using Gel mobilityshift assay
Fbx6A-s	CGA TTT ATC ATA ACA AAG CA	Fbx15 probe(6A mutation) using Gel mobilityshift assay
Fbx6A-as	TAT GCT TTG TTA TGA TAA AT	Fbx15 probe(6A mutation) using Gel mobilityshift assay
Fbx6C-s	CGA TTT ATC ATA ACA ACG CA	Fbx15 probe(6C mutation) using Gel mobilityshift assay
Fbx6C-as	TAT GCG TTG TTA TGA TAA AT	Fbx15 probe(6C mutation) using Gel mobilityshift assay
Fbx6G-s	CGA TTT ATC ATA ACA AGG CA	Fbx15 probe(6G mutation) using Gel mobilityshift assay
Fbx6G-as	TAT GCC TTG TTA TGA TAA AT	Fbx15 probe(6G mutation) using Gel mobilityshift assay
Fbx7A-s	CGA TTT ATC ATA ACA ATA CA	Fbx15 probe(7A mutation) using Gel mobilityshift assay
Fbx7A-as	TAT GTA TTG TTA TGA TAA AT	Fbx15 probe(7A mutation) using Gel mobilityshift assay
Fbx7C-s	CGA TTT ATC ATA ACA ATC CA	Fbx15 probe(7C mutation) using Gel mobilityshift assay
Fbx7C-as	TAT GGA TTG TTA TGA TAA AT	Fbx15 probe(7C mutation) using Gel mobilityshift assay
Fbx7T-s	CGA TTT ATC ATA ACA ATT CA	Fbx15 probe(7T mutation) using Gel mobilityshift assay

Fbx7T-as	TAT GAA TTG TTA TGA TAA AT	Fbx15 probe(7T mutation) using Gel mobilityshift assay
sox15A1-12059-s	aaa aaa aaa acc gcc agg acc tgc ctc ggc agg aac ctc ccc	Sox15 enhancer probe using Gel mobilityshift assay
sox15A1-12059-as	ggg gag gtt cct gcc gag gca ggt cct ggc ggt ttt ttt ttt	Sox15 enhancer probe using Gel mobilityshift assay
sox15A2-12059-s	cac cac cct taa aaa aaa aac tgc ctc ggc agg aac ctc ccc	Sox15 enhancer probe using Gel mobilityshift assay
sox15A2-12059-as	ggg gag gtt cct gcc gag gca gtt ttt ttt tta agg gtg gtg	Sox15 enhancer probe using Gel mobilityshift assay
sox15A3-12059-s	cac cac cct tcc gcc agg aca aaa aaa aaa agg aac ctc ccc	Sox15 enhancer probe using Gel mobilityshift assay
sox15A3-12059-as	ggg gag gtt cct ttt ttt ttt tgt cct ggc gga agg gtg gtg	Sox15 enhancer probe using Gel mobilityshift assay
sox15A4-12059-s	cac cac cct tcc gcc agg acc tgc ctc ggc aaa aaa aaa aaa	Sox15 enhancer probe using Gel mobilityshift assay
sox15A4-12059-as	ttt ttt ttt ttt gcc gag gca ggt cct ggc gga agg gtg gtg	Sox15 enhancer probe using Gel mobilityshift assay
sox15(10-30)-S	CCG CCA GGA CCT GCC TCG GC	Sox15 enhancer probe using Gel mobilityshift assay
sox15(10-30)-AS	GCC GAG GCA GGT CCT GGC GG	Sox15 enhancer probe using Gel mobilityshift assay
sox15(10-40)-S	CCG CCA GGA CCT GCC TCG GCA GGA ACC TCC CC	Sox15 enhancer probe using Gel mobilityshift assay
sox15(10-40)-AS	GGG GAG GTT CCT GCC GAG GCA GGT CCT GGC GG	Sox15 enhancer probe using Gel mobilityshift assay
sox15(20-40)-S	CTG CCT CGG CAG GAA CCT CCC C	Sox15 enhancer probe using Gel mobilityshift assay
sox15(20-40)-AS	GGG GAG GTT CCT GCC GAG GCA G	Sox15 enhancer probe using Gel mobilityshift assay
sox15(10-40)BS-S	gat ccC CGC CAG GAC CTG CCT CGG CAG GAA CCT CCC Cg	Sox15 enhancer probe using Gel mobilityshift assay
sox15(10-40)BS-AS	tcg acG GGG AGG TTC CTG CCG AGG CAG GTC CTG GCG Gg	Sox15 enhancer probe using Gel mobilityshift assay
Hrc-gel-s	CGG ATC ATC AAT AAA TTA CAA AGT TGA GAA AAG GGA GAC	Hrc probe using Gel mobilityshift assay
Hrc-gel-as	GTC TCC CTT TTC TCA ACT TTG TAA TTT ATT GAT GAT CCG	Hrc probe using Gel mobilityshift assay
2APRE-S	ATC CTT CCG GGA ATT CTG ATC CTT CCG GGA ATT CTG	control probe that STAT3 bind using Gel mobilityshift assay
2APRE-AS	CAG AAT TCC CGG AAG GAT CAG AAT TCC CGG AAG GAT	control probe that STAT3 bind using Gel mobilityshift assay
Nanog-2STAT3-S	ACG GTT CCT AGA AGA CAA ACG GTT CCT AGA AGA CAA	tamdem STATbinding site in Nanog dista lenhancer probe using Gel mobilityshift assay
Nanog-2STAT3-AS	TTG TCT TCT AGG AAC CGT TTG TCT TCT AGG AAC CGT	tamdem STATbinding site in Nanog dista lenhancer probe using Gel mobilityshift assay
Nanog-STAT3-S	CTC TCC CGG ACG GTT CCT AGA AGA CAA AGG CAA GCT	Nanog dista lenhancer probe using Gel mobilityshift assay
Nanog-STAT3-AS	AGC TTG CCT TTG TCT TCT AGG AAC CGT CCG GGA GAG	Nanog dista lenhancer probe using Gel mobilityshift assay
Nanog-STAT3m-S	ctc tcc cgg acg gtt cct agt cga caa agg caa gct	Nanog dista lenhancer probe(STAT mutation) using Gel mobilityshift assay
Nanog-STAT3m-AS	agc ttg cct ttg tcg act agg aac cgt ccg gga gag	Nanog dista lenhancer probe(STAT mutation) using Gel mobilityshift assay
Nanog-TCFm-S	ctc tcc cgg acg gtt cct aga aga aca agg caa gct	Nanog dista lenhancer probe(Sox mutation) using Gel mobilityshift assay
Nanog-TCFm-AS	agc ttg cct tgt tct tct agg aac cgt ccg gga gag	Nanog dista lenhancer probe(Sox mutation) using Gel mobilityshift assay
Nanog-doublem-S	ctc tcc cgg acg gtt cct agt cga aca agg caa gct	Nanog dista lenhancer probe(STAT/Sox mutation) using Gel mobilityshift assay
Nanog-doublem-AS	agc ttg cct tgt tcg act agg aac cgt ccg gga gag	Nanog dista lenhancer probe(STAT/Sox mutation) using Gel mobilityshift assay
Nanog-STAT3(AT)-S	CTC TCC CGG ACG Gtt cca tga aGA CAA AGG CAA GCT	Nanog dista lenhancer probe(STAT mutation) using Gel mobilityshift assay
Nanog-STAT3(AT)-AS	AGC TTG CCT TTG TCt tca tgg aaC CGT CCG GGA GAG	Nanog dista lenhancer probe(STAT mutation) using Gel mobilityshift assay
Nanog-STAT3(GG)-S	CTC TCC CGG ACG Ggg cct aga aGA CAA AGG CAA GCT	Nanog dista lenhancer probe(STAT mutation) using Gel mobilityshift assay
Nanog-STAT3(GG)-AS	AGC TTG CCT TTG TCt tct agg ccC CGT CCG GGA GAG	Nanog dista lenhancer probe(STAT mutation) using Gel mobilityshift assay
Nanog-TCF(TAT)-S	CTC TCC CGG ACG Gtt cct aga aGA CTA TGG CAA GCT	Nanog dista lenhancer probe(Sox mutation) using Gel mobilityshift assay
Nanog-TCF(TAT)-AS	AGC TTG CCA TAG TCt tct agg aaC CGT CCG GGA GAG	Nanog dista lenhancer probe(Sox mutation) using Gel mobilityshift assay
MsNanog-4737-4678-S	acc cgg agc tgt gcg ccc tgt acc aaa cct ttg tag aac ttg ggg taa act taa ggc tat	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4737-4678-AS	ata gcc tta agt tta ccc caa gtt cta caa agg ttt ggt aca ggg cgc aca gct ccg ggt	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4687-4628-S	tta agg cta tgg tgg cct tga ctc cgt gga ccc aga ggc aag ttt cct cct tta gag gac	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4687-4628-AS	gtc ctc taa agg agg aaa ctt gcc tct ggg tcc acg gag tca agg cca cca tag cct taa	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4637-4578-S	ttt aga gga ctc gca tgc att ttg ttt cta att tga aat gag aac cgg ctt aga gct tga	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4637-4578-AS	tca agc tct aag ccg gtt ctc att tca aat tag aaa caa aat gca tgc gag tcc tct aaa	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4587-4528-S	tag agc ttg aac cag cca gtt ctc tgg act cct ccc agc tct tac aat tcc tct ccc gga	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4587-4528-AS	tcc ggg aga gga att gta aga gct ggg agg agt cca gag aac tgg ctg gtt caa gct cta	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4537-4478-S	ctc tcc cgg acg gtt cct aga aga caa agg caa gct tac caa aat tac gtc gcc ctt ggg	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4537-4478-AS	ccc aag ggc gac gta att ttg gta agc ttg cct ttg tct tct agg aac cgt ccg gga gag	Nanog enhancer region probe using Gel mobilityshift assay

MsNanog-4487-4428-S	cgc cct tgg gac aca cct agg gtt ccc tgg tgg cat ctt ttt ttt ttc att ata aac agg	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4487-4428-AS	cct gtt tat aat gaa aaa aaa aag atg cca cca ggg aac cct agg tgt gtc cca agg gcg	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4437-4378-S	tat aaa cag gag taa att ttt gta agg gca gag ctg gta gct gag gga gag gaa ccc ttt	Nanog enhancer region probe using Gel mobilityshift assay
MsNanog-4437-4378-AS	aaa ggg ttc ctc tcc ctc agc tac cag ctc tgc cct tac aaa aat tta ctc ctg ttt ata	Nanog enhancer region probe using Gel mobilityshift assay
Nanog exon up-S	cct tac agc ttc ttt tgcattacaatgtccatggtggac	Nanog proximal enhancer probe using Gel mobilityshift assay
Nanog exon up-AS	gtccaccatggacattgtaatgcaaaagaagctgtaagg	Nanog proximal enhancer probe using Gel mobilityshift assay
Nanog exon up-O(m)-S	ccttacagcttcgggttacctacaatgtccatggtggac	Nanog proximal enhancer probe(Oct mutation) using Gel mobilityshift assay
Nanog exon up-O(m)-AS	gtccaccatggacattgtaggtaacccgaagctgtaagg	Nanog proximal enhancer probe(Oct mutation) using Gel mobilityshift assay
Nanog exon up-S(m)-S	ccttacagcttcttttgcatacccaattccatggtggac	Nanog proximal enhancer probe(Sox mutation) using Gel mobilityshift assay
Nanog exon up-S(m)-AS	gtccaccatggaattgggtatgcaaaagaagctgtaagg	Nanog proximal enhancer probe(Sox mutation) using Gel mobilityshift assay
Nanog exon up-D(m)-S	ccttacagcttcgggttaccacccaattccatggtggac	Nanog proximal enhancer probe(Oct/Sox mutation) using Gel mobilityshift assay
Nanog exon up-D(m)-AS	gtccaccatggaattgggtggtaacccgaagctgtaagg	Nanog proximal enhancer probe(Oct/Sox mutation) using Gel mobilityshift assay
hNanog gel-S	ttctccggaatggtagtctgagaagaaaaaagatta	human Nanog dista lenhancer probe using Gel mobilityshift assay
hNanog gel-AS	taatctttttttcttctcagactaccattccggagaa	human Nanog distal enhancer probe using Gel mobilityshift assay
human Nanog ex-S	atgctgccttcaagcatctgtttaacaaagcacatcttg	human Nanog proximal enhancer probe using Gel mobilityshift assay
human Nanog ex-AS	caagatgtgctttgttaaacagatgcttgaaggcagcat	human Nanog proximal enhancer probe using Gel mobilityshift assay
ECAT4mTCF(AC)-s	gga cgg ttc cta gaa gaa caa ggc aag ctt acc aaa att acg	mutagenesis in Sox binding site of Nanog distal enhancer
ECAT4mTCF(AC)-as	cgt aat ttt ggt aag ctt gcc ttg ttc ttc tag gaa ccg tcc	mutagenesis in Sox binding site of Nanog distal enhancer
ECAT4mdouble-s	gga cgg ttc cta gtc gaa caa ggc aag ctt acc aaa att acg	mutagenesis in Sox and STAT binding site of Nanog distal enhancer
ECAT4mdouble-as	cgt aat ttt ggt aag ctt gcc ttg ttc gac tag gaa ccg tcc	mutagenesis in Sox and STAT binding site of Nanog distal enhancer
mSox15F1	GAGCAAAAACTCGAGCACCG	RT-PCR in Sox15 gene
mSox15R1	TGTGCATTCTGGTTCCTTGG	RT-PCR in Sox15 gene
FGF4-RT-S	cgt ggt gag cat ctt cgg agt gg	RT-PCR in FGF4 gene
FGF4-RT-AS	cct tct tgg tcc gcc cgt tct ta	RT-PCR in FGF4 gene
NAT1U283	ATT CTT CgT TgT CAA gCC gCC AAA gTg gAg	RT-PCR in NAT1 gene
NAT1L476	AgT TgT TTg CTg Cgg AgT TgT CAT CTC gTC	RT-PCR in NAT1 gene
Fbx U1110	TCCCTATACATTGCCTCCCGACAC	RT-PCR in Fbx15 gene
Fbx L1431	AGTTCTCCATTCAAGGCCAACTTT	RT-PCR in Fbx15 gene
Ctgf-s	GGTGAGTCCTTCCAAAGCAGCTGCAAAT	RT-PCR in Ctgf gene
Ctgf-as	GCAGTTGGCTCGCATCATAGTTGGG	RT-PCR in Ctgf gene
Ebaf-s	GCTTTGCTGGGCACTCTGGGTACT	RT-PCR in Ebaf gene
Ebaf-as	CCATTCCGAACACTAGCAGGTGAGTGGA	RT-PCR in Ebaf gene
Hrc-s	GGACATTTCTACTGAGTTTGGCCACAAG	RT-PCR in Hrc gene
Hrc-as	CCTCTTCATCATCTTCCTGGCTCATGGG	RT-PCR in Hrc gene
mfg1-s	CTTTCTACACACCAGAGAGTTCATACGG	RT-PCR in mfg1 gene
mfg1-as	TGCCACATTCTTCACATTTGATCG	RT-PCR in mfg1 gene
Otx2-s	CTAAAGCAACCGCCTTACGCAGTC	RT-PCR in Otx2 gene
Otx2-as	GCGGCACTTAGCTCTTCGATTCTTAAAC	RT-PCR in Otx2 gene
Hrc-U-3152	GTCTACCACCAACCTTCCCACTCACAAC	CHIP assay in Hrc gene
Hrc-L-2920	GGTGGGTCTGGCAGAGGGTCACA	CHIP assay in Hrc gene
Hrc-U-1377	TAAGAGAGGGACCCAGAGAAAGAAAAGG	CHIP assay in Hrc gene
Hrc-L-1172	CTGTCTCTCCTCTCTGAATCTGGGACTC	CHIP assay in Hrc gene
Hrc-U-1032	AGACAGACAGACACACAGAGAGACAGGC	CHIP assay in Hrc gene
Hrc-L-842	TCACTGTTCTGAGCTTCCGTGTTTC	CHIP assay in Hrc gene
Hrc-U3721	TAGATCTGGAGGTGGTTGGTTTGGGTTC	CHIP assay in Hrc gene
Hrc-L3932	TTCAAAAGCTCTGGTGGTGACCAGCCTT	CHIP assay in Hrc gene
Hrc-U4401	CCGCCCCCTTTCCGCAGGTGCAG	CHIP assay in Hrc gene

Hrc-L4611	ATACATACCAGCCGGGTGTGCGTGTTCC	CHIP assay in Hrc gene
Nanog -5629U	GTGCACACAGAGAACAACTTGTGGG	CHIP assay in Nanog regulatory region
Nanog -5462L	CTCCAGCACTACACAGAACATAGTATCG	CHIP assay in Nanog regulatory region
Nanog-CHIP-s	TGAAATGAGAACCGGCTTAGAGCTTGAA	CHIP assay in Nanog regulatory region
Nanog-CHIP-as	CTCTCCCTCAGCTACCAGCTCTGCCCTT	CHIP assay in Nanog regulatory region
Nanog -3910U	CAGCGGGTAAAACACACTTGTATA	CHIP assay in Nanog regulatory region
Nanog -3721L	TAGCCAACTCAGTATGCAGACCAACCTG	CHIP assay in Nanog regulatory region
Nanog -2862U	AAAATGGAATCACTGGAACCCAATTCTA	CHIP assay in Nanog regulatory region
Nanog -2670L	TGTCATGTAGGTACTGGAAATTGAACCC	CHIP assay in Nanog regulatory region
Nanog -1823U	CAGCAGTTGAGGCAGTGGTGGTCGTAGT	CHIP assay in Nanog regulatory region
Nanog -1622L	CACTGCCACCAAGCCACCATTGTTA	CHIP assay in Nanog regulatory region
Nanog-exon CHIP-s	TCTTTAGATCAGAGGATGCCCCCTAAGC	CHIP assay in Nanog regulatory region
Nanog-exon CHIP-as	AAGCCTCCTACCCTACCCACCCCCTAT	CHIP assay in Nanog regulatory region
Nanog 907U	GCCCAGTACTCAGGCTTGTAAATGGTAG	CHIP assay in Nanog regulatory region
Nanog 1079L	TTAAAAGCTCAGAAAAAAGGGGACA	CHIP assay in Nanog regulatory region
mSox15-s	GAG CAA AAA CTC GAG CAC CG	In situ hybridization probe(Sox15)
mSox15-as	TGT GCA TTC TGG TTC CTT GG	In situ hybridization probe(Sox15)
sox2 in situ-s	CGC TCA TGA AGA AGG ATA AGT ACA CGC	In situ hybridization probe(Sox2)
sox2 in situ-as	TTG CAC CCC TCC CAA TTC TCT TGT ATC T	In situ hybridization probe(Sox2)
Sox15-89-S	CAC CTC CAA AGC AAG GCA GAA AAT AAG CTA GG	Luciferase reporter saasy(Sox15 enhancer)
Sox15-4188-AS	TGT CAC CCA AAG GCC CCA ATA GTC CAA T	Luciferase reporter saasy(Sox15 enhancer)
Sox15-3783-S	CAC CTG GCT TCT AGT TCA AAC TGC CCT CAT TA	Luciferase reporter saasy(Sox15 enhancer)
Sox15-7908-AS	GAG ATG GAG AAA GGA GGG ATG TGG ACC G	Luciferase reporter saasy(Sox15 enhancer)
Sox15-7738-S	CAC CAA GGC GTG CCT GCA TCT GCT GTA CTC AT	Luciferase reporter saasy(Sox15 enhancer)
Sox15-11725-AS	GGT AAG GGG AGA AAG AGG GTC TTA GCT C	Luciferase reporter saasy(Sox15 enhancer)
Sox15-11201-S	CAC CCC CAA ATC CAC CCA AGC AGA AGC TAC AC	Luciferase reporter saasy(Sox15 enhancer)
Sox15-15297-AS	TAG CTG TCC TTC GAT CCC AAG TGC TAA G	Luciferase reporter saasy(Sox15 enhancer)
Sxo15-14903-S	CAC CCT TCT GGG ATG TCT GAA GAC TGC TAC AA	Luciferase reporter saasy(Sox15 enhancer)
Sox15-19503-AS	CAT CTC TTC ACC ATG CCA CTG ACT CAA C	Luciferase reporter saasy(Sox15 enhancer)
Sox15-19107-S	CAC CAA CAG GGT GTC AGG TCG GAA TGA TTC AT	Luciferase reporter saasy(Sox15 enhancer)
Sox15-22941-AS	CAT TGA TGT GCA GCA GGT CTC CTT AGT C	Luciferase reporter saasy(Sox15 enhancer)
Sox15-12066-AS	GTG AGG GAG AGT GAC TCA GAT AAG G	Luciferase reporter saasy(Sox15 enhancer)
Sox15-11942-S	CAC CGT CCC ATT TGT ATT CTC TCC CTA CA	Luciferase reporter saasy(Sox15 enhancer)
Sox15-12767-AS	AGA ACC AAA AGT ATC AAT GTC TCC T	Luciferase reporter saasy(Sox15 enhancer)
Sox15-12652-S	CAC CAA GGA GCC CCC AAA CAG CAT AAA CA	Luciferase reporter saasy(Sox15 enhancer)
Sox15-13477-AS	GGC TGG AAC GTA GGC AAT CAG ACT C	Luciferase reporter saasy(Sox15 enhancer)
Sox15-13352-S	CAC CCT GGC TCA CGC AGG GAA GGC TCG CT	Luciferase reporter saasy(Sox15 enhancer)
Sox15-14177-AS	CAC GGG TGA GGG AGA GGG CAA CAG C	Luciferase reporter saasy(Sox15 enhancer)
Sox15-14052-S	CAC CGC CTT TGC TGA GGA GAA TCT TGA GG	Luciferase reporter saasy(Sox15 enhancer)
Sox15-14877-AS	AGC CAT CAT GTG GTT GCT AGG ATT T	Luciferase reporter saasy(Sox15 enhancer)
Sox15-14711-S	CAC CGA GTG AGT TCC AGG ACA GCC AGG GC	Luciferase reporter saasy(Sox15 enhancer)
11866-S-gw	CAC CCC AAG GAA CCA GAA TGC ACA CTA A	Luciferase reporter saasy(Sox15 enhancer)
12016-S-gw	CAC CTT TGT ATT CTC TCC CTA CAG CAG C	Luciferase reporter saasy(Sox15 enhancer)
12016-AS-gw	TGG GAC AAT AAA AGA CTC CGA AAC ATG G	Luciferase reporter saasy(Sox15 enhancer)
12037-AS-gw	GGG CCC TGG GGT TGG GTG GGG AGG TTC C	Luciferase reporter saasy(Sox15 enhancer)
12099-AS-gw	TGC CCG GGT GGG ATG GTC TAT TGG AGC C	Luciferase reporter saasy(Sox15 enhancer)
sox15-12013-s	CAC CAT TTG TAT TCT CTC CCT ACA GCA GCA ACC AGC GCC TTC ACA C	Luciferase reporter saasy(Sox15 enhancer)

sox15-12013-as	GTG TGA AGG CGC TGG TTG CTG CTG TAG GGA GAG AAT ACA AAT GGT G	Luciferase reporter saasy(Sox15 enhancer)
sox15-12059-s	CAC CAC CCT TCC GCC AGG ACC TGC CTC GGC AGG AAC CTC CCC	Luciferase reporter saasy(Sox15 enhancer)
sox15-12059-as	GGG GAG GTT CCT GCC GAG GCA GGT CCT GGC GGA AGG GTG GTG	Luciferase reporter saasy(Sox15 enhancer)
sox15-12098-s	CAC CAC CCA ACC CCA GGG CCC TTA TCT GAG TCA CTC TCC CT	Luciferase reporter saasy(Sox15 enhancer)
sox15-12098-as	AGG GAG AGT GAC TCA GAT AAG GGC CCT GGG GTT GGG TGG TG	Luciferase reporter saasy(Sox15 enhancer)
sox15-12013/12033-s	CAC CCT TCC GCC AGG ACC TG	Luciferase reporter saasy(Sox15 enhancer)
sox15-12013/12033-a	CAG GTC CTG GCG GAA GGG TG	Luciferase reporter saasy(Sox15 enhancer)
sox15-12023_12043-s	CAC CAG GAC CTG CCT CGG CAG G	Luciferase reporter saasy(Sox15 enhancer)
sox15-12023_12043-a	CCT GCC GAG GCA GGT CCT GGT G	Luciferase reporter saasy(Sox15 enhancer)
sox15-12033_12053-s	CAC CTC GGC AGG AAC CTC CCC	Luciferase reporter saasy(Sox15 enhancer)
sox15-12033_12053-a	GGG GAG GTT CCT GCC GAG GTG	Luciferase reporter saasy(Sox15 enhancer)
FGF4enha-S	ggtaccaagggctgaggtgagggagaacaatta	Luciferase reporter saasy(Sox15 enhancer)
FGF4enha-AS	gctagcctgggctatgagaccgtcttttagaaa	Luciferase reporter saasy(Sox15 enhancer)
sox15-5arm-s(NotI)	GCG GCC GCA AGA CAG GAT TAT TAG AC	5'-arm of Sox15 targeting vector
sox15-5arm-As(SpeI)	ACT AGT CCC CCA GAC GCT CCA	5'-arm of Sox15 targeting vector
sox15-3arm-s(BamHI)	GGA TCC GTC CCC TTT AGC CAA GAA	3'-arm of Sox15 targeting vector
sox15-3arm-As(XhoI)	CTC GAG TTG GTG CTC TTA ACC TCT	3'-arm of Sox15 targeting vector
sox2-5arm-sNotI	GCG GCC GCA TGC TTG GCA GTT AAG GCT TCT	5'-arm of Sox2 targeting vector
sox2-5arm-asNheI	GCT AGC AGC TCC GTC TCC ATC ATG TTA TAC	5'-arm of Sox2 targeting vector
sox2-3arm-sBamHI	GGA TCC ACC CGA CTC AGC CTC TCA ACT TAA	3'-arm of Sox2 targeting vector
sox2-3arm-asXhoI	CTC GAG CTT CCA ATC CCA AGA TAG GTA GAC	3'-arm of Sox2 targeting vector
sox15south5.s(EcoRI)	TGGGTGAATGACTCTTCTTTACTCCACA	5'southern probe in Sox15
sox15south5.as(EcoRI)	AAGGCACATCCAGTCTGAAAGGGAAGGG	5'southern probe in Sox15
sox15-3'south-sMunI	GAG TAG CAC ATG CCG CCC TGA GTT CGA C	3'southern probe in Sox15
sox15-3'south-asMun	GAC CAG AAA GAA GGC CCA GCT CGG AAT G	3'southern probe in Sox15
sox15-3'-m-tail-s	GAT GGC GCA GCA GAA CCC CAA GAT GCA C	Sox15 genotyping PCR
sox15-3'-m-tail-as	GGG GCT CCA GCA AGG GAA GTA TTA TAT G	Sox15 genotyping PCR
sox2-southern5-s(EcoRI)	GACATTGCCTTTGAGGTTAGAAGAGCTTTT	5'southern probe in Sox2
sox2-southern5-as(EcoRI)	TATTAAGTGACAGCAATGCAGGCCGGTGAC	5'southern probe in Sox2
sox2-3south-s(EcoRI)	CAGATAAGCTCACACATGGCCTGCATACCT	3'southern probe in Sox2
sox2-3south-as(EcoRI)	agttcggtttcagaccattcgccatagaag	3'southern probe in Sox2
sox2-3'-tail-s	ATT ATT CAC GTG GTA ATG AGC ACA GTC G	Sox2 genotyping PCR
sox2-3'-tail-as	GCA AGT CCT CTC TGG GTA GAA TCA TGA C	Sox2 genotyping PCR
bgeo-screening1	AATGGGCTGACCGCTTCCTCGTGCTT	Sox2, Sox15 genotyping PCR
pGV-P/AS(Bam/Sal)	GAC AGT CAT AAG TGC GGC GAC GAT	Sequence of fragment insertedinto pGV-P plasmid
pGV-P/S(Bam/Sal)	ATG TTT CAG GTT CAG GGG GAG GTG	Sequence of fragment insertedinto pGV-P plasmid

Appemdix 1(Plasmid list)

Name	Application	
pDONR/Oct3/4	Entry vector for Gateway system	Mouse Oct3/4 ORF amplified by PCR with oct3/4gw-s and oct3/4gw-as primers was was inserted into pDONR201 by BP reaction.
pDONR/Sox2	Entry vector for Gateway system	Mouse Sox2 ORF amplified by PCR with sox2-gw-s and sox2-gw-as primers was was inserted into pDONR201 by BP reaction.
pENTR/Sox15	Entry vector for Gateway system	Mouse Sox15 ORF amplified by PCR with sox15-gw-s and sox15-gw-as primers was subcloned into pENTR-D-TOPO.
pENTR/Sox7	Entry vector for Gateway system	Mouse Sox7 ORF amplified by PCR with sox7-gw-S and sox7-gw-AS primers was subcloned into pENTR-D-TOPO.
pENTR/Sox17	Entry vector for Gateway system	Mouse Sox17 ORF amplified by PCR with oct3/4gw-s and oct3/4gw-as primers was subcloned into pENTR-D-TOPO.
pENTR/Sox2-15-2	Entry vector for Gateway system	Three products amplified with sox2-gw-s and sox2HMG15-out-as, sox2HMG15-hmg-s and sox2HMG15-hmg-as, sox2HMG15-out-s and sox2-gw-as were mix, them used as templates insecond PCR with sox2-gw-s and sox2-gw-as aubcloned into pENTR/D-TOPO
pENTR/Sox15-2-15	Entry vector for Gateway system	Three products amplified with sox15-gw-s and sox15HMG2-out-as, sox15HMG2-hmg-s and sox15HMG2-hmg-as, sox15HMG2-out-s and sox15-gw-as were mix, them used as templates insecond PCR with so15-gw-s and sox15-gw-as aubcloned into pENTR/D-TOPO
pENTR/MyoD	Entry vector for Gateway system	Mouse MyoD ORF amplified by PCR with MyoD-s,gw and MyoD-as,gw primers was was subcloned into pENTR-D-TOPO.
pENTR/Myogenin	Entry vector for Gateway system	Mouse Myogenin ORF amplified by PCR with Myogenin-s,gw and Myogenin-as,gw primers was subcloned into pENTR-D-TOPO.
pENTR/Sox15C	Entry vector for Gateway system	Mouse C-terminal of Sox15 amplified by PCR with anti-sox15-s and anti-sox15-s primers was subcloned into pENTR-D-TOPO.
pENTR/Sox2C	Entry vector for Gateway system	Mouse C-terminal of Sox2 amplified by PCR with anti-sox2-s and anti-sox2-as primers was subcloned into pENTR-D-TOPO.
pCAG-IP-HA/Oct3/4	Expression vector	LR reaction with pDONR/Oct3/4 and pCAG-IP-gw-HA
pCAG-IP-HA/Sox2	Expression vector	LR reaction with pDONR/Sox2 and pCAG-IP-gw-HA
pCAG-IP-HA/Sox15	Expression vector	LR reaction with pENTR/Sox15 and pCAG-IP-gw-HA
pCAG-IP-myc/Oct3/4	Expression vector	LR reaction with pDONR/Oct3/4 and pCAG-IP-gw-myc
pCAG-IP-myc/Sox2	Expression vector	LR reaction with pDONR/Sox2 and pCAG-IP-gw-myc
pCAG-IP-myc/Sox15	Expression vector	LR reaction with pENTR/Sox15 and pCAG-IP-gw-myc
pCAG-IP-Flag/Sox7	Expression vector	LR reaction with pENTR/Sox7 and pCAG-IP-gw-Flag
pCAG-IP-Flag/Sox17	Expression vector	LR reaction with pENTR/Sox17 and pCAG-IP-gw-Flag
pCAG-IP-HA/Sox2-15-2	Expression vector	LR reaction with pENTR/Sox2-15-2 and pCAG-IP-gw-HA
pCAG-IP-HA/Sox15-2-15	Expression vector	LR reaction with pENTR/Sox15-2-15 and pCAG-IP-gw-HA
pCAG-IP-HA/MyoD	Expression vector	LR reaction with pENTR/MyoD and pCAG-IP-gw-HA
pCAG-IP-HA/Myogenin	Expression vector	LR reaction with pENTR/Myogenin and pCAG-IP-gw-HA
pDEST15/Sox15C	Expression vector	LR reaction with pENTR/Sox15C and pDEST15
pDEST17/Sox15C	Expression vector	LR reaction with pENTR/Sox15C and pDEST17
pDEST15/Sox2C	Expression vector	LR reaction with pENTR/Sox2C and pDEST15
pDEST17/Sox2C	Expression vector	LR reaction with pENTR/Sox2C and pDEST17
pCR2.1/FGF4		The PCR product amplified withFGF4enha-S and FGF4enha-AS primers in RF8 genome was subcloned into pCR2.1-TOPO.
pGV-P/FGF4	Lucifease reporter vector	A NheI/KpnI (Blunted) fragment of pCR2.1/FGF4 was introduced into the BamHI (Blunted) site of pGV-P
pCR2.1/FbxAACAATG(wt)	Gel mobility shift assay	The PCR product amplified with SELEX_N-20 FW and SELEX_N-20RV primers in AACAATG-s(wt) and -as(wt) annealing product was subcloned into pCR2.1-TOPO.
pCR2.1/FbxAACCATG(4C)	Gel mobility shift assay	AACCATG-s(4C) and -as(4C) annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(1C)	Gel mobility shift assay	Fbx1C-s and Fbx1C-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(1G)	Gel mobility shift assay	Fbx1G-s and 1G-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(1T)	Gel mobility shift assay	Fbx1T-s and 1T-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(2C)	Gel mobility shift assay	Fbx2C-s and 2C-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(2G)	Gel mobility shift assay	Fbx2G-s and 2G-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(2T)	Gel mobility shift assay	Fbx2T-s and 2T-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(3A)	Gel mobility shift assay	Fbx3A-s and 3A-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(3G)	Gel mobility shift assay	Fbx3G-s and 3G-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(3T)	Gel mobility shift assay	Fbx3T-s and 3T-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(4G)	Gel mobility shift assay	Fbx4G-s and 4G-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(4T)	Gel mobility shift assay	Fbx4T-s and 4T-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)

pCR2.1/Fbx(5T)	Gel mobility shift assay	Fbx5T-s and 5T-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(5A)	Gel mobility shift assay	Fbx5A-s and 5A-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(5C)	Gel mobility shift assay	Fbx5C-s and 5C-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(6A)	Gel mobility shift assay	Fbx6A-s and 6A-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(6C)	Gel mobility shift assay	Fbx6C-s and 6C-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(6G)	Gel mobility shift assay	Fbx6G-s and 6G-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(7A)	Gel mobility shift assay	Fbx7A-s and 7A-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(7C)	Gel mobility shift assay	Fbx7C-s and 7C-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pCR2.1/Fbx(7T)	Gel mobility shift assay	Fbx7T-s and 7T-as annealing product was introduced into the ClaI/NdeI site of pCR2.1/FbxAACAATG(wt)
pENTR/Sox15(89-4188)		The PCR product amplified with Sox15-89-S and Sxo15-4188-AS primers in RF8 genome was subcloned into pENTR/D-TOPO
pENTR/Sox15(3783-7908)		The PCR product amplified with Sox15-3783-S and Sxo15-7908-AS primers in RF8 genome was subcloned into pENTR/D-TOPO
pENTR/Sox15(7738-11725)		The PCR product amplified with Sox15-7738-S and Sxo15-11725-AS primers in RF8 genome was subcloned into pENTR/D-TOPO
pENTR/Sox15(11201-15297)		The PCR product amplified with Sox15-11201-S and Sxo15-15297-AS primers in RF8 genome was subcloned into pENTR/D-TOPO
pENTR/Sox15(14903-19503)		The PCR product amplified with Sox15-14903-S and Sxo15-19503-AS primers in RF8 genome was subcloned into pENTR/D-TOPO
pENTR/Sox15(19107-22941)		The PCR product amplified with Sox15-19107-S and Sxo15-22941-AS primers in RF8 genome was subcloned into pENTR/D-TOPO
pGV-P/Sox15(89-4188)	Lucifease reporter vector	LR reaction with pENTR/Sox15(89-4188) and pGV-P-gw
pGV-P/Sox15(3783-7908)	Lucifease reporter vector	LR reaction with pENTR/Sox15(3783-7908) and pGV-P-gw
pGV-P/Sox15(7738-11725)	Lucifease reporter vector	LR reaction with pENTR/Sox15(7738-11725) and pGV-P-gw
pGV-P/Sox15(11201-15297)	Lucifease reporter vector	LR reaction with pENTR/Sox15(11201-15297) and pGV-P-gw
pGV-P/Sox15(14903-19503)	Lucifease reporter vector	LR reaction with pENTR/Sox15(14903-19503) and pGV-P-gw
pGV-P/Sox15(19107-22941)	Lucifease reporter vector	LR reaction with pENTR/Sox15(19107-22941) and pGV-P-gw
pENTR/Sox15(11201-12066)		The PCR product amplified with Sox15-11201-S and Sxo15-12066-AS primers in pENTR/Sox15(11201-15297) was subcloned into pENTR/D-TOPO
pENTR/Sox15(11942-12767)		The PCR product amplified with Sox15-11942-S and Sxo15-12767-AS primers in pENTR/Sox15(11201-15297) was subcloned into pENTR/D-TOPO
pENTR/Sox15(12652-13477)		The PCR product amplified with Sox15-12652-S and Sxo15-13477-AS primers in pENTR/Sox15(11201-15297) was subcloned into pENTR/D-TOPO
pENTR/Sox15(13352-14177)		The PCR product amplified with Sox15-13352-S and Sxo15-14177-AS primers in pENTR/Sox15(11201-15297) was subcloned into pENTR/D-TOPO
pENTR/Sox15(14052-14877)		The PCR product amplified with Sox15-14052-S and Sxo15-14877-AS primers in pENTR/Sox15(11201-15297) was subcloned into pENTR/D-TOPO
pENTR/Sox15(14711-15297)		The PCR product amplified with Sox15-14711-S and Sxo15-15297-AS primers in pENTR/Sox15(11201-15297) was subcloned into pENTR/D-TOPO
pGV-P/Sox15(11201-12066)	Lucifease reporter vector	LR reaction with pENTR/Sox15(11201-12066) and pGV-P-gw
pGV-P/Sox15(11942-12767)	Lucifease reporter vector	LR reaction with pENTR/Sox15(11942-12767) and pGV-P-gw
pGV-P/Sox15(12652-13477)	Lucifease reporter vector	LR reaction with pENTR/Sox15(12652-13477) and pGV-P-gw
pGV-P/Sox15(13352-14177)	Lucifease reporter vector	LR reaction with pENTR/Sox15(13352-14177) and pGV-P-gw
pGV-P/Sox15(14052-14877)	Lucifease reporter vector	LR reaction with pENTR/Sox15(14052-14877) and pGV-P-gw
pGV-P/Sox15(14711-15297)	Lucifease reporter vector	LR reaction with pENTR/Sox15(14711-15297) and pGV-P-gw
pGV-P/Sox15(11201-11210)	Lucifease reporter vector	A KpnI/ApaI (Blunted) fragment of pGV-P/Sox15(7738/11725) was introduced into the BamHI (Blunted) site of pGV-P
pGV-P/Sox15(11238-11665)	Lucifease reporter vector	An ApaI/PstI Blunted) fragment of pGV-P/Sox15(11201/12066) was introduced into the BamHI (Blunted) site of pGV-P
pGV-P/Sox15(11453-11864)	Lucifease reporter vector	A SphI/BglII (Blunted) fragment of pGV-P/Sox15(11201/12066) was introduced into the BamHI (Blunted) site of pGV-P
pGV-P/Sox15(11646-12071)	Lucifease reporter vector	An ApaI/BanIII/SphI (Blunted) fragment of pGV-P/Sox15(11201/12066) was introduced into the BamHI (Blunted) site of pGV-P
pGV-P/Sox15(12034-12396)	Lucifease reporter vector	A ApaI/SacI (Blunted) fragment of pGV-P/Sox15(11201/15297) was introduced into the BamHI (Blunted) site of pGV-P
pENTR/Sox15(11866-12066)		The PCR product amplified with 11866-S-gw and 12066-AS-gw primers in pENTR/Sox15(11646-12071) was subcloned into pENTR/D-TOPO
pENTR/Sox15(12016-12037)		The PCR product amplified with 12016-S-gw and 12037-AS-gw primers in pENTR/Sox15(11646-12071) was subcloned into pENTR/D-TOPO
pENTR/Sox15(11866-12037)		The PCR product amplified with 11866-S-gw and 12037-AS-gw primers in pENTR/Sox15(11646-12071) was subcloned into pENTR/D-TOPO
pENTR/Sox15(11866-12099)		The PCR product amplified with 11866-S-gw and 12099-AS-gw primers in pENTR/Sox15(11646-12071) was subcloned into pENTR/D-TOPO
pGV-P/Sox15(11866-12066)	Lucifease reporter vector	LR reaction with pENTR/Sox15(11866-12066) and pGV-P-gw
pGV-P/Sox15(12016-12037)	Lucifease reporter vector	LR reaction with pENTR/Sox15(12016-12037) and pGV-P-gw
pGV-P/Sox15(11866-12037)	Lucifease reporter vector	LR reaction with pENTR/Sox15(11866-12037) and pGV-P-gw
pGV-P/Sox15(11866-12099)	Lucifease reporter vector	LR reaction with pENTR/Sox15(11866-12099) and pGV-P-gw
pENTR/Sox15(12013-12058)		Thefragment annealed with sox15-12013-s and sox15-12013-as primers was subcloned into pENTR/D-TOPO

pENTR/Sox15(12059-12097)		Thefragment annealed with sox15-12059-s and sox15-12059-as primers was subcloned into pENTR/D-TOPO
pENTR/Sox15(12098-12143)		Thefragment annealed with sox15-12098-s and sox15-12098-as primers was subcloned into pENTR/D-TOPO
pGV-P/Sox15(12013-12058)	Lucifease reporter vector	LR reaction with pENTR/Sox15(12013-12058) and pGV-P-gw
pGV-P/Sox15(12059-12097)	Lucifease reporter vector	LR reaction with pENTR/Sox15(12059-12097) and pGV-P-gw
pGV-P/Sox15(12098-12143)	Lucifease reporter vector	LR reaction with pENTR/Sox15(12098-12143) and pGV-P-gw
pENTR/Sox15(12013-12033)		Thefragment annealed with sox15-12013/12033-s and sox15-12013/12033-a primers was subcloned into pENTR/D-TOPO
pENTR/Sox15(12023-12043)		Thefragment annealed with sox15-12023_12043-s and sox15-12023_12043-a primers was subcloned into pENTR/D-TOPO
pENTR/Sox15(12033-12053)		Thefragment annealed with sox15-12033_12053-s and sox15-12033_12053-a primers was subcloned into pENTR/D-TOPO
pGV-P/Sox15(12013-12033)	Lucifease reporter vector	LR reaction with pENTR/Sox15(12013-12033) and pGV-P-gw
pGV-P/Sox15(12023-12043)	Lucifease reporter vector	LR reaction with pENTR/Sox15(12023-12043) and pGV-P-gw
pGV-P/Sox15(12033-12053)	Lucifease reporter vector	LR reaction with pENTR/Sox15(12033-12053) and pGV-P-gw
pCR2.1/Sox15 in situ	In situ hybridization probe	The PCR product amplified with mSox15-s and mSox15-as primers in RF8 genome was subcloned into pCR2.1-TOPO.
pBSKS/mSox15 in situ	In situ hybridization probe	An EcoRI (Blunted) fragment of pCR2.1/Sox15 in situ was introduced into the BamHI/SalI (Blunted) site of pGV-P
pCR2.1/Sox2 in situ	In situ hybridization probe	The PCR product amplified with sox2 in situ-s and sox2 in situ-as primers in RF8 genome was subcloned into pCR2.1-TOPO.
pBS-SK Nanog sox(m)		Mutagenesis with ECAT4mTCF(AC)-s and ECAT4mTCF(AC)-as using pBS-SK-ECAT4 6K(-4694/-3670) as template
pBS-SK Nanog double(m)		Mutagenesis with ECAT4mdouble-s and ECAT4mdouble-as using pBS-SK-ECAT4 6K(-4694/-3670) as template
BM4-ECAT4 5K -5041/+72-sox(m)	Lucifease reporter vector	An BamHI fragment of pBS-SK Nanog sox(m) was introduced into the BamHI site of BM4-ECAT4 5K -5041/+72
BM4-ECAT4 5K -5041/+72-double(m)	Lucifease reporter vector	An BamHI fragment of pBS-SK Nanog double(m) was introduced into the BamHI site of BM4-ECAT4 5K -5041/+72
pENTR Nanog 4735/4395-sox(m)		Mutagenesis with ECAT4mTCF(AC)-s and ECAT4mTCF(AC)-as using pENTR Nanog 4735/4395 as template
pENTR Nanog 4735/4395-double(m)		Mutagenesis with ECAT4mdouble-s and ECAT4mdouble-as using pENTR Nanog 4735/4395 as template
Nanog 4735/4395 pGV-P-sox(m)	Lucifease reporter vector	LR reaction with pENTR Nanog 4735/4395-sox(m) and pGV-P-gw
Nanog 4735/4395 pGV-P-double(m)	Lucifease reporter vector	LR reaction with pENTR Nanog 4735/4395-double(m) and pGV-P-gw
pGV-BM2 -2904/+72 oct3/4(m)		Mutagenesis with Nanog exon up-O(m)S and Nanog exon up-O(m)A using pENTR pGV-BM2 -2904/+72 as template
pGV-BM2 -2904/+72 sox2(m)		Mutagenesis with Nanog exon up-S(m)S and Nanog exon up-S(m)A using pENTR pGV-BM2 -2904/+72 as template
pGV-BM2 -2904/+72 double(m)		Mutagenesis with Nanog exon up-D(m)S and Nanog exon up-D(m)A using pENTR pGV-BM2 -2904/+72 as template
pGV-BM4 -5041/+72 oct3/4(m)	Lucifease reporter vector	A SacI/XhoI fragment of pGV-BM2 -2904/+72 oct3/4(m) was introduced into the SacI/XhoI site of pGV-BM4 -5041/+72
pGV-BM4 -5041/+72 sox2(m)	Lucifease reporter vector	A SacI/XhoI fragment of pGV-BM2 -2904/+72 sox2(m) was introduced into the SacI/XhoI site of pGV-BM4 -5041/+72
pGV-BM4 -5041/+72 double(m)	Lucifease reporter vector	A SacI/XhoI fragment of pGV-BM2 -2904/+72 double(m) was introduced into the SacI/XhoI site of pGV-BM4 -5041/+72
pCR2.1/Sox15 5'-arm		The PCR product amplified with sox15-5arm-s(NotI) and sox15-5arm-As(SpeI) primers was subcloned into pCR2.1-TOPO.
pCR2.1/Sox15 3'-arm		The PCR product amplified with sox15-3arm-s(BamHI) and sox15-3arm-As(XhoI) primers was subcloned into pCR2.1-TOPO.
pBSKS/Sox15 5'arm		A NotI/SpeI fragment of pCR2.1/Sox15 5'-arm was introduced into the NotI/SpeI site of pBS-IRES β geopA
pBSKS/Sox15 5'+3'arm	Targeting Vector	An BamHI/XhoI fragment of pCR2.1/Sox15 3'-arm was introduced into the BamHI/XhoI site of pBSKS/Sox15 5'arm
pCRXL-MC1DApA/Sox15 5'+3'arm	Targeting Vector	An NotI/XhoI fragment of pBSKS/Sox15 5'+3'arm was introduced into the Not/XhoI site of pCRXL-MC1DApA
pCR2.1/Sox2 5'arm		The PCR product amplified with sox2-5arm-sNotI and sox2-5arm-asNheI primers was subcloned into pCR2.1-TOPO.
pCR2.1/Sox2 3'arm		The PCR product amplified with sox2-3arm-sBamHI and sox2-3arm-asXhoI primers was subcloned into pCR2.1-TOPO.
pBSKS/Sox2 3'arm		A BamHI/NotI fragment of pCR2.1/Sox2 3'-arm was introduced into the NotI/SpeI site of pBS-IRES β geopA
pBSKS/Sox2 5'+3'arm	Targeting Vector	A NotI/NheI fragment of pCR2.1/Sox2 5'-arm was introduced into the NotI/SpeI site of pBSKS/Sox2 3'arm

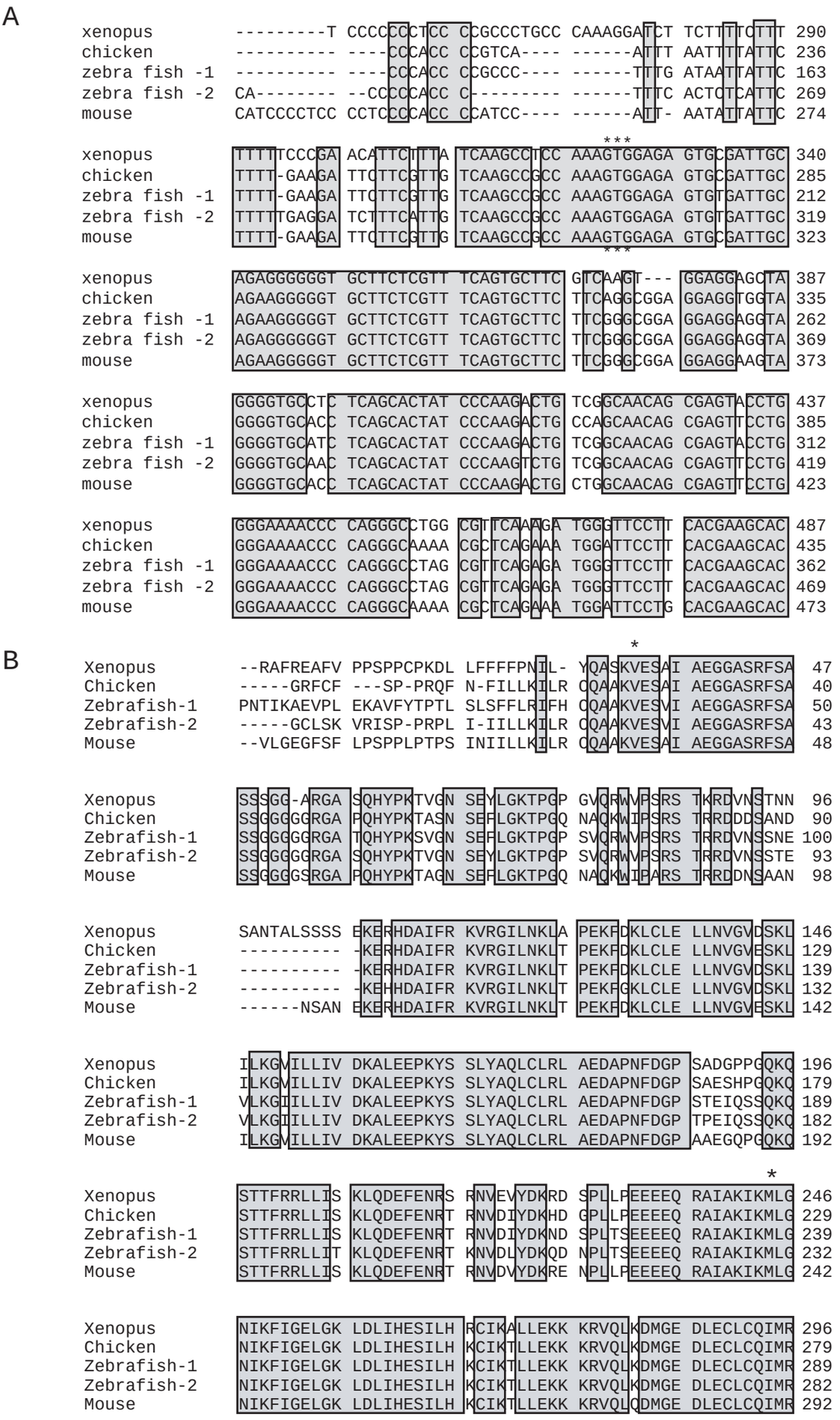


Figure 1.Sequences of vertebrate NAT1 orthologs. (A) Comparison of the nucleotide seqeunces in the 5' portion of vertebrate NAT1 orthologs with their mouse counterpart. Nucleotides conserved in all species are boxed. The GUG initiation codons are marked with asterisks. (B) Comparison of the deduced amino acid sequences of the 5' portion of vertebrate NAT1 orthologs with their mouse counterpart. Conserved amino acids are boxed. Asterisks indicate the valine encoded by the initiator GUG and the methionine encoded by the first AUG codon in the open reading frame.

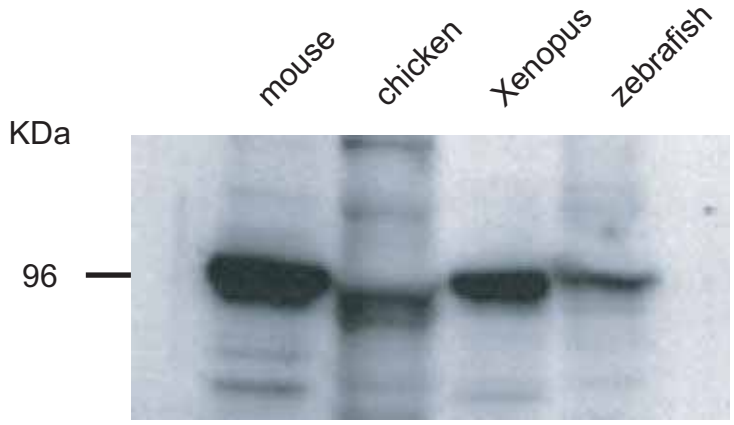


Figure 2. Western blot analysis of vertebrate NAT1 orthologs. mouse, chicken, Xenopus, and zebrafish.

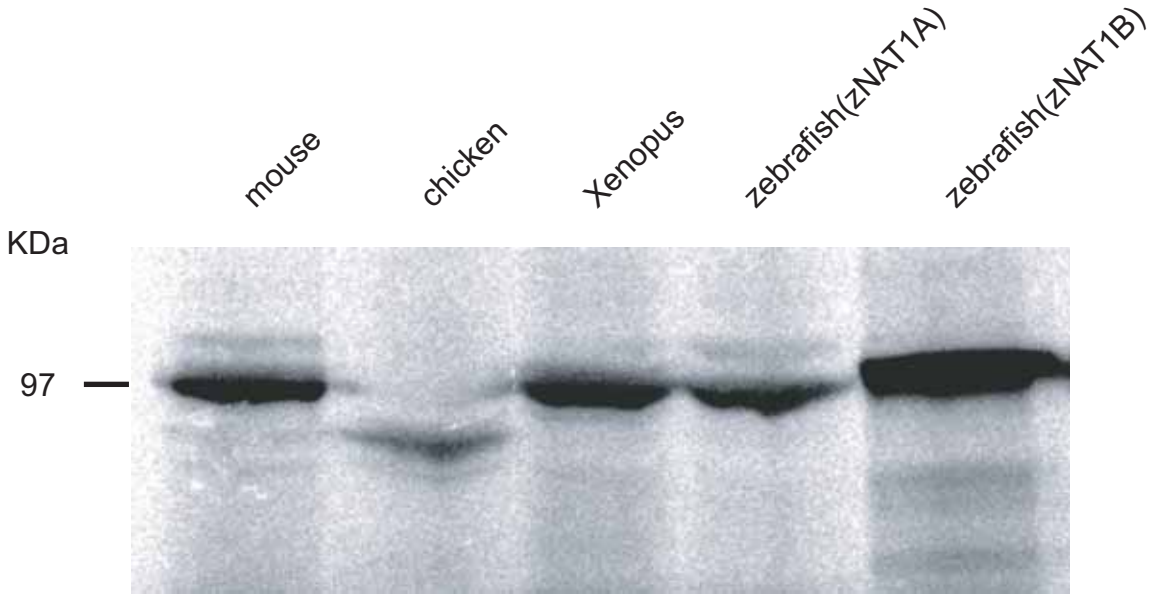
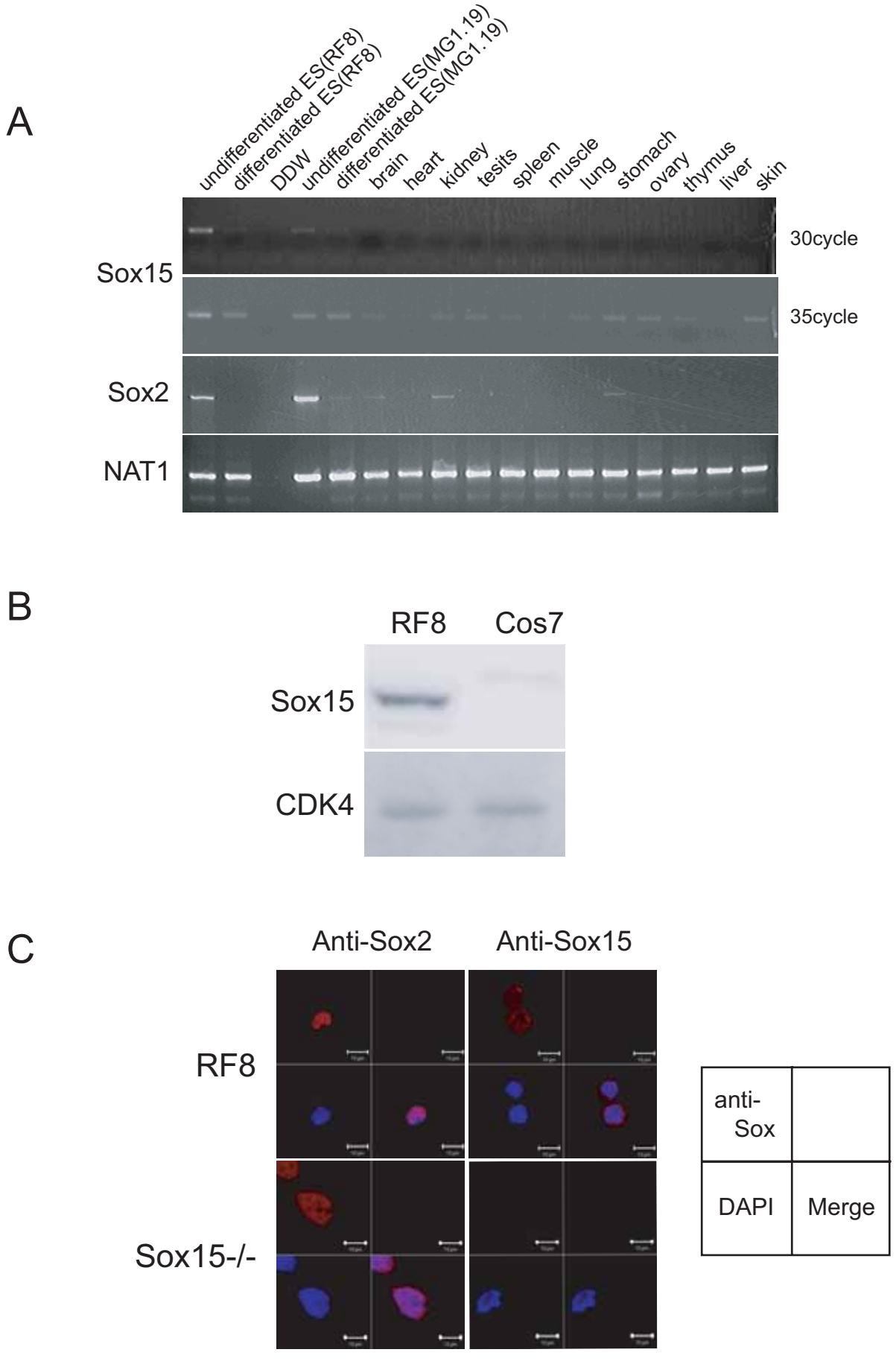


Figure 3. In vitro transcription/translation of vertebrate NAT1 orthologs. mouse, chicken, Xenopus, and zebrafish(zNAT1A), (zNAT1B)



Figuire 4. (A) RT-PCR analysis showing the expression profiles of mouse Sox15, Sox2, and NAT1. Expression of these genes was tested in two independent ES cell line, MG1.19 and RF8, as well as 12 somatic tissues from adult mice. Differentiation of ES cells was induced by retinoic acid treatment for five days. (B) Sox15 protein levels in Cos7 cells and ES cells were determined by Western blot analyses. CDK4 was used asa loading contorol. (C) Immunopstaining of Sox15 and Sox2 in wild-type ES cells and Sox15 knockout ES cells. Cells were counterstained with DAPI.

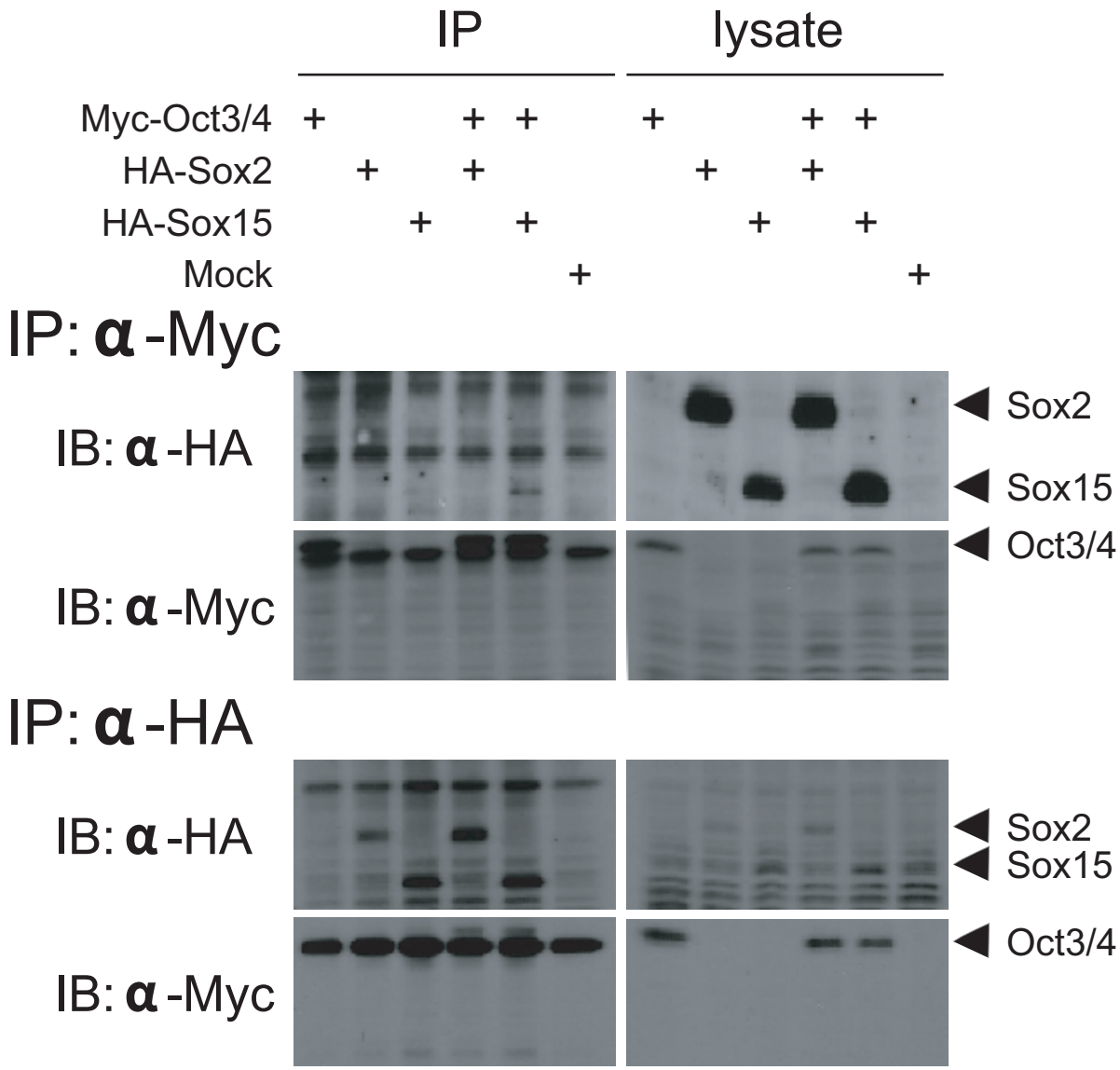
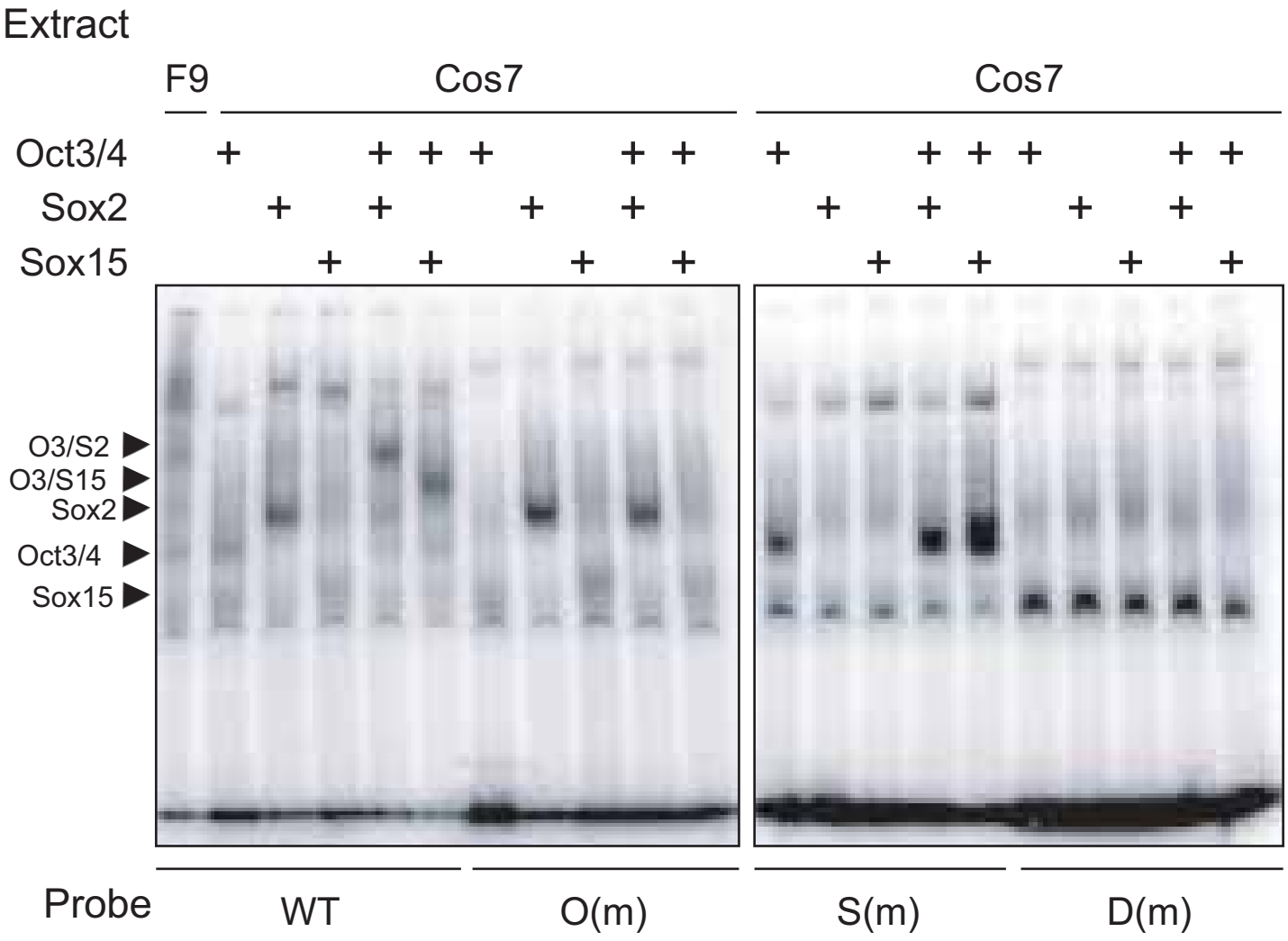
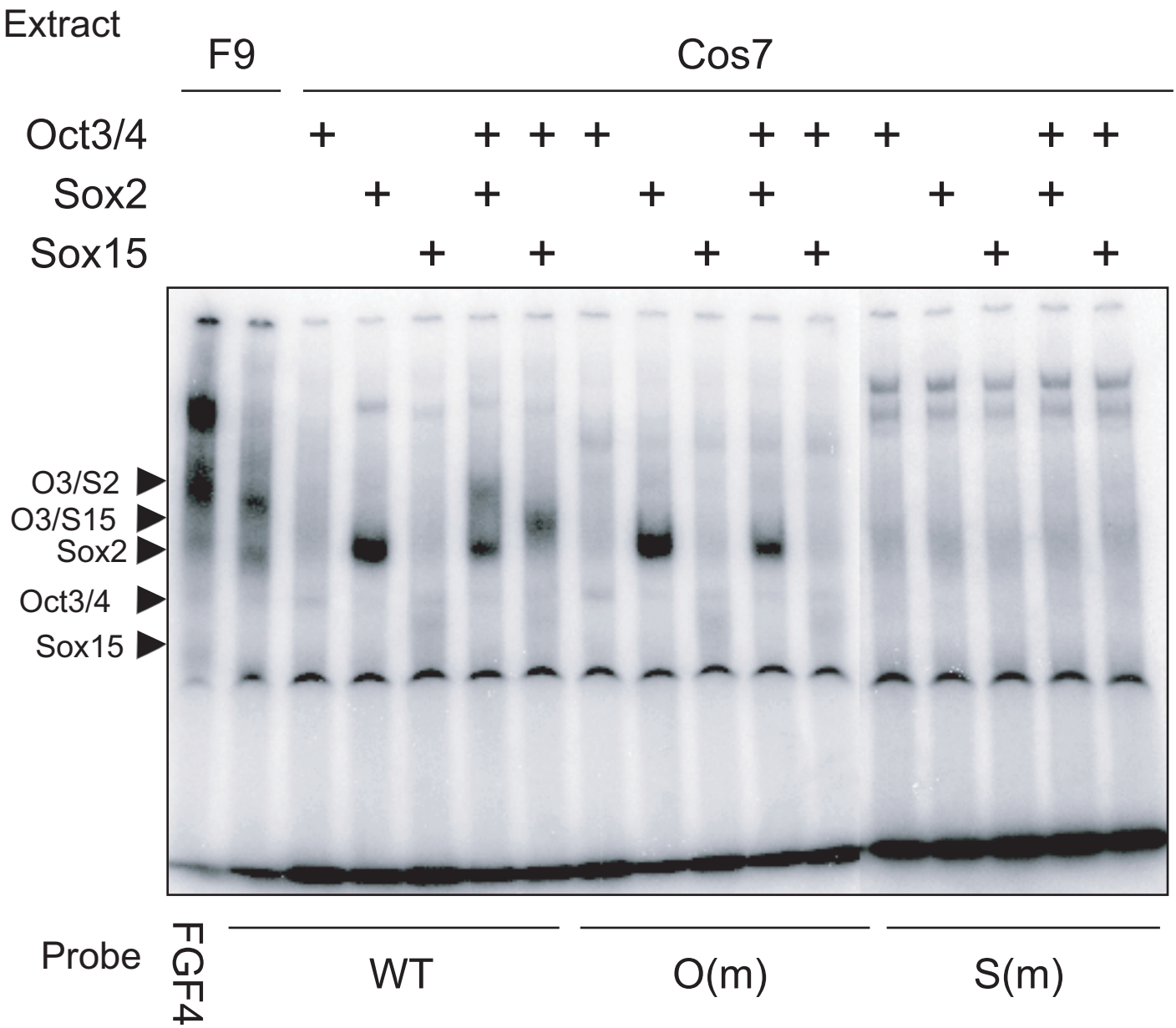


Figure 5.Westerm blot analyses showing association of Sox15 with Oct3/4. MG1.19 ES cells were transfected with HA-Sox2, HA-Sox15, and myc-Oct3/4 expression vectors. Immunoprecipitation (IP) was performed with either anti-myc or anti-HA antibody. Cell lysates and precipitated samples were separated by SDS-PAGE and immunoblotted with anti-myc or anti-HA antibody.Shown on the left are the positions of bands corresponding to each transcription factor(s).



	POU		SOX
FGF4-WT	ATTAGCAT	CCA	AACAAAG
FGF4-Oct(m)	**GGGTTACC**	CCA	AACAAAG
FGF4-SOX(m)	ATTAGCAT	CCA	A**CC**CAA**T**
FGF4-Double(m)	**GGGTTACC**	CCA	A**CC**CAA**T**

Figure 6. Gel mobility-shift assay showing binding of Sox15 to the FGF4 enhancer. A ^{32}P-labeled oligonucleotide (WT;tttaagtatcccATTAGCATccaAACAAAGagttttcta) corresponding to the FGF4 enhancer was incubated with F9 EC cell extract or Cos7 cell extract expressing Oct3/4, Sox2, or both. Oligonucleotides with mutations (underlined) in the octamer-like sequence [Oct(m);tttaagtatcccGGGTTACCccaAACAAAGagttttcta], the Sox-binding site [Sox(m);tttaagtatcccATTAGCATccaACCCAATagttttcta] and the both sites [Double(m);tttaagtatcccGGGTTACCccaACCCAATagttttcta] were also tested. Shown on the left are the positions of bands corresponding to each transcription factor.



	POU	SOX
Fbx15-WT	TTTATCAT	AACAATG
Fbx15-Oct(m)	T**CCC**TCAT	AACAATG
Fbx15-SOX(m)	TTTATCAT	AAC**C**ATG

Figure 7. Gel mobility-shift assay showing binding of Sox15 to the Fbx15 enhancer. A ^{32}P-labeled oligonucleotide (WT; ccagatgtgcTTTATCATAACAATGgaattcctaggggct) corresponding to the Fbx15 enhancer was incubated with F9 EC cell extract or Cos7 cell extract expressing Oct3/4, Sox2, or both. Oligonucleotides with mutations in the octamer-like sequence [Oct(m); ccagatgtgcTCCCTCATAACAATGgaattcctaggggct] and the Sox-binding site [Sox(m); ccagatgtgcTTTATCATAACCATGgaattcctaggggct] were also tested. As a control, an FGF4 enhancer oligonucleotide (tttaagtatcccATTAGCATccaAACAAAGagttttcta) was incubated with Cos7 cell extracts expressing Oct3/4, Sox2, or both. Shown on the left are the positions of bands corresponding to each transcription factor.

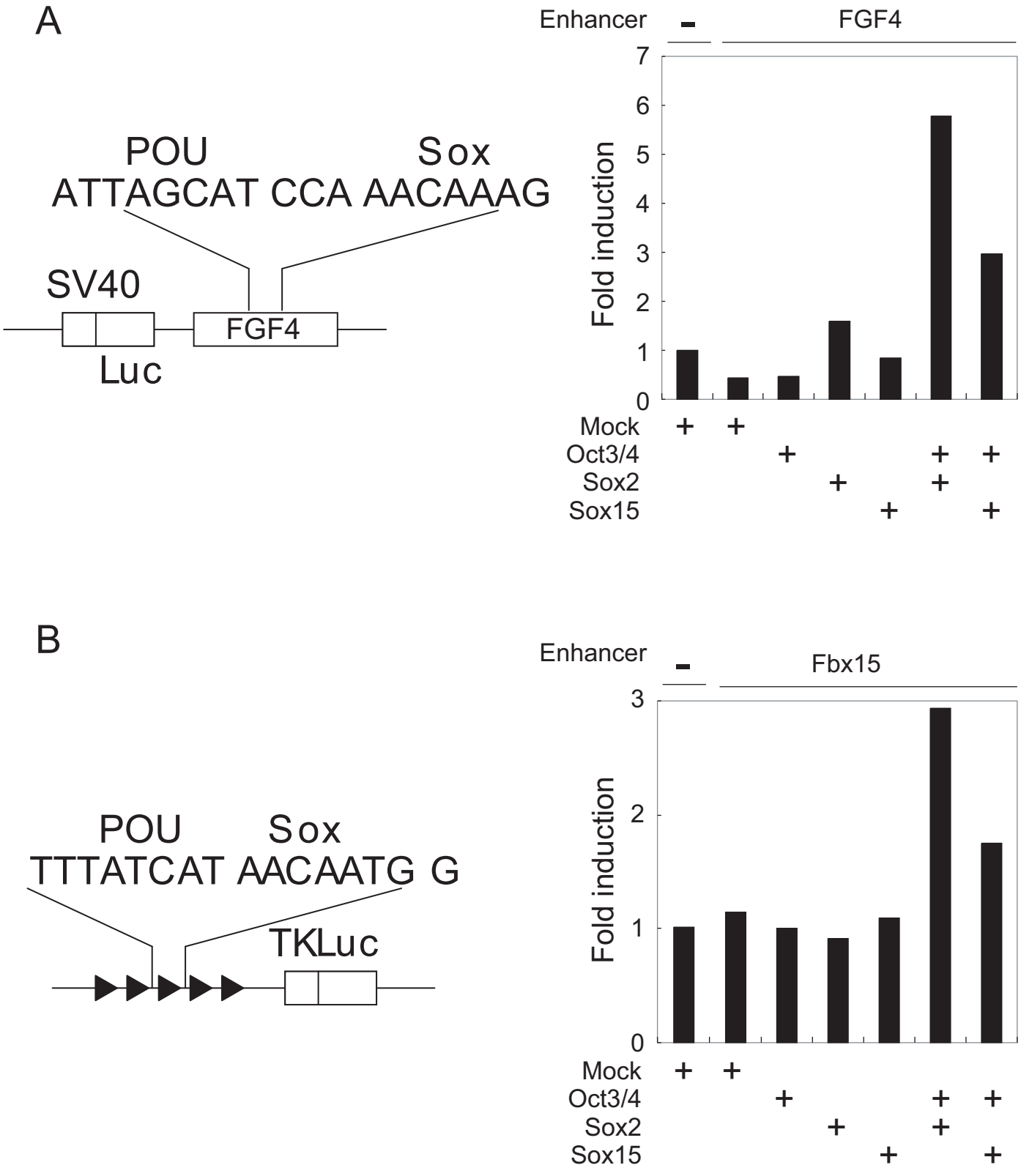


Figure 8. Activation of FGF4(A) and Fbx15(B) enhancer by Oct3/4 and Sox15. Luciferase reporter plasmids containing the FGF4 or Fbx15 enhancer was introduced into Cos7 cells together with expression vector for Sox2, Sox15 and/or Oct3/4 as indicated.

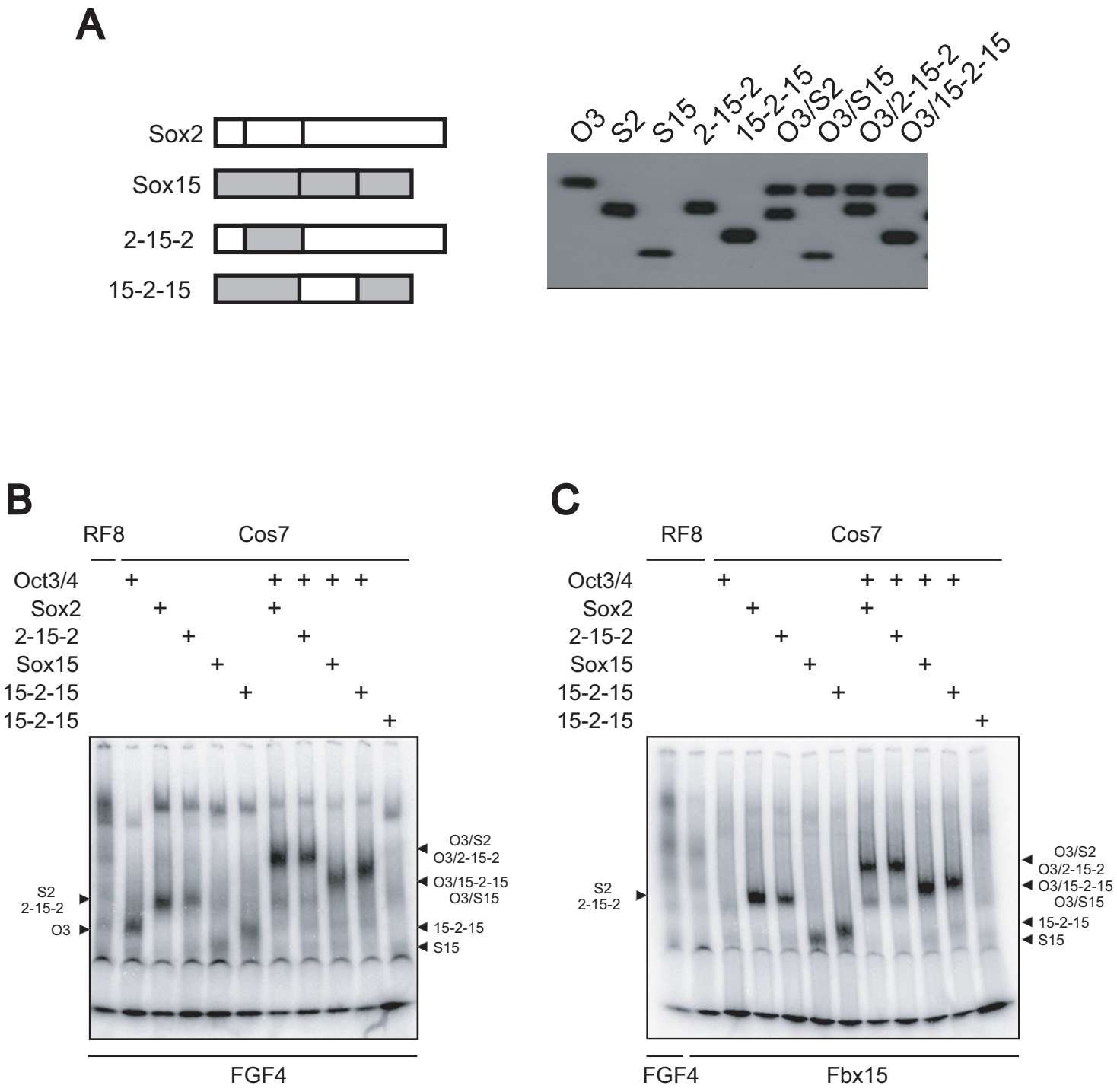
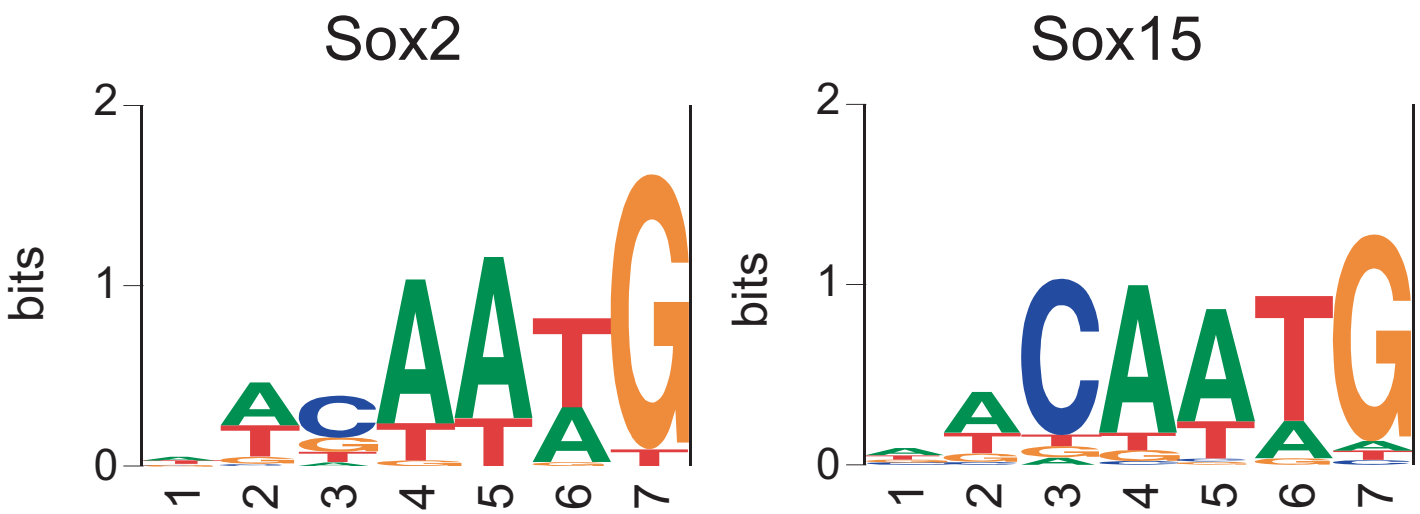


Figure 9. Expression and binding potency of Sox-2/15 chimeric proteins. (A) Western blot analysis showing the expression levels of Sox-2/15 chimeric proteins. (B & C) Electrophoresis mobility shift assay showing binding potency of Sox-2/15 chimeric proteins to the FGF4 enhancer (B) and the Fbx15 enhancer (C).

A



B

		Sox2	Sox15	O3/S2	O3/S15
WT	5'- A A C A A T G - 3'	100	100	100	100
1T	T - - - - - -	85	68	88	45
1G	G - - - - - -	107	71	79	101
1C	C - - - - - -	151	80	141	68
2T	- T - - - - -	22	34	95	68
2G	- G - - - - -	0	10	17	17
2C	- C - - - - -	8	19	34	47
3T	- - T - - - -	15	25	72	34
3A	- - A - - - -	0	6	24	7
3G	- - G - - - -	2	20	190	63
4T	- - - T - - -	0	15	0	27
4G	- - - G - - -	0	16	0	7
4C	- - - C - - -	2	1	0	0
5T	- - - - T - -	0	10	8	36
5G	- - - - G - -	0	11	77	31
5C	- - - - C - -	7	6	29	41
6A	- - - - - A -	169	92	260	120
6G	- - - - - G -	23	20	102	129
6C	- - - - - C -	20	9	128	54
7T	- - - - - - T	127	90	205	146
7A	- - - - - - A	344	215	219	171
7C	- - - - - - C	76	66	223	239

Figure 10. Comparison of DNA recognition sequence between Sox2 and Sox15.(A) DNA recognition sequences of Sox2 and Sox15 determined by SELEX. Results are shown with Sequence logo (Schneider and Stephens, 1990), which generated by WebLogo (www.bio.cam.ac.uk/cgi-bin/seqlogo/logo.cgi). The height of each Letters are sorted so that the most common one is on top. The height of the entire stack is adjusted to signify the information content (measured in Bits) of the sequence at that position.
(B) Determination of the relative importance of individual nucleotides within the Sox binding site for the interaction with Sox factors. The radio-labeled probes were generated using oligonucleotide which contained wild-type (5'-AACAATG-3') or mutated Sox sequence as templates and used for the gel shift analyses using whole cell extracts from Cos7 cells in which indicated expression vectors had been transfected. In the mutants, only mutated nucleotides are indicated.

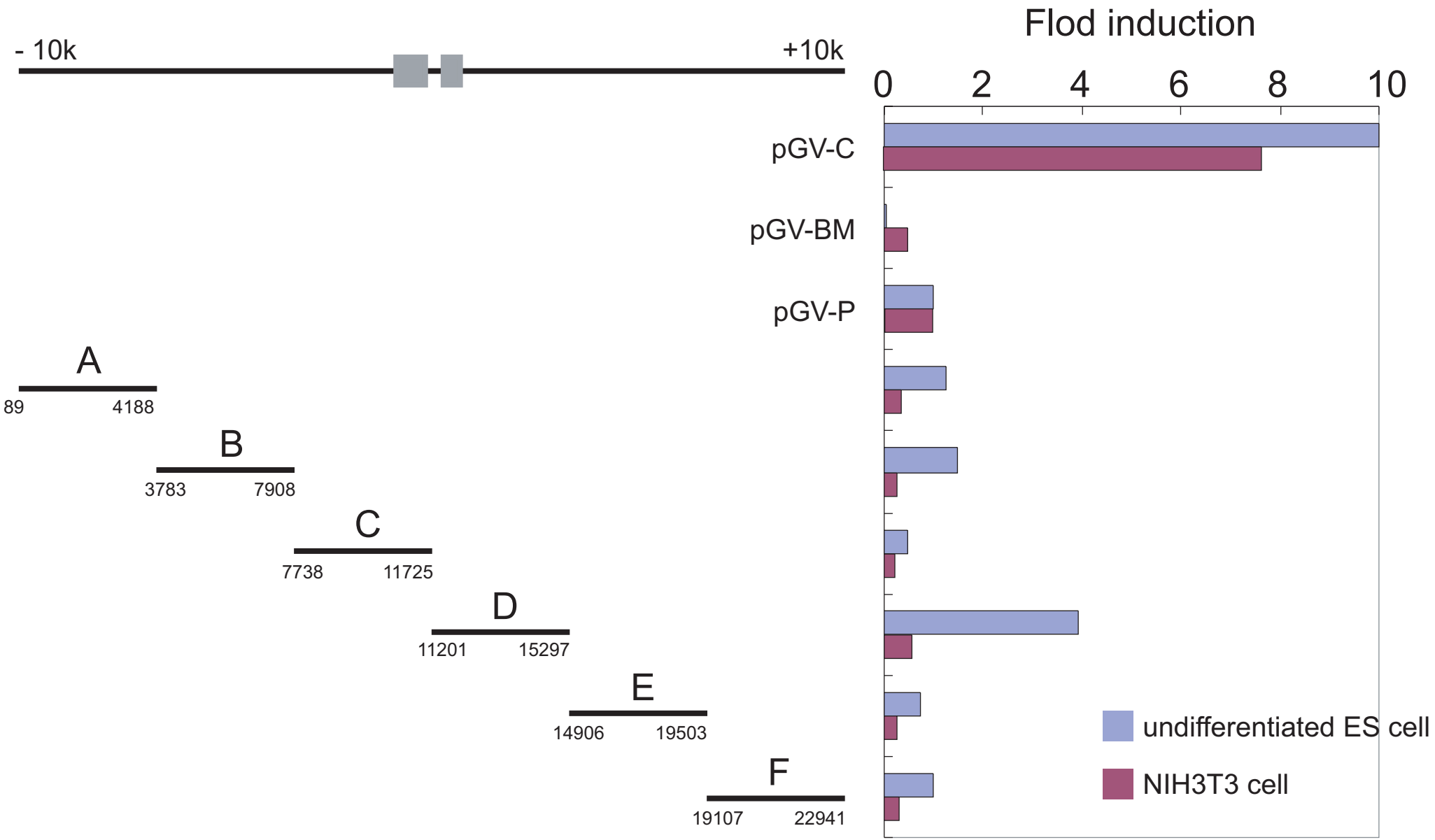


Figure 11. Enhancer analyses with the luciferase reporter genes. DNA fragmets of approximately 4.0-kb were isolated from the mouse Sox15 gene and placed in downstream of the luciferase gene driven by the minimum SV40 promoter. These reporter constructs were introduced into undifferentiated ES cells (blue) and NIH3T3 cells (purple). Shown is fold induction of normalized luciferase activity compared to that of the enhancerless construct (pGV-P).

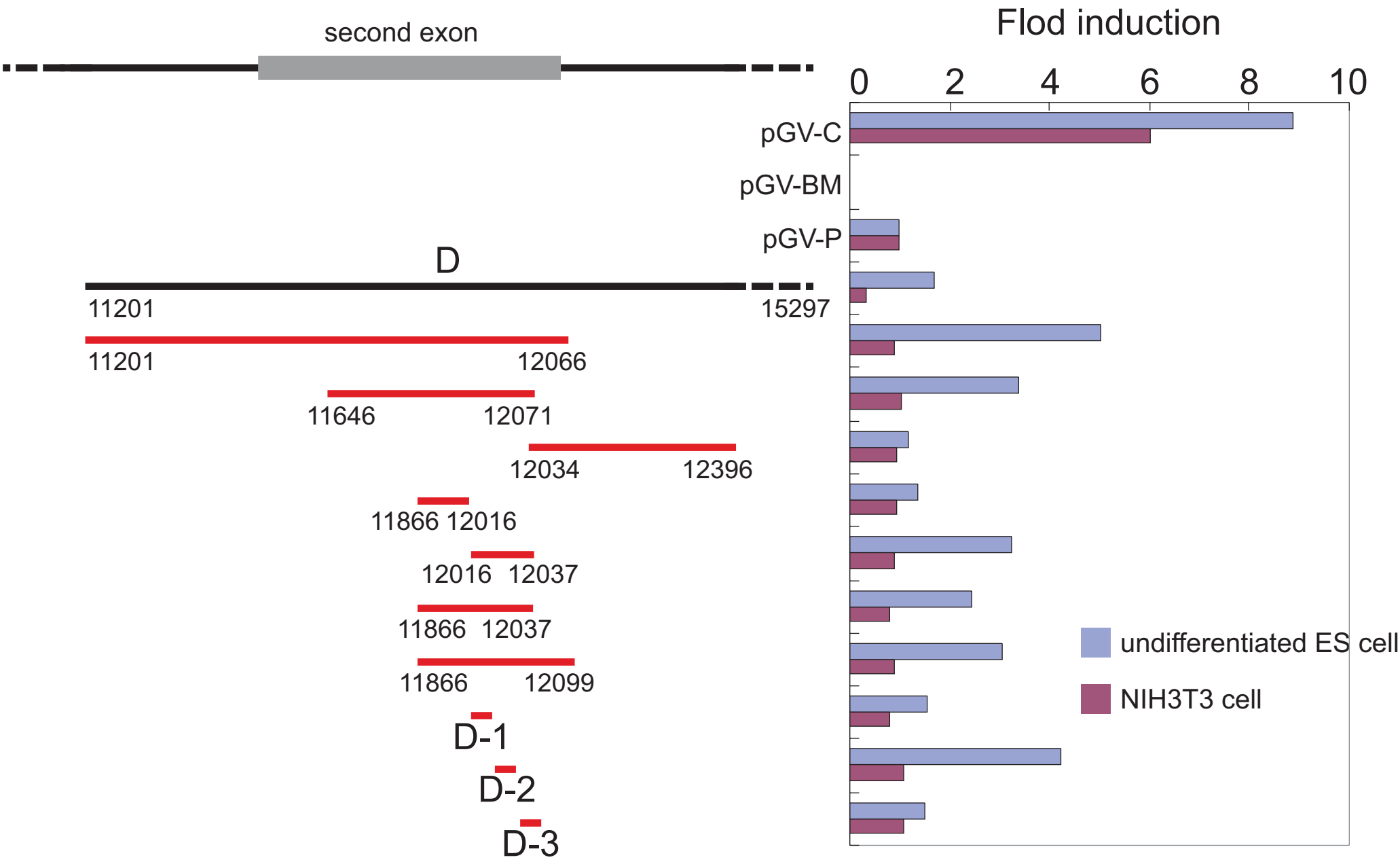


Figure 12. Enhancer analyses with the luciferase reporter genes. Shorter DNA fragmets of various sizes were isolated from the fragment D in Fig11 and placed downstream of the luciferase gene driven by the minimum SV40 promoter. These reporter constructs were introduced into undifferentiated ES cells (blue) and NIH3T3cells (purple). Shown is fold induction of normalized luciferase activity compared to that of the enhancerless construct (pGV-P).

A

CACCCTTCCGCCAGGACCTG

CCTCGGCAGGAACCTCCCC

B

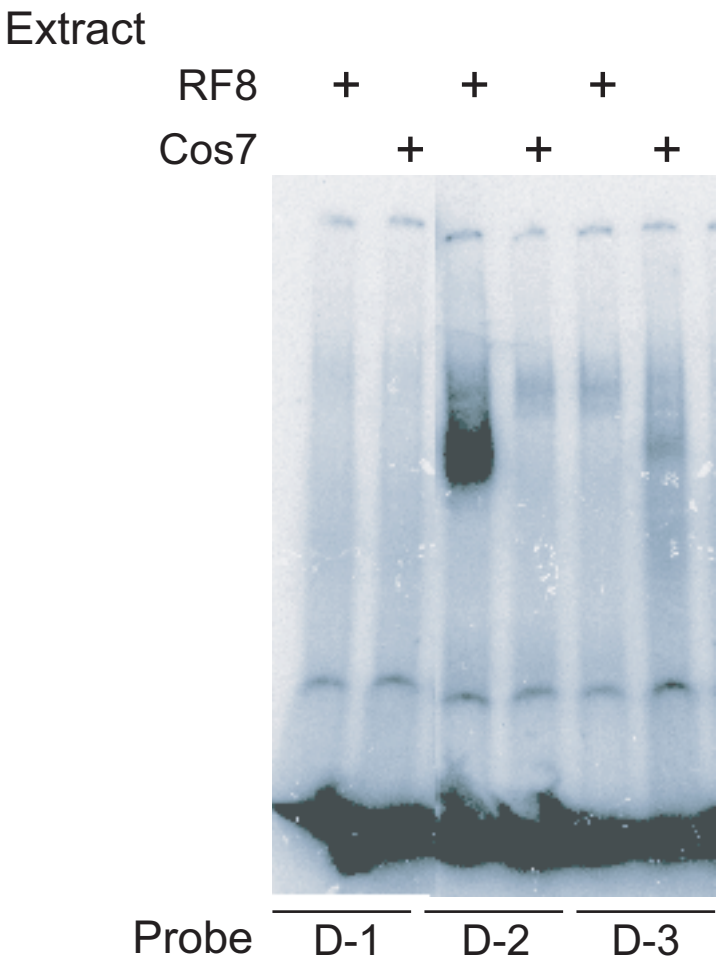


Figure 13. (A) Sequence of the Sox15 enhancer (D-2) that shows luciferase activity in undifferentiated RF8 ES cells. (B) Gel mobility shift assay. A ^{32}P-labeled oligonucleotide(D-2;CACCCTTCCGCCAGGAC-CTGCCTCGGCAGGAACCTCCCC) corresponding to the Sox15 enhancer was incubated with RF8 ES cell extract or Cos7 extract. Fragment D-1 and D-3 were used as negative control.

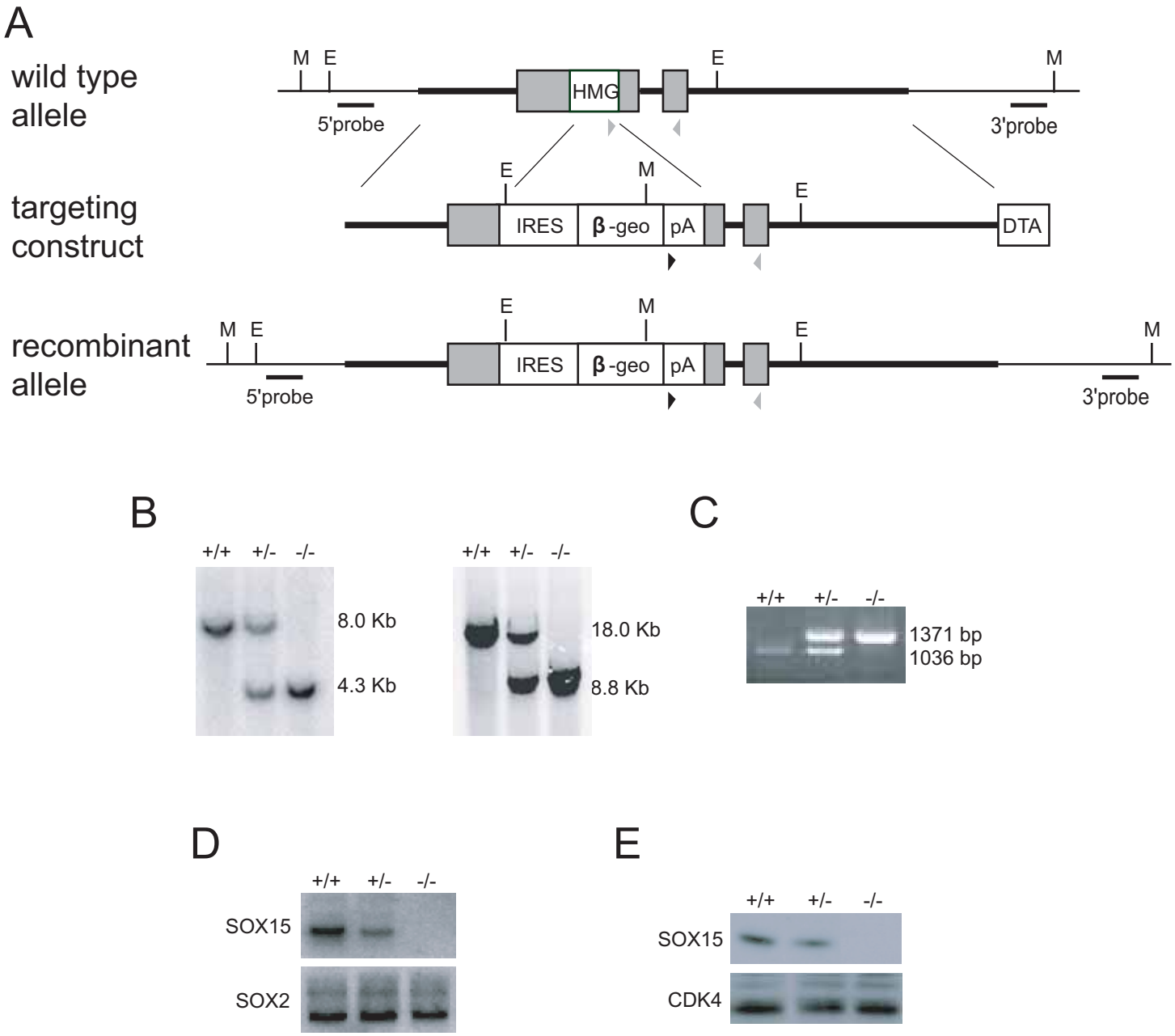


Figure 14. Targeted disruption of the mouse Sox15 gene. (A) Structure of the Sox15 genomic locus, the targeting vector and the targeted locus generated by homologous recombination. The targeting vector conatins the β-geo cassette in place of the HMG domain. The length of the diagnostic EcoRI(E) or MunI(M) restriction fragments and the locations of the 5' or 3' probe for Southern blot analysis are shown.
(B) Southern blot analysis. Specific hybridization with the 5' probe produces an 8.0-kb band from the wild-type locus and a 4.3-kb band from the target locus. Hybridization with 3' probe produced an 18.0-kb band from the wild-type locus and an 8.8-kb band from the target locus.+/+, +/-, and -/- represent genotypes of Sox15+/+, Sox15+/-, and Sox15-/-, respectively. (C) PCR analysis.PCR with the three primers shown in panel A produces a 1036-bp band from the wild-type locus and a 1371 -bp band from the target locus.
(D) Northern blot analysis. Total RNA isolated from ES cells of each genotype was separated, bloted, and hybridized to cDNA probes of Sox15 and Sox2. (E) Western blot analysis. Lysates isolated from ES cells of each genotype were separated on SDS-PAGE and immunoblotted with anti-Sox15 and anti-CDK4 antibody.

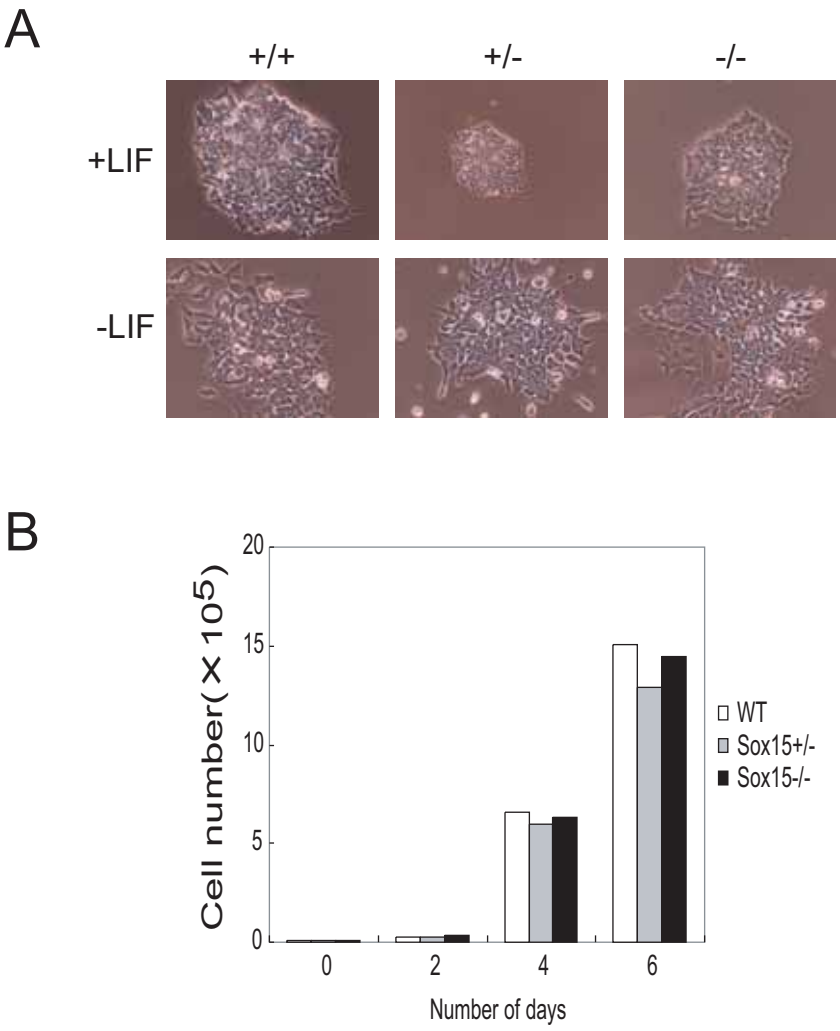


Figure 15. Analyses of Sox15 null cells. (A) Morphology of wild-type ES cells(+/+), Sox15 heterozygous cells(+/-), and Sox15 null cells(-/-) grown without STO feeder cells. (B) Effect of Sox15 deficiency on proliferation.Three genotype Cells were plated at 1×10^4 cells per well of 24-well plates. Cell numbers were counted after 2, 4, and 6 days. (C) Northern blot analyses. Total RNA was isolated from parent RF8 cells(+/+), heterozygous cells(+/-), Sox15 null cells(-/-), and rescue cells in which the Sox15 expression vectors had been introduced into Sox15-null ES cells(rescue). Expression of FGF4, Fbx15, and UTF1 was determined by northern blot analysis.

	♂	♀	Total	%	expected %
WT	7	4	11	10	6.25
S15+/-	9	11	20	19	12.5
S2+/-	6	10	16	15.2	12.5
Double+/-	13	23	36	34.3	25.0
S15-/-	7	2	9	8.6	6.25
S2+/- S15-/-	7	6	13	12.4	12.5
S2-/-	-	-	-	0	6.25
S2-/- S15+/-	-	-	-	0	12.5
Double-/-	-	-	-	0	6.25
Total	49	56	105		

Figure 16.Analysis of progeny from Sox2$^{\beta geo}$/Sox15$^{\beta geo}$ heterozyous intercross.

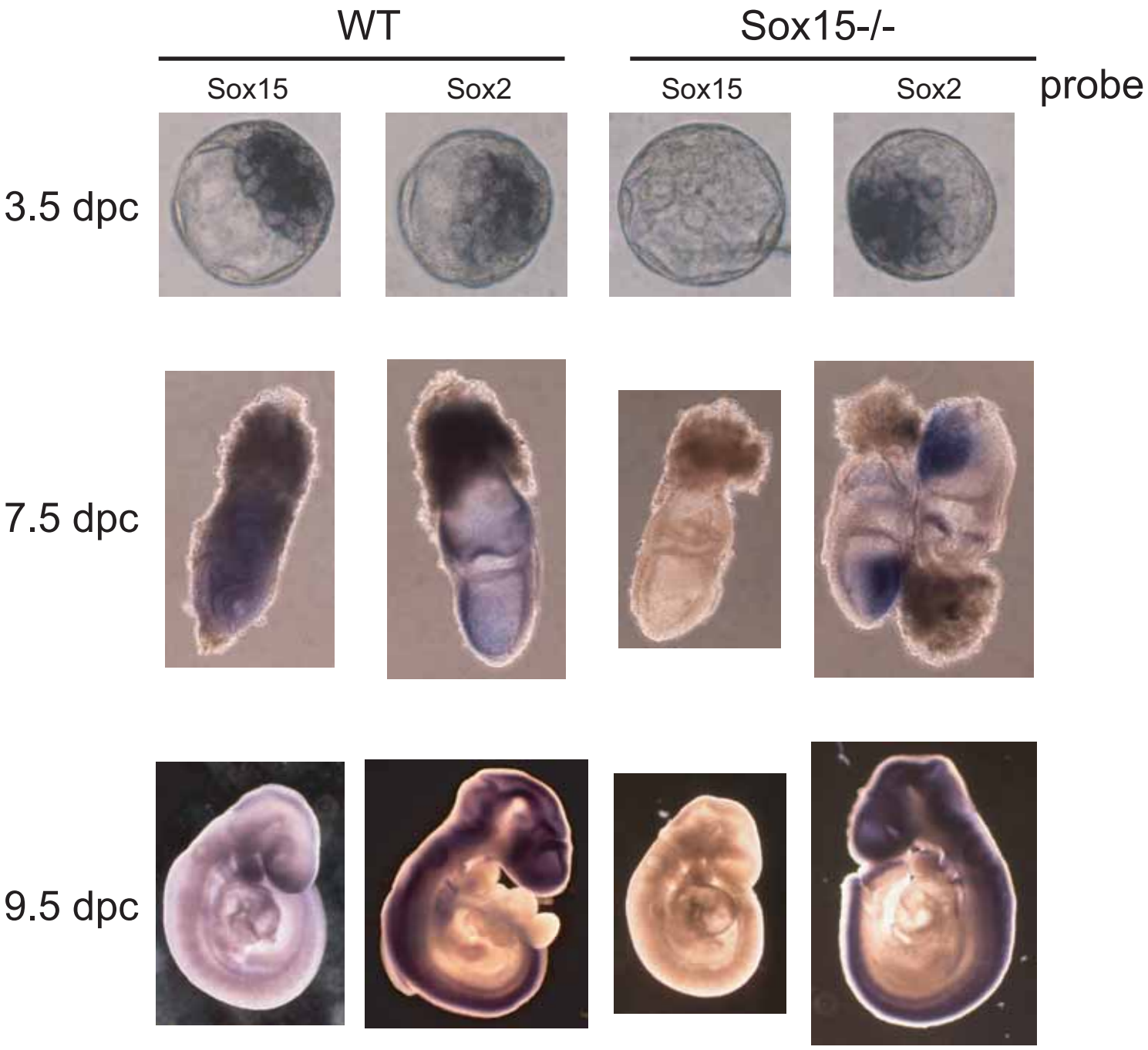


Figure 17. Expression of Sox15 RNA in mouse embryos. Whole-mount in situ hybridization in wild-type and Sox15-null mutant embryos. As a control, Sox2 probe was used.

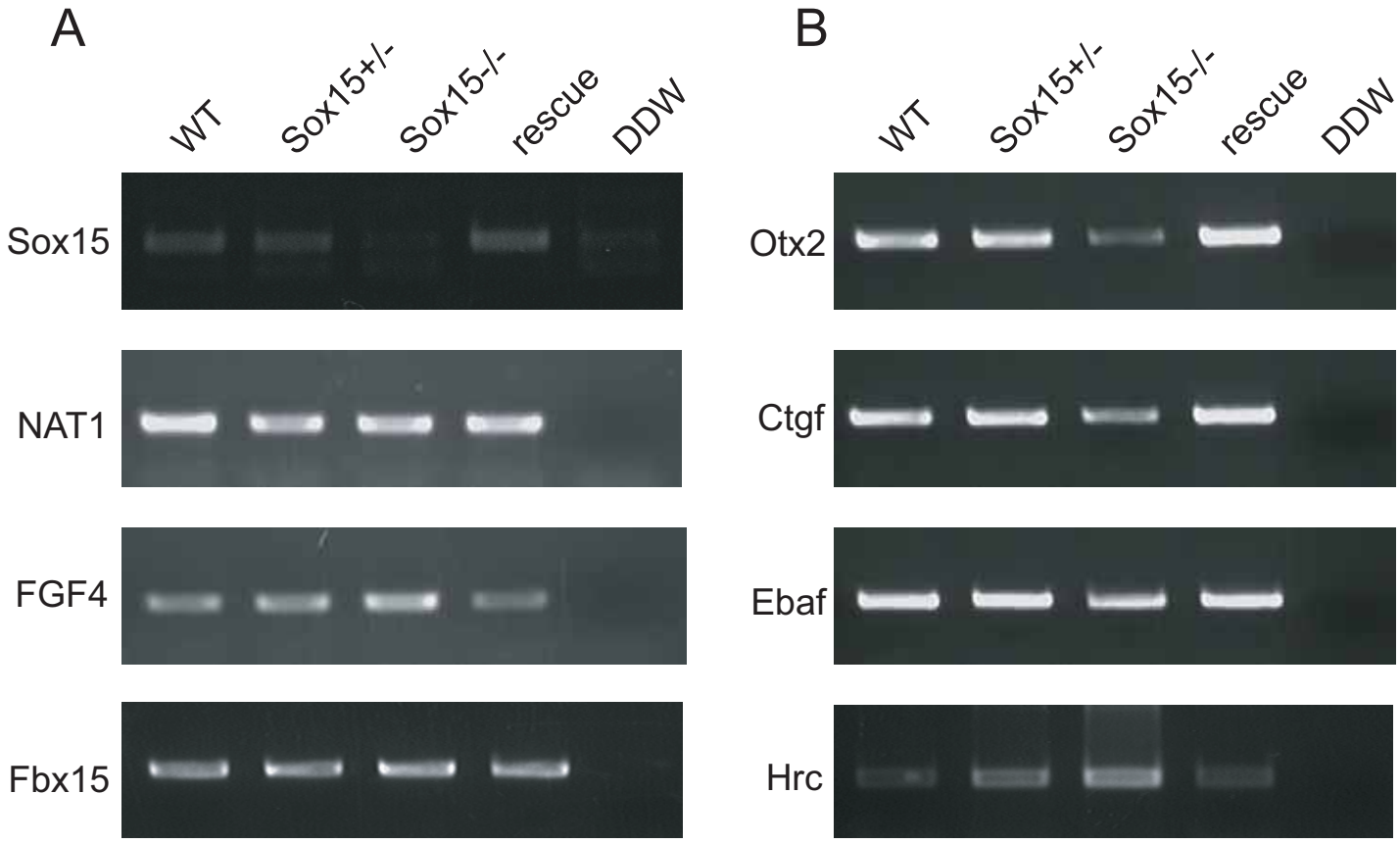


Figure 18. Identification of genes regulated by Sox15 in vivo. (A) The expression of Sox15 and two Sox2 target gene, FGF4 and Fbx15. NAT1 was used as a loading control. (B) The expression of genes that were found up- or down-regulated in Sox15-null cells by DNA microarry.

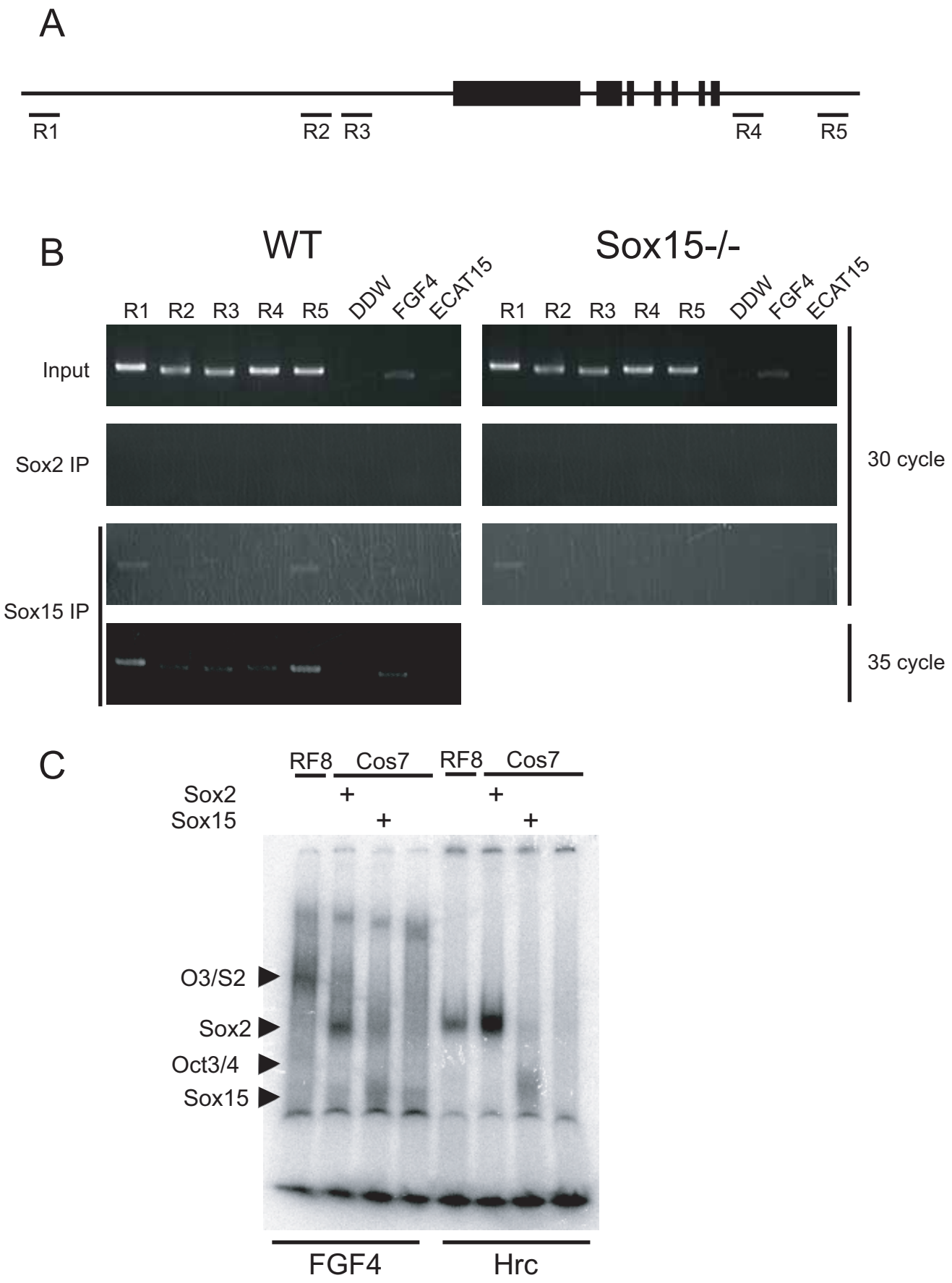


Figure 19. Identification of a Sox15 binding site in the Hrc gene. (A) Putative Sox binding sites in the mouse Hrc gene. (B) ChIP analysis. Chromatin immunoprecipitation was performed in wild-type ES cells and Sox15 knockout cells with anti-Sox15, anti-Sox2 and normal rabbit IgG. The final DNA extractions were amplified by PCR. (C) Gel mobility-shift analysis. A 32P-labeled oligonucleotide (Hrc;cggatcatcaataaatTACAAAGttga-gaaaagggagac) corresponding to the R5 site was incubated with wild-type ES cell extract or Cos7 cell extract expressing Sox15 or Sox2. As a control, an FGF4 enhancer oligonucleotide(tttaagtatcccATTAGCA-TccaAACAAAGagttttcta) was used. Shown on the left are the positions of bands corresponding to each transcription factor.

-4737 ACCCGGAGCT GTGCGCCCTG TACCAAACCT TTGTAGAACT TGGGGTAAAC

LEF/TCF1

-4687 TTAAGGCTAT GGTGGCCTTG ACTCCGTGGA CCCAGAGGCA AGTTTCCTCC

-4637 TTTAGAGGAC TCGCATGCAT TTTGTTTCTA ATTTGAAATG AGAACCGGCT

Nanog

-4587 TAGAGCTTGA ACCAGCCAGT TCTCTGGACT CCTCCCAGCT CTTACAATTC

-4537 CTCTCCCGGA CGGTTCCTAG AAGACAAAGG CAAGCTTACC AAAATTACGT

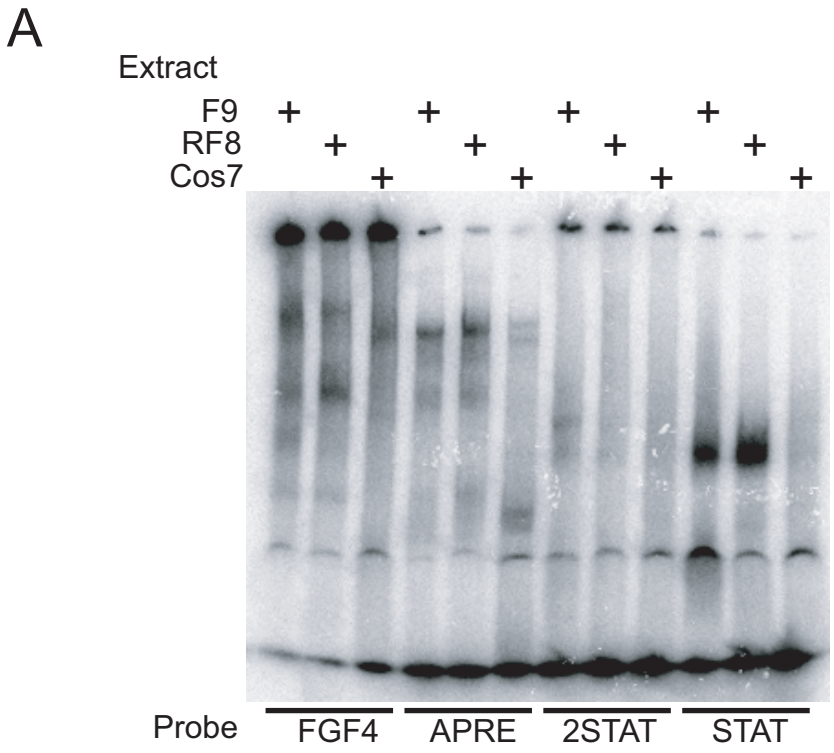
STAT3 LEF/TCF1

-4487 CGCCCTTGGG ACACACCTAG GGTTCCCTGG TGGCATCTTT TTTTTTTCAT

-4437 TATAAACAGG AGTAAATTTT TGTAAGGGCA GAGCTGGTAG CTGAGGGAGA

-4387 G

Figure 20. Sequence of Nanog enhancer. This enhancer was identified by luciferase analysis and contains putative STAT3, LEF/TCF1, and Nanog binding sites.



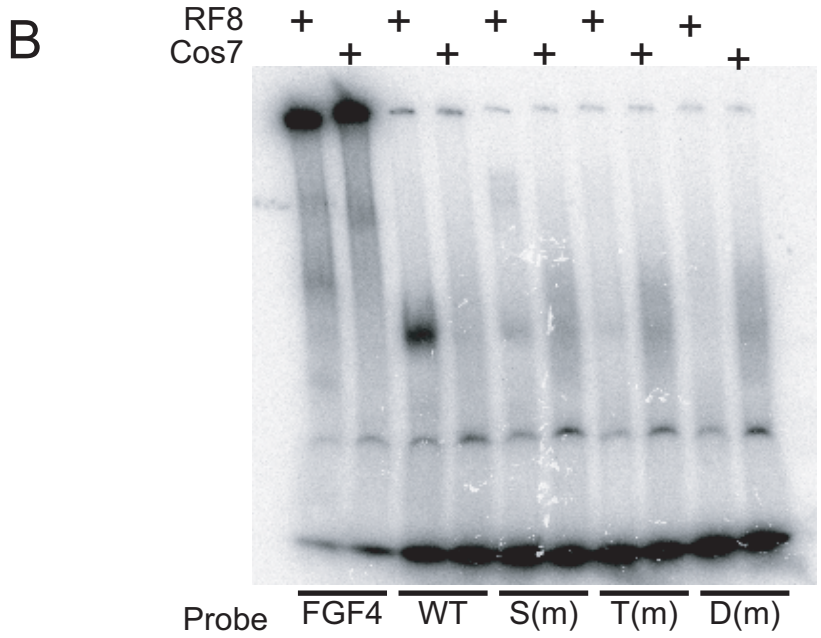


Figure 21. Gel mobility shift assay of Nanog enhancer. (A) A ^{32}P-labeled oligonucleotide(2STAT:ATCCttccggaaTTCTGATCCttccgggaaTTCTG, STAT:CTCTCCCGGACGGttcctagaaGACAAAGGCAAGCT) corresponding to a putative STAT binding site in the Nanog enhancer was incubated with F9 EC , RF8 ES cell or Cos7 extract. As a control, the FGF4 enhancer oligonucleotide(TAGAAAACTCTTTGTTTGGATGCTAATGGGATACTTAAA) and the STAT3 high affinity binding site(APRE:ATCCTTCCGGGA ATTCTGATCCTTCCGGGAATTCTG)was incubated with each extract. (B) Mutation assay. A ^{32}P-labeled oligonucleotide correspond to the wild-type Nanog enhancer (WT:CTCTCCCGGACGGttcctagaaGACAAAGGCAAGCT), a Sox binding site mutation (S(m):CTCTCCCGGACGGttcctagtcGACAAAGGCAAGCT), a TCF binding site mutation (T(m):CTCTCCCGGACGGttcctagaaGAACAAGGCAAGCT), and a double mutation (D(m):CTCTCCCGGACGGttcctagtcGAACAAGGCAAGCT) was incubated with RF8 ES cell extract or Cos7 extract. As a control, an FGF4 enhancer oligonucleotide was used.

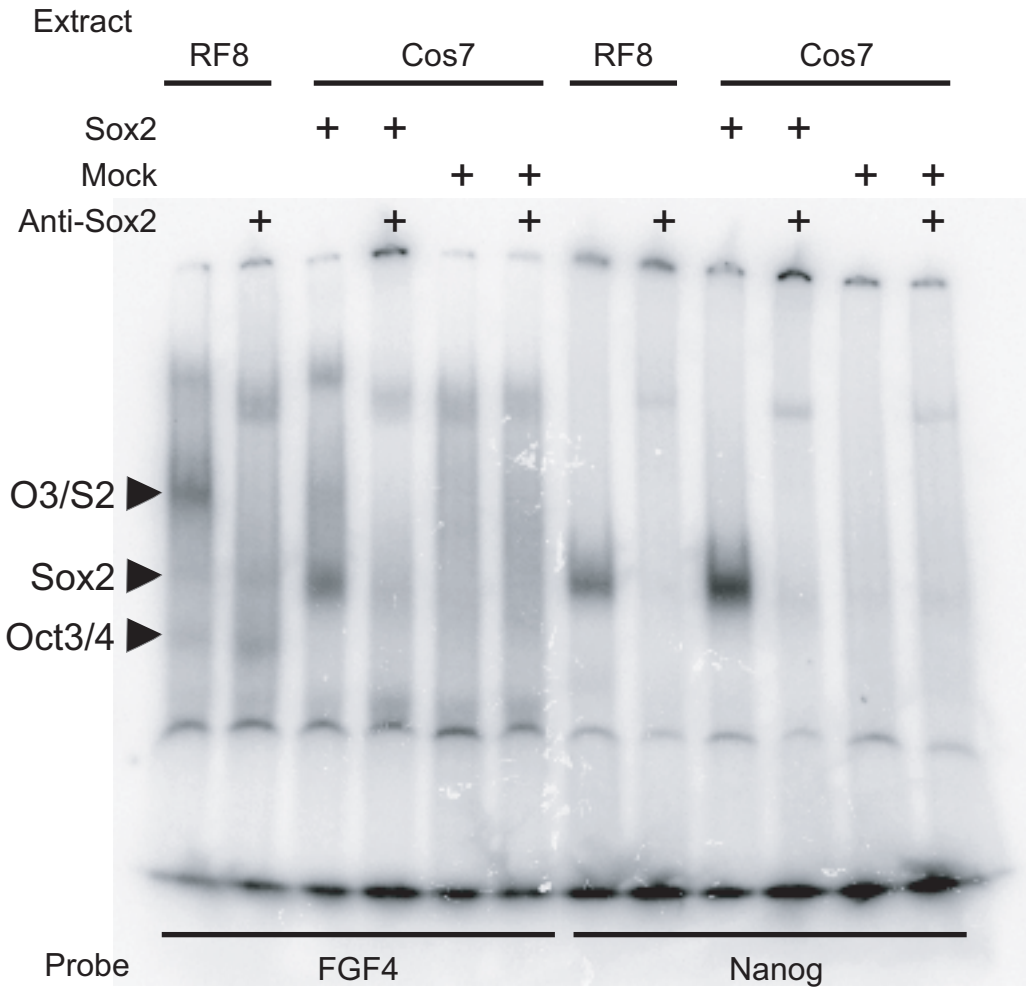


Figure 22. Supur shift assay with anti-Sox2 antibbody. A ^{32}P-labeled oligonucleotide(WT:CTCTCCCGGA-CGGttcctagaaGACAAAGGCAAGCT) corresponding to the Nanog enhancer was incubated with extracts from RF8 ES cells or Cos7 cells transfected with the Sox2 expression vector or a mock plasmid.
The reaction was performed with or without anti-Sox2 anti-serum. As a control, the FGF4 enhancer oligo-nucleotide was used.

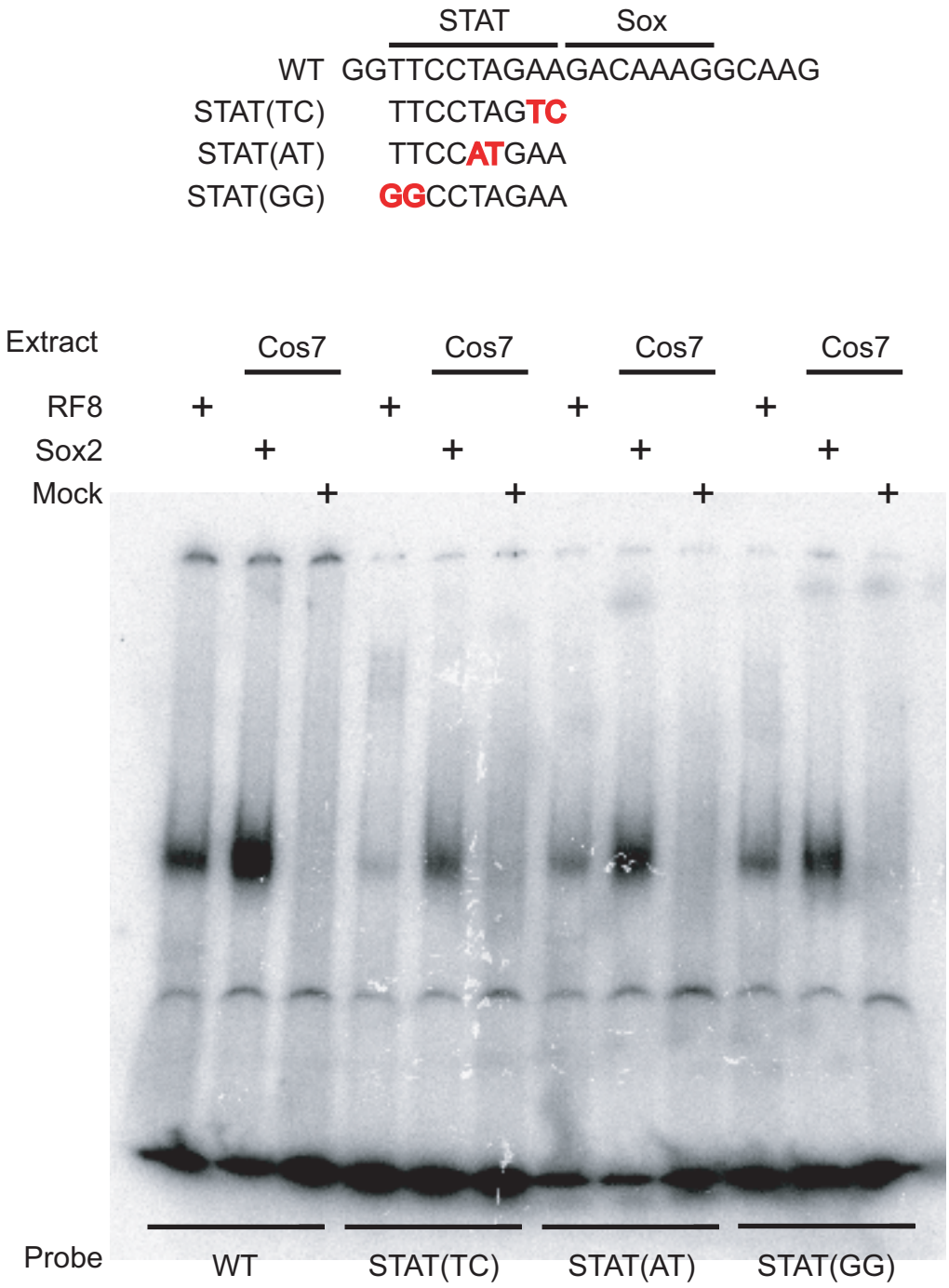


Figure 23. Gel mobility shift assay of STAT mutant probes. A ^{32}P-labeled oligonucleotides corresponding the wild-type (WT:CTCTCCCGGACGGttcctagaaGACAAAGGCAAGCT) and three indicated STAT-mutated sequences (STAT(TC):CTCTCCCGGACGGttcctagtcGACAAAGGCAAGCT, STAT(AT):CTCTCCCGGACGG-ttccatgaaGACAAAGGCAAGCT, STAT(GG):CTCTCCCGGACGGggcctagaaGACAAAGGCAAGCT) were incubated with extracts from RF8 ES cells or Cos7 cells transfected with the Sox2 expression vector or a mock plasmid.

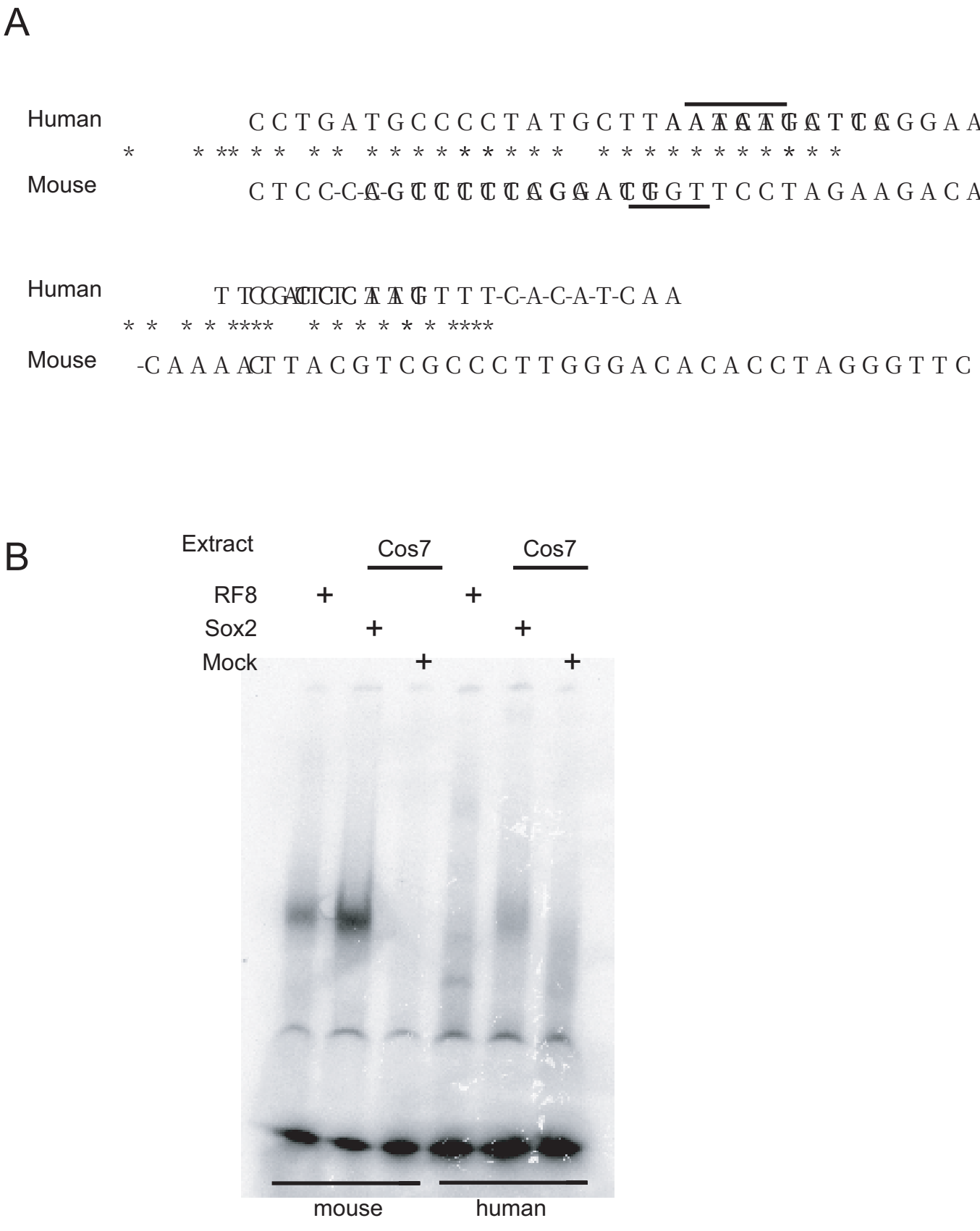
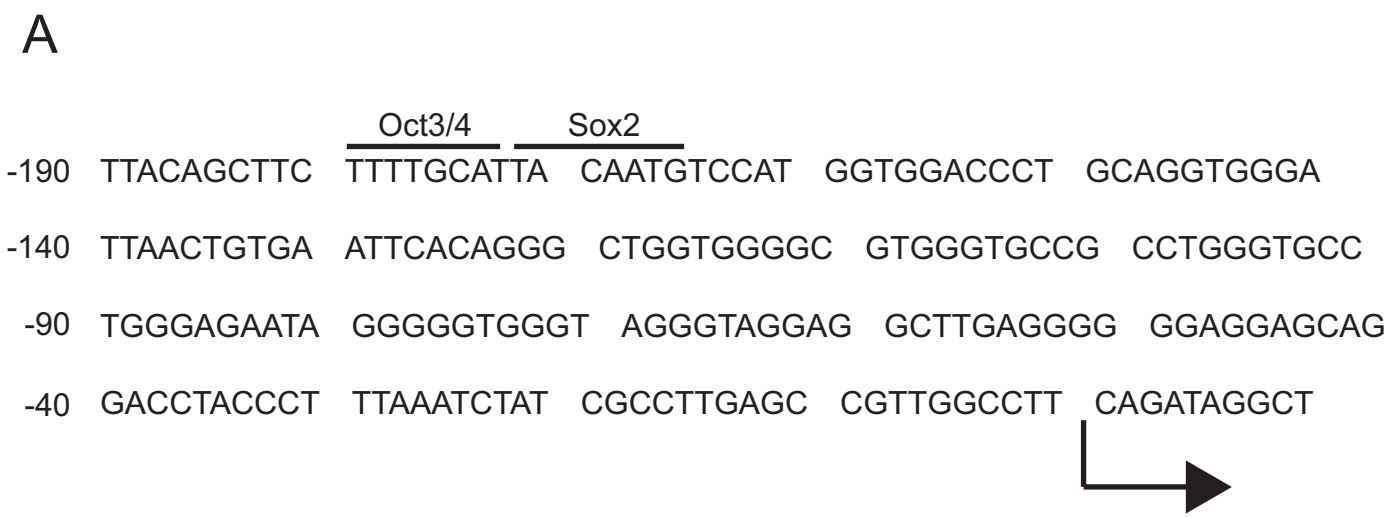


Figure 24. A putative Sox binding site in human Nanog enhancer. (A) Comparison of the mouse and human Nanog distal enhancer sequences. Underlines idicate putative Sox binding sites. (B) Gel mobility shift assay of the putative human Nanog enhancer. A ^{32}P-labeled oligonucleotides(mouse:CTCTCCCGGACGGttcctagaa-GACAAAGGCAAGCT,human:CTCTCCCGGACGGggcctagaaGACAAAGGCAAGCT) corresponding to the putative Sox binding sites were incubated with extracts from RF8 ES cells extract or Cos7 cells transfected with the Sox2 expression vector or a mock plasmid.



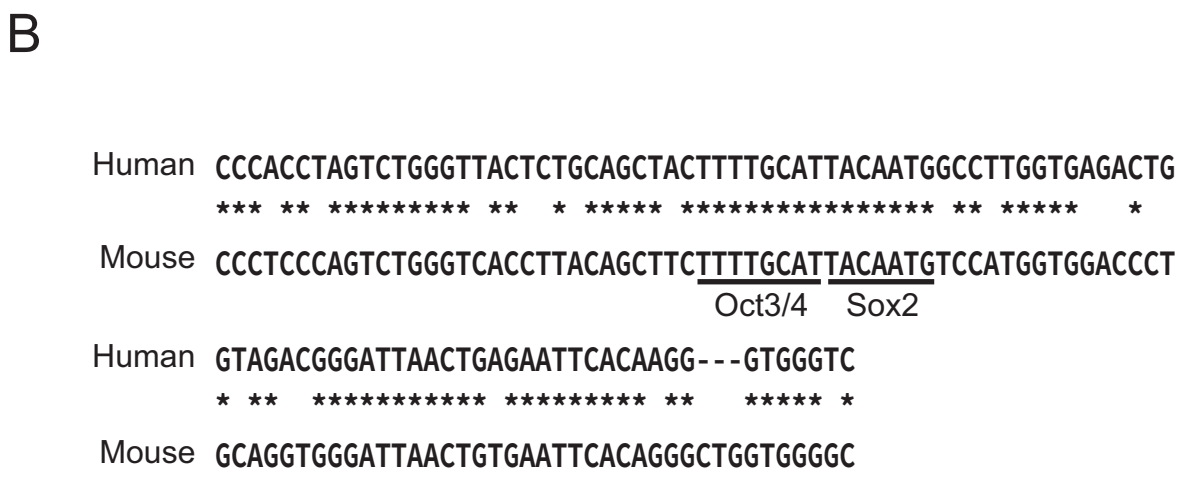


Figure 25. Oct3/4 and Sox2 binding sites in the proximal enhancer of the Nanog genes. (A) The sequence of the mouse Nanog gene upstream from the transcription initiation site (indicated by the arrow). Consensus binding sites for Oct3/4 and Sox2 are underlined. (B) Comparison between the mouse and human Nanog proximal enhancer sequences.

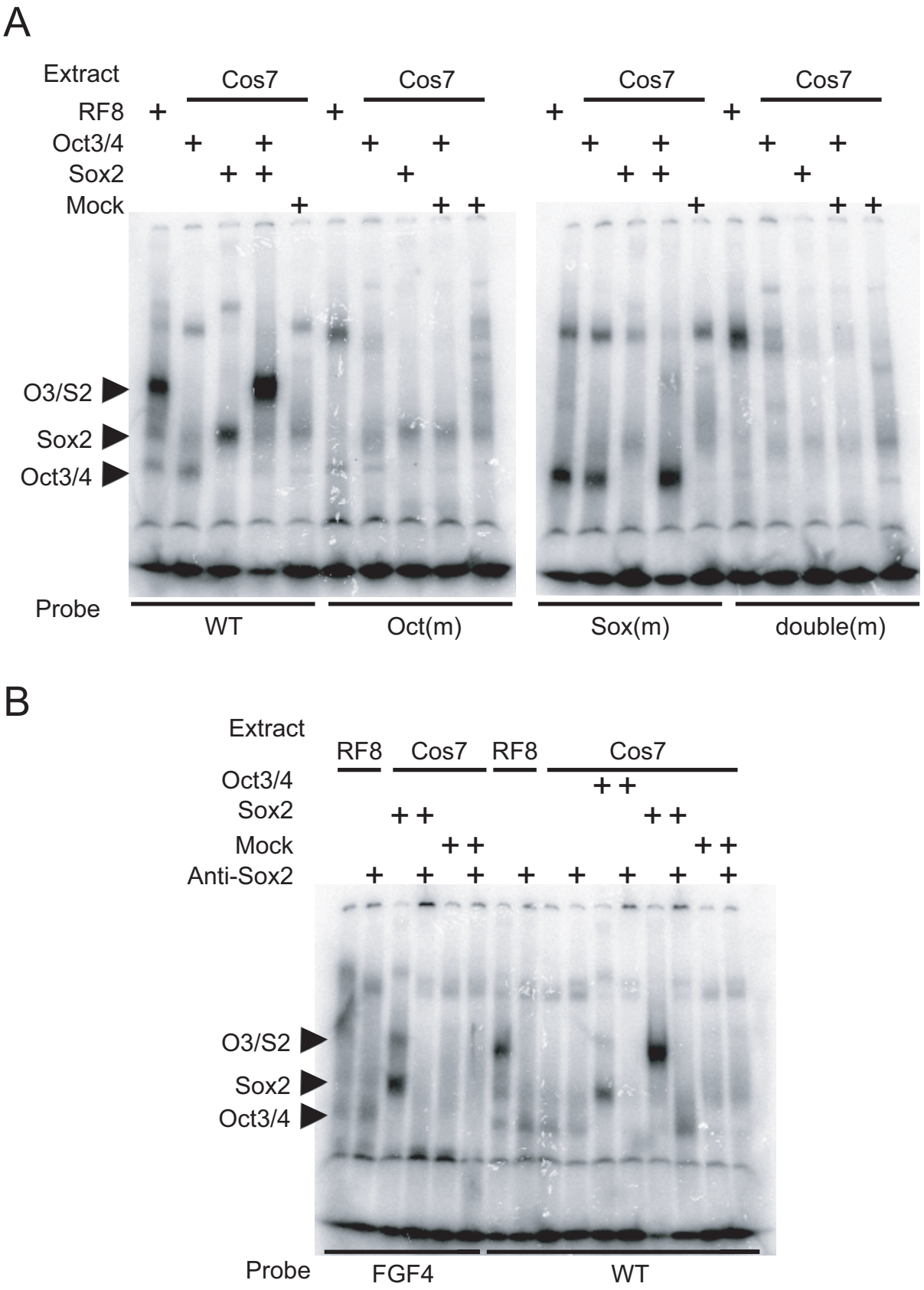


Figure 26. Gel mobility shift assay showing the binding of Oct3/4 and Sox2 to the Nanog distal enhancer. (A) Mutation assay. ^{32}P-labeled oligonucleotide(WT:CTCTCCCGGACGGttcctagaaGACAAAGGCAAGCT) corresponding to the Nanog proximal enhancer was incubated incubated with extracts from RF8 ES cells or Cos7 cells transfected with the Oct3/4, Sox2, or Oct3/4 and Sox2 expression vector or a mock plasmid. Oligonucleotides with mutations in the octamer-like sequence[O (m):CCTTACAGCTTCgggttacctacaatg-TCCATGGTGGAC], the Sox-binding site [S (m):CCTTACAGCTTC-ttttgcatacccaatTCCATGGTGGAC] and both sites [D (m):CCTTACAGCTTCgggttaccacccaatTCCATGGTGGAC] were also tested. (B) Super-shift assay. ^{32}P-labeled WT oligonucleotide was incubated with extracts from RF8 ES cells or Cos7 cells transfected with the Oct3/4, Sox2, or Oct3/4 and Sox2 expression vector or a mock plasmid. The reaction was performed with or without anti-Sox2 anti-serum. As a control, the FGF4 enhancer oligonucleotide was used.

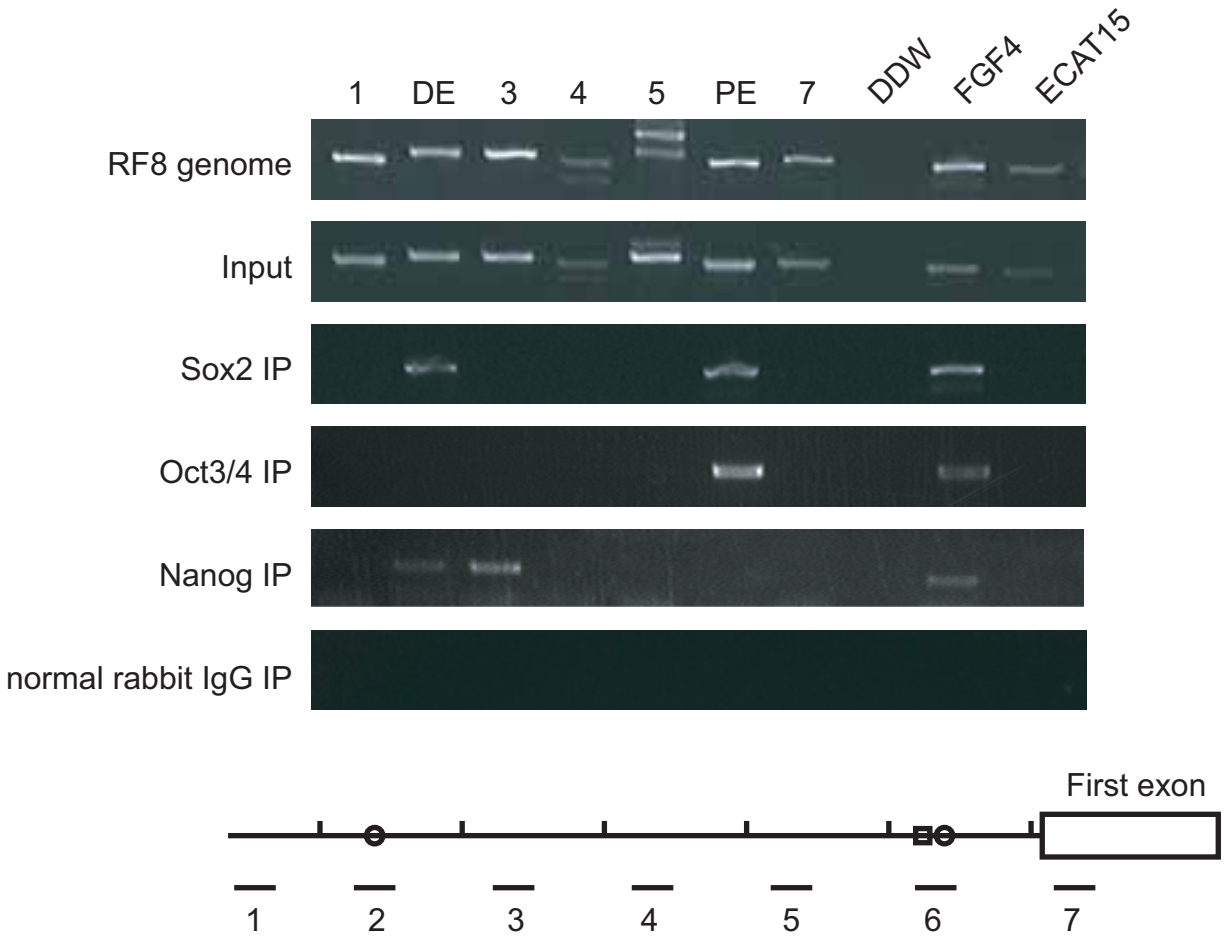


Figure 27. Chromatin immunoprecipitation alaysis of the Nanog regulatory regions. RF8 ES cells were fixed by formaldehyde, sonicated and lysed. Chromatin immunoprecipitation was performed with Sox2 anti-serum, anti-Oct3/4 antibody, Nanog anti-serum, or normal rabbit IgG. The final DNA extractions were aplied to PCR to amplify fragments (1〜7) shown at the bottom . As a positive control, FGF4 enhancer was amplified. As a negative control, an unrelated ECAT15 sequence was amplified. ○, □ shows Sox and Oct3/4 binding site, respectively.

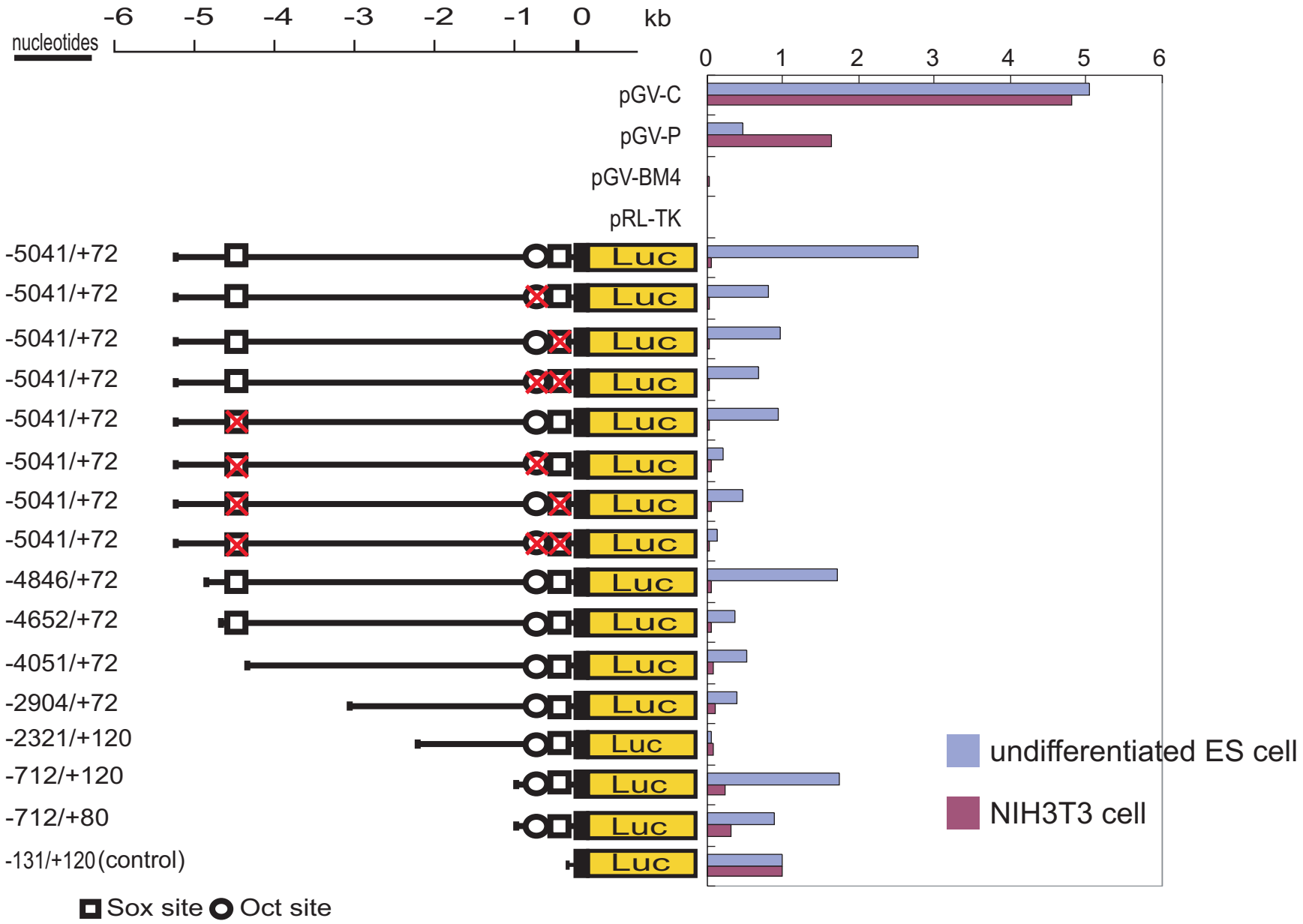


Figure 28. Effects of point mutations in the Sox or/and POU binding sites on enhancer activity. Reporter plasmids containing the Nanog promoter and enhancer sequence or mutated sequences were analyzed. These reporter constructs were introduced into undifferentiated ES cells (blue) and NIH3T3(purple). Shown are lucifease activities normalized with the valve of the shortest reporter(-131/+120).

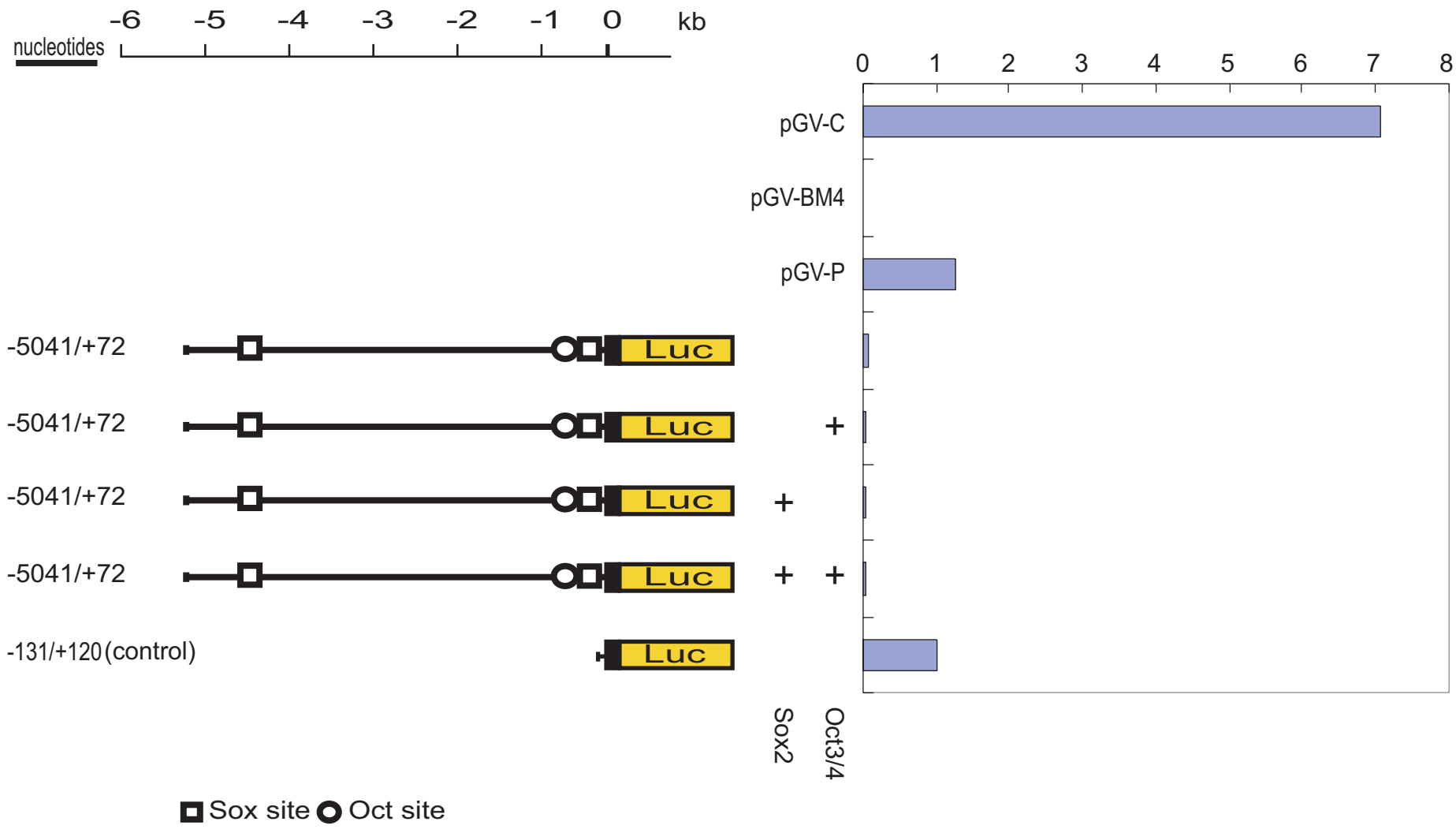


Figure 29. Failure of activation of the Nanog enhancer by Oct3/4 and Sox2 in differentiated cells. Luciferase reporter plasmids containing the FGF4 or Fbx15 enhancer were introduced into Cos7 cells together with expression vectors for Sox2 or/and Oct3/4 as indicated.

A

	♂	♀	Total
WT	14	18	32
Sox2+/-	15	10	25
Nanog+/-	11	12	23
double +/-	8	11	19

B

Strategy

Sox2+/- × Sox2+/-

↓ Tamocxifen

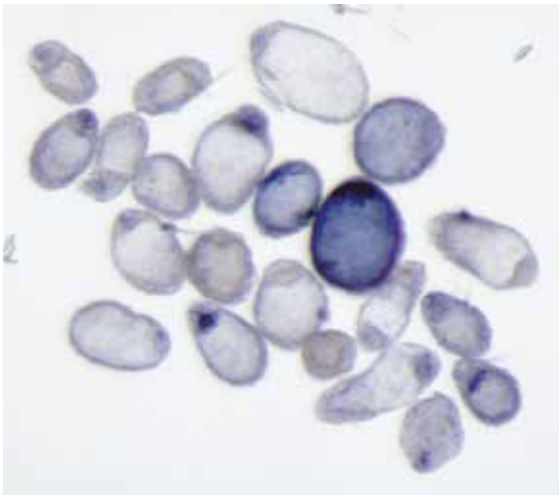
Delayed blastocyst

↓

In situ hybridization

↓

Genotyping PCR



Signal intensity

Genotype	++	+	±	−
WT	6	6	1	0
+/-	1	6	5	0
-/-	0	0	3	5

Figure 30. Interaction between Sox2, Sox15 and Nanog. (A) Analysis of Sox2$^{\beta geo}$ and Sox15$^{\beta geo}$ heterozyous intercross progeny. (B) Expression of Nanog in delayed blastocysts from Sox2$^{\beta geo}$ heterozygous intercrosses. Sox2$^{\beta geo}$ heterozygous mice were mated. Implantation was delayed by injection of Tamoxifen and Depo-provera to pregnant female mice at 2.5 dpc. After 4 days (6.5 dpc), embryos were collected and analyzed by situ hybridization with Nanog probe. Intensity of Nanog expression was scored ++ + ± − and summarized in the table according to the genotypes of the Sox2 allele.