2015

Discovery of a novel oocyte-specific KRAb-containing zinc finger protein required for early embryogenesis in cattle

Jacqelyn M. Hand

DISCOVERY OF A NOVEL OOCYTE-SPECIFIC KRAB-CONTAINING ZINC FINGER PROTEIN REQUIRED FOR EARLY EMBRYOGENESIS IN CATTLE

JACQELYN M. HAND

Dissertation submitted
to the Davis College of Agriculture, Natural Resources & Design
at West Virginia University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Reproductive Physiology

Jianbo Yao, Ph.D., Chair
Robert Dailey, Ph.D.
Vagner Benedito, Ph.D.
Jorge Flores, Ph.D.
Melanie Clemmer, Ph.D., HCLD
Michael Vernon, Ph.D., HCLD

Division of Animal and Nutritional Sciences
Morgantown, West Virginia

2015

ABSTRACT

Discovery of a Novel Oocyte-Specific KRAB-Containing Zinc Finger Protein Required for Early Embryogenesis in Cattle

Jacqelyn M. Hand

Much of the loss of potential offspring in cattle is concentrated in the early embryonic period. Maternal mRNAs that accumulate in the oocyte during oogenesis have important functional roles during the initial stages of embryonic development, before embryonic genome activation. It is well regarded that the oocyte plays an active role in regulation of key aspects of the reproductive process required for fertility. What is more, oocyte-specific transcription factors seem to be the controlling feature influencing germ cell success throughout oogenesis, fertilization, and early embryonic development. Of the remarkably diverse array of transcription factors encoded by mammalian genomes, about two-thirds encode C2H2 zinc-finger proteins.

Zinc finger proteins exclusively expressed in mammalian oocytes have not been reported. Deep sequencing of a bovine oocyte library revealed a highly abundant transcript that matches an uncharacterized gene in the NCBI database. cDNA cloning of the novel *ZNFO* gene revealed a transcript containing a 2,145 bp open reading frame that codes for a protein of 714 amino acids with a conserved KRAB domain at the N-terminus and nine zinc finger motifs at the C-terminus. The individual ZNF motifs fit the conserved two cysteine-two histidine sequence model. *ZNFO* mRNA was detectable in fetal ovaries and was undetectable in all somatic tissues analyzed, including granulosa and theca cells. Real-time PCR analysis revealed *ZNFO* mRNA is highly abundant in GV and MII stage oocytes as well as in pronuclear to 8-cell stage embryos but undetectable in blastocyst stage embryos. Immunohistochemical analysis detected ZNFO protein in oocytes throughout folliculogenesis. Identification and characterization revealed the novel *ZNFO* is a KRAB-containing maternal-effect gene found exclusively in bovine oocytes.

To elucidate the functional role of *ZNFO*, zygotes were generated by *in vitro* maturation and fertilization of oocytes and injected with small interfering RNA (siRNA) designed to knockdown *ZNFO*. Cleavage rates were not affected by ZNFO siRNA injection. However, embryonic development to 8- to 16-cell stage and blastocyst stage was reduced significantly relative to the uninjected and negative control siRNA-injected embryos. Furthermore, interaction of ZNFO with the highly conserved transcriptional repressor co-factor (KAP1) was demonstrated by GST pull-down, and evidence supporting transcriptional repression by ZNFO using a GAL4-luciferase assay. In addition, transfection studies verified that a ZNFO-GFP fusion protein localizes specifically to the nucleus, further supporting the proposed function in transcriptional regulation. These studies demonstrate that ZNFO is a maternally-derived oocyte-specific factor required for early embryonic development in cattle, which has a functional role as a transcriptional regulator required during early embryogenesis by repressing transcription, possibly controlling activation of the embryonic genome.

DEDICATION

To the one I have always admired: my father.

ACKNOWLEDGEMENTS

Foremost, I would like to acknowledge Dr. Yao. He has been the single most influential person to me over the past several years, both as my teacher and role model. He has continually provided me his wisdom, encouragement, humor, teachings, and support. I never would have imagined how fortunate I would be to have such a wonderful mentor.

I am sincerely grateful to my committee members Dr. Robert Dailey, Dr. Vagner Benedito, Dr. Melanie Clemmer, Dr. Michael Vernon, and Dr. Jorge Flores for serving on my graduate committee and for each providing me their own teachings and support throughout my graduate career. I am also grateful to Dr. E. Keith Inskeep and Dr. Scott Bowdridge for inspiring me to be passionate about science.

My deepest love and gratitude goes to my lab mates Dr. Lei Wang, Liyuan Fu, Prasanthi Koganti, Jian Wang, Xiaowei Zhang, Chieh-Hung Lin, Dr. Hao Ma, Nan Qi, Dr. Liu, and Swamy Tripurani for their unwavering support and dear friendship. Also, my undergraduates Kristen Mastrantoni and Ishan Gandhi have my genuine appreciation as well. I owe part of my success to each of you. This "Lab Family" of ours has been absolutely wonderful and will forever have a place in my heart.

I wish to thank Dr. George Smith and Dr. Kun Zhang at Michigan State University for their generous contribution to my research, as well their willingness to discuss my work and provide helpful advice.

My fellow graduate students, that became good friends, are to thank. Having a group of friends with similar interests and who are genuine people is a precious thing. I am thankful for all their emotional support, laughs, and entertainment. Special thanks to Dr. Lizzie Bowdridge.

Lastly, and ever sincerely, I wish to thank my family. My parents and sisters encouraged me with their absolute love and support, Jeremiah Nemechek was the rock that kept me grounded, and the Nemechek family provided me with continual encouragement. People come and go, and distance is difficult, but the love of a family (whether genetic or acquired) is a stronghold that lasts forever.

TABLE OF CONTENTS

CHAPTER 1: Identification and Characterization of ZNFO, a Novel Bovine Oocyte-Specific Gene Encoding a KRAB-Containing Zinc Finger Transcription Factor **48**

CHAPTER 2: Functional Analysis of the Novel Bovine ZNFO Transcription Factor Reveals it is Required for Normal Early Embryonic Development in Cattle... **80**

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BMP	Bone Morphogenetic Protein	FLOPED	Factor Located in Oocytes Permitting Embryo Development
bp	base pairs		
BTB/POZ	Broad-Complex, Tramtrack, and Bric-a-brac/Poxvirus and Zinc Finger	FSH	Follicle Stimulating Hormone
		GC	Granulosal Cell
C2H2	Cysteine2-Histidine2	GDF9	Growth Differentiation Factor 9
CASTing	Cyclic Amplification and Selection of Targets	GFP	Green Fluorescent Protein
		GO	Gene Ontology
cDNA	complementary DNA	GST	Glutathione S-Transferase
ChIP	Chromatin Immunoprecipitation	GV	Germinal Vesicle
CIBZ	CtBp-interacting BTB ZNF	GVBD	Germinal Vesicle Breakdown
CPEB1	Cytoplasmic Polyadenylation Element Binding Protein 1	H3.3	H3 Histone, Family 3A and 3B
		HP1	Heterochromatin Protein 1
DAPI	2-(4-amidinophenyl)-1H - indole-6-carboxamidine	ICM	Inner Cell Mass
		ICR	Imprint Control Region
DAZL	Deleted in Azoospermia-Like Autosomal	IGF	Insulin-like Growth Factor
		KAP1	KRAB-Associated Protein 1
DPPA3	Developmental Pluripotency-Associated 3	KITL	Kit Ligand
DNMT	DNA Methyltransferase 1	KRAB	Krüpple-Associated Box Domain
Dox	Doxycycline		
E	Embryonic day	KRIP1	KRAB-A-Interacting Protein 1
E2	Estradiol	LH	Lutenizing Hormone
EDTA	Ethylenediaminetetraacetic Acid	LHX8	LIM Homeodomain Transcription Factor 8
EGA	Embryonic Genome Activation	MII	Metaphase II
EGF	Epidermal Growth Factor	MATER	Maternal Antigen that Embryos Require
ERV	Endogenous Retrovirus		
ESC	Embryonic Stem Cell	MGA	Mid-Preimplantation Gene Activation
EST	Expressed Sequence Tag		
F	Forward	mRNA	messenger RNA
FGF	Fibroblast Growth Factor	MZT	Maternal-to-Zygotic Transition
FIGLA	Factor in the Germline Alpha		

NCBI	National Center for Biotechnology Information	SOHLH1/2	Spermatogenesis- and Oogenesis-Specific Basic Helix-Loop-Helix 1 and 2
NLS	Nuclear Localization Signal	SOX2	SRY (Sex Determining Region Y)-Box2
NOBOX	Newborn Ovary Homeobox		
NPM2	Nucleoplasmin 2	SRY	Sex Determining Region Y
NR	Nuclear Receptor Box	TET3	Tet Methylcytosine Dioxygenase 3
nt	nucleotide		
NuRD	Nucleosome Remodeling and Histone Deacetylation Complex	TetO	Tetracycline Operator sequence
		TetR	Tetracycline Repressor sequence
OBOX	Oocyte-Specific Homeobox		
OCT4	Octamer-Binding Transcription Factor 4	TF	Transcription Factor
		TFIIIA	Transcription Factor IIIA
ORF	Open Reading Frame	TGF-β	Transforming Growth Factor- β
P4	Progesterone	TIF1	Transcription Intermediary Factor 1
PBS	Phosphate Buffered Saline		
PCR	Polymerase Chain Reaction	TRIM28	Tripartite Motif-Containing Protein 28
PGC	Primordial Germ Cell		
PHD	Plant Homeodomain	TSS	TIF1 Signature Sequence
Poly-ZNF	Polydactyl-ZNF	TSS	Transcription Start Site
PRC2	Polycomb Repressive Complex	UTR	Untranslated Region
PRDM	PR-Domain Containing ZNF	YBX2	Y box Protein 2
R	Reverse	ZAR1	Zygote Arrest 1
RBCC	RING Finger, B-box Zinc Fingers, and Coiled-Coil Domain	ZFP	Zinc Finger Protein
		ZNF	Zinc Finger Transcription Factor
RING	Really Interesting New Gene	ZNFO	Novel bovine oocyte-specific zinc finger transcription factor
RNAi	RNA Interference		
RPL19	Ribosomal Protein-Like 19		
SCAN	SRE-ZP, CTfin51, AW-1 and Number 18 cDNA		
SETDB1	SET Domain, Bifurcated 1		
siRNA	short interfering RNA		

REVIEW OF LITERATURE

INTRODUCTION

The ability of the bovine embryo to reach the blastocyst stage and ultimately develop into a healthy offspring is a complex and highly regulated process. The intrinsic quality of the oocyte must first be highly regulated and includes the capacity of the oocyte to resume meiosis, cleave following fertilization, develop to the blastocyst stage and implant, and develop to term in good health [1-4]. Zinc finger (ZNF) transcription factors are known to interact with DNA through zinc finger motifs and play important roles in a variety of cellular functions, including cell growth, proliferation, development, apoptosis, and intracellular signal transduction. Limited data exist on this form of transcriptional regulation during oogenesis and early embryonic development. The KRAB-ZNF-KAP1 repression system is one of the best-characterized systems for gene-specific silencing of euchromatin in mammals. Moreover, the abundance of the KRAB domain zinc-finger proteins in the mammalian proteome [5,6] and the diverse array of DNA sequences that they recognize [7,8] potentially make this family of gene-specific silencers a master regulator of gene silencing during cellular differentiation and organism development.

Despite their vast occurrence, KRAB-ZFPs and their transcriptional targets are remarkably obscure. The substantial prevalence and diverse functional roles of zinc finger transcription factors may actually be quite overwhelming. This chapter, however, focuses first on highlighting basic knowledge on the formation of germ cells and subsequent oogenesis and folliculogenesis, followed by a summary of maternally regulated factors and events involved in the maternal-to-zygotic transition (MZT). The second part of this chapter is a review of zinc finger transcription factors with major emphasis placed on the C2H2 class of KRAB-containing ZNFs and their functional roles during germ cell- and early developmental-events. This background knowledge should stimulate thoughts on how KRAB-ZNFs may regulate the processes in the oocyte and early embryogenesis in cattle.

GERM CELL FORMATION AND OOGENESIS

1. Germ Cell Formation

     In mammals, germ cell lineage is not "pre-determined" but rather follows epigenic mechanisms in which external signals from the surrounding somatic cell line induce germ cell development[9,10]. In the other underlying mechanism of germ cell development, "pre-formation", the germ cell lineage is set aside from the somatic lineage very early in development. In many species, including *Drosophila melanogaster*, *Danio rerio*, and *Caenorhabditis elegans*, pre-formed germ line determinants are absorbed into cytoplasmic "germ plasm" of fertilized ova at the time of cell division and subsequently develop into germ cells[9-11]. Mammalian germ cells are generated in extraembryonic tissues during early development and migrate to the future gonads during cell proliferation. These germ cells are known as primordial germ cells (PGCs) and have the potential to differentiate into either sperm or ova. The remarkably unique feature of this process is that upon entering the gonads the germ cells remain bipotential despite the chromosomal content (XX or XY) until sex determination occurs under the influence of the surrounding somatic cells.

Generation of Primordial Germ Cells

     Germ cell fate is induced in pluripotent epiblast cells in response to signals from extraembryonic tissues, such as the extraembryonic ectoderm and visceral endoderm. Specification of PGCs involves the integration of three main events: a) repression of the somatic program, b) reacquisition of pluripotency, and c) genome-wide epigenetic reprogramming. An important point to note regarding specification of the germ cell lineage is that no cells of the epiblast are predetermined as PGC precursors. The epiblast cells are not irreversibly allocated to a specific cell lineage but rather can be induced to give rise to PGCs in a site-specific manner[12]. Therefore, extrinsic signals from surrounding somatic cells are required for the generation of the germ cell lineage[13].

     Primordial germ cell specification, around E5.5 in mice, begins with bone morphogenetic protein (BMP) signals that originate from the extraembryonic ectoderm (BMP4 and BMP8B)[14] and visceral endoderm (BMP2)[15]. The BMP signals act upon the adjacent proximal epiblast to

establish a region permissive for the formation of PGC precursors[9]. These cells are not considered founder PGCs because not all of their descendants will go on to become gametes. Only a subset of cells will continue on to express the transcriptional repressor proteins PR-domain containing 14 (PRDM1; also known as BLIMP1) and PRDM14. Both are PR domain-containing proteins that are considered to be key determinants of PGC specification[16-19]. PRDM1 is a zinc finger-containing DNA-binding transcriptional repressor that has been shown to be critical for repression of somatic genes, likely by shutting off the default pathway that allows epiblast cells to assume a somatic cell fate, and thereby induces the germ cell lineage[16,18]. Also, PRDM1 is required for reactivation of potential pluripotency and control of epigenetic reprogramming[20]. Similar to PRDM1, PRDM14 is required for specification, however, it does not seem to be critical for suppression of the somatic lineage[19].

Re-acquisition of pluripotency is marked around E6.25 in mice when the pluripotency-associated gene products *Pou5f1*, *Nanog*, and *Sox2* are detected in PGC precursors[21]. Around E7.25 in mice the approximately first 40 founder PGCs are established. Then, at approximately E8.0, rapid erasure of DNA methylation and reduced H3K9 dimethylation in germ cells preceded by the transient loss of DNA methyltransferases DNMT1, DNMT3A, and DNMT3B, all of which are major repressive modifications for gene expression[22]. About E8.75, germ cells substantially increase levels of repressive H3K27 trimethylation, a modification typically regarded with pluripotent cell lineages and, therefore, may be associated with reacquisition of underlying totipotency [22].

Migration and Colonization of the Gonads by PGCs

By E7.5-8.5 in mice, the PGCs are proliferating rapidly and begin to migrate as a group into the developing hindgut. Initiation of this mesoderm-to-endoderm migration still remains unknown. From about E10.5, PGCs undergo a directed migration away from the hindgut and move into the genital ridges where they undergo massive proliferation. The growth factor, kit ligand (KITL), is considered a required factor for the survival and proliferation of PGCs but is also required for PGC migration[23]. As PGCs migrate out of the hindgut, they express the adhesion molecules E-cadherin and β1-integrin that are necessary for their colonization of the genital ridges

through germ cell-germ cell interactions and germ cell-extracellular matrix interaction, respectively[24,25].

The initially bipotential genital ridges continue development along the medial surface of each of the mesonephroi until around E10.5 where the presence or absence of the Y-linked gene, *Sry*, determines organ fate. Under the influence of sex determining region-Y (SRY), XY genital ridges begin to develop as testes and, in the absence of SRY, XX genital ridges begin to develop as ovaries. If SRY is not expressed early enough or at high enough levels, it is unable to overcome the default pathway of ovarian development[26]. It is into this dynamic environment of a bipotential primordial organ reorganizing as either a testis or an ovary that the PGCs arrive. The bipotential PGCs (soon to be either oogonia or prospermatogonia) continue to proliferate by mitotis divide for 2-3 days upon entering the genital ridges. Differences between the somatic components of the ovaries and testes are evident by E12.5; however, the germ cells remain undistinguishable as either sex until E13.5.

At E12.5-13.5 in mice, commitment to either the male or female program of development is established. By committing to a female program, germ cells will migrate in a developing ovary, cease mitosis and enter prophase of the first meiotic division and thus become oocytes. They progress through the different stages of meiotic prophase I (leptotene, zygotene, pachytene, and diplotene) until they reach dictyate arrest around the time of birth [27,28] marking the transition from oogenesis to folliculogenesis.

2. Oogonial Maturation / Oogenesis

Primordial & Primary Follicle Development

During embryonic development, primordial germ cells are enclosed by presumptive follicular cells, although there is no evidence of surrounding granulosal cells, and are called oogonia, which develop in nests (also referred to as clusters or cysts). Assembly of these primordial oogonia into follicles (known as primordial follicle formation) requires a transition from the nests into primordial follicles. At birth in rodents, follicle formation occurs coincidently with a synchronous, initial wave of follicle activation and growth[29,30], and the oocytes become surrounded

by a single layer of flat squamous pre-granulosa cells. These are primordial follicles. The number of germ cell clusters then declines very rapidly after birth[31]. In contrast to rodents, most domestic species and primates (including human), follicle formation is initiated during fetal life and emerges in an asynchronous pattern over a relatively long period of time[32,33]. Over the course of this time, a portion of primordial follicles are activated and begin to develop, and, at the same time, other follicles are just beginning to be formed.

In fetal calf ovaries, primordial, primary (single layer of cuboidal granulosal cells (GCs)), and secondary (two to four layers of cuboidal GCs) follicles first appear at approximately days 90, 140, and 210, respectively, of gestation[34]. In the adult ovary, primordial follicles are located just underneath the epithelial surface in the periphery. The primordial follicle pool exists in a quiescent, naturally arrested state at prophase I of meiosis until stimulated to grow, in which case some follicles leave the arrested pool and undergo the primordial to primary follicle transition[35]. This transition is characterized by a change in GCs from squamous to proliferative cuboidal-shaped morphology and an increase in oocyte diameter[36-38]; however, the follicle itself is not yet proliferative nor does it undergo mitosis. The bovine oocyte and follicle continue to grow in parallel until the follicle reaches a diameter of ~ 3 mm, during which the oocyte increased in size from < 30 μm in the primordial follicle to > 120 μm in the tertiary (> 5 layers of GCs; formation of an antrum) follicle[35,39]. Thereafter, the growth of the oocyte plateaus at about 120-130 μm, while the follicle grows up to 15-20 mm in diameter before ovulation. During this growth phase, a series of modifications are induced that are necessary for the acquisition of meiotic and developmental competence[35,39]. Even prior to the series of complex events encompassing early embryonic development, the intrinsic quality of the oocyte must first be sufficient for fertilization.

Endocrine Factors, Growth Factors, and Communication Networks

As the basic functional unit of the ovary, the follicle is composed of somatic cells, granulosal and thecal cells, and the developing oocyte, both of which work together to control and maintain gametogenesis. Folliculogenesis is the process of growing a follicle through a series of highly regulated, sequential steps that result in either ovulation of a developed oocyte or death of the follicle (atresia). Granulosal and thecal cells are the site of action and synthesis of several

6

hormones that promote regulation of follicular development. As the follicle grows, cytodifferentiation requires the attention of various hormones and growth factors. The delicate interaction between the somatic cells and the developing oocyte is controlled through several endocrine factors: the gonadotropins [follicle stimulating hormone (FSH) and luteinizing hormone (LH)[40]], autocrine and paracrine factors (transforming growth factor-β (TGF-β) family members[41] including inhibins/activins[42]), the insulin-like growth factor (IGF) system[43], epidermal growth factor (EGF)[44], and gap junctional communication (connexins)[45]. Proliferation of these cell types is largely responsible for the development of the antral ovarian follicle.

Granulosa cells also regulate the biosynthesis of two critical steroids, estradiol (E2) and progesterone (P4), in primate, domestic, and rodent species. As the growing follicle develops, GCs differentiate and increase E2 production, as FSH promotes follicular development. Prior to reaching ovulation, the GCs develop an increased capacity to synthesize and secrete P4 under the control of LH. Thecal cells, which are separated from the outermost layer of neighboring mural GCs by a basement membrane, play a major role in androgen secretion. Thecal cells respond to LH by increasing the production of testosterone following expression of LH receptors around the tertiary follicle stage. Like GCs, thecal cells produce progestins under gonadotropin control. In contrast to secondary, preantral and antral follicles, the primordial follicle stage GCs are gonadotropin hormone-independent and are non-steroidogenic. Further, at the primordial stage, no theca cells are present; albeit, during the transition to the primary stage, thecal cells are recruited to the follicle as precursor gonadotropin-independent non-steroidogenic cells. Shortly following primary follicle assembly theca cells are recruited from the stromal-interstitial cell population resulting in a subsequent proliferation of both granulosal and thecal cells[36].

Meiotic Maturation

Oocytes gradually and sequentially acquire competence throughout folliculogenesis by synthesizing and accumulating transcripts and proteins that are necessary for successful meiotic competence during follicular development, fertilization and subsequent embryogenesis[4,46]. Oocyte meiotic maturation is a complex phenomenon, which occurs from the diplotene stage of meiotic prophase I through metaphase II (MII). On reaching the diplotene stage, which occurs around day

7

170-post conception in cattle[47], the oocyte becomes developmentally arrested. The oocyte at this stage is characterized by a single surrounding layer of pre-granulosa cells and an intact basal lamina forming the resting primordial follicle. The primordial follicle-enclosed oocytes make up a finite population of stored of oocytes, which remain non-growing and quiescent until stimulated to grow[29]. The capacity of a primordial follicle to activate and develop to the primary stage follicle is correlated with achievement of meiotic arrest in the oocyte[34], which is maintained until the follicular development of a pre-ovulatory follicle. In contrast, germ cells in a developing testis do not enter meiosis during fetal life but instead arrest in G_0/G_1 phase of the cell cycle[48].

During the pre-antral to antral transition, however, the oocyte acquires the capacity to resume meiosis[49] in response to the ovulatory LH surge and within hours germinal vesicle breakdown (GVBD) occurs. Meiotic competence is associated with the accumulation of cell cycle regulatory factors, reorganization of chromatin and microtubule configurations, as well as expulsion of the first polar body[50]. Granulosa cells in the preovulatory follicle change gene expression patterns in response to the LH surge and indirectly stimulate oocyte meiotic maturation and ovulation of a metaphase II-stage oocyte that is competent to undergo fertilization, and now therefore considered mature[50]. In most mammalian species, oocytes remain arrested at the MII stage until activated by fertilization and subsequent formation of the pronucleus (the nucleus of the egg (or sperm) prior to their fusion). The ability of the oocyte to complete meiosis is known as meiotic competence and, as described, this process is acquired gradually during growth of the follicle.

MATERNAL CONTROL OF EARLY EMBRYOGENESIS AND EMBRYONIC GENOME ACTIVATION

Early embryonic development is considered one of the most critical periods in mammalian development and comprises several important transitions including replacement of maternal RNAs with zygotic RNAs, compaction, the first lineage differentiation into the inner cell mass and trophectoderm, and, finally, implantation. Various physiological processes and biosynthetic changes regarding genomic activity take place during this early time. Among these events is the first important developmental transition that occurs following fertilization at which time the embryo switches from using transcripts derived from the maternal genome to those resulting from embryonic genome activation (EGA)[51]. Mammalian oocytes harbor a vast collection of mRNA and proteins throughout oogenesis that orchestrate subsequent embryonic development. During oocyte meiotic maturation and the early stages of embryonic development, the transcriptional machinery for this collection of molecules is silent; therefore, any events that sustain the embryonic genome prior to the onset of EGA are regulated by the translation of pre-existing maternal transcripts[52]. As development progresses, control is switched from maternal- to embryonic-derived transcription and accumulation of proteins and is referred to as the maternal-to-embryonic transition (MET). The characteristic events of MET begin at fertilization and include depletion of maternal transcripts and protein by degradation, dramatic reprogramming of both male and female genomes from a repressed chromatin state to one that is open for transcription, replacement of maternal transcripts stored in oocytes by embryonic transcripts, and finally the robust activation of the embryonic genome[53,54].

1. Maternal-Derived Factors Controlling Early Embryogenesis

Maternal factors; such as subcellular organelles, macromolecules, and maternal-effect genes; are stored in the oocyte and have accumulated throughout oogenesis. The MET is dependent on these factors for successful outcome of early embryonic development. The meiotic spindle, for example, is responsible for proper segregation of chromosomes during cell division. Therefore, defects to the spindle may lead to errors in chromosome segregation and generate aneuploidic embryos[55].

Maternal-effect genes mediate their effects through deposition of cytoplasmic transcripts or protein products during oogenesis in the female germline. These accumulated products then exert their effects on the fertilized zygotes until major activation of the embryonic genome. Needless to say, maternal-effect genes are critical for early embryonic development. The number of maternal-effect genes found to be essential for early embryogenesis is continuing to grow. Various technical strategies (such as microarray, gene-knockout, and RNA interference (RNAi)) have been used to identify products of maternal-effect genes that are stored during oogenesis and used to regulate MET. Recall that maternal factors have several prominent roles during MET[53,54], including removal of maternal RNA and protein, reprogramming of male and female genomes, and embryonic genome activation.

Oocyte-Specific Transcriptional and Post-Transcriptional Regulators

Oocytes gradually and sequentially acquire competence throughout folliculogenesis by synthesizing and accumulating transcripts and proteins that are necessary for successful follicular development, fertilization and subsequent embryogenesis[46]. These early stages are critical because many oocyte-specific genes are transcribed during the primordial to primary follicle transition and continue to be expressed throughout folliculogenesis. As identified by molecular genomic and gene knockdown studies, several oocyte/germ-specific transcription factors such as *Nobox* (Newborn ovary homeobox)[56], *Figla* (Factor in the germline alpha)[57], *Obox* (Oocyte-specific homeobox), *Sohlh1/2* (spermatogenesis and oogenesis specific basic helix-loop-helix 1 and 2)[58,59] and *Lhx8* (LIM homeodomain transcription factor 8)[60], RNA binding proteins such as *Ybx2* (Y box protein 2)[61], *Dazl* (deleted in azoospermia-like autosomal)[62] and *Cpeb1* (cytoplasmic polyadenylation element binding protein 1)[63], and growth factors *Gdf9* (Growth differentiation factor 9)[64] and *Bmp15* (bone morphogenetic protein 15)[65], which are found throughout folliculogenesis, maintain normal development of germ cells and surrounding somatic cells essential for mammalian folliculogenesis[66-69].

Several factors have been found to play key roles within the MET developmental transitions in mice: transcription factors (*Hsf1*, *Basonuclin*, and *Ctcf*) chromatin remodeling factors (*Ring1*, *Npm2*, *Trim24*, and *Brwd1*), DNA methylation machinery (*Dntm1*, *Dppa3*, and *Zfp57*),

genes involved in degradation of maternal factors (*Dicer1*, *Ago2*, and *Atg5*), pluripotency factors (*Oct4* and *Sox2*), and genes involved in the preimplantation development (*Mater*, *Zar1*, *Floped*, *Dppa3, Oct4, Npm2, Padi6,* and *Filia*). These factors highlight the functional contribution of oocyte-derived transcriptional and post-transcription regulators to early embryogenesis in mice; however, inherent species-specific differences exist. The duration and number of cell cycles to accomplish embryonic genome activation and complete MET likely varies between mice and cattle in the specificity of regulatory mechanisms and genes that control this transition[70]. Comparative genomics and functional studies in the bovine model system have uncovered the existence of a few novel bovine oocyte-specific genes that are required for early embryogenesis such as *KPNA7*[71], *Follistatin*[72], and *JY-1*[73], and the known *NOBOX*[74].

2. Post-Fertilization and Embryonic Genome Activation

Primordial germ cells (PGCs) during gametogenesis establish a set of highly regulated epigenetic marks. These marks are sex-specific and display distinct global and specific DNA methylation patterns[75,76]. As gametes mature, the haploid male and haploid female genomes become transcriptionally quiescent. Fusion of the oocyte and spermatozoon at fertilization signify one of the greatest biological events in which these highly differentiated germ cells are reprogrammed to the totipotent 1-cell zygote status. As both genomes undergo dynamic changes during MET, they are reset to support a variety of embryonic developmental events including maintenance of epigenetic modifications, depletion of maternal mRNA transcripts, activation of the newly formed embryonic genome, and cell specification.

Epigenetic Reprogramming

Chromatin organization between the maternal and paternal pronuclei shortly after fertilization is strikingly different. An evident asymmetry of DNA demethylation and histone modification patterns are signatures of this time[77]. Active DNA demethylation occurs in the male pronucleus prior to the onset of DNA replication. Protamines, that have tightly packaged the haploid sperm DNA during this wide-spread paternal demethylation, are repackaged with hyperacetylated maternal histones to form the male pronucleus[78,79]. After syngamy of the two

11

pronuclei, the zygotic genome undergoes passive demethylation until the morula stage[80]. This occurs in a DNA replication dependent fashion. Thereafter, *de novo* methylation arrangements are established to sustain successful cell lineage differentiation[78,79]. Epigenetic reprogramming, as such, is believed to resolve the discrepancy of maternal and paternal chromatin and ensure the successful transition from differentiated to totipotent zygote.

Because of the period of transcriptional quiescence in early embryos before embryonic genome activation, maternal proteins stored during oogenesis are likely required for epigenetic reprogramming in early embryos. Several maternal proteins have been described as required factors for epigenetic reprogramming including Tet methylcytosine dioxygenase 3 (Tet3) for active DNA demethylation, developmental pluripotency-associated 3 (DPPA3) for maintenance of DNA methylation, and H3.3 (H3 histone, family 3A and 3B) for reprogramming and decondensation of chromatin[53,54]. Such nuclear reprogramming is a requirement to activate the transcriptionally inactive embryonic genome.

Embryonic Genome Activation

Initially, the maternal genome regulates nearly all aspects of early development. Basic biochemical processes that implement early development events such as meiotic maturation, fertilization, the first cleavage divisions, and programming the EGA are totally dependent on the maternal mRNAs and proteins derived within the oocyte. The preservation of maternal transcripts from a period that is transcriptionally silent to one that is reliant on the resulting maternal proteins dictates the ability of the zygotic genome to activate and reprogram and sustain its own transcriptional products[81]. At the onset of EGA initiation, the destruction of maternal mRNAs begins by maternally encoded products [52,82]. Embryonic genome activation then initiates zygotic transcriptional activity within the embryonic nucleus leading to the synthesis of new mRNA and proteins[52]. In other words, initiation of gene expression and regulation of the embryonic genome becomes largely controlled by products of the embryo. Adjustments in chromatin structure can control the activity of transcription factors by permitting or restricting their access to regulatory elements of the genome but are not sufficient to activate transcription. The oocyte cytoplasm also holds an important role in transcription activation by providing active transcription factors and

RNA polymerase II. Oocyte-stored products are largely responsible for regulating this process by altering the chromatin structure to a state in which the transcription start site (TSS) of target genes is uncovered[83].

Although the primary mechanisms controlling the onset of EGA are considered to be generally conserved in mammals, differences in timing confirm it is a distinct species-specific event that takes place during the first few cell cycles post-fertilization around 2-cell stage in mice, 4- to 8-cell in human, rat and pig embryos, and 8- to 16-cell stage in sheep and bovine embryos[84-86]. Bovine *in vitro* culture systems demonstrate that a developmental block arises around the 8-cell stage in embryos treated with α-amanitin, an inhibitor of transcription[87]. Namely, progression to the 8-cell stage is transcription-independent as embryos can develop to this stage in the presence of α-amanitin, indicating a requirement for embryonic transcripts for further development past this stage. Thus, EGA is considered to be the most critical event for early developmental viability. Preceded by a minor genome transcription, EGA occurs gradually, followed by a recently proposed third successive overlapping wave of gene expression termed "mid-preimplantation gene activation" (MGA), which may play a critical role in cell polarity and the first cell lineage specification[88]. In bovine embryos, minor EGA occurs as early as the 2-cell stage[85], a transition that is necessary for genome reprogramming and acquisition of totipotency by the embryo. Therefore, proper EGA is critical for normal development to commence[81].

Due to the apparent importance of proper regulation of EGA, many studies have focused on mechanisms of EGA including gene expression profiles during the maternal-to-zygotic transition[81]. In a RNA-sequencing study of bovine oocytes and embryos, embryonic transcripts not present in oocytes were analyzed for gene ontology (GO) terms and compared at different pre-implantation stages[89]. Genes activated at the 4-cell stage or before were functionally classified in RNA processing, translation, and transport as the first transcriptional activity before the major EGA. Classification of the 8- to 16-cell stage activated genes revealed primary functional roles in the initiation of transcription and translation, as well as with the continuous degradation of maternally stored RNAs and proteins, which fits well the known idea of major EGA events[89].

At the 16-cell stage, the functions of the activated genes become diversified and targeted. Regulation of glycolysis, RNA splicing, ATP biosynthetic process, negative regulation of transcription, and transcription initiation/elongation from RNA polymerase II promoter are examples implicating the maintenance of transcription and translation and the initiation of metabolic processes. By the morula to blastocyst stages, the majority of genes activated clearly were involved in the first differentiation processes and the ongoing transcription and translation required for cell specification[89].

Tight coordination of many factors acting at several regulatory levels control the diverse range of genes expressed, both spatially and temporally, during MET and, more specifically, EGA. Each of these levels brings its own finely tuned skills, whether it be a *cis*-acting DNA sequence motif, a DNA methylation state, a set of post-translationally modified histones, or binding of a transcription factor. All of these elements and many others communicate and work together to configure the genome for the complex events of the maternal-to-embryonic transition.

ZINC FINGER TRANSCRIPTION FACTORS

1. Transcription of Eukaryotic Protein-Coding Genes and Zinc Finger Transcription Factors

Transcription Factor Repertoire of Eukaryotes

Cellular life depends on the ability to recognize and respond to an array of diverse external and internal signals and successively carry out the appropriate molecular-level function in response. The transcriptional regulatory system is an exceptionally complex process that ensures the correct expression of specific genes. As master regulatory elements, transcription factors (TFs) often are identified as controllers of many biological processes ranging from cell cycle progression and maintenance of intracellular metabolism to cellular differentiation and development. Initially thought as a relatively simple explanation for gene expression, cell differentiation, and homeostasis, updated genomic analyses have uncovered that TF mechanisms in fact, are, quite complicated and varied.

Early molecular genetic and biological investigations led to the basic knowledge of transcriptional control; many diverse proteins working to initiate successful transcription by RNA polymerase. General transcription factors and co-factors regulate the assembly of transcription-initiation complexes and the rate at which transcription is initiated, while a variety of enzymes modulate chromatin structure via changes in DNA methylation and histone modifications. Also, a multitude of sequence-specific DNA-binding TFs direct transcription initiation to specific promoters[90]. Researchers over the past 15 years have provided further complementary information regarding the function and organization of TFs since the sequencing of species complete genomes and development of whole-genome high-throughput technologies have evolved[6]. For example, analysis of the human genome predicted approximately 2,300 genes coding for the basic transcriptional machinery (nucleic acid enzymes and roughly 1,800 sequence-specific DNA-binding TFs)[5]. Most human TFs are unannotated; however, GO findings suggest the most highly represented regulatory functions include control of developmental processes and cellular processes[6].

Classification of TFs commonly is organized based on the structure of their DNA-binding domains[91]. Arranging TFs in such a system has been useful for simplifying comparisons of different modes of TF recognition, binding specificities to DNA sequences, and insights into their evolutionary histories. Three types of TF families dominate over 80% of the TF repertoire in human and mice genomes: the C2H2 zinc finger, homeodomain, and helix-loop-helix[6,92].

Zinc Finger Transcription Factors

The term "zinc finger" was first used as laboratory jargon after the discovery of a remarkable 30-residue, repeated sequence motif found in an unexpectedly abundant *Xenopus laevis* transcription factor, because it folded around a zinc ion to form a mini-domain that grasped the DNA[93]. Before current expression technology, studying TFs was challenging because TFs are present in small quantities in the cell. Fortunately, eukaryotic transcription factor IIIA (TFIIIA) turned out to be expressed at a very high level in *Xenopus* immature oocytes as a storage particle complexed with 5S RNA[94].

Zinc finger transcription factors (ZNFs) come in a variety of structural classifications based on their ligand geometry and the ligand structural properties around the zinc ion-binding site. The spatial arrangement of secondary structural elements that contribute zinc ion coordination constitute the classification of the fold groups and, therefore, share common structural features and even functionality but are not necessarily related[95]. Famously, C2H2-like ZNFs are highly abundant and the most studied. All structurally similar in that they form finger-like protrusions and are supported by a zinc ion, groups such as C2H2-like, the Gag knuckle, treble clef finger, and zinc ribbon vary by their zinc ion-binding ligands (i.e. CCHH for C2H2-like ZNFs vs. CCHC for Gag knuckle ZNFs), ligand placement within the structure, and secondary structure organization [α-helices, β-sheets, turns (zinc knuckle), and loops][95].

2. The C2H2 Class of Zinc Finger Proteins

The two cysteine and two histidine residue (C2H2) zinc finger proteins represent the largest family of transcription factors encoded by the human genome[6] and appear to be the dominant family of regulatory proteins in all mammalian genomes. C2H2 zinc finger proteins contain from

1 to more than 30 individual zinc finger motifs and are defined by the presence of the consensus sequence Φ-X-Cys-X$_{(2-4)}$-Cys-X$_3$-Φ-X$_5$-Φ-X$_2$-His-X$_{(3-4)}$-His, where X represents any amino acid and Φ represents a hydrophobic residue[96]. Twenty-five of the 30 amino acids in the repeat motif fold around a zinc ion to form a small independent domain, the "finger," and the remaining five amino acids set up the linkers between consecutive fingers. Often, the structures of small protein domains are stabilized by the formation of disulfide bonds or by binding metal ions. Two cysteine and two histidine residues tetrahedrally coordinate a zinc ion to fold the domain into the finger-like projection[97]. In addition, each repeat contains three other conserved hydrophobic residues, namely Tyr6/Phe6, Phe17, and Leu23 that possibly function to form a hydrophobic clustered core that stabilizes the compact finger module[98]. Within a 30-amino acid repeat, a high concentration of basic and polar residues lies between the second cysteine and the first histidine implicating this region as the specific nucleic acid binding region[99]. The seven conserved residues of each ZNF domain forms a tertiary structure composed of two antiparallel β-sheets followed by an α-helix that creates a left-handed $\beta\beta\alpha$-module[99]. The two cysteine ligands form the loop located within the β-hairpin, at the zinc knuckle, and the two histidine residues are found on the C-terminal end of the α-helix[95].

Binding of the zinc finger motifs occurs though sequence-specific DNA recognition to the promoter regions of genes[100]. Of all the confirmed target DNA binding sequences recognized by zinc fingers, not one, or even several, conserved consensus sequences exist. Considering the vast number of known C2H2 zinc fingers and the highly conserved structure of the C2H2 motif, it is surprising that each zinc finger protein binds a specific DNA sequence recognized uniquely by itself. However, it is variations to key amino acid residues of the finger domains, spacing, and number of zinc finger motifs that allow for such distinction and specificity[96,99]. Biochemically, the mode of DNA recognition by a finger is principally a one-to-one interaction between individual amino acids from the recognition sequence of the α-helix to individual DNA bases; specifically, amino acids at helical positions -1, 3, and 6 to three successive triplet bases on one strand of the DNA, and helical position 2 to the complementary strand[99]. To establish amino acid-DNA contact, the N-terminus of the helix must angle down into the major groove. Clearly, each finger can function as an independent module with its own triplet binding sequence. When several ZNF

motifs are linked in tandem, each with different triplet specificities, and together grasping DNA in a linear fashion, a longer and distinctly unique DNA recognition sequence arises.
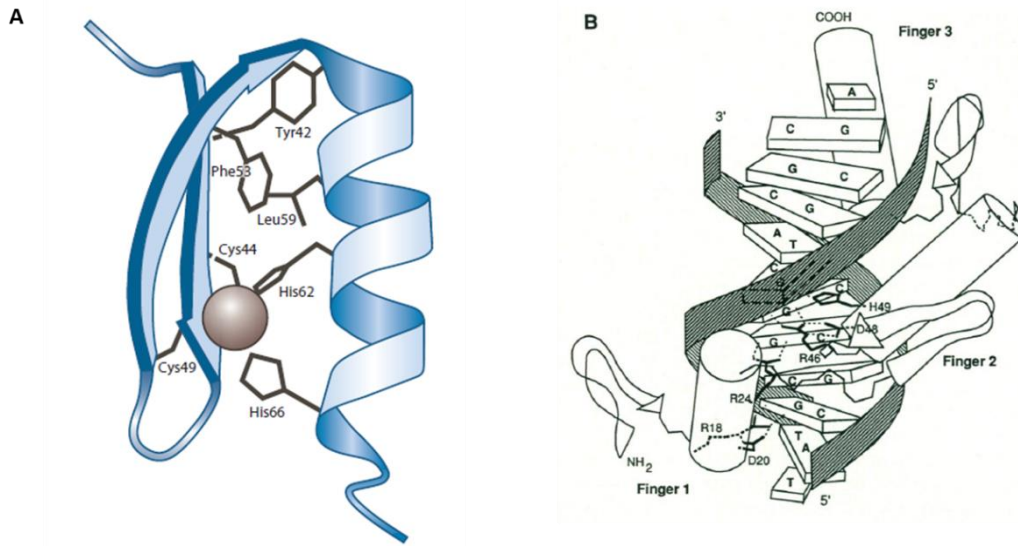
Figure 1. **Zinc Finger Structure.** (A) An individual zinc finger motif. (Copied from Klug, 2010). (B) A three-finger zinc finger factor bound to a DNA recognition site. (Copied from Pavletich & Pabo, 1991).

3. The KRAB Domain and KRAB-containing Zinc Fingers

Many C2H2 zinc finger proteins contain other conserved domains, in addition to their ZNF binding motifs, that also influence chemical distinctiveness; the BTB/POZ domain (Broad-Complex, Tramtrack, and Bric-a-brac/poxvirus and zinc finger), the SCAN domain (SRE-ZP, CTfin51, AW-1 and Number 18 cDNA), and the KRAB domain (Krüppel-Associated Box) have been found at the NH$_2$ terminus of zinc finger proteins[101].

Originally identified as a conserved motif at the amino-terminus of zinc finger proteins[102], the Krüppel-associated box (KRAB) domain is now reported to be found in nearly one-third of all C2H2 zinc finger proteins and are found highly conserved in yeast, plants, and across metazoans. The KRAB domain homology consists of approximately 75 amino acid residues and folds into two

amphipathic helices. The observed evolutionary conservation and wide distribution of the KRAB domain led to suggestions that there was an important role of this region in the transcriptional regulatory function of zinc finger proteins. Several laboratories subsequently uncovered the functional role of the KRAB domain as a potent DNA binding-dependent transcriptional repression module[103-105]. By fusing a heterologous DNA-binding domain from the yeast GAL4 protein, the KRAB domain minimal repression module of approximately 45 amino acid residues, the KRAB-A box, was shown to be necessary and sufficient for transcriptional repression[103-105], and substitutions for these conserved residues abolished repression[103].

One of the first studies to provide evidence that KRAB-containing zinc finger proteins (KRAB-ZFPs) actually bind DNA in a sequence-specific manner to regulate transcription, outside of an artificial GAL4-based transcriptional assay, found that each zinc finger motif is capable of contacting three to four nucleotides[106]. In addition, these transcriptional repressors typically use most of their collection of zinc fingers to bind DNA[106]. Considering this, a protein with 30 zinc finger domains, theoretically, could bind a DNA sequence of more than 60 nucleotides[107]. The occurrence of a sequence of such length rarely would be found in the genomes of lower eukaryotes, consistent with the knowledge that KRAB-containing proteins are found only in vertebrate tetrapods.

The KRAB domain is divided into the KRAB A-box and the KRAB B-box. The original characterization of these subdomains found that the A box alone is a considerably weaker suppression domain than the A + B boxes, but when fused to a heterologous KRAB B box, it induces repression more potently[105]. Therefore, the A box is the key repression module, and the B box enhances the repression mediated by the A box. Depending on the primary structure of the KRAB domain, mammalian KRAB-containing zinc finger proteins can be divided into three closely related subfamilies: one carrying the classical A box only (KRAB A), another carrying the classical KRAB A box together with the classical KRAB B box (KRAB A+B), and the other having the classical KRAB A box and a highly divergent KRAB B box, named b (KRAB A+b)[108]. Whether the amino-terminal domain contains either of these subfamilies, it is always known as the KRAB domain. Further, all three subfamilies effectively repress transcription through interaction

with KAP1[103,104,109-112]. Thus, it is these combination of factors that create the structural and functional versatility of the C2H2 zinc finger protein family.

4. KAP1: Structure and Mechanism

Four independent studies in 1996 identified a 100 kDa corepressor protein as an interaction partner of members of the KRAB domain-containing family of zinc finger transcription factors, named KAP1 (KRAB-associated protein 1), KRIP1 (KRAB-A-interacting protein 1), transcription intermediary factor (TIF)^{2}1β, or TRIM28 (tripartite motif-containing protein 28)[109-111]. As a member of the transcription intermediary factor (TIF1) family, which includes four tripartite motif (TRIM) proteins, TIF1α, TIF1β, TIF1γ, and TIFδ, the architecture of KAP1 includes an N-terminal TRIM known as the RBCC (RING (really interesting new gene) finger, two B-box zinc fingers, and a coiled-coil) domain. In addition, KAP1 shares a central TIF1 signature sequence (TSS) domain, a central heterochromatin protein 1 (HP1)-binding domain, and a C-terminal combination plant homeodomain (PHD) and bromodomain with other TIF1 members. Unlike other TIF1 proteins, however, KAP1 does not contain a nuclear receptor (NR) box[113].

The N-terminal RBCC domain of KAP1 is a KRAB-ZFP interaction interface spanning between amino acids 20 to 377, and is considered an absolute requirement for KAP1 recruitment to the KRAB repression module of KRAB-ZFPs. All three subdomains of RBCC directly contribute to the oligomerization and KAP1 recognition with high affinity and degree of specificity[114]. The RBCC domain binds as a homotrimer to a single KRAB domain[115]. The TSS domain is adjacent to the RBCC domain; deletion of this motif abolishes transcriptional repression activity of TIF1γ[116], although a functional role of TSS in KAP1 has yet to be identified. Also centrally located, the HP1-binding domain that contains a hydrophobic PxVxL pentapeptide that lies between amino acids 483 to 497[117]. KAP1 interacts with the chromoshadow domain of the HP1 family proteins, and this interaction is required for KAP1-mediated gene silencing[112] . It is suggested that HP1 is recruited by the KRAB-ZFP-KAP1 complex to specific loci within the genome and form heterochromatin that silences gene activity in euchromatic and pericentric heterochromatic regions[118].

The C-terminal end of KAP1 contains the tandem PHD and bromodomain (named the PB domain). Located between amino acids 618 and 835, these two domains function as a highly cooperative unit to repress transcription; the high specificity of cooperative function require both domains in order to obtain maximum repressive activity[119]. Bromodomains are commonly found in transcriptional activators, specifically involved in the recognition of acetylated histone tails[120]. Typical for bromodomain-containing proteins, the bromodomain of KAP1 has a conserved hydrophobic core and recognizes the backbone of histone tails; in contrast, KAP1 has lost the ability to contact acetylated lysine residues[119].
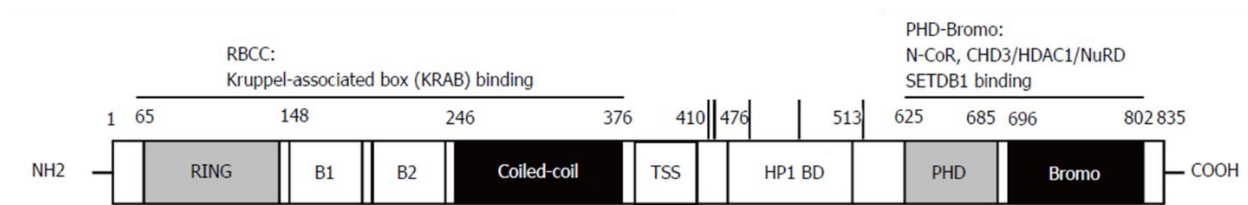
Figure 2. **KAP1 Structure.** (Copied from Cheng et al., 2014).

Molecular Mechanisms of Transcriptional Control by KAP1

KRAB-ZFP-mediated transcriptional silencing requires the presence of KAP1, a protein with several different well-studied functional domains. Logically then KAP1 could coordinate the assembly of a macromolecular complex made up of chromatin-remodeling proteins that function as effector molecules of silencing. Notably, this complex comprises histone deacetylases and the histone methyltransferase, SETDB1, which leads to histone deacetylation, deposition of H3K9me3, binding of HP1, formation of heterochromatin, and transcriptional silencing[110,111].While the RBCC domain of KAP1 acts as a high affinity interface for KRAB-ZFP-KAP1 binding, the primary function of the C-terminal PB domain of KAP1 is to interact with chromatin-modifying enzymes. The interdependence of the PHD and bromodomain for optimum repression may partially be explained by their interaction with two chromatin-modifying enzymes

that have been studied extensively: Mi-2α, an isoform of the Mi-2 proteins found in the NuRD (nucleosome remodeling and histone deacetylation) complex, and SETDB1 (SET domain, bifurcated 1), an H3K9me3-specific histone methyltransferase. NuRD complex proteins, specifically Mi-2α and HDACs, mediate nucleosome remodeling and histone deacetylation, respectively[119]. SETDB1-mediated trimethylation of histone H3 at Lys-9 creates high affinity genomic binding sites for the KAP1-HP1 complex. The PHD domain acts as an intramolecular E3 ligase of KAP1 by directing specific sumoylation modifications of particular lysine residues in the bromodomain, a post-translational modification that is required for KAP1 to recruit SETDB1[121].

And so, KAP1 has the capacity to coordinate biochemical activities required to induce and maintain the assembly of higher-order chromatin structure by epigenetically regulating gene expression through multiple transcriptional co-repressor complexes. Thus, KAP1 acts as a scaffold for chromatin-modifying complexes and chromatin remodeling activities by recruitment to the promoters of target genes and initiating ATP-dependent activities that modify chromatin. These observations suggest a model in which KRAB-ZFP-KAP1-dependent recruitment of histone modifiers for histone methylation and formation of facultative heterochromatin act to achieve gene silencing.

5. Molecular Mechanisms of KRAB-ZFP-KAP1-Mediated Transcriptional Regulation

The majority of research on KRAB-ZFP transcriptional regulation is performed on individual ZFP genes and mainly studied using artificial assays. The transcriptional repressive functions of KRAB-ZFPs are certainly well defined *in vitro*, while less is known *in vivo* and on a whole genome scale. However, recent studies have begun to examine genome-wide effects and even new models of regulation have been proposed.

A unique and delicately designed method of *in vivo* experimentation for KRAB-ZNF-KAP1 specific studies was created by the D. Trono group[122]. They created a conditional gene regulation system by fusing KRAB to the DNA binding domain of the *E. coli* tetracycline repressor (*tetR*), tTRKRAB, which is then able to bind tetracycline operator sequences (*tetO*), and induce transcriptional repression. Upon lentiviral vector transduction, tTRKRAB binds *tetO* in a doxycycline (dox)-controllable fashion creating a tightly controlled expression system in cell lines

and transgenic mice[122]. Research using this method found that KRAB-KAP1-mediated repression leads to permanent gene silencing through adjacent promoter DNA methylation only during the first few days of embryogenesis, while repression was reversible outside of this time frame[123]. KRAB-KAP1 are likely essential for *de novo* CpG methylation following the post-fertilization genome-wide erasure of methylation, placing an imperative role of ZNFs on early embryonic development.

Groner et al. were able to use the same drug-controlled *in vivo* KRAB knockdown system, but with gene-"trapped" lentiviral vectors, to study the impact of KRAB-ZFPs on gene expression[124]. They found that KRAB and its corepressor KAP1 are capable of silencing promoters located several tens of kilobases away from their primary docking sites. In addition, the silenced promoters displayed repressive chromatin marks, such as a loss of histone H3-acetylation, and increase in H3K9me3, and a drop in RNA Pol II recruitment. Furthermore, KRAB-mediated repression was established by the long-range spreading of the repressive marks of H3K9me3 and HP1β between the repressor binding site and the promoter[124]. They suggested that KRAB/KAP1 recruitment induces long-range repression through the spread of heterochromatin, and speculated that dysregulation of KRAB/KAP1-mediated epigenetic changes could be a cause of long-range epigenetic silencing in large chromosomal regions of cancerous cells.

6. Functional Roles of KRAB-ZFP-KAP1

KRAB-ZNF-KAP1-mediated regulation has been linked to essential and diverse cellular, physiological and pathological processes including development, proliferation, differentiation, metabolism, apoptosis, cell cycle, neoplastic transformation, stem cell pluripotency, early embryonic development and differentiation, genomic imprinting, response to DNA damage, control of behavioral stress. A few of these specific functional roles are discussed in more detail below. As all KRAB-containing ZNF proteins are known to interact with KAP1, to date, it is henceforth implied that any KRAB-ZNF discussed has been shown to interact with KAP1, unless stated otherwise.

Chromatin/Epigenetic

One of the best established functions of the KRAB/KAP1 system is that it protects genome integrity during early embryonic development. Several lines of evidence indicate that KAP1 and KRAB-ZNFs control endogenous retroelements during early embryogenesis. *In vitro* KAP1 deletion in embryonic stem cells (ESCs) and embryos by Cre-mediated excision lead to a significant upregulation of a range of endogenous retroviruses (ERVs)[125]. These endogenous retroelements, which contain *cis*-acting regulatory elements that can influence neighboring genes, are species-restricted, and, therefore, it is believed that their silencing in different species requires distinct sets of sequence-specific repressors. A KRAB-ZNF, ZFP809, is responsible for transcriptionally silencing murine leukemia virus, as well as a large subset of mammalian retroviruses in murine embryonic stem- and carcinoma-cells, through recognition of the sequence encoding for its primer-binding site[126]. Further, KRAB-ZNF proteins are involved in the generation of site-specific DNA methylation patterns during the early embryonic period. A group of investigators suggested a mechanism by which the site-specific KRAB-KAP1-mediated induction of heterochromatin leads to *de novo* DNA methylation during early embryogenesis at thousands of genomic loci in embryonic stem cells that are found to be methylated in adult tissues[127]. Therefore, ZNFs contribute to the genome-wide establishment of epigenetic marks that are maintained during development. Subsequently, they proposed a system in which KRAB-ZNFs, through interaction with KAP1 and SETDB1, are responsible for permanently silencing ERVs by *de novo* DNA methylation and, thereby, relieve the need for continuous expression of the ZNF *trans*-repressors[128]. All in all, control over these highly diverse and rapidly mutating genetic invaders seems to be driven largely by KAP1 and KRAB-ZNFs.

Parent-of-origin-specific expression of imprinted genes is required for normal embryonic development. Protection of the inherited, germ line-derived methylation at imprinted loci is vital, especially when egg- and sperm-derived genomes undergo extensive epigenetic reprogramming to a totipotent state. Identified during a gene trapping screen for factors downregulated upon embryonic stem cell differentiation, a maternal-zygotic effect gene, *Zfp57*, was found to be a regulator of *de novo* DNA methylation at several particular imprint control regions (ICRs). Targeted deletion of *Zfp57* in mouse oocytes revealed that it is required for the establishment and

reacquisition of *de novo* maternal methylation imprints, specifically at the *Snrpn* imprint control region[129], which is well known for its association in human Prader-Willi and Angelman syndromes. Zfp57 also maintains both paternal and maternal methylation imprints after fertilization, a period when the preimplantation embryonic genome loses most of its methylation[129]. Mutations in human *Zfp57* correlate with transient neonatal diabetes, a disease associated with imprinting defects[130]; specifically, hypomethylation of imprinted loci lead to a conserved range of clinical features. Biochemically, sequence-specific recognition of a methylated hexanucleotide motif found in all known ICRs, by ZFP57, will subsequently recruit KAP1, SETDB1, and DNA methyltransferases to the imprinted loci and, thereby, protect them from the genome-wide wave of demethylation that takes place following fertilization[131]. This ZFP57-KAP1 recruitment is essential for the maintenance of epigenetic asymmetry, including chromatin preservation and DNA methylation, of ICRs during the period of epigenetic instability that makes up the first several days of embryogenesis. Deletion of KAP1 from the maternal germ line results in embryonic lethality that is believed to occur as a result of misregulation of maternal genomic imprinting[132].

Pluripotency

Embryonic stem cells (ESCs) display a unique characteristic of pluripotency, namely, the ability to self-renew as well as the potential to differentiate into diverse cell types of the three germ layers. Considerable data indicate that ZNF proteins play an important role in maintenance of ESC pluripotency and differentiation potential as well as proliferation and cell cycle control. Of the cells commonly used as *in vitro* models in the mouse, from stem- to differentiated-cell types, the majority of KRAB-ZNF genes were found to be expressed in pluripotent stem cells and other early progenitors[133]. KAP1 and another pluripotency KRAB-ZNF, Zfx, have also been demonstrated to be required for stem cell self-renewal as part of module in a network that is distinct from the Oct4-Sox2-Nanog module[134]. Further, knockdown of KAP1 in murine stem cells resulted in differentiation to the primitive ectoderm lineage[134]. Hence, KAP1 is known to maintain pluripotency and also is required for the terminal differentiation of ESCs[134-136]. Upon knockdown of the KRAB-ZNF, *Zfp819*, high activation of ERVs is observed in ESCs suggesting it maintains

genomic integrity and downregulates endogenous retroviral elements in mouse embryonic stem cells[137]. In addition, interaction of the transcription factors Oct4, Sox2, and Nanog to the distal promoter region of *Zfp819* further implicate this ZNFs essential role in the pluripotency of stem cells.

Notably, several non-KRAB-domain-containing ZNF genes have demonstrated functional significance for pluripotency. The SCAN domain-containing protein, Zfp206, controls gene expression and differentiation of ESCs by activating transcription of *Oct4* and *Nanog* and preventing differentiation[138]. In a similar manner, *Zfp42/Rex1* is required to maintain ESCs in an undifferentiated state and promote self-renewal[139]. To continue self-renewal, ESCs must ensure an ability to rapidly proliferate. Loss-of-function either by gene deletion or siRNA-mediated knockdown of CtBp-interacting BTB (CIBZ), a BTB-containing ZNF protein, inhibits ESC proliferation and delays the progression of ESCs through the G1 to S phase transition[140]. Albeit, CIBZ-dependent ESC proliferation is in part dependent on the expression of Nanog[140]. Several pluripotency-related transcription factors, such as Nanog and even Zfp42/REX1, are heterogeneously expressed in ESC lines. Depending on the state of maintenance, these factors may be in transition between a ground state of pluripotency, in which lineage-specific genes are silenced, and a state primed to differentiate characterized by fluctuations in pluripotency factor.

Recently, the M. Saitou lab provided a model from their examination of this "meta-stable" state of pluripotency using mouse ESCs. A PR-domain ZNF protein 14 (PRDM14) ensures naïve pluripotency through a dual mechanism of interfering with the activation of fibroblast growth factor receptor (FGFR) by the *Oct4-Sox9-Nanog* complex and by repressing *de novo* DNA methyltransferase expression that alters the epigenome to a primed epiblast-like state[19]. PRDM14 is not a KRAB domain-containing ZNF; therefore, it does not recruit KAP1. However, it was shown to exert its effects by recruiting polycomb repressive complex 2 (PRC2) and binding specific key targets as a transcriptional repressor unit[19].

Early Development & Differentiation

Although the overall functions of most KRAB-ZNF genes remain elusive, functional trends seem to be becoming evident. Arisen from common ancestral genes, KRAB-ZNFs duplicated and

diverged throughout evolutionary history to display individual patterns and, therefore, have evolved specialized roles in the development and differentiation of higher organisms. KRAB-ZNFs have been implicated in nearly all aspects of differentiation including hematopoiesis, neuronal development, morphogenesis, cardiogenesis, osteoblast formation, and embryogenesis, to name a few.

Dysfunctions of KRAB-ZFPs exhibit severe mutant phenotypes during embryogenesis. The definitive endoderm layer, which is the precursor of the gut, narrows and elongates during embryogenesis and undergoes cell rearrangements and intercalation of tissues known as convergent extension. The *Chato* mutation, which in the mouse is a KRAB-ZNF gene responsible for body axis elongation in embryonic tissues, causes defects in convergent extension during development[141]. Specifically, *Chato* mutants express a wider definitive endoderm and lack of cell rearrangements[141]. Similarly, as a co-factor that is necessary for the transcriptionally repressive functions associated with KRAB-ZNFs, KAP1 has been found to be essential for early embryogenesis. Mice with a targeted deletion of KAP1 do not survive past the egg cylinder stage, prior to the onset of gastrulation, and are completely resorbed[142]. KAP1-null embryos have reduced cell number in the ectoderm, morphological alteration of the visceral endoderm, and absence of mesoderm formation[142].

As an important regulator in the homeostasis of the seminiferous epithelium, KAP1 is required for the maintenance of spermatogenesis. Depletion of KAP1 in a germ cell lineage leads to testicular degeneration, specifically by shedding of immature spermatocytes and spermatids and disappearance of the stem population[143]. One of the few KRAB-ZNF proteins to be studied *in vivo*, murine *Rsl*, is known to influence sexually dimorphic gene expression in the liver[144]. Specifically, two variant paralogs have evolved to partition regulation of their target genes by repressing transcription of male-specific liver genes[144]. As part of the adaptive immune system, a discrete subset of KRAB-ZNFs were found to be enriched in T- and B-lymphocytes, and upon tethering KAP1 to particular genomic targets, forms a complex that regulates gene networks to control T- and B-cell differentiation and responsiveness[145,146].

In spite of their numerical abundance and clearly important functional requirement, notably few KRAB-ZNFs have been assigned specific functions. The KRAB-ZNF genes discussed here stand out and are clearly essential for genomic integrity, embryonic development, differentiation, and pluripotency and, like any KRAB-ZNF-KAP1 complex, usually have more than one functional responsibility. Loss or mutation of any of these genes is not without serious consequence. Deep sequencing has uncovered hundreds of known and novel ZNF genes that have been cataloged into databases. This available sequence information, along with chromatin immunoprecipitation (ChIP), micro arrays, mass spectrometry, and other interaction methods has allowed for a broad view of the ZNF transcription factor landscape, including spatio-temporal expression and binding patterns[7,133]. A computational prediction program for specific DNA-binding recognition sites by individual C2H2 ZNF motifs has been created recently[8]. Yet, it still remains that focused experimentation of particular genes is required to fully understand the biological functional role of a KRAB-ZNF (or any transcription factor).

7. Evolutionary Conservation, Distribution, and Genomic Organization of KRAB-ZFPs

Despite the significant number of members belonging to this gene family, KRAB-ZFPs and their organismal functions are far from being completely understood. Examining the molecular mechanisms that lead to the generation of this gene family during species evolution may unveil important information for understanding their function. The KRAB-ZFP gene family is believed to represent a more recent evolutionary product as indicated by its expansion in the genome of vertebrate tetrapods; the KRAB domain is absent in the zinc finger protein sequences of fish, insects, plants, nematodes, yeast, and fungi but has been identified in the human, mouse, rat, chicken, and frog genomes[102,108]. The importance of KRAB-ZFPs is inferred from their recent origin and subsequent rapid expansion in vertebrate lineages, although their *in vivo* role in terms of both whole genome and physiological function has only recently begun to unravel.

Duplication & Divergence

The addition of the KRAB domain as a transcriptional repressor first arose as part of polydactyl ZNF (poly-ZNF) genes in tetrapod vertebrates, a distribution that suggests the emergence of the KRAB domain is a relatively recent event in evolution. KRAB-containing genes

28

are typically arranged in clusters likely reflecting a history of duplication events; albeit, many occur individually throughout the genome. The poly-ZNF genes are clustered at particular sites on chromosomes, a shared physical proximity that tends to be associated with genes closely related in sequence. The lack of degree of divergence in these clustered regions is consistent with a sporadic duplication process in which new genes arise by local duplication events affecting a few genes and then gradually disperse by subsequent genome rearrangements that break up the gene clusters over time[147].

The existence of such large and highly conserved numbers of genes undergoing repeated cycles of segmental duplications likely originated from a single ancestral gene. Zinc finger gene duplication commonly occur throughout evolution. After each duplication event a relatively low degree of sequence conservation exists as new genes diversify their coding regions to generate novel proteins. Orthologous KRAB-ZFP genes generally remain well conserved; in contrast, paralogous KRAB-ZFP genes are not necessarily under functionally selective pressure and therefore allow amino acid sequence changes via non-synonymous nucleotide substitutions[108]. Essentially all, except the most recently duplicated KRAB-ZFP genes, are found to display structural and/or functional divergence compared with parental genes[148]. Thus, after paralogous duplications, novel zinc finger proteins with new biological functions exist, perhaps to define new regulatory pathways. A possible functional consequence of this expansion could be the generation of new transcriptional regulators, as the binding specificities of the encoded zinc finger motifs is altered by accumulating changes in the amino acid sequence of the zinc finger region [108]. The binding specificities can be further modified by recombination-based additions and inactivations of entire zinc finger motifs as a result of internal duplications[108]. Hence, the evolution of transcription factors with substantially altered DNA-binding specificities arises.

Biochemical Function

The expansion of C2H2 ZNF transcription factors in eukaryotes is remarkable. KRAB-ZFPs have been recognized as important subjects of lineage-specific expansion in vertebrates. Interestingly, not only has this expansion increased in the total number of zinc finger genes throughout evolution but also in the number of DNA-binding zinc finger motifs carried in each

individual gene. In general, the average number of zinc finger motifs for a zinc finger gene of a plant (*A. thaliana*), baker's yeast (*S. cerevisiae*), a nematode (*C. elegans*), an insect (*D. melanogaster*) and humans (*H. sapiens*) is 1.0, 1.5, 2.5, 3.5, and 8, respectively[5,108]. In addition, these ZNF transcription factors may contain either the KRAB or SCAN domains, or both, which are not found in the fly or worm genomes, and increase the combinatorial pattern possibilities of these factors. Further, each ZNF binding domain may be capable of independently mediating transcriptional repression by inducing local chromatin to a closed state. This model would make sense given that many mammalian poly-ZNF proteins contain far more zinc finger repeats than seem necessary to bind a single target sequence. The repetitive nature of the zinc finger region and the rapid divergence in the binding properties within the zinc finger motif make these proteins highly adaptable[147]. These factors have been of major importance for the massive expansion in both total number of genes and complexity of zinc finger motifs within in each gene during eukaryotic evolution.

Speciation & Evolved Biological Function

The overall, ultimate *in vivo* function of the poly-ZNF family is not quite evident. Genes encoding KRAB-ZFPs are differentially expressed in various tissues during differentiation and development. It seems likely, then, that these genes have functions unique to mammalian evolution, especially involving the molecular processes that establish the phenotypic differences between vertebrates and other species[107,149]. In addition, the modular structure of KRAB-ZFP genes creates an ideal structure for rapid evolution of transcriptional regulation. As a substrate composed of consecutive zinc finger motifs, which fold autonomously, coupled to a transcriptional repression domain that operates independently of the sequence target, it may only require a few point mutations or small rearrangements to alter the transcriptional outcome of target genes. Basically, the poly-ZNF gene family's flexibility and lack of stability make it opportune to rapid adaptive evolution of transcriptional regulation[147].

Of the hundreds of poly-ZNF loci examined over many species genomes, surprisingly few ZNF encoding proteins are actually conserved between eutherians and other evolutionary groups[150]. The few that are considered "deeply conserved" between evolutionary groups were

found to be mapped to a single familial cluster and displayed an unusual noncanonical KRAB domain sequence that does not bind KAP1 and functions as a transcriptional activator[150]. And it seems the most highly conserved ZNFs are those that fit the ZNF-only and BTB/POZ-ZNF protein structures. This further confirms the history of KRAB activation and uncovers a past in which KRAB-ZNF proteins underwent independent divergence and expansion in every vertebrate lineage[150].

The diverse functional range of ZNF proteins in combination with the dichotomy between orthologous ZNF genes and paralogous expanded gene clusters suggests the possibility of more than one type of organismal function. First, because of the dramatic expansion of gene clusters and rapid divergence of KRAB-ZFP genes in mammals, it has been suggested that they function to repress transcription of endogenous retrovirus genes in an evolutionary "arms race" with their viral targets[151,152]. The second possibility is based on the strict conservation of the pattern of DNA-binding amino acids (position -1, 2, 3, and 6, as discussed above in the C2H2 ZNF section) in polydactyl ZFP proteins. This orthologous manner of speciation suggests that ZFP DNA binding activities have evolved critical biological roles, specifically those to modulate transcription of developmental genes. Most ZNF genes, whether well conserved or species-specific, seem to have a significant presence in immune, nervous, and reproductive tissues[147,149,150], especially during embryogenesis and influencing morphogenic processes[133,149], indicating that they have been recruited to regulate evolutionarily divergent biological traits in vertebrates.
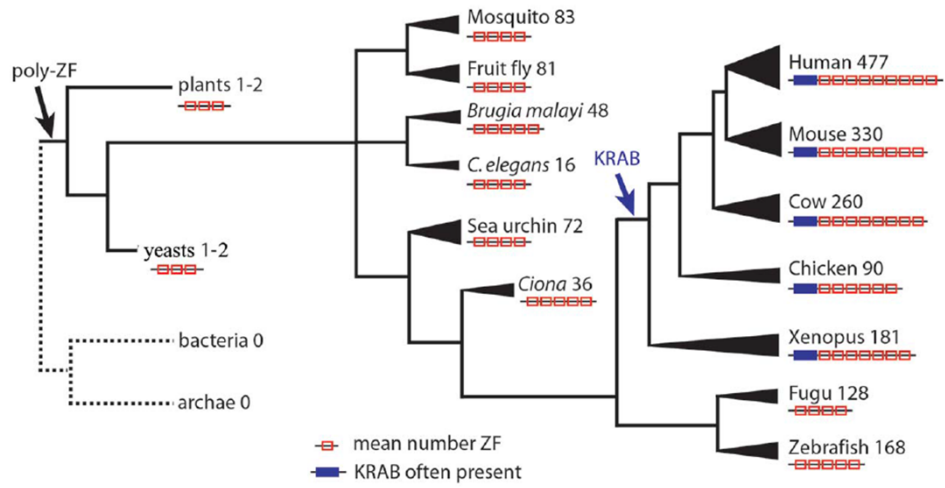
***Figure 3*. Polydactyl Zinc Finger Gene Family Across Species.** (Copied from Emerson & Thomas, 2009).

REFERENCES

1. Krisher RL. The effect of oocyte quality on development. *J Anim Sci*. 2004;82 E-Suppl:E14-23.

2. Picton H, Briggs D, Gosden R. The molecular basis of oocyte growth and development. *Mol Cell Endocrinol*. 1998;145(1-2):27-37. doi: 10.1016/S0303-7207(98)00166-X.

3. Gilchrist RB, Thompson JG. Oocyte maturation: Emerging concepts and technologies to improve developmental potential in vitro. *Theriogenology*. 2007;67(1):6-15. doi: 10.1016/j.theriogenology.2006.09.027.

4. Sirard M, Richard F, Blondin P, Robert C. Contribution of the oocyte to embryo quality. *Theriogenology*. 2006;65(1):126-136. doi: 10.1016/j.theriogenology.2005.09.020.

5. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science*. 2001;291(5507):1304-+. doi: 10.1126/science.1058040.

6. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: Function, expression and evolution. *Nature Reviews Genetics*. 2009;10(4):252-263. doi: 10.1038/nrg2538.

7. Najafabadi HS, Mnaimneh S, Schmitges FW, et al. C2H2 zinc finger proteins greatly expand the human regulatory lexicon. *Nat Biotechnol*. 2015;33(5):555-U181. doi: 10.1038/nbt.3128.

8. Persikov AV, Wetzel JL, Rowland EF, et al. A systematic survey of the cys(2)his(2) zinc finger DNA-binding landscape. *Nucleic Acids Res*. 2015;43(3):1965-1984. doi: 10.1093/nar/gku1395.

9. Lawson K, Hage W. Clonal analysis of the origin of primordial germ-cells in the mouse. *Germline Development*. 1994;182:68-84.

10. Johnson AD, Richardson E, Bachvarova RF, Crother BI. Evolution of the germ line-soma relationship in vertebrate embryos. *Reproduction*. 2011;141(3):291-300. doi: 10.1530/REP-10-0474.

11. Extavour C, Akam M. Mechanisms of germ cell specification across the metazoans: Epigenesis and preformation. *Development*. 2003;130(24):5869-5884. doi: 10.1242/dev.00804.

12. Tam P, Zhou S. The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev Biol*. 1996;178(1):124-132. doi: 10.1006/dbio.1996.0203.

13. Bowles J, Koopman P. Sex determination in mammalian germ cells: Extrinsic versus intrinsic factors. *Reproduction*. 2010;139(6):943-958. doi: 10.1530/REP-10-0075.

14. Ying Y, Liu X, Marble A, Lawson K, Zhao G. Requirement of Bmp8b for the generation of primordial germ cells in the mouse. *Molecular Endocrinology*. 2000;14(7):1053-1063. doi: 10.1210/me.14.7.1053.

15. Ying Y, Zhao G. Cooperation of endoderm-derived BMP2 and extraembryonic ectoderm-derived BMP4 in primordial germ cell generation in the mouse. *Dev Biol*. 2001;232(2):484-492. doi: 10.1006/dbio.2001.0173.

16. Ohinata Y, Payer B, O'Carroll D, et al. Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature*. 2005;436(7048):207-213. doi: 10.1038/nature03813.

17. Saitou M, Payer B, O'Carroll D, Ohinata Y, Surani M. Blimp1 and the emergence of the germ line during development in the mouse. *Cell Cycle*. 2005;4(12):1736-1740. doi: 10.4161/cc.4.12.2209.

18. Vincent S, Dunn N, Sciammas R, et al. The zinc finger transcriptional repressor Blimp1/Prdm1 is dispensable for early axis formation but is required for specification of primordial germ cells in the mouse. *Development*. 2005;132(6):1315-1325. doi: 10.1242/dev.01711.

19. Yamaji M, Seki Y, Kurimoto K, et al. Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nat Genet*. 2008;40(8):1016-1022. doi: 10.1038/ng.186.

20. Saitou M. Germ cell specification in mice. *Curr Opin Genet Dev*. 2009;19(4):386-395. doi: 10.1016/j.gde.2009.06.003.

21. Yabuta Y, Kurimoto K, Hinata Y, Seki Y, Saitou M. Gene expression dynamics during germline specification in mice identified by quantitative single-cell gene expression profiling. *Biol Reprod*. 2006;75(5):705-716. doi: 10.1095/biolreprod.106.053686.

22. Seki Y, Hayashi K, Itoh K, Mizugaki M, Saitou M, Matsui Y. Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev Biol*. 2005;278(2):440-458. doi: 10.1016/j.ydbio.2004.11.025.

23. Zama A, Hudson F, Bedell M. Analysis of hypomorphic kitl(sl) mutants suggests different requirements for KITL in proliferation and migration of mouse primordial germ cells. *Biol Reprod*. 2005;73(4):639-647. doi: 10.1095/biolreprod.105.042846.

24. Anderson R, Fassler R, Georges-Labouesse E, et al. Mouse primordial germ cells lacking beta 1 integrins enter the germline but fail to migrate normally to the gonads. *Development*. 1999;126(8):1655-1664.

25. Di Carlo A, De Felici M. A role for E-cadherin in mouse primordial germ cell development. *Dev Biol*. 2000;226(2):209-219. doi: 10.1006/dbio.2000.9861.

26. Washburn L, Albrecht K, Eicher E. C57BL/6J-T-associated sex reversal in mice is caused by reduced expression of a mus domesticus sry allele. *Genetics*. 2001;158(4):1675-1681.

27. Monk M, McLaren A. X-chromosome activity in fetal germ-cells of the mouse. *J Embryol Exp Morphol*. 1981;63(JUN):75-84.

28. Speed R. Meiosis in the fetal mouse ovary .1. an analysis at the light-microscope level using surface-spreading. *Chromosoma*. 1982;85(3):427-437. doi: 10.1007/BF00330366.

29. Peters H. Development of mouse ovary from birth to maturity. *Acta Endocrinol*. 1969;62(1):98-&.

30. Fortune JE, Eppig JJ. Effects of gonadotropins on steroid-secretion by infantile and juvenile mouse ovaries invitro. *Endocrinology*. 1979;105(3):760-768.

31. Pepling ME, Spradling AC. Mouse ovarian germ cell cysts undergo programmed breakdown to form primordial follicles. *Dev Biol*. 2001;234(2):339-351. doi: 10.1006/dbio.2001.0269.

32. Russe I. Oogenesis in cattle and sheep. *Bibl Anat*. 1983;24:77-92.

33. McNatty KP, Smith P, Hudson NL, et al. Development of the sheep ovary during fetal and early neonatal life and the effect of fecundity genes. *Journal of reproduction and fertility.Supplement*. 1995;49:123-35.

34. Yang MY, Fortune JE. The capacity of primordial follicles in fetal bovine ovaries to initiate growth in vitro develops during mid-gestation and is associated with meiotic arrest of oocytes. *Biol Reprod*. 2008;78(6):1153-1161. doi: 10.1095/biolreprod.107.066688.

35. Fair T. Follicular oocyte growth and acquisition of developmental competence. *Anim Reprod Sci*. 2003;78(3-4):203-216. doi: 10.1016/S0378-4320(03)00091-5.

36. Hirshfield AN. Development of follicles in the mammalian ovary. *International Review of Cytology-a Survey of Cell Biology*. 1991;124:43-101. doi: 10.1016/S0074-7696(08)61524-7.

37. Fortune JE. Ovarian follicular-growth and development in mammals. *Biol Reprod*. 1994;50(2):225-232. doi: 10.1095/biolreprod50.2.225.

38. Fortune JE. The early stages of follicular development: Activation of primordial follicles and growth of preantral follicles. *Anim Reprod Sci*. 2003;78(3-4):135-163. doi: 10.1016/S0378-4320(03)00088-5.

39. Hyttel P, Fair T, Callesen H, Greve T. Oocyte growth, capacitation and final maturation in cattle. *Theriogenology*. 1997;47(1):23-32. doi: 10.1016/S0093-691X(96)00336-6.

40. Danforth DR. Endocrine and paracrine control of oocyte development. *Obstet Gynecol*. 1995;172(2):747-752. doi: 10.1016/0002-9378(95)90148-5.

41. Elvin JA, Yan CN, Matzuk MM. Oocyte-expressed TGF-beta superfamily members in female fertility. *Mol Cell Endocrinol*. 2000;159(1-2):1-5. doi: 10.1016/S0303-7207(99)00185-9.

42. Knight PG, Glister C. Potential local regulatory functions of inhibins, activins and follistatin in the ovary. *Reproduction*. 2001;121(4):503-512. doi: 10.1530/rep.0.1210503.

43. Adashi EY. The IGF family and folliculogenesis. *J Reprod Immunol*. 1998;39(1-2):13-19. doi: 10.1016/S0165-0378(98)00026-6.

44. Wandji SA, Pelletier G, Sirard MA. Ontogeny and cellular-localization of I-125 labeled basic fibroblast growth-factor and I-125 labeled epidermal growth-factor binding-sites in ovaries from bovine fetuses and neonatal calves. *Biol Reprod*. 1992;47(5):807-813. doi: 10.1095/biolreprod47.5.807.

45. Kidder GM, Mhawi AA. Gap junctions and ovarian folliculogenesis. *Reproduction*. 2002;123(5):613-620. doi: 10.1530/rep.0.1230613.

46. Eppig JJ, Wigglesworth K, Pendola FL. The mammalian oocyte orchestrates the rate of ovarian follicular development. *Proc Natl Acad Sci U S A*. 2002;99(5):2890-2894. doi: 10.1073/pnas.052658699.

47. Baker TG, Hunter RHF. Interrelationships between oocyte and somatic-cells within graafian follicle of mammals. *Annales De Biologie Animale Biochimie Biophysique*. 1978;18(NB2):419-426. doi: 10.1051/rnd:19780312.

48. Hilscher B, Hilscher W, Bulthoffohnolz B, et al. Kinetics of gametogenesis .1. comparative histological and autoradiographic studies of oocytes and transitional prospermatogonia during oogenesis and prespermatogenesis. *Cell Tissue Res*. 1974;154(4):443-470.

49. Sorensen RA, Wassarman PM. Relationship between growth and meiotic maturation of mouse oocyte. *Dev Biol*. 1976;50(2):531-536. doi: 10.1016/0012-1606(76)90172-X.

50. Matzuk MM. Intercellular communication in the mammalian ovary: Oocytes carry the conversation. *Science (New York, N.Y.)*. 2002;296(5576):2178-2180.

51. Stitzel ML, Seydoux G. Regulation of the oocyte-to-zygote transition. *Science*. 2007;316(5823):407-408. doi: 10.1126/science.1138236.

52. Bettegowda A, Smith GW. Mechanisms of maternal mRNA regulation: Implications for mammalian early embryonic development. *Frontiers in Bioscience*. 2007;12:3713-3726. doi: 10.2741/2346.

53. Li L, Lu X, Dean J. The maternal to zygotic transition in mammals. *Mol Aspects Med*. 2013;34(5):919-938. doi: 10.1016/j.mam.2013.01.003.

54. Zhang K, Smith GW. Maternal control of early embryogenesis in mammals. *Reproduction Fertility and Development*. 2015;27(6):880-896. doi: 10.1071/RD14441.

55. Jones KT, Lane SIR. Molecular causes of aneuploidy in mammalian eggs. *Development*. 2013;140(18):3719-3730. doi: 10.1242/dev.090589.

56. Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM. NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. *Science*. 2004;305(5687):1157-1159. doi: 10.1126/science.1099755.

57. Soyal S, Amleh A, Dean J. FIG alpha, a germ cell-specific transcription factor required for ovarian follicle formation. *Development*. 2000;127(21):4645-4654.

58. Ballow DJ, Xin Y, Choi Y, Pangas SA, Rajkovic A. Sohlh2 is a germ cell-specific bHLH transcription factor. *Gene Expression Patterns*. 2006;6(8):1014-1018. doi: 10.1016/j.modgep.2006.04.007.

59. Pangas SA, Choi Y, Ballow DJ, et al. Oogenesis requires germ cell-specific transcriptional regulators Sohlh1 and Lhx8. *Proc Natl Acad Sci U S A*. 2006;103(21):8090-8095. doi: 10.1073/pnas.0601083103.

60. Choi Y, Ballow DJ, Xin Y, Rajkovic A. Lim homeobox gene, Lhx8, is essential for mouse oocyte differentiation and survival. *Biol Reprod*. 2008;79(3):442-449. doi: 10.1095/biolreprod.108.069393.

61. Gu W, Tekur S, Reinbold R, et al. Mammalian male and female germ cells express a germ cell-specific Y-box protein, MSY2. *Biol Reprod*. 1998;59(5):1266-1274. doi: 10.1095/biolreprod59.5.1266.

62. Ruggiu M, Speed R, Taggart M, et al. The mouse dazla gene encodes a cytoplasmic protein essential for gametogenesis. *Nature*. 1997;389(6646):73-77.

63. Racki WJ, Richter JD. CPEB controls oocyte growth and follicle development in the mouse. *Development*. 2006;133(22):4527-4537. doi: 10.1242/dev.02651.

64. Dong JW, Albertini DF, Nishimori K, Kumar TR, Lu NF, Matzuk MM. Growth differentiation factor-9 is required during early ovarian folliculogenesis. *Nature*. 1996;383(6600):531-535. doi: 10.1038/383531a0.

65. Dube JL, Wang P, Elvin J, Lyons KM, Celeste AJ, Matzuk MM. The bone morphogenetic protein 15 gene is X-linked and expressed in oocytes. *Molecular Endocrinology*. 1998;12(12):1809-1817. doi: 10.1210/me.12.12.1809.

66. Choi Y, Rajkovic A. Genetics of early mammalian folliculogenesis. *Cellular and Molecular Life Sciences*. 2006;63(5):579-590. doi: 10.1007/s00018-005-5394-7.

67. Pangas SA. Growth factors in ovarian development. *Semin Reprod Med*. 2007;25(4):225-234. doi: 10.1055/s-2007-980216.

68. Pangas SA, Rajkovic A. Transcriptional regulation of early oogenesis: In search of masters. *Hum Reprod Update*. 2006;12(1):65-76. doi: 10.1093/humupd/dmi033.

69. Andreu-Vieyra C, Lin YN, Matzuk MM. Mining the oocyte transcriptome. *Trends in Endocrinology and Metabolism*. 2006;17(4):136-143. doi: 10.1016/j.tem.2006.03.010.

70. Bettegowda A, Lee K, Smith GW. Cytoplasmic and nuclear determinants of the maternal-to-embryonic transition. *Reproduction Fertility and Development*. 2008;20(1):45-53. doi: 10.1071/RD07156.

71. Tejomurtula J, Lee K, Tripurani SK, Smith GW, Yao J. Role of importin Alpha8, a new member of the importin alpha family of nuclear transport proteins, in early embryonic development in cattle. *Biol Reprod*. 2009;81(2):333-342. doi: 10.1095/biolreprod.109.077396.

72. Lee K, Bettegowda A, Wee G, Ireland JJ, Smith GW. Molecular determinants of oocyte competence: Potential functional role for maternal (oocyte-derived) follistatin in promoting bovine early embryogenesis. *Endocrinology*. 2009;150(5):2463-2471. doi: 10.1210/en.2008-1574.

73. Bettegowda A, Yao J, Sen A, et al. JY-1, an oocyte-specific gene, regulates granulosa cell function and early embryonic development in cattle. *Proc Natl Acad Sci U S A*. 2007;104(45):17602-17607. doi: 10.1073/pnas.0706383104.

74. Tripurani SK, Lee K, Wang L, et al. A novel functional role for the oocyte-specific transcription factor newborn ovary homeobox (NOBOX) during early embryonic development in cattle. *Endocrinology*. 2011;152(3):1013-1023. doi: 10.1210/en.2010-1134.

75. Hayashi K, Surani MA. Resetting the epigenome beyond pluripotency in the germline. *Cell Stem Cell*. 2009;4(6):493-498. doi: 10.1016/j.stem.2009.05.007.

76. Lees-Murdock DJ, Walsh CP. DNA methylation reprogramming in the germ line. *Epigenetics*. 2008;3(1):5-13.

77. Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. *Hum Mol Genet*. 2005;14:R47-R58. doi: 10.1093/hmg/ddi114.

78. Seisenberger S, Peat JR, Hore TA, Santos F, Dean W, Reik W. Reprogramming DNA methylation in the mammalian life cycle: Building and breaking epigenetic barriers. *Philosophical Transactions of the Royal Society B-Biological Sciences*. 2013;368(1609):20110330. doi: 10.1098/rstb.2011.0330.

79. Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol*. 2002;241(1):172-182. doi: 10.1006/dbio.2001.0501.

80. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*. 2013;502(7472):472-479. doi: 10.1038/nature12750.

81. Lee MT, Bonneau AR, Giraldez AJ. Zygotic genome activation during the maternal-to-zygotic transition. *Annual Review of Cell and Developmental Biology, Vol 30*. 2014;30:581-613. doi: 10.1146/annurev-cellbio-100913-013027.

82. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: A play in two acts. *Development*. 2009;136(18):3033-3042. doi: 10.1242/dev.033183.

83. Ostrup O, Andersen IS, Collas P. Chromatin-linked determinants of zygotic genome activation. *Cellular and Molecular Life Sciences*. 2013;70(8):1425-1437. doi: 10.1007/s00018-012-1143-x.

84. Memili E, First N. Control of gene expression at the onset of bovine embryonic development. *Biol Reprod*. 1999;61(5):1198-1207. doi: 10.1095/biolreprod61.5.1198.

85. Memili E, First N. Zygotic and embryonic gene expression in cow: A review of timing and mechanisms of early gene expression as compared with other species. *Zygote*. 2000;8(1):87-96. doi: 10.1017/S0967199400000861.

86. Telford N, Watson A, Schultz G. Transition from maternal to embryonic control in early mammalian development - a comparison of several species. *Mol Reprod Dev*. 1990;26(1):90-100. doi: 10.1002/mrd.1080260113.

87. Meirelles F, Caetano A, Watanabe Y, et al. Genome activation and developmental block in bovine embryos. *Anim Reprod Sci*. 2004;82-3:13-20. doi: 10.1016/j.anireprosci.2004.05.012.

88. Hamatani T, Carter MG, Sharov AA, Ko MSH. Dynamics of global gene expression changes during mouse preimplantation development. *Developmental Cell*. 2004;6(1):117-131. doi: 10.1016/S1534-5807(03)00373-3.

89. Graf A, Krebs S, Heininen-Brown M, Zakhartchenko V, Blum H, Wolf E. Genome activation in bovine embryos: Review of the literature and new insights from RNA sequencing experiments. *Anim Reprod Sci*. 2014;149(1-2):46-58. doi: 10.1016/j.anireprosci.2014.05.016.

90. Lemon B, Tjian R. Orchestrated response: A symphony of transcription factors for gene control. *Genes Dev*. 2000;14(20):2551-2569. doi: 10.1101/gad.831000.

91. Luscombe NM, Austin SE, Berman HM, Thornton JM. An overview of the structures of protein-DNA complexes. *Genome Biol*. 2000;1(1):001.

92. Gray PA, Fu H, Luo P, et al. Mouse brain organization revealed through direct genome-scale TF expression analysis. *Science*. 2004;306(5705):2255-2257. doi: 10.1126/science.1104935.

93. Miller J, Mclachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor iiia from xenopus oocytes. *EMBO J*. 1985;4(6):1609-1614.

94. Pelham HRB, Brown DD. A specific transcription factor that can bind either the 5s rna gene or 5s rna. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences*. 1980;77(7):4170-4174. doi: 10.1073/pnas.77.7.4170.

95. Krishna SS, Majumdar I, Grishin NV. Structural classification of zinc fingers: Survey and summary. *Nucleic Acids Res*. 2003;31(2):532-550.

96. Klug A, Schwabe JWR. Protein motifs .5. zinc fingers. *Faseb Journal*. 1995;9(8):597-604.

97. Lee MS, Gippert GP, Soman KV, Case DA, Wright PE. 3-dimensional solution structure of a single zinc finger dna-binding domain. *Science*. 1989;245(4918):635-637. doi: 10.1126/science.2503871.

98. Klug A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annual Review of Biochemistry, Vol 79*. 2010;79:213-231. doi: 10.1146/annurev-biochem-010909-095056.

99. Pavletich NP, Pabo CO. Zinc finger dna recognition - crystal-structure of a Zif268-dna complex at 2.1-a. *Science*. 1991;252(5007):809-817. doi: 10.1126/science.2028256.

100. Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P. KRAB-zinc finger proteins: A repressor family displaying multiple biological functions. *Curr Genomics*. 2013;14(4):268-278.

101. Collins T, Stone JR, Williams AJ. All in the family: The BTB/POZ, KRAB, and SCAN domains. *Mol Cell Biol*. 2001;21(11):3609-3615. doi: 10.1128/MCB.21.11.3609-3615.2001.

102. Bellefroid EJ, Poncelet DA, Lecocq PJ, Revelant O, Martial JA. The evolutionarily conserved kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proc Natl Acad Sci U S A*. 1991;88(9):3608-3612. doi: 10.1073/pnas.88.9.3608.

103. Margolin J, Friedman J, Meyer W, Vissing H, Thiesen H, Rauscher F. Kruppel-associated boxes are potent transcriptional repression domains. *Proc Natl Acad Sci U S A*. 1994;91(10):4509-4513. doi: 10.1073/pnas.91.10.4509.

104. Witzgall R, Oleary E, Leaf A, Onaldi D, Bonventre J. The kruppel-associated box-a (krab-A) domain of zinc-finger proteins mediates transcriptional repression. *Proc Natl Acad Sci U S A*. 1994;91(10):4514-4518. doi: 10.1073/pnas.91.10.4514.

105. Vissing H, Meyer W, Aagaard L, Tommerup N, Thiesen H. Repression of transcriptional activity by heterologous krab domains present in zinc-finger proteins. *FEBS Lett*. 1995;369(2-3):153-157. doi: 10.1016/0014-5793(95)00728-R.

106. Gebelein B, Urrutia R. Sequence-specific transcriptional repression by KS1, a multiple-zinc-finger-kruppel-associated box protein. *Mol Cell Biol*. 2001;21(3):928-939. doi: 10.1128/MCB.21.3.928-939.2001.

107. Urrutia R. KRAB-containing zinc-finger repressor proteins. *Genome Biol*. 2003;4(10):231. doi: 10.1186/gb-2003-4-10-231.

108. Looman C, Abrink M, Mark C, Hellman L. KRAB zinc finger proteins: An analysis of the molecular mechanisms governing their increase in numbers and complexity during evolution. *Mol Biol Evol*. 2002;19(12):2118-2130.

109. Kim SS, Chen YM, OLeary E, Witzgall R, Vidal M, Bonventre JV. A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. *Proc Natl Acad Sci U S A*. 1996;93(26):15299-15304. doi: 10.1073/pnas.93.26.15299.

110. Friedman J, Fredericks W, Jensen D, et al. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev*. 1996;10(16):2067-2078. doi: 10.1101/gad.10.16.2067.

111. Moosmann P, Georgiev O, LeDouarin B, Bourquin J, Schaffner W. Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. *Nucleic Acids Res*. 1996;24(24):4859-4867. doi: 10.1093/nar/24.24.4859.

112. Nielsen AL, Ortiz JA, You J, et al. Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *EMBO J*. 1999;18(22):6385-6395. doi: 10.1093/emboj/18.22.6385.

113. Iyengar S, Farnham PJ. KAP1 protein: An enigmatic master regulator of the genome. *J Biol Chem*. 2011;286(30):26267-26276. doi: 10.1074/jbc.R111.252569.

114. Peng HZ, Feldman I, Rauscher FJ. Hetero-oligomerization among the TIF family of RBCC/TRIM domain-containing nuclear cofactors: A potential mechanism for regulating the switch between coactivation and corepression. *J Mol Biol*. 2002;320(3):629-644. doi: 10.1016/S0022-2836(02)00477-1.

115. Peng HZ, Begg GE, Schultz DC, et al. Reconstitution of the KRAB-KAP-1 repressor complex: A model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein-protein interactions. *J Mol Biol*. 2000;295(5):1139-1162. doi: 10.1006/jmbi.1999.3402.

116. Venturini L, You J, Stadler M, et al. TIF1 gamma, a novel member of the transcriptional intermediary factor 1 family. *Oncogene*. 1999;18(5):1209-1217. doi: 10.1038/sj.onc.1202655.

117. Lechner M, Begg G, Speicher D, Rauscher F. Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: Direct chromoshadow domain-KAP-1 corepressor interaction is essential. *Mol Cell Biol*. 2000;20(17):6449-6465. doi: 10.1128/MCB.20.17.6449-6465.2000.

118. Ryan RF, Schultz DC, Ayyanathan K, et al. KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: A potential role for kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol Cell Biol*. 1999;19(6):4366-4378.

119. Schultz D, Friedman J, Rauscher F. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: The PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the mi-2 alpha subunit of NuRD. *Genes Dev*. 2001;15(4):428-443. doi: 10.1101/gad.869501.

120. Zeng L, Zhou MM. Bromodomain: An acetyl-lysine binding domain. *FEBS Lett*. 2002;513(1):124-128. doi: 10.1016/S0014-5793(01)03309-9.

121. Ivanov A, Peng H, Yurchenko V, et al. PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain which is required for gene silencing. *Mol Biol Cell*. 2006;17.

122. Szulc J, Wiznerowicz M, Sauvain MO, Trono D, Aebischer P. A versatile tool for conditional gene expression and knockdown. *Nature Methods*. 2006;3(2):109-116. doi: 10.1038/NMETH846.

123. Wiznerowicz M, Jakobsson J, Szulc J, et al. The kruppel-associated box repressor domain can trigger de novo promoter methylation during mouse early embryogenesis. *J Biol Chem*. 2007;282(47):34535-34541. doi: 10.1074/jbc.M705898200.

124. Groner AC, Meylan S, Ciuffi A, et al. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *Plos Genetics*. 2010;6(3):e1000869. doi: 10.1371/journal.pgen.1000869.

125. Rowe HM, Jakobsson J, Mesnard D, et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature*. 2010;463(7278):237-240. doi: 10.1038/nature08674.

126. Wolf D, Goff SP. Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature*. 2009;458(7242):1201-1204. doi: 10.1038/nature07844.

127. Quenneville S, Turelli P, Bojkowska K, et al. The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development. *Cell Reports*. 2012;2(4):766-773. doi: 10.1016/j.celrep.2012.08.043.

128. Rowe HM, Friedli M, Offner S, et al. De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. *Development*. 2013;140(3):519-529. doi: 10.1242/dev.087585.

129. Li X, Ito M, Zhou F, et al. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Developmental Cell*. 2008;15(4):547-557. doi: 10.1016/j.devcel.2008.08.014.

130. Mackay DJG, Callaway JLA, Marks SM, et al. Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. *Nat Genet*. 2008;40(8):949-951. doi: 10.1038/ng.187.

131. Quenneville S, Verde G, Corsinotti A, et al. In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol Cell*. 2011;44(3):361-372. doi: 10.1016/j.molcel.2011.08.032.

132. Messerschmidt DM, de Vries W, Ito M, Solter D, Ferguson-Smith A, Knowles BB. Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science*. 2012;335(6075):1499-1502. doi: 10.1126/science.1216154.

133. Corsinotti A, Kapopoulou A, Gubelmann C, et al. Global and stage specific patterns of kruppel-associated-box zinc finger protein gene expression in murine early embryonic cells. *Plos One*. 2013;8(2):e56721. doi: 10.1371/journal.pone.0056721.

134. Hu G, Kim J, Xu Q, Leng Y, Orkin SH, Elledge SJ. A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev*. 2009;23(7):837-848. doi: 10.1101/gad.1769609.

135. Cammas F, Oulad-Abdelghani M, Vonesch JL, Huss-Garcia Y, Chambon P, Losson R. Cell differentiation induces TIF1 beta association with centromeric heterochromatin via an HP1 interaction. *J Cell Sci*. 2002;115(17):3439-3448.

136. Cammas F, Herzog M, Lerouge T, Chambon P, Losson R. Association of the transcriptional corepressor TIF1 beta with heterochromatin protein 1 (HP1): An essential role for progression through differentiation. *Genes Dev*. 2004;18(17):2147-2160. doi: 10.1101/gad.302904.

137. Tan X, Xu X, Elkenani M, et al. Zfp819, a novel KRAB-zinc finger protein, interacts with KAP1 and functions in genomic integrity maintenance of mouse embryonic stem cells. *Stem Cell Research*. 2013;11(3):1045-1059. doi: 10.1016/j.scr.2013.07.006.

138. Wang Z, Kueh JLL, Teh CH, et al. Zfp206 is a transcription factor that controls pluripotency of embryonic stem cells. *Stem Cells*. 2007;25(9):2173-2182. doi: 10.1634/stemcells.2007-0085.

139. Zhang J, Gao W, Yang H, Zhang B, Zhu Z, Xue Y. Screening for genes essential for mouse embryonic stem cell self-renewal using a subtractive RNA interference library. *Stem Cells*. 2006;24(12):2661-2668. doi: 10.1634/stemcells.2006-0017.

140. Nishii T, Oikawa Y, Ishida Y, Kawaichi M, Matsuda E. CtBP-interacting BTB zinc finger protein (CIBZ) promotes proliferation and G1/S transition in embryonic stem cells via nanog. *J Biol Chem*. 2012;287(15):12417-12424. doi: 10.1074/jbc.M111.333856.

141. Garcia-Garcia MJ, Shibata M, Anderson KV. Chato, a KRAB zinc-finger protein, regulates convergent extension in the mouse embryo. *Development*. 2008;135(18):3053-3062. doi: 10.1242/dev.022897.

142. Cammas F, Mark M, Dolle P, Dierich A, Chambon P, Losson R. Mice lacking the transcriptional corepressor TIF1 beta are defective in early postimplantation development. *Development*. 2000;127(13):2955-2963.

143. Weber P, Cammas F, Gerard C, et al. Germ cell expression of the transcriptional co-repressor TIF1 beta is required for the maintenance of spermatogenesis in the mouse. *Development*. 2002;129(10):2329-2337.

144. Krebs CJ, Larkins LK, Price R, Tullis KM, Miller RD, Robins DM. Regulator of sex-limitation (rsl) encodes a pair of KRAB zinc-finger genes that control sexually dimorihic liver gene expression. *Genes Dev*. 2003;17(21):2664-2674. doi: 10.1101/gad.1135703.

145. de Sio FRS, Massacand J, Barde I, et al. KAP1 regulates gene networks controlling mouse B-lymphoid cell differentiation and function. *Blood*. 2012;119(20):4675-4685. doi: 10.1182/blood-2011-12-401117.

146. de Sio FRS, Barde I, Offner S, et al. KAP1 regulates gene networks controlling T-cell development and responsiveness. *Faseb Journal*. 2012;26(11):4561-4575. doi: 10.1096/fj.12-206177.

147. Emerson RO, Thomas JH. Adaptive evolution in zinc finger transcription factors. *Plos Genetics*. 2009;5(1):e1000325. doi: 10.1371/journal.pgen.1000325.

148. Nowick K, Hamilton AT, Zhang H, Stubbs L. Rapid sequence and expression divergence suggest selection for novel function in primate-specific KRAB-ZNF genes. *Mol Biol Evol*. 2010;27(11):2606-2617. doi: 10.1093/molbev/msq157.

149. Lorenz P, Dietmann S, Wilhelm T, et al. The ancient mammalian KRAB zinc finger gene cluster on human chromosome 8q24.3 illustrates principles of C2H2 zinc finger evolution associated with unique expression profiles in human tissues. *BMC Genomics*. 2010;11:206. doi: 10.1186/1471-2164-11-206.

150. Liu H, Chang L, Sun Y, Lu X, Stubbs L. Deep vertebrate roots for mammalian zinc finger transcription factor subfamilies. *Genome Biology and Evolution*. 2014;6(3):510-525. doi: 10.1093/gbe/evu030.

151. Jacobs FMJ, Greenberg D, Ngan Nguyen, et al. An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature*. 2014;516(7530):242-+. doi: 10.1038/nature13760.

152. Thomas JH, Schneider S. Coevolution of retroelements and tandem zinc finger genes. *Genome Res*. 2011;21(11):1800-1812. doi: 10.1101/gr.121749.111.

CHAPTER 1

Identification and Characterization of ZNFO, a Novel Bovine Oocyte-Specific Gene Encoding a KRAB-Containing Zinc Finger Transcription Factor

Jacqelyn M. Hand[1], Lei Wang[1], Kun Zhang[2], George W. Smith[2], and Jianbo Yao[1]

[1]Laboratory of Animal Biotechnology and Genomics, Division of Animal and Nutritional Sciences, West Virginia University, Morgantown, WV 26506.

[2]Laboratory of Mammalian Reproductive Biology and Genomics, Department of Animal Science, Michigan State University, East Lansing, Michigan 48824.

AUTHOR'S CONTRIBUTIONS

JMH designed and performed most of the experiments including cloning and expression analysis of ZNFO, sequence analysis, and IHC. JMH also drafted the manuscripts. LW performed the qPCR. KZ & GWS performed the western blot. JY designed the study, analyzed the deep sequenced bovine oocyte data, and supervised the experimental work.

ABSTRACT

Initially, the maternal genome regulates nearly all aspects of early development in metazoans. Basic biochemical processes that implement early developmental events such as meiotic maturation, fertilization, the first cleavage divisions, and programming the embryonic genome are totally dependent on maternal mRNAs and proteins derived within the oocyte. At the onset of EGA initiation, the destruction of maternal mRNAs begins by maternally encoded products. To date, ZNF proteins expressed specifically in mammalian oocytes have not been reported. RNA sequencing of a bovine oocyte library uncovered a highly abundant transcript that matches an uncharacterized gene in the NCBI database. Through cDNA cloning of the novel *ZNFO* gene, a transcript containing a 2,145 bp open reading frame that codes for a protein of 714 amino acids with a conserved KRAB domain at the N-terminus and nine zinc finger motifs at the C-terminus was identified. *ZNFO* mRNA was readily detectable in fetal ovaries and was undetectable by RT-PCR in somatic tissues including granulosa and theca cells. Real-time PCR analysis revealed *ZNFO* mRNA was highly abundant in GV and MII stage oocytes as well as in pronuclear to 8-cell stage embryos but undetectable in blastocyst stage embryos ($n = 4$ pools of 10 embryos/stage; $P < 0.05$). Immunohistochemical analysis detected ZNFO protein in oocytes throughout folliculogenesis. Based on the well-conserved functions of KRAB-containing ZNF transcription factors and the current spatial and temporal observations of ZNFO, it is suggested that ZNFO may function as a transcriptional regulator during early embryonic developmental events.

INTRODUCTION

Zinc finger (ZNF) genes compose one of the largest protein superfamilies in eukaryotic organisms and uphold an essential role in transcriptional regulation. In particular, the Cys2His2 (C2H2) class of ZNFs dominate approximately 53% (~700) of the transcription factor repertoire of the mammalian genome[1]. Structurally, C2H2 ZNFs are named for the zinc finger motifs, each comprised of 28-30 amino acids, and each stabilized by a zinc ion that coordinates four highly conserved residues, two cysteines and two histidines[2]. The carboxy-terminal portion of C2H2 zinc finger proteins contain from 1 to more than 30 individual zinc finger motifs arranged in a cluster of tandem repeats. Each individual zinc finger motif is defined by the presence of the consensus sequence Φ-X-Cys-$X_{(2-4)}$-Cys-X_3-Φ-X_5-Φ-X_2-His-$X_{(3-4)}$-His, where X represents any amino acid and Φ represents a hydrophobic residue[3]. Transcriptional regulation occurs through sequence-specific DNA binding of these motifs to promoter regions of target genes[4]. Although each zinc finger domain is structurally similar, variations of key amino acid residues at particular sites, as well as zinc finger number, create chemical distinctiveness allowing for a great number of possibilities for DNA recognition[5] and, hence, the variety and presence of ZNFs in nearly all aspects of biological processes[1,3,6,7].

Over one third of C2H2-ZNF proteins contain the highly conserved Krüppel-Associated Box (KRAB) domain[6], making KRAB-ZNFs the single largest group of transcriptional mediators in the genomes of higher organisms. The KRAB domain homology consists of approximately 75 amino acid residues and folds into two amphipathic helices. The observed evolutionary conservation and wide distribution of the KRAB domain lend relevance to the importance of this region in the transcriptional regulatory function of zinc finger proteins. Recognizably, the functional role of the KRAB domain is known as a potent DNA binding-dependent transcriptional repression module[8-10]. By fusing a heterologous DNA-binding domain from the yeast GAL4 protein with the KRAB domain minimal repression module of approximately 45 amino acid residues, this KRAB-A box, was shown to be necessary and sufficient for transcriptional repression[8-10]. Furthermore, substitutions for these conserved residues abolish repression[8].

In an effort to characterize the bovine oocyte transcriptome in search of oocyte-specific factors essential for the regulation of folliculogenesis and early embryonic development in cattle, a bovine oocyte cDNA library was previously constructed[11]. Analysis of expressed sequence tag (EST) data from this library identified a novel transcript that matches an uncharacterized KRAB-containing zinc finger gene and is explicitly expressed in the bovine oocyte. Based on the analysis of highly conserved structural domain functions within this gene and the fetal ovary-specific expression, it was hypothesized that this novel C2H2 KRAB-containing zinc finger has distinct spatial and temporal expression in follicular development and the development of the early bovine embryo befitting of a maternal-effect gene. The objectives of this study were to 1) clone the novel C2H2 KRAB-containing zinc finger protein (ZNFO) and 2) determine the spatial and temporal expression of *ZNFO* mRNA and protein during oocyte maturation and early embryonic development. To date, ZNF proteins specifically expressed in mammalian oocytes have not been reported. ZNFO presents the first KRAB-ZNF protein identified exclusively in the oocyte in a mammalian species.

MATERIALS AND METHODS

Tissue Collection

Bovine tissue samples, including adult lung, spleen, stomach, brain, muscle, kidney, liver, heart, intestine, ovary, fetal testis, and fetal ovaries were collected from a local slaughterhouse. Age of fetuses from which fetal ovaries were collected was estimated by measuring crown-rump length[12]. Granulosa and theca cells were isolated from antral follicles as described by a previously established method[13]. Briefly, the theca and granulosa layers were separated from the follicles, and each other, by first carefully cutting the majority of tissue away from a follicle, freezing the follicle in liquid nitrogen, and then making a three-quarters cut around the circumference of the follicle. The outer-theca layer of tissue was then removed by light peeling with forceps, as the granulosa layer had adhered to the still-frozen follicular fluid core. All samples were frozen in liquid nitrogen following collection and stored at -80°C until use.

RNA Isolation, cDNA Synthesis, and RT-PCR Analysis

Total RNA was extracted using Tri-reagent (Ambion, Inc., Austin, TX) according to the manufacturer's instructions. Isolated total RNA was treated with TURBO[TM] DNase I (Ambion) before cDNA synthesis. Approximately 2 µg of DNase-treated total RNA was used for first strand cDNA synthesis in a 20 µl reaction including Oligo $(dT)_{18}$ primer and SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA). Concentrations of isolated RNA were determined by measuring absorbance at 260 nm. Purity of RNA was determined by calculating the ratio of absorbance at 260 nm and 280 nm, and integrity of RNA was determined by agarose gel electrophoresis. The cDNA was used as a template for PCR amplification of *ZNFO* mRNA fragments using gene-specific primers (Table 1). The RT-PCR was performed by denaturation at 95°C for 3 min followed by 35 cycles of 95°C for 30 sec, 60°C for 30 sec, and 72°C for 90 sec and final extension at 72°C for 10 min. The amplified products were separated through a 1% agarose gel containing RGB. Amplification of cDNA for bovine ribosomal protein L19 (*RPL19*) was used as a positive control for RNA quality and RT.

Cloning of ZNFO cDNA

Deep sequencing data of a bovine oocyte library revealed a match of a bovine genomic sequence that was predicted to encode a hypothetical protein corresponding to *ZNFO* in the National Center for Biotechnology Information (NCBI) Database (XM_001789794.1). Based on this predicted sequence, primers were designed (Table 1) to amplify the 5' (untranslated region) UTR end and the coding sequence through the putative translation stop codon. The amplified cDNA fragments (525 bp and 2099 bp) were cloned into pGEM-Teasy vector (Promega, Madison, WI) and sequenced. To obtain the 3' end of the cDNA sequence, 3' rapid amplification of cDNA ends (RACE) was performed using the second generation 5'/3' RACE kit (Roche Diagnostics, Indianapolis, IN) following the manufacturer's protocol. A PCR reaction was performed using a gene specific primer (Table 1). The final RACE products were cloned into pGEM-Teasy vector and sequenced. All three overlapping gene fragments were used to create a complete *ZNFO* contig (Figure 1A).

Generation of Anti-ZNFO Antibody

The ZNFO antibody was prepared commercially by GenScript Corporation (Piscataway, NJ). Polyclonal antiserum against ZNFO was raised by immunizing rabbits with a 15-amino acid synthetic peptide (KRNQGRESNREKPIC) of the predicted amino acid sequence of ZNFO. Antisera from the third bleed was used in this study.

Immunohistochemistry

Paraffin-embedded sections of calf ovary (2 mo) tissue were subjected to immunohistochemical analysis using Ultra-Sensitive ABC Peroxidase Staining kit (Pierce Chemical Co., Rockford, IL) according to the manufacturer's instructions. Briefly, approximately 12-μm serial sections were prepared and mounted onto positively-charged slides. The paraffin sections were deparaffinized in xylene and then rehydrated in graded alcohol. Antigen retrieval was performed by boiling the sections in antigen retrieval solution (10 mM Tris base, 1 mM EDTA, 0.05% Tween 20, pH 9.0) for 15 min, and allowing the slides to cool for 30 min at room temperature. After treatment with 0.3% hydrogen peroxide in methanol to eliminate endogenous

peroxidase activity, the sections were serially incubated with blocking buffer (normal goat serum in PBS) for 30 min at room temperature, rabbit polyclonal anti-ZNFO antibody (GenScript) at 5 µg/ml in blocking buffer at 4°C overnight, biotinylated anti-rabbit IgG for 1 h, followed by incubation with avidin-biotin-peroxidase complex reagent for 1 h at room temperature. Intervening PBS washes were performed after each antibody incubation. The sections were developed using a metal-enhanced DAB Substrate Kit (Pierce) for 1 min and were then counterstained with VECTOR Hematoxylin QS (Vector Laboratories, Burlingame, CA) and mounted with Cytoseal XYL (Thermo Fisher Scientific, Waltham, MA). Negative control sections were incubated in the absence of anti-ZNFO antibody.

Quantitative Real-Time RT-PCR

Oocyte and embryo samples analyzed for mRNA expression included germinal vesicle (GV)- and metaphase II (MII)-stage oocytes and pronuclear, two-cell, four-cell, eight-cell, 16-cell, and morula- and blastocyst-stage embryos (n = 5 pools of 10 embryos) generated by *in vitro* fertilization of abattoir-derived oocytes as previously described[14]. Quantitative real-time PCR set-up and standardization conditions were carried out as previously described[15]. Before RNA extraction, each sample was spiked with 250 fg of green fluorescent protein (*GFP*) synthetic RNA (polyadenylated) as an exogenous control. Copies of *GFP* RNA in each pool were determined using standard curves constructed from the plasmid pcDNA3-EGFP (Addgene, Cambridge, MA). The quantity of *ZNFO* mRNA was normalized relative to the quantity of *GFP* measured in each sample, and differences in normalized data across developmental stages were determined by one-way ANOVA using the statistical analysis package, R. Individual mean comparisons were performed using Fisher least significant difference (LSD) method. Differences of $P < 0.05$ were considered significant.

Western Blot Analysis

Protein lysate samples (10 µg/well) were separated on a 4-20% gradient ready gel (Bio-Rad, Hercules, CA), and electrophoresis was run in 1X Tris/Glycine/SDS running buffer for 2 h. Proteins were transferred onto a Immun-Blot PVDF membrane (Bio-Rad) in 1X transfer buffer (Tris/Glycine/SDS/methanol) for 1 h, 10 min. Following transfer and blocking in 5% nonfat dry

milk in PBS containing 0.1% Tween-20 (PBST) for two hours, the membrane was then incubated in 1 ug/mL ZNFO primary antibody solution in blocking buffer overnight at 4°C. After three washes with PBST, immunoreactive proteins were visualized by using a chemiluminescent horseradish peroxidase detection system (Genotech, St. Louis, MO). Specificity of the antibody was validated by Western blot analysis using purified GST-ZNFO protein (data not shown).

RESULTS

Cloning and Characterization of the Bovine **ZNFO** *Gene*

Deep sequencing analysis of the transcriptome of the bovine oocyte revealed a highly abundant transcript that matches an uncharacterized gene (LOC100141212) in the NCBI Database. Analysis of the EST data from the cattle gene index (TGI Gene Indices) showed that the transcript is represented by ESTs exclusively derived from 2-cell embryos. Through cDNA cloning a 3,595 bp transcript containing a 2,145 bp open reading frame (ORF) was obtained. The ORF and a 525 bp 5' UTR were amplified from bovine fetal ovaries (Figure 1A). Gene-specific RACE primers were designed based on the obtained sequence, and 3' RACE was performed to extend the 3'-end of the cDNA sequence (Figure 1A). All obtained fragments were cloned into pGEM-T-easy and sequenced. The complete assembled *ZNFO* cDNA sequence has been deposited in GenBank (accession number: KJ710495.1).

Analysis of tissue distribution by RT-PCR revealed that the novel transcript (*ZNFO*) is undetectable in all somatic tissues analyzed, as well as in the fetal testis germ cell counterpart, but is found specifically expressed in fetal ovaries (Fig. 2). Further RT-PCR analysis showed that *ZNFO* is not expressed in granulosa and theca cells (Fig. 3A), suggesting that *ZNFO* expression in the ovary is oocyte-specific. Expression of *ZNFO* mRNA is highly abundant in GV- and MII-stage oocytes and 2-cell through 16-cell stage embryos but is completely undetectable in morula and blastocyst stage embryos (Fig. 3B), indicating that the embryonic genome does not express this gene. Analysis of ZNFO mRNA expression in fetal ovaries of different developmental stages during gestation revealed that ZNFO mRNA can be detected in fetal ovaries as early as day 90 of gestation (Fig. 3C), a period when primordial follicles are emerging in cattle[16]. The expression of ZNFO mRNA increases steadily in fetal ovaries (day 160 and day 230) during development, suggesting a role of this gene in supporting development of primary and secondary follicles which are formed around day 140 and 210 of gestation, respectively[16]. These results indicate ZNFO is a maternal transcript abundantly present in oocytes and early embryos prior to embryonic genome activation.

Genomic Structure and Organization of ZNFO

The ZNFO ORF encodes a protein of 714 amino acids. A search of the Pfam protein database[17] in combination with visual inspection of the protein sequence revealed that ZNFO contains a conserved Krüppel-associated box (KRAB) domain at the N-terminus and nine zinc finger motifs at the C-terminus (Fig. 4A and 4B). The KRAB domain consists of 41 amino acid residues that match the minimal repression module of the conserved KRAB-A box, which is necessary for transcriptional regulation[8-10]. Each zinc finger motif fits the consensus sequence Φ-X-Cys-$X_{(2-4)}$-Cys-X_3-Φ-X_5-Φ-X_2-His-$X_{(3-4)}$-His definitive of C2H2 ZNFs[3], and the remaining seven amino acids in between each finger contain the five canonical linker residues[18,19].

Using the SUMOsp 2.0 program[20], the protein was predicted to contain two putative sumoylation sites (Lysines 13 and 260), indicating that the protein might be sumoylated, a post-translational modification event that plays a role in various cellular processes. The novel protein shares 85% and 96% sequence identity with predicted bovine ZNF708 isoform 1 and ZNF726 isoform 2, and ZNF726 isoform 4 proteins, respectively. No orthologs of this protein were found in other mammalian genomes through extensive NCBI database search. This suggests that ZNFO is a species-specific oocyte-specific gene. BLAST search of the assembled bovine genome sequence in the NCBI database using the ZNFO cDNA sequence identified an annotated bovine chromosome 18 genomic contig sequence (NW_003081470) containing the ZNFO gene which spans over 11.9 kb. Alignment of the cDNA sequence to the genomic sequence using the Splign program[21] revealed that the *ZNFO* gene contains 4 exons separated by 3 introns (Figure 1B), and all splice sites are in agreement with consensus sequences (GT-AG rule).

Analysis of the 5'flanking sequence of the ZNFO gene using TESS program (University of Pennsylvania) identified a number of putative transcription factor binding sites, such as RAR, ERα, AP-1 and Oct4 (Fig. 5). In addition, a putative E-box was identified by visual inspection, an element known to be necessary for oocyte-specific gene expression[22]. The transcription start site (TSS) was predicted using promoter prediction algorithms Tfsearch and Softberry TSSW. Twenty CpG sites from ~1600 bp flanking the predicted TSS region were identified (Fig. 5). The 3' UTR was found to contain two nuclear poly(A) signals (AAUAAA) and three cytoplasmic

polyadenylation elements ($U_5A_{1-2}U$), which are required for translational activation of maternally derived mRNAs[23].

Expression of Bovine ZNFO *mRNA and Protein during Oocyte Maturation and Early Embryonic Development*

Temporal expression of bovine *ZNFO* mRNA during oocyte maturation (GV- and MII-stage) and early embryonic development (pronuclear, 2-cell, 4-cell, 8-cell, 16-cell, morula, and blastocyst stage) was examined by quantitative real-time PCR (qPCR). The *ZNFO* transcript was abundant in GV- and MII-stage oocytes, with peak expression in MII-stage oocytes, and was also highly expressed in pronuclear stage embryos through 8-cell stage embryos (Figure 6A; $P < 0.05$). By the time of the 16-cell stage, *ZNFO* mRNA markedly declined and was barely detectable in embryos collected at morula- and blastocyst-stages (Figure 6A; n = 4 pools of 10 embryos/stage; $P < 0.05$). Immature (pre-LH surge, specifically antral follicle-housed) and mature (post-LH surge, specifically GV- and MII-stage) oocytes were examined for the presence of ZNFO protein by Western blot analysis. The immunoreactive ZNFO protein demonstrated to be approximately 84 kDa in size (Figure 6B) and was abundant in both immature and mature oocytes with the appearance of slightly less expression in zona pellucida-free immature oocytes.

Immunohistochemical localization of ZNFO protein within calf ovary sections revealed that ZNFO protein is present in oocytes of growing follicles at the primordial (Figure 7, panel A; single layer of flattened granulosal cells), primary (Figure 7, panel B; single layer of cuboidal granulosal cells), and secondary (Figure 7, panel C; multiple layers of cuboidal granulosal cells) follicle stages through antral (Figure 7, panels D and E; early antrum formation and preovulatory, respectively) follicles stages. Preovulatory follicles typically displayed some degree of immunoreactivity in the adjacent cumulus cells. No staining was detected when sections were incubated in the absence of the primary antibody (Figure 7, panel F). The expression pattern of *ZNFO* mRNA and protein during folliculogenesis and early embryogenesis is similar to many other known bovine maternal-effects genes necessary for early embryonic development[24-26] suggesting that *ZNFO* is of maternal origin.

DISCUSSION

Results of the present studies demonstrated the cloning and functional characterization of a novel oocyte-specific gene in cattle. The predicted protein encoded by this novel gene contains a highly conserved KRAB domain and nine tandem zinc finger motifs that fit the C2H2 class of zinc finger proteins, designating this novel protein as a new member of the C2H2 KRAB-containing zinc finger family. The most striking characteristic of the C2H2 ZNF family of transcription factors is their ability to repress transcription, which probably underlies the many biological processes in which they are implicated[4,6]. The results herein indicate that expression of *ZNFO* is oocyte-specific, which, to our knowledge, is the first report of a KRAB-containing zinc finger gene found strictly in the oocyte of any species.

Several factors exist that influence the ability of an oocyte to become competent including oocyte origin, follicle health, hormonal stimulation and communication between the oocyte and surrounding cumulus cells. The ability of an oocyte to become developmentally competent requires that it gain the capacity to resume meiosis, cleave following fertilization, develop to the blastocyst stage and implant, and develop to term in good health[27-30]. Oocytes gradually and sequentially acquire competence throughout folliculogenesis by synthesizing and accumulating transcripts and proteins that are necessary for successful follicular development, fertilization and subsequent embryogenesis[31]. The early stages are critical because many oocyte-specific genes are transcribed during the primordial to primary follicle transition and continue to be expressed throughout folliculogenesis. Many of these factors are not only found during follicular development but also prior to that in embryonic gonad formation and germline establishment. During prenatal development, the ovaries of mammals are endowed with a finite population of germ cells. By 90d of gestation, a time when the first primordial follicles develop following germline establishment in cattle[16], *ZNFO* was detectable. Several known oocyte-specific transcription factors have been shown to be essential for normal development of germ cells and surrounding somatic cells during mammalian folliculogenesis[32-34].

Nobox, preferentially expressed in germ cells as early as E15.5 in mice[35] and d100 in cattle[24] is present throughout folliculogenesis, including in germ cell nests, primordial, growing

60

and antral follicles[35]. Ovaries lacking *Nobox* formed apparently normal follicles up to the primordial stage but transition beyond to the primary stage was severely compromised. In addition, the knockout ovaries displayed an accelerated loss of oocytes with only a few degenerated oocytes remaining 14 days after birth[35]. Further, *Nobox* knockout ovaries under gene expression analysis revealed a downregulation of mRNA transcripts for genes preferentially expressed in oocytes, such as *Oct4*, *Gdf9*, *Bmp15*, *Rfpl4*, *H1oo*, *Zar1*, *Dnmt1*, and *Mos*, whereas genes important in germ cell migration (*Kitl* and *Kit*), apoptosis (*Bcl2*, *Bax*, *Bcl2l2*, and *Casp2*), and meiosis (*Mlh1* and *Msh5*) transcripts were unaffected and expressed as wild-type *Nobox* ovaries[35]. *Figla*, like *Nobox*, is required for early folliculogenesis as *Figla* knockouts display a block of primordial follicle development and rapid loss of oocytes shortly following[36]. Expression of the zona pellucida genes *Zp1*, *2*, and *3* are diminished in *Figla* depleted ovaries[36]; however, other important genes such as *Gdf9*, *Bmp15*, *Kit*, *Kitl*, *Cx43* (connexin 43), and *Fgf8* (fibroblast growth factor 8) are unchanged.

As essential factors for folliculogenesis, both *Nobox* and *Figla* highlight the major roles representative of many germ-specific transcription factors that are critical for successful completion of folliculogenesis and subsequent early embryonic development. Specifically, genes such as these are known as maternal-effect genes because they are maternal transcripts that accumulate and are stored in the oocyte during oogenesis and are required for successful folliculogenesis and germ cell maturation, and also for subsequent activation of the embryonic genome and early cleavage events post-fertilization[37]. Although the majority of knowledge obtained for these factors comes from mouse studies, it may still be relevant to oocyte competence in cattle, although, several substantial gaps in knowledge remain. Many genes involved in important cellular functions of follicular- and oocyte-development remain unaltered by the absence of the known key factors (i.e. *Nobox*-depleted ovaries did not interfere with meiotic-gene transcripts)[35], suggesting that other potentially novel oocyte-specific factors are necessary for oocyte survival. *ZNFO* clearly fits within the definition of a maternal-effect gene based on the observed expression profile. Abundant and continuous expression of *ZNFO* was observed from the first appearance of primordial follicles during embryonic development through folliculogenesis to the pre-ovulatory oocyte in the adult ovary. This illustrates the need for accumulating maternal stores of *ZNFO* transcripts and also suggests a possible role of ZNFO in ensuring proper follicular

development. Depletion of *ZNFO* during folliculogenesis is necessary to determine which stages of follicular development are specifically directed by this novel factor. However, because knockout studies are primarily done with rodent species for obvious reasons, and *ZNFO* is only found in the bovine genome, this type of study becomes particularly challenging and was not performed here.

Following expression in pre-ovulatory oocytes, *ZNFO* displayed a continued expression throughout early development of the bovine embryo. The ability of the bovine embryo to reach the blastocyst stage and ultimately develop into a healthy offspring is a complex and highly regulated process. Maternal transcripts are replaced and are degraded during different stages of embryogenesis by the embryonic genome, which is transcriptionally inactive before maternal-zygotic transition, and is activated at the onset of maternal-zygotic transition[38]. Therefore, the transition from maternal to embryonic control of development is characterized by degradation of maternal transcripts and proteins, sensitivity to transcriptional inhibitors (e.g. α-amanitin), and a dramatic increase in transcriptional activity from the embryonic genome[39]. This crucial transition occurs during the first few post-fertilization cell cycles in a species dependent manner. Embryonic genome activation occurs at roughly the 2-cell stage in mice, 4- to 8-cell stage in humans, and 8- to 16-cell stage in cattle[40,41]. Accumulating experimental evidence, , including relatively recent data in cattle[24-26], indicates that maternal effect genes are key regulators of folliculogenesis and subsequent early cleavage events post fertilization. The results presented here show that *ZNFO* is specifically expressed in oocytes and early embryos prior to- and during the onset of embryonic genome activation, displaying maximum expression from GV to 8-cell stages and then rapidly declining to near non-existence by morula and blastocyst stages. This specific expression pattern is similar to several oocyte-expressed genes that have been reported to be essential for initial stages of embryonic development.

Oocytes expressing the maternal-effect genes [maternal antigen that embryos require (*Mater*), Zygote Arrest 1 (*Zar1*), factor located in oocytes permitting embryonic development (*Floped*), developmental pluripotency associated 3 (*Dppa3*), octamer binding transcription factor 4 (*Oct4*) and nucleoplasmin 2 (*Npm2*)] are each found expressed in growing oocytes throughout folliculogenesis and into early embryo development. *Mater*, an oocyte antigen involved in

62

autoimmune ovarian failure, is present in early cleavage stage embryos until the blastocyst stage; however, MATER-deficient embryos become arrested at the 2-cell stage[42]. *Zar1* and *Floped* transcripts are present until the one-cell stage in wild-type embryos, but ZAR1 and FLOPED mutants are infertile and show defects in embryogenesis beginning at the one-cell stage [43,44]. Additionally, *Zar1* null zygotes failed to complete syngamy, leaving two separate haploid genomes incompletely fertilized[43]. FLOPED mutants display unequal sized blastomeres with weakened contact regions[44]. NPM2 is an oocyte-specific nuclear chaperone that mediates the assembly of nucleosomes, but *Npm2* knockout mice are subfertile or intertile[45]. Both Dppa3 and Oct4 are found throughout oogenesis, in preimplantation embryos and pluripotent stem cells. Embryos without maternally derived Dppa3 have compromised early embryonic development and rarely reach the blastocyst stage[46]. Oct4-deficient embryos are able to develop to the blastocyst stage, but the inner cell mass cells lack pluripotency leading to failure of expansion of trophectoderm layer and subsequent peri-implantation lethality[47]. Although mutation to any of the aforementioned factors maintains normal ovarian development, folliculogenesis, and fertilizable eggs, the observed effects on early cleavage events indicate that these are key maternal genes whose expression is required for normal early embryonic development. Because the patterned expression of *ZNFO* may be categorized with critical factors such as these, ZNFO is suspected to be involved in early embryonic development as a key factor. The functional role of ZNFO, however, warrants further examination.

As an apparent maternal-effect gene, it is interesting to note the OCT4 binding site within the promoter region of ZNFO. OCT4 is a transcription factor well known for its role in embryonic stem cell self-renewal and pluripotency[48]. Microinjection of *Oct4* antisense morpholino oligonucleotides into one-cell mouse embryos revealed that maternal *Oct4* is also necessary for embryonic genome activation probably by regulating genes that encode transcriptional and post-transcriptional factors[49]. *Oct4* has been confirmed in the bovine oocyte and increases expression through early embryonic development with a sharp increase following zygotic genome activation and again after compaction[50]. Considered a master regulator of early embryonic development, it would be interesting to elucidate the targets of OCT4, as that remains unknown in cattle and ZNFO seems a likely interacting partner. Also structurally important to ZNFO is the KRAB domain.

ZNFO carries the classical KRAB-A box only (KRAB A). As one of three subfamilies of KRAB domains, each effectively shown to repress transcription through interaction with <u>K</u>RAB-<u>a</u>ssociated <u>protein 1</u> (KAP1)[8,9,51-54]. The KRAB-A box domain is the minimal repression module that is necessary and sufficient for transcriptional repression[8-10]. The combination of the observed Oct4 binding site and KRAB-A box domain lead to the implied role of ZNFO as transcriptional repressor interacting in network with key oocyte-specific factors to regulate embryonic genome activation and other maternal-to-zygotic transition events.

Considering the vast diversity of tissue cell-type localization and functional roles of KRAB-containing zinc finger proteins, it is quite interesting to find that no other KRAB-ZNF has been identified that is tissue exclusive to the oocyte. In similar light, ZNFO does not seem to have any orthologs. The timing established for major EGA demonstrates an occurrence that is generally later in mono-ovulatory species, such as cattle and primates (including human), as compared with poly-ovulatory species such as the mouse, which EGA manifests much sooner. Therefore, the maternal-effect genes required to promote initial cleavage divisions and ensure successful early embryonic development in such mono-ovulatory species may be divergent from those required in the poly-ovulatory species. Further, within monotocous species, a distinct species-specific factor may be part of a common and highly conserved functional network. Oocyte-specific factors required for early embryogenesis in bovine species are poorly understood, and understanding the contribution of such factors to maternal-to-embryonic transition during early embryogenesis in cattle is limited.

Prior to the series of complex events encompassing early embryonic development, the intrinsic quality of the oocyte must first be sufficient for fertilization. Oocytes gradually and sequentially acquire competence throughout folliculogenesis by synthesizing and accumulating transcripts and proteins that are necessary for successful follicular development, fertilization and subsequent embryogenesis[31], as well as by the maturation of meiotic and cytoplasmic components[30]. The successful completion of each of the developmental events are separate and do not ensure the success of the subsequent events, so studies targeting each of the important developmental events, including meiotic maturation, fertilization, and cleavage, are required to understand how maternal-effect genes such as *ZNFO* regulate the maternal-to-zygotic transition.

FIGURES
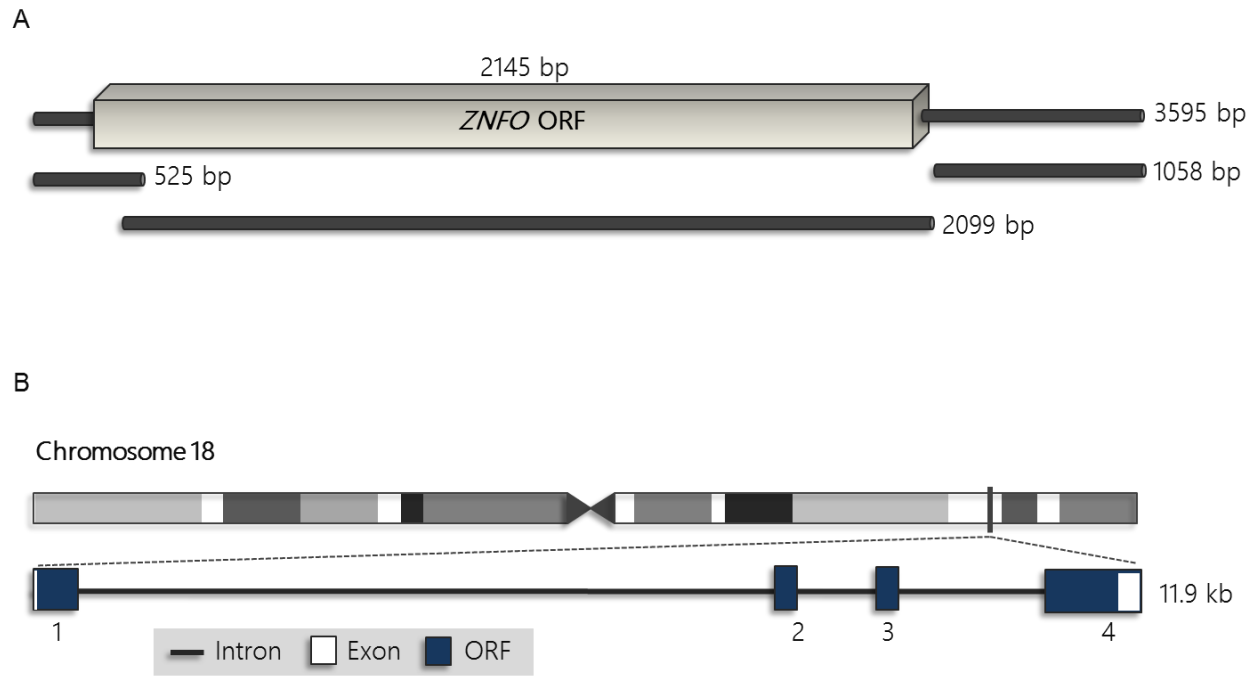
Figure 1. **Cloning strategy and sequence analysis of bovine *ZNFO* gene.** (A) Schematic representation of bovine *ZNFO* cDNA and the cloning strategy used. (B) Schematic representation of bovine *ZNFO* gene structure.
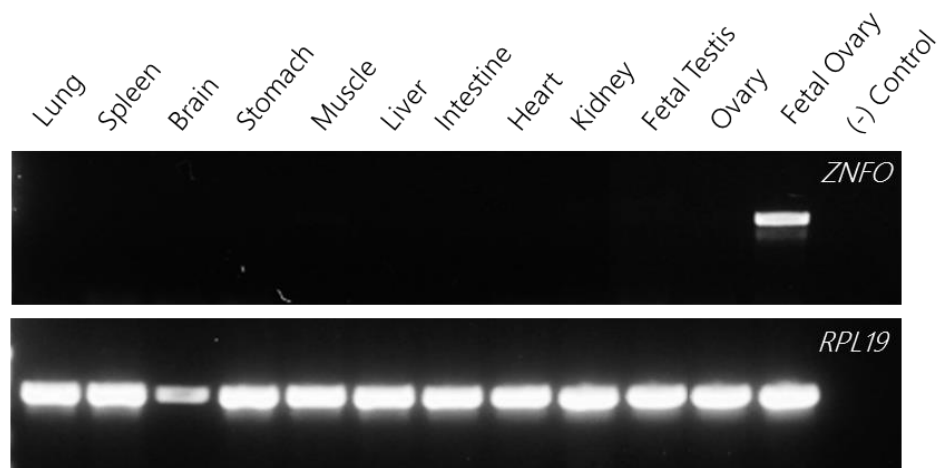
Figure 2. **Tissue distribution of *ZNFO* mRNA.** RT-PCR analysis of the novel *ZNFO* throughout various bovine tissue reveals that *ZNFO* is restricted to oocyte-rich fetal ovaries. Bovine *RPL19* was used as an internal control.
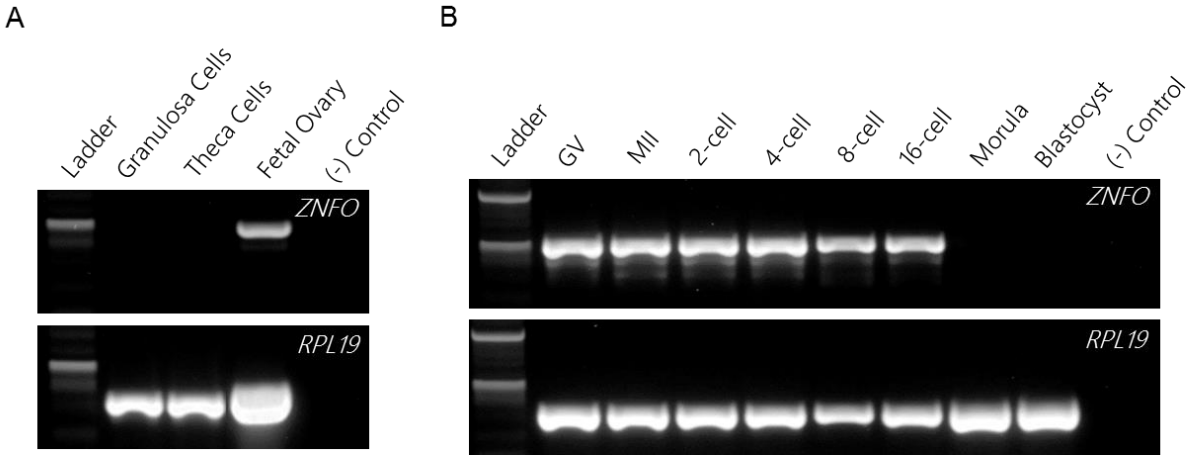
***Figure 3A & 3B*. Analysis of *ZNFO* mRNA expression by RT-PCR.** A. Bovine *ZNFO* mRNA is expressed only in oocyte-rich fetal ovaries and not surrounding granulosal cells or thecal cells. B. Expression of *ZNFO* mRNA in GV oocyte through early embryonic development. *ZNFO* is expressed from GV through 16-cell embryo. Expression of *ZNFO* is diminished in morula and blastocyst. (GV = germinal vesicle; MII = meiosis II). Bovine *RPL19* was used as an internal control.
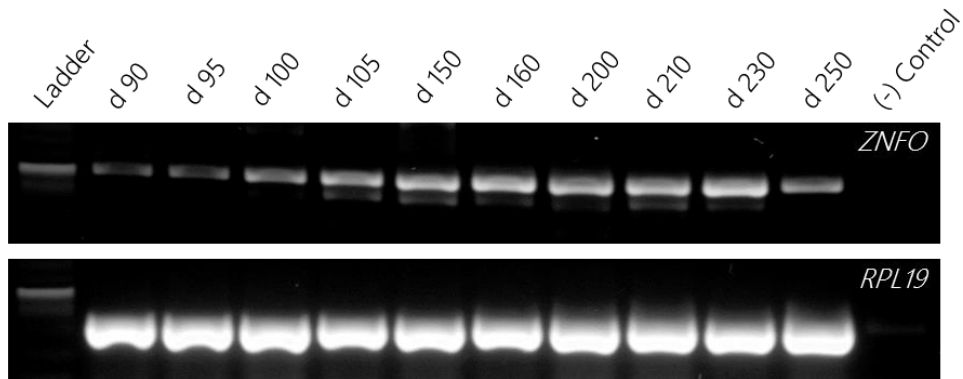
Figure 3C. **Analysis of *ZNFO* mRNA expression by RT-PCR.** C. *ZNFO* mRNA in fetal ovaries from gestational days 90-250. *ZNFO* is expressed in fetal ovaries throughout gestation. Bovine *RPL19* was used as an internal control.

A

KRAB   Zinc Finger Motifs

ZNFO

1   714

B

```
  1  MQRCVCYITSPEKTEQEEEEEMAASQGRLTFQDIAIDFTQEEWECLDLGQRELYKDVMLENYGNLASLG
 71  LVISKPDLVTFLEQMKDPWDTRRMETPAVHPAVSSHDTQGLRTQKPGLEDLFPTAKLGIYERFHLGNLHL
141  TKGWEYMRAYKEQRACYDGQNQIGTDSHNVNITSKRNQGRESNREKPISDDQFQSSPSADKCIFVSKDPH
211  HLLKHTRSLKQNVENLESHLVSTANTHSNNSEHRCLLNIHSNISGNQKLKNEGGNSQYNHFEGSFNKGLL
281  FFNQQLFSPCSKICNVDNNGRDLIQPSLFNTYGGIISVEQLYKRNKMSNALSRSSTRNNYKSIHDGMRSS
351  SCSETGHNVDQDSYLMKQQEHQFSDNSKHNKCKNIFYQSSNLTINTYKSIDIGEKTYNCYDYAKAFNQSS
421  KVIQQQNIQTKQKHYKCNTCGKVFSNSPNLSRHRKIHTGRKCFKCTACGKAFNQSSYLTEHQRIHAGEKP
491  YKCTECGKTFIYCSRVTQHQRIHTGERPYKCTECGKAFNWHLSLTVHQRTHTGEKPYKCKECGKAFISCS
561  HLTRHQRIHTGERPYKCTHCGKAFTRYSPLTQHQRIHTGERPYKCTECGKAFNWRLSLTVHQQTHTGEKS
631  YKCKECGKAFICCSHLTQHQRIHTGERPYKCTDCGKAFSRSSGLSQHQRVHTAGKSQKCKESGKGFHHSH
701  HLTHHQRIHTAEKP  714
```
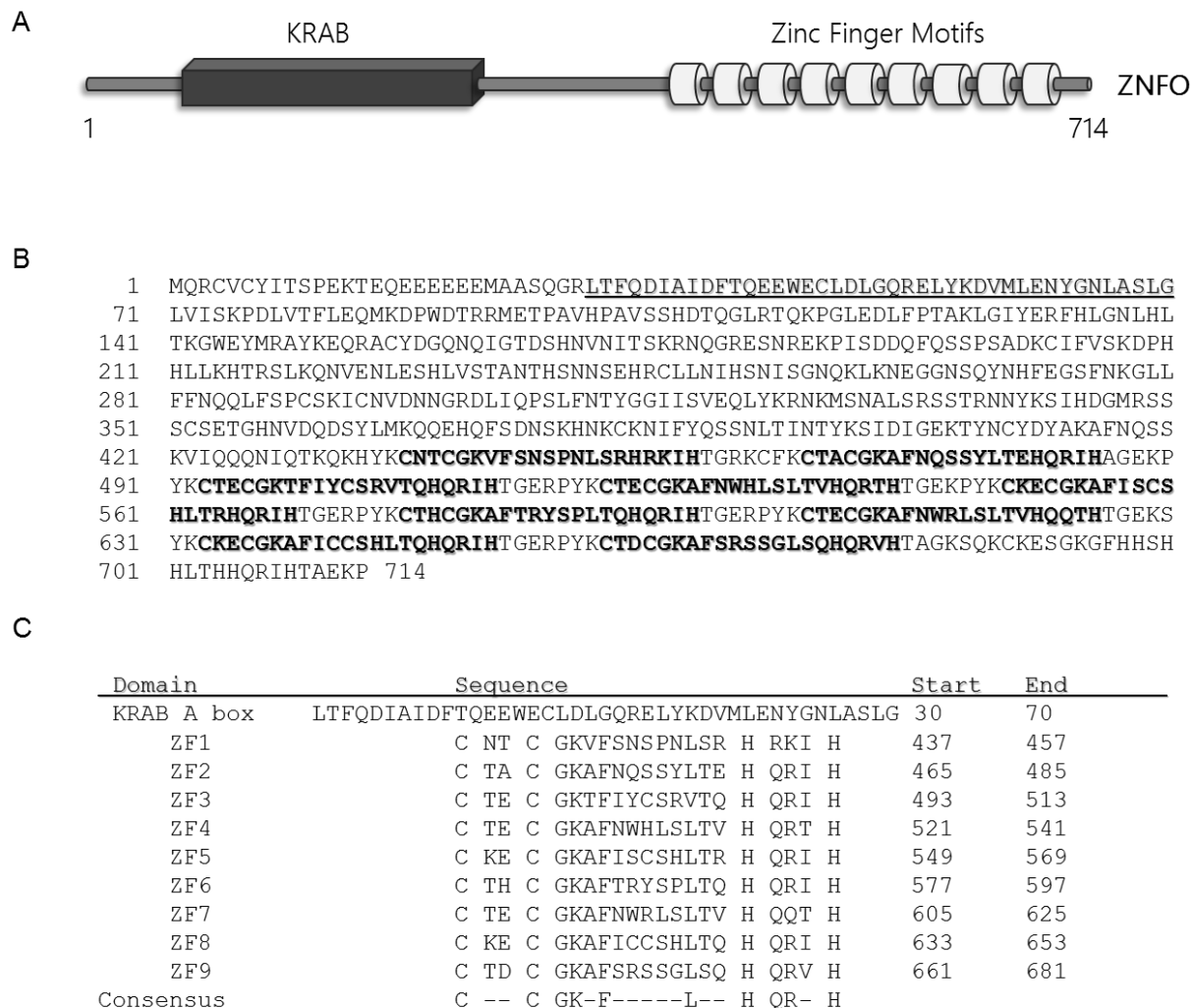
C

Domain	Sequence						Start	End
KRAB A box	LTFQDIAIDFTQEEWECLDLGQRELYKDVMLENYGNLASLG						30	70
ZF1	C	NT	C	GKVFSNSPNLSR	H	RKI	H	437	457
ZF2	C	TA	C	GKAFNQSSYLTE	H	QRI	H	465	485
ZF3	C	TE	C	GKTFIYCSRVTQ	H	QRI	H	493	513
ZF4	C	TE	C	GKAFNWHLSLTV	H	QRT	H	521	541
ZF5	C	KE	C	GKAFISCSHLTR	H	QRI	H	549	569
ZF6	C	TH	C	GKAFTRYSPLTQ	H	QRI	H	577	597
ZF7	C	TE	C	GKAFNWRLSLTV	H	QQT	H	605	625
ZF8	C	KE	C	GKAFICCSHLTQ	H	QRI	H	633	653
ZF9	C	TD	C	GKAFSRSSGLSQ	H	QRV	H	661	681
Consensus	C	--	C	GK-F-----L--	H	QR-	H		

Figure 4. **Structure of ZNFO.** (A) Secondary structure of ZNFO. (B) Primary amino acid sequence of bovine ZNFO. The predicted zinc finger domains are in *bold*, and the KRAB region is *underlined*. Numbers are shown for the amino acid sequence. (C) Alignment of the amino acid sequences of the zinc finger domains. ZNFO possesses nine C2H2 zinc fingers (ZF1–ZF9) conforming to the C2H2 consensus, $CX_{2-4}CX_3(F/C)X_5(F/L)X_2HX_{3-4}H$, in which X represents any amino acid
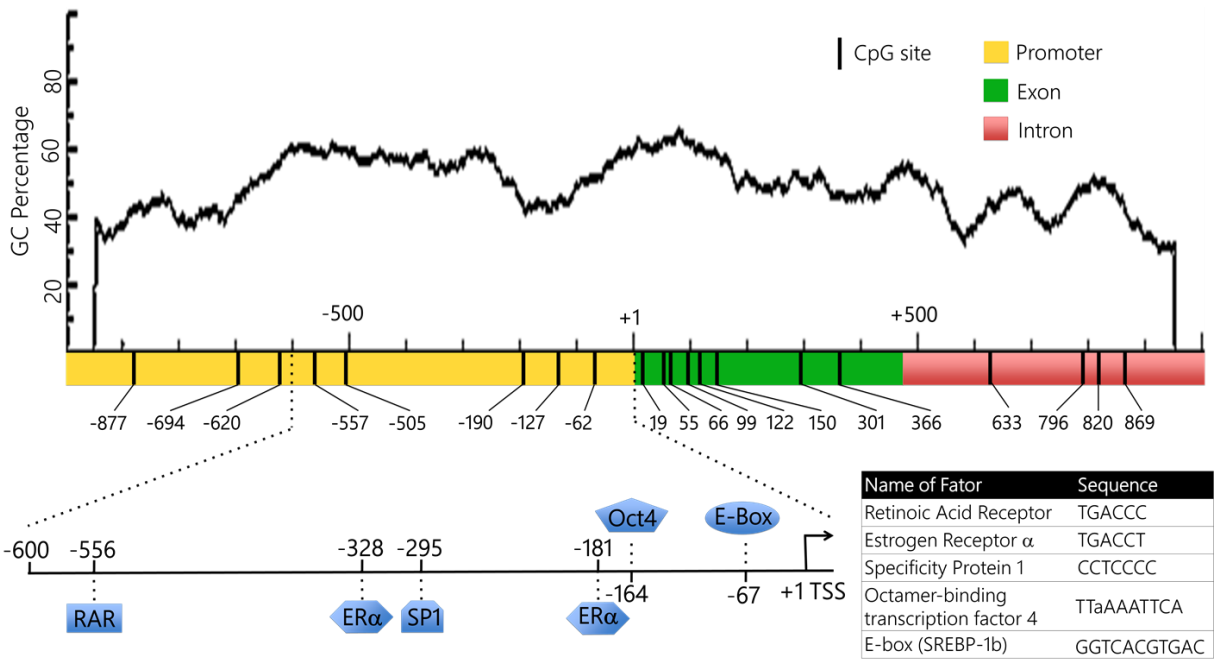
***Figure* 5. ZNFO promoter region.** Location of the transcription factor binding sites and CpG sites in the promoter region of ZNFO.
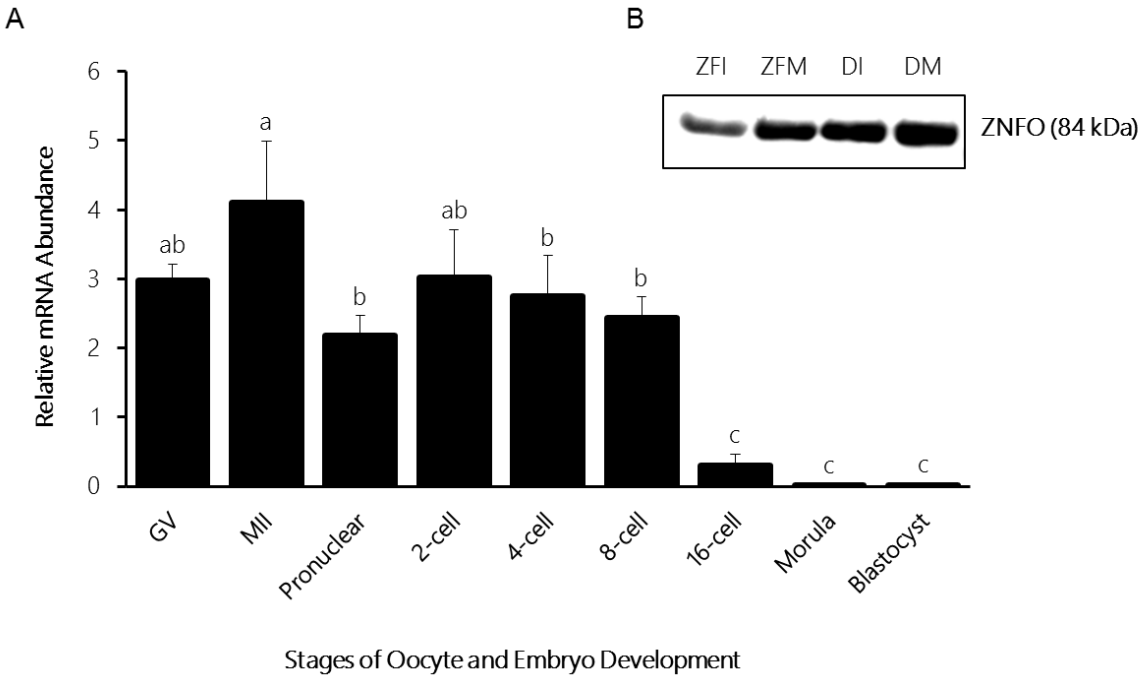
***Figure 6*. Expression of bovine *ZNFO* mRNA and protein during oocyte maturation and early embryonic development.** (A) Relative abundance of *ZNFO* mRNA in bovine oocytes and *in vitro* produced bovine early embryos: GV- and MII-stage oocytes, pronuclear, 2-cell, 4-cell, 8-cell, 16-cell, morula, and blastocyst-stage embryos. ZNFO transcript levels were normalized relative to abundance of exogenous control (*GFP*) RNA and are shown as mean ± SEM (n = 4 pools of 10 embryos/stage). Different letters indicate statistical difference ($P < 0.05$). (B) Analysis of ZNFO protein expression in bovine oocytes by Western blot analysis using an antibody specifically against ZNFO (50 oocytes per lane). ZFI, zona-free immature oocytes; ZFM, zona-free mature oocytes; DI, denuded immature oocytes; DM, denuded mature oocytes.
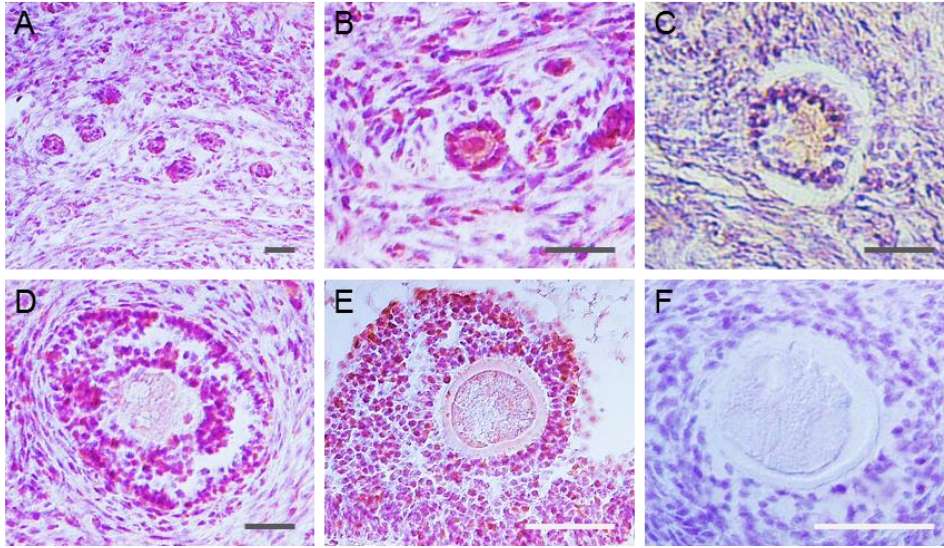
Figure 7. **Intrafollicular localization of ZNFO protein**. Immunohistochemical localization of ZNFO to the oocytes of calf ovaries in primordial (A), primary (B), secondary (C), early antral (D), and antral/pre-ovulatory (E) follicles. No signal staining was observed in oocytes incubated in the absence of anti-ZNFO antibody (F). Gray bars, 20 µM; white bars, 50 µM.

TABLES

Table 1 Primers used in this study

Primer name	Primer sequence (5'-3')	Application
ZNFO-15-F	AGGGAGCATCAGTGTGGTTTG	RT-PCR
ZNFO-539-R	CACTCCTCTTGAGTGAAGTCT	RT-PCR
ZNFO-467-F	AGGAGGAAATGGCAGCTTCTCAGG	RT-PCR
ZNFO-2537-R	CTAGGGTTTCTCTGCAGTATG	RT-PCR
ZNFO-ORF-393-F	ATGCAGAGGTGTGTTTGCTACA	RT-PCR
ZNFO-ORF-2537-R	CTAGGGTTTCTCTGCAGTATG	RT-PCR
ZNFO-Sp4-2557	AAAGCTTGGTCCTAAATATACAAT	3'RACE
ZNFO-849-F	CAGAGAGCATGTTATGATGGACA	qPCR
ZNFO-946-R	TCCCTATTTGATTCACGTCCTTG	qPCR

REFERENCES

1. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: Function, expression and evolution. *Nature Reviews Genetics*. 2009;10(4):252-263. doi: 10.1038/nrg2538.

2. Lee MS, Gippert GP, Soman KV, Case DA, Wright PE. 3-dimensional solution structure of a single zinc finger dna-binding domain. *Science*. 1989;245(4918):635-637. doi: 10.1126/science.2503871.

3. Klug A, Schwabe JWR. Protein motifs .5. zinc fingers. *Faseb Journal*. 1995;9(8):597-604.

4. Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P. KRAB-zinc finger proteins: A repressor family displaying multiple biological functions. *Curr Genomics*. 2013;14(4):268-278.

5. Pavletich NP, Pabo CO. Zinc finger dna recognition - crystal-structure of a Zif268-dna complex at 2.1-a. *Science*. 1991;252(5007):809-817. doi: 10.1126/science.2028256.

6. Urrutia R. KRAB-containing zinc-finger repressor proteins. *Genome Biol*. 2003;4(10):231. doi: 10.1186/gb-2003-4-10-231.

7. Emerson RO, Thomas JH. Adaptive evolution in zinc finger transcription factors. *Plos Genetics*. 2009;5(1):e1000325. doi: 10.1371/journal.pgen.1000325.

8. Margolin J, Friedman J, Meyer W, Vissing H, Thiesen H, Rauscher F. Kruppel-associated boxes are potent transcriptional repression domains. *Proc Natl Acad Sci U S A*. 1994;91(10):4509-4513. doi: 10.1073/pnas.91.10.4509.

9. Witzgall R, Oleary E, Leaf A, Onaldi D, Bonventre J. The kruppel-associated box-a (krab-A) domain of zinc-finger proteins mediates transcriptional repression. *Proc Natl Acad Sci U S A*. 1994;91(10):4514-4518. doi: 10.1073/pnas.91.10.4514.

10. Vissing H, Meyer W, Aagaard L, Tommerup N, Thiesen H. Repression of transcriptional activity by heterologous krab domains present in zinc-finger proteins. *FEBS Lett*. 1995;369(2-3):153-157. doi: 10.1016/0014-5793(95)00728-R.

11. Yao J, Ren X, Ireland J, Coussens P, Smith T, Smith G. Generation of a bovine oocyte cDNA library and microarray: Resources for identification of genes important for follicular development and early embryogenesis. *Physiological Genomics*. 2004;19(1):84-92. doi: 10.1152/physiolgenomics.00123.2004.

12. Richardson C, Jones P, Barnard V, Hebert C, Terlecki S, Wijeratne W. Estimation of the developmental age of the bovine fetus and newborn calf. *Vet Rec*. 1990;126(12):279-284.

13. Murdoch W, Dailey R, Inskeep E. Preovulatory changes in prostaglandins E2 and F2 alpha in ovine follicles. *J Anim Sci*. 1981;53:192-205.

14. Bettegowda A, Patel O, Ireland J, Smith G. Quantitative analysis of messenger RNA abundance for ribosomal protein L-15, cyclophilin-A, phosphoglycerokinase, beta-glucuronidase, glyceraldehyde 3-phosphate dehydrogenase, beta-actin, and histone H2A during bovine oocyte maturation and early embryogenesis in vitro. *Mol Reprod Dev*. 2006;73(3):267-278. doi: 10.1002/mrd.20333.

15. Wang L, Tripurani SK, Wanna W, Rexroad,Caird E.,,III, Yao J. Cloning and characterization of a novel oocyte-specific gene encoding an F-box protein in rainbow trout (oncorhynchus mykiss). *Reproductive Biology and Endocrinology*. 2013;11:86. doi: 10.1186/1477-7827-11-86.

16. Yang MY, Fortune JE. The capacity of primordial follicles in fetal bovine ovaries to initiate growth in vitro develops during mid-gestation and is associated with meiotic arrest of oocytes. *Biol Reprod*. 2008;78(6):1153-1161. doi: 10.1095/biolreprod.107.066688.

17. Punta M, Coggill PC, Eberhardt RY, et al. The pfam protein families database. *Nucleic Acids Res*. 2012;40(D1):D290-D301. doi: 10.1093/nar/gkr1065.

18. Wolfe SA, Nekludova L, Pabo CO. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct*. 2000;29:183-212. doi: 10.1146/annurev.biophys.29.1.183.

19. Klug A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annual Review of Biochemistry, Vol 79*. 2010;79:213-231. doi: 10.1146/annurev-biochem-010909-095056.

20. Zhao Q, Xie Y, Zheng Y, et al. GPS-SUMO: A tool for the prediction of sumoylation sites and SUMO-interaction motifs. *Nucleic Acids Res*. 2014;42(W1):W325-W330. doi: 10.1093/nar/gku383.

21. Kapustin Y, Souvorov A, Tatusova T, Lipman D. Splign: Algorithms for computing spliced alignments with identification of paralogs. *Biology Direct*. 2008;3:20. doi: 10.1186/1745-6150-3-20.

22. Yan C, Elvin J, Lin Y, et al. Regulation of growth differentiation factor 9 expression in oocytes in vivo: A key role of the E-box. *Biol Reprod*. 2006;74(6):999-1006. doi: 10.1095/biolreprod.105.050013.

23. Bettegowda A, Smith GW. Mechanisms of maternal mRNA regulation: Implications for mammalian early embryonic development. *Frontiers in Bioscience*. 2007;12:3713-3726. doi: 10.2741/2346.

24. Tripurani SK, Lee K, Wang L, et al. A novel functional role for the oocyte-specific transcription factor newborn ovary homeobox (NOBOX) during early embryonic development in cattle. *Endocrinology*. 2011;152(3):1013-1023. doi: 10.1210/en.2010-1134.

25. Tejomurtula J, Lee K, Tripurani SK, Smith GW, Yao J. Role of importin Alpha8, a new member of the importin alpha family of nuclear transport proteins, in early embryonic development in cattle. *Biol Reprod*. 2009;81(2):333-342. doi: 10.1095/biolreprod.109.077396.

26. Bettegowda A, Yao J, Sen A, et al. JY-1, an oocyte-specific gene, regulates granulosa cell function and early embryonic development in cattle. *Proc Natl Acad Sci U S A*. 2007;104(45):17602-17607. doi: 10.1073/pnas.0706383104.

27. Krisher RL. The effect of oocyte quality on development. *J Anim Sci*. 2004;82 E-Suppl:E14-23.

28. Picton H, Briggs D, Gosden R. The molecular basis of oocyte growth and development. *Mol Cell Endocrinol*. 1998;145(1-2):27-37. doi: 10.1016/S0303-7207(98)00166-X.

29. Gilchrist RB, Thompson JG. Oocyte maturation: Emerging concepts and technologies to improve developmental potential in vitro. *Theriogenology*. 2007;67(1):6-15. doi: 10.1016/j.theriogenology.2006.09.027.

30. Sirard M, Richard F, Blondin P, Robert C. Contribution of the oocyte to embryo quality. *Theriogenology*. 2006;65(1):126-136. doi: 10.1016/j.theriogenology.2005.09.020.

31. Eppig JJ, Wigglesworth K, Pendola FL. The mammalian oocyte orchestrates the rate of ovarian follicular development. *Proc Natl Acad Sci U S A*. 2002;99(5):2890-2894. doi: 10.1073/pnas.052658699.

32. Adolfo Palma G, Eduardo Arganaraz M, Daniel Barrera A, Rodler D, Angel Mutto A, Sinowatz F. Biology and biotechnology of follicle development. *Scientific World Journal*. 2012:938138. doi: 10.1100/2012/938138.

33. Pangas SA, Rajkovic A. Transcriptional regulation of early oogenesis: In search of masters. *Hum Reprod Update*. 2006;12(1):65-76. doi: 10.1093/humupd/dmi033.

34. Choi Y, Rajkovic A. Genetics of early mammalian folliculogenesis. *Cellular and Molecular Life Sciences*. 2006;63(5):579-590. doi: 10.1007/s00018-005-5394-7.

35. Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM. NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. *Science*. 2004;305(5687):1157-1159. doi: 10.1126/science.1099755.

36. Soyal S, Amleh A, Dean J. FIG alpha, a germ cell-specific transcription factor required for ovarian follicle formation. *Development*. 2000;127(21):4645-4654.

37. Zheng P, Dean J. Oocyte-specific genes affect folliculogenesis, fertilization, and early development. *Semin Reprod Med*. 2007;25(4):243-251. doi: 10.1055/s-2007-980218.

38. Schier AF. The maternal-zygotic transition: Death and birth of RNAs. *Science*. 2007;316(5823):406-407. doi: 10.1126/science.1140693.

39. BilodeauGoeseels S, Schultz G. Changes in the relative abundance of various housekeeping gene transcripts in in vitro-produced early bovine embryos. *Mol Reprod Dev*. 1997;47(4):413-420. doi: 10.1002/(SICI)1098-2795(199708)47:4<413::AID-MRD7>3.0.CO;2-N.

40. Memili E, First N. Control of gene expression at the onset of bovine embryonic development. *Biol Reprod*. 1999;61(5):1198-1207. doi: 10.1095/biolreprod61.5.1198.

41. Memili E, First N. Zygotic and embryonic gene expression in cow: A review of timing and mechanisms of early gene expression as compared with other species. *Zygote*. 2000;8(1):87-96. doi: 10.1017/S0967199400000861.

42. Tong ZB, Gold L, Pfeifer KE, et al. Mater, a maternal effect gene required for early embryonic development in mice. *Nat Genet*. 2000;26(3):267-268.

43. Wu XM, Viveiros MM, Eppig JJ, Bai YC, 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(2):187-191. doi: 10.1038/ng1079.

44. Li L, Baibakov B, Dean J. A subcortical maternal complex essential for preimplantation mouse embryogenesis. *Developmental Cell*. 2008;15(3):416-425. doi: 10.1016/j.devcel.2008.07.010.

45. Burns KH, Viveiros MM, Ren YS, et al. Roles of NPM2 in chromatin and nucleolar organization in oocytes and embryos. *Science*. 2003;300(5619):633-636. doi: 10.1126/science.1081813.

46. Payer B, Saitou M, Barton SC, et al. Stella is a maternal effect gene required for normal early development in mice. *Current Biology*. 2003;13(23):2110-2117. doi: 10.1016/j.cub.2003.11.026.

47. Nichols J, Zevnik B, Anastassiadis K, et al. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*. 1998;95(3):379-391. doi: 10.1016/S0092-8674(00)81769-9.

48. Lengner CJ, Welstead GG, Jaenisch R. The pluripotency regulator Oct4 - A role in somatic stem cells? *Cell Cycle*. 2008;7(6):725-728.

49. Foygel K, Choi B, Jun S, et al. A novel and critical role for Oct4 as a regulator of the maternal-embryonic transition. *Plos One*. 2008;3(12):e4109. doi: 10.1371/journal.pone.0004109.

50. Kurosaka S, Eckardt S, McLaughlin KJ. Pluripotent lineage definition in bovine embryos by Oct4 transcript localization. *Biol Reprod*. 2004;71(5):1578-1582. doi: 10.1095/biolreprod.104.029322.

51. Kim SS, Chen YM, OLeary E, Witzgall R, Vidal M, Bonventre JV. A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. *Proc Natl Acad Sci U S A*. 1996;93(26):15299-15304. doi: 10.1073/pnas.93.26.15299.

52. Friedman J, Fredericks W, Jensen D, et al. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev*. 1996;10(16):2067-2078. doi: 10.1101/gad.10.16.2067.

53. Moosmann P, Georgiev O, LeDouarin B, Bourquin J, Schaffner W. Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. *Nucleic Acids Res*. 1996;24(24):4859-4867. doi: 10.1093/nar/24.24.4859.

54. Nielsen AL, Ortiz JA, You J, et al. Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *EMBO J*. 1999;18(22):6385-6395. doi: 10.1093/emboj/18.22.6385.

CHAPTER 2

Functional Analysis of the Novel Bovine ZNFO Transcription Factor Reveals it is Required for Normal Early Embryonic Development in Cattle

Jacqelyn M. Hand[1], Kun Zhang[2], Kristen Mastrantoni[1], George W. Smith[2], and Jianbo Yao[1]

[1]Laboratory of Animal Biotechnology and Genomics, Division of Animal and Nutritional Sciences, West Virginia University, Morgantown, WV 26506.

[2]Laboratory of Mammalian Reproductive Biology and Genomics, Department of Animal Science, Michigan State University, East Lansing, Michigan 48824.

KEYWORDS: Zinc finger, oocyte, embryogenesis, transcriptional repressor, cattle

AUTHOR'S CONTRIBUTIONS

JMH designed and performed most of the experiments including GST pull-down and protein work, luciferase assay and cell culture studies, and cloning. JMH also drafted the manuscripts. KZ and GWS designed and performed the microinjection study. KM assisted with the cell culture work. JY designed the study and supervised the experimental work.

ABSTRACT

Zinc finger (ZNF) transcription factors interact with DNA through zinc finger motifs and play important roles in a variety of cellular functions including cell growth, proliferation, development, apoptosis, and intracellular signal transduction. One-third of ZNF proteins in metazoans contain a highly conserved N-terminal motif known as the KRAB domain, which acts as a potent, DNA-binding dependent transcriptional repression module. Identification and characterization of the novel *ZNFO* revealed it is a KRAB-containing maternal-effect gene found exclusively in bovine oocytes. To test the functional role of *ZNFO*, zygotes were generated by *in vitro* maturation and fertilization of oocytes, and injected with small interfering RNA (siRNA) designed to knockdown *ZNFO*. Cleavage rates were not affected by ZNFO siRNA injection ($P > 0.05$). However, embryonic development to 8- to 16-cell stage and blastocyst stage was significantly reduced relative to the uninjected and negative control siRNA-injected embryos (n = 3 replicates; 25-30 embryos/treatment; $P < 0.05$). Further, interaction of ZNFO with the highly conserved transcriptional repressor co-factor (KAP1) was demonstrated by GST pull-down, and evidence supporting transcriptional repression by ZNFO using a GAL4-luciferase assay. In addition, transfection studies verified that a ZNFO-GFP fusion protein localizes specifically to the nucleus, further supporting proposed function in transcriptional regulation. Results of described studies demonstrate that ZNFO is a maternally-derived oocyte-specific factor required for early embryonic development in cattle, and that ZNFO functions as a transcriptional regulator required during early embryonic developmental events by repressing transcription, possibly controlling activation of the embryonic genome.

INTRODUCTION

Early embryonic development is one of the most critical periods in mammalian development and is composed of several important transitions including replacement of maternal RNAs with zygotic RNAs, compaction, the first lineage differentiation into the inner cell mass and trophectoderm and, finally, implantation. Various physiological processes and biosynthetic changes regarding genomic activity take place during this early time. Among these events is the first important developmental transition that occurs following fertilization at which time the embryo switches from using transcripts derived from the maternal genome to those synthesized by the zygote as the result of embryonic genome activation (EGA)[1]. Mammalian oocytes accumulate a vast collection of mRNA and proteins throughout oogenesis that mediate subsequent embryonic development. During oocyte meiotic maturation and the early stages of embryonic development the transcriptional machinery for this collection of molecules is silent until EGA initiates transcriptional activity within the embryonic nucleus. The onset of EGA is a species-specific event that takes place during the first few cell cycles post-fertilization around the 8- to 16-cell stage in bovine embryos[2,3] and several cycles later than observed for the mouse (2-cell stage) or human (4- to 8-cell stage)[4,5]. Therefore, any developmental events required for early embryogenesis (i.e. meiotic maturation, fertilization, initial cleavage divisions, and programming of EGA) prior to the onset of EGA are regulated by the translation of pre-existing maternal transcripts[6]. Following onset of EGA the destruction of maternal mRNAs begins by maternally encoded products [6,7].

One third of the various conserved domains that contribute to C2H2-ZNF protein function contain the Krüppel-Associated Box (KRAB) domain[8], making KRAB-ZNFs the single largest group of transcriptional repressors in the genomes of higher organisms. The KRAB domain is a potent transcriptional repression module responsible for DNA binding-dependent gene silencing activity and is located at the amino-terminal end of most C2H2 zinc finger proteins[9,10]. When tethered to DNA via its zinc finger motifs, the KRAB domain of KRAB-ZNF proteins recruits and interacts with the corepressor protein, KRAB-associated protein 1 (KAP1)[11,12], which is an absolute requirement for KRAB-containing zinc finger proteins to bind to and mediate transcriptional repression. Upon binding to DNA, KAP1 functions as a scaffold to form a multi-

molecular complex at the promoters of target genes by recruiting various heterochromatin-inducing factors such as heterochromatin protein 1 (HP1)[13], the nucleosome remodeling and histone deacetylation (NuRD) complex[14], the histone methyltransferase SET domain, bifurcated 1 (SETDB1)[15], the nuclear receptor corepressor complex 1 (N-CoR1)[16], and during early embryonic development, *de novo* DNA methyltransferases[17]. This complex induces transcriptional silencing by condensing chromatin. As a powerful transcriptional repressor, most members of the KRAB-ZNF family have diverse functional roles in nearly all tissues and a variety of cellular functions, including cell proliferation and differentiation, metabolism, apoptosis, neoplastic transformation, cell cycle regulation, and regulation embryonic development[8,18].

Based on the observed early embryonic expression pattern characterizing *ZNFO* as a maternal-effect gene, it was hypothesized that this novel C2H2 KRAB-containing zinc finger has a distinct and essential role in the development of the early bovine embryo through a transcriptional regulation mechanism. The objectives of this study were 1) to determine the role for ZNFO in the development of early embryogenesis in cattle and 2) to confirm the transcriptional mechanism that regulates ZNFO function.

MATERIALS AND METHODS

Plasmid Construction

For the preparation of glutathione S-transferase (GST) fusion recombinant protein, full-length ZNFO cDNA was cloned into pGEX-4T1 (GE Healthcare, Salt Lake City, UT) using *Sma*I and *Xho*I sites (ZNFO:pGEX-4T1). The KAP1 ORF was amplified and subcloned into pcDNA3.1 (Invitrogen, Carlsbad, CA), containing a carboxy-terminal FLAG-tag, by *Hind*III and *Bam*HI restriction sites. The resulting FLAG-KAP1:pcDNA3.1 construct was transfected into the HEK293 cell line. For expression of a green fluorescent protein (GFP) fused to ZNFO, the full-length ZNFO sequence was amplified by PCR from the ZNFO:pGEX-4T1 target and inserted into pcDNA3-EGFP expression plasmid (Addgene, Cambridge, MA; catalog #13031) using *Kpn*I and *Xho*I restriction sites (ZNFO:pEGFP). The mutation derivative of ZNFO lacking the RHRK sequence was subcloned into pGEM-Teasy (Promega, Madison, WI) and then cloned into pcDNA3-EGFP (ZNFO_RHRK:pEGFP). Deletion constructs expressing ZNFO proteins lacking all 9 zinc fingers (ZNFOΔ1-9), the last 6 zinc fingers (ZNFOΔ4-9), and the last 3 zinc fingers (ZNFOΔ7-9) were generated by amplifying ZNFO using the same gene-specific forward primer in combination with different reverse primers that target different sites in the zinc finger region. All deletion constructs generated using the same *Kpn*I and *Xho*I restriction sites. The luciferase assay construct was generated by amplifying full-length ZNFO containing sites for *Sal*I and *Kpn*I and fusing to the carboxy-terminal end of GAL4. All clones were confirmed by sequencing. Primers for restriction-containing primers are listed in Table 1.

Expression and Purification of Glutathione S-Transferase Fusion Protein

The GST-fused ZNFO protein was transfected into *Escherichia coli* Rosetta (DE3)pLysS (Novagen, Madison, WI) strain. Bacterial cultures were induced with 2.5 mM isopropyl β-D-thiogalactopyranoside at 28°C for 8 h. Bacterial pellets were collected by centrifugation and bacteria were lysed in equilibration/wash buffer (125 mM Tris and 150 mM sodium chloride, pH 8.0 (Pierce, Rockford, IL)) by sonication on ice for 10 sec, 12 times (10 sec on, 10 sec off) at 30% amplitude. The clear lysate was incubated and passed through an immobilized glutathione column (Pierce). After washing several times with equilibration/wash buffer, the recombinant protein was

eluted by the addition of 20 mM glutathione to the buffer. The eluted protein was concentrated by Microcon centrifugal filter devices (Millipore, Bedford, MA). A portion of the proteins from various steps were electrophoresed through an SDS-PAGE gel and stained with Coomassie brilliant blue for analysis.

GST Pull-Down Assay

Equal amounts (2 µg) of GST or GST-fused ZNFO proteins were immobilized on 100 uL of glutathione beads (Pierce) in 0.3 mL of equilibration/wash buffer (above). After incubation for 1 hr at 4°C, beads were washed and incubated with FLAG-KAP1 cell lysate overnight at 4°C. The beads were washed thoroughly, then boiled in Laemmli buffer (containing β-mercaptoethanol) and resolved by SDS-PAGE, followed by Western blot analysis using FLAG-tag primary antibody and infrared fluorescent secondary antibody and visualized on an Odyssey system (Li-COR, Lincoln, NE).

Cell Culture and EGFP Reporter Assay

A HEK293 cell line was cultured in DMEM supplemented with 10% (v/v) fetal bovine serum (FBS) in a humidified 5% (v/v) CO_2 incubator at 37°C. The day before transfection, cells were seeded on coverslips in a 6-well plate. Transfection of ZNFO:pcDNA3-EGFP and the empty control plasmid (pcDNA3-EGFP) were performed with X-treme Gene 9 (Roche, Indianapolis, IN) at optimized ratios of 3:1 (uL of transfection reagent: ug of plasmid DNA). Twenty-four h after transfection, cells were washed with PBS and fixed with methanol at RT for 5 min. Seeded coverslips were placed on slides and sealed with ProLong® Gold antifade reagent with DAPI (4',6-diamidine-2'-phenylindole dihydrochloride; Life Technologies, Carlsbad, CA). Fluorescent images were taken with a Zeiss M1 microscope with an X-Cite fluorescence generator using AxioVision software version 4.8.2.

Luciferase Assay

Cells were grown to 70% confluency in 6 well plates and transfected with the reporter and gene constructs using X-treme Gene 9 (Roche) at optimized ratios of 3 uL of transfection reagent per 0.5 ug of plasmid DNA (either pBIND empty vector (GAL4-empty) or ZNFO:pBIND (GAL4-

ZNFO)) plus 300 ng luciferase reporter vector (pG5*luc*). After 48 h, cells were split into a 96 well plate and the reactions were carried out using the Dual Glo Luciferase Assay system (Promega) according to the manufacturer's instructions. Relative luciferase activity was calculated as firefly luciferase activity divided by *Renilla* luciferase activity. Luminescence was measured using Phoenix GENios Microplate Reader. All experiments were repeated in triplicate. Differences were determined by Student's t-test, or by Tukey-Kramer for the dose-response experiment, with $P < 0.05$.

Site-Directed Mutagenesis

Mutagenesis of the predicted nuclear localization signal (NLS) from the amino acid sequence RHRK to RHAA was performed by a two-step PCR method. Briefly, the fragment on either side of the RHRK DNA sequence (AGACATAGGAAA) was amplified with the RHAA DNA sequence introduced (AGACATGCCGCA) as part of the primer design (Table 1). The resulting amplicons were gel purified, combined in reaction buffer, and incubated at 95°C for 10 min and allowed to slowly cool to RT. A PCR reaction lacking primers was done with the following conditions: 60°C for 10 min, 20 cycles of 95°C for 1 min and 60°C for 10 min. The product was subjected to a final PCR reaction with primers for the full-length ORF and such conditions: 95°C for 5 min, 35 cycles of 95°C for 30 sec, 66°C for 3 min, and 72°C for 2 min 20 sec, followed by 72°C for 10 min. The complete product was gel purified and subcloned.

RNA Interference (RNAi) Experiments

Knockdown of endogenous ZNFO in bovine embryos was performed via microinjection of *ZNFO* small interfering RNA (siRNA). RNAi experiments were conducted according to previously published procedures[19-22] with modifications noted herein. The publicly available siRNA design algorithm (siRNA target finder; Ambion, Austin, TX) was used to design three distinct siRNA species targeting the open reading frame of bovine *ZNFO* mRNA (designated as siRNA species 1, 2, and 5, respectively). The candidate siRNA species were interrogated by using the basic local alignment tool program (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to rule out homology to any other known genes in the bovine expressed sequence tag and genomic database. The *ZNFO* siRNA species were generated commercially (Integrated DNA Technologies,

Coralville, IA). The sense and antisense oligonucleotide template sequences for the siRNA species are given in Table 1. Procedures for *in vitro* maturation of oocytes (obtained from abattoir-derived ovaries), *in vitro* fertilization to generate zygotes for microinjection, and for subsequent embryo culture were conducted basically as described elsewhere[23]. Presumptive zygotes collected at 16–18 h post insemination (hpi) were used in all microinjection experiments. Each individual siRNA species was validated for efficacy of *ZNFO* mRNA knockdown in early embryos. Presumptive zygotes were microinjected with approximately 20 pL of individual *ZNFO* siRNA species (25 µM concentration each). Uninjected embryos and embryos injected with a negative siRNA (universal control no. 1; Ambion) were used as control groups (n = 3 pools of 20 embryos per treatment). Efficacy of *ZNFO* siRNA in reducing ZNFO protein in early embryos was determined by microinjection of ZNFO eight-cell embryos collected 48 hpi (n = 20 embryos per group). The development of the uninjected or injected embryos (with *ZNFO* siRNA or negative control siRNA) was evaluated by recording the proportion of embryos that cleaved (48 h after insemination), reached eight- to 16-cell stage (72 h after insemination) and blastocyst stage (7 d after insemination). Each group contained 25-30 embryos per treatment (n = 3 replicates). Percent data were transformed to ArcSine. Differences in treatment means for % cleaved and % blastocyst were compared using Kruskal-Wallis; % 8-16 was normally distributed by one-way ANOVA and differences compared using Tukey-Kramer. Data are expressed as mean ± SEM from three replicates (n = 25-30 zygotes per treatment per replicate). Values with different letters across treatments indicate significant differences ($P < 0.05$).

RESULTS

ZNFO *is Required for Bovine Early Embryonic Development*

To investigate the function of ZNFO in early embryonic development, RNAi experiments were performed to reduce the expression of ZNFO in bovine embryos. Three *ZNFO* siRNA species targeting different regions of the *ZNFO* transcript were produced *in vitro*, and initial experiments were performed to test the efficacy and specificity of the siRNAs in silencing *ZNFO* (data not shown). siRNA species 5 was able to significantly reduce *ZNFO* mRNA in 4-cell embryos (Figure 1A; $P < 0.05$) relative to the uninjected and negative siRNA controls.

To determine whether knockdown of ZNFO in bovine embryos has an effect on embryonic development, *ZNFO* siRNA was microinjected into *in vitro* fertilized oocytes and the resulting cleavage rate of zygotes, and proportion of embryos developing to 8- to 16-cell stage and blastocyst stage was examined. Injection of *ZNFO* siRNA did not affect the cleavage rate (Figure 1B) but reduced the proportion of embryos developing to 8- to 16-cell stage (Figure 1C) relative to uninjected and negative control siRNA-injected embryos ($P < 0.05$). Likewise, *ZNFO* siRNA injection decreased the proportion of embryos developing to the blastocyst stage compared with the uninjected and negative control siRNA-injected embryos ($P < 0.05$; Fig 1D). These results clearly demonstrate an impaired ability of bovine zygotes to reach the blastocyst stage with knockdown of ZNFO, placing an important functional requirement of ZNFO during bovine early embryogenesis.

Interaction of Bovine *ZNFO* Protein with Bovine KAP1

The hypothesis that ZNFO functions as a transcriptional repressor is largely based on the identification of a KRAB domain located within ZNFO at the N-terminal end. The KRAB domain, and specifically KRAB-containing zinc finger proteins, are well-established interacting partners of KAP1 proteins[11,12,24,25]. The highly conserved function of KAP1 illustrates an event in which KAP1 recruits several co-factor complexes upon binding C2H2 zinc finger proteins and subsequently repress transcription by inducing heterochromatin formation[26]. Therefore, a GST pull-down assay was performed to determine whether bovine ZNFO does interact with KAP1;

ZNFO was fused to the C-terminus of GST and tested for the ability to bind KAP1. Purified GST-fused ZNFO was confirmed by SDS-PAGE analysis (Figure 2, bottom panel). The *in vitro*-binding assay showed that KAP1 had a strong binding affinity for ZNFO but not with the GST protein control (Figure 2, top panel). Furthermore, by removing the KRAB domain, KAP1-FLAG was not immunochemically detectable, illustrating that the KRAB domain is the required interaction interface for ZNFO and KAP1 binding.

Involvement of ZNFO *in Transcriptional Repression*

Based on the fact that ZNFO harbors a highly conserved KRAB domain and was just confirmed to interact with the repression associated KAP1 co-factor, it seemed quite likely that ZNFO is a functional transcriptional regulator. To test the effect of ZNFO on transcription, a reporter system was used that contains firefly luciferase driven by a viral promoter proximal to five copies of GAL4-binding sites (pG5*luc*; Fig 3A). The pG5*luc* reporter was transiently introduced into a HEK293 cell line along with a construct expressing the full-length ZNFO fused to the GAL4-DNA binding domain (GAL4-DBD) or the GAL4 expression vector alone (GAL4-empty). Compared to GAL4-empty vector, GAL4-ZNFO decreased the promoter activity of the reporter plasmid (Fig 3B). As expected, co-transfection of GAL4-ZNFO with GAL4-empty at differing concentrations did not exhibit a dose-dependent effect in luciferase activity, although addition of the GAL4-empty vector did interfere with GAL4-ZNFO activity at any concentration (Fig 3C). These results suggest that ZNFO possesses intrinsic transcriptional repressive activity.

Subcellular Localization & Nuclear Localization Signal

To gain insight into the biological function of the ZNFO protein, subcellular localization was examined by fusion of green fluorescent protein (GFP) to the C-terminus of ZNFO for EGFP reporter assay. Fluorescent imaging in HEK296 cells indicated that overexpressed ZNFO is primarily localized in the nucleus, suggesting that ZNFO functions as a nuclear protein (Figure 4.1b). Further, the subcellular localization of ZNFO was compared with a derivative form of ZNFO that was mutated at the predicted NLS (ZNFO_RHRK). Consistent with the previous observation, the wild-type ZNFO (ZNFO*wt*), which displayed the typical diffuse nucleoplasmic staining, ZNFO_RHRK also displayed a prominent nuclear staining pattern (Figure 4.1c). Both variations

were in contrast to the control empty GFP protein vector that showed clear cytoplasmic expression (Figure 4.1a). Therefore, the predicted RHRK NLS sequence does not seem to be the NLS for ZNFO.

Three constructs, each containing three different sets of zinc finger motifs, were created in an attempt to narrow the location of the region of the NLS. When the C-terminal zinc finger motifs were removed, either zinc finger motifs 7-9 or 4-9 (ZNFOΔ7-9 or ZNFOΔ4-9), subcellular localization remained specific to the nucleus (Fig 4.2e and f). Following removal of zinc fingers 1-9 (ZNFOΔ1-9), however, cytoplasmic localization was observed that is distinct from that of full-length ZNFO, ZNFOΔ4-9, or ZNFOΔ7-9 (Fig 4.2d). Although the true NLS has yet to be identified, it seems highly likely to be located within the first three zinc finger motifs of ZNFO.

DISCUSSION

The results herein indicate that expression of the oocyte-specific *ZNFO* is required for development of the early bovine embryo. Biochemical experimentation showed that ZNFO is localized in the nucleus, has the ability to interact with a key transcriptional repressor co-factor KAP1, and, indeed, represses gene transcription. Thus, these studies suggest that ZNFO may have an essential role in regulating the maternal-to-zygotic transition by inducing transcriptional silencing of genes involved in early embryonic development during- and post-EGA. Studies on the roles of KRAB-containing zinc finger transcription factors and interaction with KAP1 show that there are various effects during embryonic developmental processes including maintenance of the imprinting of genes, maturation of bone cells and hematopoiesis, and regulation of convergent extension in mouse embryo[27-29]. Although these factors have begun to uncover a few of the complex functional requirements of ZNFs in early embryonic development, they are focused on mouse and human models.

Results of the present study clearly support a functional role of ZNFO in early embryonic development in a livestock species and demonstrate that *ZNFO* knockdown dramatically impaired development to the blastocyst stage. Bovine *in vitro* culture systems demonstrate that a developmental block arises around the 8-cell stage in most embryos, and, thus, EGA is considered to be the most critical event for early developmental viability[30]. The results further suggest that ZNFO may be required during the early stages of embryonic development before activation of the embryonic genome, because *ZNFO* siRNA injection significantly reduced the development of embryos at the 8- to 16-cell stages. Before activation of the embryonic genome at the 8-cell stage the occurrence of minor EGA has been shown to take place as early as the 2-cell stage in bovine[3], a transition that is necessary for genome reprogramming and acquisition of totipotency by the embryo[31]. Although ZNFO knockdown has little effect on initial embryo cleavage divisions, it possibly plays roles in these early stages. As ZNFO appears to conform to the classic maternal-effect gene expression pattern, possibly maternal ZNFO protein, maintained from the GV or MII oocyte, continues to mediate and contribute to the minor EGA functions, because, as a protein, it would have escaped transcript knockdown at the zygotic injection. The idea that ZNFO could regulate several aspects of the maternal-to-zygotic transition (MZT) in a differential way, from

92

transcript control in the oocyte to EGA in the embryo, stems from the reported observations of other maternal-effect genes.

Oocyte maturation includes several morphological changes accompanying the progression of meiosis from prophase I to metaphase II. These changes lead the oocyte to a stage of chromosome condensation and transcriptional silencing, which will last until EGA[32]. The molecular mechanisms of KRAB-ZFP-mediated transcriptional regulation requires interaction with chromatin-remodeling factors. The universal co-repressor KAP1 acts as a scaffold for chromatin-modifying complexes and chromatin remodeling activities by recruitment to the promoters of target genes and initiating ATP-dependent activities that modify chromatin. KAP1 harbors an N-terminal RBCC region that is responsible for KRAB domain binding[11], while the central HP1-binding domain and C-terminal tandem PHD-bromodomain (PB) are required for gene silencing. The PB domain recruits factors found in the NuRD complex and SETDB1, which mediate nucleosome remodeling and histone deacetylation, and trimethylation of histone H3 at Lys-9, respectively, to create high affinity genomic binding sites for the KAP1-HP1 complex[14,15]. The KRAB-ZNF-KAP1 complex then induces heterochromatin formation following recognition of local cis-acting sequences[13,26].

In this study, ZNFO demonstrated the ability to physically interact with KAP1, with high affinity. By removal of the N-terminal KRAB domain, it was demonstrated that this region is indeed the interaction interface responsible for interaction with KAP1. ZNFO also consistently demonstrated the ability to repress transcription in the GAL4-reporter system. These results suggest that ZNFO is abundantly expressed in cells as a negative regulator of transcription by binding to the KAP1 co-factor. Further, ZNFO was localized only to the nucleus, supporting the idea that ZNFO is involved in the regulation of zygotic transcriptional activity within the embryonic nucleus around the time of EGA. In an effort to delineate the region responsible for this nuclear localization, the software-predicted NLS was mutated; however, that particular monopartite region was not responsible. What does seem to be important is the region within the first three zinc finger motifs. Perhaps a Karyopherin (Importin) signal or non-canonical NLS signal is located within that region of ZNFO. These observations suggest a model in which KRAB-ZFP-KAP1-dependent recruitment of histone modifiers for histone methylation and formation of

facultative heterochromatin act to achieve gene silencing. It is possible that ZNFO, as a transcriptional repressor, regulates the important genes required for activation of the embryonic genome.

Understanding the functions of the factors involved in the regulation of chromosome condensation and transcriptional silencing are necessary to the study of oocyte maturation. Oocyte-specific transcription factors are likely the critical switches that control oocyte maturation, fertilization, and early embryo development. The oocyte is an absolutely incredible cell; able to regulate ovarian follicular growth and then remodel upon fertilization into a totipotent zygote[33]. These remarkable transitional events are dependent on transcripts and proteins that must be synthesized prior to meiotic maturation. Oocytes that can achieve spontaneous maturation, upon being liberated from follicles and placed in culture, clearly have acquired the molecules required for resumption and completion of meiosis[34]. In order for an oocyte to become meiotically competent, several inhibitory factors are required, most of which are not well understood. As a potential transcriptional repressor, ZNFO may play a role in maintaining the arrested status during oocyte meiosis. However, further experimentation of this novel model to specifically test the functional role of maternal ZNFO in meiotic maturation and initial cleavage divisions post fertilization is warranted.

The results presented in this study describe ZNFO as an oocyte-specific C2H2 KRAB-zinc finger transcription factor that plays a key role in ensuring early embryo survival possibly by regulating transcription through its established interaction with co-repressor KAP1. Clearly, ZNFO is required for early embryonic development and is present throughout oocyte maturation and follicular development. A distinct biochemical mechanism of transcriptional regulation by ZNFO has been identified; however the important downstream interactions and effects have yet to be elucidated. The studies presented here have identified, characterized, and established a physiological necessity, as well as identified a co-factor interaction that implicates a silencing mechanism involving the novel bovine KRAB-containing ZNFO transcription factor.
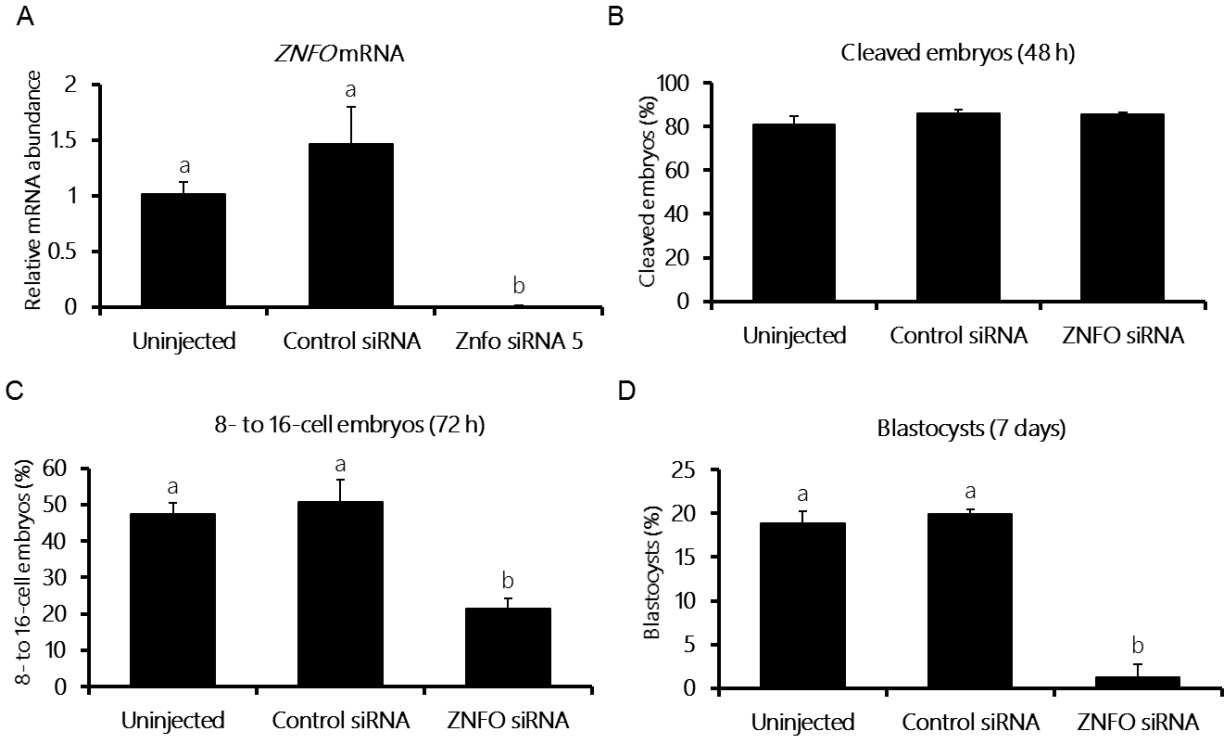
FIGURES

***Figure 1*. Effect of RNAi-induced *ZNFO* knockdown on early embryonic development.** (A). Effect of *ZNFO* siRNA microinjection on abundance of *ZNFO* mRNA in 4-cell embryos determined by real-time PCR. Data were normalized relative to abundance of GFP. Proportion of embryos that cleaved within 48 h after fertilization (B) developed to 8- to 16-cell stage (C) and developed to blastocyst stage (D). Uninjected embryos and embryos injected with a nonspecific siRNA were used as controls. Data are expressed as mean ± SEM (n = 3 replicates; 25-30 embryos/treatment). Values with different letters across treatments indicate significant differences ($P < 0.05$).
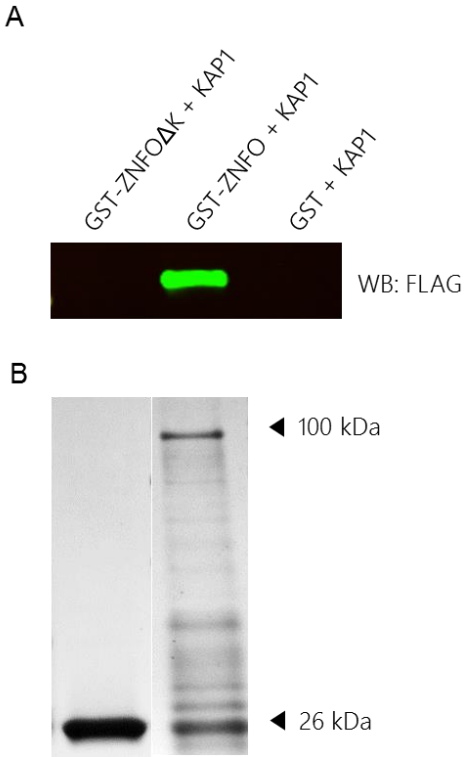
Figure 2. *In vitro* **interaction of bovine ZNFO protein and bovine KAP1 by GST pull-down assay.** (A). The GST pull-down assay. Overexpressed KAP1 cell lysate was incubated with immobilized GST proteins: either GST-ZNFO fusion protein with the KRAB domain removed (ZNFOΔK), GST-ZNFO (full length-ZNFO), and the empty vector GST protein. Western blot analysis, using antibodies against the Flag-tagged KAP1 protein, was used to detect the eluted ZNFO and KAP1 interactions. (B). Coomassie blue-stained SDS-PAGE gel confirming purified GST and GST-ZNFO fusion proteins at the correct molecular weights (arrowheads).
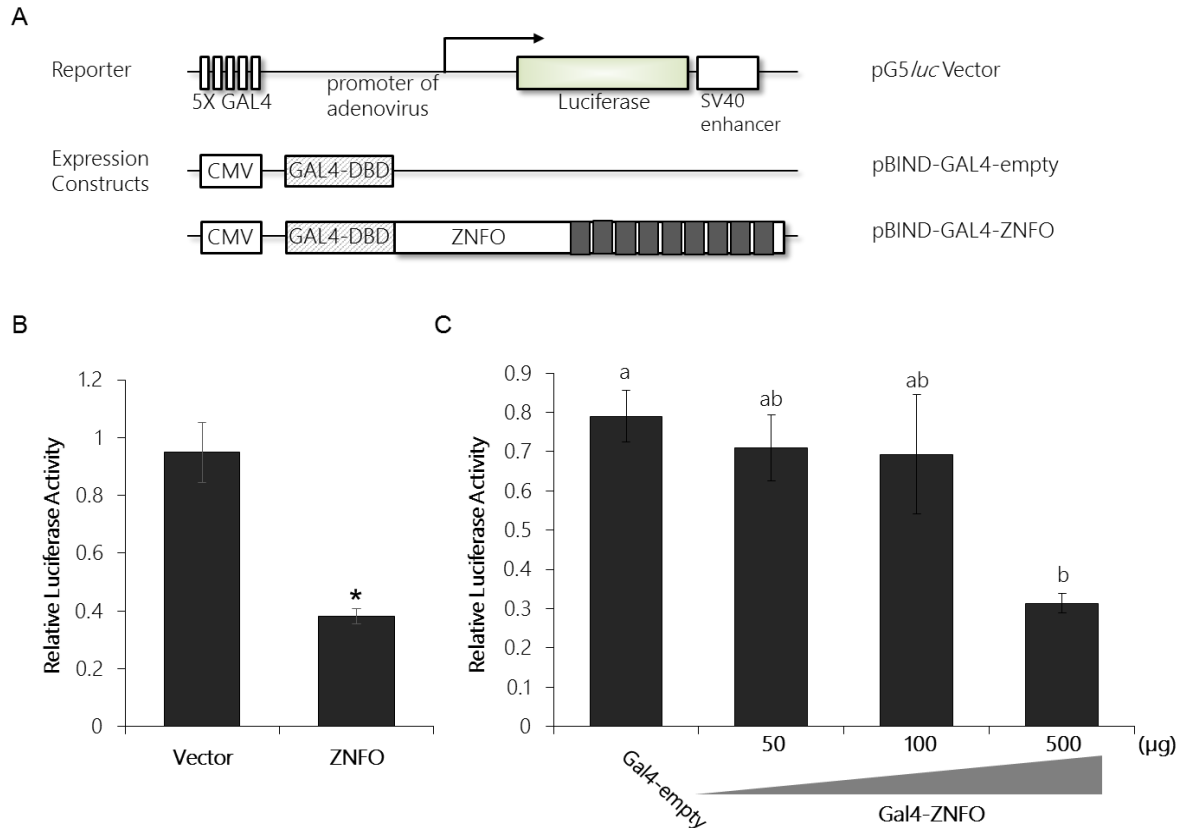
Figure 3. **Transcriptional repression by ZNFO.** (A). Schematic representation of the reporter and expression vectors used in the luciferase assay. The pG5*luc* reporter vector contains the firefly luciferase gene and five GAL4-binding sites. The ZNFO gene is fused to the GAL4-DNA-Binding Domain. *Dark shaded boxes* represent zinc finger domains. (B). Transcription repression by ZNFO. HEK293 cells were transfected with 500 ng GAL4-ZNFO or GAL4-empty expression plasmid, together with 500 ng luciferase reporter. (C). Dose response of transcription repression by ZNFO. HEK293 cells were transfected with 0, 50, 100, or 500 ng GAL4-ZNFO expression plasmid, each supplemented with GAL4-empty to 500 ng total, together with 300 ng of the luciferase reporter. Forty-eight hours later luciferase activity was measured. Relative luciferase activity was calculated as firefly luciferase activity divided by *Renilla* luciferase activity and shown relative to control GAL4-empty vector. Each graph represents mean ± s.e., and all experiments were performed in triplicate. Significant differences were determined by Student's t-test (*; B) or one way ANOVA followed by Tukey-Kramer (differing letters; C), with *P* < 0.05.
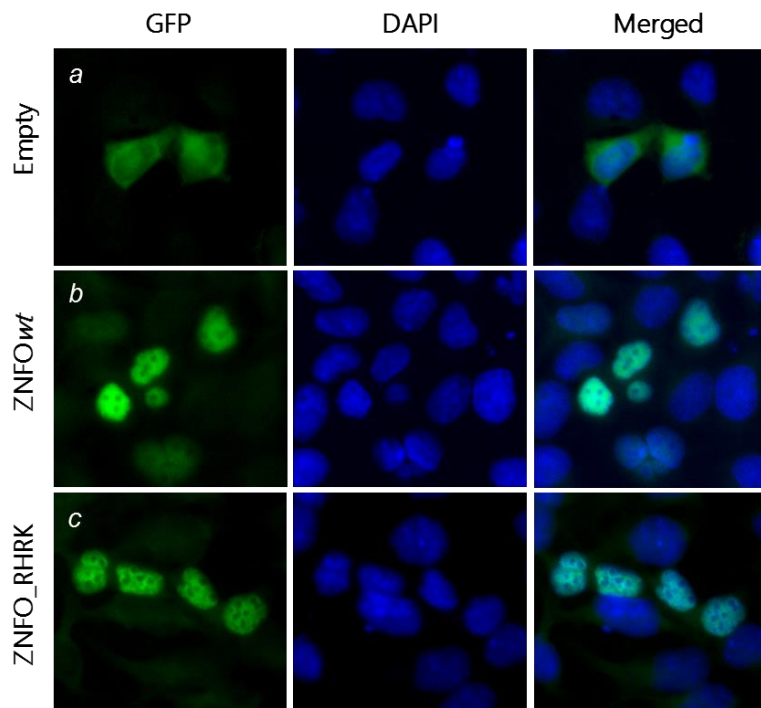
Figure 4.1. **Nuclear localization of bovine ZNFO protein**. HEK293 cells transiently transfected with (a) a GFP protein empty vector (Empty), (b) a GFP-fused ZNFO (ZNFO*wt*) or (c) a GFP-ZNFO with a mutation to the predicted NLS sequence (RHRK mutated to RHAA; ZNFO_RHRK) for fluorescent microscopic analysis. ZNFO*wt* and mutated ZNFO_RHRK both specifically localized to the nucleus, compared to the empty vector GFP control that was both cytoplasmic and nuclear.
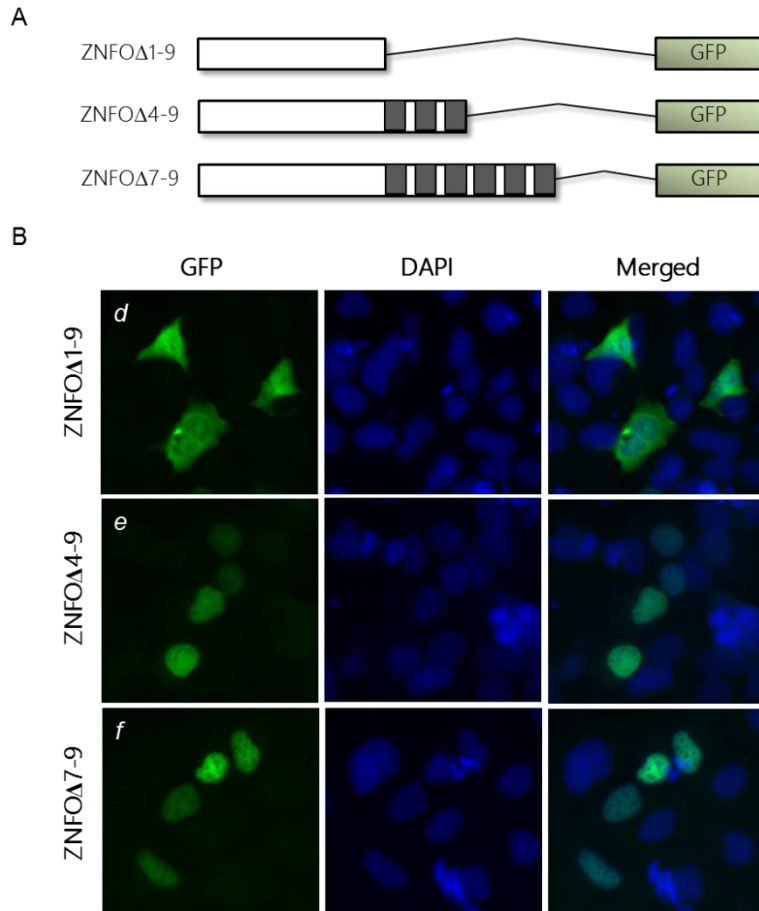
Figure 4.2. **Nuclear localization of bovine ZNFO protein**. (A) Schematic representation of the expression constructs designed for the subcellular localization assay to determine the region of the NLS. The ZNFO gene lacking all zinc finger motifs (ZNFOΔ1-9), fingers 4-9 (ZNFOΔ4-9), or motifs 7-9 (ZNFOΔ7-9) were fused to GFP. *Dark shaded boxes* represent zinc finger domains. (B) HEK293 cells transiently transfected with ZNFOΔ4-9 (e) or ZNFOΔ7-9 (f) localized specifically to the nucleus as does ZNFO*wt*, but ZNFOΔ1-9 was observed in both cytoplasmic and nuclear compartments.

TABLES

Table 1 Primers used in this study

Primer name	Primer sequence (5′-3′)	Application
ZNFO-siRNA-S1	GATTTATTCCCAACAGCAAAGCTAG	siRNA synthesis
ZNFO-siRNA-AS1	CTAGCTTTGCTGTTGGGAATAAATCTT	siRNA synthesis
ZNFO-siRNA-S2	AGAACAGAGAGCATGTTATGATGGA	siRNA synthesis
ZNFO-siRNA-AS2	TCCATCATAACATGCTCTCTGTTCTTT	siRNA synthesis
ZNFO-siRNA-S5	CAACTCTGAGCATAGGTGTCTATTA	siRNA synthesis
ZNFO-siRNA-AS5	TAATAGACACCTATGCTCAGAGTTGTT	siRNA synthesis
ZNFO-ORF-SmaI-F	CGCACACCCGGGTATGCAGAGGTGTGTTTGCTACA	Cloning – GST
ZNFO-ORF-XhoI-R	GGGCCCCTCGAGCTAGGGTTTCTCTGCAGTATG	Cloning – GST
KAP1-ORF-HindIII-F	GGGCCCAAGCTTTTGCCGCCACCATGGCGGCTTCGGCTGCGGCGG	Cloning - pcDNA
KAP1-ORF-BamHI-R + FLAG	GGGCCCGGATCCCTACTTATCGTCGTCATCCTTGTAATCGGGGCCCTCACCAGGGCCAGTA	Cloning - pcDNA
ZNFO-ORF-KpnI-F	GGGCCCGGTACCGCCGCCACCATGGAGAGGTGTGTTTGCTACA	Cloning – EGFP
ZNFO-ORF-XhoI-R	GGGCCCCTCGAGGGGTTTCTCTGCAGTATGTATTC	Cloning – EGFP
ZNFO-RHRK-KpnI-R1	ACTTAGATTTGGTGAGTTACTAAAGAC	Cloning – EGFP
ZNFO-RHRK-XhoI-F2	CTTTAGTAACTCACCAAATCTAAGTATTCATACAGGAAGGAAATGTTTCAAATGTACAGC	Cloning – EGFP
ZNFO-Δ1-9-XhoI-R	GGGCCCCTCGAGCTTGTAATGTTTCTGCTTAGTCT	Cloning - EGFP
ZNFO-Δ4-9-XhoI-R	GGGCCCCTCGAGTTTATAAGGTCTCTCCCCAGTATGA	Cloning - EGFP
ZNFO-Δ7-9-XhoI-R	GGGCCCCTCGAGTTTATAAGGTCTCTCCCCAGTATGG	Cloning - EGFP
ZNFO-ORF-SalI-F	GGGCCCGTCGACTTATGCAGAGGTGTGTTTGCTACA	Cloning – GAL4
ZNFO-ORF-KpnI-R	GGGCCCGGTACCCTAGGGTTTCTCTGCAGTATG	Cloning – GAL4

100

REFERENCES

1. Stitzel ML, Seydoux G. Regulation of the oocyte-to-zygote transition. *Science*. 2007;316(5823):407-408. doi: 10.1126/science.1138236.

2. Memili E, First N. Control of gene expression at the onset of bovine embryonic development. *Biol Reprod*. 1999;61(5):1198-1207. doi: 10.1095/biolreprod61.5.1198.

3. Memili E, First N. Zygotic and embryonic gene expression in cow: A review of timing and mechanisms of early gene expression as compared with other species. *Zygote*. 2000;8(1):87-96. doi: 10.1017/S0967199400000861.

4. TELFORD N, WATSON A, SCHULTZ G. Transition from maternal to embryonic control in early mammalian development - a comparison of several species. *Mol Reprod Dev*. 1990;26(1):90-100. doi: 10.1002/mrd.1080260113.

5. Latham K, Schultz R. Embryonic genome activation. *Frontiers in Bioscience*. 2001;6:D748-D759. doi: 10.2741/Latham.

6. Bettegowda A, Smith GW. Mechanisms of maternal mRNA regulation: Implications for mammalian early embryonic development. *Frontiers in Bioscience*. 2007;12:3713-3726. doi: 10.2741/2346.

7. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: A play in two acts. *Development*. 2009;136(18):3033-3042. doi: 10.1242/dev.033183.

8. Urrutia R. KRAB-containing zinc-finger repressor proteins. *Genome Biol*. 2003;4(10):231. doi: 10.1186/gb-2003-4-10-231.

9. Margolin J, Friedman J, Meyer W, Vissing H, Thiesen H, Rauscher F. Kruppel-associated boxes are potent transcriptional repression domains. *Proc Natl Acad Sci U S A*. 1994;91(10):4509-4513. doi: 10.1073/pnas.91.10.4509.

10. Witzgall R, Oleary E, Leaf A, Onaldi D, Bonventre J. The kruppel-associated box-a (krab-A) domain of zinc-finger proteins mediates transcriptional repression. *Proc Natl Acad Sci U S A*. 1994;91(10):4514-4518. doi: 10.1073/pnas.91.10.4514.

11. Friedman J, Fredericks W, Jensen D, et al. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev*. 1996;10(16):2067-2078. doi: 10.1101/gad.10.16.2067.

12. Moosmann P, Georgiev O, LeDouarin B, Bourquin J, Schaffner W. Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. *Nucleic Acids Res*. 1996;24(24):4859-4867. doi: 10.1093/nar/24.24.4859.

13. Lechner M, Begg G, Speicher D, Rauscher F. Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: Direct chromoshadow domain-KAP-1 corepressor interaction is essential. *Mol Cell Biol*. 2000;20(17):6449-6465. doi: 10.1128/MCB.20.17.6449-6465.2000.

14. Schultz D, Friedman J, Rauscher F. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: The PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the mi-2 alpha subunit of NuRD. *Genes Dev*. 2001;15(4):428-443. doi: 10.1101/gad.869501.

15. Schultz D, Ayyanathan K, Negorev D, Maul G, Rauscher F. SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev*. 2002;16(8):919-932. doi: 10.1101/gad.973302.

16. Underhill C, Qutob MS, Yee SP, Torchia J. A novel nuclear receptor corepressor complex, N-CoR, contains components of the mammalian SWI/SNF complex and the corepressor KAP-1. *J Biol Chem*. 2000;275(51):40463-40470. doi: 10.1074/jbc.M007864200.

17. Wiznerowicz M, Jakobsson J, Szulc J, et al. The kruppel-associated box repressor domain can trigger de novo promoter methylation during mouse early embryogenesis. *J Biol Chem*. 2007;282(47):34535-34541. doi: 10.1074/jbc.M705898200.

18. Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P. KRAB-zinc finger proteins: A repressor family displaying multiple biological functions. *Curr Genomics*. 2013;14(4):268-278.

19. Tripurani SK, Lee K, Wang L, et al. A novel functional role for the oocyte-specific transcription factor newborn ovary homeobox (NOBOX) during early embryonic development in cattle. *Endocrinology*. 2011;152(3):1013-1023. doi: 10.1210/en.2010-1134.

20. Tejomurtula J, Lee K, Tripurani SK, Smith GW, Yao J. Role of importin Alpha8, a new member of the importin alpha family of nuclear transport proteins, in early embryonic development in cattle. *Biol Reprod*. 2009;81(2):333-342. doi: 10.1095/biolreprod.109.077396.

21. Lee K, Bettegowda A, Wee G, Ireland JJ, Smith GW. Molecular determinants of oocyte competence: Potential functional role for maternal (oocyte-derived) follistatin in promoting bovine early embryogenesis. *Endocrinology*. 2009;150(5):2463-2471. doi: 10.1210/en.2008-1574.

22. Bettegowda A, Yao J, Sen A, et al. JY-1, an oocyte-specific gene, regulates granulosa cell function and early embryonic development in cattle. *Proc Natl Acad Sci U S A*. 2007;104(45):17602-17607. doi: 10.1073/pnas.0706383104.

23. Bettegowda A, Patel O, Ireland J, Smith G. Quantitative analysis of messenger RNA abundance for ribosomal protein L-15, cyclophilin-A, phosphoglycerokinase, beta-glucuronidase, glyceraldehyde 3-phosphate dehydrogenase, beta-actin, and histone H2A during bovine oocyte maturation and early embryogenesis in vitro. *Mol Reprod Dev*. 2006;73(3):267-278. doi: 10.1002/mrd.20333.

24. Kim SS, Chen YM, OLeary E, Witzgall R, Vidal M, Bonventre JV. A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. *Proc Natl Acad Sci U S A*. 1996;93(26):15299-15304. doi: 10.1073/pnas.93.26.15299.

25. Iyengar S, Farnham PJ. KAP1 protein: An enigmatic master regulator of the genome. *J Biol Chem*. 2011;286(30):26267-26276. doi: 10.1074/jbc.R111.252569.

26. Groner AC, Meylan S, Ciuffi A, et al. KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *Plos Genetics*. 2010;6(3):e1000869. doi: 10.1371/journal.pgen.1000869.

27. Li X, Ito M, Zhou F, et al. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Developmental Cell*. 2008;15(4):547-557. doi: 10.1016/j.devcel.2008.08.014.

28. Garcia-Garcia MJ, Shibata M, Anderson KV. Chato, a KRAB zinc-finger protein, regulates convergent extension in the mouse embryo. *Development*. 2008;135(18):3053-3062. doi: 10.1242/dev.022897.

29. Zheng R, Blobel GA. GATA transcription factors and cancer. *Genes & cancer*. 2010;1(12):1178-88. doi: 10.1177/1947601911404223.

30. Meirelles F, Caetano A, Watanabe Y, et al. Genome activation and developmental block in bovine embryos. *Anim Reprod Sci*. 2004;82-3:13-20. doi: 10.1016/j.anireprosci.2004.05.012.

31. Baroux C, Autran D, Gillmor CS, Grimanelli D, Grossniklaus U. The maternal to zygotic transition in animals and plants. *Control and Regulation of Stem Cells*. 2008;73:89-100. doi: 10.1101/sqb.2008.73.053.

32. Kotani T, Yamashita M. Discrimination of the roles of MPF and MAP kinase in morphological changes that occur during oocyte maturation. *Dev Biol*. 2002;252(2):271-286. doi: 10.1006/dbio.2002.0853.

33. Eppig JJ, Wigglesworth K, Pendola FL. The mammalian oocyte orchestrates the rate of ovarian follicular development. *Proc Natl Acad Sci U S A*. 2002;99(5):2890-2894. doi: 10.1073/pnas.052658699.

34. Eppig JJ, OBrien M, Wigglesworth K. Mammalian oocyte growth, and development in vitro. *Mol Reprod Dev*. 1996;44(2):260-273.

CHAPTER 3

Preliminary Studies:

Determination of the Consensus Target Sequence Recognized by ZNFO

INTRODUCTION

The term "zinc finger" was first used as laboratory jargon after the discovery of a remarkable 30-residue, repeated sequence motif found in an unexpectedly abundant *Xenopus laevis* transcription factor, because it folded around a zinc ion to form a mini-domain that grasped the DNA[1]. Zinc finger (ZNF) transcription factors are known to interact with DNA through zinc finger motifs and play important roles in a variety of cellular functions, including cell growth, proliferation, development, apoptosis, and intracellular signal transduction[2]. One-third of ZNF proteins contain a highly conserved N-terminal motif known as the KRAB domain[3], which acts as a potent, DNA-binding dependent transcriptional repression module[4]. Members of the KRAB-containing protein family bind DNA through their C2H2 zinc finger domains; two cysteine and two histidine residues tetrahedrally coordinate a zinc ion to fold the domain into the finger-like projection[5]. Within the 30-amino acid repeat of an individual zinc finger motif, a high concentration of basic and polar residues lies between the second cysteine and the first histidine implicating this region as the specific nucleic acid binding region[6].

Binding of the zinc finger motifs occurs though sequence-specific DNA recognition to the promoter regions of genes[2]. Of all the confirmed target DNA binding sequences recognized by ZNFs, not one, or even several, conserved consensus sequences exist for zinc fingers as a family. Considering the vast number of known C2H2 ZNFs and the highly conserved structure of the C2H2 motif, it may seem surprising that each zinc finger protein binds a specific DNA sequence recognized uniquely by itself. However, considering the immense functional diversity of ZNFs it is not unexpected that such sequence-specific binding diversity exists. It is variations to key amino acid residues of the finger domains, spacing, and number of zinc finger motifs that allow for such distinction and specificity[6,7].

Based on the well-conserved functions of KRAB-containing ZNF transcription factors and the observations that ZNFO functions as a transcriptional regulator required during early embryonic developmental events, it is hypothesized that ZNFO mediates downstream activity of potential targets through a *cis*-acting ZNFO consensus recognition sequence. The objective of this study was to identify potential ZNFO DNA binding elements (ZBEs) for the purpose of determining a specific molecular function of ZNFO within the developing bovine embryo.

MATERIALS AND METHODS

Plasmid Construction

For the preparation of glutathione S-transferase (GST) fusion recombinant protein, a partial ZNFO fragment containing the zinc finger motifs (1233 bp- 2145 bp of the ORF; the entire C-terminal portion) was cloned into pGEX-4T1 (GE Healthcare, Salt Lake City, UT) using *Sma*I and *Sal*I sites (ZNFOΔk:pGEX-4T1). The clone was confirmed by sequencing. Primers for restriction-containing primers are listed in Table 1.

Expression and Purification of Glutathione S-Transferase Fusion Protein

The GST-fused ZNFOΔk protein was transfected into *Escherichia coli* Rosetta (DE3)pLysS (Novagen) strain. Bacterial cultures were induced with 2.0 mM isopropyl β-D-thiogalactopyranoside at 35°C for 5 h. Bacterial pellets were collected by centrifugation and bacteria were lysed in equilibration/wash buffer (125 mM Tris and 150 mM sodium chloride, pH 8.0 (Pierce)) by sonication on ice for 10 sec, 12 times (10 sec on, 10 sec off) at 30% amplitude. The clear lysate was incubated and passed through an immobilized glutathione column (Pierce). After extensively washing with equilibration/wash buffer, the recombinant protein was maintained on the beads and stored at 4°C temporarily. A portion of the proteins from various steps were electrophoresed through an SDS-PAGE gel and stained with Coomassie brilliant blue for analysis.

Cyclic Amplification and Selection of Target (CASTing) Assay

A library of single-stranded oligonucleotides containing a 20 bp random core sequence flanked on each side by 23 bp (5'-CAGAGAGCATGTTATGATGGACA-N_{20}-CAAGGACGTGAATCAAATAGGGA-3') was generated. Double-stranded oligonucleotides were prepared by incubating 400 pmol of the library in a polymerase reaction buffer containing 1200 pmol of reverse primer (5'-TCCCTATTTGATTCACGTCCTTG-3'), 10 µM of each deoxynucleoside triphosphate, and 5 units of Taq polymerase and amplified by the following program: 3 min at 95°C, 10 min at 65°C, and 20 min at 72°C. The double-stranded oligonucleotides were purified using QIAquick nucleotide removal kit (Qiagen). The first round of capture was performed by mixing the library with 100 µl of GST-fused ZNFO bound to glutathione beads in 150 µl of a 2X binding buffer containing 50 mM Tris, 200 mM KCl, 2 mM $MgCl_2$, 10 mM DTT, 10% Glycerol, 200 uM $ZnCl_2$, 0.2% Tween20, 200 ug/mL poly (dI-dC), and 2 mg/mL BSA. After incubating for 60 min at RT with continual rotation, the beads were washed five times with cold 1X binding buffer without poly (dI-dC) and then boiled for 7 min in 30 µl sterilized H_2O. The eluted nucleotides were then amplified by PCR and subsequently used for a second round of selection. After seven rounds of amplification, PCR products were purified and exposed to Illumina paired-end sequencing.

Bioinformatic Analysis

Illumina sequencing generated ~518,000 reads. Using Trimmomatic software[8], all adapter- and flanking sequences were removed. All reads were then simultaneously analyzed by BLAST among the collective group of reads. Several groups of similar sequence reads resulted, each with 18-30 reads. A consensus was generated from each group using MEME Suite 4.10.2 software[9].

RESULTS AND DISCUSSION

To investigate the DNA-binding properties of ZNFO, cyclic amplification and selection of targets (CASTing) analysis was performed using a double-stranded oligonucleotide library containing a random core that was incubated with GST-fused ZNFO protein immobilized on

glutathione-sepharose beads for seven rounds of high affinity selection. Each successive round of selection enriched the oligonucleotide core sequence (Fig 1B). As shown in Fig 1B, the DNA-binding pattern of ZNFOΔk was enriched and then maintained with each round as compared to the GST control in which an initial weak binding signal was detected the first round and then completely diminished by the second round. The purified products were sequenced on an Illumina platform.

Many *cis*-acting sites were identified as potential ZBEs (Fig 2). The ZBEs represent consensus sequences recognized by ZNFO in order to bind and regulate potential target transcripts. Several "TATA" patterned motifs were observed. It has been shown that the KRAB domain silences both activated and basal promoter activity of TATA-containing promoters[4]. However, a preliminary EMSA was carried out using several different probes containing the predicted TATA sequences and has been unable to confirm a ZNFO protein-TATA sequence interaction.

Biochemically, the mode of DNA recognition by a finger is principally a one-to-one interaction between individual amino acids from the recognition sequence of the α-helix to individual DNA bases; specifically, amino acids at helical positions -1, 3, and 6 to three successive triplet bases on one strand of the DNA, and helical position 2 to the complementary strand[6]. Hence, each zinc finger motif is capable of contacting three to four nucleotides[10]. Each finger can function as an independent module with its own triplet binding sequence. When several ZNF motifs are linked in tandem, each with different triplet specificities, and together grasping DNA in a linear fashion, a longer and distinctly unique DNA recognition sequence arises. Considering that these transcriptional repressors typically use most of their collection of zinc fingers to bind DNA[10], a protein with 30 zinc finger domains, theoretically, could bind a DNA sequence of more than 60 nucleotides[3]. Translating that notion, the 9 zinc finger motif ZNFO is capable of binding approximately 27 nucleotides (with a range of ~18-37 nt). The oligonucleotide probe designed for this experiment only contained a 20 bp random core. Therefore, it is possible that 1) the TATA sequences identified were accurate but simply aren't complete and therefore cannot bind ZNFO correctly for EMSA confirmation or 2) the derived ZBE sequences themselves are not accurate because GST-ZNFOΔk could not locate the correct sequence as all provided sequences for

recognition were too short. Another possible reason for lack of interaction may be because the full-length ZNFO was not used here.

CONCLUSION

Results of described studies demonstrate that ZNFO is a maternally-derived oocyte-specific factor required for early embryonic development in cattle, and possesses DNA-binding ability, possibly by identified consensus sequences, but the consensus is yet to be identified. Continuation of this study, or repeating it with the full-length ZNFO protein and using an oligo probe with a larger random core, would be useful in combination with electrophoretic mobility shift assay (EMSA) to confirm a target sequence recognized by ZNFO. By determining a target sequence recognized by ZNFO, the ZBE can be aligned to the bovine genome to uncover the genes regulated by ZNFO.
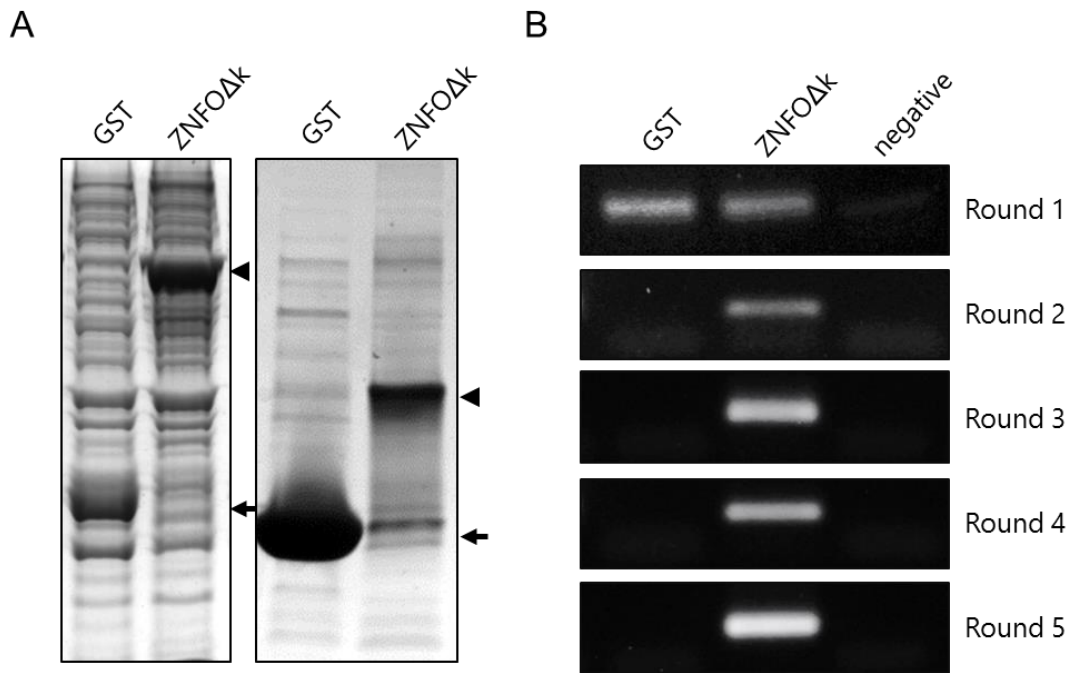
FIGURES

Figure 1. **Determination of the ZNFO consensus DNA binding sequence.** CASTing assays were performed and binding and amplification were done with GST or GST-ZNFOΔk fusion proteins. (A) GST and GST-ZNFOΔk proteins induced in *E. coli* were stained with Coomassie brilliant blue (*left*). Coomassie blue staining of GST and GST-ZNFOΔk proteins bound to glutathione beads (*right*). (B) Results from PCR amplification of bound DNA. Five of seven rounds are shown above. An arrow and arrowhead indicate GST and GST-ZNFOΔk proteins, respectively.
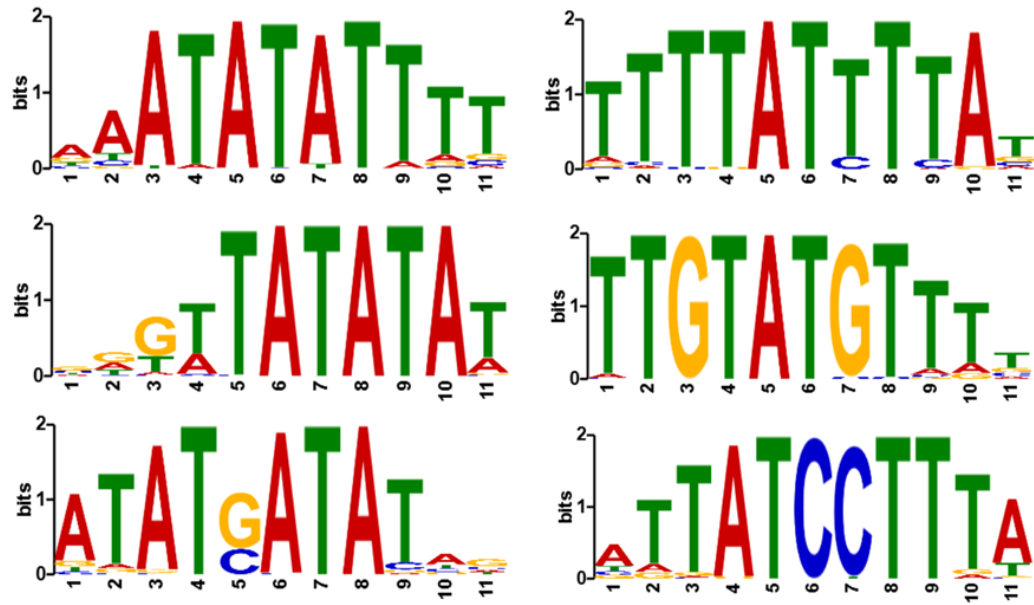
Figure 2. **Examples of potential consensus binding sequences for ZNFO recognition.**

TABLES

Table 1 Primers used in this study

Primer name	Primer sequence (5'-3')	Application
ZNFO-ΔK-SmaI-F	CGCACACCCGGGTTATGCTAAAGCTTTTAACCAGTC	Cloning – GST
ZNFO-ΔK-SalI-R	GGGCCCGTCGACCTAGGGTTTCTCTGCAGTATG	Cloning – GST

REFERENCES

1. Miller J, Mclachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor iiia from xenopus oocytes. *EMBO J*. 1985;4(6):1609-1614.

2. Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P. KRAB-zinc finger proteins: A repressor family displaying multiple biological functions. *Curr Genomics*. 2013;14(4):268-278.

3. Urrutia R. KRAB-containing zinc-finger repressor proteins. *Genome Biol*. 2003;4(10):231. doi: 10.1186/gb-2003-4-10-231.

4. Witzgall R, Oleary E, Leaf A, Onaldi D, Bonventre J. The kruppel-associated box-a (krab-A) domain of zinc-finger proteins mediates transcriptional repression. *Proc Natl Acad Sci U S A*. 1994;91(10):4514-4518. doi: 10.1073/pnas.91.10.4514.

5. Lee MS, Gippert GP, Soman KV, Case DA, Wright PE. 3-dimensional solution structure of a single zinc finger dna-binding domain. *Science*. 1989;245(4918):635-637. doi: 10.1126/science.2503871.

6. Pavletich NP, Pabo CO. Zinc finger dna recognition - crystal-structure of a Zif268-dna complex at 2.1-a. *Science*. 1991;252(5007):809-817. doi: 10.1126/science.2028256.

7. Klug A, Schwabe JWR. Protein motifs .5. zinc fingers. *Faseb Journal*. 1995;9(8):597-604.

8. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics*. 2014;30(15):2114-2120. doi: 10.1093/bioinformatics/btu170.

9. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings / ...International Conference on Intelligent Systems for Molecular Biology ; ISMB.International Conference on Intelligent Systems for Molecular Biology*. 1994;2:28-36.

10. Gebelein B, Urrutia R. Sequence-specific transcriptional repression by KS1, a multiple-zinc-finger-kruppel-associated box protein. *Mol Cell Biol*. 2001;21(3):928-939. doi: 10.1128/MCB.21.3.928-939.2001.

Discussion of Proposed Functional Roles of ZNFO

PROPOSED FUNCTIONS OF ZNFO

The first two chapters and Part I of Chapter 3 describe the data collected for ZNFO and draw conclusions specifically based on such data. Here, I would like to expand those conclusions and discuss the potential functional roles of ZNFO that I believe are sound possibilities.

Early Folliculogenesis

As demonstrated, the expression profile of *ZNFO* fits within the definition of a maternal-effect gene. Abundant expression of *ZNFO* was observed from the first appearance of primordial follicles during embryonic development through folliculogenesis to the pre-ovulatory oocyte in the adult ovary. This illustrates the need for accumulating maternal stores of *ZNFO* transcripts and also suggests a possible role of ZNFO in ensuring proper follicular development.

The early stages of follicle development are critical because many oocyte-specific genes are transcribed during the primordial to primary follicle transition and continue to be expressed throughout folliculogenesis. As identified by molecular genomic and gene knockdown studies, several oocyte/germ-specific transcription factors such as *Nobox*[1], *Figla*[2], *Sohlh1/2*[3,4] and *Lhx8*[5], and growth factors *Gdf9* (Growth differentiation factor 9)[6] and *Bmp15* (bone morphogenetic protein 15)[7], which are found throughout folliculogenesis, but are shown to be absolutely necessary for the primordial to primary transition, maintain normal development of germ cells and surrounding somatic cells essential for mammalian folliculogenesis[8-11].

Depletion of *ZNFO* during folliculogenesis is necessary to determine which stages of follicular development are specifically directed by this novel factor. However, because knockout studies are primarily done with rodent species for obvious reasons, and *ZNFO* is only found in the bovine genome, this type of study becomes particularly challenging and was not performed here.

DNA Methylation and Pluripotency

Recall that maternal factors have several prominent roles during MET[12,13], including removal of maternal RNA and protein, reprogramming of male and female genomes, and embryonic genome activation. Because of the period of transcriptional quiescence in early

116

embryos before embryonic genome activation, maternal proteins stored during oogenesis are likely required for epigenetic reprogramming in early embryos. Several maternal proteins have been described as required factors for epigenetic reprogramming including Tet3 for active DNA demethylation, DPPA3 for maintenance of DNA methylation, and H3.3 for reprogramming and decondensation of chromatin[12,13]. Such nuclear reprogramming is a requirement to activate the transcriptionally inactive embryonic genome. As a maternal-effect gene that was shown to be required for embryo survival past the onset of EGA (8- to 16-cell stage), ZNFO may very likely be involved in reprogramming the epigenome.

Dramatic methylation signature changes occur during early embryonic development. The zygotic genome undergoes passive demethylation until the morula stage[14], maintenance methylation of ICRs occurs, and thereafter *de novo* methylation arrangements are established to sustain successful cell lineage differentiation[15,16]. With such substantial and specific changes occurring in such a narrow time frame, the frame of high ZNFO activity, it is possible that ZNFO is a regulator or methylation-mediated control.

Epigenetic reprogramming is believed to resolve the discrepancy of maternal and paternal chromatin and ensure the successful transition from differentiated to totipotent zygote. In bovine embryos, minor EGA occurs as early as the 2-cell stage[17], a transition that is necessary for successful genome reprogramming and acquisition of totipotency by the embryo. Hence, ZNFO could also be considered a potential factor required for the acquisition of pluripotency or even self-renewal. Further, even though the *ZNFO* transcripts are no longer present following EGA, should ZNFO as a functional factor still be present, it may have a role to play in differentiation of the inner cell mass (ICM) from the trophectoderm. Following the gradual occurrence of EGA, a recently proposed third successive overlapping wave of gene expression termed "mid-preimplantation gene activation" (MGA) takes place, which may play a critical role in cell polarity and the first cell lineage specification[18].

Finally, the dramatic reprogramming of both male and female genomes leads to adjustments in chromatin structure from a repressed chromatin state to one that is open for transcription. If ZNFO is involved in reprogramming, it would also likely have an indirect role in

mediating other key factors required for EGA to occur by permitting or restricting their access to regulatory elements of the genome.

CONCLUSIONS

Indeed, there are many possible functions of ZNFO and several have been proposed here, but it is not entirely uncommon for a protein to have more than one critical function. Consider FIGLA (Factor In the Germline Alpha); primordial follicles cease to develop in *Figla* knockouts and expression of the zona pellucida genes *Zp1, 2,* and *3* is diminished in *Figla* depleted ovaries[2]. Therefore, FIGLA is required early in folliculogenesis for the primordial to primary transition, and, then, later in folliculogenesis for the development of the zona pellucida. Likewise, and based on the current experimental data observed, ZNFO is a repressive regulator of transcription required for early embryogenesis likely by regulating early folliculogenesis and mediating DNA methylation and pluripotency of bovine embryos.

REFERENCES

1. Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM. NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. *Science*. 2004;305(5687):1157-1159. doi: 10.1126/science.1099755.

2. Soyal S, Amleh A, Dean J. FIG alpha, a germ cell-specific transcription factor required for ovarian follicle formation. *Development*. 2000;127(21):4645-4654.

3. Ballow DJ, Xin Y, Choi Y, Pangas SA, Rajkovic A. Sohlh2 is a germ cell-specific bHLH transcription factor. *Gene Expression Patterns*. 2006;6(8):1014-1018. doi: 10.1016/j.modgep.2006.04.007.

4. Pangas SA, Choi Y, Ballow DJ, et al. Oogenesis requires germ cell-specific transcriptional regulators Sohlh1 and Lhx8. *Proc Natl Acad Sci U S A*. 2006;103(21):8090-8095. doi: 10.1073/pnas.0601083103.

5. Choi Y, Ballow DJ, Xin Y, Rajkovic A. Lim homeobox gene, Lhx8, is essential for mouse oocyte differentiation and survival. *Biol Reprod*. 2008;79(3):442-449. doi: 10.1095/biolreprod.108.069393.

6. Dong JW, Albertini DF, Nishimori K, Kumar TR, Lu NF, Matzuk MM. Growth differentiation factor-9 is required during early ovarian folliculogenesis. *Nature*. 1996;383(6600):531-535. doi: 10.1038/383531a0.

7. Dube JL, Wang P, Elvin J, Lyons KM, Celeste AJ, Matzuk MM. The bone morphogenetic protein 15 gene is X-linked and expressed in oocytes. *Molecular Endocrinology*. 1998;12(12):1809-1817. doi: 10.1210/me.12.12.1809.

8. Choi Y, Rajkovic A. Genetics of early mammalian folliculogenesis. *Cellular and Molecular Life Sciences*. 2006;63(5):579-590. doi: 10.1007/s00018-005-5394-7.

9. Pangas SA. Growth factors in ovarian development. *Semin Reprod Med*. 2007;25(4):225-234. doi: 10.1055/s-2007-980216.

10. Pangas SA, Rajkovic A. Transcriptional regulation of early oogenesis: In search of masters. *Hum Reprod Update*. 2006;12(1):65-76. doi: 10.1093/humupd/dmi033.

11. Andreu-Vieyra C, Lin YN, Matzuk MM. Mining the oocyte transcriptome. *Trends in Endocrinology and Metabolism*. 2006;17(4):136-143. doi: 10.1016/j.tem.2006.03.010.

12. Li L, Lu X, Dean J. The maternal to zygotic transition in mammals. *Mol Aspects Med*. 2013;34(5):919-938. doi: 10.1016/j.mam.2013.01.003.

13. Zhang K, Smith GW. Maternal control of early embryogenesis in mammals. *Reproduction Fertility and Development*. 2015;27(6):880-896. doi: 10.1071/RD14441.

14. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*. 2013;502(7472):472-479. doi: 10.1038/nature12750.

15. Seisenberger S, Peat JR, Hore TA, Santos F, Dean W, Reik W. Reprogramming DNA methylation in the mammalian life cycle: Building and breaking epigenetic barriers. *Philosophical Transactions of the Royal Society B-Biological Sciences*. 2013;368(1609):20110330. doi: 10.1098/rstb.2011.0330.

16. Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol*. 2002;241(1):172-182. doi: 10.1006/dbio.2001.0501.

17. Memili E, First N. Zygotic and embryonic gene expression in cow: A review of timing and mechanisms of early gene expression as compared with other species. *Zygote*. 2000;8(1):87-96. doi: 10.1017/S0967199400000861.

18. Hamatani T, Carter MG, Sharov AA, Ko MSH. Dynamics of global gene expression changes during mouse preimplantation development. *Developmental Cell*. 2004;6(1):117-131. doi: 10.1016/S1534-5807(03)00373-3.