

## Identification and Characterization of LHX8 DNA Binding Elements

Miree Park<sup>1,\*</sup>, Sanghyun Jeon<sup>1,\*</sup>, Ji-Hye Jeong<sup>1</sup>, Miseon Park<sup>2</sup>, Dong-Ryul Lee<sup>1</sup>, Tae Ki Yoon<sup>2</sup>,  
Dong Hee Choi<sup>3,¶</sup> and Youngsok Choi<sup>1,2,†</sup>

<sup>1</sup>Dept. of Biomedical Science, CHA University, Seoul 135-081, Korea

<sup>2</sup>Fertility Center of CHA Gangnam Medical Center, CHA University, Seoul 135-081, Korea

<sup>3</sup>Dept. of Obstetrics and Gynecology, CHA Bundang Medical Center, School of Medicine, CHA University, Seongnam 463-7121, Korea

**ABSTRACT** : *Lhx8* (LIM homeobox 8) gene encodes a LIM homeodomain transcriptional regulator that is preferentially expressed in germ cells and critical for mammalian folliculogenesis. However, Lhx8 DNA binding sequences are not characterized yet. We aimed to identify and characterize a cis-acting sequence of germ-cell specific transcriptional factor, Lhx8. To identify Lhx8 DNA binding element, Cyclic Amplification of Sequence Target (CAST) Analysis was performed. Electrophoretic Mobility Shift Assay (EMSA) was processed for the binding specificity of Lhx8. Luciferase assay was for the transcriptional activity of Lhx8 through identified DNA binding site. We identified a putative cis-acting sequence, TGATTG as Lhx8 DNA binding element (LBE). In addition, Lhx8 binds to the LBE with high affinity and augments transcriptional activity of luciferase reporter driven by artificial promoter containing the Lhx8 binding element. These findings indicate that Lhx8 directly regulates the transcription of genes containing Lhx8 binding element in oocytes during early folliculogenesis.

**Key words** : Lhx8, Transcription factor, Folliculogenesis

### INTRODUCTION

*LHX8* is a member of the LIM homeobox gene family (Kitanaka et al., 1998). In mice, *Lhx8* preferentially expressed in oocytes of germ cell clusters and primordial, primary, and antral follicles in the ovary (Choi et al., 2008; Pangas et al., 2006). *Lhx8* deficient female mice are infertile, whereas knockout male is fertile. *Lhx8* is required for development and differentiation of the oocytes during folliculogenesis after birth in the ovary (Choi et al., 2008). At the time of birth, histological structures of Lhx8 deficient (*Lhx8*<sup>-/-</sup>) ovaries are grossly similar to

the wild-type ovaries (Choi et al., 2008). The number and morphology of primordial follicles in *Lhx8* deficient ovaries shows no significant differences (Choi et al., 2008). However, the activation and differentiation of primordial follicles in *Lhx8* deficient ovaries are arrested at primary follicle stage, and then they lose the follicles within postnatal day 7 (Choi et al., 2008). Based on the analysis of gene profiles from newborn ovaries of wild type and *Lhx8* deficient mice, numerous oocyte-specific genes including such as zona pellucida genes 1~3 (*Zp1*~*3*), growth differentiation factor 9 (*Gdf9*), bone morphogenetic protein 15 (*Bmp15*), and *Pou5f1* (also known as *Oct4*), *Figla* (Factor in the germline alpha), *Nobox* (Newborn ovary homeobox), *Sohlh1* (Spermatogenesis and oogenesis specific basic helix-loop-helix transcription factor 1), and *Pad6* (Peptidylarginine deiminase), are mis-regulated in *Lhx8* deficient ovaries (Choi et al., 2008). However, it is unknown whether Lhx8 directly or indirectly regulates transcription of these oocyte-

† To whom correspondence should be addressed at:

Corresponding author: Youngsok Choi, Dept. of Biomedical Science, CHA University, Seoul 135-081, Korea. Phone: +82-2-3468-3504, Fax: +82-2-3468-3446, E-mail: youngsokchoi@cha.ac.kr

¶ Co-corresponding author: Dong Hee Choi, Dept. of Obstetrics and Gynecology, CHA Bundang Medical Center, School of Medicine, CHA University, Seongnam 463-7121, Korea.

\* These authors are equally contributed.

specific genes during development of the ovarian follicle. In the present study, we identified DNA sequences that Lhx8 binds and show that Lhx8 transactivates luciferase reporter containing Lhx8 DNA binding element (LBE).

## MATERIALS AND METHODS

### 1. Expression and purification of GST-Lhx8HD

To create the pET41b-Lhx8 homeodomain protein (GST-Lhx8HD) and bacterial expression construct, the insert was amplified by PCR and cloned into pET41b (Novagen, Madison, WI). We verified the sequences of the resulting GST-Lhx8HD fusion constructs to ensure that no mutations have been introduced during PCR cloning. We transformed BL21-pLysS *Escherichia coli* (Stratagene) with GST-Lhx8HD construct and expressed it by inoculating Luria-Bertani (LB) media containing 20 µg/ml kanamycin with overnight culture (1:20 dilution) at 37°C to an OD<sub>600</sub> of 0.5. Protein expression was induced at 30°C with 2mM of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to OD<sub>600</sub> of 1.0. We stored cell pellets at -80°C, thawed and lysed them with BugBuster lysis buffer (Novagen). We purified the GST-Lhx8HD fusion proteins with GST bind resin (Novagen), were dialyzed them three times in 1 liter of phosphate-buffered saline and quantified them. Purified proteins were stored at -80°C until use.

### 2. Plasmid for luciferase assay

The pGL4-promoter vector with SV40 promoter (Promega) was used for constructing luciferase reporter vectors containing three repeats of LBE (3xLBE-Luc). Overexpression vector carrying the mouse *Lhx8* were constructed by cloning the full-length of *Lhx8* cDNA into pCMV-Tag3 or Tag5 vector (Agilent Technologies).

### 3. Electrophoretic mobility shift assay (EMSA)

EMSA were performed in 20 µl reaction mixtures at room temperature at a final concentration of 10mM

Tris, pH 7.5, 50 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 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 [ $\alpha$ -<sup>32</sup>P] dCTP and Klenow polymerase (Invitrogen). Binding reactions were conducted by incubating <sup>32</sup>P-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 (Amersham Biosciences). For competition assay, purified proteins were incubated at room temperature with cold competitors for 15 min before addition of the <sup>32</sup>P-labeled probe.

### 4. Cyclic amplification of sequence target (CAST) analysis

The CAST assay was carried out essentially as previously described with some modifications (Choi et al., 2006). We used purified GST-Lhx8HD proteins and GST-bind resin (Novagen) for the CAST assay. Binding reactions for CAST were performed in DNA binding buffer (10 mM Tris, pH 7.5, 50 mM NaCl, 7.5 mM MgCl<sub>2</sub>, 1mM EDTA, 5% glycerol, 5% sucrose, 0.1% Nonidet P-40, and 5 mg/ml bovine serum albumin). 0.5 pmol of short double-stranded DNA containing a 15-bp random sequence flanked by 20-bp fixed sequences, CAST oligonucleotides was incubated with 100ng of purified GST-Lhx8HD protein for 20min 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). The inserted DNAs were sequenced and scored to get the consensus sequences.

### 5. Cell culture and reporter assays

Human embryonic kidney cells (HEK293) were grown in Dulbecco's modified Eagle's medium supplemented 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 hr before harvest. For each transfection, 200 ng of reporter construct, 200 ng of the indicated expression plasmid, and 10 ng of pRT-TK normalization plasmid were used per single well of a 6-well plate. Dual luciferase assays were carried out with total cell extracts as recommended by Promega. The transient transfection experiments were performed in triplicate, and results were normalized to the expression of *Renilla* luciferase.

### RESULTS

#### 1. Determination of the Lhx8 DNA binding sequence

We determined a consensus DNA binding sequence for Lhx8 using the CAST assay as previously described

**A**

Cast-29	1	--ATCA-CGC-ACTGATTG-----	15
Cast-33	1	--ATCA-CGC-ACTGATTG-----	15
Cast-15	1	--ATCA-GGC-AGTGATTG-----	15
Cast-19	1	--AATCA-CGG--CTGATTG-----	15
Cast-17	1	--AATCA--GT--TGATTGCT-----	15
Cast-11	1	--ATCA-AGC-GGTGATTG-----	15
Cast-23	1	--ATCA-----ACTGATTGGCT-----	15
Cast-28	1	--ATCA--GCAGCTGATTG-----	15
Cast-1	1	--ATCA-CGC-GCTGATTG-----	15
Cast-20	1	--ATCA--CATGCTGATTG-----	15
Cast-31	1	--ATCA--C--TTGATTGCGC-----	15
Cast-14	1	--ATCA--GCGGCTGATTG-----	15
Cast-3	1	--ATCA--GGCGCTGATTG-----	15
Cast-5	1	--ATCA--GCTGCTGATTG-----	15
Cast-6	1	--ATCA--GC--TGATTGCTA-----	15
Cast-21	1	--ATCA--GCGCCTGATTG-----	15
Cast-24	1	--ATCA--GCCCTGATTG-----	15
Cast-26	1	--ATCA--GCTGCTGATTG-----	15
Cast-32	1	--ATCA--GGGGTGATTG-----	15
Cast-7	1	-----TGATTGCGC-CAATCA-----	15
Cast-16	1	-----TGATTGC-----AATCGACG-	15

**B**

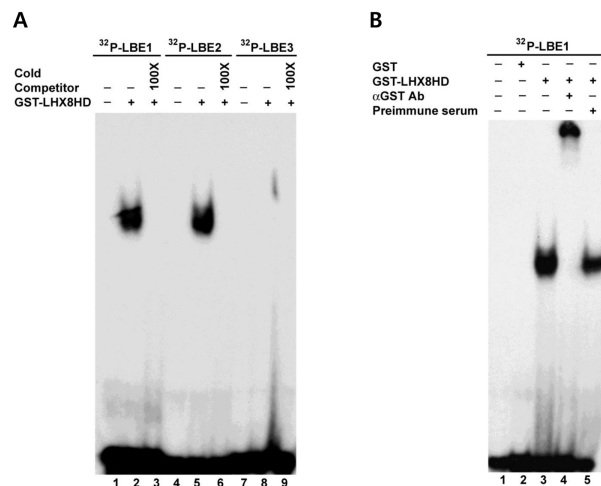
	<b>A</b>	<b>T</b>	<b>C</b>	<b>A</b>	-----	<b>T</b>	<b>G</b>	<b>A</b>	<b>T</b>	<b>T</b>	<b>G</b>
<b>T</b>	0	100	0	0		100	0	0	100	100	0
<b>A</b>	100	0	0	100		0	4	100	0	0	0
<b>G</b>	0	0	0	0		0	96	0	0	0	100
<b>C</b>	0	0	100	0		0	0	0	0	0	0

**Fig. 1. DNA binding sequence of the Lhx8.** (A) Sequences selected by recombinant GST-Lhx8HD proteins as described under "Materials and Methods" following five rounds of CAST are aligned. (B) The percentage frequencies of each nucleotide for each position are shown.

(Choi et al., 2006). We used a predicted homeodomain portion of Lhx8 (a.a 246-308) fused to GST-Lhx8HD. DNA sequences that bound GST-Lhx8HD went through five cycles of CAST. The 21 selected sequences for GST-Lhx8HD 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 96 to 100%. CAST assay revealed two groups of consensus sequence, ATCA-TGATTG (Fig. 1B), as the Lhx8 DNA Binding Element (LBE).

#### 2. Lhx8 binds to the LBE1 and LBE2 with high affinity

We next confirmed the interaction between LBE and GST-Lhx8HD using EMSA. Using a competitive binding assay (Fig. 2A), GST-Lhx8HD protein bound to <sup>32</sup>P-labeled LBE1 containing sequences, TGATTG, or LBE2 (ATCATGATTG) probes (Fig. 2A). However, GST-Lhx8HD protein did not form DNA-protein complex

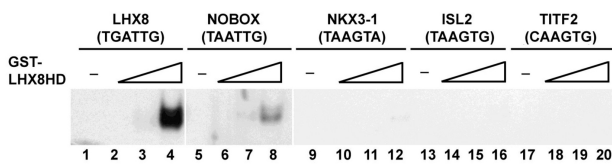


**Fig. 2. Lhx8 binds to the LBE with high affinity.** (A) <sup>32</sup>P-labeled LBE probes containing TGATTG were incubated with purified recombinant GST-Lhx8HD. A 100-fold molar excess of unlabeled TGATTG sequence was used as competitor. (B) Complex of recombinant GST-Lhx8HD, <sup>32</sup>P-labeled LBE was shifted by the antibodies against GST; TGATTG (LBE1), ATCATGATTG (LBE2), ATCA (LBE3).

with  $^{32}\text{P}$ -labeled LBE3 (ATCA) probe (Fig. 2A). The bound complex was competed out by 100-fold molar excess of unlabeled cold each competitor (Fig. 2A). To further confirm the specificity of these DNA-protein complexes, anti-GST antibodies were added to the binding reaction in the absence or presence of GST-Lhx8HD or GST protein. As shown in Fig. 2B, the anti-GST antibody further shifted the DNA-protein complex (Fig. 2B). To confirm the effect of GST, preimmune serum was added to the binding reaction in the presence of GST-Lhx8HD (Fig. 2B).

### 3. LBE-related sequences and Lhx8 complex

We examined whether the homeobox domain of Lhx8 has binding specificity with other homeobox consensus core sequences. Increasing amounts of purified GST-Lhx8HD protein incubated with  $^{32}\text{P}$ -labeled probes containing known other homeobox consensus core sequences, TAATTG for Nobox, TAAGTA for Nkx3.1, TAAGTG for Isl2, or CAAGTG for Titf2 were analyzed using EMSA. At high concentration (100 nanogram), GST-Lhx8HD protein bound to these related sequences with low affinity compared with the LBE (Fig. 3). Only the sequence, TAATTG showed low affinity with GST-Lhx8HD under high concentration in lane 8 (Fig. 3), whereas Lhx8 didn't form DNA-protein complex with other related sequences. This result suggests that Lhx8 preferentially binds to the LBE containing 6 core sequences, TGATTG.



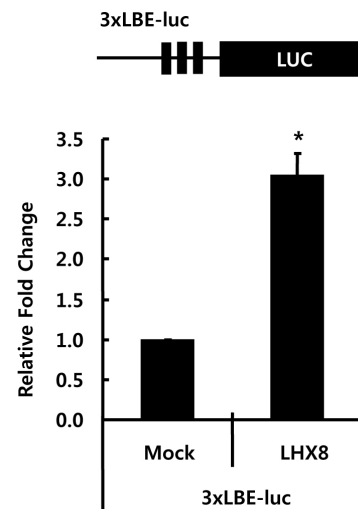
**Fig. 3. Lhx8 binds to the LBE and LBE-related sequence *in vitro*.** 10 ng (lanes 2, 6, 10, 14, 18), 50 ng (lanes 3, 7, 11, 15, 19), or 100 ng (lanes 4, 8, 12, 16, 20) of purified GST-Lhx8HD proteins were incubated with  $^{32}\text{P}$ -labeled probes containing TGATTG (lanes 1-4), TAATTG (lanes 5-8), TAAGTA (lanes 9-12), TAAGTG (lanes 13-16), or CAAGTG (lanes 17-20).

### 4. Lhx8 transactivates the reporter via the LBE

To examine whether Lhx8 can regulate the expression of LBE driven reporter gene *in vivo*, we constructed a luciferase reporter vector containing three LBEs (TGATTG), 3xLBE-pGL4 upstream of the luciferase reporter gene. This reporter vector was co-transfected with either empty vector (pCMV-Tag5) or Lhx8 overexpression vector (pCMV-Tag5) into HEK293 cells. The relative luciferase activity of reporter construct containing three copies of TGATTG was increased by 2.3-fold with the overexpression of Lhx8 (Fig. 4). This suggests that Lhx8 can activate the transcription of target genes through the LBE.

## DISCUSSION

In the previous study, we showed that numerous oocyte-specific genes, including *Gdf9* and *Nobox*, were down-regulated in ovaries that lack *Lhx8* (Choi et al., 2008). However, it is unknown whether Lhx8 directly



**Fig. 4. Transactivation of reporter genes through the LBE.** The pCMV Tag5 or pCMV Tag5-Lhx8 was co-transfected with 3xLBE-pGL4 luciferase reporter vector into HEK293 cells. Cell extracts were collected after 48h of transfection and analyzed for luciferase activity. The black box indicates the artificial 3xLBE (TGATTG). (\* $P < 0.01$ )

or indirectly regulates transcription of these oocyte-specific genes during development of the ovarian follicle. To address the direct downstream genes of Lhx8, it is required for characterization of Lhx8 as a transcription factor. In the present study, we have identified a consensus DNA binding sequence, TGATTG, as Lhx8 DNA binding elements using purified proteins in a cyclic amplification target assay (Fig. 1). The consensus DNA binding sequence of Lhx8 consists of six base pairs as core sequences of known homeoproteins including *Msx1*, *Titf2*, *Nkx3-1*, *Isl2* and *Nobox* (Choi et al., 2006; Catron et al., 1993; Guazzi et al., 1990; Kim et al., 2003). The LBE core sequence for Lhx8 is slightly different from others; CAAGTG for *Titf2*, TAAGTA for *Nkx3-1*, TAAGTG for *Isl2* and TAATTG/A for *Nobox*. This suggests that their functions might be redundant in the regulation of their down-stream targets. However, *Msx1*, *Titf2*, *Nkx3-1*, and *Isl2* except *Nobox*, are not expressed in the newborn ovaries (data not shown). In the previous study, *Nobox* is exclusively expressed in the oocyte like *Lhx8* (Rajkovic et al., 2004; Suzumori et al., 2002). We characterized a binding site and target genes of *Nobox* in the ovary (Choi et al., 2006; Choi et al., 2007). The Lhx8 DNA binding element, TGATTG, is similar to TAATTG/A, that of *Nobox*. This suggests that their functions might be overlapped in the regulation of their down-stream targets. However, each knockout mice show different phenotype (Choi et al., 2008; Rajkovic et al., 2004). *Lhx8* deficiency looks more severe than *Nobox* deficiency. It suggests that the regulatory network of Lhx8 might be independent on that of *Nobox*. These results spotlight that the transcriptional activity of Lhx8 requires direct DNA binding through specific consensus sequences, LBE.

#### ACKNOWLEDGEMENT

This work was supported by Priority Research Centers Program through the National Research Foundation of

Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0093821), by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012003799), and by the Korea Healthcare technology R&D Project, Ministry for Health, Welfare & Family Affairs, Republic of Korea (A084923).

#### REFERENCES

- Kitanaka J, Takemura M, Matsumoto K, Mori T, Wanaka A (1998) Structure and chromosomal localization of a murine LIM/homeobox gene, *Lhx8*. *Genomics* 49: 307-309.
- Choi Y, Ballow DJ, Xin Y, Rajkovic A (2008) Lim homeobox gene, *lhx8*, is essential for mouse oocyte differentiation and survival. *Biol Reprod* 79:442-449.
- Pangas SA, Choi Y, Ballow DJ, Zhao Y, Westphal H, Matzuk MM, Rajkovic A (2006) Oogenesis requires germ cell-specific transcriptional regulators *Sohlh1* and *Lhx8*. *Proc Natl Acad Sci USA* 103:8090-8095.
- Choi Y, Rajkovic A (2006) Characterization of NOBOX DNA binding specificity and its regulation of *Gdf9* and *Pou5f1* promoters. *J Biol Chem* 281:35747-35756.
- Catron KM, Iler N, Abate C (1993) Nucleotides flanking a conserved TAAT core dictate the DNA binding specificity of three murine homeodomain proteins. *Mol Cell Biol* 13:2354-2365.
- Guazzi S, Price M, De Felice M, Damante G, Mattei MG, Di Lauro R (1990) Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity. *Embo J* 1990;9:3631-3639.
- Kim DW, Kempf H, Chen RE, Lassar AB (2003) Characterization of *Nkx3.2* DNA binding specificity and its requirement for somitic chondrogenesis. *J Biol Chem* 278:27532-27539.
- Steadman DJ, Giuffrida D, Gelmann EP (2000) DNA-

binding sequence of the human prostate-specific homeodomain protein NKX3.1. *Nucleic Acids Res* 28:2389-2395.

Choi Y, Qin Y, Berger MF, Ballow DJ, Bulyk ML, Rajkovic A (2007) Microarray analyses of newborn mouse ovaries lacking Nobox. *Biol Reprod* 77:312-319.

Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM (2004) NOBOX deficiency disrupts early folli-

culogenesis and oocyte-specific gene expression. *Science* 305:1157-1159.

Suzumori N, Yan C, Matzuk MM, Rajkovic A (2002) Nobox is a homeobox-encoding gene preferentially expressed in primordial and growing oocytes. *Mech Dev* 111:137-141.

---

(Received 4 October 2012, Received in revised form 13 November 2012, Accepted 17 November 2012)