CHARACTERIZATION OF NOBOX DNA BINDING SPECIFICITY AND ITS REGULATION OF GDF9 AND POU5F1 PROMOTERS*

Youngsok Choi and Aleksandar Rajkovic

From the Department of Obstetrics and Gynecology, Baylor College of Medicine, Houston, Texas, 77030

Running title: NOBOX transactivates mouse Gdf9 and Pou5f1 promoters

Address correspondence to: Aleksandar Rajkovic, Department of Obstetrics and Gynecology, Baylor College of Medicine 1709 Dryden St., Suite 1100, Houston, Texas 77030, Tel. 713 798-1049; Fax. 713 798-2744; E-mail: rajkovic@bcm.edu

Nobox (newborn ovary homeobox gene) deficiency disrupts early folliculogenesis and the expression of oocyte-specific genes in mice. Here, we identified several cis-acting sites, TAATTG, TAGTTG and TAATTA as a NOBOX DNA binding elements (NBEs) using a library of randomly generated oligonucleotides by cyclic amplification of sequence target assay and mutation analyses. We show that NOBOX preferentially binds to the NOBOX binding elements with high affinity. In addition, we found that promoter regions of mouse Pou5f1 and Gdf9 contain one (-426) and three NOBOX binding elements (-786, -967, and -1259), respectively. NOBOX binds to these putative NOBOX binding elements with high affinity and augmented transcriptional activity of luciferase reporter driven by mouse Pou5f1 and Gdf9 promoter containing the NOBOX binding elements. In chromatin immunoprecipitation assays, DNA sequences from Pou5f1 and Gdf9 promoters co-precipitated with anti-NOBOX antibody. These results suggest that NOBOX directly regulate the transcription of Pou5f1 and Gdf9 in oocyte during early folliculogenesis.

Early ovarian folliculogenesis begins with the breakdown of germ-cell clusters and the formation of primordial follicles, the reservoir of oocytes during reproductive lifespan of mammals. In mice, formation of primordial follicles begins soon after birth. Primordial follicles are the smallest ovarian follicle units and continuously recruited to grow into primary follicles. During early folliculogenesis, transcription of numerous oocyte-specific genes such as zona pellucida 1, 2 and 3 (Zp1, Zp2, and Zp3), growth differentiation factor 9 (Gdf9), and Pou5f1 (Oct4) commences (1-6). However, the genetic mechanisms of early folliculogenesis are poorly understood (7). We recently reported a novel homeobox containing germ-cell specific transcription factor NOBOX (6,8). Nobox transcripts are detectable as early as E15.5 in the embryonic ovaries and beyond (6). Nobox mRNA and protein are preferentially expressed in oocytes of germ cell cysts and in primordial and growing oocytes throughout different stages of folliculogenesis (6,8). Nobox deletion accelerates postnatal oocyte loss and results in abnormal primordial and primary oocytes (6). Newborn ovary histology in Nobox knockout and wild type female mice is grossly similar, but differs significantly at the molecular level. Numerous genes preferentially expressed in the oocytes, including Gdf9 and Oct4 are down-regulated in oocytes that lack Nobox. However, it is unknown whether NOBOX directly or indirectly regulates transcription of these oocyte-specific genes during development of the ovarian follicle.

In the present study, we identify DNA sequences that NOBOX binds and show that NOBOX can transactivate luciferase reporters via the NOBOX DNA binding element (NBE). In addition, we describe identification of NBEs in promoters of mouse Oct4 and Gdf9. These observations suggest that NOBOX directly regulates the expression of oocyte-specific genes, Oct4 and Gdf9, during early folliculogenesis in the mouse ovary.

EXPERIMENTAL PROCEDURES

Expression and purification of GST-NXHD - To create the pET41b-Nobox homeodomain protein (GST-NXHD) and bacterial expression construct,
the insert was amplified by PCR using primers containing EcoRI and HindIII sites, and cloned into pET41b-EcoRI/HindIII (Novagen, Madison, WI). We verified the sequences of the resulting GST-NXHD fusion constructs to ensure that no mutations had been introduced during cloning. For creating GST-NXHD (N185Q) mutation proteins, the GST-NXHD construct vector was mutated using QuikChange Muti Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA) and its mutated sequences were verified. We transformed BL21-pLysS *E. coli* (Stratagene, La Jolla, CA) with GST-NXHD and GST-NXHD(N185Q) constructs and expressed them by inoculating Luria-Bertani media containing 20 μg ml⁻¹ kanamycin with overnight culture (1:20 dilution) at 37 °C to an OD₆₀₀ of 0.5. Protein expression was induced at 30 °C with 2 mM of IPTG to OD₆₀₀ of 1.0. We stored cell pellets at -80 °C, thawed and lysed them with BugBuster lysis buffer (Novagen, Madison, WI). We purified the GST-NXHD and GST-NXHD (N185Q) fusion proteins with GST bind resin (Novagen, Madison, WI), dialyzed them three times in 1 liter of PBS and quantified them. Purified proteins were stored at -80 °C until use.

**Tissue extract** - Ovaries from wild type and Nobox null newborn mice were collected and homogenized in 100 µl of lysis buffer containing a final concentration of 20 mM HEPES (pH 7.9), 25% glycerol, 420 mM NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA. Homogenized ovary tissues were incubated on ice for 30 min and then centrifuged at 4 °C for 30 min. After centrifuging, the supernatant was transferred into new tubes. Samples were stored at -80 °C until use.

**Plasmids** - The pGL3-promoter vector with SV40 promoter (Promega, Madison, WI) was used for constructing luciferase reporter vectors carrying 3xNBE and 3xmNBE. The pGL3-basic vector lacked SV40 promoter (Promega, Madison, WI) was used to subclone amplified Oct4 and Gdf9 promoters. Oct4 or Gdf9 NBEs were mutated using QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). Overexpression vector carrying the mouse Nobox were constructed by cloning the full-length of Nobox cDNA into pcDNA3 vector (Invitrogen, Carlsbad, CA) and sequenced. The pBIND vector (Promega, Madison, WI) was used to construct plasmids carrying the full-length or truncated of Nobox cDNAs. Full-length and partial NOBOX cDNAs were fused after GAL4DBD in frame and verified by sequencing. The expression of the constructs in mammalian cells was confirmed by Western blot analysis using anti-GAL4DBD monoclonal antibody (SantaCruz Biotechnology, Inc., Santa Cruz, CA).

**Electrophoretic Mobility Shift Assay (EMSA)** - EMSAs were performed in 20 µl reaction mixtures at room temperature at a final concentration of 10 mM Tris (pH 7.5), 50 mM NaCl, 1.5 mM MgCl₂, 2.5 mM dithiothreitol, 5% glycerol, 5 µg/ml poly(dI)-poly(dC), and 250 µg/ml bovine serum albumin. EMSA probes were prepared by end-filling annealed primers with [α-³²P] dCTP and Klenow polymerase (Invitrogen, Carlsbad, CA). Binding reactions were conducted by incubating ³²P-labeled probes (250,000 cpm/reaction) with 50 ng of purified proteins or 1 µg of ovarian extract in the absence and presence of polyclonal anti-GST (Amergham Biosciences, Piscataway, NJ) or polyclonal anti-NOBOX antibody (6). For competition assay, purified proteins were incubated at room temperature with cold competitors for 15 min before addition of the ³²P-labeled probe.

**Cyclic Amplification of Sequence Target (CAST) Analysis** - The CAST assay was carried out essentially as previously described with some modifications (9). Purified GST-NXHD proteins and GST-bind resin (Novagen, Madison, WI) were used for CAST assay. Binding reactions for CAST were performed in DNA binding buffer (10 mM Tris (pH 7.5), 50 mM NaCl, 7.5 mM MgCl₂, 1 mM EDTA, 5% glycerol, 5% sucrose, 0.1% Nonidet P-40, and 5 mg/ml bovine serum albumin). 0.4 pmoles of short double-stranded DNA containing a 15-bp random sequence flanked by 20-bp fixed sequences, CAST oligonucleotides (Table 1), was incubated with...
100 ng of purified GST-NXHD protein for 20 min at room temperature. Unbound DNAs were washed away with binding reaction buffer. Bound DNAs were amplified by PCR for the next round of CAST. A total of five cycles of CAST were performed. Final PCR products were cloned into the pGEM-T easy vector (Promega, Madison, WI). The inserted DNAs were sequenced and scored to get the consensus sequences.

Cell culture and Reporter Assays - Human embryonic kidney cells (HEK293) were grown in Dulbecco’s modified Eagle’s medium with 10% fetal calf serum. For transient transfection, FuGENE6 (Roche Applied Science) was used according to the manufacturer’s instructions. Following the transfection, the cells were incubated for 48 hours before harvest. For each transfection, 200 ng of reporter construct, 200 ng of the indicated expression plasmid, and 20 ng of pRT-TK normalization plasmid were used per single well of a 12-well plate. Dual luciferase assays were carried out with total cell extracts as recommended by Promega. All transfection experiments were performed in triplicate, and results were normalized to the expression of the Renilla luciferase.

Chromatin immunoprecipitation (ChIP) - Mouse newborn ovaries were collected and fixed directly by adding 27 µl of 37% formaldehyde to 1ml of Hank’s balanced salt solution (Invitrogen, Carlsbad, CA). The fixed ovaries were washed with cold 1x PBS two times, homogenized into 4 ml of homogenization buffer (50mM sodium phosphate buffer, pH 7.5, 10% glycerol, 10mM beta-mercaptoethanol, 50mM PMSF, and protein inhibitors) and prepared for immunoprecipitation using the modified protocol of chromatin immunoprecipitation (ChIP) assay kit (Upstate Biotechnology Inc., Lake Placid, NY). The samples were pre-cleared with salmon sperm DNA/protein A agarose (Upstate Biotechnology Inc., Lake Placid, NY) and incubated with either 10µl of Goat anti-Nobox antibody or IgG at 4°C overnight. Chromatin samples immunoprecipitated by salmon sperm DNA/protein beads were washed two times with 1ml of low salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, pH 8.1, 150mM NaCl), once with 1ml of high salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, pH 8.1, 500mM NaCl), once with 1ml of LiCl immune complex wash buffer (0.25 M LiCl, IGEPAL-CA630, 1% sodium deoxycholate, 1mM EDTA, 10mM Tris-HCl, pH 8.1), and two times with 1ml of TE buffer (10mM Tris-HCl, pH8.0, 1mM EDTA). Immune complexes were eluted from the antibody by adding 500 µl of elution buffer (0.1% SDS, 0.1M NaHCO₃). The complex-DNA crosslinks were reversed by adding 20 µl of 5M NaCl and heating at 65°C for 4 hours followed by adding 10 µl of 0.5M EDTA, 20 µl of 1M Tris-HCl, pH 6.5, and 2 µl of Proteinase K (10mg/ml) for one hour at 45°C. DNA was recovered by phenol/chloroform extraction and ethanol precipitation. Ethanol-precipitated DNA pellets were re-dissolved in 50 µl of TE buffer. The supernatant of an immunoprecipitation reaction done in the absence of anti-NOBOX antibody was purified and used as a control. PCR analysis was carried out using primers corresponding to promoter regions of the Oct4 and Gdf9 genes. The sequences of the PCR primers are shown in Table 1. After 33 (all primer sets) of amplification, PCR products were run on 2% agarose gels and analyzed.

RESULTS

We determined a consensus DNA binding sequence for NOBOX using the cyclic amplification of sequence target (CAST) assay as previously described (10,11). We used a predicted homeodomain portion of NOBOX (a.a. 134-214) fused to glutathione-S-transferase (GST-NXHD). DNA sequences that bound GST-NXHD went through five cycles of CAST. The 38 selected sequences for GST-NXHD were aligned (Fig. 1A) and scored (Fig. 1B). The most frequently observed sequences are shown in Fig. 1B based on the percentage occurrence of each base at each position from 58% to 97%. CAST assay revealed a
consensus sequence, TAATTG (Fig. 1B), as the NBE.

We next confirmed the interaction between NBE and GST-NXHD using EMSA. Using a competitive binding assay (Fig. 2A), GST-NXHD protein bound to the labeled probe (Fig. 2A, lane 2). The bound complex was competed out by 10-fold and 100-fold molar excess of unlabeled cold competitors containing TAATTG (Fig. 2A, lane 3 and 4 respectively). To further confirm the identity of these shifted bands, anti-GST antibody was added to the binding reaction in the absence or presence of GST-NXHD or GST protein. As shown in Fig. 2B, the anti-GST antibody further shifted the DNA-protein complex (Fig. 2B, lane 4).

A number of homeoproteins including Msx-1, Nkx2.1, Nkx3.1 and Nkx3.2 contain a conserved asparagine residue (N) in the homeodomain (Fig. 2C). This conserved asparagine residue is critical for interaction of DNA base pairs and various homeodomains (12,13). A missense mutation of this asparagine (N) to glutamine (Q) has been shown to reduce DNA binding activity in many homeoproteins (9,14-17). NOBOX also contains this conserved N at residue 185 (Fig. 2C). A point mutation of NOBOX (N185Q) was generated and used to investigate whether GST-NXHD (N185Q) can bind to the NBE sequences in vitro. Purified GST-NXHD (N185Q) was incubated with 32P-labeled NBE probe and binding was analyzed by EMSA (Fig. 2D). As shown in Fig. 2D, GST-NXHD (N185Q) failed to interact with the 32P-labeled NBE probe, but GST-NXHD formed complexes with the NBE probe. These results suggest that the homeodomain is critical for DNA binding activity.

The NBE is identical to the core binding sequences of homeoprotein, MSX1 (18), but slightly different from the binding sites of other homeoproteins such as CAAGTG for TITF1 (19), TAAGTA for NKX3-1 (20) and TAAGTG for BAPX1 (9). We examined whether NOBOX can also bind to these consensus sequences. Increasing amounts of purified GST-NXHD protein incubated with 32P-labeled probes containing these related sequences, TAATTG, CAAGTG, TAAGTA, or TAAGTG were analyzed using EMSA. At high concentration (100ng), GST-NXHD protein bound to these related sequences with low affinity compared to the NBE (Fig. 3A). In contrast, NOBOX bound to the NBE with high affinity relative to other related sequences at low concentration (Fig. 3A). These results suggest that NOBOX preferentially binds to the NBE sequence. To examine the effect of variation of each base pair of the NBE sequence, we incubated GST-NXHD with a labeled NBE probe (TAATTG) and added a 100-fold molar excess of unlabeled oligonucleotides which are different from NBE sequence by a single base pair substitution in each position. We found that the DNA-protein complex was competed out by two NBE-related sequences, TAGTTG and TAATTA, containing G in the 3rd position or A in the 6th position, respectively (Fig. 3B). In contrast, the binding specificity for the 1st, 2nd, 4th and 5th position is invariant. The variation of 3rd position was not expected, however, the variation of 6th position was consistent with the result of CAST assay, because the occurrence of the A at the 6th position is 37 % (Fig. 1B). This data suggests that TAGTTG and TAATTA are other consensus sequences for NOBOX DNA binding (Fig. 3C).

To confirm that endogeneous NOBOX from newborn ovaries binds to the NBE (TAATTG), the NBE was 32P-labeled and incubated with ovary extracts from wild type or Nobox null mice. To identify these shifted bands, anti-NOBOX antibody or pre-immune serum was added to the binding reaction. As shown in Fig. 4, the NBE-protein complex formed using 32P-labeled oligonucleotide containing TAATTG incubated with ovarian extract from wild type newborn mice, but not from Nobox null mice. Anti-NOBOX antibody shifted the DNA-protein complex (lane 3 in Fig. 4). We obtained similar results using NBE-related sequence, TAATTA.

To examine whether NOBOX can regulate the expression of NBE-driven reporter gene in vivo, we constructed two luciferase reporter vectors containing either the NBE (TAATTG), 3xNBE-pGL3, or mutant NBE (TAGGCG), 3xmNBE-pGL3 upstream of the luciferase reporter gene. These reporter vectors were co-transfected with either empty vector (pCDNA3.1) or NOBOX expression vector (pCDNA3.1-NX).
into HEK293 cells. The relative luciferase activity of reporter constructs containing three copies of TAATTG was increased by 4-fold with NOBOX overexpression (Fig. 5A). However, the level of transcriptional activation of reporter constructs containing three copies of mutated NBE (TAGGCG) was not affected by NOBOX overexpression. These data suggest that NOBOX can transactivate reporter genes through the NBE.

To investigate the contribution of the various protein regions of NOBOX to transactivation, we generated a series of truncated NOBOX proteins fused to GAL4 DNA binding domain (GAL4DBD) (Fig. 5B). These constructs were co-transfected with 5xUAS-TATA-luciferase vector into HEK293 cells. The relative fold increase of luciferase activity is shown in Fig. 5D. GAL4-CT containing the C-terminus of NOBOX significantly increased the luciferase activity. In contrast, the luciferase activity was not affected by GAL4-NT and GAL4-HD containing the N-terminus and the homeodomain of NOBOX, respectively. Therefore, the C-terminus possesses the activation domain of NOBOX.

We showed previously that Nobox deficiency disrupts the expression of several oocyte-specific genes, including Gdf9 and Oct4. Gdf9 is expressed early in the oocyte (3,6,21) and is essential for the growth and differentiation of the surrounding granulosa cells (22-24), antral follicle growth and ovulation (4,25-27). Oct4 is expressed early in the development of the germ cell, but its transcription ceases after embryonic day 14.5 (E14.5) and reappears in the oocyte at birth (5,6). Thus, we examined the promoter region of mouse Gdf9 and Oct4 for the presence of the NBEs (TAATTG, TAGTTG or TAATTA). As summarized in Table 2, the 2.0-kb mouse Gdf9 promoter contains three putative NBEs (position -1259, -967, and -786) and the 2.0-kb mouse Oct4 promoter contains one putative NBE (position -426) relative to the translational start site (+1). Interestingly, these putative NBEs contain the TAATTG sequence, but not the variant NBEs. Using a competitive binding assay (Fig. 6A and 6C) and super-shift assay (Fig. 6B and 6D), we compared GST-NXHD to these putative NBEs containing TAATTG. We found that the DNA-protein complexes were competed out by 10-fold or 100-fold molar excess of unlabeled putative NBE sequences including Oct4 NBE (-462) and Gdf9 NBEs (-1259, -967 and -786) and were super-shifted by antibody against GST.

To determine the effects of NOBOX on transactivation of mouse Oct4 and Gdf9, 293HEK cells were transiently transfected with a construct containing wild type or mutated NBE sequences. The 2.0-kb Gdf9 promoter and the 0.5-kb Oct4 promoter were cloned into luciferase reporter vector. The NBE (-462) sequences in Oct4 promoter were mutated into GGATCC (BamH I) and three NBEs (-1259, -967, and -786) in Gdf9 promoter were mutated into GGATCC (BamH I), GCTAGC (Nhe I), and GGTACC (Kpn I), respectively. These wild type or mutated constructs were transiently transfected with Nobox expression vector or empty vector as a control. A renilla luciferase expression vector was used as an internal control for normalization of transfection efficiency. The relative luciferase activities of the reporter plasmids containing wild type NBE or mutated NBE are shown in Fig. 7B. The NBE (-462) relative to the translational start site (+1). Thus, we examined the promoter regions of mouse Oct4 and Gdf9, 293HEK cells were transiently transfected with a construct containing wild type or mutated NBE sequences. The 2.0-kb Gdf9 promoter and the 0.5-kb Oct4 promoter were cloned into luciferase reporter vector. The NBE (-462) sequences in Oct4 promoter were mutated into GGATCC (BamH I) and three NBEs (-1259, -967, and -786) in Gdf9 promoter were mutated into GGATCC (BamH I), GCTAGC (Nhe I), and GGTACC (Kpn I), respectively. These wild type or mutated constructs were transiently transfected with Nobox expression vector or empty vector as a control. A renilla luciferase expression vector was used as an internal control for normalization of transfection efficiency. The relative luciferase activities of the reporter plasmids containing wild type NBE or mutated NBE are shown in Fig. 7A and 7B. Nobox expression increased the activity of Oct4 and Gdf9 luciferase construct by 3.7-fold (Fig. 7B) and 6.2-fold (Fig. 7B), respectively. Mutation of the NBE sequences abolished Nobox-dependent transactivation of Gdf9-luciferase construct (Fig. 7B), but NOBOX effect on mutated NBE containing Oct4-luciferase construct was partially decreased (Fig. 7A). These experiments suggest that Oct4 and Gdf9 are candidate genes for direct regulation by NOBOX.

To verify whether NOBOX directly binds to mouse Oct4 and Gdf9 promoters, we performed chromatin immunoprecipitation assays with protein extracts from newborn mouse ovaries or liver. Liver tissue does not express NOBOX and serves as a negative control. Extracts were cross-linked, sonicated, and immunoprecipitated with the anti-NOBOX antibody. PCR was used to amplify promoter regions containing NBE elements in the Oct4 and Gdf9 genes (P1) and regions that lacked NBE elements (C1) (Fig. 7C). NBE containing genomic sequences from the Oct4 and Gdf9 promoters co-precipitated with the anti-NOBOX antibody in the newborn ovary extracts.
In contrast, co-immunoprecipitation was not detected in the liver extracts, which do not express *Nobox* or *Oct4* (Fig. 7D).

**DISCUSSION**

Our previous study indicated that numerous oocyte-specific genes, including *Gdf9* and *Oct4*, were downregulated in ovaries that lack NOBOX (6). However, it is unknown whether NOBOX directly or indirectly regulates transcription of these oocyte-specific genes during development of the ovarian follicle. In the present study, we have identified a consensus sequence, TAATTG, as NOBOX DNA binding elements using purified proteins in a random DNA selection assay (Fig. 1). However, the sequences generated by a single base pair substitution in each position revealed that NOBOX bound two other NBE related sequences, TAGTTG and TAATTA (Fig. 3B). The variation of the 6th base was consistent with the result of CAST, but that of the 3rd base was unexpected. The TAATTG consensus site is shared with homeoprotein MSX1 (18), but TAATTA and TAGTTG appear unique to NOBOX. The NBE sequences are similar to other homeoprotein DNA binding elements such as CAAGTG for TITF1 (19), TAAGTA for NKX3-1 (20) and TAAGTG for BAPX1 (9).

NOBOX interacting partners, tissue-specific expression, and other flanking nucleotide motifs are likely required additional elements for oocyte-specific effects of NOBOX. *Mxl*, *Titf1*, *Nkx3-1* and *Bapx1* are not expressed in the newborn ovaries (data not shown), which accounts for the dominant shift of TAATTG incubated with the newborn ovary extracts to be due to NOBOX (Fig. 4). Newborn *Nobox* null ovary extracts lack the dominant shift band, in concordance with NOBOX being the main binder of the TAATTG element in wild type newborn ovaries.

Homeoproteins contain a highly conserved asparagines residue at position 51 in the homeodomain (28), and this conserved asparagine residue appears critical for the interaction of DNA base pairs and many homeodomains (12,13). NOBOX contains the conserved asparagine residue at position 51 in the homeodomain (position 185 in the protein) (Fig. 2C) and mutating this asparagines to glutamine abolished NOBOX homeodomain binding to NBEs. These results suggest that NOBOX binds DNA in a manner that is structurally similar to that of other homeoproteins.

Unlike many other homeoproteins, NOBOX protein C-terminus is proline-rich and contains a putative SH3 and WW domains (29). *Esx1*, a homeoprotein preferentially expressed in the placenta (30), also contains a proline rich region that is important in localizing ESX1 to the nucleus and also contains a proline rich SH3 domain that binds *c-abl* SH3 domain *in vitro* (31). Homeoproteins can execute their biological activity not only through protein-DNA interactions, but also protein-protein interactions (18,32-37). This suggests that NOBOX proline rich carboxyl terminus binds other, perhaps oocyte-specific proteins that are critical in NOBOX transcriptional activity.

We have shown that mouse *Gdf9* and *Oct4* promoters contain NOBOX DNA binding sites (Table 2). OCT4 and GDF9 play a crucial role for the maintenance of germ cells and the formation of follicle, yet little is known about the direct regulation of expression of *Oct4* and *Gdf9* in the oocyte. GDF9 is an oocyte-derived growth factor required for somatic cell function. The lack of *Gdf9* causes infertility secondary to an early block in follicular development at the primary one-layer follicle stage (2). *Oct4* null mutation results in peri-implantation lethality prior to the egg cylinder stage (38). However conditionally targeted *Oct4*<sup>fl ox</sup> mice whereby Cre is under the control of TNAP (Tissue non-specific alkaline phosphatase) deplete oocytes by 6 weeks of age (39). These results suggest that GDF9 and OCT4 are required for follicle formation.

*Oct4* plays a critical role in the basic biology of embryonic stem cells. Embryonic stem cells and germline share several important transcriptional regulators. The *Oct4* transcript is differentially regulated by a TATA-less minimal promoter (40) and two cis-elements, distal enhancer (DE) and proximal enhancer (PE) (41).
The DE is active in embryonic stem cells and the PE is active in the epiblast (41). Promoter elements in the Oct4 responsible for re-expression of Oct4 in the post-natal female germline are unknown.

In this study, we have shown that the promoters of mouse Oct4 (0.5kb) and Gdf9 (1.5kb) contain one (-426) and three NBEs (-786, -967 and -1259) from the ATG site (Table 2). Interestingly, these NBEs contain TAATTG sequences and they are not located in the PE or DE elements of the Oct4 promoter and the proximal region of the Gdf9 promoter. The NBEs are not conserved among other species, but several putative NBEs can be found in the proximal promoter regions of human, chimpanzee, pig, cow and sheep Oct4 and Gdf9 genes (data not shown). Our results suggest that putative NBEs are candidate cis-acting elements for NOBOX activity during folliculogenesis. We have shown that NOBOX can bind putative mouse NBEs with high affinity (Fig. 5) and that NOBOX can transactivate luciferase driven by Gdf9 (1.5kb) and Oct4 (0.5kb) promoters (Fig. 6). Because NOBOX is oocyte-specific, transactivation in oocytes may be even greater. Currently oocyte cell lines do not exist. Moreover, we have shown that antibodies against NOBOX can immunoprecipitate NBEs from the Oct4 and Gdf9 promoters (Fig. 6D). Nobox expression precedes closely the expression of Gdf9 and Oct4. All of these results implicate that NOBOX directly regulates the expression of Oct4 and Gdf9 in the oocyte during early folliculogenesis.

Oocyte-specific pathways are critical for oogenesis. Transcription of numerous genes preferentially expressed in the oocyte, and critical for ovarian folliculogenesis and early embryogenesis, commences at birth when primordial ovarian follicles form. Transcriptional networks unique to oocytes are key to understanding and generating germ cells. NOBOX is one of the few germ cell specific transcription factors that is essential in oogenesis and disrupts transcription of numerous oocyte-specific transcripts (6). Our findings indicate that transcriptional activity of NOBOX requires direct DNA binding through specific consensus sequences, the NBEs, and that NOBOX directly regulates the expression of oocyte-specific genes, Gdf9 and Oct4, through the NBE during folliculogenesis. These results also spotlight Oct4 as having a possible role in oogenesis, and future experiments are necessary to assess the function of Oct4 in early folliculogenesis and oogenesis. NOBOX transcriptional partners are unknown, and identification of NOBOX partners will be important to further investigate NOBOX specificity for genes preferentially expressed in the oocytes as well as in reprogramming somatic cells to germ cells.

REFERENCES

1. Dean, J. (2002) J Reprod Immunol 53(1-2), 171-180
2. Dong, J., Albertini, D. F., Nishimori, K., Kumar, T. R., Lu, N., and Matzuk, M. M. (1996) Nature 383(6600), 531-535
3. Dube, J. L., Wang, P., Elvin, J., Lyons, K. M., Celeste, A. J., and Matzuk, M. M. (1998) Mol Endocrinol 12(12), 1809-1817
4. Elvin, J. A., Yan, C., and Matzuk, M. M. (2000) Mol Cell Endocrinol 159(1-2), 1-5
5. Pesce, M., Wang, X., Wolgemuth, D. J., and Scholer, H. (1998) Mech Dev 71(1-2), 89-98
6. Rajkovic, A., Pangas, S. A., Ballow, D., Suzumori, N., and Matzuk, M. M. (2004) Science 305(5687), 1157-1159
7. Epifano, O., and Dean, J. (2002) Trends Endocrinol Metab 13(4), 169-173
8. Suzumori, N., Yan, C., Matzuk, M. M., and Rajkovic, A. (2002) Mech Dev 111(1-2), 137-141
9. Kim, D. W., Kempf, H., Chen, R. E., and Lassar, A. B. (2003) J Biol Chem 278(30), 27532-27539
10. Blackwell, T. K., and Weintraub, H. (1990) *Science* **250**(4984), 1104-1110
11. Chen, C. Y., and Schwartz, R. J. (1995) *J Biol Chem* **270**(26), 15628-15633
12. Fraenkel, E., and Pabo, C. O. (1998) *Nat Struct Biol* **5**(8), 692-697
13. Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Kornberg, T. B., and Pabo, C. O. (1990) *Cell* **63**(3), 579-590
14. Botfield, M. C., Jancso, A., and Weiss, M. A. (1994) *Biochemistry* **33**(20), 6177-6185
15. Choi, C. Y., Kim, Y. H., Kwon, H. J., and Kim, Y. (1995) *J Biol Chem* **270**(26), 15628-15633
16. Fraenkel, E., and Pabo, C. O. (1998) *Nat Struct Biol* **5**(8), 692-697
17. Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Kornberg, T. B., and Pabo, C. O. (1990) *Cell* **63**(3), 579-590
18. Botfield, M. C., Jancso, A., and Weiss, M. A. (1994) *Biochemistry* **33**(20), 6177-6185
19. Choi, C. Y., Kim, Y. H., Kwon, H. J., and Kim, Y. (1995) *J Biol Chem* **270**(26), 15628-15633
20. Steadman, D. J., Giuffrida, D., and Gelmann, E. P. (2000) *Nucleic Acids Res* **28**(12), 2389-2395
21. McGrath, S. A., Esquela, A. F., and Lee, S. J. (1995) *Mol Endocrinol* **9**(1), 131-136
22. Elvin, J. A., Clark, A. T., Wang, P., Wolfman, N. M., and Matzuk, M. M. (1999) *Mol Endocrinol* **13**(6), 1035-1048
23. Elvin, J. A., and Matzuk, M. M. (1998) *Rev Reprod* **3**(3), 183-195
24. Elvin, J. A., Yan, C., Wang, P., Nishimori, K., and Matzuk, M. M. (1999) *Mol Endocrinol* **13**(6), 1018-1034
25. Matzuk, M. M. (2000) *Mol Cell Endocrinol* **163**(1-2), 61-66
26. Vitt, U. A., Hayashi, M., Klein, C., and Hsueh, A. J. (2000) *Biol Reprod* **62**(2), 370-377
27. Vitt, U. A., and Hsueh, A. J. (2001) *Mol Cell Endocrinol* **183**(1-2), 171-177
28. Laughton, A. (1991) *Biochemistry* **30**(48), 11357-11367
29. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) *Faseb J* **14**(2), 231-241
30. Li, Y., Lemaire, P., and Behringer, R. R. (1997) *Dev Biol* **188**(1), 85-95
31. Yan, Y. T., Stein, S. M., Ding, J., Shen, M. M., and Abate-Shen, C. (2000) *Mol Cell Biol* **20**(2), 661-671
32. Newberry, E. P., Latifi, T., and Towler, D. A. (1999) *Biochemistry* **38**(33), 10678-10690
33. Nguyen Ba-Charvet, K. T., von Boxborg, Y., Guazzi, S., Boncinelli, E., and Godement, P. (1998) *Development* **125**(21), 4273-4282
34. Pomerantz, J. L., Kristie, T. M., and Sharp, P. A. (1992) *Genes Dev* **6**(11), 2047-2057
35. Rauskolb, C., Peifer, M., and Wieschaus, E. (1993) *Cell* **74**(6), 1101-1112
36. Spaete, R. R., and Mocarski, E. S. (1985) *J Virol* **56**(1), 135-143
37. Stern, S., Tanaka, M., and Herr, W. (1989) *Nature* **341**(6243), 624-630
38. Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998) *Cell* **95**(3), 379-391
39. Kehler, J., Tolkunova, E., Koschorz, B., Pesce, M., Gentile, L., Boiani, M., Loneli, H., Nagy, A., McLaughlin, K. J., Scholer, H. R., and Tomilin, A. (2004) *EMBO Rep* **5**(11), 1078-1083
40. Okazawa, H., Okamoto, K., Ishino, F., Ishino-Kaneko, T., Takeda, S., Toyoda, Y., Muramatsu, M., and Hamada, H. (1991) *Embo J* **10**(10), 2997-3005
41. Yeom, Y. I., Fuhrmann, G., Ovitt, C. E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., and Scholer, H. R. (1996) *Development* **122**(3), 881-894

**FOOTNOTES**
This work was supported by grants from the National Institutes of Health (HD44858) and a March of Dimes Basil O’Connor Award (5-FY02-266) to A. Rajkovic.

The abbreviations used are: NOBOX, newborn ovary homeobox protein; POU5f1, POU domain class 5 transcription factor 1; GDF9, growth differentiation factor 9; ZP, zona pellucida protein; POU, Pit-Oct-Unc family; EMSA, electrophoretic mobility shift assay; CAST, cyclic amplification of sequence target analysis; GST, glutathione-S-transferase; NX, Nobox; NXHD, Nobox homeodomain; NXNT, Nobox N-terminus; NXCT, Nobox C-terminus; HEK, human embryonic kidney; ChIP, Chromatin immunoprecipitation; NBE, NOBOX DNA binding element; mNBE, mutant NOBOX DNA binding element; N, arginine; Q, glutamine; GAL4DBD, GAL4 DNA binding domain; UAS, upstream activation sequence; PGC, primordial germ cell; DE, distal enhancer; PE, proximal enhancer; TNAP, Tissue nonspecific alkaline phosphatase.

FIGURE LEGENDS

Fig. 1. Determination of the NOBOX consensus DNA binding sequence. A, Sequences are selected by recombinant GST-NXHD proteins as described in Materials and Methods following five rounds of CAST are aligned. Underlined nucleotides represent NOBOX DNA binding element used in determination of the consensus binding sequence. B, The percentage frequencies of each nucleotide for each position are shown. The most frequently observed sequence is bold.

Fig. 2. NOBOX binds to the NBE with high affinity. A, 32P-labeled NBE probes containing TAATTG were incubated with purified recombinant GST-NXHD. A 10-fold or 100-fold molar excess of unlabeled TAATTG sequence was used as competitor. B, The complex of recombinant GST-NXHD and 32P-labeled NBE was shifted by the antibodies against GST. C, The alignment of homeodomain of NOBOX, Msx-1, Nkx2.1, Nkx3.1 and Nkx3.2. The asterisk indicates a conserved asparagine (N). D, Purified recombinant GST-NXHD and GST-NXHD(N185Q) were incubated with 32P-labeled NBE probes; The arrow indicates GST-NXHD-DNA complex (BP), super-shifted complex (SS) and free probe (FP). TAATTG containing NBE oligonucleotide sequence is shown in Table 1.

Fig. 3. NOBOX binds to the NBE and NBE-related sequences in vitro. A, 10 ng (lanes 2, 6, 10, and 14), 50 ng (lanes 3, 7, 11, and 15), or 100 ng (lanes 4, 8, 12, and 16) of purified GST-NXHD proteins were incubated with 32P-labeled probes containing TAATTG (lanes 1-4), CAAGTG (lanes 5-8), TAAGTA (lanes 9-12) or TAAAGT (lanes 13-16). NBE oligonucleotide sequence is shown in Table 1. B, Purified GST-NXHD proteins were incubated with 32P-labeled probes containing TAATTG in the either the absence (lane 2) or the presence of a 100-fold excess of unlabeled oligonucleotides containing either the NBE or sequences different from the NBE sequence at each position. C, Consensus DNA binding sequences for NOBOX are shown in the box.

Fig. 4. Wild type newborn ovarian extracts bind NBE. Gel mobility shift assay with ovarian extracts (wild type and Nobox null newborn mice) and 32P-labeled oligonucleotides containing TAATTG with no extracts (lane 1 and 5); wild type ovarian extract (lane 2, 3 and 4); Nobox null ovarian extract (lane 6, 7 and 8); no antisera (lane 2 and 6); anti-NOBOX antiserum (lane 3 and 7); or pre-immune goat serum (lane 4 and 8). TAATTG containing NBE oligonucleotide sequence is shown in Table 1.

Fig. 5. NOBOX transactivates through the NBE and the C-terminus of NOBOX contains transactivational activity. A, The pCDNA3 or pCDNA3-NX was co-transfected with either 3xNBE-pGL3 or 3xNBE-
pGL3 luciferase reporter vector into HEK293 cells. Cell extracts were collected after 48 hours of transfection and analyzed for luciferase activity. The black box indicates the wild type 3xNBE (TAATTG) and the white box indicates the mutated 3xmNBE (TAGGCG). Oligonucleotide sequence is shown in Table 1. B, Diagram of the full-length and truncated NOBOX proteins. Shown are the N-terminus (NT), homeobox domain (HD), and the C-terminus (CT). Numbers indicate the position of amino acid sequences. GAL4DBD represents GAL4 DNA binding domain. C, The full-length and truncated NOBOX proteins fused with GAL4DBD were transfected into HEK293 cells and cellular proteins were extracted in lysis buffer. Cell lysates were resolved by SDS-PAGE and visualized by Western blot analysis with anti-GAL4DBD monoclonal antibody. D, GAL4DBD vector containing the full-length and truncated NOBOX was co-transfected with 5xUAS-TATA-luciferase reporter vector into HEK293 cells. Cell extracts were collected after 48 hours of transfection and analyzed for luciferase activity.

Fig. 6. Characterization of putative NBEs in the mouse Oct4 and Gdf9 promoters using in vitro EMSA. A and C, Oligonucleotides containing predicted Oct4 and Gdf9 NBEs (Table 2) were radiolabeled and incubated with purified recombinant GST-NXHD in the absence or the presence of a 10 or 100-fold excess of unlabeled putative NBE (TAATTG) oligonucleotide. The arrow (BP) indicates GST-NXHD-DNA complex. The other arrow indicates free probe (FP). B and D, The complex of recombinant GST-NXHD and 32P-labeled NBE was shifted by the antibodies against GST. The arrow (SS) indicates super-shifted complex.

Fig. 7. NOBOX activates transcriptional activity of the promoter of mouse Oct4 and Gdf9. A, The pCDNA3-NX was HEK293 cells co-transfected with either pOct4NBE-pGL3 or pOct4mNBE-pGL3 luciferase reporter vector into HEK293 cells. Cell extracts were collected after 48 hours of transfection and analyzed for luciferase activity. The black box indicates the wild type NBE (TAATTG) and the white box indicates the mutated NBE (GGATCC). B, The pCDNA3-NX was HEK293 cells co-transfected with either pGdf9NBE-pGL3 or pGdf9mNBE-pGL3 luciferase reporter vector into HEK293 cells. Cell extracts were collected after 48 hours of transfection and analyzed for luciferase activity. The black box indicates the wild type NBE (TAATTG) and the white boxes indicate the mutated NBEs (GGATCC(-1259), GCTAGC(-967), GGTACC(-786)). C, PCR amplified regions of the chromatin of mouse Oct4 and Gdf9 are described. P1 indicates the promoter region containing NBE. C1 indicates downstream region unrelated to NBE and used as a control. D, Chromatin immunoprecipitation analysis of NOBOX binding for the mouse Oct4 and Gdf9 promoter in newborn ovary or liver extract. PCR amplifies DNA for P1 in mouse Oct4 and Gdf9 co-precipitated with Goat anti-NOBOX antibody. Input represents PCR product from chromatin pellets prior to immunoprecipitation. Samples incubated with 10µl of Goat anti-NOBOX antibody (αNOBOX) and the control sample without antibody (Bead) or IgG were used as template for PCR amplification. PCR primer are shown in Table 1.
Figure 1

A.

```
A   T   A   A   T   G   A   T   T   T
T   T   A   A   T   T   G   C   T   C
C   T   A   A   T   T   A   G   T   G
T   T   A   A   T   T   G   C   T   C
T   T   A   A   T   T   G   C   T   C
-   T   A   A   T   T   G   G   C   T
T   T   A   A   T   T   G   C   T   C
C   T   A   A   T   T   G   A   G   T
T   T   A   A   T   T   G   C   T   C
C   T   A   A   C   G   A   C   C   G
A   T   A   A   A   T   A   G   G   T
-   T   A   A   T   T   G   C   C   T
-   T   A   A   T   T   G   C   C   T
T   T   A   A   T   T   G   G   G   T
C   T   A   A   T   T   G   A   T   G
T   T   A   A   T   T   A   T   C   A
C   T   A   A   T   T   G   G   G   C
C   T   A   A   T   T   G   G   G   C
A   T   A   A   T   G   A   G   A   T
A   T   A   A   T   T   G   A   T   A
A   T   A   A   T   T   A   G   C   C
A   T   A   A   T   T   G   G   C   C
A   T   T   A   T   A   T   G   C   A
T   T   A   A   T   T   G   G   G   T
C   T   A   A   T   T   G   A   T   G
A   T   T   A   A   T   T   G   A   T   G
T   T   A   A   T   T   A   T   C   A
T   A   A   T   T   A   C   C   C   C
A   T   A   A   T   A   A   G   C   A
C   T   A   A   T   T   G   G   T   C
T   C   A   A   G   T   A   C   C   A
C   T   A   A   T   T   G   G   A   T
T   T   A   A   A   G   T   G   T   C
```

B.

```
|   | T   | A   | A   | A   | T   | T   | G   |
|---|-----|-----|-----|-----|-----|-----|-----|
| T | 40  | 94  | 5   | 3   | 92  | 84  | 3   |
| A | 26  | 3   | 95  | 97  | 0   | 5   | 37  |
| G | 0   | 0   | 0   | 0   | 5   | 11  | 58  |
| C | 34  | 3   | 0   | 0   | 2   | 31  | 34  |

(% of each base at each position)
Figure 2

A. $^{32}$P-NBE (TAATTG) Cold Competitor GST-NXHD

B. $^{32}$P-NBE (TAATTG) GST GST-NXHD $\alpha$-GST Ab

C. Nbx 137 RKKRTLYRSQELLFQEDDQYPSDKHRHISQGWQVGTPQRMVWFQNRRAKWKVEKLN 155
Nbx 1 166 NRKPRPTTTAQLALCRKTEQCVGQRGLAERAETSPGTGMLTETCVKIWFQNRRAKAKLQNAE 220
Nkx2.1 161 RKR8VLFSQAVVHELPLEPQQKLYSAPELHLSMTHLPTQVIWFQHRYKMKRQAOK 223
Nkx3.1 125 QRKSRSAFHTQVIELERFSQHYLSAPARKHKLKTETQVKIWFQHRYKTKRQKL 187
Nkx3.2 207 KKR8AFSHAQVFELELRFNHRQYLSPERHALSRLKTETQVKIWFQHRYKTKRQMAA 265

D. $^{32}$P-NBE (TAATTG) GST GST-NXHD GST-NXHD(N185Q) $\alpha$-GST Ab

SS BP FP
**Figure 4**

|            | NBE (TAATTG) |
|------------|--------------|
|            | **Wild type** | **NoBox null** |
| Ovarian extract | -  +  +  +  | -  +  +  +  +  |
| α-NoBox Ab   | -  -  +  -  | -  -  +  -  |
| Serum        | -  -  -  +  | -  -  -  +  |

![Image showing gel electrophoresis results](image)

- SS
- BP
- FP

---

32P-labeled probe
Figure 5

A.

![Bar graph showing relative fold increase](image)

B.

| Variant   | GAL4DBD | NT | HD | CT  | Length |
|-----------|---------|----|----|-----|--------|
| GAL4-NX   |         | 147| 135| 198 | 366    |
| GAL4-NT   |         | 135|     |     |        |
| GAL4-HD   |         |    | 135| 198 |        |
| GAL4-CT   |         |    |    | 198 | 527 a.a.

C.

![Protein gel image](image)

D.

![Graph showing relative fold increase](image)
Figure 6

A. $^{32}$P-labeled probe $-462$

Cold Competitor | $-$ | $-$ | $10X$ | $10X$
GST-NXHD       | $+$ | $+$ | $+$

B. $^{32}$P-labeled probe $-462$

GST           | $-$ | $+$ | $-$ | $-$
GST-NXHD      | $-$ | $-$ | $+$
α-GST Ab      | $-$ | $-$ | $+$

C. $^{32}$P-labeled probe $-1259$ | $-967$ | $-786$

Cold Competitor | $-$ | $-$ | $10X$ | $10X$ | $10X$ | $10X$
GST-NXHD       | $+$ | $+$ | $+$ | $+$ | $+$ | $+$

D. $^{32}$P-labeled probe $-1259$ | $-967$ | $-786$

GST           | $+$ | $+$ | $-$ | $-$ | $+$ | $+$ | $-$ | $+$
GST-NXHD      | $-$ | $-$ | $+$ | $+$ | $+$ | $+$ | $+$ | $+$
α-GST Ab      | $-$ | $-$ | $+$ | $+$ | $+$ | $+$ | $+$ | $+$

SS

BP

FP

SS

BP

FP

FP

FP

FP
| Name             | Sequences                                      |
|------------------|------------------------------------------------|
| CAST-oligo       | 5′-GAG TCC AGC GAA TTC TGT CG-(N)15-CG ACA GAA TTC GCT GGA CTC-3′ |
| CAST-Left        | 5′-GAG TCC AGC GAA TTC TGT CG-3′               |
| CAST-Right       | 5′-GTT GAC ACT CTC GAG GAC TC-3′              |
| NBE(TAATTG)      | 5′-ACG AGC TAC CTT ACT TAA TAG GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| AAATTG           | 5′-ACG AGC TAC CTT ACT AAA TTA GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| CAATTG           | 5′-ACG AGC TAC CTT ACTCAA TTA GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| GAATTG           | 5′-ACG AGC TAC CTT ACTGAA TTA GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| TTATTG           | 5′-ACG AGC TAC CTT ACTTTA TTA GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| TGATTG           | 5′-ACG AGC TAC CTT ACTTTA TTA GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| TATTGG           | 5′-ACG AGC TAC CTT ACTTTA TTA GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| TAATGG           | 5′-ACG AGC TAC CTT ACTTTA TAG GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| TAAATAG          | 5′-ACG AGC TAC CTT ACT TAA TAG GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| TAATTCG          | 5′-ACG AGC TAC CTT ACT TAA TTA GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| TAATAGG          | 5′-ACG AGC TAC CTT ACT TAA TTA GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| TAATTG           | 5′-ACG AGC TAC CTT ACT TAA TTA GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| Nkx2.1(CAATTG)   | 5′-ACG AGC TAC CTT ACT TAA TAGG GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| Nkx3.1(TAAGTA)   | 5′-ACG AGC TAC CTT ACT TAA TAGG GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| Nkx3.2(TAATTG)   | 5′-ACG AGC TAC CTT ACT TAA TAGG GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| 3xNBE            | 5′-ACG AGC TAC CTT ACT TAA TAG GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| 3xmNBE           | 5′-ACG AGC TAC CTT ACT TAA TAG GAC GTC ATA GTG GAA CGA CGT TTA ATT GGA CGA CGT TGA ACG CGT ACT GT-3′ |
| Gdf9-P1-Left     | 5′-GAG ACA TCA CCT TTA ACT TAG GAC AGC AGA CGA GCT GTA-3′ |
| Gdf9-P1-Right    | 5′-AGG ACA TCA CCT TTA ACT TAG GAC AGC AGA CGA GCT GTA-3′ |
| Gdf9-C1-Left     | 5′-TTT TAG TCA CTC GTA ATC GAC AGC AGA CGA GCT GTA-3′ |
| Gdf9-C1-Right    | 5′-GAC CTC GTA GCT ATC GAC AGC AGA CGA GCT GTA-3′ |
| Oct4-P1-Left     | 5′-CTA AAC TCT GGA CGA CTG GAG GT-3′ |
| Oct4-P1-Right    | 5′-ACC CCT CAC AAA CCA GTT GCT GT-3′ |
| Oct4-C1-Left     | 5′-GTA TGA GCC TAC AGG GAC ACC TTT-3′ |
| Oct4-C1-Right    | 5′-GCT CCA GGT TCT CCT GTC TAC CTC-3′ |

Top strand shown for probes used in EMSA.
**Table 2.** Sequences of putative NBE in the promoter of mouse *Oct4* and *Gdf9*

| Gene | Position | Core sequence (5'-3') |
|------|----------|-----------------------|
| Oct4 | -426     | 5'-AGTGGCCCAGAAA**TAATTG**GCACACGAACAAT-3' |
| Gdf9 | -1259    | 5'-CGATGTCTATT**TAATTG**CATTTGCTGATGT-3' |
| Gdf9 | -967     | 5'-GTGCGTTTTTTTT**TAATTG**TTTATTTTCAATT-3' |
| Gdf9 | -786     | 5'-ATTGTATGGTT**TAATTG**GCTCAAAGTTAAAG-3' |

Putative NBE sequences are bold and underlined. Top strand shown for probes used in EMSA.
