行政院國家科學委員會專題研究計畫　成果報告

小鼠胚胎幹細胞表現之新穎基因功能之分析

計畫類別：個別型計畫
計畫編號：NSC93-2314-B-040-033-
執行期間：93 年 08 月 01 日至 94 年 07 月 31 日
執行單位：中山醫學大學生物醫學科學學系

計畫主持人：王淑紅
共同主持人：李鴻

報告類型：精簡報告

處理方式：本計畫涉及專利或其他智慧財產權，1 年後可公開查詢

中　華　民　國 94 年 11 月 2 日

行政院國家科學委員會補助專題研究計畫 ■成果報告 □期中進度報告

小鼠胚胎幹細胞表現之新穎基因功能之分析

計畫類別：■ 個別型計畫　□ 整合型計畫

計畫編號：NSC 93-2314-B-040-033

執行期間：93 年 8 月 1 日至 94 年 7 月 31 日

計畫主持人：王淑紅

共同主持人：李鴻

計畫參與人員：

成果報告類型(依經費核定清單規定繳交)：■精簡報告　□完整報告

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢

□涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：中山醫學大學

中　華　民　國九十四年十月 三十日

ABSTRACT

The homeobox gene families play important roles in the transcriptional regulation of gene expression prior to and during embryo development. To identify novel homeobox genes expressed in early embryo development, we conducted a degenerate oligonucleotide polymerase chain reaction to screen the mouse embryonic stem (ES) cell cDNA library. A novel homeobox-containing gene, *Eso-1*, which is preferentially expressed in ES cells and ovaries, was identified. In this study, the full length of *Eso-1* cDNA was found to be 1710 bp with a predicted homeodomain without significant homology to reported homeobox proteins. *Eso-1* was mapped to chromosome 14A3. RT-PCR analysis showed that *Eso-1* was expressed through oogenesis and continues to be expressed through to the blastocyst stage. De novo expression of *Eso-1* started at 13.5 dpc (days postcoitum) in the ovaries, which coincides with the initiation of oogenesis. Northern blot analyses demonstrated that *Eso-1* is preferentially expressed in both ovaries and ES cells as a 1.7 kb transcript. Results from whole mount in situ hybridization revealed that the signals of *Eso-1* in oocytes increased from primordial to antral follicles. We suggest that ESO-1 protein may play a role during oocyte maturation and initiation of early embryonic development.

Key word: oogenesis, homeobox gene, cytoplasmic polyadenylationj

摘要:

同源箱基因在胚胎發育過程扮演重要的角色，我們以退化性引子進行PCR篩選胚幹細胞中所表現之新穎同源箱基因，其中一個命名為Eso-1，因為其表現於卵細胞與胚幹細胞中，經由分子層次分析此基因，發現位於小鼠14號染色體上由四個exon所組成，且其所表現出之mRNA有alternative splicing 的現象，主要的mRNA長約1710bp，由mRNA推測出之氨基酸組成，發現Eso-1基因的同源箱區與其他已知之基因之相似性很低，只有38-39%的相似性，再進一步分析Eso-1基因在發育過程中之表現趨勢，發現此基因是一個卵細胞專一性表達之基因，從13.5天之胎兒之卵巢中即有小量Eso-1mRNA的表現，自此後mRNA量慢慢增加至卵細胞成熟時期，成熟卵與授精卵之表現量最高，再往後之發育時期Eso-1的表現量慢慢減少至囊胚期，相當於囊胚期中之inner cell mass之胚幹細胞也有此基因的表現，但分化之胚幹細胞表現量明顯下降，而與胚幹細胞性質相似之P19(teratocarcinoma cell)卻沒有表現Eso-1基因，另外我們也以in situ hybridization 分析Eso-1基因在卵細胞成熟過程中之表現，發現從primodial 至antral follicle中之卵細胞都有Eso-1基因表現的訊號，我們推測此基因在卵細胞成熟或早期胚胎發育扮演重要功能。

INTRODUCTION

Mammalian oogenesis is a complex process regulated by the expression and interaction of a multitude of genes [1]. An oocyte must undergo several developmental transitions during which it synthesizes a unique set of proteins to support fertilization and early embryonic development. Thus, the genes expressed specifically at the oocyte stage may play important roles in fertilization, pre-implantation and folliculogenesis [2]. However, little is known about gene expression profiles within an oocyte during oogenesis and how these oocyte-specific gene expressions are regulated. Although it has been speculated that several hundred genes participate in oogenesis, only a few have been identified so far [3]. The elucidation of more oocyte-specific expressed genes will improve the understanding of the process of oogenesis and early embryo development.

During oogenesis, a large number of maternal proteins are synthesized. They accumulate in the oocytes prior to fertilization and are utilized in early embryogenesis [1]. During meiosis, germ cell genomes are transcriptionally silenced. Following fertilization, embryonic transcription is first detected at the late one-cell zygote stage and is required for development beyond the two-cell stage [4, 5]. Hence, the oocyte-to-embryo transition depends on the maternal transcripts and proteins that accumulated during oogenesis. The functions of maternal effect genes have been characterized by gene disruption experiments, such as Maternal Antigens That Embryos Require (*Mater*) [6], Zygotic Arrest 1 (*Zar1*) [7], *Stella* [8] and nucleoplasmin2 (*Npm2*) [9]. Early embryos lacking these maternal effect genes exhibit developmental defects and embryonic development is arrested either at the one or two cell stage. Other oocyte-specific genes have also been identified by in silico cloning, however, it is unclear whether the transcripts of these genes are translated [10-12]. As mouse oocytes grow, they accumulate transcripts that are either translated directly into proteins or stored for later activation by cytoplasmic polyadenylation. Cytoplasmic polyadeylation requires two cis-acting 3'-untranslated region (UTR) sequences: the conserved hexanucleotide polyadenylation signal (HPS), AAUAAA, and the cytoplasmic polyadenylation element (CPE), UUUUA$_{1\text{-}2}$U [13]. In general, the poly (A) tails of dormant maternal mRNAs are elongated (from 20-50 to 100-150 nucleotides) during oocyte maturation and then linked with activation of translation [14]. A number of maternal mRNAs essential for successful oogenesis and early embryogenesis in mice, such as *c-mos*, *tPA*, *cyclin B1* and *H1oo*, have been shown to require cytoplasmic polyadenylation for translational activation [15-18]. This timely controlled translation is a general process guiding the oocyte to embryo transition [19]. However, the molecular mechanisms that activate maternal effect genes and regulate their products are largely unknown. In addition to the above translational control of maternal mRNAs, transcriptional activation of oocyte-specific gene expression is also thought to be required for completion of oogenesis and early embryonic events. However, the transcriptional regulations of oocyte-specific gene expressions are still poorly understood.

Transcription factors that regulate oocyte gene expressions, such as *Zfx* (zinc-finger transcription factor on the X chromosome) [20], *Nobox* (newborn ovary homeobox-encoding gene) [21] and *Figalpha* (germ cell-specific transcription factor), should play important roles in

fertility [22, 23]. Mice lacking *Zfx* have a reduced number of oocytes and diminished fertility. Disruption of the transcription factor, NOBOX, results in infertility in female mice, while male mice are unaffected [21]. FIGalpha regulates the coordinated transcription of three ZP glycoproteins and plays an important role in primordial follicle formation [22, 23].

Homeobox gene superfamily members act as transcription factors regulating gene expressions during developmental processes [24]. In this study, we identify novel homeobox genes involved in early embryo development. We used a degenerate polymerase chain reaction (PCR) to screen a murine embryonic stem (ES) cell cDNA library and several homeobox containing genes were isolated. One of the novel homeobox-containing genes, specifically expressed in oocytes and ES cells, was cloned and named *Eso-1*. Here we report the sequences and expression profiles of *Eso-1*. In addition, the primary structure of *Eso-1* gene was deduced from a single open reading frame of cDNA that encodes a 185-amino acid protein with a predicted molecular mass of 21 kDa. The expression patterns of *Eso-1* suggest that this gene might be regulated by cytoplasmic polyadenylation during oocyte to embryo transition. ESO-1 could be a new maternal-effect transcription factor and play an important role in oogenesis and early embryo development.

RESULTS

Identification of *Eso-1* by Degenerate PCR and EST-database Analyses

To identify novel homeobox-containing genes involved in early embryonic development, we conducted a degenerate PCR screening of a murine ES cell cDNA library. A PCR product containing a novel homeodomain sequence was identified. The DNA sequence of this PCR product was used for a BLAST homology search of the database at the National Center for Biotechnology Information (NCBI; http:/www.ncbi.nlm.nih.gov). Several EST clones were found that were highly homologous to this PCR product. These clones were derived from eggs to 3.5 day mouse embryos. The majority of the clones were derived from unfertilized and fertilized eggs. In addition, one clone (AK049114) from a Riken full-length enriched ES cell cDNA library was also found to contain almost the same sequence as this PCR fragment. Based on these database sequences, full-length cDNA was cloned by RT-PCR amplification from ES RNAs and then sequenced. The sequence of this cDNA was further used for a homology search in Ensemble Genome Data Resources (Sanger Institute, http://www.ensemble.prg/Musmusculus/). The result indicated that this novel homeobox containing gene is located on mouse chromosome 14A3 (Fig. 1A). The partial cDNA fragment from PCR amplification was also used as a probe for preliminary Northern blot analysis to determine the expression pattern and transcription size of this novel homeobox gene. After preliminary Northern blot and RT-PCR analyses, this novel gene exhibited ES cell and oocyte specific gene expression, so we named this gene *Eso-1* (ES cell and oocyte expressing gene).

Based on the sequences of AK049114 and the cDNA sequences of *Eso-1* obtained, we determined that exon1 of *Eso-1* was 57 bp in length. To authenticate and search for other 5' ends of *Eso-1* gene, we performed 5'-RACE analysis using ES RNAs as templates. To increase sensitivity and specificity, two different *Eso-1* specific primers, P4 and P3, were used in two

successive rounds of nested PCR (Fig. 1B). Several RACE clones were subcloned and sequenced. Multiple 5' ends of *Eso-1* transcripts were mapped upstream of the previous 5' site of AK049114. Figure 2A indicates different transcription start sites of *Eso-1* transcripts. The full-length of *Eso-1* is longer than that of AK049114 (1654 bp). However, the 5'- and 3'-splice sites are the same in all RACE clones obtained as in AK049114. This result confirms the existence of all predicted splice sites and the 9.3 kb, 2.7 kb and 0.093 kb introns located in the *Eso-1* pre-mRNA (Fig. 1B). Comparison of sequences between *Eso-1* cDNA and *Eso-1* genomic DNA from the database revealed that *Eso-1* gene is composed of four exons with the splice sites obeying the GT-AG splicing donor-acceptor rule. Using the P1 and P4 primers, two PCR products containing the 5' ends of *Eso-1* cDNA were obtained from ES cell cDNAs (Fig. 1, B and C). The longer PCR product (686 bp) was consistently fainter than the shorter one (593 bp). This reflects the relative abundance, in vivo, of the shorter transcript. After subcloning and sequencing, the sequences of these two PCR products were identical from exon1 to exon4, but were different at intron3 (93 bp), which was only included in the longer one. This indicates the existence of alternative splicing of *Eso-1* transcripts.

Although *Eso-1* transcripts exhibit multiple transcription star sites at exon1, all *Eso-1* transcripts contain the same open reading frame, which encodes a protein with 182 predicted amino acids containing a homeodomain (Accession number:DQ224405). Comparison of the homeodomain of this novel gene with other homeodomain-containing genes revealed a maximum of 38% identity with homeodomains of the *OG9* gene [29] and other known genes [30-33], indicating that *Eso-1* does not belong to any previously isolated homeobox gene family (Fig. 2). The sequence of the homeodomain of *Eso-1* gene is located in exon2 and exon3 (Fig. 1, B and C).

Tissue- and Stage-Specific Expressions of *Eso-1*

To determine the expression patterns of *Eso-1* during early embryo development and within adult tissues, we performed RT-PCR analyses using RNAs from various adult tissues and different stages of embryos. The results showed that the expressions of *Eso-1* were detected in ES cells, adult ovaries and preimplantation embryos from eggs to blastocysts (Fig. 3, A and B). The expressions of *Eso-1* decreased from eggs to blastocysts, with a predominant expression in unfertilized and fertilized eggs (Fig. 3B). The expression of *Eso-1* in the blastocysts was retained within the inner cell mass (ICM), as demonstrated by whole mount in situ hybridization analysis (data not shown). This result was confirmed by detection of *Eso-1*expression in ES cells (Fig. 3A). Since ES is derived from the ICM of the blastocysts and retains the gene expression and pluripotency characteristics similar to the ICM. The expressions of *Eso-1* in undifferentiated and differentiated ES cells were further determined by RT-PCR. *Eso-1* was expressed at a higher level in undifferentiated ES cells than in differentiated ES cells (1 uM retinoic acid treated for 2 days) (Fig. 3C). However, the embryonic teratocarcimona cell line, P19, exhibited no *Eso-1* gene expression (Fig. 3C). To determine the expression of *Eso-1* in different stages of ovaries, we performed RT-PCR using RNAs from ovaries at various stages from fetus to adult. The expression of *Eso-1* started at 13.5 dpc and continued to the adult stages (Fig. 3D). Furthermore,

the expressions of *Eso-1* showed no difference between ovaries with or without eCG stimulation (data not shown).

The expressions of *Eso-1* were also confirmed as a 1.7 kb transcript specifically in ES cells and ovaries with northern blot analyses (Fig. 4, A and B). The size of the *Eso-1* transcript was in agreement with our predicted length (1710 bp) and the ES cell was the only cell line that we have detected to express *Eso-1* (Fig. 4C). In adult ovaries, *Eso-1* could be detected only in oocytes. No *Eso-1* signal was detected in cumulus cells isolated from ovulated oocytes by RT-PCR (Fig. 3E). Following implantation, the expression of *Eso-1* gene was not detected in 7 day to 17-day old mouse embryos (Fig. 4D). We further characterized the expression of *Eso-1* mRNA in the ovaries of 4-wk and 8-wk old mice by whole mount in situ hybridization, respectively. Antisense probes hybridized specifically to oocytes in the ovaries (Fig. 5, B and D). Following the whole mount in situ hybridization, ovaries were fixed and sectioned, and the signals of *Eso-1* mRNAs were observed. Specific signals for *Eso-1* mRNA were clearly observed in oocytes of antral and secondary follicles and faint *Eso-1* signals were detected in oocytes of primordial and primary follicles (Fig. 5F).

DISCUSSION

In the present study, we report a novel ES cell and oocyte-specific gene, *Eso-1*, which is a homeobox-containing gene. The mouse *Eso-1* was isolated by degenerate PCR screening of a murine ES cell cDNA library. To determine the full-length *Eso-1* cDNA, we performed both 5'/3'RACE reactions based on the sequences from the mouse EST database. The *Eso-1* mRNA was first detected in the ovaries of 13.5 dpc fetuses and continued to be expressed during oogenesis. This period of 13.5 dpc corresponds to the time when germ cells enter meiosis in the female gonad, suggesting that the *Eso-1* gene may have roles in ovarian development and folliculogenesis as FIGalpha does [23]. Following fertilization, the expression of *Eso-1* is gradually reduced and stops after the blastocyst stage. From the expression patterns and the existence of cytoplasmic polyadenylation, *Eso-1* should be a novel maternal-effect gene. Furthermore, the predicted ESO-1 protein is a homeodomain protein that may bind and regulate downstream gene expressions. Our results suggest that *Eso-1* might be important for folliculogenesis, oogenesis and zygotic transcription of preimplantation embryos.

REFERENCES

1. Song JL, Wessel GM. How to make an egg: transcriptional regulation in oocytes. Differentiation 2005; 73: 1-17.
2. Dean J. Oocyte-specific genes regulate follicle formation, fertility and early mouse development. J Reprod Immunol 2002; 53: 171-180.
3. Ko MS, Kitchen JR, Wang X, Threat TA, Hasegawa A, Sun T, Grahovac MJ, Kargul GJ, Lim MK, Cui Y, Sano Y, Tanaka T, Liang Y, Mason S, Paonessa PD, Sauls AD, DePalma GE, Sharara R, Rowe LB, Eppig J, Morrell C, Doi H. Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development. Development 2000; 127: 1737-1749.

4. Schultz RM. Regulation of zygotic gene activation in the mouse. Bioessays 1993; 15: 531-538.
5. Latham KE, Beddington RS, Solter D, Garrels JI. Quantitative analysis of protein synthesis in mouse embryos. II: Differentiation of endoderm, mesoderm, and ectoderm. Mol Reprod Dev 1993; 35: 140-150.
6. Tong ZB, Gold L, Pfeifer KE, Dorward H, Lee E, Bondy CA, Dean J, Nelson LM. Mater, a maternal effect gene required for early embryonic development in mice. Nat Genet 2000; 26: 267-268.
7. Wu X, Viveiros MM, Eppig JJ, Bai Y, Fitzpatrick SL, Matzuk MM. Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for the oocyte-to-embryo transition. Nat Genet 2003; 33: 187-191.
8. Payer B, Saitou M, Barton SC, Thresher R, Dixon JP, Zahn D, Colledge WH, Carlton MB, Nakano T, Surani MA. Stella is a maternal effect gene required for normal early development in mice. Curr Biol 2003; 13: 2110-2117.
9. Burns KH, Viveiros MM, Ren Y, Wang P, DeMayo FJ, Frail DE, Eppig JJ, Matzuk MM. Roles of NPM2 in chromatin and nucleolar organization in oocytes and embryos. Science 2003; 300: 633-636.
10. Rajkovic A, Yan C, Yan W, Klysik M, Matzuk MM. Obox, a family of homeobox genes preferentially expressed in germ cells. Genomics 2002; 79: 711-717.
11. Paillisson A, Dade S, Callebaut I, Bontoux M, Dalbies-Tran R, Vaiman D, Monget P. Identification, characterization and metagenome analysis of oocyte-specific genes organized in clusters in the mouse genome. BMC Genomics 2005; 6: 76.
12. West MF, Verrotti AC, Salles FJ, Tsirka SE, Strickland S. Isolation and characterization of two novel, cytoplasmically polyadenylated, oocyte-specific, mouse maternal RNAs. Dev Biol 1996; 175: 132-141.
13. Mendez R, Richter JD. Translational control by CPEB: a means to the end. Nat Rev Mol Cell Biol 2001; 2: 521-529.
14. Richter JD. Cytoplasmic polyadenylation in development and beyond. Microbiol Mol Biol Rev 1999; 63: 446-456.
15. Gebauer F, Xu W, Cooper GM, Richter JD. Translational control by cytoplasmic polyadenylation of c-mos mRNA is necessary for oocyte maturation in the mouse. Embo J 1994; 13: 5712-5720.
16. Huarte J, Belin D, Vassalli A, Strickland S, Vassalli JD. Meiotic maturation of mouse oocytes triggers the translation and polyadenylation of dormant tissue-type plasminogen activator mRNA. Genes Dev 1987; 1: 1201-1211.
17. Tay J, Hodgman R, Richter JD. The control of cyclin B1 mRNA translation during mouse oocyte maturation. Dev Biol 2000; 221: 1-9.
18. Tanaka M, Hennebold JD, Macfarlane J, Adashi EY. A mammalian oocyte-specific linker histone gene H1oo: homology with the genes for the oocyte-specific cleavage stage histone (cs-H1) of sea urchin and the B4/H1M histone of the frog. Development 2001; 128: 655-664.

19. Oh B, Hwang S, McLaughlin J, Solter D, Knowles BB. Timely translation during the mouse oocyte-to-embryo transition. Development 2000; 127: 3795-3803.
20. Luoh SW, Bain PA, Polakiewicz RD, Goodheart ML, Gardner H, Jaenisch R, Page DC. Zfx mutation results in small animal size and reduced germ cell number in male and female mice. Development 1997; 124: 2275-2284.
21. Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM. NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. Science 2004; 305: 1157-1159.
22. Liang L, Soyal SM, Dean J. FIGalpha, a germ cell specific transcription factor involved in the coordinate expression of the zona pellucida genes. Development 1997; 124: 4939-4947.
23. Soyal SM, Amleh A, Dean J. FIGalpha, a germ cell-specific transcription factor required for ovarian follicle formation. Development 2000; 127: 4645-4654.
24. Gehring WJ. Homeo boxes in the study of development. Science 1987; 236: 1245-1252.

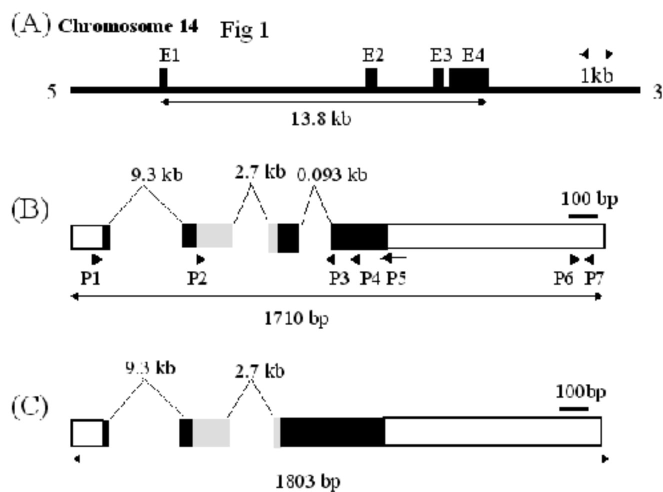

Fig 2

		Helix 1		Helix 2	Helix 3			% Homology
mESO-1	NSKPRHKFSR	DELKRLKQEF	AYAPYPDFTT	KDELARQFQC	EVSVIDNWFQ	NKRARLAPEL	(Mm)	100%
rESO-1	K-------TN	----------	KCT-------	R-------R-	Q- D-------	-----TPQ--	(Rn)	78%
CUX-1	LK ---VVLAP	E –KEA- -RAY	QQK---SPK-	I E---T- L NL	KT-TVI--- H	- Y- S- I R R- -	(Mm)	38%
BIX2	- RRR - TVY- P	SD–A- - E- Y-	Q I NM---IHQ	RE-----MGL	PE-R- Q V---	- R- SKARRQG	(Xl)	38%
CDP	L K---VVLAP	E –KE A- -RAY	QQK---SPK-	I E---T - LNL	KT-TVI --- H	-Y- S –I R R- -	(Rn)	38%
BSH9	QRRS–TT - -N	- Q I DA-ERI -	- RTQ---VY-	RE---QSTGL	TEARVQV- -S	- R----RKQ-	(Dm)	36%
OG9	HRRK–TT - -V	GQ- VE- ERV-	- AR----I S-	REH--QVIHL	PE AK- QV---	- R -- KRIKDR	(Mm)	35%

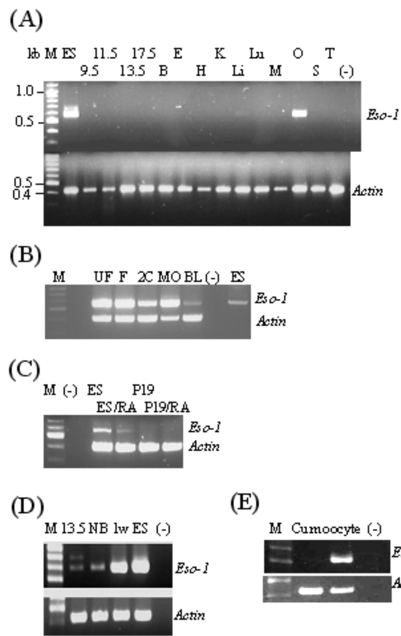

Fig 3

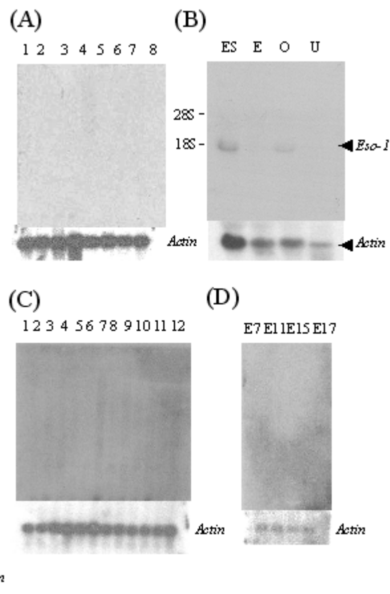

Fig 4